

DB2 Universal Database



DB2 Text Extender Verwaltung und Programmierung

Version 7

DB2 Universal Database



DB2 Text Extender Verwaltung und Programmierung

Version 7

Anmerkung

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die allgemeinen Informationen unter „Bemerkungen“ auf Seite 343 gelesen werden.

Erste Ausgabe, Juni 2000

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM DB2 Universal Database Text Extender Administration and Programming Version 7,
IBM Form SC26-9930-00,

herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 1995, 2000
© Copyright IBM Deutschland Informationssysteme GmbH 1995, 2000

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.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:
SW NLS Center
Kst. 2877
Juni 2000

Inhaltsverzeichnis

Zu diesem Buch	vii
Zielgruppe	vii
Verwendungsweise dieses Buchs	vii
Beschreibung der Syntaxdiagramme.	viii
Kommentare zum vorliegenden Buch	x

Zusammenfassung der Änderungen	xi
---	-----------

Teil 1. Einführung 1

Kapitel 1. Übersicht über den DB2 Text Extender 3

Der DB2 Text Extender in der DB2 Client/Server-Umgebung	5
Der DB2 Text Extender in der partitionierten Datenbankumgebung	8

Kapitel 2. Installation, Konfiguration und Verwaltung 9

Installation AIX-, Sun-Solaris- und HP-UX-Systemen	9
Installationsvoraussetzungen	9
Schritt 1 für AIX: Die Produktkomponenten installieren	9
Schritt 1 für Sun-Solaris: Die Produktkomponenten installieren.	11
Schritt 1 für HP-UX: Die Produktkomponenten installieren.	11
Step 2: Das DB2 Text Extender-Exemplar einrichten.	13
Schritt 3: Den DB2 Text Extender-Server starten	14
Installation unter Windows 95, Windows 98, Windows NT und Windows 2000	14
Server-Voraussetzungen	14
Client-Voraussetzungen	14
Die DB2 Extender installieren	14
Exemplare erstellen und verwalten	16
Eine Beispieldatenbank zur Installationsprüfung vorbereiten	17
Konfiguration	18
Umgebungsvariablen	18
Textkonfigurationseinstellungen	19
Die Textkonfiguration ändern	21

Einen DB2 Text Extender-Server konfigurieren und verwalten	22
Ein DB2 Text Extender-Exemplar erstellen	22
Einen DB2 Text Extender-Server starten und stoppen	23
Indizes und aktivierte Datenbanken sichern und zurückschreiben	23
Die Trace-Funktion ausführen	24

Kapitel 3. Erste Schritte 27

Ein einfaches Beispiel dafür, Text für Suchvorgänge vorzubereiten	27
Der Beispiellassistent für die Indexierung und die Suche	29

Kapitel 4. Die Sucherfordernisse planen. . . 31

Warum Textdokumente indiziert sein müssen	31
Unterstützte Dokumentformate	33
Spezielle Überlegungen zu HTML-Dokumenten.	34
Spezielle Überlegungen zu XML-Dokumenten.	34
Nicht unterstützte Dokumentformate verwenden	35
Sprachen	37
CCSIDs	37
EBCDIC	37
ASCII	38
DBCS	39
UNICODE	39
Codepage-Probleme beim Speichern und Aktivieren von Text vermeiden.	40
Sucharten	42
Linguistische Suche	43
Präzise Suche	45
Suchen nach grober Übereinstimmung oder Suchen in DBCS-Dokumenten durchführen	45
Den Indextyp ändern	46
Einen oder mehrere Textindizes für eine Tabelle erstellen.	46
Die Größe eines Indexes berechnen	47
Einen Index aktualisieren	48
Indizes in einer Umgebung mit mehreren Knoten verwalten	50

Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)	50	Die Anzahl der gefundenen Übereinstimmungen suchen und zurückgeben.	86
Attributabschnitte	51	Die Rangordnung eines gefundenen Textdokuments suchen und zurückgeben.	86
Unstrukturierte Dateien und HTML-Dokumente	54	Suchargumente angeben	87
XML-Dokumente	55	Mehrere Begriffe suchen	87
Wörterverzeichnisse, Stoppwortlisten, Abkürzungslisten und Sprachparameter.	57	Mit den Booleschen Operatoren AND und OR suchen	87
Die Stoppwort- und Abkürzungsdateien ändern	58	Nach Variationen eines Begriffs suchen	88
		Nach Teilen eines Begriffs suchen (mit Platzhalterzeichen).	88
Kapitel 5. Text für Suchvorgänge vorbereiten	59	Nach Begriffen suchen, die bereits ein Platzhalterzeichen enthalten.	89
Vor Beginn des Vorbereitens von Text für Suchvorgänge	59	Nach Begriffen in beliebiger Reihenfolge suchen	90
Den DB2 Text Extender-Befehlszeilenprozessor starten	61	Nach Begriffen im selben Satz oder Absatz suchen	90
Hilfe für den Befehlszeilenprozessor	62	Nach Begriffen in Abschnitten von strukturierten Dokumenten suchen.	90
Eine Verbindung zu einer Datenbank herstellen	63	Nach Synonymen von Begriffen suchen.	90
Eine Datenbank aktivieren	64	Eine linguistische Suche durchführen	91
Eine Texttabelle aktivieren (wahlfrei).	65	Mit dem Booleschen Operator NOT suchen	92
Beispiele	66	Suche nach grober Übereinstimmung	93
Eine Textspalte aktivieren	68	Wortgrenzen respektieren	93
Eine Kennungsspalte wird hinzugefügt	69	Nach ähnlich klingenden Wörtern suchen	93
Die Dokumentinformationen werden gesetzt	70	Thesaurus-Suche	94
Eine Protokolltabelle wird erstellt	70	Freitext- und Hybridsuche	95
Ein Index wird erstellt	71	Eine Feinanpassung für eine vorherige Suche durchführen	95
Beispiele	71	Informationen in Kennungen setzen und extrahieren	97
Eine Textspalte in einer großen Tabelle aktivieren.	72	Textdaten beim Einfügen von neuem Text setzen	97
Textspalten für einen nicht unterstützten Datentyp aktivieren	73	Informationen aus Kennungen extrahieren	98
Spalten aktivieren, die den Datentyp DATALINK enthalten	74	Informationen in Kennungen ändern.	99
Externe Textdateien aktivieren	75	Die Suchleistung verbessern.	99
Beispiele	77		
Die Sitzung beenden	78		
Kapitel 6. Vorgehensweise bei Suchvorgängen	79	Kapitel 7. Verwaltung	101
Lokation von Syntaxbeispielen von Suchfunktionen	80	Textindizes verwalten	101
Die Beispieltabelle DB2TX.SAMPLE	80	Einen Index aktualisieren	101
Kennungen für externe Dateien	83	Einen Index für externe Dateien aktualisieren.	102
Den aktuellen Funktionspfad setzen	83	Die Einstellungen eines Indexes ändern	102
Nach Text suchen	84	Den Indexstatus zurücksetzen.	103
Eine Abfrage durchführen	85	Indexereignisse löschen	104
		Einen Index reorganisieren.	105
		Nützliche Informationen abrufen.	105
		Informationen zum aktivierten Status anzeigen	106

Die Einstellungen der Umgebungsvariablen anzeigen	106
Die Textkonfigurationseinstellungen anzeigen	107
Den Status eines Indexes anzeigen	108
Fehlerereignisse anzeigen	109
Die Indexeinstellungen anzeigen	110
Die Texteinstellungen für eine Spalte anzeigen	111
Mit der DB2 Text Extender-Katalogsicht arbeiten	112
Den Textvorbereitungsprozeß umkehren	115
Eine Textspalte inaktivieren	115
Textdateien inaktivieren	116
Eine Texttabelle inaktivieren	116
Eine Datenbank inaktivieren	117

Kapitel 8. Die API-Funktionen zum Suchen und Anzeigen verwenden 119

Ihre Anwendung einrichten	119
Eine UNIX-Anwendung verbinden	119
Eine OS/2- oder Windows-Anwendung verbinden	119
Überblick über die API-Funktionen	120
Nach Text suchen	121
Eine Sucherggebnistabelle abrufen (DesGetSearchResultTable)	121
Text anzeigen	122
Anzeigeinformationen abrufen (DesGetBrowseInfo)	123
Eine Anzeigesitzung starten (DesStartBrowseSession)	123
Ein Dokument öffnen (DesOpenDocument)	123
Übereinstimmungen abrufen (DesGetMatches)	124
Ein Dokument schließen (DesCloseDocument)	125
Eine Anzeigesitzung beenden (DesEndBrowseSession)	125
Die Anzeigeinformationen freigeben (DesFreeBrowseInfo)	125

Teil 2. Referenz 127

Kapitel 9. Textvorbereitungs- und Verwaltungsbefehle für den Client 129

CHANGE INDEX SETTINGS	131
CHANGE TEXT CONFIGURATION	133
CONNECT	137

DELETE INDEX EVENTS	139
DISABLE DATABASE	140
DISABLE TEXT COLUMN	141
DISABLE TEXT FILES	142
DISABLE TEXT TABLE	143
ENABLE DATABASE	144
ENABLE TEXT COLUMN	145
ENABLE TEXT FILES	154
ENABLE TEXT TABLE	157
GET ENVIRONMENT	162
GET INDEX SETTINGS	163
GET INDEX STATUS	165
GET STATUS	166
GET TEXT CONFIGURATION	167
GET TEXT INFO	168
QUIT	169
REORGANIZE INDEX	170
RESET INDEX STATUS	171
UPDATE INDEX	172

Kapitel 10. Verwaltungsbefehle für den Server 175

TXICRT	176
TXIDROP	178
TXILIST (nur UNIX)	179
TXIPCLEAN (nur UNIX)	180
TXNADD	181
TXNCHECK	182
TXNDROP	183
TXSAMPLE	184
TXSTART	185
TXSTATUS	186
TXSTOP	187
TXTHESC	188
TXTHESN	190
TXTRACE	192
TXVERIFY	197
TXWIZARD	198

Kapitel 11. Suchfunktionen 199

Die eindeutigen Typen des DB2 Text Extender	199
Eine Zusammenfassung der DB2 Text Extender-Funktionen	200
CCSID	202
CONTAINS	203
FILE	204
FORMAT	205
HANDLE	206
HANDLE_LIST	207

INIT_TEXT_HANDLE	208	Stufe 1: Normalisierung und Begriffs- erweiterung.	269
LANGUAGE	210	Stufe 2: Erweiterte Übereinstimmung	270
NO_OF_DOCUMENTS	211	Thesaurus-Konzepte	271
NO_OF_MATCHES	212	Begriffe	273
RANK	213	Relationen	273
REFINE	214	Ngram-Thesaurus-Relation.	276
SEARCH_RESULT	215	Einen Thesaurus erstellen	276
Kapitel 12. Syntax von Suchargumenten	217	Einen Ngram-Thesaurus erstellen	281
Suchargument	219	Kapitel 16. Konfigurationsdateien.	285
Kapitel 13. API-Funktionen zum Suchen und Anzeigen.	231	Client-Konfigurationsdatei	285
DesCloseDocument	233	Server-Konfigurationsdatei.	287
DesEndBrowseSession	234	Konfigurationsdatei für die Textanalyse	291
DesFreeBrowseInfo	235	Allgemeine Optionen	292
DesGetBrowseInfo	236	Kapitel 17. Rückkehrcodes	295
DesGetMatches	239	Kapitel 18. Nachrichten	303
DesGetSearchResultTable	245	SQLSTATES, die von DB2 Text Extender- Funktionen zurückgegeben werden	303
DesOpenDocument	250	Nachrichten vom DB2 Text Extender	307
DesStartBrowseSession	253	Kapitel 19. Ursachencodes der Such- maschine	321
Kapitel 14. Beispiel-API-Programm	255	Kapitel 20. Ursachencodes für Fehler- ereignisse	325
Kapitel 15. Die Verarbeitung von linguisti- schen und präzisen Indizes auf linguisti- scher Basis	257	Teil 3. Anhänge und Schlußteil	341
Verarbeitung auf linguistischer Basis beim Indexieren	257	Bemerkungen	343
Basistextanalyse	258	Marken	346
Begriffe auf ihre Grundform reduzieren	265	Glossar	347
Stoppwortfiltern	265	Index	353
Dekomposition (Trennen von Komposita)	265	Kontaktaufnahme mit IBM	363
Verarbeitung auf linguistischer Basis für die Abfrage	266	Produktinformationen	363
Synonyme	267		
Thesaurus-Erweiterung	268		
Lauterweiterung	268		
Platzhalterzeichen für Zeichen und Wör- ter	268		
Verarbeitung auf linguistischer Basis für die Anzeige	269		

Zu diesem Buch

In diesem Buch wird beschrieben, wie der DB2 Text Extender dazu verwendet wird, eine DB2-Datenbank zum Abrufen von Textdaten vorzubereiten und zu verwalten. Außerdem wird beschrieben, wie Sie die vom DB2 Text Extender bereitgestellten SQL-Funktionen und Anwendungsprogrammierschnittstellen (APIs) für den Zugriff und die Bearbeitung dieser Datentypen verwenden können. Wenn Sie die Funktionen und APIs des DB2 Text Extender in den SQL-Anweisungen Ihres Programms verwenden, können Sie leistungsfähige und vielseitige Programme zum Textabruf erstellen.

Verweise in diesem Buch auf "DB2" beziehen sich auf DB2 UDB.

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 DB2 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.

Verwendungsweise dieses Buchs

Dieses Buch ist folgendermaßen strukturiert:

„Teil 1. Einführung“

Dieser Teil enthält eine Übersicht zum DB2 Text Extender. Zudem wird die Installation und Konfiguration beschrieben und werden Überlegungen zur Planung vorgestellt. Darüber hinaus wird auch die Vorbereitung und Verwaltung einer DB2-Datenbank beschrieben, so daß damit nach Text gesucht werden kann.

Lesen Sie diesen Teil, wenn Sie mit dem DB2 Text Extender noch nicht vertraut sind und lernen möchten, wie die DB2 Text Extender-Funktionen und -APIs zum Suchen nach Text verwendet werden.

Zu diesem Buch

„Teil 2. Referenz“

In diesem Teil werden Referenzinformationen zu Funktionen, APIs und Befehlen des DB2 Text Extender sowie zu diagnostischen Informationen, wie z. B. Nachrichten und Codes, dargestellt.

Lesen Sie diesen Teil, wenn Sie bereits mit den DB2 Text Extender-Konzepten und -Tasks vertraut sind, aber weitere Informationen zu einer spezifischen Funktion oder API, einem spezifischen Befehl, einer spezifischen Nachricht oder einem spezifischen Code des DB2 Text Extender benötigen.

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 \blacktriangleright — kennzeichnet den Anfang einer Anweisung.

Das Symbol \longrightarrow gibt an, daß die Anweisungssyntax in der nächsten Zeile fortgesetzt wird.

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

Das Symbol $\longrightarrow\blacktriangleleft$ kennzeichnet das Ende einer Anweisung.

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

\blacktriangleright —erforderliches Element— \blacktriangleleft

- Wahlfreie Elemente werden unterhalb des Hauptpfads dargestellt.

\blacktriangleright — $\underbrace{\hspace{10em}}_{\text{wahlfreies Element}}$ — \blacktriangleleft

- 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.



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. 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.

Kommentare zum vorliegenden Buch

Ihre Rückmeldungen unterstützen IBM dabei, qualitativ hochwertige Informationen bereitzustellen. Bitte senden Sie alle Kommentare ein, die Sie zu diesem Handbuch oder zu anderen DB2 Extender-Dokumentationen haben. Zum Abgeben der Kommentare können Sie eine der folgenden Methoden verwenden:

- Senden Sie Ihre Kommentare über das Web. Besuchen Sie die Web-Site unter folgender Adresse:

<http://www.ibm.com/software/data/db2/extenders>

Diese Web-Site verfügt über eine Seite für Rückmeldungen, die Sie zum Eingeben und Senden von Kommentaren verwenden können.

- Senden Sie Ihre Kommentare als E-Mail an swsdid@de.ibm.com. 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. B. Seitennummer oder Tabellennummer).
- Füllen Sie ein Formular für Leserkommentare am Ende dieses Buchs aus und schicken Sie es per Post bzw. per Fax zurück oder geben Sie es einem IBM Ansprechpartner. Die Postadresse befindet sich auf der Rückseite des Formulars. Die Faxnummer lautet +49-(0)7031-16-4892.

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

Zusammenfassung der Änderungen

Folgende Änderungen an diesem Dokument sind in Release 7.1 enthalten:

- **Suche nach Abschnitten, die einen bestimmten Typ haben**
Die unterstützten Typen sind DATE/TIME und INTEGER. Dadurch können Sie nach Dokumenten suchen, die beispielsweise im Abschnitt Publishing date (Veröffentlichungsdatum) ein späteres Datum als den 13.12.99 haben.
- **Datalink-Unterstützung**
Dokumente, auf die über eine Spalte mit einem Datalink-Datentyp verwiesen wird, können indexiert werden.
- **Weitere CCSIDs werden unterstützt**
CCSID-Einschränkungen für den Ngram-Indextyp wurden entfernt.
- **Der duale Indextyp wird nicht mehr unterstützt**
Verwenden Sie statt dessen einen linguistischen oder einen präzisen normierten Indextyp.
- **Geänderte Syntax der Funktion SEARCH_RESULT**
Die Syntax der Funktion SEARCH_RESULT wurde geändert. Die alte Syntax wird aus Gründen der Abwärtskompatibilität weiterhin unterstützt.
- **Weitere unterstützte Sprachen**
Die Liste der unterstützten Sprachen und CCSIDs wurde erweitert.
- **Die Konfigurationsfunktionen wurden erweitert**
Sie können den Tokenizer konfigurieren und die Server-Konfiguration anpassen, um beispielsweise Schwellenwerte zu reorganisieren.
- **Neuer Befehl zur Systembereinigung unter UNIX**
Es gibt einen neuen Server-Befehl TXIPCLEAN für UNIX-Plattformen, mit dem die Systembereinigung erleichtert wird, wenn der DB2 Text Extender-Server nicht erfolgreich gestartet werden kann.

Zusammenfassung der Änderungen

Teil 1. Einführung

Kapitel 1. Übersicht über den DB2 Text Extender

Der DB2 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 dazu in der Lage, nach Bildern, Video- und Sprachdaten zu suchen, und können XML-Dokumente und räumliche Daten bearbeiten.

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

Der DB2 Text Extender bietet DB2-Benutzern und Anwendungsprogrammierern eine schnelle, vielseitige und intelligente Methode zum Durchsuchen von Textdokumenten. Die Stärke des DB2 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-Datenbanken gespeichert sind, sondern Sie können auch in Textdokumenten suchen, die in Dateien gespeichert sind.

Im Zentrum des DB2 Text Extender steht die IBM Hochleistungstechnologie für die linguistische Suche, die in „Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis“ auf Seite 257 beschrieben wird. 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. 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.

Übersicht

- 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.
- Ausführen einer Freitextsuche, in der das Suchargument in natürlicher Sprache ausgedrückt wird.
- 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 DB2 Text Extender-Programmierschnittstelle können Sie sogar den Benutzern Ihrer Anwendung ermöglichen, die Dokumente anzuzeigen.

Durch die Integration der Volltextsuche in SELECT-Abfragen von DB2 steht Ihnen eine leistungsstarke Abruffunktion zur Verfügung. 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 DB2 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.

Der DB2 Text Extender in der DB2 Client/Server-Umgebung

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

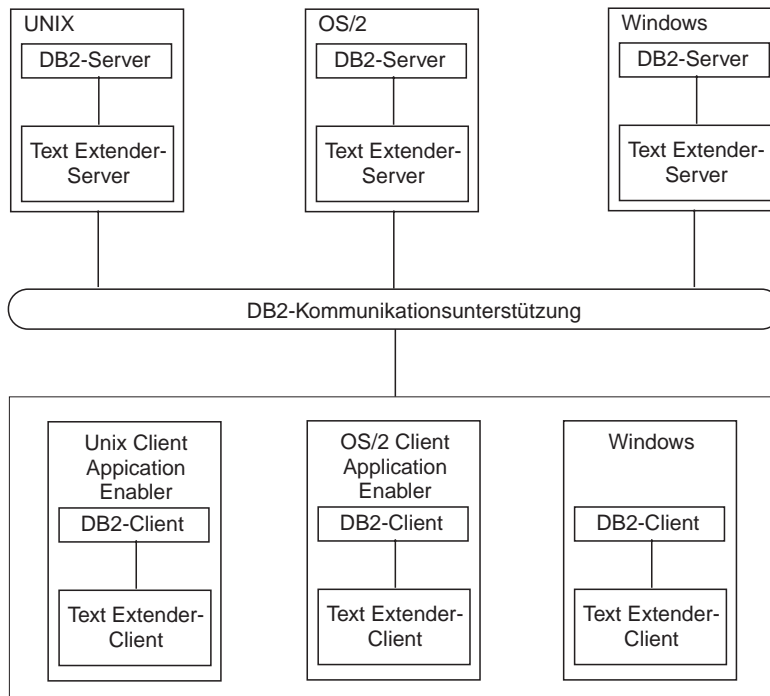


Abbildung 1. Integration des DB2 Text Extender in der DB2 Client/Server-Umgebung

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

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

Übersicht

Eine DB2 Text Extender-Installation ist flexibel und kann folgende Komponenten umfassen:

- Einen oder mehrere DB2 Text Extender-Server auf einem der Betriebssysteme, die in Abb. 1 auf Seite 5 gezeigt werden, wobei UNIX auch AIX-, SUN-Solaris- und HP-UX-Workstations einschließt.
- AIX-, SUN-Solaris-, Windows NT- und Windows 2000-Clients mit Zugriff auf einen oder mehrere ferne DB2 Text Extender-Server.
- AIX-Clients, die einen lokalen Server und Zugriff auf ferne Server haben.

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

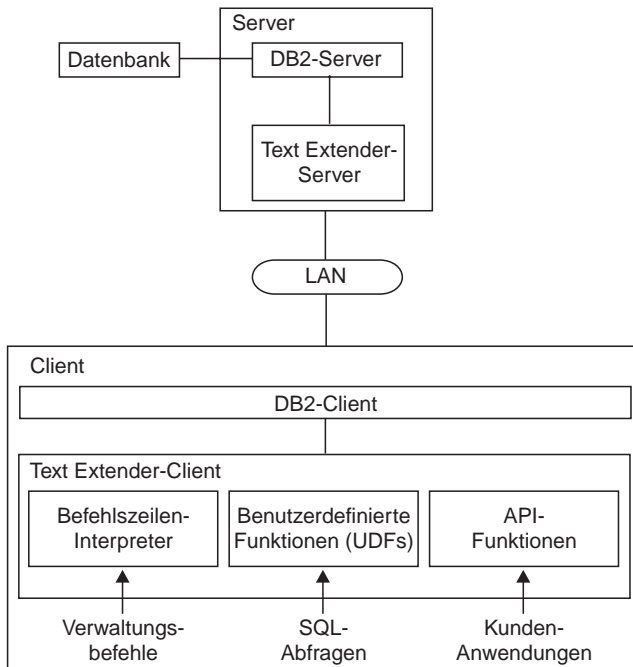


Abbildung 2. Eine DB2 Text Extender-Konfiguration

Der DB2 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.
- **SQL-Funktionen** 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 SQL-Funktionen auf dem Client, da sie als Teil einer SQL-Abfrage verwendet werden können. Tatsächlich sind sie Teil der Server-Installation und werden dort ausgeführt. Diese Funktionen können jedoch von jedem DB2-Client aus verwendet werden, ohne daß der DB2 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 Client-Dienstprogramme des DB2 Text Extender bieten Funktionen zur Textvorbereitung und zur Verwaltung sowie die API. (Diese Funktionen sind auf dem Server verfügbar.) Zur Verwendung dieser Funktionen müssen Sie den DB2 Text Extender-Client installieren. Wenn Sie die Textvorbereitung und die Verwaltung nur auf dem DB2 Text Extender-Server durchführen, müssen Sie die DB2 Text Extender-Client-Dienstprogramme nur auf Clients installieren, die die API-Funktionen verwenden.

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

Der DB2 Text Extender in der partitionierten Datenbankumgebung

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

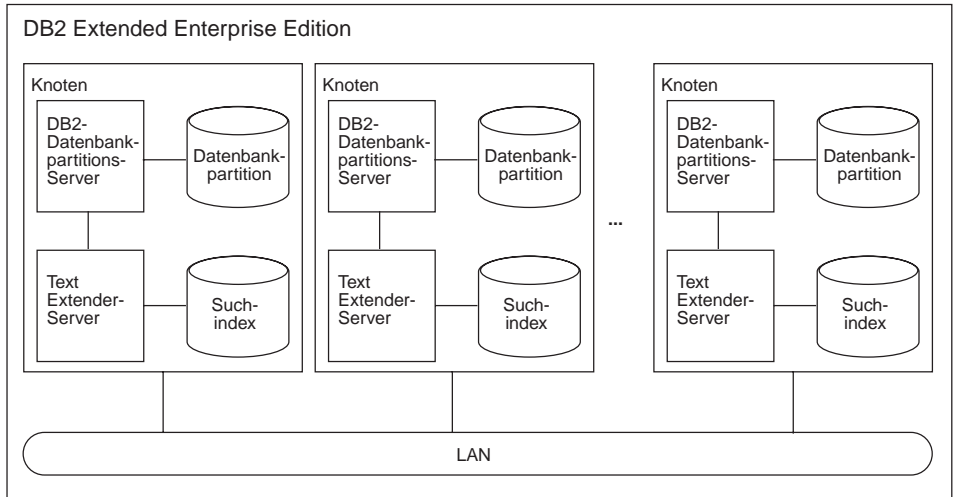


Abbildung 3. Integration des DB2 Text Extender in die DB2-Umgebung für partitionierte Datenbanken

Jeder Datenbankpartitions-Server, als *Knoten* bekannt, kann einer separaten Maschine zugeordnet sein, oder sie können sich alle auf einer einzelnen Maschine befinden. 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 DB2 Text Extender unterstützt beide Konfigurationen.

Eine *Knotengruppe* ist eine benannte Untergruppe von einem oder mehreren Datenbankpartitions-Servern. Wenn eine DB2 Text Extender-fähige Tabelle einer Mehrpartitions-knotengruppe zugeordnet ist, befinden sich die Text-indizes, die vom DB2 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 DB2 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 DB2 Text Extender über eine Datenbankpartition. Jede Datenbankpartition kann verwendet werden, um DB2 Text Extender-Anforderungen auszugeben. Die Anforderung wird an alle betroffenen Datenbankpartitionen weitergeleitet.

Kapitel 2. Installation, Konfiguration und Verwaltung

Installation AIX-, Sun-Solaris- und HP-UX-Systemen

Der DB2 Text Extender Version 7 für AIX, Sun-Solaris und HP-UX kann in einer Umgebung für nicht partitionierte Datenbanken ausgeführt werden. Unter AIX und Sun-Solaris kann er auch die Unterstützung für partitionierte Datenbanken nutzen, die von DB2 Extended Enterprise Edition bereitgestellt wird.

Im folgenden sind die Installationsschritte aufgeführt:

1. Installieren Sie die Produktkomponenten auf der Zielmaschine.
2. Richten Sie das DB2 Text Extender-Exemplar ein.
3. Starten Sie den DB2 Text Extender-Server.

Installationsvoraussetzungen

Bevor Sie den DB2 Text Extender installieren, sollten Sie sicherstellen, daß die folgenden Installationsvoraussetzungen für die zugehörigen Server- und Client-Komponenten erfüllt sind.

Server-Voraussetzungen

Die Server-Komponenten erfordern die Installation von DB2 Version 7.1 (oder höher).

Client-Voraussetzungen

Die Client-Komponenten erfordern die Installation von DB2 Version 7.1 (oder höher) Personal Edition, DB2 Version 7.1 (oder höher) Client Application Enabler oder DB2 Version 7.1 (oder höher) Software Developer's Kit.

Schritt 1 für AIX: Die Produktkomponenten installieren

Im Verlauf der Installation wird automatisch eine Gruppe mit dem Namen smadmin erstellt.

1. Melden Sie sich auf der Ziel-Client- oder -Server-Maschine als Root an.
2. Verwenden Sie in einer Schnittstelle mit Bedienungsführung SMIT (System Management Interface Tool) zur Ausführung von `installp`, oder geben Sie den Befehl `installp` ein.
3. Geben Sie den Befehl `smit install_latest` ein. Das Menü **Software Installation and Maintenance Tool** wird geöffnet.
4. Geben Sie die Lokation des Pakets im Feld EINGABE-Einheit/Verzeichnis für Software ein.
5. Klicken Sie auf den Knopf Aktion oder drücken Sie die Eingabetaste. Dadurch wird das angegebene Installationsverzeichnis bestätigt.

Installation auf UNIX-Systemen

- Legen Sie im Feld SOFTWARE die installiert werden muß fest, ob die Client- oder die Server-Komponenten installiert werden müssen.
- Klicken Sie auf den Knopf Aktion oder drücken Sie die Eingabetaste. Sie werden dazu aufgefordert, die Installationsparameter zu bestätigen. Drücken Sie dazu die Eingabetaste.
- Legen Sie bei Bedarf im Feld SOFTWARE die installiert werden muß fest, welche Wörterverzeichnisse Sie auf dem Client installieren möchten. Wenn Sie kein Wörterverzeichnis angeben, wird das Wörterverzeichnis für amerikanisches Englisch installiert.

In Tabelle 1 sind die DB2 Text Extender-Wörterverzeichnisse aufgelistet. Wenn Sie alle Wörterverzeichnisse im Client installieren möchten, machen Sie im Feld SOFTWARE die installiert werden muß die Eingabe
db2tx_07_01.dic.

Tabelle 1. DB2 Text Extender-Wörterverzeichnisse

Wörterverzeichnis	Sprache
db2tx_07_01.dic.Ar_AA	Arabisch
db2tx_07_01.dic.Da_DK	Dänisch
db2tx_07_01.dic.NL_NL	Niederländisch
db2tx_07_01.dic.En_GB	Englisch (Großbritannien)
db2tx_07_01.dic.En_US	Englisch (Vereinigte Staaten)
db2tx_07_01.dic.Fi_FI	Finnisch
db2tx_07_01.dic.Fr_FR	Französisch
db2tx_07_01.dic.Fr_CA	Französisch (Kanada)
db2tx_07_01.dic.De_DE	Deutsch
db2tx_07_01.dic.De_CH	Deutsch (Schweiz)
db2tx_07_01.dic.Iw_IL	Hebräisch
db2tx_07_01.dic.Is_IS	Isländisch
db2tx_07_01.dic.It_IT	Italienisch
db2tx_07_01.dic.No_NO	Norwegisch
db2tx_07_01.dic.Pt_PT	Portugiesisch
db2tx_07_01.dic.Ru_RU	Russisch
db2tx_07_01.dic.Es_ES	Spanisch und Katalanisch
db2tx_07_01.dic.Sv_SE	Schwedisch
db2tx_07_01.dic.Th_TH	Thailändisch

- Melden Sie sich ab.

Schritt 1 für Sun-Solaris: Die Produktkomponenten installieren

Im Verlauf der Installation wird automatisch eine Gruppe mit dem Namen smadmin erstellt.

1. Melden Sie sich auf der Ziel-Client- oder -Server-Maschine als Root an.
2. Verwenden Sie den Befehl tar, um aus der Fixpack-Datei das installierbare Paket zu erhalten:

```
tar -xvf IBMdb2tx.tar
```

3. Verwenden Sie den Befehl pkgadd, um die Dateien aus Ihrem Verzeichnis auf die Zielmaschine zu übertragen. Dafür können Sie eine der beiden folgenden Methoden auswählen:

- Installieren Sie den DB2 Text Extender-Client und -Server mit dem Wörterverzeichnis in amerikanischem Englisch:

```
pkgadd -d . db2txclnt db2txsrv db2txcom db2txmsgs db2txmsc db2txdius
```

- Geben Sie an, ob Sie einen DB2 Text Extender-Client oder -Server installieren möchten und welche Wörterverzeichnisse hinzugefügt werden sollen:

```
./db2tx_install
```

4. Melden Sie sich ab.

Schritt 1 für HP-UX: Die Produktkomponenten installieren

Im Verlauf der Installation wird automatisch eine Gruppe mit dem Namen smadmin erstellt.

1. Melden Sie sich auf der Ziel-Client- oder -Server-Maschine als Root an.
2. Verwenden Sie den Befehl tar, um aus der Fixpack-Datei das installierbare Paket zu erhalten:

```
tar -xvf b2tx.tar
```

3. Führen Sie das Programm swinstall aus:

```
swinstall
```

Dadurch werden die Fenster Software Selection und Specify Source geöffnet.

4. Ändern Sie bei Bedarf den 'Source Host Name'.
5. Geben Sie Ihr tar-Verzeichnis ein, wobei Sie IBMdb2tx als Wert für das Feld Source Depot Path verwenden.
6. Wählen Sie im Feld Change Software View die Option Products aus.
7. Wählen Sie OK aus, um zum Fenster Software Selection zurückzukehren, das eine Liste mit der verfügbaren zu installierenden Software enthält.
8. Führen Sie jetzt eine der nachfolgenden Anweisungen aus:
 - Heben Sie die Option DB2TX Product hervor, um den DB2 Text Extender mit allen Wörterverzeichnissen und der Dokumentation in allen verfügbaren Sprachen zu installieren.

Installation auf UNIX-Systemen

- Gehen Sie wie folgt vor, um Teile des DB2 Text Extender zu installieren:
 - a. Wählen Sie im Menü Actions die Option Open Item aus.
 - b. Heben Sie alle Dateigruppen, die Sie installieren möchten, hervor und wählen Sie anschließend im Menü Actions die Option Choose Mark for Install aus. Sie können auswählen, ob Sie den DB2 Text Extender-Server oder -Client installieren möchten:
 - Text Extender-Server - DB2TX.db2txsrv
 - Text Extender-Client - DB2TX.db2txclnt
 - c. Wählen Sie bei Bedarf die Wörterverzeichnisse aus, die Sie installieren möchten. In Tabelle 2 sind die DB2 Text Extender-Wörterverzeichnisse aufgelistet.

Tabelle 2. DB2 Text Extender-Wörterverzeichnisse

Wörterverzeichnis	Sprache
db2txdiaa	Arabisch
db2txdidk	Dänisch
db2txdinl	Niederländisch
db2txdigb	Englisch (Großbritannien)
db2txdius	Englisch (Vereinigte Staaten)
db2txdifi	Finnisch
db2txdifr	Französisch
db2txdica	Französisch (Kanada)
db2txdide	Deutsch
db2txdich	Deutsch (Schweiz)
db2txdiil	Hebräisch
db2txdiis	Isländisch
db2txdiit	Italienisch
db2txdino	Norwegisch
db2txdipt	Portugiesisch
db2txdies	Russisch
db2txdiru	Spanisch und Katalanisch
db2txdise	Schwedisch
db2txdith	Thailändisch

9. Wählen Sie im Menü Actions die Option Install aus, um mit der Installation des Produkts zu beginnen und das Fenster Install Analysis zu öffnen.

10. Wählen Sie im Fenster Install Analysis den Knopf OK aus, wenn im Feld Status die Nachricht Ready angezeigt wird.
11. Wählen Sie im Fenster Confirmation den Knopf Yes aus, um zu bestätigen, daß Sie die DB2 für HP-UX-Softwareprodukte installieren möchten.
12. Zeigen Sie das Fenster Install an, um während der Installation der Software Verarbeitungsdaten zu lesen. Die Installation ist vollständig abgeschlossen, wenn im Feld Status die Nachricht Ready angezeigt und das Fenster Note geöffnet wird.

Das Programm `swinstall` lädt die Dateigruppe und führt dafür die Steuerprozeduren aus.

Das Betriebssystem HP-UX bietet detaillierte Hilfe für das Programm `swinstall`. Diese Hilfe ist verfügbar, wenn folgender Befehl eingegeben wird:

```
man swinstall
```

13. Melden Sie sich ab.

Step 2: Das DB2 Text Extender-Exemplar einrichten

1. Stellen Sie sicher, daß Sie als Root aktiv sind.
2. Geben Sie unter AIX folgendes ein:

```
cd /usr/lpp/db2ext_07_01/instance
```

Geben Sie unter Solaris und HP-UX folgendes ein:

```
cd /opt/IBMDB2TX/V7.1/instance
```

3. Wenn Sie über ein bereits vorhandenes Exemplar verfügen, geben Sie folgendes ein:

```
./txiupdt exemplar_id db2instanceid
```

Dadurch wird Ihr Exemplar mit den neuen Dateien aktualisiert.

Geben Sie andernfalls zum Erstellen eines neuen DB2 Text Extender-Exemplars folgendes ein:

```
./txicrt exemplar_id db2instanceid
```

Hierbei steht *exemplar_id* für eine Benutzer-ID eines bestehenden DB2-Exemplars oder eine Benutzer-ID, die einer DB2-Exemplar-ID zugeordnet werden soll.

Durch Ausführen des Befehls `txicrt` wird das Verzeichnis `/u/exemplar_id/db2tx` erstellt. Erstellen Sie im Verzeichnis `/u/exemplar_id/db2tx` keine weiteren Dateien oder Verzeichnisse; diese gehen beim Löschen des Exemplars verloren.

Installation auf UNIX-Systemen

Zusätzliche Parameter sind dann erforderlich, wenn Sie eine Umgebung mit partitionierten Datenbanken verwenden. Weitere Informationen finden Sie im Abschnitt „TXICRT“ auf Seite 176.

4. Melden Sie sich ab.

Schritt 3: Den DB2 Text Extender-Server starten

1. Melden Sie sich als *exemplar_id* an.
2. Stoppen Sie DB2 und starten Sie es erneut.
3. Geben Sie folgenden Befehl ein, um den DB2 Text Extender-Server zu starten:
txstart
4. Melden Sie sich ab.

Installation unter Windows 95, Windows 98, Windows NT und Windows 2000

Der DB2 Text Extender wird als Teil des DB2 Extender-Pakets installiert.

Server-Voraussetzungen

Die Server-Komponenten für die DB2 Extender können auf einem Windows NT- oder Windows 2000-Server installiert werden.

Die Server-Komponenten, die in einer Umgebung für nicht partitionierte Datenbanken ausgeführt werden, erfordern die Installation von DB2 Version 7.1 (oder höher).

Die Server-Komponenten, die in einer Umgebung mit partitionierten Datenbanken ausgeführt werden, erfordern die Installation von DB2 Extended Enterprise Edition Version 7.1 (oder höher).

Client-Voraussetzungen

Die Client-Komponenten für die DB2 Extender können unter Windows 95, Windows 98, Windows NT oder Windows 2000 installiert werden.

Die Client-Komponenten erfordern die Installation von DB2 Version 7.1 (oder höher) Personal Edition, DB2 Version 7.1 (oder höher) Client Application Enabler oder DB2 Version 7.1 (oder höher) Software Developer's Kit für Windows.

Die DB2 Extender installieren

Gehen Sie wie folgt vor, um die DB2 Extender unter Windows zu installieren:

1. Legen Sie die CD-ROM in das CD-ROM-Laufwerk Ihrer Client-Workstation ein.
2. Wenn die Funktion zum automatischen Starten von CDs aktiviert ist, wird das DB2 Extender LaunchPad automatisch gestartet. Klicken Sie

andernfalls auf den Knopf Start auf der Task-Leiste. Wählen Sie Ausführen aus. Geben sie im Feld Öffnen folgendes ein:

```
d:\db2ext\winnt95\setup
```

Hierbei ist "d" der Laufwerksbuchstabe des CD-ROM-Laufwerks.

3. Vom LaunchPad aus können Sie die Readme-Datei für den DB2 Extender anzeigen. Sehen Sie sich die Installationsvoraussetzungen an oder installieren Sie das Produkt.
4. Wählen Sie den Knopf Installieren auf der linken Seite aus, um mit der Installation zu beginnen.
5. Im ersten Fenster können Sie auswählen, welche Version der DB2 Extender installiert werden soll: 'Client/Server für EEE', 'Client/Server für nicht-EEE' oder nur 'Client'.
6. Im nächsten Fenster können Sie den Ordner auswählen, in dem die Extender installiert werden sollen.
7. Wählen Sie anschließend den gewünschten Installationstyp aus: 'Standard', 'Minimal' oder 'Benutzer'.

Die **Standard**installation enthält den DB2 Extender-Client/-Server, Beispieldateien, SDK, Online-Dokumentation in amerikanischem Englisch und das DB2 Text Extender-Wörterverzeichnis für amerikanisches Englisch.

Die **Minimal**installation installiert nur den Client/Server und das DB2 Text Extender-Wörterverzeichnis für amerikanisches Englisch.

Bei der **Benutzer**installation können Sie aus den folgenden Komponenten auswählen:

- Wählen Sie **DB2 Extender-Server (EEE)** aus, wenn Sie die DB2 Extender in einer Umgebung mit partitionierten Datenbanken installieren (DB2 Extended Enterprise Edition). Hierdurch wird auch die Komponente **DB2 Extender-Client** installiert.
- Wählen Sie **DB2 Extender-Server (nicht-EEE)** aus, wenn Sie die DB2 Extender in einer Umgebung für nicht partitionierte Datenbanken installieren. Hierdurch wird auch die Komponente **DB2 Extender-Client** installiert.
- Wählen Sie **DB2 Extender-Client** aus, wenn Sie die DB2 Extender-Client-Umgebung installieren.
- Wählen Sie **DB2 Extender SDK** aus, wenn Sie Kopffdateien und Bibliotheken für die Entwicklung von DB2 Extender-Anwendungen installieren wollen.
- Wählen Sie **Beispieldateien** aus, wenn Sie die Beispielprogramme und Beispielmultimediateien installieren wollen.
- Wählen Sie **Online-Dokumentation** aus, wenn Sie die Extender-Veröffentlichungen im Softcopy-Format installieren wollen. Wenn Sie diese Auswahl getroffen haben, können Sie auf den Knopf Ändern im Beschreibungsabschnitt klicken, um die verfügbaren sprachspezifischen

Installation auf Windows-Systemen

Veröffentlichungen anzuzeigen. Wählen Sie die für die Installation gewünschten Veröffentlichungen aus der Liste aus.

- Wählen Sie **Text Extender-Wörterverzeichnisse** aus, wenn Sie die linguistische Unterstützung für den DB2 Text Extender in ausgewählten Sprachen installieren möchten. Wenn Sie diese Auswahl getroffen haben, können Sie auf den Knopf Ändern im Beschreibungsabschnitt klicken, um die verfügbaren sprachspezifischen Wörterverzeichnisse anzuzeigen. Wählen Sie die für die Installation gewünschten Wörterverzeichnisse aus der Liste aus. Amerikanisches Englisch ist immer erforderlich.

Falls Sie nicht über genügend Plattenspeicherplatz für die ausgewählten Komponenten verfügen, drücken Sie den Knopf Zurück, um das Installationsverzeichnis erneut auszuwählen. Drücken Sie den Knopf Weiter, um fortzufahren.

8. Wenn Sie den Extender-Server in einer Umgebung mit partitionierten Datenbanken installieren und sich auf einer Exemplareignermaschine befinden, werden Sie dazu aufgefordert, einen TCP/IP-Anschlußbereich anzugeben. Auf allen Knoten, die dem aktuellen DB2-Exemplar zugeordnet sind, muß der gemeinsame Name existieren.

Wenn Sie die DB2 Extender auf einer Maschine installieren, die keine Exemplareignermaschine ist, werden Sie dazu aufgefordert, den Namen der Exemplareignermaschine einzugeben. Der TCP/IP-Anschlußbereich und die Datenpfadinformationen werden von der Exemplareignermaschine abgerufen.

9. Klicken Sie auf den Knopf Weiter, um die ausgewählten Komponenten zu installieren.

Wenn das Installationsprogramm gestartet wird, überprüft es, ob die Workstation die Softwarevoraussetzungen für die DB2 Extender erfüllt.

10. Fahren Sie die Workstation herunter und starten Sie sie erneut.

Exemplare erstellen und verwalten

Sie können mehrere Exemplare des DB2 Extender-Servers erstellen und verwenden. Mehrere Exemplare sollten Sie erstellen, wenn Sie mehrere Exemplare des DB2-Servers für Windows erstellt haben. Jedes Exemplar des DB2 Extender-Servers wird einem Exemplar des DB2-Servers für Windows zugeordnet.

Bei der Installation eines DB2 Extender unter Windows in einer Umgebung mit nicht partitionierten Datenbanken wird ein standardmäßiges DB2 Extender-Exemplar mit dem Namen DB2 erstellt.

Bei der Installation der DB2 Extender unter Windows in einer Umgebung mit partitionierten Datenbanken wird ein standardmäßiges DB2 Extender-Exemplar mit dem Namen DB2MPP erstellt. In einer Umgebung mit partitio-

nierten Datenbanken wird für den DB2 Text Extender kein Standardexemplar erstellt. Verwenden Sie den Befehl `txicrt`, um ein Exemplar zu erstellen. Weitere Informationen finden Sie im Abschnitt „Ein DB2 Text Extender-Exemplar erstellen“ auf Seite 22.

Eine Beispieldatenbank zur Installationsprüfung vorbereiten

Der DB2 Text Extender bietet Dienstprogramme zum Vorbereiten einer Beispieldatenbank. Sie sind nützlich, um sofort nach der Installation rasch Text zum Testen der Suchfunktionen des DB2 Text Extender vorzubereiten. Zudem dienen sie für allgemeine Testzwecke zu einem späteren Zeitpunkt.

Diese Dienstprogramme befinden sich in den folgenden Verzeichnissen:

- Unter OS/2, Windows NT und Windows 2000: `DMBMMPATH\samples` wobei `DMBMMPATH` der Installationspfad ist.
- Unter AIX, HP-UX und SUN-Solaris: `$HOME/db2tx/samples`

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

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 auf einer Client-Workstation arbeiten. Dieser Befehl führt folgende Schritte aus:

Er stellt eine Verbindung zur angegebenen Datenbank her.

Er erstellt die Tabelle `db2tx.sample`.

Er importiert englische Beispieldokumente zum Auffüllen der Tabelle.

Er aktiviert die Textspalte in der Tabelle mit den folgenden Indextypen:

präzise (PRECISE)

präzise normalisiert (PRECISE NORMALIZED)

linguistisch (LINGUISTIC)

linguistisch abschnittsfähig (LINGUISTIC SECTIONS_ENABLED)

Ngram (NGRAM)

Ngram unter Beachtung der Groß-/Kleinschreibung (NGRAM

CASE_ENABLED)

Er wartet darauf, daß der Textindex erstellt wird.

Konfiguration

In diesem Abschnitt werden die DB2 Text Extender-Umgebungsvariablen und -Konfigurationsdaten beschrieben. Mit beiden können Sie Standardwerte für viele Parameter angeben, die vom DB2 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 106 beschrieben wird.

DB2INSTANCE	Der DB2-Exemplarname.
DB2DBDFT	Name der Standarddatenbank. Der Name der DB2-Datenbank, die angenommen wird, wenn kein Datenbankname angegeben ist.
DB2TX_INSTOWNER	Der DB2 Text Extender-Exemplarname. Dies ist der Anmeldenname des Benutzers, der Eigner des Exemplars ist. Er ist erforderlich, wenn ein DB2 Text Extender-Client eine Verbindung zu einem UNIX-Server herstellt.
DB2TX_INSTOWNERHOMEDIR	Das Ausgangsverzeichnis des Exemplareigners (nur UNIX).

In einer Windows-Umgebung sind die DB2 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 DB2 Text Extender-Umgebungsvariablen während der Installation zur Datei `CONFIG.SYS` hinzugefügt.

In einer UNIX-Umgebung werden die DB2 Text Extender-Umgebungsvariablen (mit Ausnahme von `DB2DBDFT`) in den folgenden Profil-Shell-Prozeduren angegeben:

db2txcshrc für die C-Shell

db2txprofile für die Korn-Shell und die Bourne-Shell

Wenn Sie ein Exemplar erstellen, wird die Datei `db2txprofile` bzw. die Datei `db2txcshrc` in `.profile` kopiert.

Wenn Sie ein Exemplar erstellen, wird `db2txprofile` zu `.profile` und zu `db2profile` des Exemplareigners hinzugefügt. Die Shell-Prozedur wird unter `$DB2TX_INSTOWNERHOMEDIR/db2tx` gespeichert.

Sie benötigen außerdem die folgenden Umgebungsvariablen:

LANG	Gibt die Spracheinstellung an.
NLSPATH	Gibt den Pfad für die DB2 Text Extender-Nachrichtendatei an.
DB2ENVLIST	Liste von Umgebungsvariablen, die von DB2 an die DB2 Text Extender-Funktionen übergeben werden.
RAHUSER	Gibt den Namen des Exemplareigners an, der für den DB2-Befehl <code>db2a11</code> 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 Profil `db2tx` vom Exemplareigner ein, anstelle zu versuchen, diese Umgebungsvariablen manuell zu setzen.

Hinweis

Wenn Sie Umgebungsvariablen ändern, müssen Sie DB2 und den DB2 Text Extender stoppen und erneut starten, damit diese Änderungen wirksam werden.

Textkonfigurationseinstellungen

Alle Datenbanken verfügen über Konfigurationseinstellungen, die aus folgenden Elementen bestehen:

- Textmerkmale
- Indexmerkmale
- Verarbeitungsmerkmale

Diese Merkmale werden festgelegt, wenn die Datenbank für die Verwendung durch den DB2 Text Extender aktiviert wird. Durch den Befehl `ENABLE DATABASE` werden entweder die von Ihnen in diesem Befehl angegebenen Einstellungen oder die hier beschriebenen Anfangswerte verwendet. Sie können diese Standardeinstellungen anzeigen und ändern. Siehe „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 107 und „Die Textkonfiguration ändern“ auf Seite 21.

Konfiguration

Die Textmerkmale

„Kapitel 4. Die Sucherfordernisse planen“ auf Seite 31 beschreibt die Dokumentformate, -sprachen und -CCSIDs, die vom DB2 Text Extender unterstützt werden. Für diese Einstellungen werden von verschiedenen Befehlen Standardwerte benötigt.

FORMAT Anfangswert: TDS

LANGUAGE Anfangswert: Die Angabe LANGUAGE, die für die Datenbank gesetzt wurde

CCSID Anfangswert: Die Angabe CCSID, die für die Datenbank gesetzt wurde

Die Indexmerkmale

DIRECTORY Verzeichnis, das zum Speichern des Indexes verwendet werden soll.

Anfangswert für UNIX-Systeme:

DB2TX_INSTOWNERHOMEDIR/db2tx/indexes.

Anfangswert für OS/2, Windows NT und Windows 2000:

DMBMMPATH\instance\exemplarname \db2tx\indexes, wobei *DMBMMPATH* der Pfad ist, in dem der DB2 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 „Sucharten“ auf Seite 42 finden Sie dazu eine Beschreibung.

Anfangswert: NGRAM

UPDATEFREQ

Frequenz für die regelmäßige Indexaktualisierung. In „Einen Index aktualisieren“ auf Seite 48 finden Sie dazu eine Beschreibung.

Anfangswert: NONE

Die Verarbeitungsmerkmale

UPDATEINDEX

Einstellung, mit der festgelegt wird, wann die erste Indexaktualisierung durchgeführt wird: entweder gleich während des Aktivierungsschritts, später gemäß den Einstellungen für die Aktualisierungsfrequenz (NOUPDATE) oder aufgrund eines expliziten Befehls UPDATE INDEX.

Anfangswert: UPDATE

COMMITCOUNT

Einstellung, mit der festgelegt wird, nach wie vielen Einfügungs- oder Aktualisierungsanweisungen der DB2 Text Extender eine DB2-Anweisung COMMIT ausgibt. Siehe „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 72.

Anfangswert: 0

Die Textkonfiguration ändern

Während der ersten Installation des DB2 Text Extender werden die Standardwerte für die Textkonfiguration festgelegt. Um die aktuellen Werte der Textkonfiguration aufzurufen, siehe „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 107.

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   /myfs/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
```

Einen Server konfigurieren und verwalten

Einen DB2 Text Extender-Server konfigurieren und verwalten

Der nächste Schritt bei einem neu installierten DB2 Text Extender besteht im Konfigurieren des DB2 Text Extender-Servers, damit dieser von DB2 Text Extender-Clients verwendet werden kann. Dies wird normalerweise durch den DB2 Text Extender-Server-Administrator vorgenommen und umfaßt folgende Schritte:

1. Ein DB2 Text Extender-Exemplar erstellen.
2. Einen DB2 Text Extender-Server starten.

Andere Aufgaben zur Verwaltung eines DB2 Text Extender-Servers sind im folgenden aufgeführt:

1. Indizes und aktivierte Datenbanken des DB2 Text Extender sichern und zurückschreiben.
2. Die Trace-Funktion ausführen.

Informationen zur Befehlssyntax sind in „Kapitel 10. Verwaltungsbefehle für den Server“ auf Seite 175 enthalten.

Ein DB2 Text Extender-Exemplar erstellen

Bevor Sie die Arbeit mit dem DB2 Text Extender beginnen können, müssen Sie ein DB2 Text Extender-Exemplar erstellen. Jedes Exemplar, das Sie erstellen, bietet eine isolierte Umgebung, 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-Exemplars, das dem Text-Extender-Exemplar zugeordnet werden soll.

Unter Windows NT und Windows 2000 wird während der Installation ein Standardexemplar erstellt. Um ein weiteres Exemplar zu erstellen, geben Sie folgenden Befehl ein:

```
txicrt db2_exemplarname
```

Um Exemplare in einer Umgebung mit mehreren Knoten zu erstellen, siehe „TXICRT“ auf Seite 176.

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

```
txidrop exemplarname db2_exemplarname
```

Um ein Exemplar unter Windows NT und Windows 2000 zu löschen, geben Sie folgenden Befehl ein:

```
txidrop db2_exemplarname
```

Um festzustellen, mit welchem Exemplar Sie arbeiten wollen, definieren Sie die entsprechende Umgebungsvariable durch Eingeben des folgenden Befehls:

```
set DB2INSTANCE = exemplarname
```

Stoppen Sie DB2 und den DB2 Text Extender anschließend und starten Sie sie erneut.

Einen DB2 Text Extender-Server starten und stoppen

Um den DB2 Text Extender-Server zu starten, melden Sie sich unter der Benutzer-ID des Exemplareigners an (nur auf UNIX-Systemen), und geben Sie anschließend folgenden Befehl ein:

```
txstart
```

Um den Status des DB2 Text Extender-Servers anzuzeigen, geben Sie folgenden Befehl ein:

```
txstatus
```

Um den DB2 Text Extender-Server zu stoppen, geben Sie den folgenden Befehl ein:

```
txstop
```

Hinweis für UNIX-Benutzer

Wenn Sie beim Starten eines DB2 Text Extender-Servers Probleme haben, sollten Sie gemeinsam benutzte Ressourcen bereinigen. Siehe „TXIPCLEAN (nur UNIX)” auf Seite 180.

Indizes und aktivierte Datenbanken sichern und zurückschreiben

Sie können aktivierte Datenbanken und Textindizes, die der DB2 Text Extender erstellt hat, sichern und zurückschreiben.

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

1. Stellen Sie fest, welche Tabellen durch den DB2 Text Extender aktiviert wurden. Geben Sie dazu folgenden Befehl ein:

```
db2tx GET STATUS
```

Einen Server konfigurieren und verwalten

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

```
db2tx GET INDEX SETTINGS tabellenname
```

3. Stoppen Sie den DB2 Text Extender-Server. Geben Sie dazu folgenden Befehl ein:

```
TXSTOP
```

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  
    DMBMPATH\db2tx\txins000 (OS/2, Windows NT und Windows 2000)
```

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

6. Starten Sie den DB2 Text Extender-Server erneut:

```
TXSTART
```

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

1. Stoppen Sie den DB2 Text Extender-Server:
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 DB2 Text Extender-Server erneut:

```
TXSTART
```

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 192 beschrieben. Eine Beschreibung weiterer Optionen finden Sie ebenfalls dort.

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. B. 2.2-6.1,3.*, wobei die Zahlen folgende Bedeutung haben:

- 2** gibt den DB2 UDB DB2 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
```

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

```
txtrace format name_der_speicherauszugsdatei name_der_formatierten_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
```

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

```
txtrace off
```

Einen Server konfigurieren und verwalten

Kapitel 3. Erste Schritte

Verwenden Sie dieses Kapitel, um sich mit den Grundlagen dafür vertraut zu machen, wie Text für Suchvorgänge vorbereitet werden kann. Hierbei wird davon ausgegangen, daß Sie mit einem aktiven DB2 Text Extender-System arbeiten, das bereits installiert und konfiguriert wurde und für das ein DB2 Text Extender-Exemplar erstellt und gestartet wurde.

Hinweis

In diesem Kapitel werden nur die Grundlagen dafür beschrieben, wie Text für Suchvorgänge vorbereitet werden kann. Lesen Sie den Abschnitt „Vor Beginn des Vorbereitens von Text für Suchvorgänge“ auf Seite 59, bevor Sie Ihren eigenen Text für die Suche vorbereiten.

Es gibt zwei Arten, wie Sie sich mit dem Thema vertraut machen können, wie Text für Suchvorgänge vorbereitet werden kann:

- Sie können ein einfaches Beispiel durchführen und dabei die Befehle manuell eingeben.
- Sie können einen Beispielasistenten für die Indexierung und die Suche ausführen.

Ein einfaches Beispiel dafür, Text für Suchvorgänge vorzubereiten

1. Starten Sie auf Windows-Systemen den DB2-Befehlszeilenprozessor.

Geben Sie in der Eingabeaufforderung des Betriebssystems folgenden Befehl ein:

```
=>db2 cmd
```

2. Starten Sie den DB2 Text Extender-Befehlszeilenprozessor.

Geben Sie auf UNIX-Systemen in der Eingabeaufforderung des Betriebssystems folgenden Befehl ein:

```
=>db2tx
```

Geben Sie auf Windows-Systemen in der DB2-Eingabeaufforderung den folgenden Befehl ein:

```
db2=>db2tx
```

3. Stellen Sie eine Verbindung zu einer Datenbank her.

Wählen Sie eine Datenbank aus, die Text enthält, den Sie für Suchvorgänge vorbereiten wollen. Geben Sie den folgenden Befehl ein, um eine Verbindung zu der betreffenden Datenbank herzustellen:

Erste Schritte

```
db2tx=>CONNECT TO mydatabase
```

4. Aktivieren Sie die Datenbank für die Textsuche.

Geben Sie den folgenden Befehl ein, um die Datenbank zu aktivieren, zu der eine Verbindung hergestellt wurde:

```
db2tx=>ENABLE DATABASE
```

5. Aktivieren Sie eine Texttabelle für die Textsuche (wahlfrei).

Sie müssen nur dann eine Texttabelle aktivieren, wenn Sie für die gesamte Tabelle einen einzigen Index erstellen möchten.

Ein Index oder mehrere Indizes?

Im Abschnitt „Einen oder mehrere Textindizes für eine Tabelle erstellen“ auf Seite 46 wird erläutert, daß eine Tabelle für Suchvorgänge vorbereitet werden kann, indem entweder ein einziger Textindex für die gesamte Tabelle erstellt wird oder indem mehrere Indizes (einer pro Textspalte) erstellt werden.

- Um einen einzigen Index für die gesamte Tabelle zu erstellen, führen Sie an dieser Stelle den Befehl `ENABLE TEXT TABLE` aus, um einen leeren Index zu erstellen. Anschließend führen Sie den Befehl `ENABLE TEXT COLUMN` (siehe nächster Schritt) mehrmals aus, für jede Textspalte ein Mal, um den einzelnen Index zu füllen.
- Sie können alternativ dazu für jede Textspalte einen eigenen Index erstellen, indem Sie den Schritt zur Erstellung eines Indexes für die gesamte Tabelle überspringen (d. h., Sie führen den Befehl `ENABLE TEXT TABLE` nicht aus) und anschließend den Befehl `ENABLE TEXT COLUMN` mehrmals ausführen, um Indizes zu erstellen und diese zu füllen (für jede Textspalte einen).

Im vorliegenden Beispiel wird kein Index für die gesamte Tabelle erstellt. Fahren Sie statt dessen mit dem nächsten Schritt fort, um einen Index für eine Textspalte zu erstellen.

6. Aktivieren Sie eine Textspalte für die Textsuche.

Geben Sie den folgenden Befehl ein, um den DB2 Text Extender zur Suche in der Textspalte `mycolumn` in der Tabelle `db2tx.sample` zu aktivieren und den Namen `myhandle` zu der Kennungsspalte zuzuordnen, die durch diesen Befehl erstellt wird.

```
db2tx=>ENABLE TEXT COLUMN db2tx.sample mycolumn HANDLE myhandle
```

Durch diesen Befehl wird ein Textindex erstellt. Für den zu indizierenden Typ von Dokumenten und für die Indexmerkmale werden Standardwerte verwendet.

7. Prüfen Sie den Status des Indexes, den Sie gerade erstellen.

Geben Sie folgenden Befehl ein:

```
db2tx=>GET INDEX STATUS db2tx.sample HANDLE myhandle
```

8. Verlassen Sie den DB2 Text Extender-Befehlszeilenprozessor.

Geben Sie folgenden Befehl ein:

```
db2tx=>QUIT
```

9. Starten Sie den DB2-Befehlszeilenprozessor.

Geben Sie in der Eingabeaufforderung des Betriebssystems folgenden Befehl ein:

```
=>db2
```

10. Suchen Sie nach Text.

Jetzt können Ihre Dokumente durchsucht werden. Versuchen Sie es mit dem nachfolgend aufgeführten Befehl SELECT, durch den alle Stellen gefunden werden, an denen searchterm in dem Text vorkommt, den Sie gerade indexiert haben:

```
db2=>SELECT COUNT (*)
        FROM sample
        WHERE DB2TX.CONTAINS (myhandle, 'searchterm') = 1
```

Der Beispiellassistent für die Indexierung und die Suche

Ein Indexierungs- und Suchassistent wird als Beispiel dafür bereitgestellt, wie Indizes für ausgewählte Tabellen oder Spalten erstellt und diese Indizes durchsucht werden. Dieser Beispiellassistent führt Sie durch die Schritte, die für die Suche in ausgewählten Spalten einer Tabelle unter Verwendung der Suchfunktionen des DB2 Text Extender erforderlich sind.

Zum Starten des Beispiellassistenten geben Sie folgenden Befehl ein:

```
txwizard
```

Sie können den Java-Quellcode verwenden, der mit dem Beispiellassistenten geliefert wird, um entweder Ihre eigene Textvorbereitungs- und Suchanwendung zu erstellen oder zu lernen, wie die DB2 Text Extender-Funktionen verwendet und innerhalb einer Java-Umgebung aufgerufen werden können.

Unter AIX ist der Java-Quellcode des Assistenten in folgendem Verzeichnis gespeichert:

```
/usr/lpp/db2tx_06_01/samples/wizard
```

Unter Windows NT und Windows 2000 ist er in folgendem Verzeichnis gespeichert:

```
%DMBMPATH%\samples\wizard
```

Kapitel 4. Die Sucherfordernisse planen

Bevor Sie die in „Kapitel 5. Text für Suchvorgänge vorbereiten“ auf Seite 59 aufgeführten Schritte ausführen, müssen Sie folgendes herausfinden:

- In welchem Format, welcher Codepage und welcher Sprache Ihre Dokumente vorliegen
- Wie Codepage-Probleme vermieden werden können
- Welche Art von Suchfunktionalität benötigt wird
- Welcher Plattenspeicherplatz benötigt wird
- Was ein Textindex ist und ob ein gemeinsamer Index für eine DB2-Tabelle oder ein separater Index für jede Tabellenspalte verwendet werden soll

In diesem Kapitel wird beschrieben, wie und warum diese Informationen gesammelt werden sollten.

Es gibt verschiedene Indextypen, aus denen Sie wählen können: linguistisch, präzise 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 indiziert 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. B. „und“, „von“ und „der/die/das“, werden nicht indiziert. (Bei Ngram-Indizes wird kein Stoppwortfiltern durchgeführt.) Der DB2 Text Extender verwendet eine Liste dieser Wörter, bekannt als *Stoppwörter*, um zu verhindern, daß sie indiziert werden. Das Abfragesystem durchsucht den Index nach den angeforderten Begriffen, um zu ermitteln, welche Textdokumente diese Begriffe enthalten.

Warum Textdokumente indexiert sein müssen

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 58). Da aber nur eine Datei für das gesamte System zur Verfügung steht, sollten Sie sie nur einmal ändern, wenn Sie den DB2 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.

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.

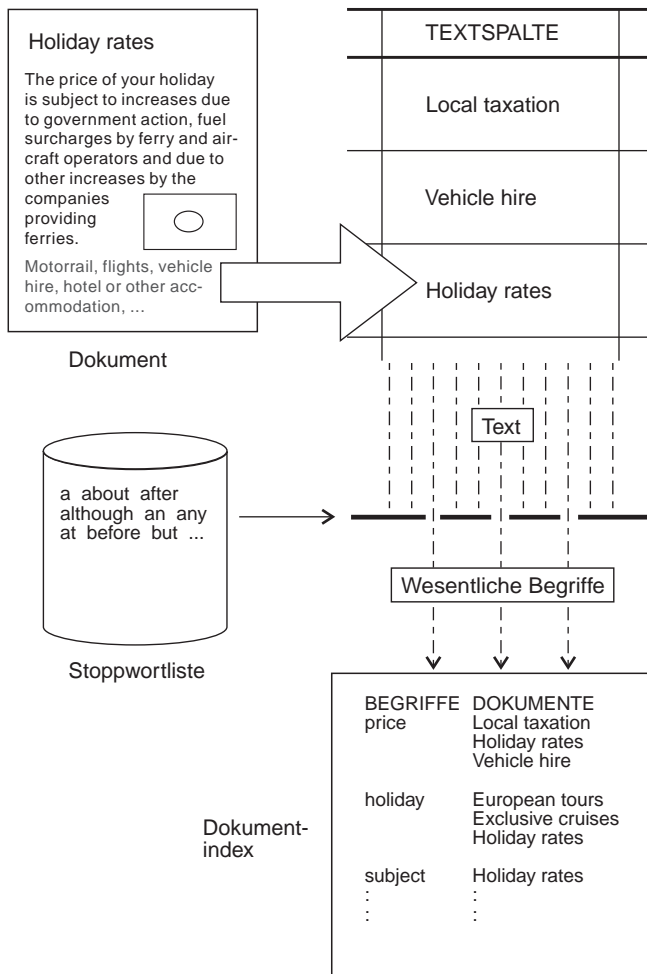


Abbildung 4. Nur wesentliche Begriffe werden indexiert

Unterstützte Dokumentformate

Der DB2 Text Extender muß das Format (oder den Typ) von Textdokumenten kennen, die Sie durchsuchen wollen, z. B. HTML oder ASCII. Diese Informationen werden beim Indexieren von Textdokumenten benötigt.

Folgende Textdokumentformate werden unterstützt:

HTML	Hypertext Markup Language
XML	Extended Markup Language
ASCII_SECTIONS	ASCII strukturiert, mit Abschnitten
TDS	ASCII unstrukturiert

Dokumentformate

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

Spezielle Überlegungen zu HTML-Dokumenten

Wie Umlaute und diakritische Zeichen in HTML-Dokumenten behandelt werden, hängt von der Codepage des betreffenden Dokuments ab:

- Für die Codepages 37, 273, 277, 278, 280, 284, 297, 437, 500, 819, 850, 858, 860, 863, 865, 871, 923, 924 und 1252 gilt folgendes:
 - Für Umlaute und Sonderzeichen wird Entitätsschreibweise verwendet.
Beispiel:
ä für ä.
 - Nur die Zeichen sind gültig, die in Codepage 819 (ASCII) oder 500 (EBCDIC) einen Codepunkt haben.
 - Während der Indexierung führen Dokumente, die sprachspezifische Zeichen wie ä enthalten, zu Fehlern bei der Worterkennung, wenn die Codepage des betreffenden Dokuments weder 819 noch 500 ist.
 - Wenn Sie Dokumente zu einem Ngram-Index hinzufügen, muß der Index unter Verwendung der Codepage 819, 500 oder UTF8 erstellt worden sein.
- Für alle anderen Codepages für Einzelbytezeichensätze gilt folgendes:
 - Entitäten werden nicht aufgelöst.
 - Sonderzeichen müssen in sprachspezifischen Codepunkten geschrieben werden.

Spezielle Überlegungen zu XML-Dokumenten

XML-Dokumente müssen eine Verschlüsselungszeichenfolge enthalten, die während der Syntaxanalyse verwendet wird. Die folgenden Verschlüsselungen werden unterstützt:

UTF8, utf8, utf-8, ibm-1208, utf_8,
UTF16_BigEndian, utf-16be, utf16
UTF16_LittleEndian, utf-16le
LATIN_1, latin1, latin-1, ascii, ibm-819, iso-8859-1, Latin-1
ibm-912, iso-8859-2
ibm-913, iso-8859-3
ibm-914, iso-8859-4
ibm-915, iso-8859-5

ibm-1089, iso-8859-6
 ibm-813, iso-8859-7
 ibm-916, iso-8859-8i
 ibm-920, iso-8859-9
 ibm-0037, ebcdic-cp-us, ebcdic-cp-ca, ebcdic-cp-nl, ebcdic-cp-dk, ebcdic-cp-no,
 ebcdic-cp-fi, ebcdic-cp-se, ebcdic-cp-it, ebcdic-cp-es, ebcdic-cp-gb
 ibm-297, ebcdic-cp-fr, ebcdic-cp-ar1, ebcdic-cp-he, ebcdic-cp-ch, ebcdic-cp-roece,
 ebcdic-cp-yu, ebcdic-cp-is, ebcdic-cp-ar2
 ibm-954, euc-jp eucjis
 ibm-943, shift_jis, sjis, shiftjis, shift-jis
 ibm-950 , big_5, big5
 ibm-949, iso2022kr, euc-kr
 ibm-878, koi8-r

Die Standard-Codepage wird bei XML-Dokumenten ignoriert.

Wenn Sie XML-Dokumente zu einem Ngram-Index hinzufügen, muß dieser unter Verwendung der Codepage UTF8 erstellt worden sein. Wenn der Index über Abschnittsunterstützung verfügt, wird bei der Indexierung nur die Abschnittskennung auf Blattebene verwendet. Es gibt keine Unterstützung für die Auflösung von Text in verschachtelten Abschnitten.

Nicht unterstützte Dokumentformate verwenden

Geben Sie für nicht unterstützte Dokumentformate 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 Typ nicht unterstützt wird, bietet der DB2 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. Der Benutzer-Exit muß sowohl in der Server-Konfigurationsdatei (DESSR.VINI) als auch in der Client-Konfigurationsdatei (DESCL.INI) registriert sein. Aktualisieren Sie die Option USEREXIT im Abschnitt [DOCUMENTFORMAT] mit dem Namen des Benutzer-Exit.

Um den Benutzer-Exit zu aktivieren, editieren Sie die folgenden ASCII-Dateien:

Windows NT und Windows 2000:
 %DMBMPATH%\instance\%DB2INSTANCE%\db2tx\desc1.ini
 %DMBMPATH%\instance\%DB2INSTANCE%\db2tx\txinsnnn\dessrv.ini

UNIX:
 \$DB2TX_INSTOWNERHOMEDIR/db2tx/desc1.ini
 \$DB2TX_INSTOWNERHOMEDIR/db2tx/txinsnnn/dessrv.ini

Fügen Sie zu diesen Dateien jeweils die folgenden Anweisungen hinzu:

```
[DOCUMENTFORMAT]
USEREXIT=name_der_ausführbaren_datei
```

Dokumentformate

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.

Verwenden Sie die folgende Syntax, um den Benutzer-Exit aufzurufen:

```
<name_des_benutzer_exit> -sourcefile <quellendateiname>
                          -targetfile <zieldateiname>
                          -sourceccsid <quellen_ccsid>
                          -targetccsid <ziel_ccsid>
                          -sourceformat <quellenformat>
                          -targetformat <zielformat>
```

quellendateiname

Die vom Benutzer-Exit-Programm umzuwandelnde Datei. Der Dateiname ist vollständig qualifiziert und befindet sich im Arbeitsverzeichnis, das entweder im Client-Profil oder im Server-Exemplar angegeben wurde.

zieldateiname

Die Datei, die die Ausgabe des Benutzer-Exit enthält. Diese Datei wird anschließend für die Verarbeitung durch den DB2 Text Extender verwendet. Der Dateiname ist vollständig qualifiziert und zeigt auf das Arbeitsverzeichnis, das entweder im Client-Profil oder im Server-Exemplar angegeben wurde. Die Einträge im Client-Profil werden für den API-Aufruf EhwGetMatches, die im Server-Exemplar für den API-Aufruf EhwUpdateIndex verwendet.

quellen_ccsid

Die Codepage der Quellendatei. Dies ist die Standard-Codepage.

ziel_ccsid

Die vom DB2 Text Extender erwartete Codepage. Diese Codepage ist 850.

quellenformat

Das Format der Quellendatei. Dies ist das Standardformat.

zielformat

Das Format der vom DB2 Text Extender erwarteten Datei. Momentan werden nur das unstrukturierte Dateiformat (TDS) oder, bei abschnittsfähigen Indizes, ASCIISECTION unterstützt.

Der Benutzer-Exit muß die folgenden Werte zurückgeben können:

- 0** Formatumwandlung verlief erfolgreich.
- >0** Formatumwandlung verlief nicht erfolgreich. Während der Indexierung werden die Fehlnachrichten in die Dokumentfehltabelle geschrieben. Verwenden Sie den Befehl `desmsgix` zum Anzeigen der Fehlnachrichten.

Sprachen

Der DB2 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. In Tabelle 5 auf Seite 57 wird eine Liste der Sprachparameter gezeigt, die beim Aktivieren einer Textspalte oder externer Dokumente angegeben werden können.

CCSIDs
Hinweis

Lesen Sie vor dem Angeben einer CCSID beim Aktivieren einer Textspalte die Informationen in Abschnitt „Codepage-Probleme beim Speichern und Aktivieren von Text vermeiden“ auf Seite 40.

Dokumente können indiziert werden, wenn sie sich in einer der CCSIDs in der nachfolgend aufgeführten Tabelle befinden.

Anmerkung: Die CCSIDs 861, 865 und 4946 werden nicht von DB2 unterstützt. Um Dokumente, die diese CCSIDs haben, zu indizieren, speichern Sie die Dokumente in einer Spalte mit einem binären Datentyp (BLOB oder FOR BIT DATA).

EBCDIC

37	Amerikanisches Englisch (Kanada)
273	Deutsch
277	Dänisch, Norwegisch
278	Finnisch, Schwedisch
280	Italienisch
284	Spanisch (Lateinamerika)
285	Britisches Englisch
297	Französisch
420	Arabisch
424	Hebräisch
437	Amerikanisches Englisch
500	Internationales lateinisches Alphabet 1
871	Isländisch

CCSIDs

875	Griechisch
1025	Russisch
1112	Lettisch
1122	Estnisch
1123	Ukrainisch

ASCII

813 AIX, HP, SUN	Griechisch
819 AIX, HP, SUN	Lateinisches Alphabet 1
850 AIX, OS/2	Lateinisches Alphabet 1
855 OS/2	Bulgarisch
860 OS/2	Portugiesisch
861 Siehe Anmerkung	Isländisch
862 OS/2	Hebräisch
864 OS/2	Arabisch
863 OS/2	Kanadisch
865 Siehe Anmerkung	Dänisch, Norwegisch
866 OS/2	Russisch
869 OS/2	Griechisch
915 AIX, OS/2, HP	Russisch
916 AIX	Hebräisch
921 AIX, OS/2, WIN	Lettisch
922 AIX, OS/2, WIN	Estnisch
1064 AIX	Arabisch
1089 AIX, HP	Arabisch
1124 AIX	Ukrainisch
1125 OS/2	Ukrainisch
1129	Vietnamesisch
1130	Vietnamesisch
1131	Vietnamesisch
1250 WIN	Kroatisch, Weißrussisch
1251 WIN	Russisch

1252 WIN	Lateinisches Alphabet 1
1253 WIN	Tschechisch
1255 WIN	Hebräisch
1256 WIN	Arabisch
1257 WIN	Griechisch
1258 WIN	Vietnamesisch

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

1208	UTF8
13488	UCS2

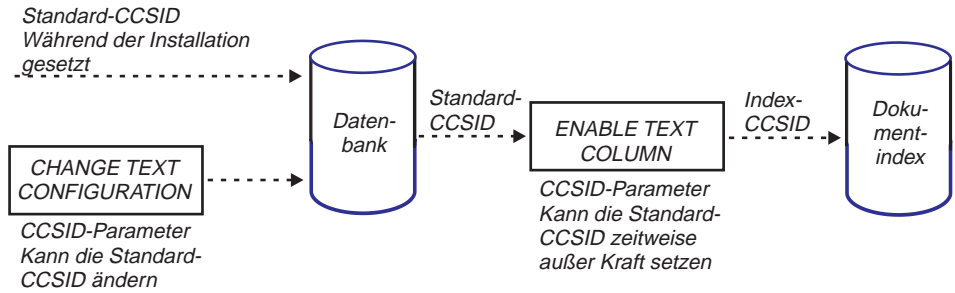


Abbildung 6. Setzen der CCSID eines Textindexes durch den DB2 Text Extender

Während der Suche wird die CCSID der Datenbank verwendet, um die CCSID des Suchbegriffs zu interpretieren.

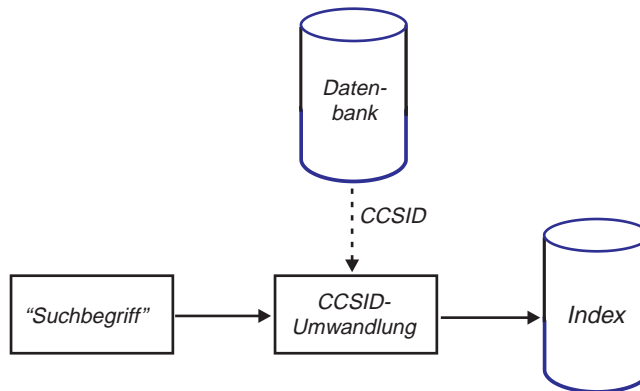


Abbildung 7. Bei der Suche wird die Datenbank-CCSID verwendet

Im folgenden ist aufgeführt, wie die Codepage-Einstellungen geprüft werden können:

- Verwenden Sie den folgenden DB2-Befehl, um die Datenbank-Codepage zu prüfen:
db2 get db cfg
- Verwenden Sie den folgenden DB2 Text Extender-Befehl, um die Standardindex-Codepage zu prüfen:
db2tx get text cfg

Um Codepage-Probleme zu vermeiden, ist es wichtig, diese Codepage-Einstellungen richtig zu koordinieren:

- **Beispiel 1:** Umgebung 850, Dokument 850, Datenbank 850, Index 850
Richtig. DB2 geht richtigerweise davon aus, daß das Dokument dieselbe Codepage wie die Umgebung hat. Da die Datenbank dieselbe Codepage wie die Umgebung hat, nimmt DB2 keine Umwandlung vor und speichert

das Dokument mit Codepage 850 in der Datenbank. Wenn Sie anschließend die Textspalte aktivieren und die Standard-Codepage-Einstellung (850, die Codepage der Datenbank) verwenden, wird das Dokument korrekt in den 850-Codepage-Index indiziert.

- **Beispiel 2:** Umgebung 1252, Dokument 1252, Datenbank 850, Index 850
Richtig. DB2 geht richtigerweise davon aus, daß das Dokument dieselbe Codepage wie die Umgebung hat. Beim Speichern des Dokuments in der Datenbank wird eine 1252-zu-850-Codepage-Umwandlung vorgenommen. Wenn Sie anschließend die Textspalte aktivieren und die Standard-Codepage-Einstellung 850 (die Codepage der Datenbank) verwenden, wird das Dokument korrekt in den 850-Codepage-Index indiziert.
- **Beispiel 3:** Umgebung 1252, Dokument 850, Datenbank 850 (Index ANY)
Fehler. DB2 geht fälschlicherweise davon aus, daß das Dokument (Codepage 850) dieselbe Codepage wie die Umgebung (Codepage 1252) hat. Beim Speichern des Dokuments in der Datenbank wird eine falsche 1252-zu-850-Codepage-Umwandlung vorgenommen.
- **Beispiel 4:** Umgebung 1252, Dokument 850, Datenbank 1252, Index 850
Richtig. DB2 geht fälschlicherweise davon aus, das das Dokument dieselbe Codepage wie die Umgebung hat. Da aber die Codepage der Umgebung dieselbe wie die der Datenbank ist, nimmt DB2 keine Umwandlung vor und speichert das Dokument in Codepage 850 in der Datenbank. Wenn Sie anschließend die Textspalte aktivieren, müssen Sie jedoch für Dokumente Codepage 850 angeben, damit das jeweilige Dokument korrekt in den 850-Codepage-Index indiziert wird.
- **Beispiel 5:** Umgebung 1252, Dokument 1252, Datenbank 850, Index 1252
Möglicher Fehler. DB2 geht richtigerweise davon aus, daß das Dokument wie auch die Umgebung die Codepage 1252 hat. Das Dokument wird beim Speichern in der Datenbank in Codepage 850 umgewandelt. Der potentielle Fehler tritt auf, wenn Sie anschließend die ursprüngliche Codepage (1252) des Dokuments angeben, wenn Sie die Textspalte für den Codepage-850-Index aktivieren. Die richtige Vorgehensweise bestünde darin, die Standard-Codepage-Einstellung 850 (die Codepage der Datenbank) zu verwenden.

Sucharten

Sie können zu einer Spalte, die zu durchsuchenden Text enthält, einen der folgenden Indextypen und verschiedene Optionen zuordnen: *linguistisch*, *präzise* und *Ngram*. Sie müssen entscheiden, welcher Indextyp erstellt werden soll, bevor Sie solche Spalten für die Verwendung durch den DB2 Text Extender vorbereiten. Eine detaillierte Beschreibung, wie die einzelnen Indextypen die Verarbeitung auf linguistischer Basis beeinflussen, finden Sie in „Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis“ auf Seite 257.

Zusammenfassung

- Für die **Suche nach linguistischen Wortvariationen** wird ein **linguistischer Index** verwendet.
Dabei werden Wortvariationen auf der Basis von Normalisierung und Wortstambildung und unter Verwendung eines Wörterverzeichnisses gesucht; verwendet den geringsten Plattenspeicherplatz.
- Für die Durchführung **exakter Suchen** wird ein **präziser Index** verwendet.
Dabei werden Begriffe exakt so gesucht, wie sie eingegeben werden; die Indexierung und die Suche sind schneller; mehr Plattenspeicherplatz wird verwendet. Wenn die Option NORMALIZED angegeben ist, wird bei Suchvorgängen die Groß-/Kleinschreibung beachtet.
- Für die **Suche nach Zeichenvariationen** wird ein **Ngram-Index** verwendet.
Dabei werden selbst falsch geschriebene Wörter gefunden; wenn die Option CASE_ENABLED für Suchvorgänge unter Beachtung der Groß-/Kleinschreibung verwendet wird, benötigt der Index mehr Speicherbereich und die Suchvorgänge können länger dauern.
- Für die **Suchvorgänge in DBCS-Dokumenten** wird ein **Ngram-Index** verwendet.
Einzige Wahlmöglichkeit bei DBCS-Dokumenten; kann auch für SBCS-Dokumente vom Typ TDS verwendet werden.

Der DB2 Text Extender bietet eine breite Palette von Suchoptionen, wobei allerdings nicht alle für alle Indextypen verfügbar sind. Prüfen Sie Tabelle 8 auf Seite 225 und Tabelle 9 auf Seite 226, bevor Sie sich entscheiden, welcher Indextyp verwendet werden soll.

Linguistische Suche

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 18 auf Seite 258 faßt zusammen, wie Begriffe für das Indexieren extrahiert werden, wenn Sie den linguistischen Index verwenden.

Sucharten

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. 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. 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.

Variationen von Wörter, die nicht richtig geschrieben sind, können nicht auf eine Grundform reduziert werden. Ein Beispiel hierfür ist das Wort röstete, das korrekt in seiner Grundform, dem Verb rösten, indexiert ist. Ein Suchbegriff wie röstete oder rösteten wird auf korrekte Weise auf die Grundform rösten normalisiert, und der Begriff wird im Index gefunden. Ein Suchbegriff wie rostete kann nicht auf rösten normalisiert werden, wodurch der Begriff nicht im Index gefunden werden kann.

- 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 indiziert. “Ich fahre mit meinem Auto schnell” wird indiziert als fahren und auto. Die Wörter “Ich” und “meinem” werden als Stoppwörter gelöscht, aber zusätzlich wird das Adverb “schnell” durch das Wortartfiltern entfernt.

Präzise Suche

In einem präzisen Index werden die Begriffe in einem Textdokument exakt so indiziert, 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 19 auf Seite 258 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 Indizieren und Abfragen schneller ist. Da jede unterschiedliche Form und Schreibweise jedes Begriffs indiziert wird, wird mehr Plattenspeicherplatz benötigt als für den linguistischen Index.

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

Wort- und Satztrennung
Stoppwortfiltern

Suchen nach grober Übereinstimmung oder Suchen in DBCS-Dokumenten durchführen

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. h., Sie können Zeichenfolgen finden, die dem angegebenen Such-

Sucharten

begriff 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 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.

Eine Liste der von Ngram-Indizes unterstützten CCSIDs finden Sie im Abschnitt „CCSIDs“ auf Seite 37. Ein Ngram-Index unterstützt eine Liste von Basis-CCSIDs. Bei allen anderen CCSIDs werden die Daten von der entsprechenden CCSID zu UTF8 zugeordnet.

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 12. Syntax von Suchargumenten“ auf Seite 217.

Den Indextyp ändern

Wenn Sie feststellen, daß der verwendete Indextyp nicht geeignet ist, löschen Sie zuerst den betreffenden Index, indem Sie die Textspalte oder -tabelle inaktivieren, und erstellen Sie ihn anschließend erneut, indem Sie die Textspalte oder -tabelle erneut aktivieren.

Einen oder mehrere Textindizes für eine Tabelle erstellen

„Kapitel 5. Text für Suchvorgänge vorbereiten“ auf Seite 59 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

Einem oder mehreren Textindizes für eine Tabelle erstellen

gilt, oder ob Sie mehrere Textindizes erstellen, jeweils einen für jede indexierte Textspalte. Eine Tabelle, die einen separaten Index für jede Textspalte hat, wird als *Mehrindextabelle* bezeichnet.

Hinweis

Wenn Sie externe Dateien indexieren wollen (siehe „Externe Textdateien aktivieren“ auf Seite 75), muß die zugeordnete Tabelle eine Mehrindextabelle sein.

Das Verwenden von mehreren Indizes bringt folgende Vorteile mit sich:

- Erstellen eines unterschiedlichen Indextyps für jede Textspalte
Dadurch erlangen Sie Flexibilität in den Merkmalen, die einer Textspalte zugeordnet sind, wie z. B. wann der zugehörige Index regelmäßig aktualisiert wird und in welchem Verzeichnis der Index gespeichert wird. Eine Beschreibung dieser Merkmale finden Sie im Abschnitt „ENABLE TEXT COLUMN“ auf Seite 145.
- Indexieren von Spalten zu unterschiedlichen Zeitpunkten
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 Zeitpunkten indexieren.

Wenn Sie die Flexibilität, die eine Mehrindextabelle bietet, nicht benötigen, hat ein gemeinsamer Index den Vorteil, daß der DB2 Text Extender einfacher zu verwalten ist; wenn Sie eine Texttabelle aktivieren, definieren Sie die Indexierungsparameter, die als Standardwerte für alle zugehörigen Textspalten verwendet werden. Daneben können Sie die Spalten bei Bedarf mit Hilfe eines einzigen Befehls inaktivieren, indem Sie die Texttabelle inaktivieren.

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 Speicherbereich, da ihr Inhalt mit Steuerzeichen aufgebaut ist. Als Richtlinie gilt: reservieren Sie für einen linguistischen Index 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. Für einen Ngram-Index benötigen Sie fast zwei Mal so viel Plattenspeicherplatz.

Die Größe eines Indexes berechnen

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.

Einen Index aktualisieren

Wenn ein Textdokument zu einer Datenbank hinzugefügt wird oder wenn ein bestehendes Dokument in einer Datenbank geändert wird, muß das Dokument indiziert 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.

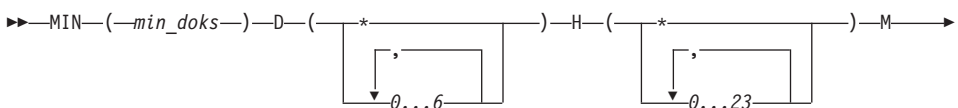
Informationen dazu, welche Dokumente neu, geändert oder gelöscht sind, werden automatisch durch Auslöser in einer Protokolltabelle gespeichert. Die in der Protokolltabelle aufgelisteten Dokumente werden das nächste Mal, wenn eine Indexaktualisierung stattfindet, indiziert.

Mit dem Befehl `UPDATE INDEX` können Sie einen Index direkt auf eine Anforderung hin aktualisieren. Normalerweise aktualisieren Sie einen Index jedoch 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 benötigte Zeit hängt von vielen Faktoren ab, z. 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.

Indizes in einer Umgebung mit mehreren Knoten verwalten

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 DB2 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 unter Windows NT und Windows 2000 den Befehl `NET TIME`, um die Systemzeiteinstellungen zu synchronisieren.

Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)

Abschnittsunterstützung ermöglicht das Indexieren und Suchen spezifischer Abschnitte in einem strukturierten Dokument, beispielsweise Titel, Autor oder Beschreibung. Die Dokumente können im XML-Format oder im HTML-Format bzw. als unstrukturierte Dokumente mit HTML-ähnlichen Befehlen vorliegen. Die Befehle der Formatierungssprache und die entsprechenden Abschnittsnamen werden in einem *Dokumentmodell* definiert. Das Dokumentmodell legt fest, welche Abschnitte in den betreffenden Dokumenten indexiert und daher für Suchvorgänge verfügbar sind. Die Abschnittsnamen sind beschreibend und werden in Abfragen für den jeweiligen Abschnitt verwendet.

Um die Abschnittsunterstützung verfügbar zu machen, müssen Sie die Option `INDEXPROPERTY SECTIONS_ENABLED` auswählen, wenn Sie die Textspalte, die die Dokumente enthält, aktivieren.

Strukturierte Dokumente (Abschnittsunterstützung)

In einer *Dokumentmodelldatei* werden alle für das Server-Exemplar definierten Dokumentmodelle aufgelistet. Wenn ein Server-Exemplar erstellt wird, wird ein Beispiel einer Dokumentmodelldatei, DESMODEL.INI, automatisch im Unterverzeichnis für das Server-Exemplar erstellt. Diese Datei liegt in der Codepage ASCII vor.

Die Dokumentmodelldaten werden in das Indexverzeichnis kopiert. Wenn Sie die Dokumentmodelldatei für das Server-Exemplar nach Erstellen des Indexes ändern, wirkt sich dies nicht auf die Abschnittsunterstützung für den erstellten Index aus.

Beispielsweise kann eine Suche in einem Index mit Abschnittsunterstützung nach McDaniel im Abschnitt 'Author' (Autor) wie nachfolgend dargestellt aussehen. Der betreffende Abschnitt, in diesem Fall 'Author', wird immer mit dem Präfix des Modellnamens versehen.

```
db2 "select count (*)
     from db2tx.htmltable
     where db2tx.contains(myhandle,'MODEL myhtmlmodel SECTION (author)
     "Schmidt") = 1
```

Attributabschnitte

Abschnitte können von verschiedenen Typen sein. Abschnitte mit unverschlüsseltem Text haben keinen Typ. Abschnitte mit einem deklarierten Typ werden als *Attributabschnitte* bezeichnet. Die unterstützten Attributtypen sind im folgenden aufgeführt:

- DATE
- TIME
- FLOAT
- INTEGER

Sie können Wertebereiche verwenden, um Dokumente zu durchsuchen, die Attributabschnitte enthalten. Der betreffende Abschnitt muß den entsprechenden Abschnittstyp haben. Im folgenden Beispiel muß der Abschnitt ABC den Typ TIME haben, damit die Abfrage unterstützt wird:

```
alle Dokumente mit Werten zwischen
20:00:00 und 22:00:00 im Abschnitt ABC abrufen
```

Ein Abschnitt kann über mehr als einen Attributwert verfügen. Werte werden durch Leerzeichen oder durch den Beginn einer neuen Zeile getrennt. Attributabschnitte können nicht verschachtelt werden.

Strukturierte Dokumente (Abschnittsunterstützung)

Unterstützte Datumsformate

Sie können Daten, wie beispielsweise das Erstellungs- oder Veröffentlichungsdatum, in die Liste mit unterstützten Abschnitten in einem Dokumentmodell aufnehmen. Die folgenden Datumsformate werden unterstützt:

Tabelle 3. Unterstützte Datumsformate

Formatname	Format	Beispiel
ISO	JJJJ-MM-TT	1986-12-25 100-12-25 10123-2-17
	TT-MM-JJJJ	01-10-1999 5-8-1990
IBM USA-Standard (USA)	MM/TT/JJJJ	08/18/1999 5/13/1977
	JJJJ/MM/TT	1999/08/18
IBM Europäischer Standard (EUR)	TT.MM.JJJJ	20.10.1999 1.5.1980

Die Suche in einem Index mit Abschnittsunterstützung nach allen Dokumenten, die an einem bestimmten Datum veröffentlicht wurden, könnte folgendermaßen aussehen:

```
dessrch -s SERVER -x TESTIX -section sample/PublishingDate -term 1999-09-09
```

Unterstützte Zeitformate

Die folgenden Zeitformate werden unterstützt:

Tabelle 4. Unterstützte Zeitformate

Formatname	Format	Beispiel
ISO	hh:mm:ss	19:00:23
	hh:mm	19:00
	hh.mm.ss	19.00.23
	hh.mm	19.00
	hh	19
	hhmm	1900
	hhmmss	190013

Führende Nullen können bei Stunden weggelassen werden. Werte, die sich außerhalb der folgenden Bereiche befinden, sind nicht gültig.

Stunden (h)	Minuten (m)	Sekunden (s)
0-23	0-59	0-59

Strukturierte Dokumente (Abschnittsunterstützung)

Unterstützte Gleitkommaformate

Die folgenden Formate werden unterstützt:

- Eine Ziffernfolge im Bereich von 1.175494351e-38F bis 3.402823466e+38F
- Wissenschaftliche Schreibweise mit e oder E.

Verwenden Sie Leerzeichen oder neue Zeilen, um Gleitkommawerte zu trennen.

Im folgenden sind Beispiele für richtige und falsche Formate für Gleitkommawerte aufgeführt:

Richtiges Format	Falsches Format
1000 1 1.0e3 1.0E3 1E3 .1E4 0.1E4	1 000 1,000 e3

Unterstützte Formate für ganze Zahlen

Ganze Zahlen im Bereich von -2147483648 bis 2147483647 werden unterstützt. Verwenden Sie Leerzeichen, um ganze Zahlen zu trennen. Dies bedeutet, daß die Zeichenfolge "42 234" in einem Abschnitt mit dem Typ INTEGER in die ganzzahligen Werte "42" und "234", nicht "42234", aufgelöst wird.

Im folgenden sind Beispiele für richtige und falsche Formate von ganzzahligen Werten aufgeführt:

Richtiges Format	Falsches Format
1000	1.000 1,000 1 000
1000000	100.000 100,00 100 000

Hinweis

Sprachspezifische Trennzeichen und Formate für Währungseinheiten werden nicht unterstützt.

Strukturierte Dokumente (Abschnittsunterstützung)

Unstrukturierte Dateien und HTML-Dokumente

Bei unstrukturierten Dateien werden die Abschnitte mit HTML-ähnlichen Befehlen formatiert, wie z. B. <title> und <subject>. Ein Dokument mit formatierten Abschnitten könnte folgendermaßen aussehen:

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

Eine Dokumentmodelldatei für unstrukturierte Dateien oder HTML-Dokumente könnte wie das nachfolgend aufgeführte Beispiel aussehen. Die Modell- und Abschnittsnamen können nur die Zeichen A-Z, a-z und 0-9 enthalten. Bei Modellnamen muß die Groß-/Kleinschreibung immer beachtet werden. Bei Abschnittsnamen kann die Groß-/Kleinschreibung zu beachten sein; diese Einstellung geben Sie beim Erstellen des betreffenden Modells an.

```
;list of document models
;model always starts with 'modelname' and the name of the model
[MODELS]
modelname=sample
modelname=sample2
modelname=sample3

; a 'sample' document model definition
; left - section name identifier
; right - section name tag
[sample]
Title = title
Author = author
Subject = subject
Abstract = abstract
Content = content
PublishingDate=date

[sample2]
Title = title
Author = author
Subject = subject
[sample3]
Title = title
Author = author
Abstract = abstract
Docnum = docnum
```

Strukturierte Dokumente (Abschnittsunterstützung)

Wenn ein Dokument einen formatierten Abschnitt enthält, der nicht im Dokumentmodell definiert wurde, wird der Inhalt dieses Abschnitts zur Indexierung und Suche in den vorher definierten Abschnitt aufgenommen. Ein Dokument enthält beispielsweise die folgenden formatierten Abschnitte:

```
<title>IBM Dictionary of Computing
<subject> Computers, Reference, ....
<author>McDaniel, George
<abstract> Contains up-to-the-minute coverage of information processing systems,
communication products and facilities, personal computers, and office systems, as
well as the full range of IBM hardware and software products.
```

Das Dokumentmodell book ist folgendermaßen definiert:

```
[MODELS]
modelname=book
[book]
Title = title
Author = author
Abstract = abstract
```

Der Abschnitt `<subject>` ist nicht im Dokumentmodell book enthalten. Wenn das Dokument indexiert wird, wird der Inhalt des Abschnitts `<subject>` mit dem Inhalt des Abschnitts `<title>` indexiert. Er ist auch zur Suche innerhalb des Abschnitts `<title>` verfügbar.

Wenn Sie beim Erstellen des Indexes eine Liste mit Modellen angeben, ist das Standardmodell das erste darin. Sie können das Standardmodell unter Verwendung des Befehls `desmodix` ändern.

XML-Dokumente

Bei abschnittsfähigen Indizes müssen XML-Dokumente richtig strukturiert sein und ein Stammelement enthalten. Der Name des Stammelements muß einschließlich der Groß-/Kleinschreibung einem der definierten Modellnamen entsprechen. Die Modellbeschreibung in der Dokumentmodelldatei muß eine Untermenge des Dokumentmodells sein, das in der DTD-Datei (DTD=Document Type Definition, Dokumenttypendefinition) für das Dokument definiert wurde.

Die Modellbeschreibung muß mit dem Stammelement anfangen. Für jedes XML-Element, das Sie als Abschnitt verwenden möchten, müssen Sie die zugehörige vollständige Hierarchie in die Modellbeschreibung aufnehmen. Wenn ein Abschnitt den Typ DATE hat, muß dieser Abschnitt ein Blatt in der Dokumentmodellbaumstruktur sein. Die Verschachtelung von Attributabschnitten wird nicht unterstützt.

Strukturierte Dokumente (Abschnittsunterstützung)

Eine Modellbeschreibung für XML-Dokumente könnte folgendermaßen aussehen:

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

Ein XML-Dokument könnte wie nachfolgend dargestellt aussehen. In diesem Beispiel wird auch gezeigt, wie die nicht im Modell definierten Abschnitte indiziert werden.

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

<LETTER>
  <HEADER>This tag has been skipped in the definition, to this text will
    be added to the section named LETTER
  </HEADER>
  <DATE>
    01.01.2000 03.02.2000
  </DATE>
  <ADDRESS>
    Text will be added to the section named LETTER/address.
  <CITY>
    Text will be added to section named LETTER/address/City.
  </CITY>
</ADDRESS>
<CONTENT>
  Text will be added to the section named LETTER/Content.

  <NOSECTION>Text will be added to the section named LETTER/Content
    because NOSECTION is not defined.
</NOSECTION>
<GREETINGS>
  Text will be added to section named LETTER/Content/Greetings.
</GREETINGS>
</CONTENT>
</LETTER>
```

Wörterverzeichnisse, Stoppwortlisten, Abkürzungslisten und Sprachparameter

In Tabelle 5 werden die unterstützten Sprachen und die Namen der Dateien gezeigt, 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 verwenden.

In dieser Tabelle wird auch gezeigt, welchen LANGUAGE-Parameter Sie angeben müssen, wenn Sie eine Textspalte oder externe Dokumente aktivieren. Dadurch wird dem DB2 Text Extender mitgeteilt, in welcher Sprache die Dokumente geschrieben sind, damit das richtige Wörterverzeichnis für die vorkommende linguistische Verarbeitung verwendet werden kann.

Tabelle 5. Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden

Sprache	Dateiname	Parameter LANGUAGE	Codepage
Arabisch	arabic	ARABIC	864
Portugiesisch (Brasilien)	brazil	BRAZILIAN	850
Französisch (Kanada)	canadien	CAN_FRENCH	850
Katalanisch	catala	CATALAN	850
Dänisch	dansk	DANISH	850
Niederländisch	nederlnd	DUTCH	850
Finnisch	suomi	FINNISH	850
Französisch	francais	FRENCH	850
Deutsch	deutsch	GERMAN	850
Hebräisch	hebrew	HEBREW	862
Isländisch	islensk	ICELANDIC	850
Italienisch	italiano	ITALIAN	850
Norwegisch (Bokmal)	norbook	BM_NORWEGIAN	850
Norwegisch (Nynorsk)	norntn	NN_NORWEGIAN	850
Portugiesisch	portugal	PORTUGUESE	850
Russisch	russian	RUSSIAN	866
Spanisch	espana	SPANISH	850
Schwedisch	svensk	SWEDISH	850
Deutsch (Schweiz)	dschweiz	SWISS_GERMAN	850

Wörterverzeichnisse, Stoppwortlisten und Abkürzungslisten

Tabelle 5. Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden (Forts.)

Sprache	Dateiname	Parameter LANGUAGE	Codepage
Thailändisch	thai	THAI	850
Englisch (Großbritannien)	uk	UK_ENGLISH	850
Englisch (USA)	us	US_ENGLISH	850

Die Dateien werden durch ihre Erweiterung voneinander unterschieden.

Inhalt	Erweiterung
Wörterverzeichnis	DIC
Stoppwortliste	STW
Abkürzungsliste	ABR

Die Stoppwort- und Abkürzungsdateien ändern

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 31.

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.

Kapitel 5. Text für Suchvorgänge vorbereiten

Durch „Kapitel 3. Erste Schritte“ auf Seite 27 werden Sie damit vertraut gemacht, wie Texte für Suchvorgänge des DB2 Text Extender vorbereitet werden, indem Sie schrittweise durch ein einfaches Beispiel geführt werden. In diesem Kapitel wird das Vorbereiten eines Textes für Suchvorgänge detaillierter beschrieben. Zudem sind darin alle Aspekte aufgeführt, die vor Beginn in Betracht gezogen werden sollten.

Im folgenden sind die Schritte zum Vorbereiten eines Textes für Suchvorgänge aufgeführt:

1. Bereiten Sie sich umfassend vor.
2. Starten Sie den DB2 Text Extender-Befehlszeilenprozessor.
3. Stellen Sie eine Verbindung zu einer Datenbank her.
4. Aktivieren Sie eine Datenbank für die Textsuche.
5. Aktivieren Sie eine Texttabelle für die Textsuche (dieser Schritt ist nicht erforderlich, wenn Sie einen Index pro Textspalte erstellen).
6. Aktivieren Sie eine Textspalte für die Textsuche.

Vor Beginn des Vorbereitens von Text für Suchvorgänge

Hinweis

Lesen Sie diesen Abschnitt sorgfältig durch. Darin werden die Optionen aufgelistet, mit denen Sie *vor Beginn* des Vorbereitens von Text für Suchvorgänge vertraut sein sollten.

- **UNIX-Benutzer müssen ihr Profil konfigurieren.**

Wenn Sie kein DB2 Text Extender-Exemplareigner sind, finden Sie im Abschnitt „Konfiguration“ auf Seite 18 Informationen, wie Sie Ihr Profil einrichten.

- **Einen Index für die gesamte Tabelle erstellen?**

Sie müssen sich entscheiden, ob Sie einen einzigen Index für eine ganze Texttabelle oder einen separaten Index für jede Textspalte erstellen möchten. Der Abschnitt „Einen oder mehrere Textindizes für eine Tabelle erstellen“ auf Seite 46 enthält Informationen, die Ihre Entscheidung erleichtern.

Vor Beginn des Vorbereitens von Text für Suchvorgänge

- **Kennen Sie Ihre Dokumente.**

Wenn Sie Dokumente für Suchvorgänge vorbereiten, müssen Sie deren CCSIDs, Sprachen und Textformate angeben. Weitere Informationen hierzu finden Sie im Abschnitt „Kapitel 4. Die Sucherfordernisse planen“ auf Seite 31.

- **Entscheiden Sie sich für die Textindextypen, die Sie benötigen.**

Der von Ihnen benötigte Indextyp wird durch die Arten von Suchvorgängen bestimmt, die Sie durchführen möchten (präzise Suche, Suche nach grober Übereinstimmung usw.). Daneben spielt eine Rolle, ob Ihre Dokumente im SBCS oder im DBCS vorliegen. Weitere Informationen hierzu finden Sie im Abschnitt „Kapitel 4. Die Sucherfordernisse planen“ auf Seite 31.

- **Treffen Sie eine Entscheidung dazu, wo Sie Indizes speichern möchten.**

Wenn Sie Dokumente für Suchvorgänge vorbereiten, erstellt der DB2 Text Extender einen Textindex. Sie müssen angeben, in welchem Verzeichnis Sie diesen Index speichern möchten. Stellen Sie sicher, daß dafür ausreichend Plattenspeicherplatz vorhanden ist (siehe „Die Größe eines Indexes berechnen“ auf Seite 47).

- **Prüfen Sie den Namen der Standarddatenbank.**

Der Name der Standarddatenbank in der DB2-Umgebungsvariablen DB2DBDFT (siehe „Umgebungsvariablen“ auf Seite 18) ist der Name, der vom DB2 Text Extender verwendet wird, wenn Sie während des Vorbereitens von Text für Suchvorgänge keinen Datenbanknamen angeben.

- **Definieren Sie die Textkonfiguration.**

Die Textkonfiguration bestimmt die Standardeinstellungen für die Index-CCSID, für die Sprache und das Format der Dokumente, für den Indextyp, für die Indexaktualisierungsfrequenz, für den Tabellenbereichsnamen und für das Indexverzeichnis.

Sie können diese Einstellungen außer Kraft setzen, wenn Sie den Text für Suchvorgänge vorbereiten. Es ist allerdings einfacher, die Standardeinstellungen im Vorfeld richtig zu definieren. Die anfänglichen Textkonfigurationseinstellungen bei der Installation des DB2 Text Extender werden in „Textkonfigurationseinstellungen“ auf Seite 19 beschrieben. Verwenden Sie die Informationen im Abschnitt „CHANGE TEXT CONFIGURATION“ auf Seite 133, um die Installationseinstellungen zu ändern und eigene Standardwerte zu definieren.

- **Definieren Sie die Abschnittsunterstützung.**

Wenn Sie Suchvorgänge auf einen bestimmten Abschnitt eines Dokuments einschränken müssen, finden Sie Informationen dazu, wie Modelle in der Dokumentmodelldatei angegeben werden, im Abschnitt „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50.

Vor Beginn des Vorbereitens von Text für Suchvorgänge

- **Ändern Sie die Stoppwort- und Abkürzungslisten.**

Lesen Sie die Abschnitte „Warum Textdokumente indexiert sein müssen“ auf Seite 31 und „Die Stoppwort- und Abkürzungsdateien ändern“ auf Seite 58, um sich mit dem Konzept von *Stoppwort*- und Abkürzungslisten vertraut zu machen und zu entscheiden, ob Sie sie vor Beginn der Indexierung ändern möchten.

Wenn Sie die Informationen gesammelt und die im Abschnitt „Vor Beginn des Vorbereitens von Text für Suchvorgänge“ auf Seite 59 beschriebenen Entscheidungen getroffen haben, sind Sie dazu bereit, Ihren Text für Suchvorgänge vorzubereiten.

Den DB2 Text Extender-Befehlszeilenprozessor starten

Zusammenfassung	
Zeitpunkt	Wahlfrei. Bei Beginn jeder Sitzung.
Befehl	db2tx
Berechtigung	Beliebig

Sie können DB2 Text Extender-Client-Befehle in der Eingabeaufforderung des Betriebssystems eingeben. Die Befehle ähneln DB2-Befehlen, aber statt ihnen db2 voranzustellen, müssen Sie db2tx voranstellen.

Anstatt allen DB2 Text Extender-Befehlen db2tx voranzustellen, können Sie auch den DB2 Text Extender-Befehlszeilenprozessor starten. Dies hat den Vorteil, daß Ihre Datenbankverbindung nicht nach jedem Befehl unterbrochen wird; die Verbindung bleibt bestehen. (Wenn Sie ohne Befehlszeilenprozessor arbeiten, wird bei jeder Ausführung eines db2tx-Befehls automatisch erneut eine Verbindung hergestellt. Sie müssen aber sicherstellen, daß der Name der Datenbank, zu der die Verbindung hergestellt werden soll, in der Umgebungsvariablen DB2DBDFT definiert ist.)

1. **(Nur unter Windows): Starten Sie den DB2-Befehlszeilenprozessor.**

Geben Sie in der Eingabeaufforderung Ihres Betriebssystems folgenden Befehl ein:

```
=>db2 cmd
```

2. **Starten Sie den DB2 Text Extender-Befehlszeilenprozessor.**

Geben Sie auf UNIX-Systemen in der Eingabeaufforderung des Betriebssystems folgenden Befehl ein:

```
=>db2tx
```

Den DB2 Text Extender-Befehlszeilenprozessor starten

Geben Sie auf Windows-Systemen in der DB2-Eingabeaufforderung folgenden Befehl ein:

```
db2=>db2tx
```

Die db2tx-Eingabeaufforderung wird angezeigt:

```
db2tx=>
```

Alle nachfolgenden Befehle werden als DB2 Text Extender-Befehle interpretiert.

Geben Sie folgenden Befehl ein, um diesen Modus zu verlassen:

```
db2tx=>quit
```

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 DB2 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 von einem Client aus arbeiten, müssen Sie in der DB2 Text Extender-Befehlsgebung sein und das vollständige Format der Anweisung CONNECT verwenden, wobei Sie Ihre Benutzer-ID und Ihr Kennwort angeben.

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

Hilfe für den Befehlszeilenprozessor

Um eine Liste mit Befehlen anzuzeigen, geben Sie folgendes ein:

```
db2tx ?
```

Um die Syntax zu einem individuellen Befehl aufzurufen, geben Sie folgendes ein:

```
db2tx ? befehl
```

Beispiel:

```
db2tx ? CHANGE TEXT CONFIGURATION
```

Eine Verbindung zu einer Datenbank herstellen

Zusammenfassung

Zeitpunkt	Wahlfrei. Zum Starten einer Sitzung mit einer anderen Datenbank als der in der Umgebungsvariablen DB2DBDFT angegebenen Standarddatenbank.
Befehl	CONNECT
Berechtigung	CONNECT für die Datenbank.

Bevor Sie in einer DB2 Text Extender-Sitzung weitere Befehle ausgeben können, müssen Sie eine Verbindung zu einer Datenbank herstellen. Wenn Sie einen Befehl ausgeben, ohne mit einer Datenbank verbunden zu sein, stellt der DB2 Text Extender automatisch eine Verbindung zur Standarddatenbank her, die in der Umgebungsvariablen DB2DBDFT angegeben ist. Sie können explizit eine Verbindung zu einer Datenbank herstellen, indem Sie den DB2 Text Extender-Befehl CONNECT TO verwenden.

```
db2tx CONNECT TO sample-database
```

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

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

Anmerkung: Wenn Sie anschließend einen DB2 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.

Eine Datenbank aktivieren

Eine Datenbank aktivieren

Zusammenfassung

Zeitpunkt	Einmal für jede Datenbank, die Textspalten enthält, die durchsucht werden sollen.
Befehl	db2tx=>ENABLE DATABASE
Berechtigung	SYSADM oder DBADM

Dieser Befehl hat keine Parameter. Er bereitet die verbundene Datenbank für die Verwendung durch den DB2 Text Extender vor.

Dieser Befehl deklariert auch die DB2 Text Extender-Funktionen und die eindeutigen Typen des DB2 Text Extender für DB2. Hierbei handelt es sich um SQL-Funktionen, die Sie später für die Suche nach Text verwenden. Sie werden in „Kapitel 11. Suchfunktionen“ auf Seite 199 beschrieben. Diese Deklarationen gelten für alle zukünftigen Sitzungen.

Eine Katalogsicht, TEXTINDEXES, wird erstellt, die die aktivierten Textspalten protokolliert. Siehe „Mit der DB2 Text Extender-Katalogsicht arbeiten“ auf Seite 112.

Dieser Befehl erstellt Textkonfigurationsdaten für die Datenbank. Dazu gehören Standardwerte für den Index, den Text und Verarbeitungsmerkmale. Sie werden in „Textkonfigurationseinstellungen“ auf Seite 19 beschrieben.

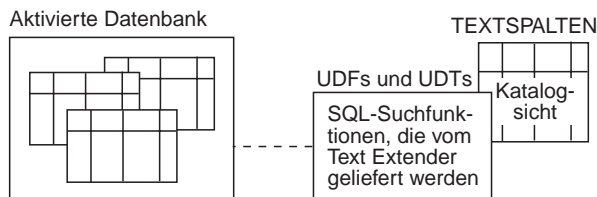


Abbildung 8. Eine Datenbank aktivieren

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 117.

Hinweise

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.

Wenn Sie sich später entscheiden, eine aktivierte Datenbank zu löschen, sollten Sie sie zunächst inaktivieren, um sicherzustellen, daß die deklarierten DB2 Text Extender-Funktionen, die Katalogsicht usw. gelöscht sind.

Eine Texttabelle aktivieren (wahlfrei)

Zusammenfassung

Zeitpunkt	Wahlfrei. Verwenden Sie diesen Befehl ausschließlich, um einen gemeinsamen Index für alle Textspalten in einer Tabelle zu erstellen. Siehe „Einen oder mehrere Textindizes für eine Tabelle erstellen“ auf Seite 46.
Befehl	db2tx=>ENABLE TEXT TABLE ... (siehe Beispiele)
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. h. einen separaten Index für jede Textspalte.

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. 9 und Abb. 10 auf Seite 68 gezeigt.

Während dieses Schritts erstellt der DB2 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.

Eine Texttabelle aktivieren (wahlfrei)

Hinweis

Wenn eine Einstellung, z. B. der Indextyp, für die meisten Textspalten gleich sein soll, verwenden Sie die Textkonfigurationsdaten, um Standardeinstellungen anzugeben. Siehe „Textkonfigurationseinstellungen“ auf Seite 19.

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 DB2 Text Extender aktiviert wurde.

Um einen Index, der mit `ENABLE TEXT TABLE` erstellt wurde, zu löschen, siehe „Eine Texttabelle inaktivieren“ auf Seite 116.

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.MYTABLE`:

```
db2tx ENABLE TEXT TABLE db2tx.mytable
```

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.mytable
      INDEXTYPE linguistic
      UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)
      DIRECTORY DB2TX_INSTOWNERHOMEDIR/db2tx/indexes
```

Für einen OS/2-, einen Windows NT- oder einen Windows 2000-Server:

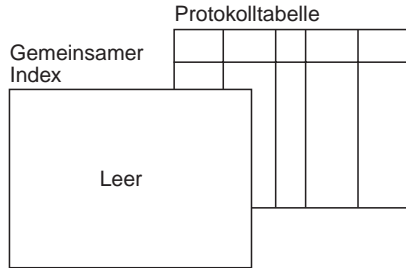
```
db2tx ENABLE TEXT TABLE db2tx.mytable
      INDEXTYPE linguistic
      UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)
      DIRECTORY d:\dmb\db2tx\indexes
```


Eine Texttabelle aktivieren (wahlfrei)

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.

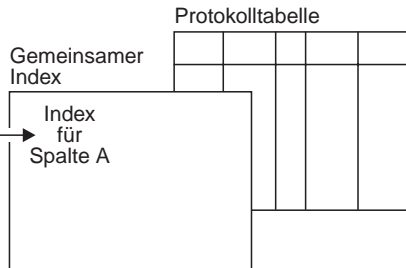
1. Texttabelle aktivieren

	A		B
	Text		
			Text



2. Textspalte A aktivieren

	A		B
	Text		
			Text



3. Textspalte B aktivieren

	A		B
	Text		
			Text

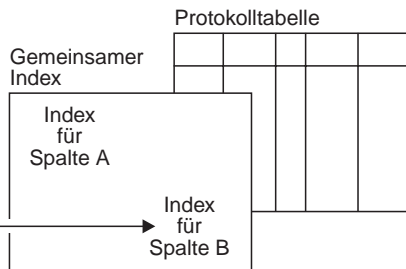
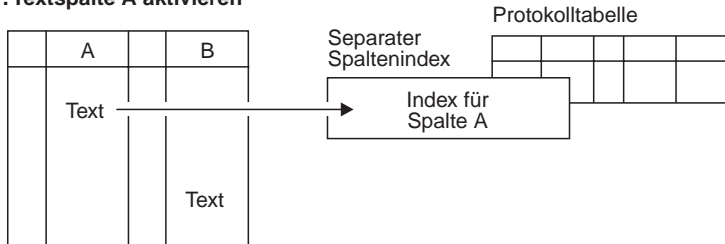


Abbildung 9. Einen gemeinsamen Index für alle Textspalten in einer Tabelle erstellen

Eine Textspalte aktivieren

1. Textspalte A aktivieren



2. Textspalte B aktivieren

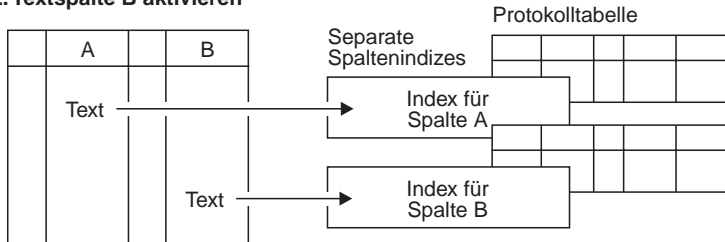


Abbildung 10. Einen separaten Index für jede Textspalte erstellen

Eine Textspalte aktivieren

Zusammenfassung

Zeitpunkt	Einmal für jede Spalte, die zu durchsuchenden Text enthält.
Befehl	db2tx=>ENABLE TEXT COLUMN ... (siehe Beispiele)
Berechtigung	ALTER, SELECT, UPDATE für die Tabelle

Hinweis

Wenn eine Einstellung, z. B. der Indextyp, für die meisten Textspalten gleich sein soll, verwenden Sie die Textkonfigurationsdaten, um Standardeinstellungen anzugeben. Siehe „Textkonfigurationseinstellungen“ auf Seite 19.

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 und wird eine Protokolltabelle sowie ein Index erstellt.

Eine Kennungsspalte wird hinzugefügt

Während dieses Schritts fügt der DB2 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.

Hinweis

Diese Version des DB2 Text Extender kann nicht mit Indizes arbeiten, die unter Verwendung eines früheren Releases erstellt wurden. Sie müssen alle DB2 Text Extender-fähigen Datenbanken und Tabellen inaktivieren und anschließend erneut aktivieren.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
Daten	Daten	Daten	Daten	Text

Abbildung 11. Struktur der Tabelle DB2TX.MYTABLE vor der Aktivierung

Die Spalte mit Textblöcken 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 DB2 Text Extender vorbereiten.

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

DB2TX.SAMPLE

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

Abbildung 12. Struktur der Tabelle DB2TX.MYTABLE nach der Aktivierung

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

Eine Textspalte aktivieren

Die Dokumentinformationen werden gesetzt

Sie geben den Typ von Textdokumenten an, die Sie normalerweise in dieser Textspalte speichern: deren Format (z. B. ASCII), deren Sprache und deren CCSID. Standardwerte für diese Informationen können in den Textkonfigurationseinstellungen angegeben werden. Siehe „Textkonfigurationseinstellungen“ auf Seite 19.

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 verfügbar). 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 Plattenspeicherplatzes 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 indiziert 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 102 gezeigt.

Wenn während der Indexierung Fehler auftreten, d. h., wenn z. 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 109 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 72.

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 DB2 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 DB2 Text Extender während dieses Schritts einen separaten Index für jede Textspalte. 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 Textspalte zulassen wollen, indem Sie die Textspalte unterschiedlichen Indextypen, z. 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 entscheiden haben, die Konfiguration einer Knotengruppe zu ändern und eine Neuverteilung der Tabellenzeilen zu starten, müssen Sie die Textindizes und Protokolltabellen erneut erstellen, d. h., Sie müssen alle DB2 Text Extender-fähigen Tabellen dieser Knotengruppe inaktivieren und erneut aktivieren.

Beispiele

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

```
db2tx ENABLE TEXT COLUMN db2tx.mytable comment
        HANDLE          commenthandle
```

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

Eine Textspalte aktivieren

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.mytable   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.mytable   comment
                                HANDLE      commenthandle
                                INDEXTYPE   linguistic
                                UPDATEFREQ   min(100) d(1,2,3,4,5) h(12,15) m(00)
                                UPDATEINDEX  UPDATE
                                DIRECTORY    DB2TX_INSTOWNERHOMEDIR/db2tx/indexes
```

Für einen OS/2-, einen Windows NT- oder einen Windows 2000-Server:

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

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. 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-Dienstprogramm REORGANIZE, um zu prüfen, ob die Tabelle reorganisiert werden muß. 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 indiziert.
2. Reorganisieren Sie die Tabelle unter Verwendung des DB2-Dienstprogramms REORGANIZE.
3. Erstellen Sie den Index, indem Sie UPDATE INDEX ausführen.

Wenn Sie eine Textspalte oder externe Dateien aktivieren, fügt der DB2 Text Extender eine Kennungsspalte zur Tabelle hinzu und initialisiert die Kennungswerte, wodurch DB2-Protokolleinträge geschrieben werden. Wenn eine ungewöhnlich große Anzahl an Protokolleinträgen geschrieben wird, verfügt DB2 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-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 einen Zwischenbefehl COMMIT ausführt, indem Sie den Konfigurationsparameter COMMITCOUNT verwenden, der in „Konfiguration“ auf Seite 18 beschrieben wird. Der von Ihnen angegebene Wert gibt an, nach wie vielen Einfügungs- oder Aktualisierungsanweisungen der DB2 Text Extender eine DB2-Anweisung COMMIT ausgibt. Hierdurch wird die erforderliche Größe für Protokolltabellen verringert, allerdings wird auch die Zeit erhöht, die für den Aktivierungsschritt erforderlich ist.

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 DB2 Text Extender aktiviert werden zu können. Wenn sich die Dokumente in einer Spalte eines anderen Typs befinden, z. B. ein benutzerdefinierter eindeutiger Typ (UDT), müssen Sie eine Funktion liefern, die den Benutzertyp als Eingabe nimmt und CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB als Ausgabetyt liefert.

Eine Textspalte aktivieren

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 benutzerdefinierten eindeutigen Typ (User-defined Distinct Type, 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 DB2 Text Extender zu aktivieren:

1. Erstellen Sie eine UDF (User-defined Function, benutzerdefinierte Funktion) 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
...

```

Spalten aktivieren, die den Datentyp DATALINK enthalten

Der DB2 Text Extender unterstützt den Datentyp `DATALINK`. Die Data Link-Technologie von DB2 bietet referentielle Integrität zwischen in DB2 gespeicherten Daten und Daten in externen Dateien. Dies geschieht durch einen neuen Datentyp `DATALINK`, der Verweise auf ein Objekt enthält, das außerhalb der Datenbank (extern) gespeichert ist. Ein in einer DB2-Tabelle gespeicherter `DATALINK`-Wert verschlüsselt den Namen des Data Links-Servers, der die Datei und den Dateinamen enthält, in Form einer URL (Uniform Resource Location).

Wenn sich die `DATALINK`-Daten in externen Dateien, nicht in einer DB2-Tabelle, befinden, befolgen Sie die Beschreibung im Abschnitt „Externe Textdateien aktivieren“ auf Seite 75.

Es ist wichtig, daß Sie vor der Aktivierung einer Textspalte, die den `DATALINK`-Datentyp enthält, das Installationsprüfbeispiel ausführen (siehe „Eine Beispieldatenbank zur Installationsprüfung vorbereiten“ auf Seite 17).

Beispiel für AIX: Die folgende Anweisung stellt dar, wie ein Data Link-Wert unter Verwendung des Datentyps `DATALINK` unter AIX in Ihre Textspalte eingefügt wird:


```
db2 "insert into mytable values ('Peter',
                                dlvalue ('http://dlfs.boeblingen.ibm.com/test/documents/letter1',
                                          'URL', 'this is a letter from Peter'))"
```

In diesem Beispiel ist *dlfs.boeblingen.ibm.com* der Name des Datei-Servers, *test* ist der Name des Dateisystems und *documents/letter1* ist der Pfad und der Dateiname.

Stellen Sie sicher, daß das Dateisystem mit demselben Namen an den DB2-Server angehängt ist, der auch in der URL angegeben wurde.

Beispiel für Windows NT oder Windows 2000: Die folgende Anweisung stellt dar, wie ein Data Link-Wert unter Verwendung des Datentyps DATALINK unter Windows in Ihre Textspalte eingefügt wird:

```
db2 "insert into mytable ('Peter',
                           dlvalue ('unc:\\dlfs.boeblingen.ibm.com\cdrive\documents\letter1',
                                     'URL', 'this is a letter from Peter'))"
```

In diesem Beispiel ist *cdrive* der gemeinsame Name des Laufwerks, auf dem die Datei gespeichert ist.

Wenn Sie während der Indexierung von Dokumenten, die in einer Datalink-Spalte gespeichert sind, auf Probleme stoßen, stellen Sie sicher, daß Sie von Ihrem DB2-Server aus unter Verwendung der von DB2 bereitgestellten APIs auf diese Dokumente zugreifen können.

Externe Textdateien aktivieren

Diese Option wird nur bereitgestellt, um die Kompatibilität zu früheren Releases zu gewährleisten. Die bevorzugte Methode zum Indexieren von Dateien, die nicht in DB2-Tabellen gespeichert sind, ist die Verwendung einer Spalte mit einem DATALINK-Datentyp. Siehe „Spalten aktivieren, die den Datentyp DATALINK enthalten“ auf Seite 74.

Zusammenfassung

Zeitpunkt	Wahlfrei. Einmal für jede Tabelle, die den zu durchsuchenden externen Dateien zugeordnet ist.
Befehl	db2tx=>ENABLE TEXT FILES ... (siehe Beispiele)
Berechtigung	ALTER, SELECT, UPDATE für die Tabelle

Der DB2 Text Extender kann nicht nur in Text suchen, der in DB2-Tabellen gespeichert ist, sondern auch in Textdokumenten, die in Dateien gespeichert sind. Dieser Vorbereitungsschritt ist erforderlich, wenn Sie nach Text in exter-

Externe Textdateien aktivieren

nen 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 68 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 72.
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`-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 83 finden Sie dazu eine Beschreibung.

Sie können zusätzliche Parameter, wie z. 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.

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
```

Die Sitzung beenden

Die Sitzung 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 DB2 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 101.

Wenn das Indexieren der Dokumente beendet ist, können Sie beginnen, Informationen abzurufen, wie in „Kapitel 6. Vorgehensweise bei Suchvorgängen“ auf Seite 79 beschrieben.

Hinweis

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

Geben Sie folgenden Befehl ein, um den DB2 Text Extender-Befehlsprozessor zu beenden:

```
db2tx=>quit
```

Kapitel 6. Vorgehensweise bei Suchvorgängen

Der DB2 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 und werden hier als DB2 Text Extender-Funktionen bezeichnet.

Ein Beschreibung der Syntax dieser Funktionen finden Sie in „Kapitel 11. Suchfunktionen“ auf Seite 199.

Bevor Sie mit der Suche beginnen, lesen Sie den Abschnitt „Sucharten“ auf Seite 42 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 Beispielfunktionen des DB2 Text Extender
- Die Beispieltabelle `DB2TX.SAMPLE`
- Kennungen für externe Dateien
- Setzen des Funktionspfads, um SQL Zugriff auf die DB2 Text Extender-Funktionen zu geben
- Suchen nach Text unter Verwendung von `CONTAINS`, `NO_OF_MATCHES` und `RANK`
- Angeben von Suchargumenten in DB2 Text Extender-Funktionen 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 `SEARCH_RESULT`

Lokation von Syntaxbeispielen von Suchfunktionen

Lokation von Syntaxbeispielen von Suchfunktionen

Der DB2 Text Extender stellt eine Eingabedatei für den Befehlszeilenprozessor mit dem Namen `txsample.udf` bereit. Darin sind Beispiele von DB2 Text Extender-Suchfunktionen enthalten, die in bezug auf die Beispieltabelle ausgeführt werden. Verwenden Sie diese Datei, um Beispiele der Syntax der DB2 Text Extender-Textvorbereitungs- und -Suchfunktion 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

Die Beispieltabelle DB2TX.SAMPLE wird in vielen der Beispiele zur Suchsyntax verwendet. Sie können die Beispiele selbst unter Verwendung des DB2 Text Extender testen.

Informationen zum Erstellen der Tabelle DB2TX.SAMPLE finden Sie im Abschnitt „Eine Beispieldatenbank zur Installationsprüfung vorbereiten“ auf Seite 17.

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

Tabelle 6. 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.

Die Beispieltabelle DB2TX.SAMPLE

Tabelle 6. Ein Auszug aus der Beispieltabelle DB2TX.SAMPLE (Forts.)

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 7	SKY at TOROLAB4	ODBC & Stored Procedures	1995-07-25 -20.42.27	<p>There are two sets of sample programs explaining the use of Stored Procedures using CLI (ODBC).</p> <p>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.</p> <p>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.</p>
doc 8	ADAMACHE at TOROLAB2	DB2SYS.DLL access violation	1995-07-25 -21.13.22	<p>Did you have a previous beta version installed? If so, did you remove it using Software Installer?</p> <p>Did you remove the database directories (SQLDBDIR and SQL00001, etc.) from previous beta drivers?</p>
doc 9	ADAMACHE at TOROLAB2	CREATE DB = SYS3175: db2sysc.exe in db2eng.dll	1995-07-25 -21.40.09	<p>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</p> <p>This happens because the directory format has changed between beta4 and beta5. Our DB2/2 installation does not migrate the sqlbdir directory between beta drivers. You should remove all occurrences of sqlbdir and sql000x directories and \sqllib\db2\sqlbdir directory.</p> <p>What you should do is delete the previous beta with Software Installer, remove all occurrences of sqlbdir and sql000x directories and \sqllib\db2\sqlbdir directory, and then install the new code.</p>

Die Beispieltabelle DB2TX.SAMPLE

Tabelle 6. Ein Auszug aus der Beispieltabelle DB2TX.SAMPLE (Forts.)

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
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 13. 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 DB2 Text Extender vorbereiten, indem Sie den Befehl `ENABLE TEXT COLUMN` verwenden.

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 14. Die Tabelle DB2TX.SAMPLE nach dem Aktivieren

Die Tabelle enthält jetzt eine zusätzliche 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..'
doc 2	After rechecking ...	X'..handle..'

Kennungen, die durch ENABLE TEXT COLUMN erstellt wurden

Eingefügte Zeile:

doc 11	I have installed ...	X'..handle..'
--------	----------------------	---------------

Kennungen, die durch einen Einfügeauslöser erstellt wurden

Abbildung 15. Die Kennung für eine eingefügte Zeile wird von einem Auslöser erstellt

Eine Kennung enthält die folgenden Informationen:

- Dokument-ID

- Name und Standort des zugeordneten Indexes

- Dokumentinformationen: CCSID, Format und Sprache

Die SQL-Funktionen, die der DB2 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 DB2 Text Extender kann nicht nur in Text suchen, der in DB2-Tabellen gespeichert ist, sondern auch in Textdateien, die an anderen Orten gespeichert sind. „Externe Textdateien aktivieren“ auf Seite 75 beschreibt den Vorbereitungsschritt, der es ermöglicht, in Textdokumenten zu suchen, die nicht in DB2-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

►► SET CURRENT FUNCTION PATH [] DB2TX, ... ◀◀

Den aktuellen Funktionspfad setzen

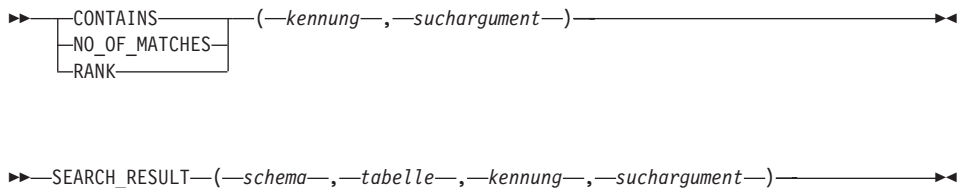
Verwenden Sie die SQL-Anweisung SET CURRENT FUNCTION PATH, um DB2TX zu den aktuellen Pfadnamen hinzuzufügen, so daß SQL die DB2 Text Extender-Funktionen finden kann. Wenn Sie sich gegen diese Vorgehensweise entscheiden, können Sie die Funktionsnamen explizit qualifizieren, indem Sie beispielsweise DB2TX.CONTAINS für die Funktion CONTAINS eingeben.

Die Beispiele in diesem Kapitel verwenden die qualifizierte Form für DB2 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 DB2 Text Extender gelieferten SQL-Funktionen 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 99 beschrieben.

Jede dieser DB2 Text Extender-Funktionen sucht im Textindex nach Stellen, an denen das Suchargument vorkommt. Angenommen, es befinden sich 100 000 Textdokumente in der Tabelle, dann wird die Funktion 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 Funktion eine interne Liste aller Dokumente, die den Suchbegriff enthalten, erstellt.

Nachfolgende Aufrufe der Funktion stellen fest, ob das betreffende Dokument sich in der Liste befindet.

Hinweis

Wenn Sie die DB2 Text Extender-Funktionen für die Suche in einer Tabelle verwenden, stellen Sie sicher, daß Sie die Kennungsspalte, und nicht die Textspalte, an die Funktion 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. B. von der Anzahl an Dokumenten, die indiziert 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. 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.

```
db2=>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` indiziert 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:

```
db2=>SELECT DATE, SUBJECT
        FROM "db2tx.sample"
        WHERE DB2TX.CONTAINS (COMMENTHANDLE, "compress") = 1
```

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

```
DB2 "SELECT DB2TX.file(COMMENTHANDLE)
      FROM db2tx.sample"
WHERE DB2TX.CONTAINS (COMMENTHANDLE, \"compress\") = 1
```

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.

```
db2=>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 Funktion RANK abrufen.

Beispiel:

```
db2=>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 DB2 Text Extender-Funktionen 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:

```
db2=>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 92.)

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

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,
        '"compress" | "compiler"') = 1
```

Sie können mehrere Begriffe unter Verwendung von Booleschen Operatoren kombinieren:

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,
        '"compress" | "compiler" & "DB2"') = 1
```

Wenn Sie mehr als einen Booleschen Operator verwenden, wertet der DB2 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 beispielsweise

Suchargumente angeben

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

ausgewertet als

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

Im folgenden Beispiel müssen Sie die runden Klammern einfügen:

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

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 DB2 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 DB2 Text Extender auch nach Variationen der Begriffe, z. 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.

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.

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 (Wörter, die mit Platzhalterzeichen versehen sind), nicht auf eine Grundform reduziert werden werden. Wenn Sie also nach `pass%` 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 DB2 Text Extender verwendet zwei Platzhalterzeichen: das Prozentzeichen (%) und das Unterstreichungszeichen (_):

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

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE, '%"name"') = 1
```

Bei diesem Suchbegriff werden Textdokumente gefunden, die beispielsweise "username" (Benutzername), "file_name" (Dateiname) und "table-name" (Tabellenname) enthalten.

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

```
db2=>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.

```
db2=>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:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              '"10!% interest" 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.

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"hard disk"') = 1
```

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

```
db2=>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 Sucharguments, mit dem Textdokumente gefunden werden, in denen die Suchbegriffe im selben Satz auftreten:

```
db2=>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:

```
db2=>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 Dokumentmodell-datei beschrieben ist. Weitere Informationen finden Sie in „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50.

```
db2=>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 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 beispiels-

weise nach dem englischen Wort "book" suchen, kann es auch nützlich sein, nach dessen Synonymen zu suchen. Geben Sie dazu folgenden Befehl ein:

```
db2=>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. h., das Suchargument wird folgendermaßen interpretiert:

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

Die Synonyme befinden sich in einem Wörterverzeichnis, das mit dem DB2 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:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'SYNONYM FORM OF UK_ENGLISH "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 DB2 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. Diese Funktionen werden in „Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis“ auf Seite 257 beschrieben.

Ein Beispiel dafür ist die Suche nach einer Pluralform, z. 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.

Suchargumente angeben

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.

```
db2=>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 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. 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 90 beschrieben) verwenden und auch nicht mit SYNONYM FORM OF (in „Nach Synonymen von Begriffen suchen“ auf Seite 90 beschrieben).

Sie können den Operator NOT nur mit einem Suchbegriff verwenden, d. h., Sie können die Operatoren &, | und NOT nicht frei kombinieren (siehe „Suchargumentsyntax“ auf Seite 219).

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:

```
db2=>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, compress, 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 DB2 Text Extender während der Suche Wortgrenzen respektiert. Beispiel:

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                           'BOUND "korean-expression"') = 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.

Suchargumente angeben

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'SOUNDS LIKE "Meyer"') = 1
```

Diese Suche findet Vorkommen von “Meyer”, “Mayer”, “Meier” und “Maier”.

Thesaurus-Suche

Die Thesaurus-Suche ist ein weiteres Beispiel für die leistungsfähigen Erweiterungsfunktionen für Suchbegriffe des DB2 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 271. Das Beispiel in Abb. 18 auf Seite 272 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.

```
db2=>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”.

```
db2=>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:

```
db2=>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:

```
db2=>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 Sucharguments 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

Eine Feinanpassung für eine vorherige Suche durchführen

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.

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.

```
db2=>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.

```
db2=>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.

```
db2=>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.

```
db2=>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
```

Eine Feinanpassung für eine vorherige Suche durchführen

```
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.

```
db2=>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 DB2 Text Extender-Funktionen können Sie die Textdaten in Kennungen setzen oder ändern.

Textdaten beim Einfügen von neuem Text setzen

```
►►—INIT_TEXT_HANDLE—┌(—format—,—sprache—)—————►►
                    └(—CCSID—,—format—,—sprache—,—dateiname—)┘
```

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.

Informationen in Kennungen setzen und extrahieren

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.

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`.

```
db2=>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:

```
db2=>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.

```
db2=>SELECT COUNT(*)
      FROM DB2TX.SAMPLE
      WHERE DB2TX.FORMAT(COMMENTHANDLE) = 'TDS'
```


Die Suchleistung verbessern

Beispiel:

```
db2=>SET CURRENT FUNCTION PATH = db2tx
```

```
db2=>SELECT author  
       FROM db2tx.sample T1,  
       table(DB2TX.SEARCH_RESULT('DB2TX','SAMPLE','MYDOCHANDLE','compress')) T2
```

Hierbei ist T1.handle = T2.mydochandle.

Wenn Sie den aktuellen Pfad nicht gesetzt haben, sieht die letzte Zeile des Beispiels folgendermaßen aus:

```
where cast(T1.handle as varchar(60)) = cast(T2.mydochandle as varchar (60))
```

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.

Kapitel 7. Verwaltung

In diesem Kapitel wird beschrieben, wie Textindizes verwaltet und nützliche Information abgerufen werden sowie der Textvorbereitungsprozeß umgekehrt wird.

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 im Abschnitt „Eine Textspalte aktivieren“ auf Seite 68.)

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.mytable
```

Dieses Beispiel aktualisiert den Index für eine Spalte einer Mehrindextabelle:

```
db2tx UPDATE INDEX db2tx.mytable HANDLE commenthandle
```

Textindizes verwalten

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 DB2 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 75). 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 DB2 Text Extender in Protokolltabellen überwacht werden, da diese Aktualisierungen nicht innerhalb des Bereichs von DB2 auftreten.

Damit die Aktualisierungen an externen Dateien in einem DB2 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 „Einen Index aktualisieren“ auf Seite 48. 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.mytable
                        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.mytable
                        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

Es können Situationen auftreten, die verhindern, daß Sie in einem Index suchen oder einen Index aktualisieren. In „Den Status eines Indexes anzeigen“ auf Seite 108 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.

Textindizes verwalten

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

```
db2tx RESET INDEX STATUS db2tx.mytable
```

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.mytable 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 indiziert 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.mytable
```

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.mytable 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 DB2 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.mytable
```

Dieses Beispiel reorganisiert den Index einer Mehrindextabellenspalte:

```
db2tx REORGANIZE INDEX db2tx.mytable HANDLE commenthandle
```

Nützliche Informationen abrufen

In diesem Abschnitt werden die Befehle 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 Texteingstellungen für eine Spalte

Nützliche Informationen abrufen

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 benutztem Index aktiviert

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

Textspaltenname	Kennungsspaltenname
-----	-----
COMMENT	COMMENTHANDLE

Tabelle DB2TX.TEST ist als Tabelle mit Mehrindex aktiviert

Textspaltenname	Kennungsspaltenname
-----	-----
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 DB2 Text Extender-Umgebungsvariablen.

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

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 19 beschrieben. Um sie zu ändern, siehe „Die Textkonfiguration ändern“ auf Seite 21.

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
```

Nützliche Informationen abrufen

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.mytable
```

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.mytable HANDLE commenthandle
```

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

Knoten 1	
Suchstatus	= Suche verfügbar
Aktualisierungsstatus	= 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 20. Ursachencodes für Fehlerereignisse“ auf Seite 325.

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.

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 104 beschrieben.

Fehlerereignisse anzeigen

Wenn Probleme während der Indexierung auftreten, z. 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 20. Ursachencodes für Fehlerereignisse“ auf Seite 325 beschrieben.

Nützliche Informationen abrufen

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.mytable
```

Dieses Beispiel ruft die Indexeinstellungen für den Index einer Mehrindex-tabellenspalte ab:

```
db2tx GET INDEX SETTINGS db2tx.mytable  
                           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.

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.mytable
```

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

```
db2tx GET TEXT INFO db2tx.mytable 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.

Nützliche Informationen abrufen

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 DB2 Text Extender-Katalogsicht arbeiten

Der DB2 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 DB2 Text Extender aktiviert sind.

In DB2TX.TEXTINDEXES werden immer dann neue Einträge erstellt, wenn eine Tabelle bzw. eine Spalte aktiviert wird. Gleiches gilt auch für externe Dateien. Einträge werden immer dann 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 7 zeigt den Inhalt der Katalogsicht.

Tabelle 7. DB2 Text Extender-Katalogsicht

Spaltenname	Datentyp	Kann Nullwerte enthalten	Beschreibung
TABLESCHEMA	CHAR(8)	Nein	Das Schema der Tabelle, für die dieser Eintrag gilt.
TABLERNAME	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.

Mit der DB2 Text Extender-Katalogsicht arbeiten

Tabelle 7. DB2 Text Extender-Katalogsicht (Forts.)

Spaltenname	Datentyp	Kann Nullwerte enthalten	Beschreibung
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	Indextyp: LINGUISTIC, PRECISE, NGRAM.
MINIMUM	INTEGER	Ja	Die kleinste Anzahl von Indexaktualisierungsanforderungen, die erforderlich sind, bevor eine Indexaktualisierung ausgeführt wird. Siehe „Einen Index aktualisieren“ auf Seite 48. 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 „Einen Index aktualisieren“ auf Seite 48. 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 „Einen Index aktualisieren“ auf Seite 48. 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 „Einen Index aktualisieren“ auf Seite 48. 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.

Mit der DB2 Text Extender-Katalogsicht arbeiten

Tabelle 7. DB2 Text Extender-Katalogsicht (Forts.)

Spaltenname	Datentyp	Kann Nullwerte enthalten	Beschreibung
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.
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 Zugriffsfunktion, 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 Zugriffsfunktion, 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 Leistungsfunktionen. 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.

Tabelle 7. DB2 Text Extender-Katalogansicht (Forts.)

Spaltenname	Datentyp	Kann Nullwerte enthalten	Beschreibung
NODENUMBER	INTEGER	Nein	Die Knotennummer der Tabellenpartition.

Den Textvorbereitungsprozeß umkehren

Wenn Text für die Verwendung durch den DB2 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.mytable
      HANDLE      commenthandle
```

Wenn Sie eine Textspalte inaktivieren, passiert folgendes:

- Wenn es sich um eine Mehrindextabelle handelt, d. 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. 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.

Den Textvorbereitungsprozess umkehren

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.mytable
      HANDLE      commenthandle
```

Wenn Sie externe Textdateien inaktivieren, passiert folgendes:

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

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.mytable
```

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 indiziert 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 DB2 Text Extender-Katalogsicht, die beim Aktivieren der Datenbank erstellt wurde.
- Die Deklaration der SQL-Funktionen und der eindeutigen Typen (UDTs) des DB2 Text Extender 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 indiziert 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.

Kapitel 8. Die API-Funktionen zum Suchen und Anzeigen verwenden

In diesem Kapitel wird beschrieben, wie die Such- und Anzeigefunktionen der DB2 Text Extender-API verwendet werden. Eine detaillierte Beschreibung dieser Funktionen finden Sie in „Kapitel 13. API-Funktionen zum Suchen und Anzeigen“ auf Seite 231. Beispiele für Programme, die die API-Funktionen verwenden, finden Sie in „Kapitel 14. Beispiel-API-Programm“ auf Seite 255. In demselben Kapitel wird eine Beispielanzeigefunktion, `DesBrowseDocument`, beschrieben.

Hinweis

Bevor Sie eine Suche durchführen, sollten Sie „Sucharten“ auf Seite 42 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.

Ihre Anwendung einrichten

Ein Anwendungsprogramm, das die DB2 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 DB2 Text Extender-API verwenden.

Schließen Sie in Ihre Anwendung die Datei `des_ext.h` ein, die im Unterverzeichnis `include` des DB2 Text Extender-Installationsverzeichnisses geliefert wird.

Um Ihr Anwendungsprogramm mit der DB2 Text Extender-API zu verwenden, verbinden Sie Ihr Programm mit der API.

Eine UNIX-Anwendung verbinden

Sie müssen die Bibliothek `libdesc1.a` mit Ihrer Anwendung verbinden. Diese Bibliothek befindet sich im Unterverzeichnis `lib` des DB2 Text Extender-Installationsverzeichnisses.

Eine OS/2- oder Windows-Anwendung verbinden

Alle DB2 Text Extender-API-Funktionen befinden sich in der Bibliothek für dynamisches Verbinden `desc1api.dll`. Die externen API-Funktionsaufrufe

Ihre Anwendung einrichten

werden in der Importbibliothek `desclapi.lib` aufgelöst. Sie werden in den entsprechenden Unterverzeichnissen `dll` und `lib` des DB2 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 DB2 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 API-Funktionen für folgende Aufgaben verwenden:

- **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 121 näher beschrieben. Siehe auch „Kapitel 14. Beispiel-API-Programm“ auf Seite 255.

- **Text anzeigen**

Verwenden Sie die folgenden Funktionen in der angegebenen Reihenfolge:

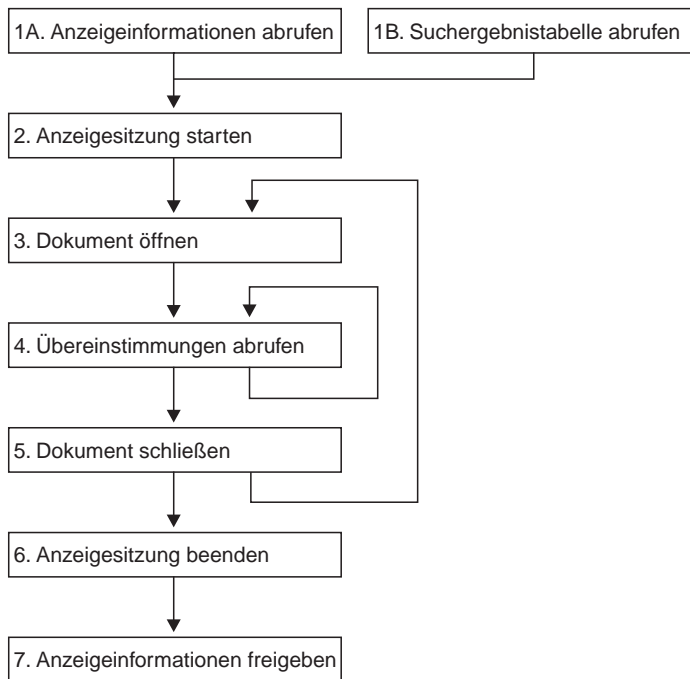


Abbildung 16. Reihenfolge für die Verwendung der API-Funktionen

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 122 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.

Nach Text suchen

Sie können diese Informationen auch unter Verwendung der Funktionen 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 DB2 Text Extender-Funktionen, wenn die Abfrage eine reine Textabfrage ist. Die Funktion geht direkt zum DB2 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. 16 auf Seite 121 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.

Anzeigeeinformationen abrufen (DesGetBrowseInfo)

Die Funktion DesGetBrowseInfo empfängt ein Suchargument und eine Kennung. Sie gibt einen Zeiger auf Anzeigeeinformationen zurück, die von DesStartBrowseSession benötigt werden. Zu den Anzeigeeinformationen gehört eine Liste aller Begriffe, die hervorgehoben werden sollen.

Eine andere Methode zum Abrufen von Anzeigeeinformationen besteht darin, die Anzeigeeoption 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 Anzeigeeinformationen 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 Anzeigeeinformationen, 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 Anzeigeeinformationen von DesGetBrowseInfo oder DesGetSearchResultTable
- Eine Benutzer-ID
- Ein Kennwort

Ausgabe

Diese Funktion gibt eine Anzeigesitzungskennung zurück.

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 270. DES_FAST verwendet die Basistextanalyse, ohne die Verwendung eines Wörterverzeichnisses, um zu ermitteln, welche Begriffe hervorgehoben werden sollen. DES_EXTENDED verwendet die erweiterte Übereinstimmung.

Text anzeigen

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 240 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

Teil 2. Referenz

Kapitel 9. Textvorbereitungs- und Verwaltungsbefehle für den Client

Befehl	Zweck	Seite
CHANGE INDEX SETTINGS	Ändert die Merkmale eines Indexes	131
CHANGE TEXT CONFIGURATION	Ändert die Textkonfigurationseinstellungen	133
CONNECT	Stellt eine Verbindung zu einer Datenbank her	137
DELETE INDEX EVENTS	Löscht Indexereignisse aus einer Protokolltabelle	139
DISABLE DATABASE	Inaktiviert eine Datenbank für die Verwendung durch den DB2 Text Extender	140
DISABLE TEXT COLUMN	Inaktiviert eine Textspalte für die Verwendung durch den DB2 Text Extender und löscht den zugeordneten Index	141
DISABLE TEXT FILES	Inaktiviert Textdateien für die Verwendung durch den DB2 Text Extender und löscht den zugeordneten Index	142
DISABLE TEXT TABLE	Inaktiviert eine Tabelle für die Verwendung durch den DB2 Text Extender und löscht die Indizes, die der Tabelle zugeordnet sind	143
ENABLE DATABASE	Bereitet eine Datenbank für die Verwendung durch den DB2 Text Extender vor	144
ENABLE TEXT COLUMN	Bereitet eine Textspalte für die Verwendung durch den DB2 Text Extender vor und erstellt einen individuellen Textindex für die Spalte	145
ENABLE TEXT FILES	Bereitet Textdateien für die Verwendung durch den DB2 Text Extender vor und erstellt einen individuellen Textindex für die Dateien	154
ENABLE TEXT TABLE	Erstellt einen gemeinsamen Textindex für eine Tabelle	157
GET ENVIRONMENT	Zeigt die aktuellen Einstellungen der Umgebungsvariablen an	162
GET INDEX SETTINGS	Zeigt die Merkmale eines Indexes an	163
GET INDEX STATUS	Zeigt Statusinformationen für einen Index an	165
GET STATUS	Zeigt den aktivierten Status von Datenbanken, Tabellen und Spalten an	166
GET TEXT CONFIGURATION	Zeigt die Textkonfigurationseinstellungen an	167
GET TEXT INFO	Zeigt die Textdaten für eine Textspalte an	168
QUIT	Verläßt den Modus für den DB2 Text Extender-Befehlszeilenprozessor	169

Befehl	Zweck	Seite
REORGANIZE INDEX	Reorganisiert einen Index, um die Sucheffizienz zu verbessern	170
RESET INDEX STATUS	Setzt den Status eines Indexes zurück, um zu ermöglichen, daß er erneut verwendet wird	171
UPDATE INDEX	Aktualisiert einen Textindex	172

In diesem Kapitel wird die Syntax von Textvorbereitungs- und Verwaltungsbefehlen für den Client beschrieben. In „Kapitel 5. Text für Suchvorgänge vorbereiten“ auf Seite 59 und in „Kapitel 7. Verwaltung“ auf Seite 101 wird die Verwendung dieser Befehle beschrieben.

Bevor Sie diese Befehle verwenden, starten Sie den DB2 Text Extender-Befehlszeilenprozessor, indem Sie den Befehl `db2tx` eingeben. Dieser Befehl aktiviert einen interaktiven Eingabemodus, in dem alle nachfolgenden Befehle als DB2 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
```


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 133.

Berechtigung

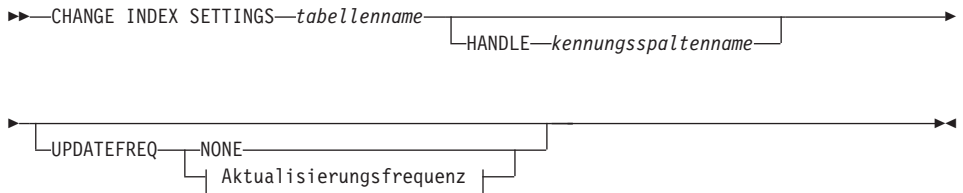
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

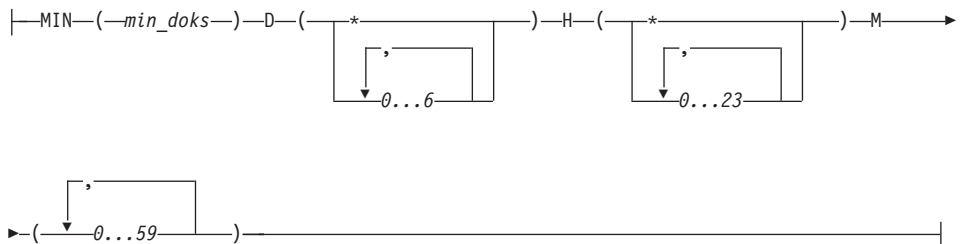
SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Aktualisierungsfrequenz:



CHANGE INDEX SETTINGS (Befehl)

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalte enthält, deren Indexaktualisierungsfrequenz geändert werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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. h., wenn der Index unter Verwendung des Befehls `ENABLE TEXT COLUMN` erstellt wurde.

Wenn jedoch der Index unter Verwendung des Befehls `ENABLE TEXT TABLE` erstellt wurde, d. 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 „Einen Index aktualisieren“ auf Seite 48 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 DB2 Text Extender werden in „Textkonfigurationseinstellungen“ auf Seite 19 beschrieben.

Um diese Einstellungen für eine bestimmte Datenbank zu ändern, **nachdem** die Datenbank aktiviert wurde, verwenden Sie „CHANGE INDEX SETTINGS“ auf Seite 131.

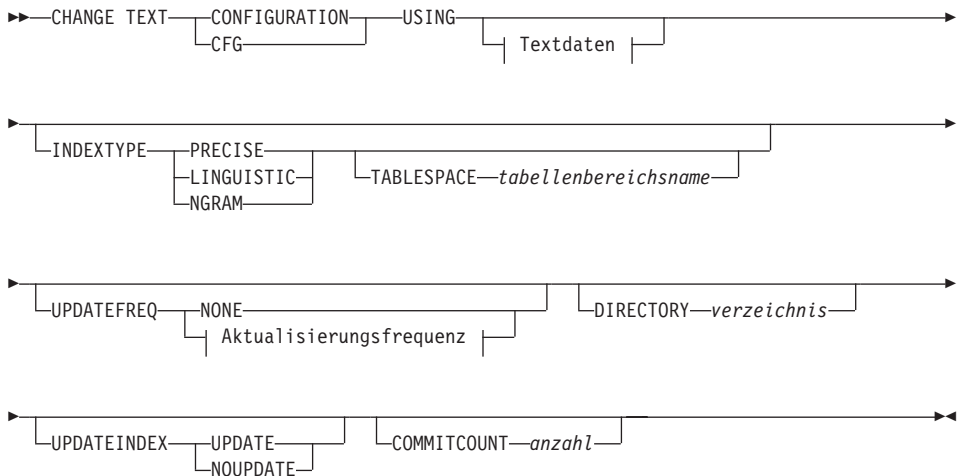
Berechtigung

Sie müssen über folgende Berechtigung verfügen:

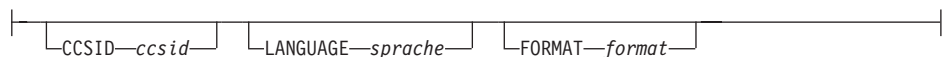
SYSADM-Berechtigung

DBADM-Berechtigung

Befehlsyntax

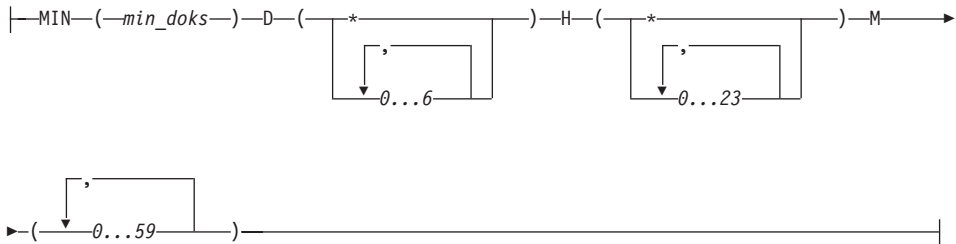


Textdaten:



CHANGE TEXT CONFIGURATION (Befehl)

Aktualisierungsfrequenz:



Befehlsparameter

INDEXTYPE

Um den Standardindextyp zu ändern, eine der folgenden Optionen auswählen. Weitere Informationen finden Sie in „Sucharten“ auf Seite 42.

PRECISE

Die Begriffe werden exakt so indiziert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC

Die Begriffe werden linguistisch verarbeitet, bevor sie indiziert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

NGRAM

Die Begriffe werden indiziert, indem Zeichengruppen syntaktisch analysiert werden, statt ein Wörterverzeichnis zu verwenden. Dieser Wörterverzeichnisstyp 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 die indexspezifischen Tabellen zu enthalten, die vom DB2 Text Extender erstellt wurden (wie beispielsweise die Protokolltabellen). 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 darauf, 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 „Einen Index aktualisieren“ auf Seite 48 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, indiziert 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 1 000 000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen eine DB2-Zwischenanweisung COMMIT 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.

CHANGE TEXT CONFIGURATION (Befehl)

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 37.

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. In „Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis“ auf Seite 257 wird beschrieben, wie Wörterverzeichnisse verwendet werden.

Die unterstützten Sprachen werden in Tabelle 5 auf Seite 57 aufgelistet.

FORMAT `format`

Der Typ des gespeicherten Textdokuments, z. B. WordPerfect oder ASCII. Der DB2 Text Extender benötigt die Informationen beim Indexieren von Dokumenten. Die unterstützten Dokumentformate werden in „Unterstützte Dokumentformate“ auf Seite 33 aufgelistet.

CONNECT (Befehl)

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 DB2 Text Extender-Befehlszeilenprozessor starten“ auf Seite 61.

DELETE INDEX EVENTS

Dieser Befehl löscht Indexierungsereignisse aus der Protokolltabelle eines Indexes für eine bestimmte Kennungsspalte oder Tabelle.

Berechtigung

Keine

Befehlssyntax

```

▶▶—DELETE INDEX EVENTS—tabellenname—┬──────────────────────────────────┬──▶
                                     |
                                     └──HANDLE—kennungsspaltenname──┘
  
```

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, deren Fehlerereignisse aus der Protokolltabelle gelöscht werden sollen. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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.

Wenn Sie die Protokolltabellen bereinigen, sollten Sie Indexereignisse löschen, nachdem Sie den Grund für das betreffende Ereignis geprüft und nach Möglichkeit die Fehlerquelle beseitigt haben.

DISABLE DATABASE (Befehl)

DISABLE DATABASE

Dieser Befehl setzt alle Vorbereitungsarbeiten, die der DB2 Text Extender für eine Datenbank durchgeführt hat, zurück und inaktiviert alle Texttabellen für die Verwendung durch den DB2 Text Extender.

Berechtigung

Sie müssen über folgende Berechtigung verfügen:

SYSADM-Berechtigung

DBADM-Berechtigung

Befehlssyntax

▶—DISABLE DATABASE—◀

Befehlsparameter

Keine

Hinweise zur Verwendung

Dieser Befehl setzt die verbundene Datenbank zurück, so daß sie nicht länger vom DB2 Text Extender durchsucht werden kann, d. h., alle DB2 Text Extender-Texttabellen und -Textspalten in der Datenbank werden inaktiviert. Alle Änderungen, die in der Datenbank durchgeführt wurden, um DB2 Text Extender-Texttabellen, -Textspalten und externe Dateien zu aktivieren, werden zurückgesetzt, alle zugehörigen Textindizes werden gelöscht, die DB2 Text Extender-Katalogsicht TEXTCOLUMNS in der Datenbank wird gelöscht und alle DB2 Text Extender-Auslöser werden gelöscht.

DISABLE TEXT COLUMN

Dieser Befehl inaktiviert eine Textspalte für die Verwendung durch den DB2 Text Extender.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax

►►—DISABLE TEXT COLUMN—*tabellenname*—HANDLE—*kennungsspaltenname*—►►

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die zu inaktivierende Spalte enthält. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

HANDLE *kennungsspaltenname*

Der Name der Kennungsspalte, die für die Verwendung durch den DB2 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 (Befehl)

DISABLE TEXT FILES

Dieser Befehl inaktiviert eine Gruppe von externen Textdateien für die Verwendung durch den DB2 Text Extender.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

- ALTER-Berechtigung
- SELECT-Berechtigung
- UPDATE-Berechtigung

Befehlssyntax

►—DISABLE TEXT FILES—*tabellenname*—HANDLE—*kennungsspaltenname*—►

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 Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname

Der Name der Kennungsspalte, die für die Verwendung durch den DB2 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

Dieser Befehl inaktiviert alle Textspalten in einer Tabelle für die Verwendung durch den DB2 Text Extender.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax

►►—DISABLE TEXT TABLE—*tabellenname*—◄◄

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die zu inaktivierende Spalte enthält. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

Hinweise zur Verwendung

Dieser Befehl inaktiviert alle Textspalten in der Tabelle für die Verwendung durch den DB2 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 (Befehl)

ENABLE DATABASE

Der Befehl aktiviert die aktuelle Datenbank zum Speichern von Textdaten.

Berechtigung

Sie müssen über folgende Berechtigung verfügen:

SYSADM-Berechtigung

DBADM-Berechtigung

Befehlssyntax

▶—ENABLE DATABASE—▶

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 137).

Dieser Befehl bereitet die verbundene Datenbank für die Verwendung durch den DB2 Text Extender vor. Dieser Schritt muß ausgeführt werden, bevor eine DB2 Text Extender-Texttabelle oder -Textspalte in der Datenbank aktiviert werden kann.

ENABLE DATABASE erstellt die DB2 Text Extender-Katalogsicht DB2TX.TEXTINDEXES, die in „Mit der DB2 Text Extender-Katalogsicht arbeiten“ auf Seite 112 beschrieben wird, und die Katalogsicht DB2TX.TEXTCOLUMNS, die für „Leistungsabfragen“ verwendet wird.

Der Befehl erstellt außerdem Textkonfigurationseinstellungen, die in „Textkonfigurationseinstellungen“ auf Seite 19 beschrieben werden.

Darüber hinaus werden einige andere Aufgaben ausgeführt, wie z. B. die Deklaration von eindeutigen Typen des DB2 Text Extender und von DB2 Text Extender-Funktionen.

ENABLE TEXT COLUMN

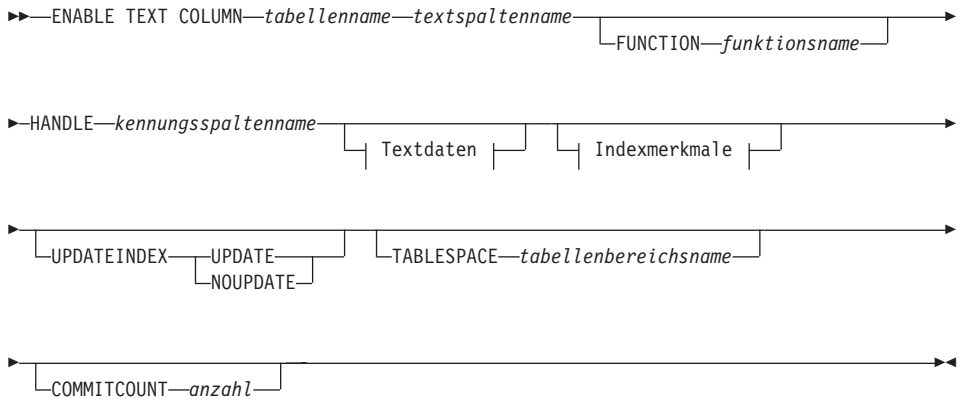
Dieser Befehl aktiviert eine Textspalte für die Verwendung durch den DB2 Text Extender.

Berechtigung

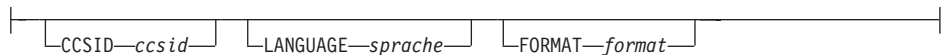
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

- ALTER-Berechtigung
- SELECT-Berechtigung
- UPDATE-Berechtigung

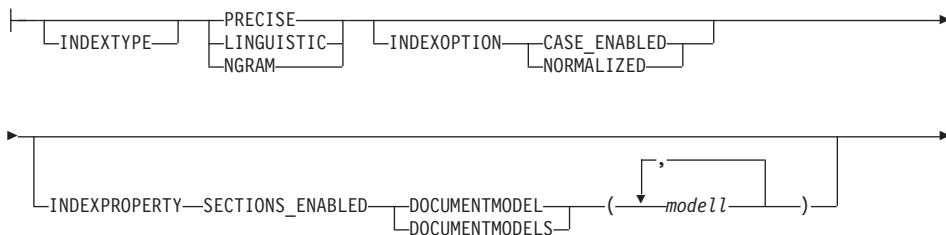
Befehlsyntax



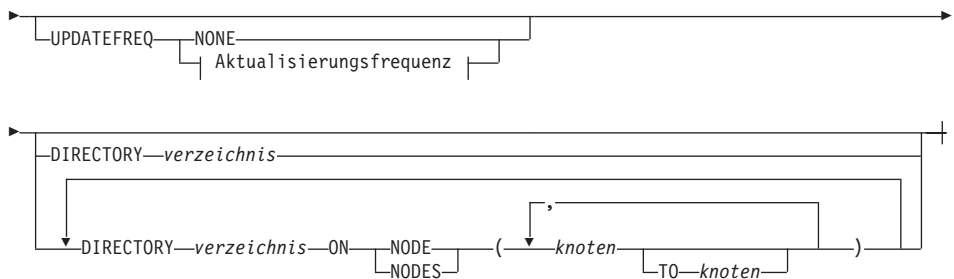
Textdaten:



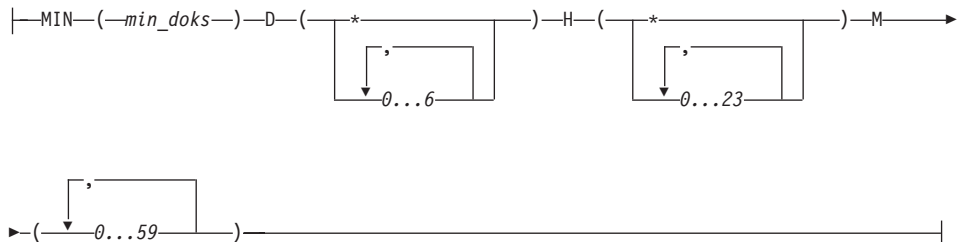
Indexmerkmale:



ENABLE TEXT COLUMN (Befehl)



Aktualisierungsfrequenz:



Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die zu aktivierende Spalte enthält. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

textspaltenname

Der Name der Spalte, die für die Verwendung durch den DB2 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

Der Name einer benutzerdefinierten Funktion, die vom DB2 Text Extender 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 73.

HANDLE kennungsspaltenname

Der Name einer Kennungsspalte, die für die Verwendung durch die Funktionen des DB2 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 Sie beim Aktivieren einer Textspalte für einen Ngram-Index eine CCSID angeben, muß diese dieselbe wie die CCSID der Datenbank sein. Zudem muß die während der Suche verwendete CCSID (die CCSID der Datenbank) mit dieser CCSID übereinstimmen.

Verwenden Sie folgenden Befehl, um die Standard-CCSID zu finden:
`db2tx get text cfg`

Die Standardeinstellung bei der Installation ist die Datenbank-CCSID. 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 37.

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. In „Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis“ auf Seite 257 wird beschrieben, 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 Tabelle 5 auf Seite 57 aufgelistet.

ENABLE TEXT COLUMN (Befehl)

FORMAT format

Der Typ des gespeicherten Textdokuments, z. B. WordPerfect oder ASCII. Der DB2 Text Extender benötigt die Informationen beim Indizieren von Dokumenten. Die unterstützten Dokumentformate werden in „Unterstützte Dokumentformate“ auf Seite 33 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 „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50.) 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 indiziert. Wenn kein Modell für eine erkannte DTD definiert wurde, wird das Dokument nicht indiziert. XML-Dokumente können verschachtelten Abschnitte enthalten.

Für diese Formate müssen Sie die Strukturinformationen in einer Dokumentmodelldatei angeben. Siehe „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50. Wenn das Format TDS und INDEXPROPERTY SECTION_ENABLED angegeben sind, wird davon ausgegangen, daß das Dokumentformat ASCII_SECTIONS ist.

Befehle, die nicht in der Dokumentmodelldatei definiert sind, werden auf normale Art und Weise gemäß dem Indextyp indiziert.

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 „Sucharten“ auf Seite 42.

PRECISE

Die Begriffe werden exakt so indiziert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC

Die Begriffe werden linguistisch verarbeitet, bevor sie indiziert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

NGRAM

Die Begriffe werden indiziert, 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.

Dokumente im XML-Format werden nicht für Ngram-Indizes 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 indiziert werden, daß eine Suche unter Beachtung der Groß-/Kleinschreibung möglich ist. Weitere Informationen finden Sie in „Suchen nach grober Übereinstimmung oder Suchen in DBCS-Dokumenten durchführen“ auf Seite 45.

ENABLE TEXT COLUMN (Befehl)

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) *modell*

Merkmale eines ausgewählten Indextyps.

SECTIONS_ENABLED gibt an, daß der ausgewählte Indextyp Informationen zur Dokumentstruktur enthalten kann.

DOCUMENTMODEL/DOCUMENTMODELS *modell* 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 „Einen Index aktualisieren“ auf Seite 48 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 indiziert 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 `"/txinsnnn"`, wobei `nnn` die Knotennummer ist.

Dies ist ein Verzeichnis auf dem System, auf dem der DB2 Text Extender-Server läuft. Wenn das Verzeichnis noch nicht besteht, wird es erstellt. Wenn es bereits besteht, stellen Sie auf UNIX-Plattformen sicher, daß das Verzeichnis über Lese-/Schreibberechtigungen für den Exemplareigner verfügt. Prüfen Sie darüber hinaus die abgeschirmte Benutzer-ID (Fenced User ID), wenn eine angegeben wurde, die sich von der unterscheidet, die bei der Erstellung des DB2-Exemplars angegeben wurde.

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.

ENABLE TEXT COLUMN (Befehl)

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 indiziert werden sollen, ohne auf die nächste regelmäßige Indexierung zu warten, die durch UPDATEFREQ gesetzt ist.

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.

TABLESPACE tabellenbereichsname

Der Name des Tabellenbereichs für den Index, der intern für die Kennungsspalte erstellt wird. Der Tabellenbereich muß vorher erstellt worden sein.

COMMITCOUNT anzahl

Ein Wert zwischen 500 und 1 000 000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen DB2 eine Zwischenanweisung COMMIT ausgeben 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 Funktionen des DB2 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. h. Einfügungen, Aktualisierungen und Löschungen. Einfügungs-, Aktualisierungs- und Löschauslöser werden für die Textspalte definiert, um die Protokolltabelle automatisch auf dem neuesten Stand zu halten.

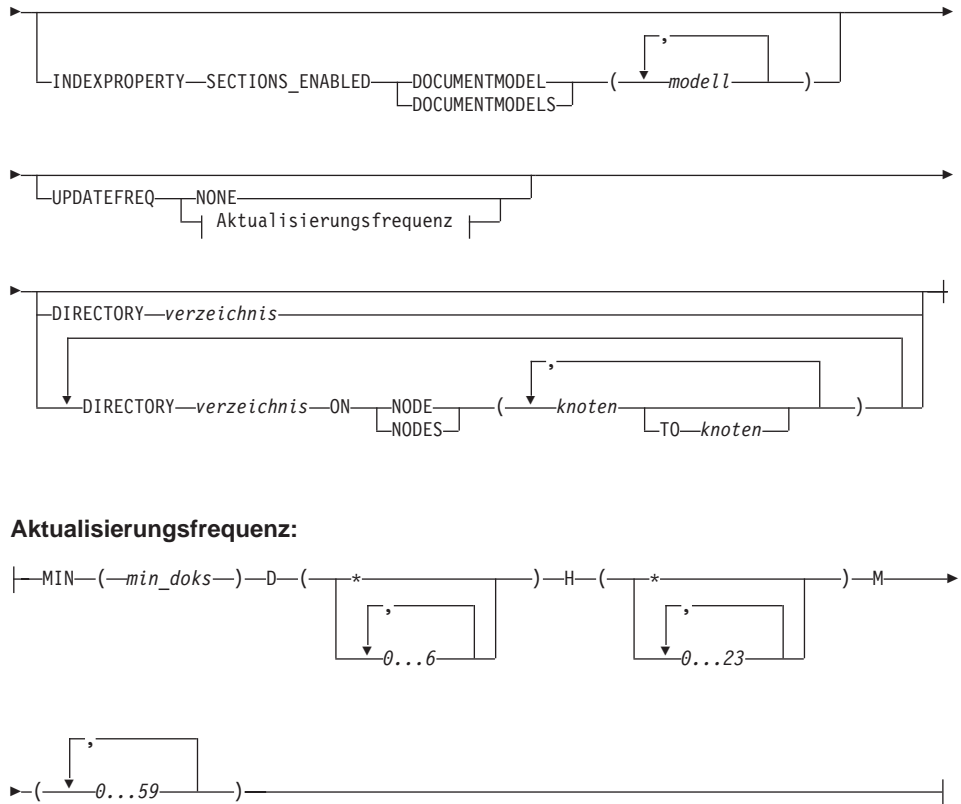
Wenn die Textspalte, die Sie aktivieren, zu einer Tabelle gehört, die Teil einer Knotengruppe mit mehreren Knoten ist, muß das Indexverzeichnis, das Sie angeben, auf allen physischen Knoten verfügbar sein. Wenn Sie das Standardverzeichnis verwenden, das in der Textkonfiguration angegeben ist, stellen 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 DB2 Text Extender aktivierte Tabelle enthält, müssen Sie die Tabelle erneut indexieren.

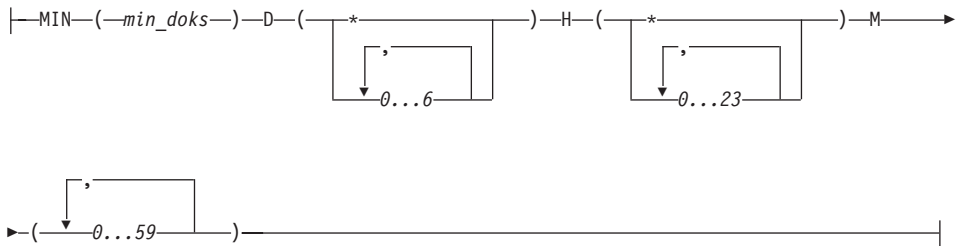
Hinweis

Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 72.

ENABLE TEXT FILES (Befehl)



Aktualisierungsfrequenz:



Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die den zu indexierenden externen Textdateien zugeordnet werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

Eine Beschreibung der anderen Befehlsparameter finden Sie in „ENABLE TEXT COLUMN“ auf Seite 145.

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 Funktionen des DB2 Text Extender verwendet. Ein Index wird erstellt, der diesen Dateien zugeordnet ist.

ENABLE TEXT FILES (Befehl)

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 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 102.

Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 72.

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

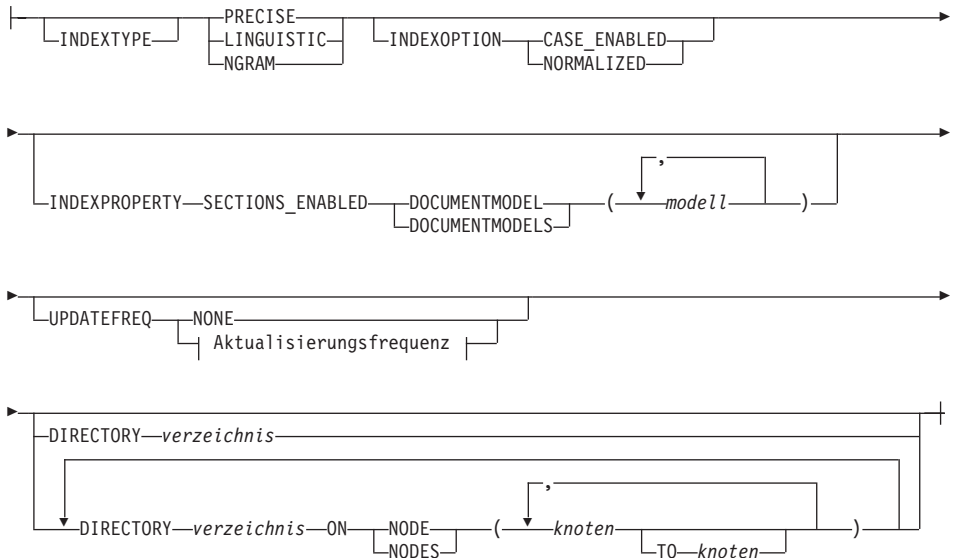
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

- ALTER-Berechtigung
- SELECT-Berechtigung
- UPDATE-Berechtigung

Befehlssyntax

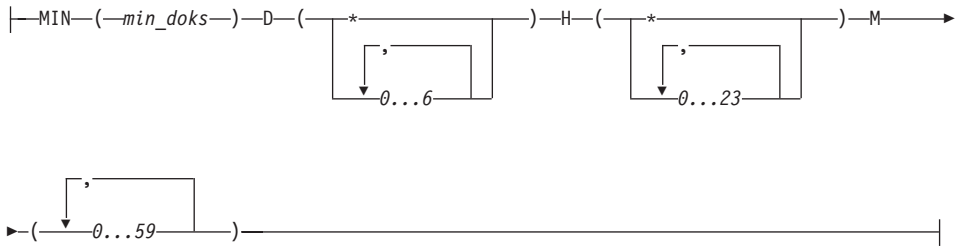
►►—ENABLE TEXT TABLE—*tabellenname*—
└─ Indexmerkmale ─┘

Indexmerkmale:



ENABLE TEXT TABLE (Befehl)

Aktualisierungsfrequenz:



Befehlsparameter

tabellenname

Der Name der Texttabelle, die in der verbundenen Datenbank aktiviert werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname entspricht Ihrer Benutzer-ID.

INDEXTYPE

Der Typ des zu erstellenden Indexes. Weitere Informationen finden Sie in „Sucharten“ auf Seite 42.

PRECISE

Die Begriffe werden exakt so indiziert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC

Die Begriffe werden linguistisch verarbeitet, bevor sie indiziert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

NGRAM

Die Begriffe werden indiziert, indem Zeichengruppen syntaktisch analysiert werden, statt ein Wörterverzeichnis zu verwenden. Dieser Wörterverzeichnisstyp ist verbindlich, wenn die Dokumente, die Sie indizieren, 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 „Suchen nach grober Übereinstimmung oder Suchen in DBCS-Dokumenten durchführen“ auf Seite 45.

INDEXPROPERTY SECTIONS_ENABLED DOCUMENTMODEL(S) modell

Merkmale eines ausgewählten Indextyps.

SECTIONS_ENABLED gibt an, daß der ausgewählte Indextyp Informationen zur Dokumentstruktur enthalten kann.

DOCUMENTMODEL/DOCUMENTMODELS *modell* 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 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 „Einen Index aktualisieren“ auf Seite 48 beschrieben.

ENABLE TEXT TABLE (Befehl)

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 indiziert 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 Verzeichnis auf dem System, auf dem der DB2 Text Extender-Server läuft. Wenn das Verzeichnis noch nicht besteht, wird es erstellt. Wenn es bereits besteht, stellen Sie auf UNIX-Plattformen sicher, daß das Verzeichnis über Lese-/Schreibberechtigungen für den Exemplareigner verfügt. Prüfen Sie darüber hinaus die abgeschirmte Benutzer-ID (Fenced User ID), wenn eine angegeben wurde, die sich von der unterscheidet, die bei der Erstellung des DB2-Exemplars angegeben wurde.

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. 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 DB2 Text Extender indexiert die Textdokumente, die in der Protokolltabelle aufgelistet sind, regelmäßig, so häufig, wie durch das Schlüsselwort UPDATEFREQ angegeben.

GET ENVIRONMENT (Befehl)

GET ENVIRONMENT

Dieser Befehl zeigt die Einstellungen der Umgebungsvariablen an.

Berechtigung

Keine

Befehlssyntax

▶—GET ENVIRONMENT—◀

Befehlsparameter

Keine

Hinweise zur Verwendung

Folgende Umgebungsvariablen werden angezeigt:

DB2INSTANCE

Der DB2-Exemplarname.

DB2DBDFT

Der Standarddatenbankname.

DB2TX_INSTOWNER

Der DB2 Text Extender-Exemplarname. Nur dann erforderlich, wenn Sie eine Verbindung zu einem UNIX-Server herstellen.

DB2TX_INSTOWNERHOMEDIR

Das Ausgangsverzeichnis des Exemplareigners. Nur dann erforderlich, wenn Sie eine Verbindung zu einem UNIX-Server herstellen.

GET INDEX SETTINGS (Befehl)

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

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 indextierten Dokumente
- Die Anzahl der indextierten Dokumente im Primärindex
- Die Anzahl der indextierten Dokumente im Sekundärindex
- Fehlerereignisse

Berechtigung

Keine

Befehlssyntax

```

▶▶—GET INDEX STATUS—tabellenname—┬──────────────────────────────────────────┬──────────────────▶
                                     └─HANDLE—kennungsspaltenname—┘
  
```

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Status angezeigt werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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 (Befehl)

GET STATUS

Dieser Befehl zeigt Informationen zum aktivierten Status von Datenbanken, Tabellen oder Textspalten an.

Berechtigung

Keine

Befehlssyntax

▶—GET STATUS—◀

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 133.

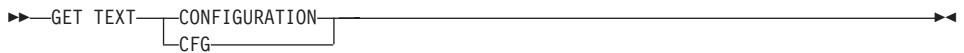
Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

SELECT-Berechtigung

Befehlssyntax

```
▶▶ GET TEXT CONFIGURATION
```



Befehlsparameter

Keine

Hinweise zur Verwendung

Ein Beispiel für die Informationen zur Textkonfiguration finden Sie in „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 107.

GET TEXT INFO (Befehl)

GET TEXT INFO

Dieser Befehl zeigt die Textdateneinstellungen für Textspalten an:

- CCSID
- Sprache
- Format

Berechtigung

Keine

Befehlssyntax

```
▶▶—GET TEXT INFO—tabellenname—┐  
└—HANDLE—kennungsspaltenname—┘▶▶
```

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Textdateneinstellungen angezeigt werden sollen. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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 DB2 Text Extender und gibt die Steuerung an das Betriebssystem zurück.

Berechtigung

Keine

Befehlssyntax

»—QUIT—«

Befehlsparameter

Keine

Hinweise zur Verwendung

Die Verbindung zur Datenbank wird beendet.

REORGANIZE INDEX (Befehl)

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 DB2 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

```
►►—REORGANIZE INDEX—tabellenname—┐—————┐—————►
                                └—HANDLE—kennungsspaltenname—┘
```

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, deren Index reorganisiert werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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.

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 109).

Berechtigung

Keine

Befehlssyntax

```

▶▶—RESET INDEX STATUS—tabellenname—┬──────────────────────────────────┬──────────────────▶
                                     |HANDLE—kennungsspaltenname—|
    
```

Befehlsparameter

tabellenname

Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Status zurückgesetzt werden soll. Der Name muß einen expliziten Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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 (Befehl)

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.

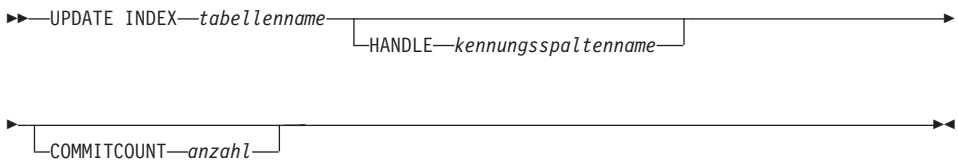
Damit Aktualisierungen von externen Dateien im Index widergespiegelt werden, erzwingen Sie, daß ein "Änderungseintrag" in die Protokolltabelle gestellt wird, indem Sie eine UPDATE-Anweisung für die entsprechende Kennungsspalte ausgeben. Ein Beispiel finden Sie in „Einen Index für externe Dateien aktualisieren“ auf Seite 102.

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 Schemennamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemenname 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.

COMMITCOUNTanzahl

Ein Wert zwischen 500 und 1 000 000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen DB2 eine Zwischenanweisung COMMIT ausgeben 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 10. 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 DB2 Text Extender-Servers und zum Erstellen einer Beispieldatenbank und einer Beispieltabelle ausführen können. „Einen DB2 Text Extender-Server konfigurieren und verwalten“ auf Seite 22 beschreibt die Verwendung dieser Befehle.

Befehl	Zweck	Seite
TXICRT	Ein DB2 Text Extender-Exemplar erstellen	176
TXIDROP	Ein DB2 Text Extender-Exemplar löschen	178
TXILIST (nur UNIX)	DB2 Text Extender-Exemplare auflisten (nur UNIX)	179
TXIPCLEAN (nur UNIX)	Bereinigt Systemressourcen nach einem Systemfehler (nur UNIX)	180
TXNADD	DB2 Text Extender-Server hinzufügen	181
TXNCHECK	Die DB2 Text Extender-Datenbankintegrität prüfen	182
TXNDROP	DB2 Text Extender-Server löschen	183
TXSAMPLE	Erstellt und aktiviert eine Beispieltabelle	184
TXSTART	Die DB2 Text Extender-Services starten	185
TXSTATUS	Den Status des Suchservices anzeigen	186
TXSTOP	Die DB2 Text Extender-Services stoppen	187
TXTHESC	Eine Thesaurus-Definitionsdatei kompilieren	188
TXTHESN	Eine Ngram-Thesaurus-Definitionsdatei kompilieren	190
TXTRACE	Trace-Informationen erstellen	192
TXVERIFY	Eine Beispieldatenbank erstellen und aktivieren	197
TXWIZARD	Ein Beispielprogramm zur Indexierung und Suche (nur AIX, Windows NT und Windows 2000)	198

TXICRT (Befehl)

TXICRT

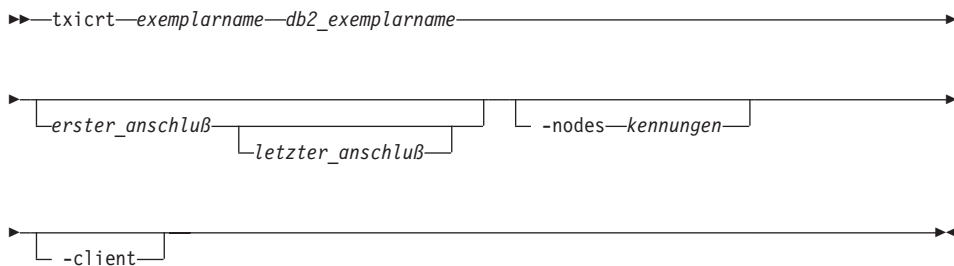
Dieser Befehl erstellt ein DB2 Text Extender-Exemplar.

Berechtigung

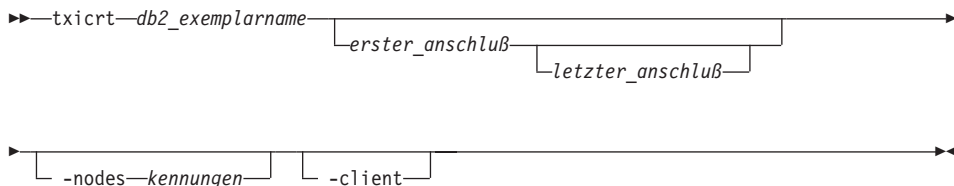
(Nur UNIX) Root

Befehlsyntax

Für UNIX:



Für Windows NT und Windows 2000:



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 DB2 Text Extender-Client haben. Wenn Sie diesen Parameter nicht angeben, konfiguriert der Befehl TXICRT ein Client/Server-Exemplar des DB2 Text Extender.

Hinweise zur Verwendung

Das Aktivieren von Texttabellen oder Spalten ist nur auf Knoten möglich, auf denen Sie ein DB2 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_07_01/instance`.

Auf Windows NT- oder Windows 2000-Systemen wird während der Installation in einer Umgebung mit nicht partitionierten Datenbanken ein Standardexemplar erstellt.

TXIDROP (Befehl)

TXIDROP

Dieser Befehl löscht ein DB2 Text Extender-Exemplar zusammen mit all seinen Indizes.

Berechtigung

(Nur UNIX) Root

Befehlssyntax

Für UNIX:

```
►—txidrop—exemplarname—db2_exemplarname—◄◄
```

Für Windows NT und Windows 2000:

```
►—txidrop—exemplarname—◄◄
```

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_07_01/instance`.

TXILIST (nur UNIX)

Dieser Befehl listet die Exemplare für die aktuelle DB2 Text Extender-Version auf.

Berechtigung

SYSADM, DBADM

Befehlssyntax

▶▶—txilist—————▶▶

Befehlsparameter

Keine

Hinweise zur Verwendung

Wenn mehr als eine Version des DB2 Text Extender installiert ist, listet dieser Befehl nur die Exemplare für die DB2 Text Extender-Version auf, mit der Sie momentan arbeiten, d. h. die Version, die durch das aktuelle Verzeichnis bestimmt ist.

TXIPCLEAN (Befehl, nur UNIX)

TXIPCLEAN (nur UNIX)

Dieser Befehl bereinigt Systemressourcen für den angemeldeten Benutzer, d. h., er beendet den Prozeß dessdem mit dem Befehl 'Kill', bereinigt gemeinsam benutzten Speicher und segmentiert Nachrichtenwarteschlangen und Semaphors.

Berechtigung

Mitglied der Gruppe smadmin

Befehlssyntax

►—txipclean—◄

Befehlsparameter

Keine

Hinweise zur Verwendung

Dieser Befehl ist mit Vorsicht zu verwenden; er bereinigt nicht nur die DB2 Text Extender-Ressourcen, sondern auch alle gemeinsam benutzten Ressourcen des angemeldeten Benutzers. Melden Sie sich immer als Exemplareigner an, um die Bereinigung der Exemplarressourcen zu beschränken.

Es kann Situationen geben, in denen Sie den DB2 Text Extender nicht richtig starten oder stoppen können. Versuchen Sie in diesem Fall, erst DB2 und den DB2 Text Extender zu stoppen, bevor Sie den Befehl TXIPCLEAN ausführen.

TXNCHECK (Befehl)

TXNCHECK

Dieser Befehl prüft die DB2 Text Extender-Integrität für eine bestimmte Datenbank.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

```
▶▶ txncheck datenbankname ▶▶
```

Befehlsparameter

datenbankname

Der Name der zu überprüfenden Datenbank. Wenn kein Datenbankname angegeben ist, wird der Name der DB2-Standarddatenbank verwendet.

Hinweise zur Verwendung

Eine DB2 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 DB2 Text Extender-fähigen Tabellen aufgrund einer neu verteilten Knotengruppe erneut indexiert werden müssen.

TXNDROP

Dieser Befehl löscht einen DB2 Text Extender-Server für einem bestimmten Knoten.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

►►—txndrop—*knotennummer*—————►◄

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.

TXSAMPLE (Befehl)

TXSAMPLE

Dieser Befehl erstellt eine Beispieltabelle, lädt englische Beispieldokumente in die Spalte COMMENT und aktiviert die Textspalte. Siehe „Eine Beispieldatenbank zur Installationsprüfung vorbereiten“ auf Seite 17.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

```
► txsample—datenbankname—┬───┬───►  
                           └──benutzer-ID—kennwort──┘
```

Befehlsparameter

datenbankname

Der Name der bereits durch den DB2 Text Extender aktivierten 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

Wenn Ihre Datenbank-CCSID nicht 850 ist, verwenden Sie das Beispiel für den Aktivierungsschritt für den Indextyp Ngram und ändern Sie darin die CCSID, damit sie mit Ihrer Datenbank-CCSID übereinstimmt.

Hinweis

Dieser Befehl kann auch auf einer Client-Workstation verwendet werden.

TXSTATUS (Befehl)

TXSTATUS

Dieser Befehl zeigt an, ob der DB2 Text Extender aktiv ist.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

```
►► txstatus [NODENUM knotennummer] ◄◄
```

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 DB2 Text Extender-Services.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

```
▶▶ txstop [NODENUM knotennummer] ▶▶
```

Befehlsparameter

NODENUM knotennummer

Die Knotennummer des Servers, dessen Status überprüft werden soll.

Wenn keine Knotennummer angegeben ist, werden alle Server gestoppt.

Hinweise zur Verwendung

Mit diesem Befehl wird DB2 nicht gestoppt.

TXTHESC (Befehl)

TXTHESC

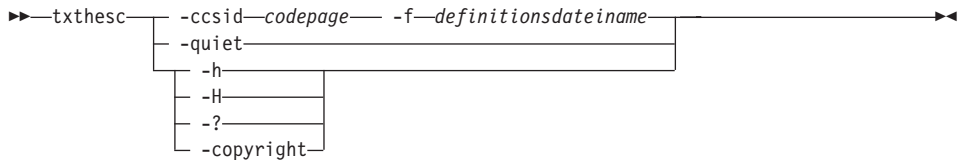
Dieser Befehl kompiliert eine Thesaurus-Definitionsdatei. Dieser Thesaurus kann nur für Suchen in präzisen oder linguistischen Indizes verwendet werden.

Informationen zum Kompilieren einer Ngram-Thesaurus-Definitionsdatei finden Sie im Abschnitt „TXTHESN“ auf Seite 190.

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

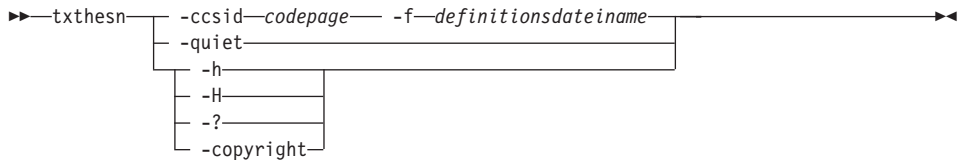
TXTHESN

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 *lkn*, 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 „CCSIDs“ auf Seite 37.

-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 281 beschrieben ist.

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
```

TXTRACE (Befehl)

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.

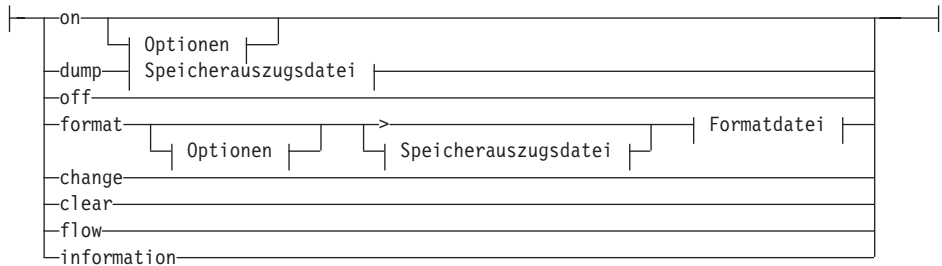
Berechtigung

(Nur UNIX) SYSADM, DBADM

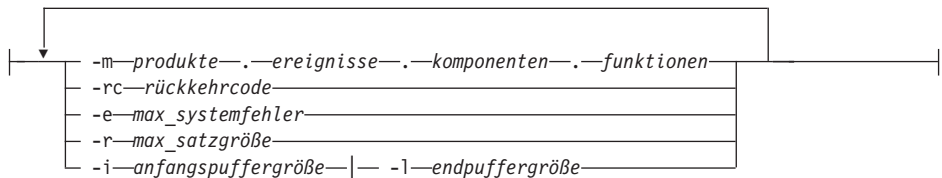
Befehlssyntax



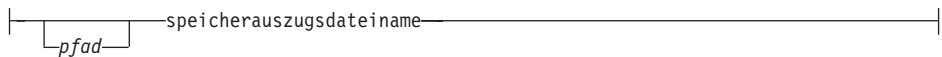
Parameter:



Optionen:



Speicherauszugsdatei:



Formatdatei:**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

TXTRACE (Befehl)

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 DB2 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 DB2 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

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. B. -e 50) angeben, wenn

TXTRACE (Befehl)

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 24.

TXVERIFY

Dieser Befehl erstellt eine Beispieldatenbank und aktiviert sie.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax

```
►►—txverify—datenbankname—┬──────────────────────────────────────────┬──►  
                             └benutzer-ID—kennwort—┘
```

Befehlsparameter**datenbankname**

Der Name der Datenbank, in der die Beispieldatenbank 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.

TXWIZARD (Befehl)

TXWIZARD

Dieser Befehl ist nur auf den Systemen AIX, Windows NT und Windows 2000 verfügbar. Er wird als Beispiel dafür bereitgestellt, wie Indizes für ausgewählte Tabellen oder Spalten erstellt und wie diese Indizes durchsucht werden.

Befehlssyntax

► txwizard ◄

Befehlsparameter

Keine

Hinweise zur Verwendung

Dieser Beispiellassistent führt Sie durch die Schritte, die für die Suche in ausgewählten Spalten einer Tabelle unter Verwendung der Suchfunktionen des DB2 Text Extender erforderlich sind. Sie können den Java-Quellcode verwenden, der mit dem Beispiellassistenten geliefert wird, um entweder Ihre eigene Textvorbereitungs- und Suchanwendung zu erstellen oder zu lernen, wie die DB2 Text Extender-Funktionen verwendet und innerhalb einer Java-Umgebung aufgerufen werden können.

Unter AIX ist der Java-Quellcode des Assistenten in folgendem Verzeichnis gespeichert:

```
/usr/lpp/db2tx_06_01/samples/wizard
```

Unter Windows NT und Windows 2000 ist er in folgendem Verzeichnis gespeichert:

```
%DMBMPATH%\samples\wizard
```

Kapitel 11. Suchfunktionen

Der DB2 Text Extender bietet SQL-Funktionen zum Suchen in Textdokumenten, die in DB2 gespeichert sind, und zum Arbeiten mit den Ergebnissen einer Suche. Einige der Parameter dieser Funktionen sind Datentypen mit dem Namen *eindeutige Typen*, die mit dem DB2 Text Extender geliefert werden.

In diesem Kapitel werden die SQL-Funktionen und eindeutigen Typen des DB2 Text Extender beschrieben.

Der DB2 Text Extender liefert eine Eingabedatei für den DB2-Befehlszeilenprozessor mit dem Namen `txsample.udf`. Darin sind Beispiele von DB2 Text Extender-Funktionen enthalten, die mit der Beispieltabelle ausgeführt werden, die in „Eine Beispieldatenbank zur Installationsprüfung vorbereiten“ auf Seite 17 beschrieben ist. Verwenden Sie diese Datei, um Beispiele der Syntax der Textvorbereitungs- und Suchfunktionen anzuzeigen.

Die eindeutigen Typen des DB2 Text Extender

Eindeutiger Typ	Quellendatentyp	Kommentare
DB2TEXTH	VARCHAR(60) FOR BIT DATA	<p>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.</p> <p>Kennungen werden in Spalten gespeichert, die der DB2 Text Extender erstellt und der jeweiligen Textspalte zuordnet.</p> <p>Die Funktionen <code>HANDLE</code> und <code>INIT_TEXT_HANDLE</code> geben diesen Datentyp zurück.</p>

Eindeutige Typen des DB2 Text Extender

Eindeutiger Typ	Quellendatentyp	Kommentare
DB2TEXTFH	VARCHAR(210) FOR BIT DATA	<p>Dateikennung. Eine Zeichenfolge variabler Länge, die Informationen enthält, die für die Indexierung einer externen Textdatei - eine Datei, die außerhalb der Steuerung von DB2 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.</p> <p>Dateikennungen werden in Spalten gespeichert, die der DB2 Text Extender erstellt und der jeweiligen Gruppe von externen Dateien zuordnet.</p> <p>Die Funktionen FILE und INIT_TEXT_HANDLE geben diesen Datentyp zurück.</p>
DB2TEXTHLISTP	VARCHAR(16) FOR BIT DATA	<p>Kennungslistenzeiger. Ein Zeiger auf eine Liste von Kennungen, die Textdokumenten zugeordnet sind, die in einer Suche gefunden wurden.</p> <p>Die Funktion HANDLE_LIST gibt diesen Datentyp zurück.</p>
DB2TEXTFHLISTP	VARCHAR(16) FOR BIT DATA	<p>Kennungslistenzeiger. Ein Zeiger auf eine Liste von Kennungen, die externen Dateien zugeordnet sind, die in einer Suche gefunden wurden.</p> <p>Die Funktion HANDLE_LIST gibt diesen Datentyp zurück.</p>

Eine Zusammenfassung der DB2 Text Extender-Funktionen

Suchfunktion	Zweck	Seite
CCSID	Gibt die CCSID aus der Kennung zurück	202
CONTAINS	Führt eine Suche nach Text in einem bestimmten Dokument durch	203
FILE	Gibt den Pfad und Namen einer Datei in einer bestehenden Kennung zurück oder ändert ihn	204

Zusammenfassung der DB2 Text Extender-Funktionen

Suchfunktion	Zweck	Seite
FORMAT	Gibt die Dokumentformateinstellung in einer Kennung zurück oder ändert sie	205
HANDLE ¹	Gibt eine Kennung aus einer Liste von Kennungen zurück	206
HANDLE_LIST ¹	Sucht eine Liste von Kennungen und gibt sie zurück	207
INIT_TEXT_HANDLE	Gibt eine teilweise initialisierte Kennung zurück, die Informationen, wie z. B. Format- und Spracheinstellungen, enthält	208
LANGUAGE	Gibt die Spracheinstellung in einer Kennung zurück oder ändert sie	210
NO_OF_DOCUMENTS ¹	Gibt die Anzahl von Dokumenten zurück, die in einer Kennungsliste aufgelistet sind	211
NO_OF_MATCHES	Sucht die Anzahl der gefundenen Übereinstimmungen und gibt sie zurück	212
RANK	Sucht den Rangordnungswert eines gefundenen Textdokuments und gibt ihn zurück	213
REFINE	Empfängt ein Suchargument und ein Suchargument zur Feinanpassung und gibt ein kombiniertes Suchargument zurück	214
SEARCH_RESULT	Gibt eine Zwischentabelle mit den Suchergebnissen des angegebenen Suchbegriffs zurück	215

Beispiele zur Verwendung von DB2 Text Extender-Funktionen sind in „Kapitel 6. Vorgehensweise bei Suchvorgängen“ auf Seite 79 enthalten.

1. Diese Suchfunktionen stammen aus einem früheren Release des DB2 Text Extender. Aus Kompatibilitätsgründen werden sie weiterhin unterstützt, ihre Funktionalität wurde jedoch durch die Suchfunktion PDF SEARCH_RESULT ersetzt.

CCSID (Funktion)

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 37 beschrieben. Er wird durch den Befehl ENABLE TEXT COLUMN für jede Textspalte gesetzt.

Funktionssyntax

► CCSID (—*kennung*—) ◄

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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

►►—CONTAINS—(—*kennung*—,—*suchargument*—)—————►►

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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 12. Syntax von Suchargumenten“ auf Seite 217.

FILE (Funktion)

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

►►FILE(—*kennung*—)◄◄

►►FILE(—*kennung*—, —*dateiname*—)◄◄

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-Exemplareigner zulassen.

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 DB2TEXTFH oder DB2TEXTH.

Dies ist der Formatparameter, der für die Indexierung des entsprechenden Textdokuments verwendet wird. Die unterstützten Dokumentformate werden in „Unterstützte Dokumentformate“ auf Seite 33 aufgelistet.

Funktionssyntax

►►—FORMAT⁽¹⁾—(—*kennung*—)—————►►

Anmerkungen:

- 1 Gibt einen Formatwert vom Typ VARCHAR(30) zurück.

►►—FORMAT⁽¹⁾—(—*kennung*—, —*format*—)—————►►

Anmerkungen:

- 1 Gibt eine Kennung vom Typ DB2TEXTFH oder DB2TEXTH zurück.

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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 (Funktion)

HANDLE

Die Funktion HANDLE gibt eine Kennung vom Typ DB2TEXTFH oder DB2TEXTH 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 DB2 Text Extender. Aus Kompatibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde jedoch durch die Suchfunktion SEARCH_RESULT ersetzt.

Funktionssyntax

►—HANDLE—(—*kennungsliste*—,—*integer*—)—————►

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. 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 DB2 Text Extender. Aus Kompatibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde jedoch durch die Suchfunktion SEARCH_RESULT ersetzt.

Funktionssyntax

►►—HANDLE_LIST—(—*kennung*—,—*suchargument*—)—————►►

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 12. Syntax von Suchargumenten“ auf Seite 217.

INIT_TEXT_HANDLE (Funktion)

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-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

▶▶ INIT_TEXT_HANDLE(—*format*—,—*sprache*—)————▶▶

▶▶ INIT_TEXT_HANDLE(—*CCSID*—,—*format*—,—*sprache*—,—*dateiname*—)————▶▶

Funktionsparameter

format

Eine Zeichenfolge vom Typ VARCHAR(30), die die neue Einstellung für das Dokumentformat angibt. Die unterstützten Formate werden in „Unterstützte Dokumentformate“ auf Seite 33 aufgelistet.

sprache

Eine Zeichenfolge vom Typ VARCHAR(30), die die neue Einstellung für die Dokumentsprache angibt. Die unterstützten Sprachen werden in Tabelle 5 auf Seite 57 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-Exemplareigner in den Dateizugriffsberechtigungen enthalten sein. Für OS/2- und Windows-Benutzer müssen die Dateizugriffsberechtigungen die Anmeldebenedutzer-IDs enthalten.

LANGUAGE (Funktion)

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 DB2TEXTFH oder DB2TEXTH.

Dies ist der Sprachparameter, der für die Indexierung des entsprechenden Textdokuments verwendet wird. Die unterstützten Sprachen werden in Tabelle 5 auf Seite 57 aufgelistet.

Funktionssyntax

►►—LANGUAGE⁽¹⁾—(—*kennung*—)——►►

Anmerkungen:

- 1 Gibt einen Sprachwert vom Typ VARCHAR(30) zurück.

►►—LANGUAGE⁽¹⁾—(—*kennung*—, —*sprache*—)——►►

Anmerkungen:

- 1 Gibt eine Kennung vom Typ DB2TEXTFH oder DB2TEXTH zurück.

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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 DB2 Text Extender. Aus Kompatibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde jedoch durch die Suchfunktion SEARCH_RESULT ersetzt.

Funktionssyntax

►►—NO_OF_DOCUMENTS—(—*kennungsliste*—)—————◄◄

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 Funktionen 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 (Funktion)

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

►►NO_OF_MATCHES(—*kennung*—,—*suchargument*—)◄◄

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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 12. Syntax von Suchargumenten“ auf Seite 217.

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

►►—RANK—(—*kennung*—,—*suchargument*—)—————►►

Funktionsparameter

kennung

Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH oder DB2TEXTH 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 12. Syntax von Suchargumenten“ auf Seite 217.

REFINE (Funktion)

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

►►REFINE(—suchargument—,—suchargument—)◄◄

Funktionsparameter

suchargument

Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe „Kapitel 12. Syntax von Suchargumenten“ auf Seite 217.

Das Suchargument darf keinen der Suchparameter IS ABOUT, THESAURUS oder EXPAND enthalten.

SEARCH_RESULT

Die Funktion SEARCH_RESULT gibt das Ergebnis einer Suche in einer Zwischentabelle zurück. Diese Funktion 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 Funktion ist schneller als CONTAINS oder RANK, wenn große Tabellen verarbeitet werden.

Funktionssyntax

►►—SEARCH_RESULT—(—*schema*—,—*tabelle*—,—*kennung*—,—*suchargument*—)—►►

Funktionsparameter

schema

Der Schemenname der Tabelle, in der Sie suchen.

tabelle

Der Name der Tabelle, in der Sie suchen.

kennung

Der Name der Kennungsspalte, die der Spalte entspricht, die die zu durchsuchenden Dokumente enthält.

suchargument

Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe „Kapitel 12. Syntax von Suchargumenten“ auf Seite 217.

Beispiele

Ein Beispiel finden Sie in „Die Suchleistung verbessern“ auf Seite 99. Außerdem können Sie sich die Beispielfunktionen des DB2 Text Extender ansehen, die in der Datei geliefert werden, die in „Lokation von Syntaxbeispielen von Suchfunktionen“ auf Seite 80 beschrieben ist.

SEARCH_RESULT (Funktion)

Kapitel 12. 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 87 und in der Datei `txsample.udf` aufgeführt. Darin sind Beispiele von DB2 Text Extender-Funktionen enthalten, die mit der Beispieltabelle ausgeführt werden, die in „Eine Beispieldatenbank zur Installationsprüfung vorbereiten“ auf Seite 17 beschrieben ist.

Zu den DB2 Text Extender-Funktionen, 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. h. die Rangordnung, die Anzahl an Übereinstimmungen und die Kennung, enthält.

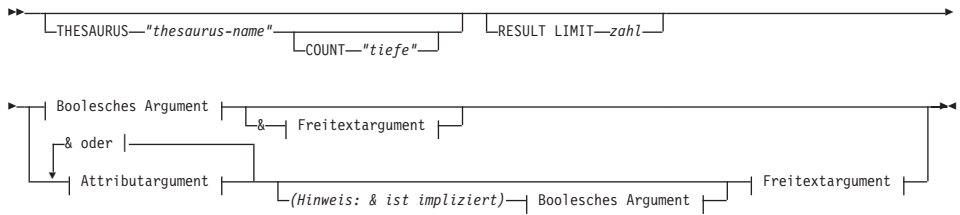
Syntax von Suchargumenten

Zu den API-Funktionen, die Suchargumente verwenden, gehören:

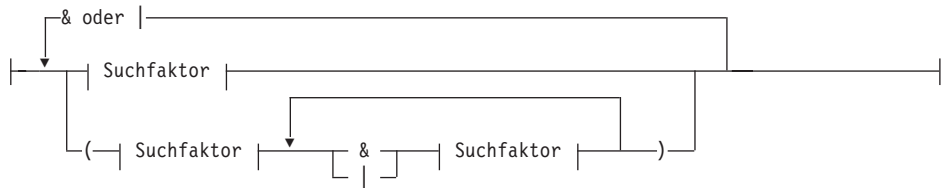
- `DesGetBrowseInfo`. Diese Funktion verwendet ein Suchargument für die Suche in Text, der durch eine Kennung identifiziert ist. Sie gibt einen Zeiger auf Anzeigeinformationen zurück, die von `DesStartBrowseSession` zur Hervorhebung der Begriffe benötigt werden.
- `DesGetSearchResultTable`. Diese Funktion 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

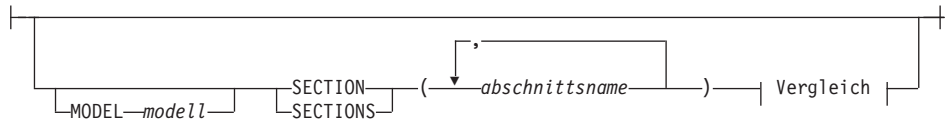
Suchargumentensyntax



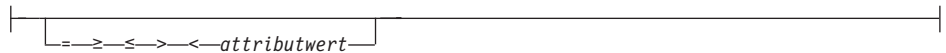
Boolesches Argument:



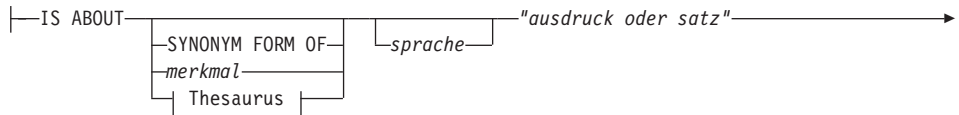
Attributargument:



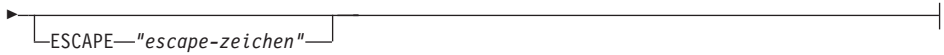
Vergleich:



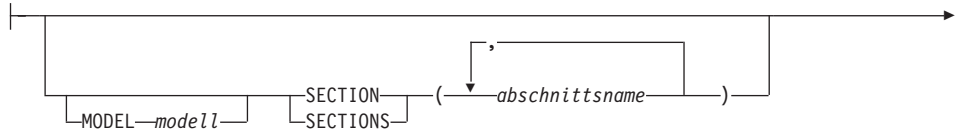
Freitextargument:



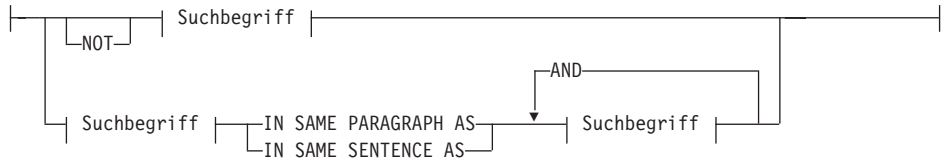
Syntax von Suchargumenten



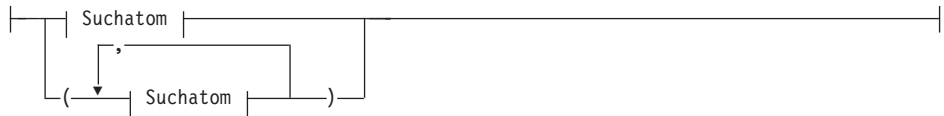
Suchfaktor:



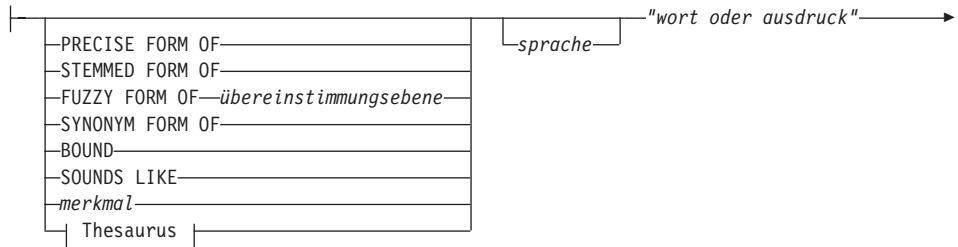
Suchelement:

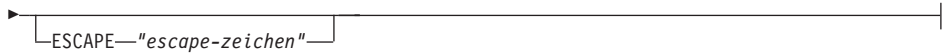


Suchbegriff:



Suchatom:





Thesaurus (wenn THESAURUS angegeben ist):



Beispiele

Beispiele werden in „Suchargumente angeben“ auf Seite 87 geliefert.

Suchparameter

IS ABOUT

Eine Option, mit der Sie ein Freitextsuchargument angeben können, d. h. ein Ausdruck oder ein Satz in natürlicher Sprache, der das zu suchende Konzept beschreibt. Siehe „Freitext- und Hybridsuche“ auf Seite 95.

MODEL *modell*

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 „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50 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 angegeben werden, die in „Mit strukturierten Dokumenten arbeiten (Abschnittsunterstützung)“ auf Seite 50 beschrieben ist. Ein Abschnittsname kann mit den Platzhalterzeichen % und _ angegeben werden.

Abschnitte können innerhalb von anderen Abschnitten verschachtelt sein. Beispiel:

```
play/Act/Title=play/act/title
```

Syntax von Suchargumenten

Einschränkungen: Die Suche in verschachtelten Abschnitten ist nur in Dokumenten möglich, die in Spalten gespeichert sind, die im XML-Format aktiviert sind. Bei Ngram-Indizes kann nur der Name eines Abschnitts gesucht werden und das XML-Format wird nicht unterstützt.

<attributwert

Ein Wert, der zusammen mit einem vorangehenden Vergleichsoperator für die Attribute verwendet wird, die in der vorangehenden Abschnittsliste aufgeführt sind. Eine Abfrage, die anfordert, daß der Wert eines Attributs in einem bestimmten Bereich ist, kann zwei Vergleichsoperatoren innerhalb einer einzigen Attributbedingung verwenden.

Eine Kombination von Operatoren, die dieselbe Vergleichsart (wie \geq in der ersten und $>$ in der zweiten) für dieselbe Bedingung verwenden, ist ungültig. Das Angeben von zwei Vergleichen mit einem Operator = ist ebenfalls ungültig.

- = Erfordert einen Gleichheitsvergleich des Attributs im indexierten Dokument mit dem folgenden Attributwert.
- \geq Erfordert einen "größer-gleich"-Vergleich des Attributs im indexierten Dokument mit dem folgenden Attributwert.
- $>$ Erfordert einen "größer-als"-Vergleich des Attributs im indexierten Dokument mit dem folgenden Attributwert.
- \leq Erfordert einen "kleiner-gleich"-Vergleich des Attributs im indexierten Dokument mit dem folgenden Attributwert.
- $<$ Erfordert einen "kleiner-als"-Vergleich des Attributs im indexierten Dokument mit dem folgenden Attributwert.

Abschnitte können innerhalb von anderen Abschnitten verschachtelt sein. Beispiel:

```
play/Act/Title=play/act/title
```

Einschränkungen: Die Suche in verschachtelten Abschnitten ist nur in Dokumenten möglich, die in Spalten gespeichert sind, die im XML-Format aktiviert sind. Bei Ngram-Indizes kann nur der Name eines Abschnitts gesucht werden und das XML-Format wird nicht unterstützt.

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. 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 271.

Verwenden Sie bei einem Ngram-Thesaurus den in „Einen Ngram-Thesaurus erstellen“ auf Seite 281 beschriebenen Member-Relationsnamen. Verwenden Sie bei benutzerdefinierten Member-Relationen `:RELATION n`, wobei *n* für die Member-Relationsnummer steht, die in `:RELATED (nummer)` angegeben wurde.

TERM OF *"wort oder ausdrück"*

Der Suchbegriff oder Mehrwortsuchbegriff, zu dem andere Suchbegriffe aus dem Thesaurus hinzugefügt werden sollen.

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"
```

Syntax von Suchargumenten

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"
```

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 8 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 9 auf Seite 226 näher beschrieben.

Tabelle 8. Linguistische Optionen

Schlüsselwort für Suchatom	Indextyp				
	Linguistisch	Präzise	Präzise, Normalisiert	Ngram	Ngram, Groß-/Kleinschreibung möglich
PRECISE FORM OF		X	X		O
STEMMED FORM OF	X			O	O
FUZZY FORM OF				O	O
IS ABOUT	O	O	O		
SYNONYM FORM OF	O	O	O		
EXPAND	O	O	O		
SOUNDS LIKE	O	O	O		
IN SAME SENTENCE AS	O	O	O	O	O
IN SAME PARAGRAPH AS	O	O	O	O	O
BOUND				O	O

X = Standardeinstellung, O = Funktion verfügbar

Syntax von Suchargumenten

Tabelle 9. Suchbegriffsoptionen für Ngram-Indizes

Schlüsselwort für Suchatom	Verarbeitung von Suchbegriffen				
	Groß-/Kleinschreibung		Wortstamm- bildung	Übereinstimmung	
	wird beachtet	wird nicht beachtet		Genau	Grob
PRECISE FORM OF	wenn Groß-/Kleinschreibung möglich	X		X	
STEMMED FORM OF		X	X		
FUZZY FORM OF		X			X

X = Standardeinstellung

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 Indizes wird bei dieser Form der Suche die Groß-/Kleinschreibung beachtet, d. 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 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. 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 econoy 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.

Syntax von Suchargumenten

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.

sprache Eine Variable, die festlegt, welches Wörterverzeichnis bei der Verarbeitung auf linguistischer Basis von Textdokumenten während der Indizierung und Abfrage verwendet wird. Dies gilt nicht nur für linguistische Indizes, sondern auch für präzise Indizes, da sie ein Wörterverzeichnis zur Verarbeitung von Stopwörtern verwenden.

Zur Verarbeitung auf linguistischer Basis gehört die Synonymverarbeitung und die Wortstammverarbeitung.

Die unterstützten Sprachen werden in Tabelle 5 auf Seite 57 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 DB2 Text Extender kann entweder die präzise Form eines Wortes oder Ausdrucks suchen oder eine Variation davon. Wenn Sie nicht eine der Optionen in Tabelle 8 auf Seite 225 angeben, werden die linguistischen Standardoptionen verwendet, gemäß dem verwendeten Indextyp.

Geben Sie die doppelten Anführungszeichen zweimal ein, wenn Sie nach einer Zeichenfolge suchen, die doppelte Anführungszeichen enthält. Beispielsweise können Sie wie folgt nach dem Text der Begriff "Platzhalter" suchen:

```
"der Begriff ""Platzhalter"""
```

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 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.

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.

Syntax von Suchargumenten

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 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.

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

Bei einem präzisen Index ignoriert, bei einem normalisierten präzisen Index, der englische Dokumente enthält, jedoch verfügbar.

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 13. API-Funktionen zum Suchen und Anzeigen

Der DB2 Text Extender bietet C-Funktionen für das Durchsuchen von Textdokumenten und das Anzeigen der gefundenen Dokumente. Diese Funktionen bilden die Anwendungsprogrammierschnittstelle (API) des DB2 Text Extender. In diesem Kapitel werden die API-Funktionen in alphabetischer Reihenfolge beschrieben.

„Kapitel 8. Die API-Funktionen zum Suchen und Anzeigen verwenden“ auf Seite 119 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.	233
DesEndBrowseSession	Schließt eine Anzeigesitzung und gibt den Speicher frei, der von DesStartBrowseSession zugeordnet wurde.	234
DesFreeBrowseInfo	Gibt den Speicher frei, der von DesGetBrowseInfo zugeordnet wurde.	235
DesGetBrowseInfo	Sucht unter Verwendung eines Sucharguments im Dokument nach Text und erstellt Anzeigeinformationen.	236
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.	239
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.	245
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.	250

API-Funktionen zum Suchen und Anzeigen

Funktion	Zweck	Seite
DesStartBrowseSession	Öffnet eine Anzeigesitzung unter Verwendung der Anzeigeeinformationen von DesGetBrowseInfo und gibt eine Anzeigesitzungskennung für die Verwendung durch andere Anzeigefunktionen zurück.	253

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 DB2 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 10. Argumente für DesCloseDocument

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	<i>BrowseSession</i>	Eingabe	Anzeigesitzungskennung.
DESHANDLE	<i>DocumentHandle</i>	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 (API-Funktion)

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 11. Argumente für DesEndBrowseSession

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	<i>BrowseSession</i>	Eingabe	Anzeigesitzungskennung.

Verwendung

Diese Funktion gibt den Speicher, der durch DesGetBrowseInfo für die Anzeigesitzung zugeordnet wurde, nicht frei. Dieser Speicher enthält Anzeigeeinformationen, 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.

DesFreeBrowseInfo

Zweck

Gibt den Speicher frei, der durch DesGetBrowseInfo für die Anzeigeeinformationen zugeordnet wurde.

Syntax

```
DESRETURN
DesFreeBrowseInfo
( DESBROWSEINFO      BrowseInfo );
```

Funktionsargumente

Tabelle 12. Argumente für DesFreeBrowseInfo

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSEINFO	<i>BrowseInfo</i>	Eingabe	Anzeigeeinformationen.

Rückkehrcodes

RC_SUCCESS

RC_INVALID_PARAMETER

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesGetBrowseInfo Speicher für die Anzeigeeinformationen zugeordnet haben.

DesGetBrowseInfo (API-Funktion)

DesGetBrowseInfo

Zweck

Empfängt ein Suchargument für das Durchsuchen von Text, der durch eine Kennung identifiziert ist. Die Funktion gibt einen Zeiger auf Anzeigeeinformationen zurück, die von DesStartBrowseSession zum Hervorheben der gefundenen Begriffe benötigt wird.

Syntax

```
DESRETURN  
DesGetBrowseInfo  
    (SQLHDBC          hdbc,  
     SQLCHAR         *pHandle,  
     DESUSHORT       HandleLength,  
     char            *pSearchArgument,  
     DESSMALLINT     ArgumentLength,  
     DESBROWSEINFO  *pBrowseInfo,  
     DESMESSAGE      *pErrorMessage);
```

Funktionsargumente

Tabelle 13. Argumente für DesGetBrowseInfo

Datentyp	Argument	Verwendung	Beschreibung
SQLHDBC	<i>hdbc</i>	Eingabe	Eine Datenbankverbindungskennung.
SQLCHAR *	<i>pHandle</i>	Eingabe	Zeiger auf eine Kennung, die aus der Datenbank extrahiert wurde.
DESUSHORT	<i>HandleLength</i>	Eingabe	Länge von pHandle. DES_NTS kann hier nicht verwendet werden.
char *	<i>pSearchArgument</i>	Eingabe	Zeiger auf ein Textsuchargument, das die Informationen angibt, die Sie suchen wollen.
DESSMALLINT	<i>ArgumentLength</i>	Eingabe	Entweder die Länge von pSearchArgument (ohne Nullbyteabschlußzeichen) oder DES_NTS.
DESBROWSEINFO *	<i>pBrowseInfo</i>	Ausgabe	Zeiger auf Anzeigeeinformationen, die die Daten enthalten, die zum Anzeigen eines Textdokuments erforderlich sind. Dieser Zeiger wird an DesStartBrowseSession übergeben.

Tabelle 13. Argumente für DesGetBrowseInfo (Forts.)

Datentyp	Argument	Verwendung	Beschreibung
DESMESSAGE *	<i>pErrorMessage</i>	Ausgabe	<p>Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der DB2 Text Extender einen Fehlercode und eine Fehlermeldung zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu.</p> <p>Wenn <i>pErrorMessage</i> der Nullzeiger ist, wird keine Fehlermeldung zurückgegeben.</p>

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 DB2 Text Extender ordnet Speicher für Anzeigeeinformationen 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 12. Syntax von Suchargumenten“ auf Seite 217 beschrieben.

Rückkehrcodes

```
RC_SUCCESS
RC_NO_BROWSE_INFO

RC_ALLOCATION_ERROR
RC_FILE_IO_PROBLEM
RC_INTERNAL_ERROR
RC_INVALID_PARAMETER
```

DesGetBrowseInfo (API-Funktion)

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 DB2 Text Extender-Funktion 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 240. 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.

Syntax

```

DESRETURN
DesGetMatches
    (DESBROWSESESSION BrowseSession,
    DESHANDLE DocumentHandle,
    DESMATCHINFO *pMatchInfo,
    DESULONG *pMatchInfoLength,
    DESMESSAGE *pErrorMessage);

```

Funktionsargumente

Tabelle 14. Argumente für DesGetMatches

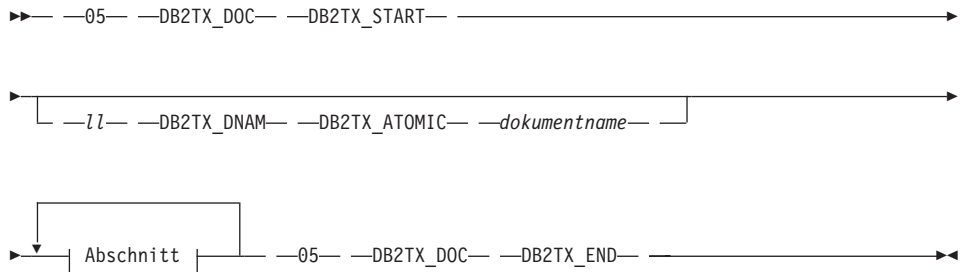
Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	<i>BrowseSession</i>	Eingabe	Anzeigesitzungskennung.
DESHANDLE	<i>DocumentHandle</i>	Eingabe	Dokumentkennung, die von DesOpenDocument zurückgegeben wird.
DESMATCHINFO *	<i>pMatchInfo</i>	Ausgabe	Zeiger auf einen Puffer, der die empfangenen Datenstromabschnitte enthält. DesGetMatches ordnet diesen Puffer zu.
DESULONG *	<i>pMatchInfoLength</i>	Ausgabe	Die Länge des Datenstromabschnitts, auf den durch pMatchInfo gezeigt wird.

DesGetMatches (API-Funktion)

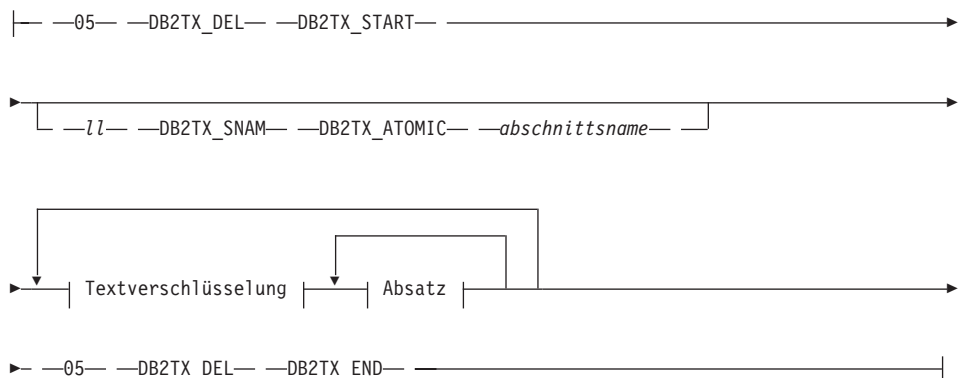
Tabelle 14. Argumente für DesGetMatches (Forts.)

Datentyp	Argument	Verwendung	Beschreibung
DESMESSAGE *	<i>pErrorMessage</i>	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der DB2 Text Extender einen Fehlercode und eine Fehlermeldung zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn <i>pErrorMessage</i> der Nullzeiger ist, wird keine Fehlermeldung zurückgegeben.

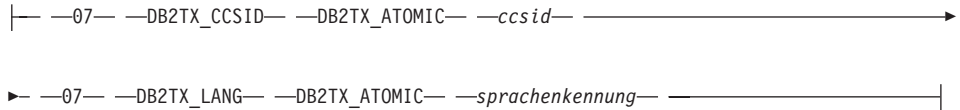
Datenstromsyntax



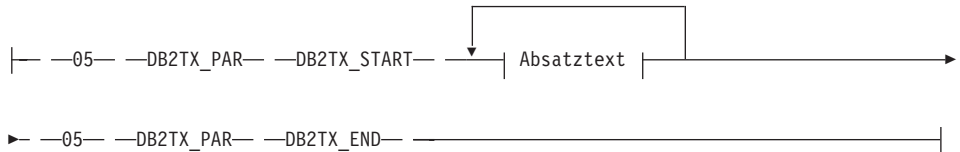
Abschnitt:



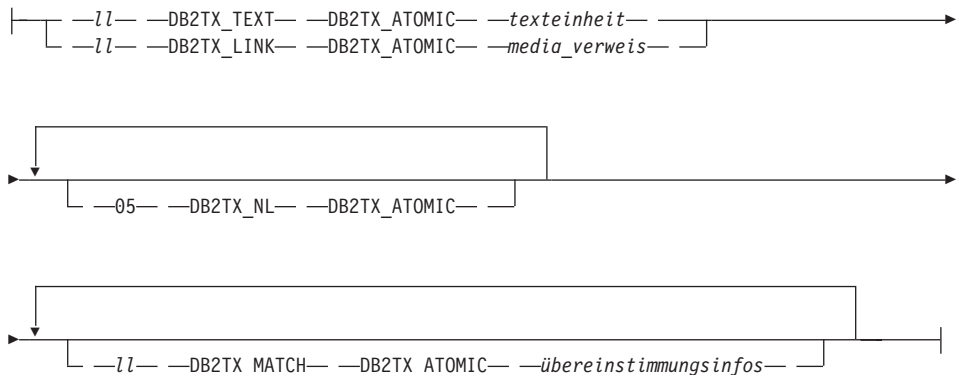
Textverschlüsselung:



Absatz:



Absatztext:



Jedes Segment im Syntaxdiagramm, wie z. B. 05 DB2TX_DOC DB2TX_START beginnt mit einem Längenfeld vom Typ Integer, das in diesem Diagramm entweder eine explizite Zahl, z. 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.

DesGetMatches (API-Funktion)

Zu jedem Segment gehört eine der folgenden, aus einem Byte bestehende Typenkennungen:

DB2TX_START

Gibt den Start eines Segments an, z. B. Dokument oder Absatz.

DB2TX_END

Gibt das Ende eines Segments an.

DB2TX_ATOMIC

Gibt an, daß das nachfolgende Element autark ist, z. 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 DB2 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 DB2 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 DB2 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 DB2 Text Extender geliefert wird, definiert symbolische Namen für alle Sprachenkennungen, die vom DB2 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.

DesGetMatches (API-Funktion)

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

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. Anzeigeeinformationen 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 15. 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.
char *	pTableName	Eingabe	Zeiger auf den Namen der zu durchsuchenden Basistabelle.
DESSMALLINT	TableNameLength	Eingabe	Entweder die Länge von pTableName (ohne Nullbyteabschlußzeichen) oder DES_NTS.

DesGetSearchResultTable (API-Funktion)

Tabelle 15. Argumente für DesGetSearchResultTable (Forts.)

Datentyp	Argument	Verwendung	Beschreibung
char *	pColumnName	Eingabe	Zeiger auf den Namen der Spalte, die in der beabsichtigten Textsuche adressiert werden soll. Die Spalte muß vom Typ DESTEXTH sein.
DESMALLINT	ColumnNameLength	Eingabe	Entweder die Länge von pColumnName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pSearchArgument	Eingabe	Zeiger auf das Textsuchargument.
DESMALLINT	ArgumentLength	Eingabe	Entweder die Länge von pSearchArgument (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pResultSchema	Eingabe	Zeiger auf das Schema, das die Ergebnistabelle enthält.
DESMALLINT	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. 17 auf Seite 247.
DESMALLINT	ResultTableNameLength	Eingabe	Entweder die Länge von pResultTableName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
DESSEARCHOPTION	SearchOption	Eingabe	<p>Eine Option, die festlegt, ob die Rangordnungsinformationen, die Anzahl an Übereinstimmungen oder nur die Kennungen der übereinstimmenden Textdokumente abgerufen werden.</p> <p>DES_RANK DES_MATCH DES_RANKANDMATCH DES_TEXTHANDLEONLY</p> <p>Diese Option legt den Inhalt der Ergebnistabelle fest, wie in „Verwendung“ auf Seite 247 beschrieben.</p>
DESBROWSEOPTION	BrowseOption	Eingabe	Reserviert.

Tabelle 15. Argumente für DesGetSearchResultTable (Forts.)

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 DB2 Text Extender einen Fehlercode und eine Fehlermeldung zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn pErrorMessage der Nullzeiger ist, wird keine Fehlermeldung 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 17. 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 12. Syntax von Suchargumenten“ auf Seite 217 beschrieben.

Wenn der Wert von *BrowseOption BROWSE* ist, gibt der DB2 Text Extender Anzeigeinformationen aus der DB2 Text Extender-Suchmaschine zurück, die sich auf dem Server befindet. *pBrowseInfo* zeigt auf die Anzeigeinformationen,

DesGetSearchResultTable (API-Funktion)

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

DesGetSearchResultTable (API-Funktion)

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 16. Argumente für DesOpenDocument

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	<i>BrowseSession</i>	Eingabe	Anzeigesitzungskennung.
SQLCHAR *	<i>pHandle</i>	Eingabe	Zeiger auf eine Kennung, die aus der Datenbank extrahiert wurde.
DESUSHORT	<i>HandleLength</i>	Eingabe	Länge von <i>pHandle</i> (DES_NTS kann nicht verwendet werden).
DESMATCHMODE	<i>MatchMode</i>	Eingabe	Modus zum Ermitteln, ob ein Wörterverzeichnis zum Suchen von Hervorhebungsinformationen verwendet wird. DES_FAST Kein Wörterverzeichnis verwenden DES_EXTENDED Ein Wörterverzeichnis verwenden
DESHANDLE *	<i>pDocumentHandle</i>	Ausgabe	Eine Dokumentkennung für das iterative Aufrufen von DesGetMatches.

Tabelle 16. Argumente für DesOpenDocument (Forts.)

Datentyp	Argument	Verwendung	Beschreibung
DESMESSAGE *	<i>pErrorMessage</i>	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der DB2 Text Extender einen Fehlercode und eine Fehlermeldung zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn <i>pErrorMessage</i> der Nullzeiger ist, wird keine Fehlermeldung 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 269. 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 DB2 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.

DesOpenDocument (API-Funktion)

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
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 Anzeigeeinformationen, entweder von DesGetBrowseInfo oder von DesGetSearchResultTable, und gibt eine Anzeigesitzungskennung für die Verwendung durch andere Anzeigefunktionen zurück.

Syntax

```

DESRETURN
DesStartBrowseSession
(
    DESBROWSEINFO
    char
    DESSMALLINT
    char
    DESSMALLINT
    DESBROWSESESSION
    DESMESSAGE
    BrowseInfo,
    *pUserId,
    UserIdLength,
    *pPassword,
    PasswordLength,
    *pBrowseSession,
    *pErrorMessage);
  
```

Funktionsargumente

Tabelle 17. Argumente für DesStartBrowseSession

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSEINFO	<i>BrowseInfo</i>	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 *	<i>pUserId</i>	Eingabe	Benutzer-ID für die Datenbank.
DESSMALLINT	<i>UserIdLength</i>	Eingabe	Länge der Benutzer-ID für die Datenbank.
char *	<i>pPassword</i>	Eingabe	Kennwort für die Datenbank.
DESSMALLINT	<i>PasswordLength</i>	Eingabe	Länge des Kennworts für die Datenbank.
DESBROWSESESSION *	<i>pBrowseSession</i>	Ausgabe	Eine Kennung für eine Anzeigesitzung zur Verwendung durch andere Anzeigefunktionen.

DesStartBrowseSession (API-Funktion)

Tabelle 17. Argumente für DesStartBrowseSession (Forts.)

Datentyp	Argument	Verwendung	Beschreibung
DESMESSAGE *	<i>pErrorMessage</i>	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der DB2 Text Extender einen Fehlercode und eine Fehlermeldung zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn <i>pErrorMessage</i> der Nullzeiger ist, wird keine Fehlermeldung 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 Anzeigeeinformationen 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.

Kapitel 14. Beispiel-API-Programm

Der DB2 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. 16 auf Seite 121 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, Windows NT und Windows 2000; 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. B. Suchergebnisse, zu speichern; sie hat die folgende Struktur:

Spalte	Datentyp
HANDLE	DB2TX.DB2TEXTH oder DB2TX.DB2TEXTFH
RANK	DOUBLE
MATCHES	INTEGER

Kapitel 15. Die Verarbeitung von linguistischen und präzisen Indizes auf linguistischer Basis

Der DB2 Text Extender bietet die Verarbeitung auf linguistischer Basis in folgenden Abfragebereichen:

- **Indexierung.** Wenn der DB2 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 DB2 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 DB2 Text Extender Dokumente indexiert und abrufen, 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
- Stopwortfiltern
- Dekomposition (Trennen von Komposita)

Tabelle 18 auf Seite 258 zeigt eine Zusammenfassung, wie Begriffe indexiert werden, wenn der Indextyp **linguistisch** ist und keine zusätzlichen Indexmerkmale angefordert wurden.

Verarbeitung auf linguistischer Basis beim Indexieren

Tabelle 18. 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 19 eine Zusammenfassung, wie Begriffe indexiert werden, wenn der Indextyp **präzise** ist.

Tabelle 19. 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

Basistextanalyse

Der DB2 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. B. "\$14,225.23", "Extender-Server" und "10/22/90".

Verarbeitung auf linguistischer Basis beim Indexieren

Die folgenden Zeichen werden als Teil eines Begriffs angesehen:

Akzente

Währungssymbole

Nummertrennzeichen (wie z. 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. B. l'aventure in Französisch.
- Nationale Formate für Datum- und Zeitangaben und für Zahlen.
- Alternativen, z. B. Verantwortlichkeit/Zuständigkeit, die z. B. im Englischen mit dem Zeichen “/” angegeben werden.
- Abschließende Apostrophe in italienischen Wörtern, wie z. 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 beim linguistischen Indextyp standardmäßig durchgeführt. (In einem präzisen Index bleibt 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.

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.

Verarbeitung auf linguistischer Basis beim Indexieren

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

Wenn der Suchbegriff Platzhalterzeichen 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 endet.

Der DB2 Text Extender bietet zwei Arten der Satzenderkennung:

- Universal Unicode Tokenizer für Sprachen außer Arabisch und Hebräisch.
Dies stellt die einfachere und zugleich schnellere Methode dar. Der Tokenizer sucht einen Punkt, ein Ausrufe- oder ein Fragezeichen, dem ein Tokenzeichen wie beispielsweise ein Buchstabe vorangeht und ein Leerzeichen, ein Tabulatorzeichen oder ein Zeichen für den Beginn einer neuen Zeile folgt. Um sicherzustellen, daß dies tatsächlich das Ende eines Satzes und nicht nur eine Abkürzung ist, die mit einem Punkt endet, wird ein sprachspezifisches Abkürzungsverzeichnis geprüft.
- Tokenizer auf POE-Basis für Arabisch und Hebräisch
Dieser Tokenizer ist in linguistischer Hinsicht fortschrittlicher, macht aber auch eine höhere Verarbeitungsleistung erforderlich. Er findet die Satzenden in erster Linie durch Interpunktionsabgleich, aber auch durch Hinweise über bestimmte Eingabearten und die Zahl der Wörter.

Tokenizer auf POE-Basis zur Satzerkennung

Der Tokenizer auf POE-Basis bestimmt Grenzen von Sätzen (oder von Satzfragmenten) mit Hilfe von Interpunktionsregeln und sprachspezifischer Verarbeitung, die auch die Verarbeitung von Abkürzungen umfaßt. Die Funktionsstufe unterscheidet sich aber von Sprache zu Sprache in starkem Maß. Die meisten Sprachen, die Einzelbyte-Codepages verwenden, verfügen über ein zugehöriges Abkürzungsaddendum (Abbreviation Addenda Dictionary), das mit POE geliefert wird. Da bei Doppelbytesprachen normalerweise keine Abkürzungen mit Punkten vorkommen, sind für solche Sprachen keine Abkürzungsaddenda verfügbar.

Die Ermittlung eines Satzendes erfolgt in erster Linie über einen Interpunktionsabgleich. In der folgenden Tabelle werden die beendenden Interpunktionszeichen und die zugehörigen globalen Schriftzeichenkennungen (Graphic Character Global Identifieries, GCGIDs) aufgelistet.

Verarbeitung auf linguistischer Basis beim Indexieren

GCGID von SBCS-Zeichen	GCGID von DBCS-Zeichen	Beschreibung
SP110000	SP110080	Punkt
SP020000	SP020080	Ausrufezeichen
SP150000	SP150080	Fragezeichen
SP140000	Nicht zutreffend	Semikolon (griechisches Fragezeichen)
Nicht zutreffend	JQ730080	Doppelbytekreispunkt

Es wird davon ausgegangen, daß ein beendendes Interpunktionszeichen, wie beispielsweise ein Punkt, ein Ausrufezeichen oder ein Fragezeichen, das Ende eines Satzes markiert, wenn nicht einer der folgenden Fälle vorliegt:

- Dem beendenden Interpunktionszeichen folgt ein abschließendes Interpunktionszeichen, das in der folgenden Tabelle aufgeführt ist, wie beispielsweise eine abschließende runde Klammer.

GCGID von SBCS-Zeichen	GCGID von DBCS-Zeichen	Beschreibung
SP070000	SP070080	Abschließende runde Klammer
SP040000	SP040080	Doppeltes Anführungszeichen
SP050000	SP050080	Einfaches Anführungszeichen
SP180000	SP070083	Doppeltes Winkelanführungszeichen
Nicht zutreffend	SM140080	Abschließende geschweifte Klammer
Nicht zutreffend	SM080080	Abschließende eckige Klammer
Nicht zutreffend	JQ720080	Einfaches eckiges Anführungszeichen
Nicht zutreffend	JQ720081	Doppeltes eckiges Anführungszeichen
Nicht zutreffend	SP200080	Abschließendes einfaches Hakenanführungszeichen
Nicht zutreffend	SP220080	Abschließendes doppeltes Hakenanführungszeichen
Nicht zutreffend	SP070081	Abschließende Schaleneckklammer

Verarbeitung auf linguistischer Basis beim Indexieren

GCGID von SBCS-Zeichen	GCGID von DBCS-Zeichen	Beschreibung
Nicht zutreffend	SP070082	Abschließendes einfaches Winkelanführungszeichen
Nicht zutreffend	SP070084	Abschließende Eckklammer
Nicht zutreffend	SP370080	Vertikales abschließendes einfaches eckiges Anführungszeichen
Nicht zutreffend	SP370081	Vertikales abschließendes doppeltes eckiges Anführungszeichen
Nicht zutreffend	SP250084	Vertikale abschließende eckige Klammer
Nicht zutreffend	SP250080	Vertikale abschließende runde Klammer
Nicht zutreffend	SP350080	Vertikale abschließende geschweifte Klammer
Nicht zutreffend	SP250081	Vertikale abschließende Schaleneckklammer
Nicht zutreffend	SP250083	Vertikales abschließendes doppeltes Winkelanführungszeichen
Nicht zutreffend	SP250082	Vertikales abschließendes einfaches Winkelanführungszeichen

Anmerkung: Die als 'nicht zutreffend' markierten Elemente werden von POE nicht als abschließende Interpunktionszeichen betrachtet. Vertikale abschließende Interpunktionszeichen werden nur in Chinesisch unterstützt.

Beispiel:

...dieser Satz endet mit zwei runden Klammern.))

In diesem Beispiel wird die zweite runde Klammer als Ende des Satzes erkannt. Im Deutschen wird ein abschließendes Anführungszeichen nicht als Markierung für das Satzende betrachtet, wenn es von einem Komma gefolgt wird.

Verarbeitung auf linguistischer Basis beim Indexieren

- Dem beendenden Interpunktionszeichen folgt ein weiteres beendendes Zeichen. Beispiel:

Dies ist ein starker Ausruf!!!

Das letzte Ausrufezeichen wird als Ende des Satzes erkannt.

- Dem beendenden Interpunktionszeichen geht entweder ein numerisches Zeichen oder ein Interpunktionszeichen voraus und ihm folgt ein numerisches Zeichen. Dadurch wird verhindert, daß Zeichenfolgen wie '1.25' und '.314' das Ende eines Satzes bilden.
- Das beendende Interpunktionszeichen ist ein Punkt und gleichzeitig Teil einer Abkürzung, die am Ende eines Satzes nicht zulässig ist. Für alle Sprachen wird eine eingeschränkte Abkürzungsverarbeitung ausgeführt.
- Das beendende Interpunktionszeichen ist ein Punkt und ihm folgt kein Zeichen für einen Leerraum, wie z. B. ein Leerzeichen oder ein Zeichen für den Beginn einer neuen Zeile. Dadurch wird vermieden, daß Überschriften wie 'III.IV' als Satzende erkannt werden.

Der Tokenizer auf POE-Basis führt auch Abkürzungsverarbeitung durch, um zu ermitteln, ob ein Punkt Teil einer Abkürzung oder eine Markierung für ein Satzende ist. Zum Abkürzungsaddendum können auch Abkürzungen hinzugefügt werden. Wenn an den Tokenizer auf POE-Basis kein Wörterverzeichnis übermittelt wird, werden alle einzelnen Buchstaben, denen ein Punkt folgt, als Abkürzungen markiert; darüber hinaus findet keine weitere Abkürzungsverarbeitung statt.

Ob eine Textstelle eine Abkürzung ist, ist in vielen Fällen nicht eindeutig sicher, da eine Abkürzung mit einem normalen Wort verwechselt werden kann, dem ein Punkt folgt. Betrachten Sie beispielsweise die Zeichen "Gen." in den folgenden Sätzen:

Im Gen. wird an dieses Wort ein 's' angehängt.
Der Grund für seine Erbkrankheit lag in einem Gen.

Aber selbst wenn klar ist, daß es sich bei einem bestimmten Textteil um eine Abkürzung handelt, ist immer noch unklar, ob danach der Satz endet. Einige Abkürzungen bilden nie das Ende eines Satzes, während dies für andere in einigen Fällen zutreffen kann. Betrachten Sie beispielsweise die Verwendung der Abkürzung "usw." in den folgenden Sätzen:

Kräftige Farben wie beispielsweise blau, rot, grün usw.
werden von diesem Maler bevorzugt verwendet.
Diese Firma produziert Maschinenteile, Kugellager usw.

Da Abkürzungen mehrdeutig sein können und einige Abkürzungen nicht am Ende eines Satzes vorkommen können, versucht POE, die gefundenen Abkürzungen zu klassifizieren.

Verarbeitung auf linguistischer Basis beim Indexieren

Wenn festgestellt wird, daß ein Punkt Teil einer Abkürzung ist, die in einigen Fällen einen Satz beendet, wird eine weitere Verarbeitung durchgeführt. Wenn POE feststellt, daß die betreffende Abkürzung nicht am Ende eines Satzes steht, wird das Token, das den Punkt darstellt, mit dem Token des Abkürzungstexts verknüpft. Andernfalls bleibt das Token, das den Punkt darstellt, separat.

Bei der Abkürzungsverarbeitung auf POE-Basis werden drei Gruppen von Kriterien verwendet, um zu ermitteln, ob ein Punkt Teil einer Abkürzung ist:

- Alle einzelnen Buchstaben, denen ein Punkt folgt, werden als Abkürzungen betrachtet. Abkürzungen mit einem einzigen Buchstaben werden als mögliche Satzenden klassifiziert.
- Die im Eingabeabkürzungsaddendum enthaltenen Wörter gelten in allen Fällen als Abkürzungen. Ob eine im jeweiligen Addendum enthaltene Abkürzung am Ende eines Satzes stehen kann, wird durch die Informationen festgelegt, die diesem Wort in diesem Addendum zugeordnet sind. Beispielsweise ist "Mr." in der Version des Abkürzungsaddendums in amerikanischem Englisch als Abkürzung markiert, die nicht am Ende eines Satzes stehen kann. Dagegen ist "etc." eine Abkürzung, die in einigen Fällen am Ende eines Satzes stehen kann.
- Alle Wörter mit einer Länge von zwei bis sechs Zeichen, auf die ein Punkt folgt und die nicht in einem der Eingabewörterverzeichnisse oder Abkürzungsaddenda enthalten sind, gelten auch als Abkürzungen. Dadurch werden Fälle wie "Abhdlng. f. Comp.wiss." berücksichtigt. Abkürzungen, die durch Nachschlagen im Wörterverzeichnis bzw. Addendum ermittelt werden, werden in allen Fällen als mögliche Satzenden behandelt.

Wenn eine Abkürzung als mögliches Satzende ermittelt wurde, untersucht POE den Text, der der Abkürzung folgt, um zu ermitteln, ob die Abkürzung am Ende des aktuellen Satzes steht. Dies geschieht durch die Prüfung, ob das nächste Wort mit einem Großbuchstaben beginnt.

Wenn einer Abkürzung zwei oder mehr Datenelemente für den Beginn einer neuen Zeile, eines neuen Satzes oder eines neuen Absatzes folgen, geht POE davon aus, daß das Ende eines Satzes erreicht wurde. Wenn der nachfolgende Text ein umgekehrtes Fragezeichen oder ein umgekehrtes Ausrufezeichen ist, wird in die Ausgabe auch eine Markierung für das Satzende eingefügt.

Wenn POE feststellt, daß der Punkt Teil einer Abkürzung ist, die nicht das Ende eines Satzes bildet, wird die Suche nach einem Satzbegrenzer fortgesetzt. Andernfalls werden andere Ausnahmeregeln für beendende Interpunktionszeichen (nach beendener oder abschließender Interpunktion) geprüft, bevor das Satzende markiert wird.

Begriffe auf ihre Grundform reduzieren

In einem linguistischen 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.

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. Die Stoppwortlisten, die in verschiedenen Sprachen geliefert werden, können geändert werden.

Ein Ngram-Index verfügt über keine Stoppwortliste.

Dekomposition (Trennen von Komposita)

In germanischen Sprachen, wie Deutsch oder Niederländisch, werden viele Komposita verwendet, wie z. 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. Bei einem linguistischen Index werden Komposita während der Indexierung getrennt. Bei der Suche werden Komposita getrennt, wenn Sie einen linguistischen Index haben.

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.

Verarbeitung auf linguistischer Basis beim Indexieren

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. 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:

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 Abruftrate bei Suchen erhöht wird, d. 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 Booleschen OR-Ausdruck in der Abfragesprache des DB2 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

Verarbeitung auf linguistischer Basis für die Abfrage

- 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 DB2 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 265)
 - Normalisierung (siehe „Begriffe auf eine Standardform normalisieren“ auf Seite 259).
 - Stoppwörter (siehe „Stopwortfiltern“ auf Seite 265).
- 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. B. Substantiv, Verb, etc.) wie der Quellbegriff. 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 épître
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. B. die Relation "Narrower Term" (engerer Begriff)), oder assoziativ (z. 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 271 wird die Thesaurus-Erweiterung detaillierter beschrieben.

Der Suchbegriff (Startbegriff) wird bei der Thesaurus-Suche nicht normalisiert. Die aus der Thesaurus-Suche resultierenden Wörter werden entsprechend des Indextyps auf ihre Grundform reduziert.

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 258 beschrieben.

Begriffserweiterung

Die Begriffserweiterung ist die Umkehrung der Reduzierung eines Begriffs auf seine Grundform. Wenn der Index linguistisch ist, werden die Suchbegriffe auf ihre Grundform reduziert, bevor die Suche beginnt.

Ebenso werden, wenn Sie einen linguistischen 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	geht	geht
ging	ginge	gingen	gingest	ginget	gingst	gingt	geht

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 `schwim%` 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 `schwim%` 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.

Verarbeitung auf linguistischer Basis für die Anzeige

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.
- 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.

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.

Thesaurus-Konzepte

Weitere zugehörige Begriffe können erreicht werden, indem iterativ alle Verbindungen verfolgt werden, die von den Knoten ausgehen, die im vorherigen Schritt erreicht wurden. Abb. 18 zeigt ein Beispiel der Struktur eines sehr kleinen Thesaurus.

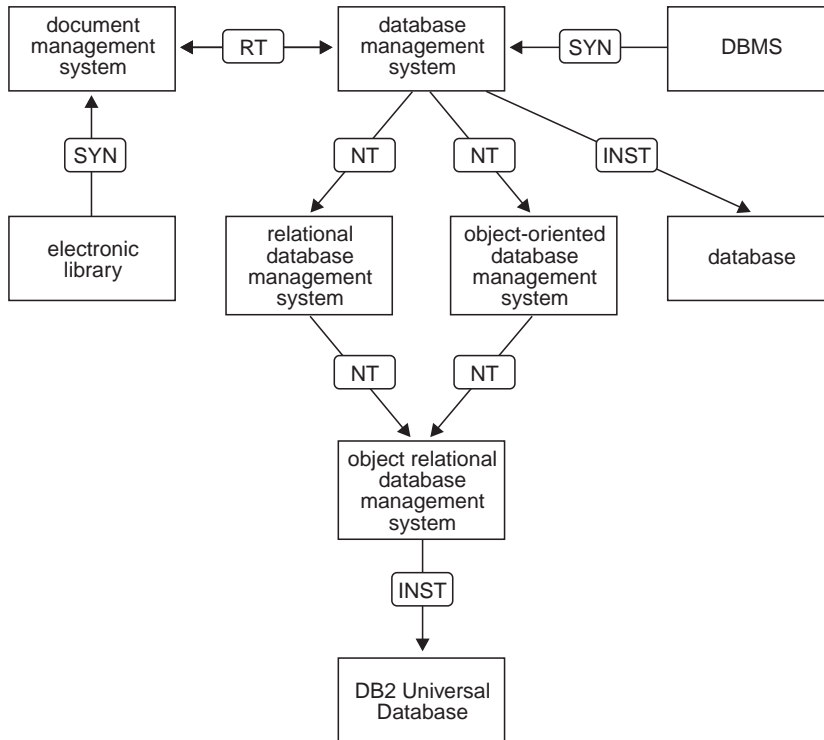


Abbildung 18. Ein als Netz angezeigter Thesaurus

Mit dem DB2 Text Extender können Sie einen Suchbegriff erweitern, indem Sie zusätzliche Begriffe aus einem Thesaurus, den Sie zuvor erstellt haben, hinzufügen. In „Kapitel 12. Syntax von Suchargumenten“ auf Seite 217 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 geliefert. Sie unterscheiden sich nicht nur in den zugrundeliegenden Konzepten, sondern es sind auch unterschiedliche

Quellenformate erforderlich. Sie sollten zunächst entscheiden, welchen Indextyp Sie verwenden wollen, bevor Sie beginnen, die Thesauri für Ihre Anwendungen zu definieren.

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

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 DB2 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 12. Syntax von Suchargumenten“ auf Seite 217 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 exponentiell wächst, wenn die Tiefe erhöht wird.

Thesaurus-Konzepte

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, rheumatology, radiotherapy, psychiatry, preventive medicine, pathology, odontology, nutrition, nuclear medicine, neurology, nephrology, medical check up, industrial medicine, hematology, general medicine, epidemiology, clinical trial, cardiology, cancerology

DB2 Text Extender-Thesaurus-Relationen

Dies sind die Relationsarten, die von einem DB2 Text Extender-Thesaurus geboten werden:

- Assoziativ
- Synonym
- Hierarchisch
- OTHER

In einem DB2 Text Extender-Thesaurus sind keine vordefinierten Relationen vorhanden. Sie können jeder Relation einen Namen geben, z. 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 DB2 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 276 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. Beispiele:

katzentier USE katze
rechtsanwalt UF verteidiger

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. Beispiele:

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.

Thesaurus-Konzepte

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 281.

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 281.

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 Wörterverzeichnis

```
DB2TX_INSTOWNER /db2tx/samples
```

Eine kompilierte Version dieses Thesaurus und seiner SGML-Eingabedatei wird im Verzeichnis 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 281 weitere ausgeschlossene Dateierweiterungen.

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

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

Hierbei kann *dateiname* nur die Zeichen a-z, A-Z und 0-9 enthalten.

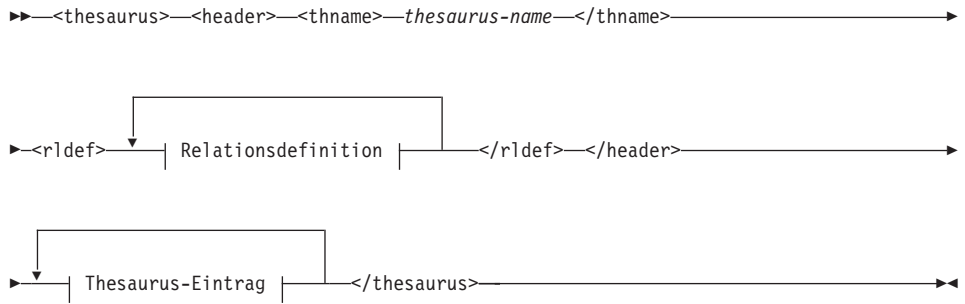
Momentan wird nur die CCSID 850 unterstützt.

`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.

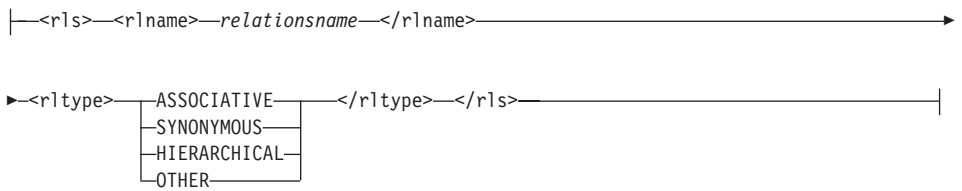
In „Kapitel 12. Syntax von Suchargumenten“ auf Seite 217 finden Sie Informationen, wie ein Thesaurus in einer Abfrage verwendet wird.

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.

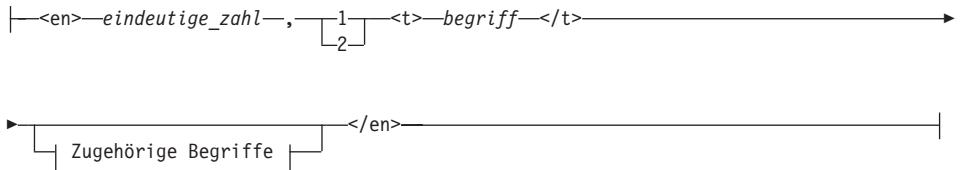
Thesaurus-Konzepte



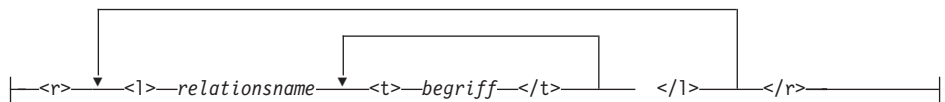
Relationsdefinition:



Thesaurus-Eintrag:



Zugehörige Begriffe:



relationsname kann nur die Zeichen a-z, A-Z und 0-9 enthalten.

Abb. 19 auf Seite 279 zeigt die SGML-Definition des Thesaurus, der in Abb. 18 auf Seite 272 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 19. Die Definition eines einfachen Thesaurus (Teile- 1 von 2)

Thesaurus-Konzepte

```
<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 Universal Database</t>
</en>

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

Abbildung 19. Die Definition eines einfachen Thesaurus (Teile- 2 von 2)

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`, `lkn`
- Für temporäre Dateien: `: wnf`, `wnv`, `gnf`, `gnv`, `M!1`, `M!2`, `N!1`, `N!2`, `R!1`, `R!2`, `S!1`, `S!2`, `Mnn`, `Nnn`, `Rnn`, `Snn`, `$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 DB2 Text Extender-Thesauri verwenden, verwenden Sie nicht die Erweiterungen, die unter „Einen Thesaurus erstellen“ auf Seite 276 aufgelistet sind.

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

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

Nachfolgend ist eine Liste mit den Codepages aufgeführt, die von einem Ngram-Thesaurus unterstützt werden:

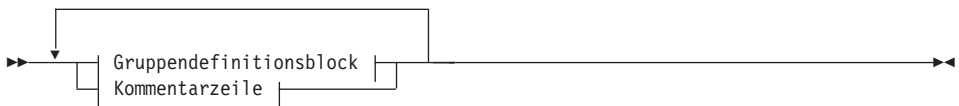
932	AIX, OS/2	Japanisch
942	OS/2	Japanisch
943	OS/2, Windows	Japanisch
949	OS/2	Koreanisch
950	AIX, HP-UX, OS/2, SUN-Solaris, Windows	Traditionelles Chinesisch

Thesaurus-Konzepte

970	AIX, HP-UX, SUN-Solaris	Koreanisch
1381	OS/2, Windows	Vereinfachtes Chinesisch
1363	Windows	Koreanisch
1383	AIX, HP-UX, SUN-Solaris	Vereinfachtes Chinesisch
850	AIX, OS/2	Lateinisches Alphabet 1
1252	Windows	Lateinisches Alphabet 1

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.

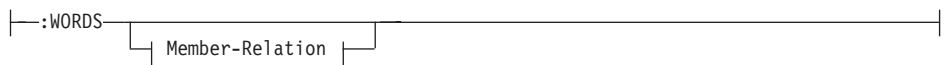
Geben Sie den Inhalt des Thesaurus unter Verwendung des folgenden Syntaxdiagramms an:



Gruppendefinitionsblock:



Startzeile für Block:



Member-Relation:



Member-Begriffsdefinition:

```
|—member-begriff—|
```

Definition für zugehörigen Begriff:

```
|
| .RELATED_TO—| zugehöriger_begriff—|
| .SYNONYM_OF—|
| .RELATED_TO—(—zahl—)|
```

Kommentarzeile:

```
|—#—beliebiger_kommentar—|
```

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 12. Syntax von Suchargumenten“ auf Seite 217 finden Sie Informationen, wie die Thesaurus-Erweiterung in einer Abfrage verwendet wird.

Kapitel 16. Konfigurationsdateien

In diesem Kapitel werden die Konfigurationsdateien beschrieben. Diese werden automatisch generiert, wenn ein DB2 Text Extender-Exemplar mit dem Befehl `txicrt` oder ein Client-Profil mit dem Befehl `descrc1` erstellt wird. Diese Dateien werden in Codepage 819 generiert.

Sie können diese Dateien zur Optimierung Ihres Systems bearbeiten, müssen jedoch sicherstellen, daß Sie dabei die richtige Codepage verwenden. Bei den Abschnitts- und Optionsnamen wird die Groß-/Kleinschreibung nicht beachtet. Als Kommentarbegrenzer wird ein Semikolon verwendet.

Bei Booleschen Werten gelten die Werte TRUE, YES, ON und 1 unabhängig davon als wahr, ob sie groß- oder kleingeschrieben werden; alle anderen Werte gelten als falsch.

Client-Konfigurationsdatei

Dateiname DESCL.INI
Standort UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx
Windows NT und Windows 2000:
%DMBMMPATH%*exemplar*\%DB2INSTANCE%\db2tx

Die aktualisierten Optionen werden bei der nächsten Funktion `StartSession` aktiv.

Abschnitt	Option	Standardwert	Beschreibung
[INSTANCE]	DESWORKCL		Zeigt auf ein Arbeitsverzeichnis, das für temporäre Dateien verwendet wird.
	DESNLPSCl		Zeigt auf das Ressourcenverzeichnis.
	DESDTDPATHSRV		Zeigt auf das Verzeichnis, in dem die DTDs für die XML-Dokumente gespeichert sind.
	DESDTDPATHCL		Zeigt auf den Pfad, in dem DTD-Dateien gespeichert werden müssen. Für XML-Dateien verwendet.

Client-Konfigurationsdatei

Abschnitt	Option	Standardwert	Beschreibung
[BUFFER]	BUFFERSEGMENTSIZ	32 000	Die Größe der Blocksegmente in Byte, die für die Pufferung verwendet werden.
	BUFFERSEGMENTCOUNT	3	Die Höchstzahl an Segmenten, die verwendet wird, bevor Puffer in temporäre Dateien ausgelagert werden. Ein Puffersegment ist durch den Parameter BUFFERSEGMENTSIZ definiert.
[DOCUMENTFORMAT]	USEREXIT		<p>Der Name des Benutzer-Exit, der zur Arbeit mit nicht unterstützten Dokumentformaten verwendet wird. Geben Sie entweder einen Dateinamen an, wenn der Benutzer-Exit in einem Verzeichnis gespeichert ist, das in der Anweisung PATH enthalten ist, oder einen vollständig qualifizierten Dateinamen.</p> <p>Weitere Informationen zum Benutzer-Exit zur Formatumwandlung finden Sie im Abschnitt „Nicht unterstützte Dokumentformate verwenden“ auf Seite 35.</p>
	FORMATRECOGNITION	TRUE	<p>Löst die Formaterkennung der Dokumentformate aus.</p> <p>TRUE: Formaterkennung ist eingeschaltet</p> <p>FALSE: Formaterkennung ist ausgeschaltet</p>

Abschnitt	Option	Standardwert	Beschreibung
	UseExitForAllFormats	FALSE	<p>Stellt fest, wann der Benutzer-Exit für die Arbeit mit Dokumentformaten aufgerufen wird, die nicht in deslsdef.h aufgelistet sind. Für die Option USEREXIT müssen Sie einen Wert angeben.</p> <p>TRUE: Der Benutzer-Exit wird immer aufgerufen. Wenn der Wert gesetzt ist, wird FORMATRECOGNITION ignoriert.</p> <p>FALSE: Der Benutzer-Exit wird für alle Dokumentformate über dem Wert von EHW_USER_FORMATS aufgerufen.</p>

Server-Konfigurationsdatei

Dateiname DESSRV.INI

Standort UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx/TXINSnnn

Windows NT und Windows 2000:

%DMBMPATH%*exemplar*\%DB2INSTANCE%\db2tx\txinsnnn

Diese Datei ist nur einmal pro Server-Exemplar vorhanden.

Die aktualisierten Optionen werden beim nächsten Start des Server-Exemplars aktiviert.

Tabelle 20. Optionen der Server-Konfigurationsdatei

Abschnitt	Option	Standardwert	Beschreibung
[INSTANCE]	DESWORKSRV		Zeigt auf ein Arbeitsverzeichnis, das für temporäre Dateien verwendet wird.
	DESNLPSSRV		Zeigt auf das Ressourcenverzeichnis.

Server-Konfigurationsdatei

Tabelle 20. Optionen der Server-Konfigurationsdatei (Forts.)

Abschnitt	Option	Standardwert	Beschreibung
[DAEMON]	MaxMtEntries	30 unter AIX, Windows, OS/390 10 unter Solaris	Die Höchstzahl an Indizes, die zu einem bestimmten Zeitpunkt parallel verwendet werden. Verringern Sie diese Zahl, wenn nur geringe Ressourcen wie Semaphors oder gemeinsam benutzte Speicher zur Verfügung stehen. Verfügbare Ressourcen sind plattformabhängig, weshalb die Standardwerte ebenfalls plattformabhängig sind.
	MaxIndexEntries	1000	Die Höchstzahl an verwendeten Indizes. Verringern Sie diese Zahl, wenn nur geringe Ressourcen wie gemeinsam benutzte Speicher zur Verfügung stehen.
[BUFFER]	BUFFERSEGMENTSIZE	32 000	Die Größe der Blocksegmente in Byte, die für die Pufferung verwendet werden. Wird von EhwUpdate verwendet.
	BUFFERSEGMENTCOUNT	3	Die Höchstzahl an Segmenten, die während des Prozesses der Indexaktualisierung verwendet wird, bevor Puffer in temporäre Dateien ausgelagert werden. Erhöhen Sie diese Zahl, wenn Ihre Dokumentensammlungen große Dokumente enthalten.
	BUFFERSORTSIZE	20 000 000	Die Größe des Puffers in Byte, der für das Sortieren temporärer Arbeitsdateien verwendet wird.

Tabelle 20. Optionen der Server-Konfigurationsdatei (Forts.)

Abschnitt	Option	Standardwert	Beschreibung
[DOCUMENTFORMAT]	USEREXIT		<p>Der Name des Benutzer-Exit, der zur Arbeit mit Dokumentformaten verwendet wird, die nicht in deslsdef.h aufgelistet sind. Geben Sie entweder einen Dateinamen an, wenn der Benutzer-Exit in einem Verzeichnis gespeichert ist, das in der Anweisung PATH enthalten ist, oder einen vollständig qualifizierten Dateinamen.</p> <p>Weitere Informationen zum Benutzer-Exit zur Formatumwandlung finden Sie im Abschnitt „Nicht unterstützte Dokumentformate verwenden“ auf Seite 35.</p>
	FORMATRECOGNITION	TRUE	<p>Löst die Formaterkennung der in deslsdef.h aufgelisteten Dokumentformate aus.</p> <p>TRUE: Formaterkennung ist eingeschaltet</p> <p>FALSE: Formaterkennung ist ausgeschaltet</p>

Server-Konfigurationsdatei

Tabelle 20. Optionen der Server-Konfigurationsdatei (Forts.)

Abschnitt	Option	Standardwert	Beschreibung
	UseExitForAllFormats	FALSE	<p>Stellt fest, wann der Benutzer-Exit für die Arbeit mit Dokumentformaten aufgerufen wird, die nicht in desldef.h aufgelistet sind. Für die Option USEREXIT müssen Sie einen Wert angeben.</p> <p>TRUE: Der Benutzer-Exit wird immer aufgerufen. Wenn der Wert gesetzt ist, wird FORMATRECOGNITION ignoriert.</p> <p>FALSE: Der Benutzer-Exit wird für alle Dokumentformate über dem Wert von EHW_USER_FORMATS aufgerufen.</p>
[LINGPREC] Für alle Indizes mit 'linguistisch' oder 'präzise' als Basistyp.	UPDATETHRESHOLD	4 000 000	<p>Ein Prozeß zur Indexaktualisierung wird intern in mehrere Aktualisierungs- und Reorganisationsläufe aufgeteilt. Dieser Wert gibt die Zahl der Wörter an, die in einem Aktualisierungsschritt gesammelt werden.</p>
	UPDATESLICE	1	<p>Die Zahl der Aktualisierungsläufe, die durchgeführt werden, bevor ein interner Reorganisationsprozeß startet. Ein Aktualisierungslauf wird durch die Option UPDATETHRESHOLD definiert.</p>

Tabelle 20. Optionen der Server-Konfigurationsdatei (Forts.)

Abschnitt	Option	Standardwert	Beschreibung
[NGRAM] Für alle Indizes mit DBCS-Unterstützung.	UPDATETHRESHOLD	10 000 000	Gesamtgröße (in Byte) von Dokumenten, die während eines Aktualisierungslaufs zu einem Index hinzugefügt werden. Wenn der Schwellenwert überschritten wird, wird automatisch ein Reorganisationsprozeß gestartet.
	UPDATESLICE	10 000	Die Höchstzahl an Dokumenten in einem Sekundärindex. Diese Anzahl wird nach jedem Aktualisierungslauf geprüft. Wenn die Anzahl der Dokumente größer als dieser Wert ist, wird automatisch ein Reorganisationsprozeß gestartet.

Konfigurationsdatei für die Textanalyse

Dateiname DESTAF.TCF

Standort UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx/TXINSnnn

Windows NT und Windows 2000:

%DMBMPATH%*exemplar*\%DB2INSTANCE%\db2tx\txinsnn

Diese Datei ist nur einmal pro Server-Exemplar vorhanden.

Die Konfigurationsdatei für die Textanalyse enthält Abschnitte zur Aktualisierungs- und Abfrageverarbeitung für alle Indextypen mit Ausnahme von Ngram-Indizes. Sie sollten nur die hier beschriebenen Optionen ändern. Wenn Sie einen der anderen Einträge ändern, führt die möglicherweise zu Fehlern während der Aktualisierungs- oder Abfrageverarbeitung.

Konfigurationsdatei für die Textanalyse

Allgemeine Optionen

Die folgenden Optionen gelten für:

- [TAF_Application:UPDATE:PREC)] für den Indextyp PRECISE
- [TAF_Application:UPDATE:LING)] für den Indextyp LINGUISTIC
- [TAF_Application:UPDATE:NORM)] für den Indextyp PRECISE NORMALIZED
- [TAF_Application:QUERY:PREC)]
- [TAF_Application:QUERY:LING)]
- [TAF_Application:QUERY:NORM)]

Option	Standardwert	Beschreibung
FilterSet	Keiner, erforderlich	Name des Abschnitts in der Konfigurationsdatei, der die Regeln für das Stoppwortfiltern enthält, z. B. StwUseAllUpCs, StwIgnAllUpCs.
Normalize	No	Gibt an, ob für den Index Normalisierung ausgeführt wird. Wenn diese Option auf 'Yes' (Ja) gesetzt ist, werden durch die Normalisierung alle Zeichen in Kleinbuchstaben umgewandelt und diakritische Zeichen durch Basiszeichen ersetzt. Wenn diese Option auf 'No' (Nein) gesetzt ist, werden die Einträge für die Optionen NormalizeUmlauts und ConvertAllUpperToLower ignoriert.
NormalizeUmlauts	Yes	Gibt an, ob Umlaute ersetzt werden sollen, z. B. ä durch ae. Dies gilt nur für Dokumente in Deutsch, Dänisch, Schwedisch oder Norwegisch.
ConvertAllUpperToLower	No	Gibt an, ob Wörter, die komplett in Großbuchstaben geschrieben sind, in Kleinbuchstaben umgewandelt werden.

[TAF_Filter:STWUSEALLUPCS] [TAF_Filter:STWIGNALLUPCS]

Die Optionen in diesen Abschnitten beschreiben die Stoppwortfiltergruppen. Der Abschnitt [TAF_Filter:STWUSEALLUPCS] behandelt alle Wörter, die komplett in Großbuchstaben geschrieben sind, als Wörter mit gemischter Groß- und Kleinschreibung. Im Abschnitt [TAF_Filter:STWIGNALLUPCS] wird für alle Wörter in Großbuchstaben eine Sonderbehandlung definiert.

Option	Standardwert	Beschreibung
FilterType	TAF_FILTER_STW	Abschnitt in der Konfigurationsdatei, der Regeln für das Stoppwortfiltern enthält.
MinTokenLength	3	Mindestzahl an Zeichen, die für ein Token erforderlich sind. Wenn die Länge des Token geringer als dieser Wert ist, wird es nicht indexiert. Dies ist ein ganzzahliger Wert ≥ 1 .
MinUpperCaseTokenLength	2	Mindestzahl an Zeichen, die für ein Token erforderlich sind, das komplett in Großbuchstaben geschrieben ist. Wenn die Länge des Token geringer als dieser Wert ist, wird es nicht indexiert. Dies ist ein ganzzahliger Wert ≥ 1 .
UseAlternateTerritories	No	<p>Gibt an, ob ein alternatives Wörterverzeichnis für die Stoppwortverarbeitung verwendet werden soll, wenn das sprachspezifische Wörterverzeichnis nicht gefunden werden kann.</p> <p>Wenn diese Option auf 'Yes' (Ja) gesetzt ist, wird das erste Stoppwortwörterverzeichnis verwendet, das mit der verwendeten Sprache übereinstimmt. Wenn die aktuelle Sprache beispielsweise en-gb ist, es aber keine Stoppwortdatei en-gb.tsw gibt, wird die Stoppwortdatei en-us.tsw verwendet.</p> <p>Wenn diese Option auf 'No' (Nein) gesetzt ist, wird kein Stoppwortfiltern durchgeführt.</p>

Konfigurationsdatei für die Textanalyse

[TAF_AnnotationType:TAF_ThesaurusEntry]

In diesem Abschnitt werden die Standardwerte für Abfragen gesetzt, für die Thesaurus-Erweiterung erforderlich ist.

Schlüsselwort	Standardwert	Beschreibung
ThesBaseName	desthes	Setzt den Standardnamen für den Thesaurus, der für Abfragen verwendet wird, für die Abfrageerweiterung erforderlich ist. Der Name ist eine Zeichenfolge.
Depth	1	Setzt den Standardpfad für das Durcharbeiten des Thesaurus, der für Abfragen verwendet wird, für die Thesaurus-Erweiterung erforderlich ist. Dies ist ein ganzzahliger Wert ≥ 1 .

[TAF_Plugin:TAF_POE]

Zum Ändern der Standardeinstellung für die Sprachumgebung, die beim Start der Textsuchmaschine vorab geladen ist, können Sie in der Option PreloadLanguages eine Sprache definieren.

Option	Standardwert	Beschreibung
PreloadLanguages	EN-US	Gibt eine Reihe sprachspezifischer Ressourcen an, die beim Systemstart vorab geladen werden sollen. Wenn eine Anwendung beispielsweise italienische Dokumente bearbeitet, werden die Ressourcen für Italienisch geladen. Der Wert dieses Felds hat die Form XX-YY, wobei XX für die Sprache und YY für das Gebiet steht.

Kapitel 17. Rückkehrcodes

In diesem Kapitel werden die Codes aufgelistet, die von der DB2 Text Extender-API als Antwort auf einen Funktionsaufruf zurückgegeben werden. Sie sind in alphabetischer Reihenfolge aufgelistet.

Alle DB2 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 DB2 Text Extender geliefert wird.

Die DB2 Text Extender-API fängt Fehlersituationen ab und berichtet Fehlerbedingungen mit einem Rückkehrcode.

Anwendungen, die DB2 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 13. API-Funktionen zum Suchen und Anzeigen“ auf Seite 231 aufgelistet.

In einigen Fällen kann eine falsche Eingabe, z. B. ein veralteter Sitzungszeiger, zu einer abnormalen Beendigungsbedingung in den API-Services führen, die nicht durch den DB2 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 DB2 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 Anzeigeeinformationen, die von DesGetSearchResultTable oder DesGetBrowseInfo zurückgegeben wurden und als Eingabe für DesStartBrowseSession verwendet werden, sind nicht gültig.

Aktion: Überprüfen Sie, ob ein Programmierfehler die Anzeigeeinformationen überschreibt.

RC_INVALID_BROWSE_OPTION

Erläuterung: Die Anzeigeeoption in DesGetSearchResultTable ist nicht gültig.

Aktion: Stellen Sie sicher, daß die Option BROWSE oder NO_BROWSE ist.

Rückkehrcodes

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 Fehlermeldung, die vom DB2 Text Extender zurückgegeben wurde, um die Ursache festzustellen.

RC_INVALID_SEARCH_OPTION

Erläuterung: Die Suchoption in DesGetSearch-ResultTable 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 DB2 Text Extender wurden keine Anzeigeeinformationen 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 12. Syntax von Suchargumenten“ auf Seite 217 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. 17 auf Seite 247 gezeigt.

RC_SE_BROWSER_TIME_OUT

Erläuterung: Der Anzeigeprozess wurde gestartet, jedoch erfolgte keine Antwort innerhalb der zulässigen Zeit. Der DB2 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 DB2 Text Extender-Server ist fehlgeschlagen. Der Fehler wurde möglicherweise durch fehlenden Speicherbereich oder durch eine inkorrekte Installation des DB2 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 8 auf Seite 225 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 DB2 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 DB2 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.

Rückkehrcodes

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 DB2 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 DB2 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 DB2 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 DB2 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 DB2 Text Extender-Funktion aufgerufen, die durch den Administrator verhindert wurde.

Aktion: Bitten Sie Ihren Administrator um Unterstützung. Es ist möglicherweise erforderlich, den DB2 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 DB2 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. 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 DesEnd-BrowseSession.

Überprüfen Sie, ob ein Programmierfehler eine falsche Kennung erstellt.

RC_SE_INDEX_DELETED

Erläuterung: Der DB2 Text Extender-Index, auf den zugegriffen werden soll, ist gelöscht.

Aktion: Kontaktieren Sie den DB2 Text Extender-Administrator, um den Index erneut zu erstellen.

RC_SE_INDEX_NOT_ACCESSIBLE

Erläuterung: Auf den DB2 Text Extender-Index kann nicht zugegriffen werden und der aktuelle Aufruf kann nicht verarbeitet werden.

Aktion: Bitten Sie den DB2 Text Extender-Administrator, die Zugriffsmöglichkeit auf den Index zu prüfen.

RC_SE_INDEX_SUSPENDED

Erläuterung: Der DB2 Text Extender hat eine Anforderung in bezug auf einen Index empfangen, der in einer anderen Sitzung oder in der aktuellen Sitzung ausgeschlossen wurde.

Aktion: Bitten Sie den DB2 Text Extender-Administrator, den Status des Indexes zu prüfen.

RC_SE_INSTALLATION_PROBLEM

Erläuterung: Der DB2 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 txinsnnn`, 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 beabsichtigte Aktion durch den Administrator, z. B. das Löschen einer DB2 Text Extender-Indexdatei
- Falsche Einstellung in der Textkonfiguration

Aktion: Beenden Sie die Anwendung. Prüfen Sie beim Administrator folgendes:

- Alle Dateien im aktuellen DB2 Text Extender-Index sind vorhanden
- Die Textkonfigurationseinstellungen sind richtig

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 Daten-

Rückkehrcodes

bank zugreifen oder der Benutzer ist für die Texttabelle nicht berechtigt.

Aktion: Überprüfen Sie, daß die Eingabe für die Funktion, z. 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 DB2 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 Anzeigeeinformationen 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 Anzeigeeinformationen anfordern, werden keine Anzeigeeinformationen zurückgegeben. Für die Anzeigeeinformationen 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 DB2 Text Extender nicht verwaltet werden kann.

RC_SE_REQUEST_IN_PROGRESS

Erläuterung: Ein DB2 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 DB2 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 DB2 Text Extender-Client kann derzeit keine Sitzung mit dem angeforderten DB2 Text Extender-Server aufbauen, oder die DB2 Text Extender-Server-Kommunikationsverbindung wurde unterbrochen und kann nicht erneut hergestellt werden.

Der DB2 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 Server.

RC_SE_SERVER_CONNECTION_LOST

Erläuterung: Die Verbindung zwischen Client und Server wurde unterbrochen und kann nicht erneut hergestellt werden.

Die DB2 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 DB2 Text Extender-API-Services konnten keine Sitzung mit dem angeforderten DB2 Text Extender-Server aufbauen.

Der DB2 Text Extender-Server wurde möglicherweise nicht gestartet.

Aktion: Überprüfen Sie, daß der DB2 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 DB2 Text Extender-Abfragen.

RC_SE_UNEXPECTED_ERROR

Erläuterung: Ein Fehler ist aufgetreten, der möglicherweise durch eine inkorrekte Installation des DB2 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.

Rückkehrcodes

Aktion: Geben Sie einen Abschnittsnamen an, der Teil des angegebenen Modells oder des Standardmodells ist.

RC_SE_UNKNOWN_SERVER_NAME

Erläuterung: Der Name des DB2 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 DB2 Text Extender-Server-Workstation oder durch eine inkorrekte Installation des DB2 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 DB2 Text Extender zurückgegeben wurde, nach weiteren Informationen, z. 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 18. Nachrichten

In diesem Kapitel werden folgende Themen beschrieben:

- **SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden:** Diese Nachrichten können angezeigt werden, wenn Sie DB2 Text Extender-Funktionen verwenden.
- **Nachrichten vom DB2TX-Befehlszeilenprozessor:** Diese Nachrichten können angezeigt werden, wenn Sie unter Verwendung des Befehlszeilenprozessors DB2TX Befehle eingeben. Vor jeder Nachrichtennummer steht das Präfix DES.

SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden

Die SQL-Funktionen, die vom DB2 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 Funktion 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 Funktion REFINE gibt das erste Suchargument anstelle des kombinierten Sucharguments zurück.

Aktion: Verringern Sie die Länge eines oder beider Suchargumente und wiederholen Sie anschließend die Abfrage.

SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden

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 DB2 Text Extender-Abfragen.

01H14 **Ein Wörterverzeichnis einer Sprache für die Verarbeitung auf linguistischer Basis fehlt.**

Erläuterung: Die linguistischen Services des DB2 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 8 auf Seite 225 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 des DB2 Text Extender 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-Katalogansicht SYSCAT.FUNCTIONS in der Spalte IMPLEMENTATION nach der Funktion, 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 DB2 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 des internen Textkennzeichens ist falsch.**

Erläuterung: Eine Kennung mit einem falschen Format wurde als Argument für eine DB2 Text Extender-Funktion verwendet.

Aktion: Stellen Sie sicher, daß die Kennung

SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden

nicht durch INIT_TEXT_HANDLE erstellt wurde.

38705 *udf_name* Falsche Deklaration der benutzerdefinierten Funktion.

Erläuterung: Der spezifische Name einer DB2 Text Extender-Funktion wurde in der Prozedur, in der Funktionen deklariert werden, geändert. Namen von DB2 Text Extender-Funktionen können geändert werden, aber nicht deren spezifische Namen.

Aktion: Überprüfen Sie die Prozedur DESCVD-FDDL, die die Deklarationen von DB2 Text Extender-Funktionen enthält, um sicherzustellen, daß noch immer die korrekten Namen verwendet werden. Vergleichen Sie die Namen mit denen im Originalprogrammdateiträger.

38706 *attribut* Dieses Attribut kann nicht erkannt werden.

Erläuterung: Ein Versuch wurde unternommen, eine CCSID, ein Format oder eine Sprache auf einen unbekanntem Wert zu setzen.

Aktion: Die korrekten Werte finden Sie in „Kapitel 4. Die Sucherfordernisse planen“ auf Seite 31.

38707 Die angeforderte Funktion wird noch nicht unterstützt.

Erläuterung: Die angegebene Funktion wird noch nicht unterstützt.

Aktion: Prüfen Sie die angegebene Funktion.

38708 *rückkehrcode* Suche ist fehlgeschlagen.

Erläuterung: Ein Fehler ist während der Verarbeitung einer Suchanforderung aufgetreten.

Aktion: Eine Beschreibung des Rückkehrcodes finden Sie in „Kapitel 17. Rückkehrcodes“ auf Seite 295.

38709 Nicht genügend Speicher verfügbar.

Erläuterung: Zum Ausführen der DB2 Text Extender-Funktion 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.

SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden

38717 **Der angegebene Thesaurus konnte nicht gefunden werden.**

Erläuterung: Der angegebene Thesaurus kann nicht gefunden werden.

Aktion: Prüfen Sie den angegebenen Thesaurus-Namen.

38718 **Der angegebene Relationsname konnte im Thesaurus nicht gefunden werden.**

Erläuterung: Die angegebene Relation ist im angegebenen Thesaurus nicht vorhanden.

Aktion: Stellen Sie sicher, daß die angegebene Relation vorhanden ist.

38719 **Ein Verarbeitungsfehler bei der Suche ist aufgetreten. Ursachen-code: rc.**

Erläuterung: Die Suche konnte aus dem angegebenen Grund nicht durchgeführt werden.

Aktion: Versuchen Sie, das im Ursachencode dargestellte Problem zu lösen. Wenn der angegebene Grund Ihnen nicht weiterhilft und keine weiteren Informationen in der Datei `desdiag.log` enthalten sind, erstellen Sie einen Trace und melden Sie diese Informationen Ihrem örtlichen IBM Ansprechpartner.

38720 **Fehler bei Zuordnung des gemeinsam benutzten Speichers aufgetreten.**

Erläuterung: Das System kann keinen Zugriff auf gemeinsam benutzten Speicher erhalten.

Aktion: Prüfen Sie Ihre Systemkonfiguration und erhöhen Sie die gemeinsam benutzten Ressourcen, oder prüfen Sie die aktuelle Verwendung gemeinsam benutzter Ressourcen (ipcs) und bereinigen Sie Ressourcen, die nicht mehr benötigt werden.

38721 **Fehler bei Semaphorerstellung/-zugriff aufgetreten.**

Erläuterung: Das System kann kein Semaphor erstellen bzw. darauf Zugriff erhalten.

Aktion: Prüfen Sie Ihre Systemkonfiguration und erhöhen Sie die gemeinsam benutzten Ressourcen, oder prüfen Sie die aktuelle Verwendung gemeinsam benutzter Ressourcen (ipcs) und bereinigen Sie Ressourcen, die nicht mehr benötigt werden.

38722 **Ein Suchvorgang ist nicht zurückgeliefert worden.**

Erläuterung: Ein Fehler ist während der Verarbeitung einer Suchanforderung aufgetreten.

Aktion: Prüfen Sie Ihre Systemkonfiguration `descfgcl` und stellen Sie sicher, daß alle Knoten aktiv sind.

38723 **Die CCSIDs des Indexes und der Abfrage stimmen nicht überein.**

Erläuterung: Die Datenbank-CCSID, die für die Abfragezeichenfolge verwendet wird, entspricht nicht der CCSID des Textindexes.

Aktion: Inaktivieren Sie den Textindex und erstellen Sie ihn unter Verwendung der CCSID der Datenbank erneut.

38724 **Der Abschnitts- oder Modellname ist nicht korrekt.**

Erläuterung: Der angegebene Abschnitts- oder Modellname in der Abfrage ist nicht richtig.

Aktion: Prüfen Sie den Abschnitts- oder Modellnamen.

38726 **Fehler beim Lesen der Modell-datei.**

Erläuterung: Die Modelldefinitionsdatei wurde nicht gefunden oder kann nicht geöffnet werden.

Aktion: Stellen Sie sicher, daß die Modelldefinitionsdatei im Indexverzeichnis vorhanden ist.

Nachrichten vom DB2 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 linguisti-

Nachrichten vom DB2 Text Extender

sche Angabe wird verwendet, wie in Tabelle 8 auf Seite 225 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 Funktionsextraktion 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 8 auf Seite 225.

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 DB2 Text Extender-Client kann keine Sitzung mit dem angeforderten Server aufbauen.

Aktion: Überprüfen Sie, daß der DB2 Text Extender-Server gestartet wurde. Andernfalls führen Sie TXSTART aus.

DES0709W Das Wörterverzeichnis für die Sprache *sprache* ist nicht installiert.

Erläuterung: Der DB2 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 DB2 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 '*knoten*' für den Knoten ist nicht in der Knotengruppendifinition enthalten.

Erläuterung: Die angegebene Knotennummer ist ungültig.

Aktion: Überprüfen Sie die DB2-Knotennummer.

DES0701N Der Zahlenwert '*knoten*' 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 'modell' 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 für den Parameter einen Wert an.

DES0711N Ein interner Fehler im Text Extender ist aufgetreten. Diagnoseinformationen: message.

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 ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie den Parameter unter Verwendung einer gültigen Länge an.

DES0713N Die Länge des Parameters 'parameter' ist falsch: %d1.

Erläuterung: Der angegebene Parameter ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie den 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 Schemenname und der Typenname der Textspalte oder das

Nachrichten vom DB2 Text Extender

Ergebnis einer Zugriffsfunktion. Der Datentyp für eine Textspalte wird vom DB2 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 „Unterstützte Dokumentformate“ auf Seite 33.

DES0717N **Sprache *sprache* wird nicht unterstützt.**

Erläuterung: *sprache* ist eine Sprache, die entweder nicht vom DB2 Text Extender unterstützt wird, oder die ausgewählte Sprache wird nicht vom angegebenen Indextyp unterstützt.

Aktion: Überprüfen Sie die Liste der unterstützten Sprachen in Tabelle 5 auf Seite 57.

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. 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 DB2 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 DB2 Text

Extender aktiviert werden kann.

Aktion: Stellen Sie sicher, daß der Tabellename 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 *kennungsspalte* 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

Nachrichten vom DB2 Text Extender

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.

DES0736N ***kennungsspalte* ist bereits eine Kennungsspalte in Tabelle *schema.tabelle*.**

Erläuterung: Sie versuchen, einen vorhandenen Kennungsspaltenamen 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. B. `DB2TX_INSTOWNER` und `DB2TX_INSTOWNERHOMEDIR`, korrekt gesetzt sind.

DES0740N **Kennungsspalte '*spalte*' kann nur für eine Textspalte des Typs *spaltentyp* wiederverwendet werden.**

Erläuterung: Die Kennungsspalte wurde bereits für einen anderen Kennungsspaltenname 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 '*commit-zählung*' 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 '*indexwert*' ist beim DB2 Text Extender nicht bekannt.

Erläuterung: Der angegebene Wert für 'UPDATE INDEX' 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 'UPDATE FREQUENCY' ist ungültig.

Aktion: Geben Sie einen gültigen Wert für 'UPDATE FREQUENCY' an.

Nachrichten vom DB2 Text Extender

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 'tabellenname' 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 Umgebungsvariable

riablen *umgebungsvar* zu lesen, aber diese Umgebungsvariable ist nicht definiert.

Aktion: Definieren Sie die erforderliche Umgebungsvariable.

DES0774N Die Länge des Werts *länge* für die Variable '*variable*' ist nicht im gültigen Bereich.

Erläuterung: Die Wert für die Länge des Parameters ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie unter Verwendung einer gültigen Länge einen Parameter an.

DES0775N Der Wert für das Indexverzeichnis '*verzeichnis*' 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 '*tabellenbereich*' 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: Prüfen 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: Prüfen 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 Tabellename 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.

Nachrichten vom DB2 Text Extender

Aktion: Wenn der Tabellename 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 *tabellename* 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 DB2 Text Extender aktiviert.

Aktion: Aktivieren Sie die Tabelle für den DB2 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 '*aliasname*' ist kein bekannter Aliasname für eine Datenbank. Nur *nn* 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 DB2 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 Befehl ist ein Tabellenname erforderlich.

Aktion: Geben Sie den entsprechenden Tabellennamen ein.

DES0832N Unerwartetes Befehlsende. Der Datenbankname fehlt.

Erläuterung: Für den Befehl ist ein Datenbankname erforderlich.

Aktion: Geben Sie den entsprechenden Datenbanknamen ein.

DES0833N Unerwartetes Befehlsende. Der Spaltenname fehlt.

Erläuterung: Für den Befehl ist ein Spaltenname erforderlich.

Aktion: Geben Sie den entsprechenden Spaltennamen ein.

Nachrichten vom DB2 Text Extender

DES0899N **Unbekannter DB2TX-Befehl:**
befehl.

Erläuterung: Der angegebene Befehl wird vom DB2 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 DB2 Text Extender verwendet, ist ein Fehler aufgetreten.

Aktion: Überprüfen Sie den Ursachencode der Suchmaschine in „Kapitel 19. Ursachencodes der Suchmaschine“ auf Seite 321. 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 DB2 Text Extender ist aufgetreten.

Aktion: Verwenden Sie die vom DB2 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 DB2 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 19. Ursachencodes der Suchmaschine

In diesem Kapitel werden die Ursachencodes aufgelistet, die von der Suchmaschine des DB2 Text Extender zurückgegeben werden.

Tabelle 21. Ursachencodes der Suchmaschine

Ursachencodes	Werte
RC_DONE	0
RC_CONTINUATION_MODE_ENTERED	1
RC_END_OF_INFORMATION	2
RC_EMPTY_LIST	3
RC_MORE_INFORMATION	4
RC_INDEX_GROUP_SEARCH_ERROR	7
RC_INDEX_SPECIFIC_ERROR	8
RC_DICTIONARY_NOT_FOUND	9
RC_PROCESSING_LIMIT_EXCEEDED	12
RC_UNKNOWN_SERVER_NAME	16
RC_INCORRECT_AUTHENTICATION	17
RC_DATASTREAM_SYNTAX_ERROR	18
RC_QUERY_SCOPE_TOO_COMPLEX	20
RC_QUERY_TOO_COMPLEX	22
RC_MEMBER_OF_INDEX_GROUP	23
RC_UNKNOWN_INDEX_NAME	24
RC_INCORRECT_HANDLE	25
RC_INDEX_NOT_MEMBER_OF_GROUP	26
RC_UNKNOWN_SESSION_POINTER	27
RC_UNKNOWN_COMMUNICATION_TYPE	29
RC_UNKNOWN_SERVER_INFORMATION	30
RC_INVALID_MASKING_SYMBOL	31
RC_UNEXPECTED_ERROR	32
RC_SERVER_NOT_AVAILABLE	33
RC_INDEX_ALREADY_OPENED	35
RC_MAX_NUMBER_OF_OPEN_INDEXES	36
RC_MAX_NUMBER_OF_RESULTS	37

Ursachencodes der Suchmaschine

Tabelle 21. Ursachencodes der Suchmaschine (Forts.)

Ursachencodes	Werte
RC_CCS_NOT_SUPPORTED	41
RC_LANGUAGE_NOT_SUPPORTED	42
RC_CONFLICT_WITH_INDEX_TYPE	43
RC_MAX_INPUT_SIZE_EXCEEDED	46
RC_SERVER_BUSY	47
RC_SERVER_CONNECTION_LOST	48
RC_SERVER_IN_ERROR	49
RC_REQUEST_IN_PROGRESS	50
RC_UNKNOWN_INDEX_TYPE	51
RC_INCORRECT_INDEX_NAME	52
RC_INCORRECT_LS_EXECUTABLES	53
RC_INCORRECT_LIBRARY_ID	54
RC_INDEX_ALREADY_EXISTS	55
RC_MAX_NUMBER_OF_INDEXES	56
RC_INCORRECT_LOCATION	57
RC_LOCATION_IN_USE	58
RC_UNKNOWN_CONDITION	59
RC_INDEX_DELETED	60
RC_INDEX_SUSPENDED	61
RC_INDEX_NOT_ACCESSIBLE	62
RC_MAX_NUMBER_OF_BUSY_INDEXES	63
RC_CONFLICTING_TASK_RUNNING	64
RC_NOT_ENOUGH_MEMORY	65
RC_MAX_OUTPUT_SIZE_EXCEEDED	68
RC_COMMUNICATION_PROBLEM	70
RC_NO_ACTION_TAKEN	71
RC_EMPTY_INDEX	72
RC_EMPTY_QUERY	73
RC_INSTALLATION_PROBLEM	74
RC_FUNCTION_DISABLED	75
RC_FUNCTION_IN_ERROR	76
RC_IO_PROBLEM	77
RC_WRITE_TO_DISK_ERROR	78
RC_SERVER_VERSION_NOT_CURRENT	79

Table 21. Ursachencodes der Suchmaschine (Forts.)

Ursachencodes	Werte
RC_FUNCTION_NOT_SUPPORTED	80
RC_RESULT_ALREADY_RANKED	81
RC_RESULT_VIEW_EXISTS	82
RC_INDEX_NOT_OPEN	83
RC_NO_RANKING_DATA_AVAILABLE	84
RC_LINGUISTIC_SERVICE_FAILED	85
RC_THESAURUS_PROBLEM	86
RC_INVALID_IDENTIFIER	88
RC_DOCUMENT_MODEL_ALREADY_EXISTS	89
RC_UNKNOWN_DOCUMENT_SECTION_NAME	90
RC_DOCMOD_READ_PROBLEM	91
RC_UNKNOWN_DOCUMENT_MODEL_NAME	92
RC_SECTION_NAME_ALREADY_EXISTS	94
RC_SECTION_TAG_ALREADY_EXISTS	95
RC_MAX_NUMBER_OF_TASKS	96
RC_LS_NOT_EXECUTABLE	97
RC_LS_FUNCTION_FAILED	98
RC_CAPACITY_LIMIT_EXCEEDED	99
RC_DOCUMENT_NOT_ACCESSIBLE	100
RC_DOCUMENT_CURR_NOT_ACCESSIBLE	101
RC_DOCUMENT_NOT_TO_INDEX	102
RC_DOCUMENT_NOT_FOUND	103
RC_DOCUMENT_IN_ERROR	104
RC_DOCUMENT_NOT_SUPPORTED	105
RC_CROSSIDX_SEARCH_NOT_ALLOWED	110
RC_DOCUMENT_GROUP_NOT_FOUND	111
RC_INVALID_ATTRIBUTE_VALUE	112
RC_INVALID_SECTION_TYPE	113
RC_INCORRECT_RELEVANCE_VALUE	120
RC_NO_RAT_EXPANSION	130
RC_DOCUMENT_NOT_IN_VIEW	131

Ursachencodes der Suchmaschine

Kapitel 20. Ursachencodes für Fehlerereignisse

In diesem Kapitel werden die Fehlerereignisse aufgelistet, die auftreten können, wenn der DB2 Text Extender Dokumente indiziert. Sie können beispielsweise in folgenden Fällen auftreten:

- Dokumente können nicht indiziert werden.
- Dokumente werden indiziert, 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 indiziert. Eine der Indexdateien konnte nicht geöffnet werden.

281

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht gelesen werden.

441

Das Dokument wurde nicht indiziert. Diese Nachricht tritt nur bei Ngram-Indizes auf. Die Codepage des Dokuments unterscheidet sich von der, mit der der Index erstellt wurde. Dies kann für HTML- und XML-Dokumente auftreten, wenn der Index nicht in Codepage UTF8 erstellt wurde.

Ursachencodes für Fehlerereignisse

500

Das Dokument wurde nicht indiziert. Die Bibliotheksfunktionen konnten nicht geladen werden. Stellen Sie sicher, daß die DLL verfügbar und der Ressourcenpfad gültig ist.

501

Das Dokument wurde nicht indiziert. 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 indiziert. Ein Fehler ist aufgetreten, während der Dokumentinhalt in der Bibliotheksfunktion LIB_read_doc_content gelesen wurde.

503

Das Dokument wurde nicht indiziert. Ein Fehler ist in der Bibliotheksfunktion LIB_access_doc aufgetreten.

504

Das Dokument wurde nicht indiziert. 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.

507

Der Aufruf der Bibliotheksservices LIB_read_doc_content ist mit einem unerwarteten Rückkehrcode fehlgeschlagen.

508

Der Aufruf der Bibliotheksservices LIB_close_doc hat einen Fehler RC_TERMINATION zurückgegeben.

545

Das Dokument wurde nicht indiziert. Eine der temporären Indexdateien konnte nicht geöffnet werden.

546

Das Dokument wurde nicht indiziert. Eine der temporären Indexdateien konnte nicht geschlossen werden.

548

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

549, 550

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung.

551-564

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

565

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung.

566-587

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden.

588-590

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

591-604

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden.

605

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung.

606-623

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

624

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung.

625-631

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht

Ursachencodes für Fehlerereignisse

geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

632

Eine der während der Indexierung erstellten temporären Dateien konnte nicht mit Schreibzugriff geöffnet werden. Prüfen Sie die Zugriffsberechtigungen.

633

Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschlossen werden.

634

Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschrieben werden. Stellen Sie sicher, daß das Indexarbeitsverzeichnis über ausreichend Plattenspeicherplatz verfügt.

635

Eine der während der Indexierung erstellten temporären Dateien konnte nicht gelesen werden.

636

Eine der während der Indexierung erstellten temporären Dateien konnte nicht mit Lesezugriff geöffnet werden. Prüfen Sie die Zugriffsberechtigungen.

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.

661

Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschlossen 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.

664

Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschlossen 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.

668-669

Das Dokument wurde nicht indiziert. Es gab ein Problem beim Abgleich von Abschnittsbefehlen im Dokument mit denen, die in der Dokumentmodelldatei definiert wurden.

670-672

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

673

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung.

674

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

675-687

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

688, 690

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung. Versuchen Sie die Verwendung von kleineren Werten in der Konfigurationsdatei.

689

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

691-695

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin

Ursachencodes für Fehlerereignisse

geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

696-707

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

708

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung. Versuchen Sie die Verwendung von kleineren Werten in der Konfigurationsdatei.

709-718

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

719-721

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

722, 729

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung. Versuchen Sie die Verwendung von kleineren Werten in der Konfigurationsdatei.

730, 732, 733, 735-738

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der Platte ausreichend Speicherbereich für den Index vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

731, 739-742, 744-746, 749, 755-758, 760-761, 767

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

743, 748, 750-754, 759, 765-766, 768-770

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

747, 763, 764

Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung. Verringern Sie die Arbeitsbelastung. Versuchen Sie die Verwendung von kleineren Werten in der Konfigurationsdatei.

815

Es gibt zwei mögliche Ursachen:

- Eine der Ressourcendateien wurde nicht gefunden, die zur Unterstützung der Sprache benötigt wird, die für das Dokument verwendet wird, das den Fehler verursacht hat
- Die vom betreffenden Dokument angeforderte Sprache wird vom DB2 Text Extender nicht unterstützt

831

Das Dokument wurde nicht indiziert. Kein Text konnte gefunden werden. Die Dokumentlänge ist 0 Byte.

860

Fehler beim Öffnen einer Datei. Einige Wörterverzeichnisse oder die Thesaurus-Dateien konnten nicht gefunden werden. Prüfen Sie Ihren Ressourcenpfad für die Wörterverzeichnisdateien. Wenn Sie während der Suche Pfaddaten für Ihre Thesaurus-Dateien angegeben haben, prüfen Sie den Standort und den Dateinamen.

954-956

Interner Fehler. Schicken Sie die Informationen im Diagnoseprotokoll an Ihren IBM Ansprechpartner.

957-967

Das Dokument wurde nicht indiziert. Eine der Indexdateien konnte nicht geöffnet, gelesen oder geschlossen werden bzw. es konnte nicht darin geschrieben werden. Stellen Sie sicher, daß auf der für den Index verwendeten Platte ausreichend Speicherbereich vorhanden ist und daß die Zugriffsberechtigungen korrekt sind.

1000

Ein Fehler ist während des Öffnens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen.

1001

Ein Fehler ist während des Anhängens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen.

1002

Ein Fehler ist während des Lesens einer Datei aufgetreten. Möglicherweise ist die Datei beschädigt.

Ursachencodes für Fehlerereignisse

1003

Ein Fehler ist während des Schreibens einer Datei aufgetreten. Prüfen Sie den Plattenspeicherplatz und die Zugriffsberechtigungen.

1005

Ein Fehler ist während des Lesens einer Datei aufgetreten (Positionierung in der Datei). Möglicherweise ist die Datei beschädigt.

1006

Ein Fehler ist während des Umbenennens einer temporären Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen.

1007

Ein Fehler ist während des Erstellens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen.

1008

Ein Fehler ist während des Komprimierens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen.

1009

Ein Fehler ist während des Schließens einer Datei aufgetreten. Möglicherweise ist die Datei beschädigt.

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.

1014

Ein Fehler ist während des Kopierens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen und den Plattenspeicherplatz.

1017

Der Indexname ist unbekannt. Stellen Sie sicher, daß die Schreibung korrekt ist.

1019

Ein Fehler ist während des Löschens einer Datei aufgetreten. Prüfen Sie die Zugriffsberechtigungen. Diese Fehlernachricht kann als "Folgefehler"

auftreten - lesen Sie die Einträge in der Diagnosedatei um festzustellen, ob ein vorangehender Fehlereintrag weitere Informationen bietet.

1020

Allgemeiner Dateifehler. Prüfen Sie die Zugriffsberechtigungen.

1070-1074

Das Dokument wurde nicht indiziert. Die angegebene Codepage ist entweder allgemein oder nur für den Index ungültig, auf den zugegriffen wird.

1085

Das Dokument wurde nicht indiziert. Ein Fehler ist beim Lesen der Indizierungswarteschlange aufgetreten.

1086

Das Dokument wurde nicht indiziert. Die Indizierungswarteschlange ist leer.

1116-1117

Das Dokument wurde nicht indiziert. Informationen aus der Initialisierungsdatei des Server-Exemplars konnten nicht verarbeitet werden. Stellen Sie sicher, daß die Einträge in der Initialisierungsdatei gültig sind und daß die Anwendung auf die Datei zugreifen kann.

1129

Kein Dokument wurde indiziert. Das Starten der Hintergrundverarbeitung ist fehlgeschlagen.

1158

Beim Umbenennen einer Datei ist ein Fehler aufgetreten. Prüfen Sie die Zugriffsberechtigungen und den Plattenspeicherplatz.

1162

Die Indexdateien eines Ngram-Indexes sind möglicherweise beschädigt.

1163, 1164

Das Dokument wurde aufgrund eines unerwarteten Fehlers nicht indiziert.

1165

Das Dokument wurde aufgrund einer unerwarteten Dateiendebedingung nicht indiziert.

1176

Keine weiteren Dokumente können für einen Ngram-Index indiziert werden. Es gibt eine Überlaufbedingung für Dokumentnummern (Überlauf von 'LONG'). Wenn es viele Löschoptionen oder wiederholte Aktualisierungen desselben Dokuments gab, versuchen Sie zum Lösen des Problems einen Aufruf an EhwReorg. Wenn dies nicht der Fall war, verwenden Sie einen zweiten Index für neue Dokumente.

Ursachencodes für Fehlerereignisse

1177

Das Dokument wurde nicht indiziert. Es war für den Ngram-Indexierer zu umfangreich.

1189

Das Dokument wurde nicht indiziert. Es gab ein Problem mit der Grenzreihenfolge (spezifisch für die koreanische Sprache).

1198-1200

Das Dokument wurde nicht indiziert. Es gab ein Problem mit dem Indexzugriff. Der Index könnte beschädigt sein.

1201

Das Dokument wurde nicht indiziert. Die Codepage des Dokuments konnte nicht in die indexspezifische Codepage umgewandelt werden. Dieser Fehler tritt nur bei Ngram-Indizes in der Codepage UTF8 auf.

1202

Das Dokument wurde nicht indiziert. Die Codepage des Dokuments konnte aufgrund von ungültigen Daten im Dokument nicht in die indexspezifische Codepage umgewandelt werden. Dieser Fehler tritt nur bei Ngram-Indizes in der Codepage UTF8 auf.

1500-1505

In der Komponente zur Dokumentanalyse gibt es Probleme. Sie konnte entweder nicht initialisiert werden (prüfen Sie LIBPATH und den Inhalt der Konfigurationsdatei) oder schlug aufgrund von internen Problemen fehl. Weitere Informationen enthält die Diagnosedatei.

1904

Das Dokument wurde nicht indiziert. Beim Zugriff auf das Dokumentmodell eines abschnittsfähigen Indexes gibt es ein Problem. Prüfen Sie die Zugriffsberechtigungen und stellen Sie sicher, daß die Datei vorhanden ist.

2000

Das Dokument wurde nicht indiziert. Der Dokumenttyp wird nicht unterstützt. Die Bibliotheksfunktion Lib_access_doc gab einen ungültigen Dokumenttyp zurück.

2001

Das Dokument wurde nicht indiziert. Eine falsche Reihenfolge von Feldern wurde im Datenstrom des Dokuments festgestellt.

2002

Das Dokument wurde nicht indiziert. Ein falsch strukturiertes Feld wurde im Datenstrom des Dokuments festgestellt.

2003

Das Dokument wurde nicht indiziert. Nur ein Textkapitel ist für ein Dokument im DB2 Text Extender-Textformat zulässig.

2005

Das Dokument wurde nicht indiziert. Eine Sprache, die im Datenstrom des Dokuments angegeben ist, wird nicht unterstützt.

2006

Das Dokument wurde nicht indiziert. Eine CCSID, die im Datenstrom des Dokuments angegeben ist, wird nicht unterstützt.

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 indiziert, da darauf nicht zugegriffen werden konnte.

2009

Das Dokument wurde nicht indiziert, da es im Gebrauch war und darauf nicht zugegriffen werden konnte.

2010

Das Dokument wurde nicht indiziert. Die angegebene CCSID ist nicht korrekt.

2011

Das Dokument wurde nicht indiziert, 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 indiziert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein strukturiertes Feld enthält eine falsche Längenangabe.

2013

Das Dokument wurde nicht indiziert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein falsches Steuerelement wurde im Dokument festgestellt.

2014

Das Dokument wurde nicht indiziert, 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 indiziert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein doppelter Dokumentparameter wurde gefunden.

Ursachencodes für Fehlerereignisse

2016

Das Dokument wurde nicht indiziert, 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 indiziert. Es ist weder ein WordPerfect-Dokument noch eine WordPerfect-Datei.

2021

Das Dokument wurde nicht indiziert. Es ist eine WordPerfect-Datei, aber kein WordPerfect-Dokument.

2022

Das Dokument wurde nicht indiziert. Diese Version von WordPerfect wird nicht unterstützt.

2023

Das Dokument wurde nicht indiziert. 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.

2028

Der Parser hat einen Nicht-Dokumenttext zurückgegeben. Prüfen Sie den Dateiinhalt, insbesondere hinsichtlich formatspezifischer Wörter. Prüfen Sie, ob das Dokumentformat unterstützt wird. Wenn die automatische Formaterkennung fehlschlägt, stellen Sie sicher, daß der richtige Parser aufgerufen wird.

2030

Das Dokument wurde nicht indiziert. 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 indiziert. Ein unerwartetes Dateiende wurde in einem Microsoft Word-Dokument festgestellt.

2032

Das Dokument wurde nicht indiziert. 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 für Macintosh-Dokument; es kann nicht verarbeitet werden. Speichern Sie das Dokument im Word für 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-Steurelement, auf das keine Zahl folgt.

2042

Das Dokument wurde nicht indexiert. Es enthält ein .SO LEN-Steurelement, 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-Steurelement ist zulässig. Sichern Sie jedes ECTF-Dokument jeweils in einer separaten Datei.

2044

Das Dokument wurde nicht indexiert. Auf ein .SO HDE-Steurelement muß ein Anfangsbefehl und ein Endbefehl folgen.

2046

Das Dokument wurde nicht indexiert. Das Dokument enthält Text vor dem .SO DOC-Steurelement.

2047

Das Dokument wurde nicht indexiert. Das Dokument enthält Text vor einem .SO PID-Steurelement.

2048

Das Dokument wurde nicht indexiert. Ein Endbefehl fehlt nach einem Anfangsbefehl.

Ursachencodes für Fehlerereignisse

2050

Das Dokument wurde nicht indiziert. Falsche Befehle wurde nach einem .SO HDE-Steurelement festgestellt.

2051

Das Dokument wurde nicht indiziert. Ein Zeilenende wurde nach einem .SO-Steurelement festgestellt.

2052

Das Dokument wurde nicht indiziert. Ein unerwartetes Textende wurde festgestellt.

2060

Das Dokument wurde nicht indiziert. Entweder ist es kein AmiPro-Dokument oder es handelt sich um eine Version von AmiPro, die nicht unterstützt wird.

2061

Das Dokument wurde nicht indiziert. Die Länge eines Steurelements in einem AmiPro-Dokument ist zu lang.

2062

Das Dokument wurde nicht indiziert. Diese Version von AmiPro wird nicht unterstützt. Nur AmiPro Architecture Version 4 wird unterstützt.

2063

AmiPro Style Sheets wurden nicht indiziert.

2064

Das Dokument wurde nicht indiziert. Ein falscher Zeichensatz wurde festgestellt. Nur Lotus Character Set 82 (Windows ANSI) wird unterstützt.

2065

Das Dokument wurde nicht indiziert. 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 indiziert. Entweder ist es kein RTF-Dokument oder es handelt sich um eine Version von RTF, die nicht unterstützt wird.

2081

Das Dokument wurde nicht indiziert. Ein RTF-Steuerwort, das zu lang ist, wurde festgestellt.

2083

Das Dokument wurde nicht indiziert. Die Macintosh-Codepage wird nicht unterstützt.

2084

Das Dokument wurde nicht indiziert. Es ist ein RTF-Dokument, aber diese Version von RTF wird nicht unterstützt. Nur RTF Version 1 wird unterstützt.

2090

Das Dokument wurde nicht indiziert. Es handelt sich dabei um ein HTML-Dokument, das einen Befehl enthält, der vom Parser als zu lang betrachtet wird.

2093

Das Dokument wurde nicht indiziert. Es handelt sich dabei um ein XML-Dokument, das vom XML-Parser zurückgewiesen wurde.

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 indiziert werden, da es leer ist oder keinen Text enthält. Überprüfen Sie, ob das Dokument nur Grafiken enthält.

2102

Das Dokument kann nicht indiziert werden, da es entweder kennwortgeschützt oder verschlüsselt ist.

2105

Der Dokumenttyp ist bekannt, aber der Filter ist nicht verfügbar.

2106

Das Dokument kann nicht indiziert werden, da es leer ist.

2107

Das Dokument kann nicht indiziert werden, da es nicht geöffnet werden kann. Den Dokumentzugriff prüfen.

2112

Das Dokument kann nicht indiziert werden, da es eine ausführbare Datei ist.

2113

Das Dokument kann nicht indiziert werden, da es komprimiert ist.

Ursachencodes für Fehlerereignisse

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.

Teil 3. Anhänge und Schlußteil

Bemerkungen

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.

Verweise in diesen Informationen auf Web-Sites anderer Anbieter dienen lediglich als Benutzerinformationen und stellen keinerlei Billigung des Inhalts dieser Web-Sites dar. Das über diese Web-Sites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Web-Sites geschieht auf eigene Verantwortung.

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

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 Canada Limited
Office of the Lab Director
1150 Eglinton Ave. East
North York, Ontario
M3C 1H7
CANADA

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.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer gesteuerten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Garantie, daß diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen. Diese Daten stellen deshalb keine Leistungsgarantie dar.

Informationen über Produkte anderer Hersteller als IBM wurden von den Herstellern dieser Produkte zur Verfügung gestellt, bzw. aus von ihnen veröffentlichten Ankündigungen oder anderen öffentlich zugänglichen Quellen entnommen. IBM hat diese Produkte nicht getestet und übernimmt im Hinblick auf Produkte anderer Hersteller keine Verantwortung für einwandfreie Funktion, Kompatibilität oder andere Ansprüche. Fragen hinsichtlich des Leistungsspektrums von Produkten anderer Hersteller als IBM sind an den jeweiligen Hersteller des Produkts zu richten.

Die oben genannten Erklärungen bezüglich der Produktstrategien und Absichtserklärungen von IBM stellen die gegenwärtige Absicht der IBM dar, unterliegen Änderungen oder können zurückgenommen werden, und repräsentieren nur die Ziele der IBM.

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 Quellsprache 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.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

© (Name Ihrer Firma) (Jahr). Teile des vorliegenden Codes wurden aus Beispielprogrammen der IBM Corp. abgeleitet. © Copyright IBM Corp. _Jahr/Jahre angeben_. Alle Rechte vorbehalten.

Marken

Folgende Namen sind in gewissen Ländern Marken oder Dienstleistungsmarken der IBM Corporation:

ACF/VTAM	IBM
AISPO	IMS/ESA
AIX	LAN Distance
AIXwindows	MVS
AnyNet	MVS/ESA
APPN	MVS/XA
AS/400	Net.Data
BookManager	OS/2
CICS	OS/390
C Set++	OS/400
C/370	PowerPC
DATABASE 2	QBIC
DataHub	QMF
DataJoiner	RACF
DataPropagator	RS/6000
DataRefresher	S/370
DB2	SP
DB2 Connect	SQL/400
DB2 Extenders	System/370
DB2 OLAP Server	System/390
DB2 Universal Database	SystemView
Distributed Relational Database Architecture	VisualAge
DRDA	VM/ESA
Extended Services	VSE/ESA
FFST	VTAM
First Failure Support Technology	WebExplorer
	WIN-OS/2

Folgende Namen sind in gewissen Ländern Marken oder eingetragene Marken anderer Unternehmen:

Microsoft, Windows, Windows NT und das Logo von Windows sind in gewissen Ländern Marken oder eingetragene Marken der Microsoft Corporation.

Java und alle auf Java basierenden Marken sind in gewissen Ländern Marken von Sun Microsystems, Inc.

UNIX ist in gewissen Ländern eine eingetragene Marke der Open Group.

Andere Namen von Unternehmen, Produkten oder Dienstleistungen können Marken anderer Unternehmen sein.

Glossar

In diesem Glossar werden viele der Begriffe und Abkürzungen definiert, die in diesem Handbuch verwendet werden.

A

Abrufen. Das Finden eines Textdokuments unter Verwendung eines Sucharguments in einer der Suchfunktionen des DB2 Text Extender.

Aktivieren. Eine Datenbank, eine Texttabelle oder eine Textspalte für die Verwendung durch den DB2 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 DB2 Text Extender.

Anzeigen. Text auf einem Computerbildschirm sichtbar machen.

API. Anwendungsprogrammierschnittstelle (Application Programming Interface).

Auslöser. Ein Mechanismus, der immer automatische Informationen über Dokumente, die indiziert 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, DB2 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 DB2 Text Extender stellt Verwaltungs- und Suchfunktionen, z. 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. 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 DB2 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 DB2 Text Extender beim Indexieren die Abschnitte innerhalb von Dokumenten. 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 Dokumentmodelldatei angeben.

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 DB2 Text Extender-Umgebung. Sie können mehrere Exemplare des DB2 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. Der Typ eines Dokument, z. 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 *Boolescher Suche* und *Freitextsuche*.

I

Inaktivieren. Eine Datenbank, eine Texttabelle oder eine Textspalte in den Zustand zurückbringen, den sie vor der Aktivierung für den DB2 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 es exakte oder linguistische Formen von Dokumentbegriffen enthält. Siehe *Präziser Index*, *Linguistischer 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 DB2 Text Extender zu Verwaltungszwecken erstellt wurde. Eine Katalogsicht enthält Informationen zu den Tabellen und Spalten, die zur Verwendung durch den DB2 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 DB2 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 physisch 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* und *Ngram-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. 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 Überein-

stimmung in SBCS-Dokumenten unterstützt. Siehe auch *Linguistischer Index* und *Präziser 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 physisch 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* und *Ngram-Index*.

Profil. Siehe *Umgebungsprofil*.

Protokolltabelle. Eine vom DB2 Text Extender erstellte Tabelle, die Informationen dazu enthält, welche Textdokumente indiziert 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. 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* und *Linguistischer 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 eines Fehlers zu finden.

U

Übereinstimmung. Das Vorkommen eines Suchbegriffs in einem Textdokument.

UDF. Benutzerdefinierte Funktion.

UDT. Benutzerdefinierter eindeutiger Typ.

Umgebungsprofil. Eine mit dem DB2 Text Extender gelieferte Prozedur, die Einstellungen für *Umgebungsvariablen* enthält.

Umgebungsvariable. Eine Variable, mit der Standardwerte für die Werte für die DB2 Text Extender-Umgebung geliefert werden.

V

Vorkommen. Synonym zu *Übereinstimmung*.

W

Wörterverzeichnis. Eine Sammlung von sprachbezogenen linguistischen Informationen, die der DB2 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 DB2 Text Extender verarbeitet werden kann.

Index

Sonderzeichen

& (AND), Operator in Suchargumenten

Suchargumentsyntax 223

Verwendung 87

| (OR), Operator in Suchargumenten

Suchargumentsyntax 223

Verwendung 87

A

Abfrage, Verarbeitung auf linguistischer Basis 266

Abkürzungen

Editieren einer Abkürzungsdatei 58

Listen 57

Abrufen der Suchergebnistabelle,

API-Funktion

Beschreibung 245

Verwendung 121

Abrufen von Anzeigeeinformationen,

API-Funktion

Beschreibung 236

Verwendung 123

Abrufen von Übereinstimmungen,

API-Funktion

Beschreibung 239

Verwendung 124

Abschnitte in Dokumenten

Aktivieren der Abschnittsunterstützung 50

Attributabschnitte 51

Attributwert in Suchsyntax 222

DESMODEL.INI 51

Dokumentmodelldatei, Inhalt 54

HTML-Dokumente 54

MODEL, Schlüsselwort in der Suchsyntax 221

SECTION, Schlüsselwort in der Suchsyntax 221

Suchbeispiel 90

unstrukturierte Dokumente 54

unterstützte Datumsformate 52

unterstützte Formate für ganze

Zahlen 53

unterstützte Gleitkommatformate 53

unterstützte Zeitformate 52

XML-Dokumente 55

Aktualisieren eines Textindexes

Ändern der Frequenz 103

regelmäßig 48

UPDATEFREQ in CHANGE

INDEX SETTINGS 132

UPDATEFREQ in CHANGE

TEXT CONFIGURATION 135

Aktualisierungsfrequenz

ändern 103

Beschreibung 48

GET INDEX SETTINGS,

Befehl 163

Standardwert in den Textkonfigurationseinstellungen 20

Syntax 48

UPDATEFREQ in CHANGE

INDEX SETTINGS 132

UPDATEFREQ in CHANGE

TEXT CONFIGURATION 135

Aktualisierungsstatus, anzeigen

Beispiel und Ausgabe 108

Syntax 165

Aktualisierungsstatus, zurücksetzen

Beispiel 103

Syntax 171

AmiPro, Dokumentformat 33

Analyse von Text

für die Anzeige 269

für die Indexierung 258

AND

Boolescher Operator 87

Schlüsselwort in einem Suchargument 224

Anwendungsprogrammierschnittstelle (API)

Abrufen des Zeigers für Hervorhebungsinformationen 239

Abrufen einer Suchergebnistabelle 121

Abrufen von Anzeigeeinformationen 123

Abrufen von Übereinstimmungen 124

Anzeigefunktionen 119

Beenden einer Anzeigesitzung 125

DesCloseDocument, Funktion 233

Anwendungsprogrammierschnittstelle (API) (*Forts.*)

DesEndBrowseSession, Funktion 234

DesFreeBrowseInfo, Funktion 235

DesGetBrowseInfo, Funktion 236

DesGetMatches, Funktion 239

DesGetSearchResultTable, Funktion 245

DesOpenDocument, Funktion 250

DesStartBrowseSession, Funktion 253

Freigeben des Speichers für Anzeigeeinformationen 125

Hervorhebungsinformationen 239

nach Text suchen 121

Nachrichten 303

Öffnen eines Dokuments für die Anzeige 123

Programmbeispiel 255

Referenz 231

Rückkehrcodes 295

Schließen eines Dokuments 125

Starten einer Anzeigesitzung 123

Suchfunktionen 119

Überblick 120

Zusammenfassung 231

Anzeigefunktionen 119

Anzeigen

den eigenen Browser verwenden 122

Programmbeispiel 255

Verarbeitung auf linguistischer Basis 269

ASCII, Dokumentformat 33

ASCII unstrukturiert, Dokumentformat 33

Assistent, Beispiel

Beschreibung 29

TXWIZARD, Befehl 198

Auslöser

Beschreibung 32

erstellen 70

B

Basistextanalyse

- Begriffe, die nichtalphanumerische Zeichen enthalten 258
- für die Hervorhebung 269
- Normalisierung 259
- Satzerkennung 260
- zum Indexieren von Begriffen 258

Bedarf an Plattenspeicherplatz für

- Indizes 47

Beenden einer Anzeigesitzung, API-Funktion

- Beschreibung 234
- Verwendung 125

Befehle

- CHANGE INDEX SETTINGS 131
- CHANGE TEXT CONFIGURATION 133
- CONNECT 137
- DB2TX 130
- DELETE INDEX EVENTS 139
- DISABLE DATABASE 140
- DISABLE TEXT COLUMN 141
- DISABLE TEXT FILES 142
- DISABLE TEXT TABLE 143
- ENABLE DATABASE 144
- ENABLE TEXT COLUMN 145
- ENABLE TEXT FILES 154
- ENABLE TEXT TABLE 157
- GET ENVIRONMENT 162
- GET INDEX SETTINGS 163
- GET INDEX STATUS 165
- GET STATUS 166
- GET TEXT CONFIGURATION 167
- GET TEXT INFO 168
- QUIT 169
- REORGANIZE INDEX 170
- RESET INDEX STATUS 171
- TXICRT 176
- TXIDROP 178
- TXILIST 179
- TXIPCLEAN 180
- TXNADD 181
- TXNCHECK 182
- TXNDROP 183
- TXSAMPLE 184
- TXSTART 185
- TXSTATUS 186
- TXSTOP 187
- TXTHESC 188
- TXTHESN 190
- TXTRACE 192

Befehle (Forts.)

- TXVERIFY 197
 - TXWIZARD 198
 - UPDATE INDEX 172
 - Zusammenfassung, Client-Befehle 129
 - Zusammenfassung, Server-Befehle 175
 - Befehlszeilenprozessor
 - DB2TX, Befehl 130
 - Hilfe 62
 - QUIT, Befehl 169
 - starten 61
 - Begriffserweiterung für die Hervorhebung 269
 - Beispiel-API-Programm 255
 - Beispielassistent 29
 - Beispieldatenbank
 - erstellen (TXVERIFY) 17
 - Beispielfunktionen des DB2 Text Extender
 - aktiv 80
 - Beispieltabelle
 - Beschreibung 80
 - löschen 117
 - TXSAMPLE, Befehl 184
 - Benutzer-Exit, Dokumentformatumwandlung 35
 - Boolesche Operatoren
 - & (AND) und | (OR) 87
 - NOT 92
 - Boolesches Suchargument 223
 - BOUND, Schlüsselwort 225
 - BOUND-Suche, Beispiel 93
- ## C
- CASE_ENABLED, Schlüsselwort
 - in ENABLE TEXT COLUMN 149
 - in ENABLE TEXT TABLE 159
 - CCSID
 - Beschreibung 37
 - Codepage-Probleme vermeiden 40
 - Extrahieren aus einer Kennung 98
 - Funktion 202
 - GET TEXT INFO, Befehl 168
 - in CHANGE TEXT CONFIGURATION 136
 - in ENABLE TEXT COLUMN 147
 - Initialisieren in Kennungen 208
 - Liste 37
 - Standardwert in der Textkonfiguration 20

- CHANGE INDEX SETTINGS, Befehl
 - Syntax 131
 - Verwendung 102
- CHANGE TEXT CONFIGURATION, Befehl
 - Syntax 133
 - Verwendung 21
- Client/Server-Umgebung 5
- COMMITCOUNT, Konfigurationsparameter
 - Beschreibung 72
 - in CHANGE TEXT CONFIGURATION 135
 - in ENABLE TEXT COLUMN 152
 - in ENABLE TEXT TABLE 173
 - Protokollbereich beibehalten 73
 - Standardwert in den Textkonfigurationseinstellungen 21
- CONNECT, Befehl
 - Syntax 137
 - Verwendung 63
- CONTAINS, Funktion
 - Beispiel 85
 - Syntax 203
- COUNT, Schlüsselwort 223

D

- DATALINK, Datentyp 74
- Datenbank
 - aktivieren 64
 - CONNECT, Befehl 137
 - DISABLE DATABASE, Befehl 140
 - ENABLE DATABASE, Befehl 144
 - GET STATUS, Befehl 166
 - inaktivieren 117
 - sichern und zurückschreiben 23
 - Statusinformationen, anzeigen 106
 - Verbindung herstellen 63
- Datenstromsyntax 240
- Datentypen von Textdokumenten
 - Aktivieren von Spalten, die den Datentyp DATALINK enthalten 74
 - Funktion zum Umsetzen 73, 74
 - unterstützt 146
- Datumsformate, Abschnittsunterstützung 52
- DB2 Extender
 - Beispiel für die Verwendung 4
- DB2DBDFT, Umgebungsvariable 18
- DB2INSTANCE, Umgebungsvariable 18

- DB2TEXTFH, eindeutiger Typ 199
- DB2TEXTFHLISTP, eindeutiger Typ 200
- DB2TEXTH, eindeutiger Typ 199
- DB2TEXTHLISTP, eindeutiger Typ 200
- DB2TX_, Umgebungsvariablen anzeigen 106
Beschreibung 18
- DB2TX, Befehlszeilenprozessor Syntax 130
Verwendung 61
- DB2TX_INSTOWNER, Umgebungsvariable 18
- DB2TX_INSTOWNERHOMEDIR, Umgebungsvariable 18
- DB2TX.SAMPLE, Tabelle Beschreibung 80
Dienstprogramm zum Erstellen 17
löschen 117
- DBCS-Dokumente, suchen in 45
- Dekomposition von Komposita 265
- DELETE INDEX EVENTS, Befehl Beispiel 104
Syntax 139
- DES_BROWSE, Option in DesGetSearchResultTable 246
- DES_EXT.H, Kopfdatei 119
- DES_EXTENDED, Option in DesOpenDocument 250
- DES_FAST, Option in DesOpenDocument 250
- DES_MATCH, Option in DesGetSearchResultTable 246
- DES_NOBROWSE, Option in DesGetSearchResultTable 246
- DES_RANK, Option in DesGetSearchResultTable 246
- DES_RANKANDMATCH, Option in DesGetSearchResultTable 246
- DES_TEXTHANDLEONLY, Option in DesGetSearchResultTable 246
- DESCL.INI 285
- DesCloseDocument, Funktion Beschreibung 233
Verwendung 125
- DesEndBrowseSession, Funktion Beschreibung 234
Verwendung 125
- DesFreeBrowseInfo, Funktion Beschreibung 235
Verwendung 125
- DesGetBrowseInfo, Funktion Beschreibung 236
- DesGetBrowseInfo, Funktion (*Forts.*) Verwendung 123
- DesGetMatches, Funktion Beschreibung 239
Verwendung 124
- DesGetSearchResultTable, Funktion Beschreibung 245
Verwendung 121
- DESMODEL.INI 51
- DesOpenDocument, Funktion Beschreibung 250
Verwendung 123
- DESRESTB, zum Erstellen einer Ergebnistabelle 247
- DESSAMP1, Beispielprogramm 255
- DESSRV.INI 287
- DesStartBrowseSession, Funktion Beschreibung 253
Verwendung 123
- DIRECTORY, Schlüsselwort Anzeigen der aktuellen Einstellung 110
in CHANGE TEXT CONFIGURATION 135
in ENABLE TEXT COLUMN 151
in ENABLE TEXT TABLE 160
Standardwert in den Textkonfigurationseinstellungen 20
- DISABLE DATABASE, Befehl Syntax 140
Verwendung 117
- DISABLE TEXT COLUMN, Befehl Syntax 141
Verwendung 115
- DISABLE TEXT FILES, Befehl Syntax 142
Verwendung 116
- DISABLE TEXT TABLE, Befehl Syntax 143
Verwendung 116
- Dokument Anzeigen der Einstellungen 111
CCSID 37
Format, Beschreibung 33
Format in CHANGE TEXT CONFIGURATION 136
Format in ENABLE TEXT COLUMN 148
Format umwandeln 35
GET TEXT INFO, Befehl 168
indexieren 31
Informationen 111
Sprache 37
Struktur 50
- Dokument (*Forts.*) Umsetzen von Datentypen 73
unterstützte Datentypen 146
unterstützte Formate 33
- Dokumentmodell Ändern der Dokumentmodelldatei 50
Attributwert in Suchsyntax 222
Beschreibung 50
MODEL, Schlüsselwort in der Suchsyntax 221
SECTION, Schlüsselwort in der Suchsyntax 221
- Dokumentmodelldatei, Inhalt 54
- E**
- Eindeutige Typen 199
- Einschränkungen für Suchargumente 229
- ENABLE DATABASE, Befehl Syntax 144
Verwendung 64
- ENABLE TEXT COLUMN, Befehl Syntax 145
Verwendung 68
- ENABLE TEXT FILES, Befehl Syntax 154
Verwendung 75
- ENABLE TEXT TABLE, Befehl Syntax 157
Verwendung 65
- Ereignis, Ursachencodes 325
- Ergebnistabelle 247
- Erkennen von Sätzen 260
- Erste Schritte 27
- Erstellen einer Beispieldatenbank 197
- Erstellen einer Beispieldatenbank TXSAMPLE, Befehl 184
- Erstellen eines DB2 Text Extender-Exemplars TXICRT, Befehl 176
- Erweiterte Übereinstimmung 270
- Erweiterung von Begriffen für die Hervorhebung 269
- Escape-Zeichen Syntax 229
Verwendung 89
- Exemplar löschen TXIDROP, Befehl 178
Vorgehensweise 22
- Exemplare Befehl für Liste, TXILIST 179
erstellen 22
löschen 22
Umgebungsvariablen 18

- EXPAND, Schlüsselwort 223
- Extender
 - Beispiel für die Verwendung 4
- Externe Dateien
 - Abrufen oder Ändern eines Dateinamens in einer Kennung 204
 - aktivieren 75
 - Ändern des Pfades/Namens in Kennungen 99
 - Aspekte der Indexaktualisierung 102
 - DISABLE TEXT FILES, Befehl 142
 - ENABLE TEXT FILES, Befehl 154
 - Extrahieren des Pfades/Namens aus einer Kennung 98
 - FILE, Funktion 204
 - inaktivieren 116
 - Initialisieren von Kennungen 208
 - Kennungen 83
- F**
 - Fehlerereignisse
 - anzeigen 109
 - aufzeichnen 70
 - DELETE INDEX EVENTS 139
 - GET INDEX STATUS, Befehl 165
 - löschen 104
 - Ursachencodes 325
 - Fehlersuche 24
 - Feinanpassung einer vorherigen Suche 95
 - FFT, Dokumentformat 33
 - FILE, Funktion
 - Beispiel 98
 - Syntax 204
 - FORMAT, Funktion
 - Beispiel 98
 - Syntax 205
 - Format von Textdokumenten 33
 - Ändern in Kennungen 99
 - Beschreibung 33
 - Extrahieren aus einer Kennung 98
 - FORMAT, Funktion 205
 - FORMAT, Schlüsselwort 136, 148
 - GET TEXT INFO, Befehl 168
 - in CHANGE TEXT CONFIGURATION 136
 - Format von Textdokumenten 33
 - (Forts.)
 - in ENABLE TEXT COLUMN 148
 - Initialisieren in Kennungen 208
 - Liste der unterstützten 33
 - nicht unterstütztes umwandeln 35
 - Standardwert in der Textkonfiguration 20
 - Freitextsuche
 - Beispiel 95
 - FUNCTION, Schlüsselwort
 - Beschreibung 73
 - in ENABLE TEXT COLUMN 146
 - Funktion
 - API-Funktionen 119
 - Pfad für DB2 Text Extender-Funktionen setzen 83
 - SET CURRENT FUNCTION PATH, Anweisung 83
 - Suchfunktionen 79
 - zum Umsetzen von Datentypen 73
 - Funktionen
 - Angaben von Suchargumenten 87
 - Beschreibung 79
 - CCSID 202
 - CONTAINS 203
 - Feinanpassung einer vorherigen Suche 95
 - FILE 204
 - FORMAT 205
 - Funktionspfad 83
 - HANDLE 206
 - HANDLE_LIST 207
 - INIT_TEXT_HANDLE 208
 - LANGUAGE 210
 - nach Text suchen 84
 - NO_OF_DOCUMENTS 211
 - NUMBER_OF_MATCHES 212
 - RANK 213
 - Referenz 199
 - REFINE 214
 - SEARCH_RESULT 215
 - Setzen und Extrahieren von Informationen in Kennungen 97
 - Überblick 200
 - Verbessern der Suchleistung 99
 - zurückgegebene SQLSTATES 303
- FUZZY FORM OF, Schlüsselwort 225
- G**
 - Ganze Zahlen, Formate, Abschnittsunterstützung 53
 - GET ENVIRONMENT, Befehl
 - Beispiel und Ausgabe 106
 - Syntax 162
 - GET INDEX SETTINGS, Befehl
 - Beispiel und Ausgabe 110
 - Syntax 163
 - GET INDEX STATUS, Befehl
 - Beispiel und Ausgabe 108
 - Syntax 165
 - GET STATUS, Befehl
 - Beispiel und Ausgabe 106
 - Syntax 166
 - GET TEXT CONFIGURATION, Befehl
 - Beispiel und Ausgabe 107
 - Syntax 167
 - GET TEXT INFO, Befehl
 - Beispiel und Ausgabe 111
 - Syntax 168
 - Gleitkommaformate, unterstützte 53
 - Große Tabellen, aktivieren 72
 - Grundform, Begriffe reduzieren 265
 - GUI, Beispiel 29
- H**
 - HANDLE, Funktion
 - Syntax 206
 - Verwendung 99
 - HANDLE_LIST, Funktion
 - Syntax 207
 - Verwendung 99
 - Hervorhebungsinformationen
 - Datenstrom 124
 - Datenstromsyntax 239
 - Hilfe für Befehle 62
 - HTML, Dokumentformat 33
 - HTML, strukturierte Dokumente 50
 - HTML-Dokumente, Abschnittsunterstützung 54
 - Hybridsuche, Beispiel 95
- I**
 - IN SAME PARAGRAPH AS, Schlüsselwort 224
 - IN SAME SENTENCE AS, Schlüsselwort 224
 - Include-Datei, des_ext.h 119

Index

Aktualisieren für externe Dateien 102
Aktualisierungsfrequenz 48
Ändern der Aktualisierungsfrequenz 103
Ändern der aktuellen Einstellungen 102
Ändern des Indextyps 46
Anzeigen der aktuellen Einstellungen 110
CASE_ENABLED, Option 46
CHANGE INDEX SETTINGS, Befehl 131
CHANGE TEXT CONFIGURATION, Befehl 133
Erstellen von verschiedenen Typen für eine Textspalte 71
GET INDEX SETTINGS, Befehl 163
GET INDEX STATUS, Befehl 165
GET TEXT CONFIGURATION, Befehl 167
Größenberechnung 47
INDEXOPTION in ENABLE TEXT COLUMN 149
INDEXOPTION in ENABLE TEXT TABLE 159
INDEXTYPE in CHANGE TEXT CONFIGURATION 134
INDEXTYPE in ENABLE TEXT COLUMN 149
INDEXTYPE in ENABLE TEXT TABLE 158
linguistisch (LINGUISTIC) 43
mehrere, verwenden 46
Ngram (NGRAM) 45
planen 31
präzise (PRECISE) 45
regelmäßige Indexaktualisierung 48
reorganisieren 105
sichern und zurückschreiben 23
sofortige Indexaktualisierung 101
Standardtyp in den Textkonfigurationseinstellungen 20
Tabelle mit einem gemeinsam benutzten Index 46
TABLESPACE in CHANGE TEXT CONFIGURATION 134
Textkonfiguration ändern 21
Typen 42
Überblick 31

Index (Forts.)

UPDATE INDEX, Befehl 172
verwalten 101
Indexieren, Verarbeitung auf linguistischer Basis 257
Indexierungsereignisse
Ursachencodes 325
Indexierungsereignisse, löschen
Beispiel 104
Syntax 139
Indexmerkmale
anzeigen 110
in ENABLE TEXT COLUMN 145
in ENABLE TEXT FILES 154
in ENABLE TEXT TABLE 157
Standardwerte in den Textkonfigurationseinstellungen 20
INDEXOPTION, Schlüsselwort
in CHANGE TEXT CONFIGURATION 134
in ENABLE TEXT COLUMN 149
in ENABLE TEXT TABLE 159
INDEXPROPERTY, Schlüsselwort
in ENABLE TEXT COLUMN 150
in ENABLE TEXT TABLE 159
Indexstatus, anzeigen
Beispiel und Ausgabe 108
Syntax 165
Indexstatus, zurücksetzen
Beispiel 103
Syntax 171
Indextyp, ändern
ändern 46
Erstellen von verschiedenen Typen für eine Textspalte 71
INDEXTYPE, Schlüsselwort
in CHANGE TEXT CONFIGURATION 134
in ENABLE TEXT COLUMN 149
in ENABLE TEXT TABLE 158
Informationen zu Textdokumenten
Anzeigen der aktuellen Einstellung 111
CCSID 37
Format 33
GET TEXT INFO, Befehl 168
Sprache 37
Typen 33
INIT_TEXT_HANDLE, Funktion
Beispiel 97
Syntax 208

Initialisieren einer Kennung
INIT_TEXT_HANDLE, Funktion 208
Vorgehensweise 97
Installationsprüfung 17

K

Katalogsicht
erstellen 64
Inhalt 112
löschen 117
Kennung
Abrufen aus einer Liste von Kennungen 206
Ändern von Format und Sprache 99
Beschreibung 82
CCSID, Funktion 202
eindeutiger Typ, DB2TEXTFH 199
eindeutiger Typ, DB2TEXTH 199
Extrahieren von CCSID, Format und Sprache 98
FORMAT, Funktion 205
für externe Dateien 83
initialisieren 97
LANGUAGE, Funktion 210
Setzen und Extrahieren von Informationen 97
Verwendung von Listen zur Verbesserung der Leistung 99
Kennungslistenzeiger (eindeutiger Typ DB2TEXTFHLISTP) 200
Kennungslistenzeiger (eindeutiger Typ DB2TEXTHLISTP) 200
Knoten
Befehl zum Hinzufügen von Knoten, TXNADD 181
Befehl zum Löschen von Knoten, TXNDROP 183
Befehl zum Prüfen von Knoten, TXNCHECK 182
Knotengruppen und Tabellenbereiche 70
Umgebung mit mehreren Knoten 50
zuordnen 8
Kompilieren einer Ngram-Thesaurus-Definitionsdatei 190
Kompilieren einer Thesaurus-Definitionsdatei 188
Komposita, trennen 265
Konfiguration 18
Konfigurationsdateien 285

Konfigurationstabelle
 anzeigen 107
 CHANGE TEXT CONFIGURATION, Befehl 133
 erstellen 64
 GET TEXT CONFIGURATION, Befehl 167
 Kopfdatei, des_ext.h 119

L

LANGUAGE, Funktion
 Beispiel 98
 Syntax 210

LANGUAGE, Schlüsselwort 136, 147

Lauterweiterung
 Beispiel 93
 Beschreibung 268

Leistung, verbessern 99

Linguistischer Index
 Beschreibung 43
 Suchoption, Standardwerte 225

Logischer Knoten, zuordnen 8

LOGPRIMARY, LOGSECOND und LOGFILSIZ, Parameter in DB2 72

M

Mehrere Indizes, verwenden 46

Mehrpartitions-knotengruppe 8

Microsoft, Dokumentformat 33

N

Nach Text suchen
 Abrufen der Anzahl der gefundenen Übereinstimmungen 86
 Abrufen der Rangordnung eines gefundenen Dokuments 86
 Durchführen einer Abfrage 85
 Erstellen einer Liste von gefundenen Dokumenten 207
 Feinanpassung einer vorherigen Suche 95
 Programmbeispiel 255
 REFINE, Funktion 214
 SEARCH_RESULT, Funktion 215
 Syntax 219
 Überblick 84
 Verbessern der Leistung 99
 Verwendung der API 121

Nachrichten 303

Ngram-Index
 Beschreibung 45
 CASE_ENABLED, Option 46
 Suchoption, Standardwerte 225

NO_OF_DOCUMENTS, Funktion
 Syntax 211

NODE, Schlüsselwort
 in ENABLE TEXT COLUMN 152
 in ENABLE TEXT TABLE 160

Normalisierung von Begriffen 259

NORMALIZED, Schlüsselwort
 in ENABLE TEXT COLUMN 150

NOT
 Boolescher Operator 92
 Schlüsselwort in einem Suchargument 224

NUMBER_OF_MATCHES, Funktion,
 Syntax 212

O

Öffnen eines Dokuments, API-Funktion
 Beschreibung 250
 Verwendung 123

ON NODE, Schlüsselwort
 in ENABLE TEXT COLUMN 152
 in ENABLE TEXT TABLE 160

OR, Boolescher Operator 87

P

Parallelismus, Konzepte 8

Physischer Knoten, zuordnen 8

Plattenspeicherplatz für Indizes 47

Platzhalterzeichen
 in einem Suchbegriff 88
 Platzhalterzeichen für Wörter 268
 Verarbeitung auf linguistischer Basis 268

Platzhalterzeichen für Zeichen 268

Präzise Suche 45

Präziser Index
 Beschreibung 45
 Suchoption, Standardwerte 225

PRECISE FORM OF, Schlüsselwort 225

Protokollbereich, nicht genügend 72

Protokolltabelle
 Beschreibung 32
 erstellen 70
 Extrahieren von Fehlerereignissen 109
 Zuordnen zum Tabellenbereich 70

Q

QUIT, Befehl
 Syntax 169

QUIT, Befehl (*Forts.*)
 Verwendung 78

R

Rangordnung
 aus DesGetSearchResultable 121
 in einem Suchergebnis 86

RANK, Funktion
 Beispiel 86
 Syntax 213

Reduzieren von Begriffen auf ihre Grundform 265

REFINE, Funktion
 Beispiel 95
 Syntax 214

Regeln für Suchargumente 229

REORGANIZE INDEX, Befehl
 Beispiel 105
 Syntax 170

RESET INDEX STATUS, Befehl
 Beispiel 103
 Syntax 171

RESULT LIMIT, Schlüsselwort 223

Rückkehrcodes 295

S

Satzerkennung 260

Satztrennung 45

Schließen eines Dokuments, API-Funktion
 Beschreibung 233
 Verwendung 125

Schnittstelle, Beispiel 29

SEARCH_RESULT, Funktion
 Beispiel 99
 Syntax 215

Server
 Anzeigen des Status 186
 konfigurieren und verwalten 22
 starten 185
 stoppen 187
 Trace-Funktion 24
 TXICRT, Befehl 176
 TXIDROP, Befehl 178
 TXILIST, Befehl 179
 TXIPCLEAN, Befehl 180
 TXNADD, Befehl 181
 TXNCHECK, Befehl 182
 TXNDROP, Befehl 183
 TXSAMPLE, Befehl 184
 TXSTART, Befehl 185
 TXSTATUS, Befehl 186
 TXSTOP, Befehl 187
 TXTRACE, Befehl 192
 TXVERIFY, Befehl 197

- Server (*Forts.*)
 - TXWIZARD, Befehl 198
 - SET CURRENT FUNCTION PATH, Anweisung 83
 - Shell-Profile 18
 - Sichern und Zurückschreiben 23
 - Spalte
 - aktivieren 68
 - Aktivieren für verschiedene Indextypen 71
 - Aktivieren in einer großen Tabelle 72
 - DISABLE TEXT COLUMN, Befehl 141
 - ENABLE TEXT COLUMN, Befehl 145
 - inaktivieren 115
 - Speicherplatz freigeben für Anzeigeformen, API-Funktion
 - Beschreibung 235
 - Verwendung 125
 - Sprache von Textdokumenten
 - Ändern in Kennungen 99
 - Beschreibung 37
 - Extrahieren aus einer Kennung 98
 - GET TEXT INFO, Befehl 168
 - in einem Suchargument 91
 - Initialisieren in Kennungen 208
 - LANGUAGE, Funktion 210
 - Liste 37
 - Standardwert in der Textkonfiguration 20
 - Sprachparameter, Liste 57
 - SQLSTATES, die von DB2 Text Extender-Funktionen zurückgegeben werden 303
 - Starten des DB2 Text Extender-Servers 185
 - Starten einer Anzeigesitzung, API-Funktion
 - Beschreibung 253
 - Verwendung 123
 - Status des DB2 Text Extender-Servers 186
 - Status eines Indexes
 - anzeigen 108
 - Anzeigen des aktuellen Status 165
 - zurücksetzen 103
 - zurücksetzen nach einem Fehler 103
 - STEMMED FORM OF, Schlüsselwort 225
 - Stoppen des DB2 Text Extender-Servers 187
 - Stoppwörter
 - als Teil der Basistextanalyse 265
 - Beschreibung 31
 - Editieren einer Stoppwortdatei 58
 - Listen 57
 - Struktur von Dokumenten
 - Aktivieren der Abschnittsunterstützung 50
 - Attributwert in Suchsyntax 222
 - MODEL, Schlüsselwort in der Suchsyntax 221
 - SECTION, Schlüsselwort in der Suchsyntax 221
 - Suchbeispiel 90
 - Suchargument
 - & (AND), Operator 223
 - | (OR), Operator 223
 - AND, Schlüsselwort 224
 - angeben 87
 - Attributwert 222
 - Beschreibung 217
 - BOUND, Schlüsselwort 225
 - BOUND-Suche 93
 - COUNT, Schlüsselwort 223
 - EXPAND, Schlüsselwort 223
 - Freitextsuche 95
 - FUZZY FORM OF, Schlüsselwort 225
 - Hybridsuche 95
 - IN SAME PARAGRAPH AS 224
 - IN SAME SENTENCE AS 224
 - MODEL, Schlüsselwort 221
 - NOT, Schlüsselwort 224
 - PRECISE FORM OF, Schlüsselwort 225
 - RESULT LIMIT, Schlüsselwort 223
 - SECTION, Schlüsselwort 221
 - STEMMED FORM OF, Schlüsselwort 225
 - Suche nach grober Übereinstimmung 93, 225
 - Suchen mit & und | 87
 - Suchen mit NOT 92
 - Suchen nach ähnlich klingenden Wörtern 93
 - Suchen nach Begriffen im selben Absatz 90
 - Suchen nach Begriffen im selben Satz 90
 - Suchen nach Begriffen in beliebiger Reihenfolge 90
 - Suchargument (*Forts.*)
 - Suchen nach Begriffen in Dokumentabschnitten 90
 - Suchen nach Begriffen in verschiedenen Sprachen 91
 - Suchen nach mehreren Begriffen 87
 - Suchen nach Synonymen 90
 - Suchen nach Teilen eines Begriffs 88
 - Suchen nach Variationen eines Begriffs 88
 - SYNONYM FORM OF, Schlüsselwort 225
 - Syntax 219
 - TERM OF, Schlüsselwort 223
 - THESAURUS, Schlüsselwort 222
 - Thesaurus-Suche 94
 - unter Verwendung von Platzhalterzeichen 88
 - Zusammenfassung der Regeln und Einschränkungen 229
 - Suche nach grober Übereinstimmung 43
 - Suche nach grober Übereinstimmung, Beispiel 93
 - Suchfunktionen 119
 - Suchmaschine, Ursachencodes 321
 - Suchstatus, anzeigen
 - Beispiel und Ausgabe 108
 - Syntax 165
 - Suchstatus, zurücksetzen
 - Beispiel 103
 - Syntax 171
 - Synonyme
 - Beschreibung 267
 - in einem Suchargument 90
 - SYNONYM FORM OF, Schlüsselwort 225
- ## T
- Tabelle mit einem gemeinsam benutzten Index
 - Beschreibung 46, 47
 - ENABLE TEXT TABLE, Befehl 157
 - erstellen 65
 - Tabellenbereich 70
 - Tabellenbereiche und Knotengruppen 70
 - TERM OF, Schlüsselwort 223
 - TEXTINDEXES, Katalogsicht
 - erstellen 64
 - Inhalt 112

- TEXTINDEXES, Katalogsicht (*Forts.*)
löschten 117
- Textkonfigurationseinstellungen
ändern 21
anzeigen 107
Installationsstandardwerte 19
- Textmerkmale
CCSID 37
Format 33
in ENABLE TEXT COL-UMN 145
in ENABLE TEXT FILES 154
Sprache 37
Standardwerte in der Textkonfiguration 20
- Texttabelle
Aktivieren einer Spalte in einer großen Tabelle 72
DISABLE TEXT TABLE, Befehl 143
ENABLE TEXT TABLE, Befehl 157
sichern und zurückschreiben 23
- Thesaurus-Suche
Beispiel 94
Erstellen eines Ngram-Thesaurus 281
Erstellen eines Thesaurus 276
Kompilieren einer Ngram-Thesaurus-Definitionsdatei 190
Kompilieren einer Thesaurus-Definitionsdatei 188
Konzepte 271
Syntax 222
THESAURUS, Schlüsselwort 222
TXTHESC, Befehl 188
TXTHESN, Befehl 190
- Tiefe von Begriffen in einem Thesaurus, angeben 223
- Trace-Funktion
einstellen 24
TXTRACE, Befehl 192
- TXICRT, Befehl
DB2 Text Extender-Exemplar erstellen 22
Syntax 176
- TXIDROP, Befehl
Syntax 178
- TXILIST, Befehl
Syntax 179
- TXIPCLEAN, Befehl
Syntax 180
- TXNADD, Befehl
Syntax 181
- TXNCHECK, Befehl
Syntax 182
- TXNDROP, Befehl
Syntax 183
- TXSAMPLE, Befehl
Syntax 184
Verwendung 17
- TXSAMPLE.UDF
aktiv 80
- TXSTART, Befehl
Syntax 185
Verwendung 23
- TXSTATUS, Befehl
Syntax 186
Verwendung 23
- TXSTOP, Befehl
Syntax 187
Verwendung 23
- TXTHESC, Befehl
Syntax 188
- TXTHESN, Befehl
Syntax 190
- TXTRACE, Befehl
Syntax 192
Verwendung 24
- TXVERIFY
Erstellen einer Beispieldatenbank 17
- TXVERIFY, Befehl
Syntax 197
- TXWIZARD, Befehl
Syntax 198
- Typen von Textindizes
CASE_ENABLED, Option 46
CHANGE TEXT CONFIGURATION, Befehl 134
GET INDEX SETTINGS, Befehl 163
INDEXTYPE in CHANGE TEXT CONFIGURATION 134
INDEXTYPE in ENABLE TEXT COLUMN 149
INDEXTYPE in ENABLE TEXT TABLE 158
linguistisch (LINGUISTIC) 43
Ngram (NGRAM) 45
präzise (PRECISE) 45
Standardwert in den Textkonfigurationseinstellungen 20
Suchoption, Standardwerte 225
- U**
Übereinstimmung
aus DesGetSearchResultTable 121
- Übereinstimmung (*Forts.*)
DesGetMatches, Funktion 239
in einem Suchergebnis 86
NUMBER_OF_MATCHES, Funktion 212
Übereinstimmung, erweitert 270
Übersicht über den DB2 Text Extender 3
UDTs 199
Umgebung, Client/Server 5
Umgebung mit mehreren Knoten 50
Umgebungsvariablen 18
anzeigen 106
Beschreibung 18
GET ENVIRONMENT, Befehl 162
Unstrukturierte Dokumente, Abschnittsunterstützung 54
Unterstützung nationaler Zeichensätze 40
Unterstützung partitionierter Datenbanken 8
UPDATE INDEX, Befehl
Beispiel 101
Syntax 172
UPDATEFREQ, Schlüsselwort in CHANGE INDEX SETTINGS 132
in CHANGE TEXT CONFIGURATION 135
in ENABLE TEXT COL-UMN 150
in ENABLE TEXT TABLE 159
UPDATEINDEX, Schlüsselwort
Anzeigen der aktuellen Einstellung 110
GET INDEX SETTINGS, Befehl 163
in CHANGE TEXT CONFIGURATION 135
in ENABLE TEXT COL-UMN 152
Standardwert in den Textkonfigurationseinstellungen 20
Ursachencodes von der Suchmaschine 321
- V**
Variablen
Anzeigen von Umgebungsvariablen 106
Beschreibung der Umgebungsvariablen 18
GET ENVIRONMENT, Befehl 162

Verarbeitung auf linguistischer Basis

- Basistextanalyse 258
- Begriffserweiterung 269
- beim Indexieren 257
- Beschreibung 257
- erweiterte Übereinstimmung 270
- für die Abfrage 266
- für die Anzeige 269
- Lauterweiterung 268
- Platzhalterzeichen 268
- Platzhalterzeichen für Wörter 268
- Platzhalterzeichen für Zeichen 268
- Reduzieren von Begriffen auf ihre Grundform 265
- Stoppwortfiltern 265
- Synonyme 267
- Trennen von Komposita 265

Verarbeitungsmerkmale

- Standardwerte in den Textkonfigurationseinstellungen 20

Verbindung zu einer Datenbank herstellen

- CONNECT, Befehl 137
- Vorgehensweise 63

Verwaltung

- Abkürzungsdatei, editieren 58
- Aktualisieren eines Indexes für externe Dateien 102
- Ändern der Indexeinstellungen 102
- Ändern von Stoppwort- und Abkürzungsdateien 58
- Anzeigen der Einstellungen für Textdaten 111
- Anzeigen der Indexeinstellungen 110
- Anzeigen des Indexstatus 108
- Anzeigen des Status von Datenbank, Tabelle und Spalte 306
- Befehlszeilenprozessor 130
- Befehlszusammenfassung, Client 129
- Befehlszusammenfassung, Server 175
- Beispieldatenbank, Dienstprogramm zum Erstellen 17
- CHANGE INDEX SETTINGS, Befehl 131
- CHANGE TEXT CONFIGURATION, Befehl 133
- CONNECT, Befehl 137

Verwaltung (Forts.)

- DB2 Text Extender-Exemplar erstellen 22
- DB2 Text Extender-Exemplar löschen 22
- DB2TX, Befehl 130
- DELETE INDEX EVENTS 139
- DELETE INDEX EVENTS, Befehl 139
- DISABLE DATABASE, Befehl 140
- DISABLE TEXT COLUMN, Befehl 141
- DISABLE TEXT FILES, Befehl 142
- DISABLE TEXT TABLE, Befehl 143
- Erstellen einer Beispieldatenbank 197
- GET ENVIRONMENT, Befehl 162
- GET INDEX SETTINGS, Befehl 163
- GET INDEX STATUS, Befehl 165
- GET STATUS, Befehl 166
- GET TEXT CONFIGURATION, Befehl 167
- GET TEXT INFO, Befehl 168
- Inaktivieren einer Datenbank 117
- Inaktivieren einer Textspalte 115
- Inaktivieren einer Texttafel 116
- Inaktivieren von Textdateien 116
- Kompilieren einer Ngram-Thesaurus-Definitionsdatei 190
- Kompilieren einer Thesaurus-Definitionsdatei 188
- Löschen von Indexereignissen 104
- QUIT, Befehl 169
- Reorganisieren eines Indexes 105
- REORGANIZE INDEX, Befehl 170
- RESET INDEX STATUS, Befehl 171
- Server-Status anzeigen 23
- sichern und zurückschreiben 23
- sofortiges Aktualisieren eines Indexes 101
- Starten des DB2 Text Extender-Servers 185

Verwaltung (Forts.)

- Statusinformationen, abrufen 105
- Stoppen des DB2 Text Extender-Servers 187
- Stoppwortdatei, ändern 58
- Textkonfiguration ändern 21
- Trace-Funktion 24
- TXICRT, Befehl 176
- TXIDROP, Befehl 178
- TXILIST, Befehl 179
- TXIPCLEAN, Befehl 180
- TXNADD, Befehl 181
- TXNCHECK, Befehl 182
- TXNDROP, Befehl 183
- TXSAMPLE, Befehl 184
- TXSTART, Befehl 185
- TXSTATUS, Befehl 186
- TXSTOP, Befehl 187
- TXTHESC, Befehl 188
- TXTHESN, Befehl 190
- TXTRACE, Befehl 192
- TXVERIFY, Befehl 197
- TXWIZARD, Befehl 198
- Umgebungsvariablen 18
- UPDATE INDEX, Befehl 172
- Verwalten von Textindizes 101
- Zurücksetzen des Indexstatus 103
- Zusammenfassung der Befehle, Client 129
- Zusammenfassung der Befehle, Server 175
- Verzeichnis für den Index
- GET INDEX SETTINGS, Befehl 163
- Vorkommen eines Suchbegriffs 212

W

- WordPerfect, Dokumentformat 33
- Wörterverzeichnisdateinamen 57
- Worttrennung 45

X

- XML, Dokumentformat 33
- XML, strukturierte Dokumente 50
- XML-Dokumente, Abschnittsunterstützung 55

Z

- Zeitformate, Abschnittsunterstützung 52
- Zugriffsfunktion
- Beschreibung 73
- in ENABLE TEXT COLUMN 146

Kontaktaufnahme mit IBM

Bei technischen Problemen lesen Sie bitte die entsprechenden Korrekturmaßnahmen im Handbuch *Troubleshooting Guide* und führen Sie diese aus, bevor Sie sich mit der IBM Kundenunterstützung in Verbindung setzen. Mit Hilfe dieses Handbuchs können Sie Informationen sammeln, die die DB2-Kundenunterstützung zur Fehlerbehebung verwenden kann.

Wenn Sie weitere Informationen benötigen oder eines der DB2 Universal Database-Produkte bestellen möchten, setzen Sie sich mit einem IBM Ansprechpartner in einer lokalen Geschäftsstelle oder einem IBM Software-Vertriebspartner in Verbindung.

Telefonische Unterstützung erhalten Sie unter der folgenden Nummer:

- Unter 0180 3/313 233 erreichen Sie Hallo IBM, wo Sie Antworten zu allgemeinen Fragen erhalten.

Produktinformationen

Telefonische Unterstützung erhalten Sie über folgende Nummern:

- Unter 0180 3/313 233 erreichen Sie Hallo IBM, wo Sie Antworten zu allgemeinen Fragen erhalten.
- Unter 0180/55 090 können Sie Handbücher telefonisch bestellen.

<http://www.ibm.com/software/data/>

Auf den DB2-World Wide Web-Seiten erhalten Sie aktuelle DB2-Informationen wie Neuigkeiten, Produktbeschreibungen, Schulungspläne und vieles mehr.

<http://www.ibm.com/software/data/db2/library/>

Mit **DB2 Product and Service Technical Library** können Sie auf häufig gestellte Fragen, Berichtigungen, Handbücher und aktuelle technische DB2-Informationen zugreifen.

Anmerkung: Diese Informationen stehen möglicherweise nur auf Englisch zur Verfügung.

<http://www.elink.ibm.com/pbl/pbl/>

Auf der Web-Site für die Bestellung internationaler Veröffentlichungen (International Publications) finden Sie Informationen zum Bestellverfahren.

<http://www.ibm.com/education/certify/>

Das 'Professional Certification Program' auf der IBM Web-Site stellt Zertifizierungstestinformationen für eine Reihe von IBM Produkten, u. a. auch DB2, zur Verfügung.

<ftp://software.ibm.com>

Melden Sie sich als *anonymous* an. Im Verzeichnis /ps/products/db2 finden Sie Demo-Versionen, Berichtigungen, Informationen und Tools zu DB2 und vielen zugehörigen Produkten.

<comp.databases.ibm-db2>, <bit.listserv.db2-1>

Über diese Internet-Newsgroups können DB2-Benutzer Ihre Erfahrungen mit den DB2-Produkten austauschen.

Für CompuServe: GO IBMDB2

Geben Sie diesen Befehl ein, um auf IBM DB2 Family Forums zuzugreifen. Alle DB2-Produkte werden über diese Foren unterstützt.

In Anhang A des Handbuchs *IBM Software Support Handbook* finden Sie Informationen dazu, wie Sie sich mit IBM in Verbindung setzen können. Rufen Sie die folgende Web-Seite auf, um auf dieses Dokument zuzugreifen:

<http://www.ibm.com/support/>. Wählen Sie anschließend die Verbindung zum IBM Software Support Handbook am unteren Rand der Seite aus.

Anmerkung: In einigen Ländern sollten sich die IBM Vertragshändler an die innerhalb ihrer Händlerstruktur vorgesehene Unterstützung wenden, nicht an die IBM Unterstützungsfunktion.

Antwort

**DB2 Universal Database
DB2 Text Extender
Verwaltung und Programmierung
Version 7**

IBM Form SC12-2893-00

Anregungen zur Verbesserung und Ergänzung dieser Veröffentlichung nehmen wir gerne entgegen. Bitte informieren Sie uns über Fehler, ungenaue Darstellungen oder andere Mängel.

Zur Klärung technischer Fragen sowie zu Liefermöglichkeiten und Preisen wenden Sie sich bitte entweder an Ihre IBM Geschäftsstelle, Ihren IBM Geschäftspartner oder Ihren Händler.

Unsere Telefonauskunft "HALLO IBM" (Telefonnr.: 01803/31 32 33) steht Ihnen ebenfalls zur Klärung allgemeiner Fragen zur Verfügung.

Kommentare:

Danke für Ihre Bemühungen.

Sie können ihre Kommentare betr. dieser Veröffentlichung wie folgt senden:

- Als Brief an die Postanschrift auf der Rückseite dieses Formulars
- Als E-Mail an die folgende Adresse: comment@tcvm.vnet.ibm.com

Name

Adresse

Firma oder Organisation

Rufnummer

E-Mail-Adresse

Antwort
SC12-2893-00



IBM Deutschland Informationssysteme GmbH
SW NLS Center

70548 Stuttgart



Printed in Ireland

SC12-2893-00

