

DB2[®] Universal Database



DB2 Text Extender

Administración y programación

Versión 7

DB2[®] Universal Database



DB2 Text Extender

Administración y programación

Versión 7

Nota!

Antes de utilizar esta información y el producto al que da soporte, asegúrese de leer la información general bajo "Notas" en la página 339.

Primera edición, junio de 2000

Esta edición se aplica a DB2 Universal Database Text Extender, una característica de la Versión 7 de DB2 Universal Database (DB2 UDB), 5645-DB2, y a todos los releases y modificaciones posteriores hasta que se indique lo contrario en nuevas ediciones.

Esta edición sustituye a SC26-9646-00 que queda obsoleta.

Este manual es la traducción del original inglés *DB2 Universal Database DB2 Text Extender Administration and Programming Version 7 (SC26-9930-00)*.

Este documento contiene información propiedad de IBM. Se proporciona bajo un contrato de licencia y está protegido por la ley de propiedad intelectual. La información contenida en esta publicación no incluye ninguna garantía del producto y cualquier declaración proporcionada en este manual no debe interpretarse como tal.

Pida las publicaciones a su representante de IBM o a la sucursal de IBM que opera en su localidad o llamando al 1-800-879-2755 en Estados Unidos o al 1-800-IBM-4YOU en Canadá.

Cuando se envía información a IBM, se otorga a IBM un derecho no exclusivo de utilizar o distribuir la información en cualquier manera que encuentre adecuada sin incurrir en ninguna obligación con el remitente.

© Copyright International Business Machines Corporation 1995, 2000. Reservados todos los derechos.

Contenido

Acerca de este manual	vii	Cambio de configuración del texto	20
Quién debe utilizar este manual	vii	Configuración y mantenimiento de un	
Cómo utilizar este manual	vii	servidor DB2 Text Extender	21
Cómo leer los diagramas de sintaxis	viii	Creación de una instancia de DB2 Text	
Cómo enviar sus comentarios	ix	Extender	21
		Inicio y detención de un servidor DB2 Text	
Resumen de Cambios	xi	Extender	22
		Hacer copia de seguridad y restaurar los	
		índices y las base de datos habilitadas	22
		Rastreo de anomalías	23
Parte 1. Guía	1		
		Capítulo 3. Iniciación	25
Capítulo 1. Visión general de DB2 Text		Ejemplo sencillo de preparación de texto para	
Extender	3	la búsqueda	25
DB2 Text Extender en el entorno de		Asistente de ejemplo para la indexación y la	
cliente/servidor DB2	4	búsqueda	27
DB2 Text Extender en el entorno de bases de			
datos particionadas	7		
		Capítulo 4. Planificación para las	
Capítulo 2. Instalación, configuración y		necesidades de la búsqueda.	29
mantenimiento	9	Por qué es necesario indexar los documentos	29
Instalación en sistemas AIX, Sun-Solaris y		Formatos de documentos que se soportan	31
HP-UX	9	Documentos HTML, consideraciones	
Requisitos para la instalación.	9	especiales	32
Paso 1 para AIX: Instalación de los		Documentos XML, consideraciones	
componentes del producto	9	especiales	32
Paso 1 para Sun-Solaris: Instalación de los		Utilización de formatos de documentos no	
componentes del producto	10	soportados	33
Paso 1 para HP-UX: Instalación de los		Idiomas	35
componentes del producto	11	CCSID	35
Paso 2: Establecimiento de la instancia de		EBCDIC	35
DB2 Text Extender	13	ASCII	36
Paso 3: Inicio del servidor DB2 Text		DBCS	37
Extender	13	UNICODE	37
Instalación en Windows 95, Windows 98,		Evitar problemas de página de códigos al	
Windows NT y Windows 2000	14	almacenar y habilitar texto	37
Requisitos del servidor	14	Tipos de búsqueda	40
Requisitos del cliente	14	Búsqueda lingüística	41
Instalación de DB2 Extenders	14	Búsqueda exacta	43
Creación y gestión de instancias	16	Efectuar una búsqueda similar o una	
Preparación de una base de datos de		búsqueda en documentos DBCS	43
ejemplo para la verificación de la		Cambio del tipo de índice	44
instalación	16	Creación de uno o varios índices de texto	
Configuración	17	para una tabla	44
Variables de entorno	17	Cálculo del tamaño de un índice	45
Valores de la configuración del texto	19	Actualización de un índice	46

Mantenimiento de los índices en un entorno de múltiples nodos.	48	Establecimiento de la vía de acceso de la función actual	82
Trabajar con documentos estructurados (soporte de secciones).	48	Búsqueda de texto	82
Secciones de atributos.	49	Realización de una consulta.	83
Archivos planos y documentos HTML	51	Búsqueda y devolución del número de coincidencias encontradas	84
Documentos XML	53	Búsqueda y devolución del rango de un documento de texto encontrado	84
Diccionarios, listas de palabras de parada, listas de abreviaturas y parámetros de idiomas	54	Especificación de los argumentos de búsqueda	85
Modificación de los archivos de palabras de parada y de abreviaturas.	55	Búsqueda de varios términos	85
Capítulo 5. Preparación del texto para la búsqueda	57	Búsqueda con los operadores booleanos AND y OR	85
Preparación anterior a la preparación del texto para la búsqueda	57	Búsqueda de variaciones de un término.	86
Inicio del procesador de línea de mandatos de DB2 Text Extender	59	Búsqueda de fragmentos de un término (enmascaramiento de caracteres)	86
Ayuda del procesador de línea de mandatos	60	Búsqueda de términos que ya contienen un carácter de máscara	87
Conexión a una base de datos	60	Búsqueda de términos en cualquier orden	88
Habilitación de una base de datos.	61	Búsqueda de términos en la misma sentencia o párrafo.	88
Habilitación de una tabla de texto (opcional)	62	Búsqueda de términos en secciones de documentos estructurados	88
Ejemplos	63	Búsqueda de sinónimos de términos	89
Habilitación de una columna de texto	66	Realización de una búsqueda lingüística	89
Se añade una columna de descriptores de contexto	67	Búsqueda con el operador booleano NOT	90
Se establece la información de los documentos	68	Búsqueda similar	91
Se crea una tabla de registro cronológico	68	Respeto de los límites de la palabra-frase	91
Se crea un índice	69	Búsqueda de palabras con sonido similar	92
Ejemplos	69	Búsqueda en tesauruso	92
Habilitación de una columna de texto en una tabla grande	70	Búsqueda de texto libre e híbrida	93
Habilitación de columnas de texto de un tipo de datos no soportado	71	Redefinición de la búsqueda anterior.	93
Habilitación de columnas que contienen el tipo de datos DATALINK	72	Establecimiento y extracción de información en descriptores de contexto	95
Habilitación de archivos de texto externos	73	Establecimiento de información de texto al insertar nuevo texto	95
Ejemplos	74	Extracción de información de descriptores de contexto	96
Finalización de la sesión	76	Cambio de la información de los descriptores de contexto	97
Capítulo 6. Cómo buscar	77	Mejora del rendimiento de la búsqueda	97
Dónde encontrar ejemplos de sintaxis de las funciones de búsqueda	77	Capítulo 7. Administración	99
Tabla de ejemplo DB2TX.SAMPLE	78	Mantenimiento de los índices de texto	99
Descriptores de contexto para archivos externos	81	Actualización de un índice	99
		Actualización de un índice para archivos externos	100
		Cambio de los valores de un índice.	101
		Restablecimiento del estado del índice	102
		Supresión de sucesos de índice	102

Reorganización de un índice	103
Obtención de información útil	104
Visualización de información del estado habilitado	104
Visualización de los valores de las variables de entorno	105
Visualización de los valores de configuración del texto	105
Visualización del estado de un índice	106
Visualización de sucesos de error	108
Visualización de los valores de índice	109
Visualización de los valores de texto para una columna	110
Trabajar con la vista del catálogo de DB2 Text Extender	111
Anulación del proceso de preparación de texto	113
Inhabilitación de una columna de texto	113
Inhabilitación de los archivos de texto	114
Inhabilitación de una tabla de texto	115
Inhabilitación de una base de datos	115

Capítulo 8. Utilización de las funciones

API para buscar y examinar	117
Configuración de la aplicación	117
Enlace de una aplicación UNIX	117
Enlace de una aplicación OS/2 o Windows.	117
Visión general de las funciones API.	118
Búsqueda de texto	119
Obtención de una tabla del resultado (DesGetSearchResultTable)	119
Examen de texto	120
Obtención de información de examen (DesGetBrowseInfo)	120
Inicio de una sesión de examen (DesStartBrowseSession)	121
Apertura de un documento (DesOpenDocument).	121
Obtención de coincidencias (DesGetMatches)	122
Cierre de un documento (DesCloseDocument).	122
Finalización de una sesión de examen (DesEndBrowseSession).	123
Liberalización de la información de examen (DesFreeBrowseInfo)	123

Parte 2. Consulta. 125

Capítulo 9. Mandatos de preparación y administración de texto para el cliente.	127
CHANGE INDEX SETTINGS	129
CHANGE TEXT CONFIGURATION	131
CONNECT	135
DELETE INDEX EVENTS	137
DISABLE DATABASE	138
DISABLE TEXT COLUMN.	139
DISABLE TEXT FILES	140
DISABLE TEXT TABLE	141
ENABLE DATABASE	142
ENABLE TEXT COLUMN	143
ENABLE TEXT FILES	152
ENABLE TEXT TABLE	155
GET ENVIRONMENT	160
GET INDEX SETTINGS.	161
GET INDEX STATUS	163
GET STATUS	164
GET TEXT CONFIGURATION	165
GET TEXT INFO	166
QUIT	167
REORGANIZE INDEX	168
RESET INDEX STATUS	169
UPDATE INDEX	170

Capítulo 10. Mandatos de administración para el servidor 173

TXICRT	174
TXIDROP	176
TXLIST (sólo para UNIX)	177
TXIPCLEAN (sólo para UNIX)	178
TXNADD	179
TXNCHECK	180
TXNDROP	181
TXSAMPLE.	182
TXSTART	183
TXSTATUS	184
TXSTOP	185
TXTHESC	186
TXTHESN	188
TXTRACE	190
TXVERIFY	195
TXWIZARD.	196

Capítulo 11. Funciones de búsqueda . . . 197

Los tipos diferenciados de DB2 Text Extender	197
Un resumen de las funciones de DB2 Text Extender	198
CCSID	200

CONTAINS	201	Proceso lingüístico para examen	266
FILE	202	Etapa 1: Normalización y ampliación de	
FORMAT	203	términos	266
HANDLE	204	Etapa 2: Comparación ampliada	266
HANDLE_LIST	205	Conceptos del tesaurus	268
INIT_TEXT_HANDLE	206	Términos	270
LANGUAGE	208	Relaciones	270
NO_OF_DOCUMENTS	209	Relaciones de tesaurus Ngram	272
NO_OF_MATCHES	210	Creación de un tesaurus	273
RANK	211	Creación de un tesaurus Ngram	278
REFINE	212	Capítulo 16. Archivos de configuración	281
SEARCH_RESULT	213	Archivo de configuración de cliente.	281
Capítulo 12. Sintaxis de los argumentos		Archivo de configuración del servidor	283
de búsqueda	215	Archivo de configuración de análisis de	
Argumento de búsqueda	217	prueba	287
Capítulo 13. Funciones API para la		Opciones comunes	287
búsqueda y examen	229	Capítulo 17. Códigos de retorno	291
DesCloseDocument	231	Capítulo 18. Mensajes	299
DesEndBrowseSession	232	Indicaciones SQL devueltas por funciones de	
DesFreeBrowseInfo	233	DB2 Text Extender	299
DesGetBrowseInfo	234	Mensajes de DB2 Text Extender	303
DesGetMatches	237	Capítulo 19. Códigos de razón de la	
DesGetSearchResultTable	243	aplicación de búsqueda	317
DesOpenDocument	248	Capítulo 20. Códigos de razón de	
DesStartBrowseSession	251	sucesos de error	321
Capítulo 14. Programa API de ejemplo	253	<hr/>	
Capítulo 15. Proceso lingüístico para los		Parte 3. Apéndices	337
índices lingüístico y exacto.	255	Notas	339
Proceso lingüístico al indexar	255	Marcas registradas	342
Análisis básico de texto	256	Glosario	343
Reducción de términos a su formato base		Índice	349
(lematización)	262	Ponerse en contacto con IBM	359
Filtro de palabras de parada	262	Información sobre el producto	359
Descomposición (división de términos			
compuestos)	263		
Proceso lingüístico para la recuperación	264		
Sinónimos	264		
Ampliación de tesaurus	265		
Ampliación de sonido	265		
Enmascaramiento de caracteres y palabras	265		

Acerca de este manual

Este manual describe cómo utilizar DB2 Text Extender para preparar y mantener una base de datos DB2(R) para recuperar datos de texto. También describe cómo se pueden utilizar las funciones SQL proporcionadas por DB2 Text Extender y las interfaces de programas de aplicación (API) para acceder y manipular estos tipos de datos. Mediante la incorporación de las funciones de DB2 Text Extender a las sentencias de SQL del programa y la incorporación de las API, puede crear programas poderosos y versátiles de recuperación de texto.

Las referencias que hace este manual a "DB2" son en realidad a DB2 UDB.

Quién debe utilizar este manual

Este manual está dirigido a los administradores de bases de datos DB2 que están familiarizados con los conceptos, herramientas y técnicas de administración de DB2.

Este manual también está dirigido a los programadores de aplicaciones DB2 que están familiarizados con SQL y con uno o varios lenguajes de programación que se pueden utilizar para los programas de aplicación DB2.

Este manual es para las personas que van a trabajar con DB2 Text Extender. Las personas que trabajan con DB2 Image, Audio y Video Extenders deben consultar la publicación *DB2 Image, Audio, and Video Extenders Administration and Programming*.

Cómo utilizar este manual

Este manual está estructurado de la siguiente manera:

“Parte 1. Guía”

Esta parte proporciona una visión general de DB2 Text Extender, describe cómo instalarlo y configurarlo, y explica las consideraciones que hay que tomar en cuenta en la planificación. También describe cómo preparar y mantener una base de datos DB2 para que pueda buscar texto.

Lea esta parte si no conoce DB2 Text Extender y desea saber cómo utilizar las funciones de DB2 Text Extender y las API para buscar texto.

“Parte 2. Consulta”

Acerca de este manual

Esta parte presenta información de consulta para las funciones de DB2 Text Extender, las API, los mandatos y la información de diagnóstico como, por ejemplo, mensajes y códigos.

Lea esta parte si está familiarizado con los conceptos y las tareas de DB2 Text Extender, pero necesita información acerca de una función específica de DB2 Text Extender, una API, un mandato, un mensaje o un código.

Cómo leer los diagramas de sintaxis

En este manual, la sintaxis se describe utilizando la estructura definida de la siguiente manera:

- Lea los diagramas de sintaxis de izquierda a derecha y de arriba a abajo, siguiendo la trayectoria de la línea. El símbolo ►— indica el principio de una sentencia.

El símbolo —► indica que la sintaxis de la sentencia continúa en la línea siguiente.

El símbolo ►— indica que una sentencia continúa de la línea anterior.

El símbolo —►◄ indica el final de una sentencia.

- Los elementos obligatorios aparecen en la línea horizontal (la trayectoria principal).

►—elemento obligatorio—►◄

- Los elementos opcionales aparecen por debajo de la trayectoria principal.

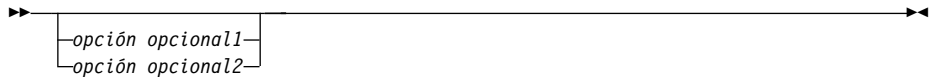
►—
└─elemento opcional┘—►◄

- Si se puede elegir entre dos o más elementos, aparecen en una pila.

Si *debe* elegir uno de los elementos, un elemento de la pila aparece en la trayectoria principal.

►—
└─opción obligatoria1┘
└─opción obligatoria2┘—►◄

Si se puede optar por no seleccionar ningún elemento, toda la pila aparece por debajo de la trayectoria principal.



Una flecha de repetición por encima de la pila indica que puede realizar más de una elección de los elementos de la pila.



- Las palabras clave aparecen en mayúsculas; deben escribirse tal como aparecen. Las variables aparecen en minúsculas (por ejemplo, *vía*fuente). Representan nombres o valores suministrados por el usuario en la sintaxis.
- Si aparecen signos de puntuación, paréntesis, operadores aritméticos u otros símbolos parecidos, debe entrarlos como parte de la sintaxis.

Cómo enviar sus comentarios

La información que envían los usuarios ayuda a IBM a proporcionar información de calidad. Envíe cualquier comentario que desee sobre este manual u otra documentación de DB2 Extenders. Puede utilizar cualquiera de los métodos siguientes para facilitar sus comentarios:

- Envíe sus comentarios desde la Web. Visite el sitio Web:
<http://www.ibm.com/software/data/db2/extenders>

El sitio Web tiene una página de envío de información que puede utilizar para entrar y enviar sus comentarios.

- Envíe sus comentarios por correo electrónico a swsdid@de.ibm.com. Asegúrese de incluir el título del manual, el número de publicación del mismo, la versión del producto y, si procede, la ubicación específica del texto sobre el que envía información (por ejemplo, un número de página o un número de tabla).
- Rellene la Hoja de comentarios del lector que encontrará al final de este manual y remítala por correo, fax o entréguela al representante de IBM. La dirección postal está en la parte posterior de la hoja. El número de fax es +49-(0)7031-16-4892.

Cuando se envía información a IBM, se otorga a IBM un derecho no exclusivo de utilizar o distribuir la información en cualquier manera que encuentre adecuada, sin incurrir en ninguna obligación con el remitente.

Resumen de Cambios

Los cambios en este documento para el Release 7.1 son:

- **Búsqueda en secciones que tienen un tipo especial**
Los tipos soportados son la fecha/hora y entero. Esto permite al usuario buscar documentos que tienen, por ejemplo, una fecha posterior a 12/13/99 en la sección Fecha de publicación.
- **Soporte de enlace de datos**
Los documentos a los que se hace referencia mediante una columna que tiene el tipo de datos Datalink pueden indexarse.
- **Más CCSID soportados**
Los límites de CCSID para el tipo de índice Ngram se han eliminado.
- **Ya no se soporta el tipo de índice dual**
Utilice el tipo de índice lingüístico en su lugar o un tipo de índice exacto normalizado.
- **Sintaxis de search_result cambiada**
La sintaxis de la función search_result se ha cambiado. La sintaxis anterior se sigue soportando por compatibilidad retroactiva.
- **Más idiomas soportados**
La lista de idiomas y CCSID soportados se ha ampliado
- **Las posibilidades de configuración se han ampliado**
Puede configurar el señalizador y personalizar la configuración del servidor, por ejemplo, para reorganizar los umbrales.
- **Nuevo mandato de limpieza del sistema para UNIX**
Hay un nuevo mandato de servidor TXIPCLEAN para plataformas UNIX para facilitar la limpieza del sistema si el servidor DB2 Text Extender no puede iniciarse satisfactoriamente.

Resumen de los cambios

Parte 1. Guía

Capítulo 1. Visión general de DB2 Text Extender

DB2 Text Extender forma parte de la familia de DB2 Extenders. Permite que los programadores incluyan consultas de SQL para documentos de texto en sus aplicaciones.

Los demás Extenders pueden buscar datos de imágenes, de vídeo y de voz y pueden manejar documentos XML y datos espaciales.

DB2 Text Extender añade la posibilidad de recuperar texto completo en las consultas de SQL utilizando las características disponibles en DB2 UDB que permiten al usuario almacenar documentos de texto no estructurados en bases de datos.

DB2 Text Extender ofrece a los usuarios y programadores de aplicaciones DB2 UDB un método rápido, versátil e inteligente para buscar en estos documentos de texto. La eficacia de DB2 Text Extender reside en la posibilidad de buscar en miles de documentos de texto grandes a gran velocidad y no tan sólo encontrar lo que se ha pedido directamente, si no también variaciones de palabras y sinónimos.

La búsqueda no se restringe únicamente a los documentos de texto almacenados en bases de datos DB2 UDB , también se puede buscar en documentos de texto almacenados en archivos.

En el corazón de DB2 Text Extender se encuentra la tecnología de búsqueda lingüística de alto rendimiento de IBM descrita en el “Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto” en la página 255. Permite que las aplicaciones accedan y recuperen documentos de texto de diversas maneras. Las aplicaciones pueden:

- Buscar documentos que contengan un texto específico, los sinónimos de una palabra o una frase, o buscar palabras cercanas en el espacio como, por ejemplo, en la misma oración o párrafo.
- Efectuar búsquedas mediante comodines, utilizando un comodín al principio, en medio o al final para enmascarar palabras y caracteres.
- Buscar documentos en varios idiomas en diversos formatos de documento.
- Efectuar una búsqueda “similar” de palabras que tengan una ortografía similar a la del término de búsqueda. Esto resulta útil para encontrar palabras incluso si se han escrito incorrectamente.
- Realizar una búsqueda de texto libre, en la cual el argumento de búsqueda se exprese en idioma natural.

Visión general

- Buscar palabras que suenen como el término de búsqueda.

Puede integrar la búsqueda de texto en las consultas de datos de gestión. Por ejemplo, puede codificar una consulta de SQL de una aplicación para buscar documentos de texto creados por un autor específico, dentro de un rango de fechas y que contengan una palabra o una frase en particular. Mediante la utilización de la interfaz de programación de DB2 Text Extender, también puede permitir que los usuarios de las aplicaciones examinen los documentos.

Mediante la integración de la búsqueda de texto completo en consultas SELECT de DB2 UDB , dispone de una poderosa función de recuperación. La siguiente sentencia de SQL muestra un ejemplo:

```
SELECT * FROM MiTablaTexto
  WHERE version = '2'
  AND DB2TX.CONTAINS (
      DB2BOOKS_HANDLE,
      "autorización"
      IN SAME PARAGRAPH AS "tabla"
      AND SYNONYM FORM OF "suprimir") = 1
```

DB2TX.CONTAINS es una de las distintas funciones de búsqueda de DB2 Text Extender. DB2BOOKS_HANDLE es el nombre de una columna de descriptores de contexto que hace referencia a la columna DB2BOOKS que contiene los documentos de texto en los que se ha de buscar. El resto de la sentencia es un ejemplo de un argumento de búsqueda que busca la palabra autorización, que aparezca en el mismo párrafo que las palabras tabla, y suprimir o cualquier sinónimo de suprimir.

DB2 Text Extender en el entorno de cliente/servidor DB2

La Figura 1 en la página 5 muestra cómo se integra DB2 Text Extender en el entorno de cliente/servidor DB2.

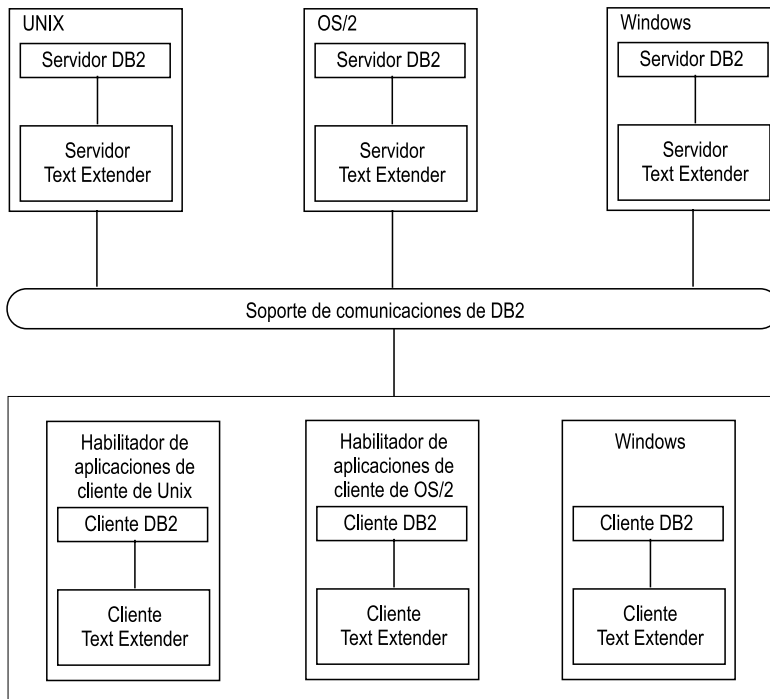


Figura 1. Integración de DB2 Text Extender en el entorno de cliente/servidor DB2

Para ver una lista de los protocolos de soporte de comunicaciones DB2 (TPC/IP o NETBIOS, por ejemplo) para un cliente, consulte la publicación *DB2 Quick Beginnings Guide* para la plataforma adecuada.

La parte principal de DB2 Text Extender se instala en la misma máquina que el servidor DB2. Sólo puede instalarse una instancia de servidor DB2 Text Extender con una instancia de servidor DB2.

La instalación de DB2 Text Extender es flexible y puede constar de:

- Uno o varios servidores DB2 Text Extender en cualquier sistema operativo que aparece en la Figura 1, donde UNIX incluye estaciones de trabajo AIX(R), SUN-Solaris y HP-UX.
- Clientes AIX, SUN-Solaris, Windows NT y Windows 2000 con acceso a uno o varios servidores DB2 Text Extender remotos.
- Clientes AIX que contienen un servidor local y que tienen acceso a servidores remotos.

La Figura 2 en la página 6 muestra una configuración típica de DB2 Text Extender. Para ejecutar DB2 Text Extender desde un cliente, primero debe instalar un cliente DB2 y algunos programas de utilidad de DB2 Text

Visión general

Extender. Estos programas de utilidad constituyen el “cliente” DB2 Text Extender, aunque no se trate de un cliente en el sentido estricto de la palabra. El cliente se comunica con el servidor a través de la conexión cliente DB2.

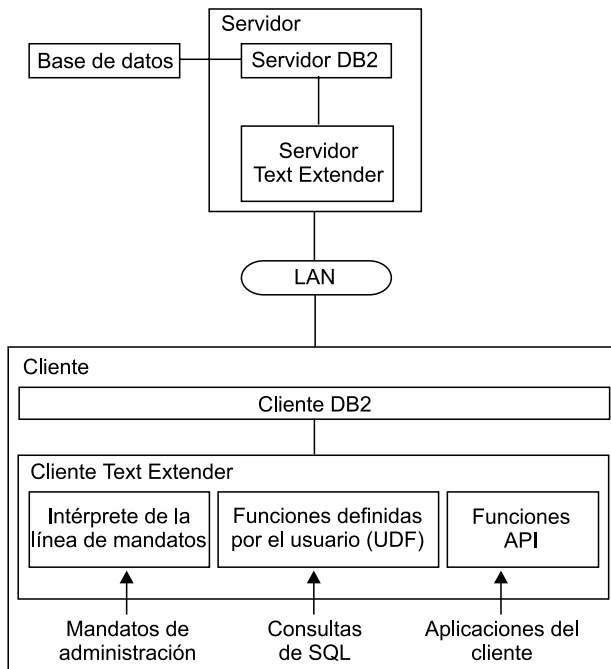


Figura 2. Una configuración de DB2 Text Extender

DB2 Text Extender tiene los componentes principales siguientes:

- **Un intérprete de línea de mandatos.** Hay mandatos disponibles que le permiten preparar el texto en columnas para la búsqueda y mantener índices de texto.
- **Funciones de SQL.** Dispone de funciones que puede incluir en consultas de SQL para la búsqueda en texto y para averiguar, por ejemplo, cuántas veces aparece el término de búsqueda en el texto. La figura muestra las funciones SQL en el cliente para mayor claridad, porque pueden utilizarse como parte de una consulta de SQL. De hecho, forman parte de la instalación del servidor y se ejecutan allí. Sin embargo, estas funciones pueden utilizarse desde cualquier cliente DB2 sin necesidad de instalar el cliente DB2 Text Extender.
- **Una interfaz de programación de aplicaciones (API)** que consta de funciones que pueden llamarse en programas C para la búsqueda en texto y la visualización de los resultados de la búsqueda.

Consejo

Los programas de utilidad del cliente DB2 Text Extender ofrecen funciones de preparación de texto, funciones de administración y la API. (Estas funciones están disponibles en el servidor.) Para utilizar estas funciones, debe instalar el cliente DB2 Text Extender. Si sólo efectúa la preparación y la administración de texto en el servidor DB2 Text Extender, sólo debe instalar los programas de utilidad del cliente DB2 Text Extender en los clientes que utilizan las funciones API.

Si sólo necesita la posibilidad de búsqueda en el cliente utilizando sentencias de SQL de DB2 UDB, no necesita instalar el cliente DB2 Text Extender. Toda la comunicación la maneja DB2 UDB y la aplicación de búsqueda de DB2 Text Extender sólo se ejecuta en el servidor.

DB2 Text Extender en el entorno de bases de datos particionadas

La Figura 3 muestra cómo se integra DB2 Text Extender en el entorno de bases de datos particionadas de DB2.

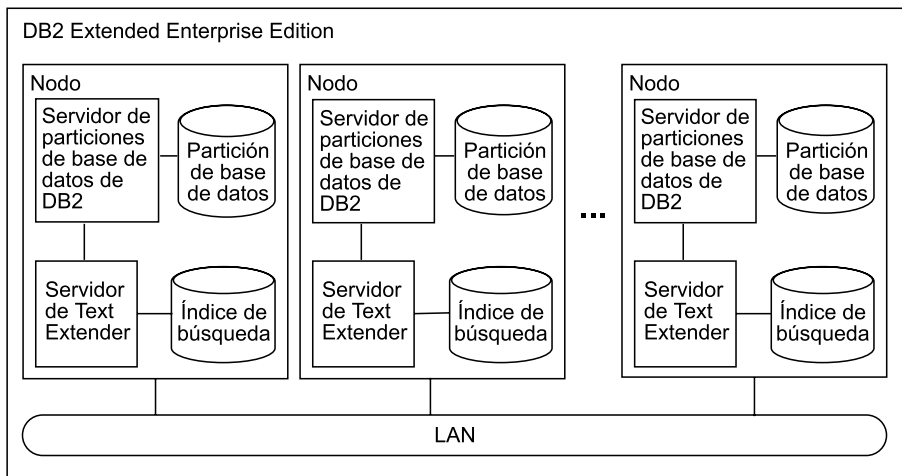


Figura 3. Integración de DB2 Text Extender en el entorno de bases de datos particionadas de DB2

Cada servidor de particiones de base de datos, conocido como *nodo*, puede asignarse a una máquina separada o puede ubicarse en una sola máquina. Los servidores de particiones de bases de datos ubicados en máquinas separadas se conocen como *nodos físicos*. Los que están ubicados en una sola máquina se conocen como *nodos lógicos*. DB2 Text Extender soporta ambas configuraciones.

Visión general

Un *grupo de nodos* es un subconjunto con nombre de uno o varios servidores de particiones de bases de datos. Si se asigna una tabla habilitada por DB2 Text Extender a un grupo de nodos de múltiples particiones, los índices de texto que DB2 Text Extender crea, están ubicados en el mismo nodo que cada partición de tabla correspondiente. Esto asegura que no se tenga que mover los datos entre nodos durante la indexación.

Un servidor DB2 Text Extender se ejecuta en cada partición de un grupo de nodos y gestiona las filas indexadas de la partición correspondiente. Se interactúa con DB2 Text Extender mediante una partición de base de datos. Cada partición de base de datos puede utilizarse para emitir peticiones DB2 Text Extender. Las peticiones se dirigen a todas las particiones de base de datos implicadas.

Capítulo 2. Instalación, configuración y mantenimiento

Instalación en sistemas AIX, Sun-Solaris y HP-UX

El DB2 Text Extender Versión 7 para AIX, Sun-Solaris y HP-UX puede funcionar en un entorno de bases de datos no particionadas. En AIX y Sun-Solaris también puede utilizar el soporte de bases de datos particionadas proporcionado por DB2 Extended Enterprise Edition.

Los pasos de la instalación son los siguientes:

1. Instalación de los componentes del producto en la máquina de destino.
2. Establecimiento de la instancia de DB2 Text Extender.
3. Iniciación del servidor DB2 Text Extender.

Requisitos para la instalación

Antes de instalar DB2 Text Extender, asegúrese de que cumple los siguientes requisitos de instalación para los componentes de servidor y cliente.

Requisitos del servidor

Los componentes del servidor necesitan la instalación de DB2 Versión 7.1 (o superior).

Requisitos del cliente

Los componentes del cliente necesitan la instalación de DB2 Versión 7.1 (o superior) Personal Edition, Habilitador de aplicaciones de cliente de DB2 Versión 7.1 (o superior) o DB2 Versión 7.1 (o superior) Software Developer's Kit.

Paso 1 para AIX: Instalación de los componentes del producto

Durante la instalación, se crea automáticamente un grupo denominado `smadmin`.

1. Inicie la sesión en el cliente de destino o la máquina servidor como usuario `root`.
2. Utilice System Management Interface Tool (SMIT) para ejecutar `installp` desde una interfaz con indicador de mandatos, o entre el mandato `installp`.
3. Entre el mandato `smit install_latest`. Se abre el menú de Software Installation and Maintenance Tool.
4. Escriba la ubicación del paquete en el campo ENTRE dispositivo / directorio para software.
5. Pulse el botón REALIZAR o pulse INTRO. Esto confirma el directorio de instalación.

Instalación en sistemas UNIX

6. En el campo SOFTWARE A INSTALAR, identifique si se han de instalar los componentes de cliente o de servidor.
7. Pulse el botón REALIZAR o pulse INTRO. Se le solicita que confirme los parámetros de instalación. Para confirmar, pulse INTRO.
8. En el campo SOFTWARE A INSTALAR, identifique los diccionarios que desea instalar en el cliente (si va a instalar alguno). Si no especifica ningún diccionario, se instalará el inglés de EE.UU.

La Tabla 1 lista los diccionarios de DB2 Text Extender. Para instalar todos los diccionarios en el cliente, entre db2tx_07_01.dic en el campo SOFTWARE A INSTALAR.

Tabla 1. Diccionarios de DB2 Text Extender

Diccionario	Idioma
db2tx_07_01.dic.De_DE	Alemán
db2tx_07_01.dic.De_CH	Alemán (Suiza)
db2tx_07_01.dic.Ar_AA	Árabe
db2tx_07_01.dic.Da_DK	Danés
db2tx_07_01.dic.Es_ES	Español y catalán
db2tx_07_01.dic.Fi_FI	Finlandés
db2tx_07_01.dic.Fr_FR	Francés
db2tx_07_01.dic.Fr_CA	Francés (Canadá)
db2tx_07_01.dic.Iw_IL	Hebreo
db2tx_07_01.dic.NI_NL	Holandés
db2tx_07_01.dic.En_GB	Inglés (Reino Unido)
db2tx_07_01.dic.En_US	Inglés (EE.UU)
db2tx_07_01.dic.Is_IS	Islandés
db2tx_07_01.dic.It_IT	Italiano
db2tx_07_01.dic.No_NO	Noruego
db2tx_07_01.dic.Pt_PT	Portugués
db2tx_07_01.dic.Ru_RU	Ruso
db2tx_07_01.dic.Sv_SE	Sueco
db2tx_07_01.dic.Th_TH	Tailandés

9. Finalice la sesión.

Paso 1 para Sun-Solaris: Instalación de los componentes del producto

Durante la instalación, se crea automáticamente un grupo denominado smadmin.

1. Inicie la sesión en el cliente de destino o la máquina servidor como usuario root.
2. Utilice tar para extraer el paquete instalable del archivo de paquete de arreglo:

```
tar -xvf IBMdb2tx.tar
```
3. Utilice pkgadd para transferir los archivos del directorio a la máquina de destino. Puede elegir entre dos maneras de hacerlo:
 - Instalar el cliente y el servidor DB2 Text Extender con el diccionario inglés de EE.UU.:

```
pkgadd -d . db2txclnt db2txsrv db2txcom db2txmsgs db2txmsc db2txdius
```
 - Indicar si desea instalar un cliente o un servidor DB2 Text Extender y los diccionarios que desea instalar:

```
./db2tx_install
```
4. Finalice la sesión.

Paso 1 para HP-UX: Instalación de los componentes del producto

Durante la instalación, se crea automáticamente un grupo denominado smadmin.

1. Inicie la sesión en el cliente de destino o la máquina servidor como usuario root.
2. Utilice tar para extraer el paquete instalable del archivo de paquete de arreglo:

```
tar -xvf b2tx.tar
```
3. Ejecute el programa swinstall:

```
swinstall
```

Esto abre las ventanas Selección de software y Especificar fuente.

4. Si es necesario, cambie el Nombre de sistema principal fuente.
5. Entre el directorio tar utilizando IBMdb2tx como valor para el campo Vía de acceso de depósito fuente.
6. En el campo Cambiar vista de software, seleccione Productos.
7. Seleccione Aceptar para volver a la ventana Selección de software, que contiene una lista del software disponible para instalar.
8. Ahora, efectúe una de las acciones siguientes:
 - Para instalar DB2 Text Extender con todos los diccionarios y la documentación en todos los idiomas proporcionados, resalte Producto DB2TX.
 - Para instalar parte de DB2 Text Extender:
 - a. Seleccione Abrir elemento en el menú Acciones.

Instalación en sistemas UNIX

- b. Resalte cada catálogo de archivos que desee instalar y seleccione Elegir marca para instalar en el menú Acciones. Puede seleccionar la instalación del servidor o del cliente DB2 Text Extender:

Servidor Text Extender - DB2TX.db2txsrv
Cliente Text Extender - DB2TX.db2txclnt

- c. Seleccione los diccionarios que se han de instalar (si desea instalar alguno). La Tabla 2 lista los diccionarios de DB2 Text Extender.

Tabla 2. Diccionarios de DB2 Text Extender

Diccionario	Idioma
db2txdide	Aleman
db2txdich	Alemán (Suiza)
db2txdiaa	Árabe
db2txdidk	Danés
db2txdies	Español y catalán
db2txdifi	Finlandés
db2txdifr	Francés
db2txdica	Francés (Canadá)
db2txdiil	Hebreo
db2txdinl	Holandés
db2txdigb	Inglés (Reino Unido)
db2txdius	Inglés (EE.UU)
db2txdiis	Islandés
db2txdiit	Italiano
db2txdino	Noruego
db2txdipt	Portugués
db2txdiru	Ruso
db2txdise	Sueco
db2txdith	Tailandés

9. Seleccione Instalar en el menú Acciones para empezar la instalación del producto y abrir la ventana Análisis de la instalación.
10. Seleccione el botón Aceptar de la ventana Análisis de la instalación cuando el campo Estado visualice el mensaje Listo.
11. Seleccione el botón Sí en la ventana Confirmación para confirmar que desea instalar los productos de software de DB2 para HP-UX.

12. Visualice la ventana Instalación para leer los datos del proceso mientras se instala el software. La instalación finaliza cuando el campo Estado indica Listo y se abre la ventana Nota.

El programa `swinstall` carga el catálogo de archivos y ejecuta los scripts de control para el mismo.

El sistema operativo HP-UX proporciona ayuda detallada para `swinstall`. Obtendrá ayuda escribiendo:

```
man swinstall
```

13. Finalice la sesión.

Paso 2: Establecimiento de la instancia de DB2 Text Extender

1. Asegúrese de que esté activo como usuario root.
2. Para AIX, entre:

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

Para Solaris y HP-UX, entre:

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

3. Si tiene una instancia existente, entre:

```
./txiupdt idinstancia db2instanceid
```

para actualizar la instancia con los nuevos archivos.

De lo contrario, para crear una nueva instancia de DB2 Text Extender, entre:

```
./txicrt idinstancia db2instanceid
```

donde *idinstancia* es el ID de usuario de una instancia de DB2 existente o el ID de usuario que se va a asociar con un ID de instancia de DB2.

La ejecución de `txicrt` crea el directorio `/u/idinstancia/db2tx`. No cree archivos ni directorios adicionales bajo `/u/idinstancia/db2tx`; estos archivos se perderán si suprime la instancia.

Necesita parámetros adicionales si utiliza un entorno de bases de datos particionadas. Para obtener más información, consulte el apartado "TXICRT" en la página 174.

4. Finalice la sesión.

Paso 3: Inicio del servidor DB2 Text Extender

1. Inicie la sesión como *idinstancia*.
2. Detenga y reinicie DB2.
3. Para iniciar el servidor DB2 Text Extender, entre:

```
txstart
```

Instalación en sistemas UNIX

4. Finalice la sesión.

Instalación en Windows 95, Windows 98, Windows NT y Windows 2000

DB2 Text Extender se instala como parte del paquete DB2 Extenders.

Requisitos del servidor

Los componentes del servidor para DB2 Extenders se pueden instalar en un servidor Windows NT o Windows 2000.

Los componentes del servidor que se van a utilizar en un entorno de bases de datos no particionadas necesitan la instalación de DB2 Versión 7.1 (o superior).

Los componentes del servidor que se van a utilizar en un entorno de bases de datos particionadas necesitan la instalación de DB2 Extended Enterprise Edition Versión 7.1 (o superior).

Requisitos del cliente

Los componentes del cliente para DB2 Extenders se pueden instalar en Windows 95, Windows 98, Windows NT o Windows 2000.

Los componentes del cliente necesitan la instalación de DB2 Versión 7.1 (o superior) Personal Edition, Habilitador de aplicaciones de cliente de DB2 Versión 7.1 (o superior) o DB2 Versión 7.1 (o superior) Software Developer's Kit para Windows.

Instalación de DB2 Extenders

Para instalar DB2 Extenders en Windows:

1. Inserte el CD-ROM en la unidad de CD-ROM de la estación de trabajo cliente.
2. Si está habilitada la característica de inicio automático del CD, el Área de ejecución de DB2 Extender se inicia automáticamente. De lo contrario, pulse el botón Inicio en la barra de tareas. Seleccione EJECUTAR. Escriba:
`d:\db2ext\winnt95\setup`

en el campo que se abre, donde "d" es la letra de la unidad de CD-ROM.
3. En el Área de ejecución, puede ver el archivo ReadMe de DB2 Extender. Consulte los requisitos previos a la instalación o instale el producto.
4. Seleccione el botón Instalar situado a la izquierda, para empezar la instalación.
5. En la primera ventana, seleccione la versión de DB2 Extenders que se ha de instalar: Cliente/Servidor para EEE, Cliente/Servidor para no EEE o Sólo cliente.

6. En la ventana siguiente, seleccione la carpeta en la que se instalarán los Extenders.
7. Después, seleccione el tipo de instalación que desea: Típica, Personalizada o Compacta.

La instalación **Típica** incluye los archivos de ejemplo del cliente/servidor DB2 Extenders, el SDK, la documentación en línea en inglés de EE.UU. y el diccionario inglés de EE.UU. de DB2 Text Extender.

La instalación **Compacta** sólo instala el cliente/servidor y el diccionario inglés de EE.UU. de DB2 Text Extender.

La instalación **Personalizada** le permite seleccionar entre los componentes siguientes:

- Seleccione **Servidor DB2 Extenders (EEE)** si está instalando DB2 Extenders en un entorno de bases de datos particionadas (DB2 Extended Enterprise Edition). Esto también instala el componente **Cliente DB2 Extenders**.
- Seleccione **Servidor DB2 Extenders (no EEE)** si está instalando DB2 Extenders en un entorno de bases de datos no particionadas. Esto también instala el componente **Cliente DB2 Extenders**.
- Seleccione **Cliente DB2 Extenders** si está instalando el entorno de cliente DB2 Extenders.
- Seleccione **SDK de DB2 Extenders** si desea instalar los archivos de cabecera y las bibliotecas para desarrollar aplicaciones DB2 Extenders.
- Seleccione **Archivos de ejemplo** si desea instalar los programas de ejemplo y los archivos de soporte de almacenamiento de los ejemplos.
- Seleccione **Documentaciones en línea** si desea instalar las publicaciones de Extender en formato de copia software. Una vez efectuada esta selección, puede pulsar el botón **Cambiar** en la sección **Descripción** para ver las publicaciones específicas del idioma disponibles. Seleccione las que desea instalar, en la lista.
- Seleccione **Diccionarios de Text Extender** si desea instalar el soporte lingüístico para los idiomas seleccionados para DB2 Text Extender. Una vez efectuada la selección, puede pulsar el botón **Cambiar** en la sección **Descripción** para ver los diccionarios específicos del idioma que están disponibles. Seleccione los que desearía instalar, en la lista. El Inglés de EE.UU. siempre es obligatorio.

Si no tiene suficiente espacio de disco para los componentes seleccionados, pulse **Retroceder** para volver a seleccionar el diccionario de instalación. De lo contrario, pulse **Siguiente**.

8. Si está instalando el servidor Extender (EEE) en un entorno de bases de datos particionadas y está en una máquina que es propietaria de la instancia, se le solicita que especifique un rango de puertos TCP/IP. Debe existir el mismo nombre de compartimiento en todos los nodos asociados a la instancia de DB2 actual.

Instalación en sistemas Windows

Si está instalando DB2 Extenders en una máquina que no es propietaria de la instancia, se le solicita que entre el nombre de la máquina que es propietaria de la instancia. Los rangos de puertos TCP/IP y la información de la vía de acceso de datos se obtendrá de la máquina propietaria de la instancia.

9. Pulse el botón **Siguiente** para instalar los componentes que ha seleccionado.

Al iniciarse, el programa de instalación comprueba que la estación de trabajo cumple los requisitos de software de DB2 Extenders.

10. Concluya y reinicie la estación de trabajo.

Creación y gestión de instancias

Puede crear y utilizar múltiples instancias del servidor DB2 Extenders. Debe crear múltiples instancias si ha creado múltiples instancias del servidor DB2 para Windows. Cada instancia del servidor DB2 Extenders se asocia con una instancia del servidor DB2 para Windows.

Se crea una instancia de DB2 Extenders por omisión denominada DB2 al instalar un expansor de DB2 en Windows en un entorno de bases de datos no particionadas.

Se crea una instancia de DB2 Extenders por omisión denominada DB2MPP al instalar DB2 Extenders en Windows en un entorno de bases de datos particionadas. En un entorno de bases de datos particionadas, no se crea ninguna instancia por omisión para DB2 Text Extender. Para crear una instancia, utilice el mandato `txicrt`. Para obtener más información, consulte el apartado “Creación de una instancia de DB2 Text Extender” en la página 21.

Preparación de una base de datos de ejemplo para la verificación de la instalación

DB2 Text Extender ofrece programas de utilidad para preparar una base de datos de ejemplo. Son útiles para preparar rápidamente el texto para probar las posibilidades de búsqueda de DB2 Text Extender inmediatamente después de la instalación y para efectuar una prueba general en cualquier momento posterior.

Estos programas de utilidad están en los directorios siguientes:

- En OS/2(R), Windows NT y Windows 2000: `DMBMMPATH\samples` donde `DMBMMPATH` es la vía de acceso de instalación.
- En AIX, HP-UX y SUN-Solaris: `$HOME/db2tx/samples`

Para preparar una base de datos de ejemplo, en la estación de trabajo servidor o cliente:

1. En el indicador del sistema operativo, ejecute:

```
TXVERIFY nombre-base-datos [id-usuario] [contraseña]
```

Este mandato crea y habilita una base de datos de ejemplo.

2. En el indicador del sistema operativo, ejecute:

```
TXSAMPLE nombre-base de datos [id-usuario] [contraseña]
```

El ID de usuario y la contraseña sólo son obligatorios si se trabaja desde una estación de trabajo cliente. Este mandato hace lo siguiente:

- Conecta a la base de datos especificada

- Crea la tabla `db2tx.sample`

- Importa los documentos en inglés de ejemplo para llenar la tabla

- Habilita la columna de texto de la tabla, con los siguientes tipos de índice:

 - exacto (precise)

 - exacto normalizado (precise normalized)

 - lingüístico (linguistic)

 - lingüístico con secciones habilitadas (linguistic sections-enabled)

 - Ngram

 - Ngram sensible a las mayúsculas y minúsculas (Ngram case-enabled)

- Espera a que se cree el índice de texto.

Configuración

Esta sección describe las variables de entorno de DB2 Text Extender y la información de configuración. Ambas le permiten especificar los valores por omisión para muchos parámetros que DB2 Text Extender necesita.

Variables de entorno

Las variables de entorno establecen los valores por omisión de los parámetros de entorno. Para visualizar los valores actuales de las variables de entorno, utilice el mandato `GET ENVIRONMENT`, descrito en el apartado “Visualización de los valores de las variables de entorno” en la página 105.

DB2INSTANCE	Nombre de instancia de DB2 UDB .
DB2DBDFT	Nombre de base de datos por omisión. El nombre de DB2 UDB base de datos que se asume si no se especifica ningún nombre de base de datos.
DB2TX_INSTOWNER	Nombre de instancia de DB2 Text Extender. Se trata del nombre de inicio de sesión del usuario propietario de la instancia. Es necesario si un cliente DB2 Text Extender se conecta a un servidor UNIX.

Configuración

DB2TX_INSTOWNERHOMEDIR

Directorio inicial del propietario de la instancia (sólo para UNIX).

En un entorno Windows, las variables de entorno de DB2 Text Extender se especifican como variables del sistema en la página de propiedades del sistema. Se añaden a los valores del sistema durante la instalación.

En un entorno OS/2, las variables de entorno de DB2 Text Extender se añaden al archivo CONFIG.SYS durante la instalación.

En un entorno UNIX, las variables de entorno de DB2 Text Extender excepto DB2DBDFT, se especifican en los scripts de shell de perfil siguientes:

db2txcshrc para el shell C

db2txprofile para el shell Korn y el shell Bourne.

Cuando se crea una instancia, el archivo db2txprofile o db2txcshrc se copia en .profile.

Cuando se crea una instancia, db2txprofile se añade a .profile y a db2profile del propietario de la instancia. El script de shell se almacena en SDB2TX_INSTOWNERHOMEDIR/db2tx.

También se necesitan las variables de entorno siguientes:

LANG	Especifica el valor del idioma
NLSPATH	Especifica la vía de acceso para el archivo de mensajes de DB2 Text Extender
DB2ENVLIST	Lista de variables de entorno que DB2 ha pasado a las funciones de DB2 Text Extender
RAHUSER	Especifica el nombre del propietario de la instancia utilizado por el mandato db2all de DB2. Sólo se utiliza en un entorno DB2 EEE.

Si es el propietario de la instancia, estas variables de entorno se establecen automáticamente cuando se crea la instancia. Si no es el propietario de la instancia, incluya el perfil db2tx del propietario de la instancia, en lugar de intentar establecer manualmente estas variables de entorno.

Consejo

Siempre que cambie alguna variable de entorno, para que los cambios entren en vigor, debe detener y reiniciar DB2 y DB2 Text Extender.

Valores de la configuración del texto

Cada base de datos tiene valores de configuración de texto que consisten en:

- Las características del texto
- Las características del índice
- Las características del proceso

Se establecen cuando se habilita la base de datos para que DB2 Text Extender la utilice. El mandato `ENABLE DATABASE` toma los valores que especifica en el mandato o toma los valores iniciales que se describen aquí. Puede visualizar y cambiar estos valores por omisión; consulte los apartados “Visualización de los valores de configuración del texto” en la página 105 y “Cambio de configuración del texto” en la página 20.

Las características del texto

El “Capítulo 4. Planificación para las necesidades de la búsqueda” en la página 29 describe los formatos, los idiomas y los CCSID de documento soportados por DB2 Text Extender. Los valores por omisión para los mismos son obligatorios para varios mandatos.

FORMAT	Valor inicial: TDS
LANGUAGE	Valor inicial: LANGUAGE que se ha establecido para la base de datos
CCSID	Valor inicial: El CCSID que se ha establecido para la base de datos

Las características del índice

DIRECTORY Directorio que se ha de utilizar para almacenar el índice.

Valor inicial para sistemas UNIX:

DB2TX_INSTOWNERHOMEDIR/db2tx/indexes

El valor inicial para OS/2, Windows NT y Windows 2000:

DMBMMPATH\instance\nombre-instancia\db2tx\indexes
donde *DMBMMPATH* es la vía de acceso en la que DB2 Text Extender está instalado. Para clientes Windows, **DIRECTORY** debe establecerse en el mismo nombre que un directorio de la máquina servidor.

INDEXTYPE Tipo de índice que se ha de utilizar. Consulte el apartado “Tipos de búsqueda” en la página 40 para ver una descripción.

Valor inicial: NGRAM

UPDATEFREQ

Frecuencia para la actualización periódica del índice. Consulte el apartado “Actualización de un índice” en la página 46 para ver una descripción.

Configuración

Valor inicial: NONE

Las características del proceso

UPDATEINDEX

Valor para determinar cuándo se produce la primera actualización del índice: inmediatamente durante el paso de habilitación, después según los valores de frecuencia de actualización (NOUPDATE) o como resultado de un mandato UPDATE INDEX explícito.

Valor inicial: UPDATE

COMMITCOUNT

Valor para determinar cuántas sentencias de inserción o de actualización deben realizarse para que DB2 Text Extender emita una sentencia de compromiso DB2 UDB . Consulte el apartado “Habilitación de una columna de texto en una tabla grande” en la página 70.

Valor inicial: 0

Cambio de configuración del texto

Cuando se instala DB2 Text Extender por primera vez, se establecen los valores por omisión para la configuración del texto. Para visualizar los valores actuales de configuración del texto, consulte el apartado “Visualización de los valores de configuración del texto” en la página 105.

Para cambiar la configuración del texto que se ha de utilizar como valor por omisión al crear los índices, entre:

```
db2tx CHANGE TEXT CFG USING valores
```

Ejemplos:

Para cambiar el tipo de índice por omisión y el directorio de índices por omisión para los índices posteriores:

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

Para cambiar la frecuencia de actualización por omisión para los índices de modo que se actualicen a las 12:00 o a las 15:00, de lunes a viernes, si hay un mínimo de 100 documentos de texto en cola:

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

Para detener la actualización periódica de un índice:

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

Configuración y mantenimiento de un servidor DB2 Text Extender

El siguiente paso para DB2 Text Extender recién instalado es configurar el servidor DB2 Text Extender para que los clientes DB2 Text Extender puedan utilizarlo. Normalmente, esto lo hace un administrador de servidor DB2 Text Extender e implica:

1. Crear una instancia de DB2 Text Extender
2. Iniciar un servidor DB2 Text Extender

Otras tareas de mantenimiento de un servidor DB2 Text Extender son:

1. Hacer copia de seguridad y restaurar los índices y las bases de datos habilitadas de DB2 Text Extender
2. Rastrear las anomalías

Consulte el “Capítulo 10. Mandatos de administración para el servidor” en la página 173 para ver la sintaxis del mandato.

Creación de una instancia de DB2 Text Extender

Antes de poder empezar a trabajar con DB2 Text Extender, debe crear una instancia de DB2 Text Extender. Cada instancia que crea ofrece un entorno aislado en el que puede mantener índices, almacenándolos en directorios separados.

Para crear una instancia en un entorno de un solo nodo UNIX, entre:

```
txicrt nombre-instancia nombre-instancia-db2
```

donde *nombre-instancia* es el ID de usuario del propietario de la instancia y *nombre-instancia-db2* es el nombre de la instancia DB2 UDB que se ha de asociar al mismo.

En Windows NT y Windows 2000 se crea una instancia por omisión durante la instalación. Para crear una instancia adicional, entre:

```
txicrt nombre-instancia-db2
```

Para crear instancias en entornos de múltiples nodos, consulte el apartado “TXICRT” en la página 174.

Para desechar una instancia en sistemas UNIX, entre:

```
txidrop nombre-instancia nombre-instancia-db2
```

Para desechar una instancia en Windows NT y Windows 2000, entre:

```
txidrop nombre-instancia-db2
```

Configuración y mantenimiento de un servidor

Para determinar la instancia con la que desea trabajar, establezca la variable de entorno adecuada:

```
set DB2INSTANCE = nombre-instancia
```

después detenga y reinicie DB2 y DB2 Text Extender.

Inicio y detención de un servidor DB2 Text Extender

Para iniciar el servidor DB2 Text Extender, inicie la sesión del ID de usuario propietario de la instancia (sólo en sistemas UNIX) y, a continuación, entre:

```
txstart
```

Para visualizar el estado del servidor DB2 Text Extender, entre:

```
txstatus
```

Para detener el servidor DB2 Text Extender, entre:

```
txstop
```

Consejo para los usuarios de UNIX

Si tiene problemas para iniciar un servidor DB2 Text Extender, intente limpiar los recursos compartidos. Consulte el apartado “TXIPCLEAN (sólo para UNIX)” en la página 178.

Hacer copia de seguridad y restaurar los índices y las base de datos habilitadas

Puede hacer copia de seguridad y restaurar las bases de datos habilitadas y los índices de texto que DB2 Text Extender ha creado.

Para **hacer copia de seguridad**:

1. Averigüe las tablas que DB2 Text Extender ha habilitado. Para hacerlo, entre

```
db2tx GET STATUS
```
2. Averigüe los nombres de los directorios de índices utilizados por la base de datos. Para hacerlo, entre

```
db2tx GET INDEX SETTINGS nombre-tabla
```
3. Detenga el servidor DB2 Text Extender. Para hacerlo, entre

```
TXSTOP
```
4. Haga copia de seguridad de los directorios de índice y sus subdirectorios `index` y `work`.
5. Haga copia de seguridad del archivo `desmastr.dat` que está ubicado en:
 directorio_inicial_propietario_instancia/db2tx/txins000 (UNIX)
 o
 DMBMPATH\db2tx\txins000 (OS/2, Windows NT y Windows 2000)

Configuración y mantenimiento de un servidor

Tenga en cuenta que, para las bases de datos particionadas, hay varios directorios de instancias denominados `txins000`, `txins001`, ..., de modo que sería más conveniente hacer copia de seguridad de todo el directorio `.../db2tx` en lugar de los archivos de los directorios individuales.

6. Reinicie el servidor DB2 Text Extender:

```
TXSTART
```

Para **restaurar**:

1. Detenga el servidor DB2 Text Extender:

```
TXSTOP
```

2. Guarde el archivo `desmastr.dat` existente.
3. Restaure la copia de seguridad del archivo `desmastr.dat`.
4. Restaure las copias de seguridad de los directorios de índices en la misma vía de acceso que antes.
5. Reinicie el servidor DB2 Text Extender:

```
TXSTART
```

Rastreo de anomalías

Si debe informar de un error a un representante de IBM, es posible que le pida que active el rastreo para poder grabar la información en un archivo que pueda utilizarse para localizar el error. Utilice el recurso de rastreo únicamente si se lo indica un representante del Centro de soporte de IBM o el representante de soporte técnico.

La activación del rastreo afecta al rendimiento del sistema, por lo tanto utilícelo solamente cuando se produzcan condiciones de error.

Para activar el rastreo, entre:

```
txtrace on [opciones]
```

La sintaxis y las listas de los sucesos y los componentes se proporcionan en el apartado "TXTRACE" en la página 190. Allí, también se describen otras opciones.

Puede filtrar el rastreo especificando una "máscara" que haga que el rastreo acepte o rechace cada registro de rastreo basándose en su ID. El valor por omisión es rastrearlo todo.

Una máscara está compuesta de cuatro partes separadas por puntos, por ejemplo: `2.2-6.1,3.*` donde:

- 2** indica DB2 UDB DB2 Text Extender.
- 2-6** sólo incluye las entradas con un ID de suceso entre 2 y 6.
- 1,3** sólo incluye los sucesos informados por los componentes 1 y 3.

Configuración y mantenimiento de un servidor

- * incluye todas las funciones de los componentes.

Puede excluir los errores del sistema cuya gravedad sea inferior a una determinada y puede especificar si se han de conservar los primeros o los últimos registros cuando se llena el almacenamiento intermedio de rastreo.

Para reproducir el error y grabar la información de rastreo en binario en un archivo de vuelco, entre:

```
txtrace dump nombarch-vuelco
```

Para producir una versión formateada del archivo de vuelco, entre:

```
txtrace format nombarch-vuelco nombarch-formateado
```

También puede grabar la información de rastreo directamente desde la memoria compartida en un archivo formateado mientras el rastreo está activado:

```
txtrace format > archivo-formateado
```

Después de haber grabado la información de rastreo en un archivo, desactive el rastreo:

```
txtrace off
```

Capítulo 3. Iniciación

Lea este capítulo para familiarizarse con las técnicas básicas de preparación de texto para la búsqueda. Se supone que está trabajando con un sistema DB2 Text Extender en ejecución, uno que se ha instalado y configurado y en el que se ha creado e iniciado una instancia de DB2 Text Extender.

Consejo

Este capítulo sólo describe las técnicas básicas para preparar el texto para la búsqueda. Antes de preparar su propio texto para la búsqueda, lea el apartado “Preparación anterior a la preparación del texto para la búsqueda” en la página 57.

Hay dos métodos para familiarizarse con la preparación del texto para la búsqueda:

- Puede realizar un ejemplo sencillo, entrando los mandatos manualmente o
- Puede ejecutar un asistente de indexación y búsqueda de ejemplo.

Ejemplo sencillo de preparación de texto para la búsqueda

1. **En sistemas Windows, inicie el procesador de línea de mandatos de DB2**

En el indicador del sistema operativo, entre:

```
=>db2 cmd
```

2. **Inicie el procesador de línea de mandatos de DB2 Text Extender**

En sistemas UNIX, en el indicador del sistema operativo, entre:

```
=>db2tx
```

En sistemas Windows, en el indicador de DB2, entre:

```
db2=>db2tx
```

3. **Conéctese a una base de datos**

Elija una base de datos que contenga el texto que desea preparar para la búsqueda. Para conectarse a la base de datos, entre:

```
db2tx=>CONNECT TO mibasedatos
```

4. **Habilite la base de datos para la búsqueda de texto**

Para habilitar la base de datos conectada, entre:

```
db2tx=>ENABLE DATABASE
```

5. **Habilite una tabla de texto para la búsqueda de texto (opcional)**

Debe habilitar una tabla de texto sólo si desea crear un único índice para toda la tabla.

¿Un solo índice o varios?

El apartado “Creación de uno o varios índices de texto para una tabla” en la página 44 explica que puede preparar una tabla para la búsqueda creando un índice de texto para toda la tabla o creando varios índices, uno para cada columna de texto.

- Para crear un índice para toda la tabla, debe ejecutar `ENABLE TEXT TABLE` en este punto para crear un índice vacío y, a continuación, debe ejecutar `ENABLE TEXT COLUMN` (consulte el paso siguiente) varias veces, una para cada columna de texto, para rellenar el índice.
- Alternativamente, para crear un índice separado para cada columna de texto, puede saltarse el paso de crear un índice para toda la tabla, es decir puede saltarse `ENABLE TEXT TABLE` y ejecutar `ENABLE TEXT COLUMN` varias veces para crear y rellenar los índices, uno para cada columna de texto.

Para este ejemplo, no creará un índice para toda la tabla. En su lugar, continúe en el paso siguiente para crear un índice para una columna de texto.

6. Habilite una columna de texto para la búsqueda de texto

Entre el mandato siguiente para permitir que DB2 Text Extender busque en la columna de texto `mycolumn` en la tabla `db2tx.sample` y para asignar el nombre `myhandle` a la columna de descriptores de contexto que este mandato crea.

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

Este mandato crea un índice de texto. Se utilizan los valores por omisión para el tipo de documentos que se están indexando y para las características de índice.

7. Compruebe el estado del índice que está creando

Entre:

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

8. Salga del procesador de línea de mandatos de DB2 Text Extender

Entre:

```
db2tx=>QUIT
```

9. Inicie el procesador de línea de mandatos de DB2

En el indicador del sistema operativo, entre:

```
=>db2
```


10. Búsqueda de texto

Ahora, se puede buscar en los documentos. Pruebe este mandato SELECT que busca todas las apariciones de searchterm en el texto que acaba de indexar:

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

Asistente de ejemplo para la indexación y la búsqueda

Se proporciona un asistente de indexación y de búsqueda como ejemplo de creación de índices en las tablas o las columnas seleccionadas y de búsqueda en esos índices. Este asistente de ejemplo le conduce por los pasos necesarios para posibilitar la búsqueda en las columnas seleccionadas de una tabla utilizando las funciones de búsqueda de DB2 Text Extender.

Para iniciar el asistente de ejemplo, entre:

```
txwizard
```

Puede utilizar el código fuente Java que se proporciona con el asistente de ejemplo para crear su propia aplicación de preparación y búsqueda de texto o aprender cómo se pueden utilizar y llamar las funciones de DB2 Text Extender en un entorno Java.

En AIX, el código fuente Java del asistente se almacena en el directorio

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

En Windows NT y Windows 2000, se almacena en el directorio

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

Capítulo 4. Planificación para las necesidades de la búsqueda

Antes de empezar los pasos descritos en el “Capítulo 5. Preparación del texto para la búsqueda” en la página 57, debe averiguar:

- El formato y la página de códigos que tienen los documentos y el idioma en que están escritos
- Cómo evitar problemas de página de códigos
- La clase de funcionalidad de búsqueda que necesitará
- El espacio de disco que necesitará
- Lo que es un índice de texto y si desea un índice común para una tabla DB2 o un índice separado para cada columna de tabla.

Este capítulo describe porqué y cómo reunir esta información.

Hay varios tipos de índice entre los que puede elegir: lingüístico, exacto y Ngram. La elección del tipo de índice es significativa. Por ejemplo, si elige *lingüístico* como tipo de índice, puede buscar variaciones de palabras y sinónimos del término de búsqueda. El tipo de índice también afecta al rendimiento de la indexación y al tamaño del índice. También puede utilizar las posibilidades de búsqueda de más de un tipo de índice creando varios índices, cada uno de ellos con un tipo de índice diferente, por columna de texto.

Por qué es necesario indexar los documentos

Un sistema rápido de recuperación de información no explora secuencialmente los documentos de texto; tardaría demasiado. En su lugar, opera en un índice de texto creado previamente. Puede decirse que un índice de texto consiste en los términos significativos extraídos de los documentos de texto; cada término se almacena junto con información sobre el documento que lo contiene.

Un índice de texto sólo contiene información importante; las palabras insignificantes como, por ejemplo, “y”, “de” y “que”, no se indexan. (Para índices Ngram se realiza un filtrado sin palabras de parada.) DB2 Text Extender utiliza una lista de estas palabras, conocidas como *palabras de parada* para evitar que se indexen. El sistema de recuperación busca en el índice los términos pedidos para averiguar los documentos de texto que los contienen.

Por qué es necesario indexar los documentos

Consejo

Si necesita modificar la lista de palabras de parada, hágalo sólo una vez y en el momento de la instalación.

Una lista de las palabras de parada por idioma se almacena en un archivo que puede modificar (consulte el apartado “Modificación de los archivos de palabras de parada y de abreviaturas” en la página 55), pero, debido a que sólo hay un archivo para todo el sistema, sólo debe cambiarlo una vez al configurar DB2 Text Extender por primera vez. Si luego cambia el archivo, los índices existentes no reflejarán el cambio.

Digamos, por ejemplo, que algunos documentos contienen el nombre de una revista semanal denominada “Ahora”. Si elimina esta palabra de las palabras de parada, se indexará y podrá encontrarse en búsquedas posteriores. Sin embargo, los índices que haya creado antes de eliminar la palabra de parada no contendrán la palabra “ahora” y una búsqueda de la misma no sería satisfactoria.

Si decide cambiar las palabras de parada y desea que este cambio se refleje en todos los índices, debe volverlos a crear todos.

La indexación es un proceso que consta de dos pasos. El primer paso es registrar en una *tabla de registro cronológico* los documentos de texto que se deben indexar. Esto se realiza automáticamente con los *desencadenantes* de DB2 siempre que se inserta, actualiza o suprime un documento de texto de una columna.

El segundo paso es indexar los documentos de texto listados en la tabla de registro cronológico. Esto puede hacerse periódicamente. Los términos de los documentos que se han insertado o cambiado en la columna se añaden al índice. Los términos de los documentos que se han suprimido de la columna se eliminan del índice.

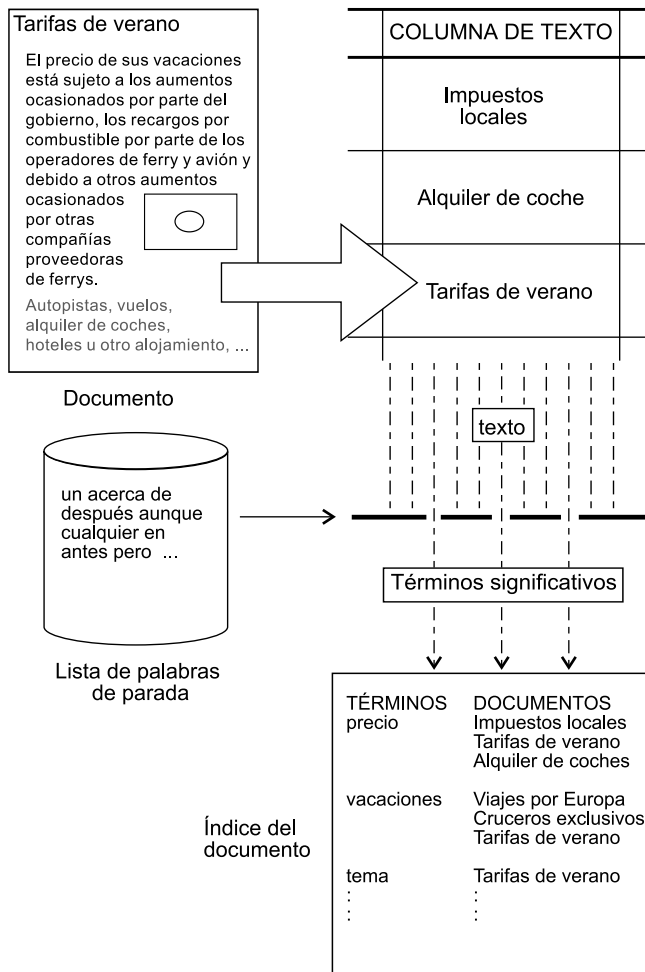


Figura 4. Indexación de los términos significativos solamente

Formatos de documentos que se soportan

DB2 Text Extender necesita conocer el formato (o tipo) de los documentos de texto como, por ejemplo, HTML o ASCII, que tiene intención de buscar. Esta información es necesaria para indexar los documentos de texto.

Los formatos de los documentos de texto que se soportan son:

HTML

Hypertext Markup Language

XML

Extended Markup Language

ASCII_SECTIONS

ASCII estructurado que contiene secciones

TDS

ASCII plano

Formatos de documentos

AMI	AmiPro Architecture Versión 4
FFT	IBM Final Form Text: Arquitectura del contenido del documento
MSWORD	Microsoft Word, Versiones 5.0 y 5.5
RFT	IBM Revisable Form Text: Arquitectura del contenido del documento
RTF	Microsoft Rich Text Format (RTF), Versión 1
WP5	WordPerfect (OS/2 y Windows), Versiones 5.0, 5.1 y 5.2

Documentos HTML, consideraciones especiales

El trato que se da a las diéresis y a los caracteres diacríticos de los documentos HTML depende de la página de códigos del documento:

- Para las páginas de códigos 37, 273, 277, 278, 280, 284, 297, 437, 500, 819, 850, 858, 860, 863, 865, 871, 923, 924 y 1252 se aplica lo siguiente:
 - Se utiliza la notación de entidad para las diéresis y los caracteres especiales, por ejemplo, &auuml; para ä.
 - Sólo son válidos los caracteres que tienen un elemento de código en la página 819 (ASCII) ó 500 (EBCDIC).
 - Durante la indexación, los documentos que contienen caracteres específicos de idioma como, por ejemplo ä, conducen a errores en el reconocimiento de palabras si la página de códigos del documento no es 819 ni 500.
 - Si añade documentos a un índice Ngram, éste debe haberse creado utilizando la página de códigos 819, 500 ó UTF8.
- Para todas las demás páginas de códigos del juego de caracteres de un solo byte, se aplica lo siguiente:
 - Las entidades no se resuelven.
 - Los caracteres especiales deben estar escritos en elementos de códigos específicos del idioma.

Documentos XML, consideraciones especiales

Los documentos XML deben contener una serie de codificación y ésta se utiliza durante el análisis. Se da soporte a las codificaciones siguientes:

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-roeece, 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

Se pasa por alto la página de códigos por omisión para los documentos XML.

Si está añadiendo documentos XML a un índice Ngram, el índice debe haberse creado utilizando la página de códigos UTF8. Si el índice da soporte a las secciones, sólo se utiliza el identificador de sección de nivel de hoja en la indexación, no se da soporte a la resolución de texto de secciones anidadas.

Utilización de formatos de documentos no soportados

Para los formatos de documentos no soportados, especifique un ID numérico. Los valores válidos son de 1 a 100. Este valor se pasa como formato fuente a la salida de usuario que convierte el formato original a TDS.

Si, durante la indexación, hay un documento que no es de ningún tipo soportado, DB2 Text Extender proporciona una salida que graba el documento en un disco y llama a un programa que el usuario proporciona para extraer el texto a uno de los formatos soportados. La salida de usuario debe estar registrada en el archivo de configuración del servidor (DESSRV.INI) y en el archivo de configuración del cliente (DESCL.INI). Actualice la opción USEREXIT en la sección [DOCUMENTFORMAT] con el nombre de la salida de usuario.

Para habilitar la salida de usuario, edite los archivos ASCII siguientes:

Windows NT y Windows 2000:
%DMBMPATH%\instance%\DB2INSTANCE%\db2tx\descl.ini
%DMBMPATH%\instance%\DB2INSTANCE%\db2tx\txinsnn\dessrv.ini

UNIX:
\$DB2TX_INSTOWNERHOMEDIR/db2tx/descl.ini
\$DB2TX_INSTOWNERHOMEDIR/db2tx/txinsnn/dessrv.ini

añadiendo las sentencias siguientes:

```
[DOCUMENTFORMAT]  
USEREXIT=nombre_de_ejecutable
```

Formatos de documentos

donde <nombre_de_ejecutable> es el nombre de la salida de usuario. Puede especificar un nombre de archivo completamente calificado o, si la salida de usuario se almacena en un directorio que está en la sentencia PATH, puede especificar simplemente el nombre de archivo.

Para llamar a la salida de usuario utilice la sintaxis siguiente:

```
<nombre_salida_usuario> -sourcefile <nombrearchivofuente>  
                        -targetfile <nombrearchivodestino>  
                        -sourceccsid <ccsidfuente>  
                        -targetccsid <ccsiddestino>  
                        -sourceformat <formatofuente>  
                        -targetformat <formatodestino>
```

nombrearchivofuente

El archivo que el programa de salida de usuario ha de convertir. El nombre de archivo está completamente calificado y está ubicado en el directorio de trabajo especificado en el perfil de cliente o en la instancia del servidor.

nombrearchivodestino

El archivo que contiene la salida producida por la salida de usuario. Después, DB2 Text Extender utiliza este archivo para el proceso. El nombre de archivo está completamente calificado y apunta al directorio de trabajo especificado en el perfil de cliente o en la instancia de servidor. Las entradas del perfil de cliente se utilizan para la llamada a la API EhwGetMatches y las de la instancia del servidor para la llamada a la API EhwUpdateIndex.

ccsidfuente

La página de códigos del archivo fuente. Se trata de la página de códigos por omisión.

ccsiddestino

La página de códigos que DB2 Text Extender espera. La página de códigos es 850.

formatofuente

El formato del archivo fuente. Se trata del formato por omisión.

formatodestino

El formato del archivo esperado por DB2 Text Extender. Actualmente, sólo se soporta el formato de archivo plano (TDS) o, para índices habilitados para secciones, ASCIISECTION.

La salida de usuario debe poder devolver los valores siguientes:

- 0** La conversión de formato ha sido satisfactoria.
- >0** La conversión de formato no ha sido satisfactoria. Durante la

indexación, los mensajes de error se graban en la tabla de errores del documento. Utilice el mandato `desmsgix` para visualizar los mensajes de error.

Idiomas

DB2 Text Extender también necesita saber el idioma en que está escrito un documento para poder utilizar el diccionario correcto para el proceso lingüístico que se produce. La Tabla 5 en la página 54 muestra una lista de parámetros de los idiomas que puede especificar al habilitar una columna de texto o los documentos externos.

CCSID

Consejo

Antes de especificar un CCSID al habilitar una columna de texto, consulte el apartado “Evitar problemas de página de códigos al almacenar y habilitar texto” en la página 37

Pueden indexarse los documentos si están en uno de los CCSID de la tabla de más abajo.

Nota: DB2 UDB no da soporte a los CCSID 861, 865 y 4946. Para indexar documentos que tienen estos CCSID, almacene los documentos en una columna con un tipo de datos binario (BLOB o FOR BIT DATA).

EBCDIC

37	Inglés de EE.UU. y Canadá
273	Austriaco, alemán
277	Danés, noruego
278	Finlandés, sueco
280	Italiano
284	Español, Latinoamericano
285	Inglés del Reino Unido
297	Francés
420	Árabe
424	Hebreo
437	Inglés de EE.UU.

CCSID

500	Latin-1 internacional
871	Islandés
875	Griego
1025	Ruso
1112	Letón
1122	Estonio
1123	Ucraniano

ASCII

813 AIX, HP, SUN	Griego
819 AIX, HP, SUN	Latin-1
850 AIX, OS/2	Latin-1
855 OS/2	Búlgaro
860 OS/2	Portugués
861 Ver nota	Islandés
862 OS/2	Hebreo
864 OS/2	Árabe
863 OS/2	Canadiense
865 Ver nota	Danés, noruego
866 OS/2	Ruso
869 OS/2	Griego
915 AIX, OS/2, HP	Ruso
916 AIX	Hebreo
921 AIX, OS/2, WIN	Letón
922 AIX, OS/2, WIN	Estonio
1064 AIX	Árabe
1089 AIX, HP	Árabe
1124 AIX	Ucraniano
1125 OS/2	Ucraniano
1129	Vietnamita
1130	Vietnamita
1131	Vietnamita

1250 WIN	Croata, Bieloruso
1251 WIN	Ruso
1252 WIN	Latin-1
1253 WIN	Checo
1255 WIN	Hebreo
1256 WIN	Árabe
1257 WIN	Griego
1258 WIN	Vietnamita

DBCS

932 AIX, OS/2	Japonés, SBCS/DBCS combinados
942 OS/2	Japonés, SBCS/DBCS combinados
943 OS/2, WIN	Japonés, SBCS/DBCS combinados
5039 HP	Japonés, SBCS/DBCS combinados
954 AIX, HP, SUN	Japonés
949 OS/2	Coreano
970 AIX, HP, SUN	Coreano
1363 WIN	Coreano
948 OS/2	Chino (tradicional), SBCS/DBCS combinados
950 AIX, HP, OS/2, SUN, WIN	Chino (tradicional), SBCS/DBCS combinados
964 AIX, HP, SUN	Chino (tradicional), SBCS/DBCS combinados
1381 OS/2, WIN	Chino (simplificado), SBCS/DBCS combinados
1383 AIX, HP, SUN	Chino (simplificado), SBCS/DBCS combinados
1386 AIX, OS/2, WIN	Chino (simplificado), SBCS/DBCS combinados
4946 Ver nota	Latin-1 (CP850)
5039 HP	Japonés

UNICODE

1208	UTF8
13488	UCS2

Evitar problemas de página de códigos al almacenar y habilitar texto

Las áreas siguientes tienen valores de páginas de códigos:

- El entorno de la aplicación activa

- Cada documento
- Cada base de datos DB2
- Cada índice de DB2 Text Extender

Cuando almacena documentos en una columna de base de datos DB2 que tiene un tipo de datos de caracteres como, por ejemplo, VARCHAR y CLOB, DB2 supone que cada documento tiene la misma página de códigos que el entorno de la aplicación activa y convierte el documento desde esa página de códigos a la página de códigos de la base de datos. La página de códigos de la base de datos es ya la misma que la del entorno de la aplicación activa (no tiene lugar ninguna conversión) o es la página de códigos que ha especificado al crear la base de datos y que era diferente a la página de códigos de la aplicación (se efectúa la conversión).

Cuando almacena datos en una base de datos DB2, en una columna que tiene un tipo de datos binario como, por ejemplo, BLOB o FOR BIT DATA, DB2 no convierte los datos y los documentos conservan sus CCSID originales.

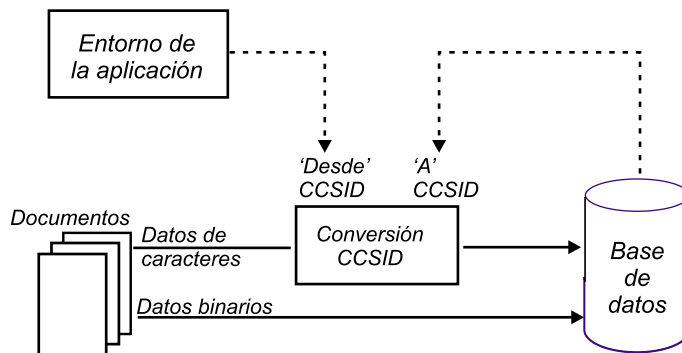


Figura 5. Cómo establece DB2 el CCSID de un documento de base de datos

Cuando habilita una columna de texto para que DB2 Text Extender la utilice, es decir, cuando utiliza el mandato `ENABLE TEXT COLUMN` para crear un índice para la búsqueda, la página de códigos del índice se establece en la página de códigos de la base de datos (por omisión) o en el valor por omisión actual que puede establecerse utilizando el mandato `CHANGE TEXT CONFIGURATION` o en la página de códigos que ha especificado en el mandato `ENABLE TEXT COLUMN`.

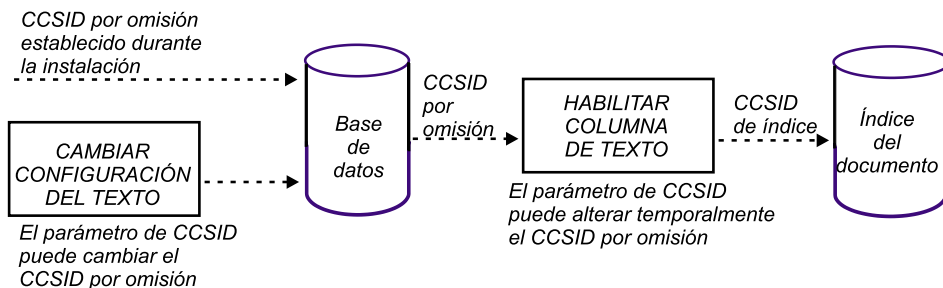


Figura 6. Cómo establece DB2 Text Extender el CCSID de un índice de texto

Durante la búsqueda, se utiliza el CCSID de la base de datos para interpretar el CCSID de la serie de búsqueda.

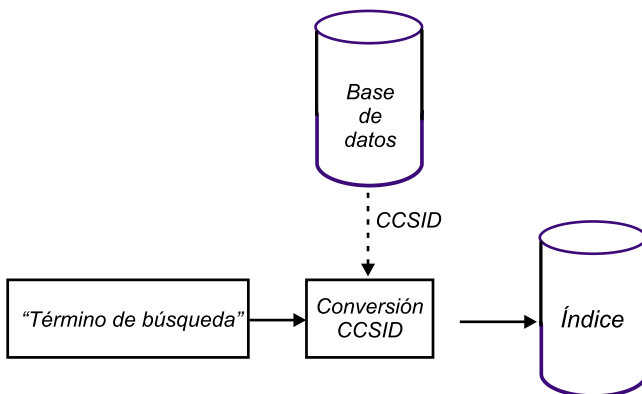


Figura 7. La búsqueda utiliza el CCSID de la base de datos

Así es cómo se comprueban los valores de la página de códigos:

- Para comprobar la página de códigos de la base de datos, utilice el mandato DB2:
db2 get db cfg
- Para comprobar la página de códigos del índice por omisión, utilice el mandato DB2 Text Extender:
db2tx get text cfg

Para evitar problemas de página de códigos es importante coordinar sus valores correctamente:

- **Ejemplo 1:** Entorno 850, Documento 850, Base de datos 850, Índice 850
Correcto. DB2 supone correctamente que el documento tiene la misma página de códigos que el entorno y, puesto que la base de datos tiene la misma página de códigos que el entorno, DB2 no realiza ninguna conversión, almacenando el documento con la página de códigos 850 en la

base de datos. Cuando luego se habilita la columna de texto y se toma el valor de página de códigos por omisión (850, la página de códigos de la base de datos), el documento se indexa correctamente en el índice de página de códigos 850.

- **Ejemplo 2:** Entorno 1252, Documento 1252, Base de datos 850, Índice 850
Correcto. DB2 supone correctamente que el documento tiene la misma página de códigos que el entorno y realiza una conversión de la página de códigos 1252 a 850 al almacenar el documento en la base de datos. Cuando luego habilita la columna de texto y toma el valor de página de códigos por omisión 850 (la página de códigos de la base de datos), el documento se indexa correctamente en el índice de página de códigos 850.
- **Ejemplo 3:** Entorno 1252, Documento 850, Base de datos 850, (NINGÚN índice)
Error. DB2 supone incorrectamente que el documento (página de códigos 850) tiene la misma página de códigos que el entorno (página de códigos 1252) y realiza una conversión incorrecta desde la página de códigos 1252 a 850 al almacenar el documento en la base de datos.
- **Ejemplo 4:** Entorno 1252, Documento 850, Base de datos 1252, Índice 850
Correcto. DB2 supone incorrectamente que el documento tiene la misma página de códigos que el entorno, pero, puesto que la página de códigos del entorno es la misma que la de la base de datos, DB2 no realiza ninguna conversión, almacenando el documento en la base de datos, en la página de códigos 850. Sin embargo, cuando luego habilite la columna de texto, debe especificar la página de códigos de documento 850 para que el documento se indexe correctamente en el índice de página de códigos 850.
- **Ejemplo 5:** Entorno 1252, Documento 1252, Base de datos 850, Índice 1252
Posible error. DB2 supone correctamente que el documento tiene la misma página de códigos 1252 que el entorno y convierte el documento a la página de códigos 850 al almacenarlo en la base de datos. El posible error se produce si luego especifica la página de códigos original del documento 1252 cuando habilita la columna de texto para el índice de página de códigos 850. La acción correcta sería tomar el valor de página de códigos por omisión 850 (la página de códigos de la base de datos).

Tipos de búsqueda

Puede asignar uno de estos tipos de índice y varias opciones a la columna que contiene el texto que se ha de buscar: *lingüístico*, *exacto* y *Ngram*. Debe decidir el tipo de índice que se ha de crear antes de preparar cualquiera de estas columnas para que DB2 Text Extender la utilice. Para ver una descripción más detallada sobre la forma en que cada tipo de índice afecta al proceso lingüístico, consulte el “Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto” en la página 255.

Resumen

- Para la **búsqueda de variaciones lingüísticas de palabras**, utilice un **índice lingüístico**.

Busca variaciones de palabras basándose en la normalización y la reducción a la raíz y en el uso de un diccionario; utiliza el menor espacio de disco.

- Para **realizar búsquedas exactas**, utilice un **índice exacto**.

Busca el término exactamente tal como se ha entrado; la indexación y la búsqueda son más rápidas; utiliza más espacio de disco. Si es NORMALIZED, las búsquedas son sensibles a las mayúsculas y minúsculas.

- Para la **búsqueda de variaciones de caracteres**, utilice un **índice Ngram**.

Encuentra palabras aunque se hayan escrito incorrectamente; si se utiliza CASE_ENABLED para permitir la búsqueda sensible a las mayúsculas y minúsculas, el índice utiliza más espacio y las búsquedas tardan más.

- Para la **búsqueda en documentos DBCS**, utilice un **índice Ngram**.

Es la única elección para los documentos DBCS, pero también puede utilizarse para documentos SBCS de tipo TDS.

DB2 Text Extender ofrece una amplia variedad de opciones de búsqueda, aunque no todas están disponibles para todos los tipos de índice. Consulte la Tabla 8 en la página 223 y la Tabla 9 en la página 223 antes de decidir el tipo de índice que va a utilizar.

Búsqueda lingüística

Para un índice lingüístico, se aplica el proceso lingüístico al analizar el texto de cada documento para la indexación. Esto significa que las palabras se reducen a su formato base antes de almacenarse en un índice; el término “ratones”, por ejemplo, se almacena en el índice como ratón.

Para una consulta en el índice lingüístico, se aplica el mismo proceso lingüístico a los términos de búsqueda antes de buscar en el índice de texto. Por lo tanto, si busca “ratones”, se reduce a su formato base ratón antes de empezar la búsqueda. La Tabla 18 en la página 256 resume la forma en que se extraen los términos para la indexación cuando se utiliza un índice lingüístico.

La ventaja de este tipo de índice es que cualquier variación de un término de búsqueda coincide con cualquier otra variación que aparezca en uno de los documentos de texto indexados. El término de búsqueda ratón coincide con

Tipos de búsqueda

los términos del documento “ratón”, “ratones”, “RATONES” (mayúsculas), etcétera. De manera similar, el término de búsqueda Ratones coincide con los mismos términos del documento.

Este tipo de índice necesita el menor espacio de disco. Sin embargo, la indexación y la búsqueda pueden tardar más que para un índice exacto.

Los tipos de proceso lingüístico disponibles dependen del idioma del documento. A continuación se ofrece una lista de los tipos:

- Separación de palabras y frases.
- Proceso del principio de frase.
- Sin guiones.
- Normalización de los términos a un formato estándar en el que no hay letras mayúsculas y en el que las letras acentuadas como, por ejemplo, “ü” se cambian por un formato sin acentos. Por ejemplo, la palabra alemana “Tür” (puerta) se indexa como tuer.
- La reducción de los términos a su formato base. Por ejemplo, “comprado” se indexa como comprar, “ratones” como ratón.

Consejo

Los fragmentos de palabras (palabras enmascaradas por caracteres comodín) no pueden reducirse a su formato base. Por lo tanto, si busca `estuv%`, no encontrará la palabra “estuvo”, porque se reduce a su formato base en el índice. Para encontrarla, debe buscar `esta%`.

Las variaciones de palabras que no se han escrito correctamente no se pueden reducir al formato base. Por ejemplo, la palabra en alemán `röstete` que se indexa correctamente en su formato base, el verbo `rösten`. Un término de búsqueda `röstete` o `rösteten` se normaliza correctamente al formato base `rösten` y el término se encuentra en el índice. Un término de búsqueda `rostete` no se puede normalizar a `rösten` y el término no se encuentra en el índice.

- Descomposición de palabras, cuando las palabras compuestas como la palabra en alemán “Wetterbericht” (informe del tiempo) se indexan no sólo como `wetterbericht`, sino también como `wetter` y `bericht`.
- Filtro de palabras de parada en que los términos que no son importantes no se indexan. “Un informe de todos los animales” se indexa como `informe` y `animal`.
- Filtro de parte de la alocución, que es similar al filtro de palabras de parada; sólo se indexan nombres, verbos y adjetivos. “Yo conduzco mi coche rápidamente” se indexa como `conducir` y `coche`. Las palabras “Yo” y

“mi” se eliminan como palabras de parada, pero además el filtro de parte de la alocución elimina el adverbio “rápidamente”.

Búsqueda exacta

En un índice exacto, los términos de los documentos de texto se indexan exactamente tal como aparecen en el documento. Por ejemplo, el término de búsqueda ratón puede encontrar “ratón” pero no “ratones” y tampoco “Ratón”; la búsqueda en un índice exacto es sensible a las mayúsculas y minúsculas.

En una consulta, se aplica el mismo proceso a los términos de consulta, que después se comparan con los términos encontrados en el índice. Esto significa que los términos encontrados son exactamente iguales al término de búsqueda. Puede utilizar caracteres de máscara para ampliar la búsqueda; por ejemplo, el término de búsqueda `experiment%` puede encontrar “experimental”, “experimentado”, etcétera.

La Tabla 19 en la página 256 ofrece algunos ejemplos de cómo se extraen los términos del texto del documento para la indexación cuando se utiliza un índice exacto.

La ventaja de este tipo de índice es que la búsqueda es más precisa y la indexación y recuperación más rápidas. Puesto que se indexa cada formato diferente y ortografía de cada término, se necesita más espacio de disco que para un índice lingüístico.

Los procesos lingüísticos utilizados para indexar documentos de texto para un índice exacto son:

- Separación de palabras y frases

- Filtro de palabras de parada.

Efectuar una búsqueda similar o una búsqueda en documentos DBCS

Un índice Ngram analiza el texto analizando los juegos de caracteres. Este análisis no se basa en ningún diccionario.

Si el texto contiene caracteres DBCS, debe utilizar un índice Ngram. Ningún otro tipo de índice soporta los caracteres DBCS.

Este tipo de índice soporta la búsqueda “similar”, lo que significa que puede buscar series de caracteres que sean similares al término de búsqueda especificado. Por ejemplo, una búsqueda de `Extender` encuentra la palabra escrita incorrectamente `Extendrrs`. También puede especificar un grado de similitud obligatorio.

Nota: Incluso aunque utilice la búsqueda similar, los tres primeros caracteres deben coincidir.

Tipos de búsqueda

Para realizar una búsqueda sensible a las mayúsculas y minúsculas en un índice Ngram, no es suficiente especificar la palabra clave `PRECISE FORM OF` en la consulta. Esto es debido a que, normalmente, el índice Ngram no distingue entre los caracteres en mayúsculas y minúsculas indexados. Sin embargo, puede crear un índice Ngram sensible a las mayúsculas y minúsculas, especificando la opción `CASE_ENABLED` cuando se crea el índice. Después, en la consulta, especifique la palabra clave `PRECISE FORM OF`.

Cuando se utiliza la opción `CASE_ENABLED`, el índice necesita más espacio y las búsquedas pueden tardar más.

Consulte el apartado “CCSID” en la página 35 para ver una lista de los CCSID soportados por los índices Ngram. Un índice Ngram da soporte a una lista de CCSID nativos. Para todos los demás CCSID, los datos se correlacionan de este CCSID con UTF8.

Aunque el tipo de índice Ngram se ha diseñado para utilizarse al indexar documentos DBCS, también se puede utilizar para documentos SBCS. Sin embargo, sólo da soporte a documentos TDS.

Tenga en cuenta que no soporta todas las opciones de sintaxis de búsqueda. Consulte el resumen de normas y restricciones en el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Cambio del tipo de índice

Si decide que el tipo de índice que utiliza no es el adecuado, primero suprima el índice inhabilitando la columna de texto o la tabla de texto y, después, vuelva a crear el índice volviendo a habilitar la columna de texto o la tabla de texto.

Creación de uno o varios índices de texto para una tabla

El “Capítulo 5. Preparación del texto para la búsqueda” en la página 57 describe cómo preparar las tablas para que pueda buscar texto en ellas. Sin embargo, antes de realizar esta preparación, debe decidir si va a crear un índice de texto que sea común a todas las columnas de texto indexado de una tabla o varios índices de texto, uno para cada columna de texto indexado. Una tabla que tiene un índice separado para cada columna de texto se conoce como una *tabla de múltiples índices*.

Consejo

Si piensa indexar varios archivos externos (consulte el apartado “Habilitación de archivos de texto externos” en la página 73), la tabla asociada debe ser una tabla de múltiples índices.

La utilización de múltiples índices tiene las ventajas siguientes:

- Creación de un tipo de índice diferente para cada columna de texto
Esto proporciona flexibilidad a las características que se asocian a la columna de texto como, por ejemplo, cuándo se actualiza su índice periódicamente y el directorio en el que se almacena el índice. Consulte el apartado “ENABLE TEXT COLUMN” en la página 143 para ver una descripción de estas características.
- Indexación de columnas en momentos distintos
La indexación es una actividad que puede tardar y consumir recursos. Si tiene una tabla de múltiples índices, puede dilatar esta actividad en el tiempo indexando las columnas en momentos distintos.

Si no necesita la flexibilidad ofrecida por la tabla de múltiples índices, un índice común facilita el mantenimiento de DB2 Text Extender; cuando se habilita una tabla de texto, se establecen los parámetros de indexación que se utilizan como valores por omisión para todas sus columnas de texto. También, si necesita inhabilitar las columnas, puede hacerlo utilizando un solo mandato que inhabilite la tabla de texto.

Cálculo del tamaño de un índice

El espacio de disco que se necesita para un índice depende del tamaño y el tipo de los datos que se han de indexar y el tipo de índice. Los documentos de texto escritos con procesadores de texto necesitan menos espacio porque la mayoría de su contenido se toma con caracteres de control. Como directriz, para un índice lingüístico reserve aproximadamente un espacio de disco de 0,7 por el tamaño de los documentos que se indexan, después multiplique esa cifra por 2 para reservar espacio temporal para reorganizar el índice. Para un índice Ngram necesitará como mínimo el doble de espacio de disco.

Si tiene varios índices grandes, debe almacenarlos en dispositivos de disco separados, especialmente si tiene acceso simultáneo a los índices durante la actualización o búsqueda de índice.

Actualización de un índice

Actualización de un índice

Cuando se añade un documento de texto a una base de datos o cuando se cambia un documento existente en una base de datos, debe indexarse el documento para mantener sincronizado el contenido del índice con el contenido de la base de datos. Cuando se suprime un documento de texto de una base de datos, sus términos deben eliminarse del índice.

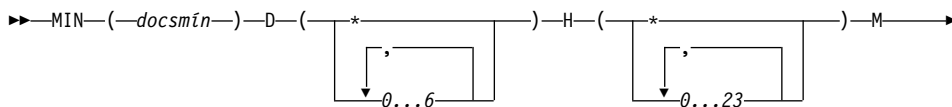
Los desencadenantes almacenan automáticamente la información acerca de los documentos nuevos, cambiados y suprimidos en una tabla de registro cronológico. Los documentos listados en la tabla de registro cronológico se indexan la siguiente vez que tiene lugar una actualización de índice.

El mandato `UPDATE INDEX` le permite actualizar un índice inmediatamente, a demanda. Sin embargo, normalmente se actualiza un índice automáticamente a intervalos especificados en la variable de entorno `DB2TXUPDATEFREQ`. La variable de entorno determina los valores por omisión. Los valores por omisión pueden alterarse temporalmente cuando se crea un índice utilizando los mandatos `ENABLE TEXT COLUMN` o `ENABLE TEXT TABLE`. Se puede cambiar la frecuencia de actualización para un índice existente utilizando el mandato `CHANGE`.

La frecuencia de actualización del índice se especifica mediante el momento en que se ha de realizar una actualización y el número mínimo de documentos de texto que deben estar en cola. Si no hay suficientes documentos en la tabla de registro cronológico en el día y hora dadas, el índice no se actualiza.

Debe planificar la indexación periódica con cuidado; la indexación de documentos de texto es una tarea que tarda y consume recursos. El tiempo empleado depende de muchos factores como, por ejemplo, el número de documentos de texto que se han añadido o cambiado desde la última actualización del índice, el tamaño de los documentos y la potencia del procesador.

Sintaxis





MIN docsmín

El número mínimo de documentos de texto que deben estar en cola antes de que pueda actualizarse el índice.

D El día o los días de la semana en que se actualiza el índice:

* Cada día

0 Domingo

1 Lunes

2 Martes

3 Miércoles

4 Jueves

5 Viernes

6 Sábado

H La hora o las horas del día o los días especificados en que se actualiza el índice:

* Cada hora

0...23 A la hora especificada

M El minuto o los minutos de la hora u las horas especificadas en que se actualiza el índice:

0...59 En el minuto especificado

Ejemplo: $\text{min}(100) \text{ d}(1,2,3,4,5) \text{ h}(12,15) \text{ m}(0)$

Si, a las 12:00 o a las 15:00, de lunes a viernes, hay como mínimo 100 documentos en cola, el índice se actualiza.

Puede combinar varias especificaciones de frecuencia:

Ejemplo: $\text{min}(1) \text{ d}(*) \text{ h}(22) \text{ m}(0) ; \text{min}(100) \text{ d}(1,2,3,4,5) \text{ h}(12,15) \text{ m}(0)$

La actualización del índice está planificada de lunes a viernes a las 12:00 y 15:00 como antes pero, además, cada día a las 22:00 se actualiza el índice incluso si sólo hay un documento de texto en la tabla de registro cronológico.

Mantenimiento de los índices en un entorno de múltiples nodos

Mantenimiento de los índices en un entorno de múltiples nodos

Si la columna de texto que está habilitando pertenece a una tabla que forma parte de un grupo de nodos de múltiples nodos, el directorio de índices que especifique debe estar disponible en todos los nodos físicos. Si utiliza el directorio por omisión especificado en la configuración del texto, asegúrese de que la vía de acceso esté disponible en todos los nodos del grupo de nodos. Si no es conveniente, puede especificar una vía de acceso específica para cada nodo en el mandato `ENABLE TEXT COLUMN`.

Si cambia la configuración de los nodos de un grupo de nodos que contiene una tabla que está habilitada para DB2 Text Extender, debe volver a indexar la tabla. Puede comprobar si se ha cambiado la configuración de nodos utilizando el mandato `TXNCHECK`.

Consejo

Si está utilizando DB2 UDB Extended Enterprise Edition y tiene tablas particionadas entre varios nodos, puede existir una diferencia entre la hora del sistema de los nodos (por ejemplo, cuando uno de los nodos está en horario de verano). Para evitar problemas, asegúrese de que la diferencia entre los valores de hora del sistema de los nodos está dentro del valor de `MAX_TIME_DIFF` de DB2. En Windows NT y Windows 2000, utilice el mandato `NET TIME` para sincronizar los valores de hora del sistema.

Trabajar con documentos estructurados (soporte de secciones)

El soporte de secciones le permite indexar y buscar en secciones específicas de un documento estructurado, por ejemplo, en el título, en el autor o en la descripción. Los documentos pueden estar en formato XML o HTML o en documentos de archivos planos con códigos como HTML. El usuario define los códigos de marcación y sus nombres de sección correspondientes en un *modelo de documento*. El modelo de documento define las secciones de los documentos que se indexan y, por lo tanto, están disponibles para la búsqueda. Los nombres de las secciones son nombres descriptivos utilizados en las consultas en esa sección.

Para que el soporte de secciones esté disponible, debe especificar `INDEXPROPERTY SECTIONS_ENABLED` cuando habilite la columna de texto que contiene los documentos.

Un *archivo de modelos de documentos* lista todos los modelos de documentos definidos para la instancia del servidor. Cuando se crea una instancia del servidor, se crea automáticamente un archivo de modelos de documentos de

Documentos estructurados (soporte de secciones)

ejemplo, DESMODEL.INI, en el subdirectorio de la instancia del servidor. El archivo está en la página de códigos ASCII .

La información de modelo de documento se copia en el directorio de índices. Si cambia el archivo de modelos de documentos para la instancia del servidor después de crear el índice, no afecta al soporte de secciones para el índice creado.

Por ejemplo, una búsqueda de McDaniel en la sección Autor en un índice con soporte de secciones puede ser parecida a la siguiente. La sección, en este caso Autor, siempre lleva el prefijo del nombre de modelo.

```
db2 "select count (*)
     from db2tx.htmltable
     where db2tx.contains(midescriptor,'MODEL mimodelohtml
     SECTION (autor) "Schmidt") = 1
```

Secciones de atributos

Las secciones pueden ser de diferentes tipos. Las secciones de texto plano no tienen ningún tipo. Las secciones con un tipo declarado se conocen como *secciones de atributos*. Los tipos de atributo soportados son:

- DATE
- TIME
- FLOAT
- INTEGER

Puede utilizar rangos de valores para buscar en documentos que contienen secciones de atributos. La sección debe tener el tipo de sección adecuado. En el ejemplo siguiente, la sección ABC debe tener el tipo TIME para poder dar soporte a la consulta:

recuperar todos los documentos que contengan valores entre 20:00:00 y 22:00:00 en la sección ABC

Una sección puede tener más de un valor de atributo. Separe los valores con un blanco o una nueva línea. No puede anidar secciones de atributos.

Formatos de fecha soportados

Puede incluir fechas, por ejemplo, la fecha de creación o la fecha de publicación en la lista de secciones soportadas de un modelo de documento. Se da soporte a los formatos de fecha siguientes:

Documentos estructurados (soporte de secciones)

Tabla 3. Formatos de fechas soportados

Nombre de formato	Formato	Ejemplo
ISO	AAAA-MM-DD	1986-12-25 100-12-25 10123-2-17
	DD-MM-AAAA	01-10-1999 5-8-1990
Estándar de EE.UU de IBM (USA)	MM/DD/AAAA	08/18/1999 5/13/1977
	AAAA/MM/DD	1999/08/18
Estándar europeo de IBM (EUR)	DD.MM.AAAA	20.10.1999 1.5.1980

La búsqueda en un índice con soporte de secciones de todos los documentos publicados en una fecha determinada puede parecerse a la siguiente:

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

Formatos de hora soportados

Se da soporte a los formatos de hora siguientes:

Tabla 4. Formatos de hora soportados

Nombre de formato	Formato	Ejemplo
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

Puede omitir los ceros iniciales para las horas. Los valores fuera de los rangos siguientes no son válidos:

Horas	Minutos	Segundos
0-23	0-59	0-59

Formatos flotantes soportados

Se da soporte a los siguientes formatos:

- Una secuencia de dígitos en el rango de 1.175494351e-38F a 3.402823466e+38F
- Notación científica que utiliza e o E.

Utilice blancos o líneas nuevas para separar los valores flotantes.

Documentos estructurados (soporte de secciones)

Los siguientes ejemplos muestran formatos correctos e incorrectos de valores flotantes:

Formato correcto	Formato incorrecto
1000 1 1.0e3 1.0E3 1E3 .1E4 0.1E4	1 000 1,000 e3

Formatos de enteros soportados

Se da soporte a los enteros en el rango de -2147483648 a 2147483647. Utilice blancos para separar los enteros. Esto significa que la serie “42 234” de una sección de tipo INTEGER se resuelve en los valores enteros “42” y “234” y no en “42234”.

Los ejemplos siguientes muestran formatos correctos e incorrectos para valores de enteros:

Formato correcto	Formato incorrecto
1000	1.000 1,000 1 000
1000000	100.000 100,00 100 000

Nota

No se da soporte a los separadores específicos de idiomas ni a los formatos monetarios.

Archivos planos y documentos HTML

Para archivos planos, las secciones se marcan utilizando códigos parecidos a HTML como, por ejemplo, <title> y <subject>. Un documento con secciones marcadas puede parecer lo siguiente:

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

Un archivo de modelos de documento para archivos planos o documentos HTML puede parecerse al siguiente. Los nombres de modelos y de secciones

Documentos estructurados (soporte de secciones)

sólo pueden contener los caracteres A-Z, a-z y 0-9. Los nombres de modelos son siempre sensibles a las mayúsculas y minúsculas. Los nombres de sección pueden ser sensibles a las mayúsculas y minúsculas o no; se especifica el valor cuando se crea el modelo.

```
;lista de modelos de documentos
;el modelo siempre empieza por 'modelname' y el nombre del modelo
[MODELS]
modelname=ejemplo
modelname=ejemplo2
modelname=ejemplo3

; un 'ejemplo' de definición de modelo de documento
; izquierda - identificador del nombre de sección
; derecha - código del nombre de sección
[ejemplo]
Title = title
Author = author
Subject = subject
Abstract = abstract
Content = content
PublishingDate=date

[ejemplo2]
Title = title
Author = author
Subject = subject
[ejemplo3]
Title = title
Author = author
Abstract = abstract
Docnum = docnum
```

Si un documento contiene una sección marcada que no está definida en el modelo de documento, el contenido de la sección se incluye en la sección definida con anterioridad para la indexación y la búsqueda. Por ejemplo, un documento contiene las secciones de marcación siguientes:

```
<title> IBM Dictionary of Computing
<subject> Computers, Reference, ....
<author> McDaniel, George
<abstract> Contiene la cobertura de información hasta el momento de
sistemas de proceso de información,
productos de comunicación y recursos, personal computers y sistemas de
oficina, así
como todo el rango de productos de hardware y software de IBM.
```

El modelo de documento, book, se define como:

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

Documentos estructurados (soporte de secciones)

La sección <subject> no se incluye en el modelo de documento book. Cuando se indexa el documento, el contenido de la sección del tema se indexa con el contenido de la sección de título. También están disponibles para la búsqueda dentro de la sección de título.

Si especifica una lista de modelos cuando crea el índice, el modelo por omisión es el primero de la lista. Puede cambiar el modelo por omisión utilizando el mandato `desmodix`.

Documentos XML

Para los índices con secciones habilitadas, los documentos XML deben estar estructurados correctamente y contener un elemento raíz. El nombre del elemento raíz debe ser igual a uno de los nombres de modelo definidos y las mayúsculas y minúsculas deben coincidir. La descripción del modelo en el archivo de modelos de documentos debe ser un subconjunto del modelo de documento definido en el archivo de DTD (definición de tipo de documento) para el documento.

La descripción del modelo debe empezar por el elemento raíz. Para cada elemento XML que desee utilizar como sección, debe incluir su jerarquía completa en la descripción del modelo. Si una sección es de tipo fecha, debe ser una hoja del árbol de modelos de documentos. No se da soporte a anidar las secciones de atributos.

La descripción de un modelo para documentos XML puede ser parecida a la siguiente:

```
; lista de modelos de documentos
[MODELS]
modelname = LETTER
; ejemplo de definición de modelo XML
; lado izquierdo = identificador de nombre de sección que codifica toda
;                  la vía de acceso
; lado derecho   = códigos de nombre de sección que especifican el
;                  identificador para cada elemento de la vía de acceso
;                  del árbol hasta el nodo especificado. El delimitador
;                  de códigos es /.
[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
```

Un documento XML puede parecerse al siguiente. También muestra cómo se indexan las secciones que no están definidas en el modelo.

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

Documentos estructurados (soporte de secciones)

```
<LETTER>
<HEADER>Se ha saltado este código en la definición, para este texto
        se añadirá a la sección llamada LETTER
</HEADER>
<DATE>
        01.01.2000  03.02.2000
</DATE>
<ADDRESS>
        El texto se añadirá a la sección llamada LETTER/address.
        <CITY>
                El texto se añadirá a la sección llamada LETTER/address/City.
        </CITY>
</ADDRESS>
<CONTENT>
        El texto se añadirá a la sección llamada LETTER/Content.

        <NOSECTION>El texto se añadirá a la sección llamada LETTER/Content
                porque NOSECTION no está definida.
        </NOSECTION>
        <GREETINGS>
                El texto se añadirá a la sección llamada LETTER/Content/Greetings.
        </GREETINGS>
</CONTENT>
</LETTER>
```

Diccionarios, listas de palabras de parada, listas de abreviaturas y parámetros de idiomas

La Tabla 5 muestra los idiomas soportados y los nombres de los archivos que se proporcionan como diccionarios, las listas de palabras de parada y las listas de abreviaturas. Los archivos de diccionarios están en formato binario y no se pueden cambiar. Los archivos de palabras de parada y los archivos de abreviaturas (si existen) están en formato de archivo plano y pueden cambiarse. Si cambia cualquiera de estos archivos, asegúrese de que utiliza la página de códigos para el idioma.

Esta tabla muestra también el parámetro de idioma que debe especificar al habilitar una columna de texto o documentos externos. Esto indica a DB2 Text Extender en qué idioma están escritos los documentos para que pueda utilizar el diccionario correcto para el proceso lingüístico que se produce.

Tabla 5. Funciones lingüísticas utilizadas para los diversos idiomas

Idioma	Nombre de archivo	Parámetro LANGUAGE	Página de códigos
Árabe	arabic	ARABIC	864
Portugués del Brasil	brazil	BRAZILIAN	850
Francés de Canadá	canadien	CAN_FRENCH	850
Catalán	catala	CATALAN	850

Diccionarios, listas de palabras de parada y listas de abreviaturas

Tabla 5. Funciones lingüísticas utilizadas para los diversos idiomas (continuación)

Idioma	Nombre de archivo	Parámetro LANGUAGE	Página de códigos
Danés	dansk	DANISH	850
Holandés	nederlnd	DUTCH	850
Finlandés	suomi	FINNISH	850
Francés	francais	FRENCH	850
Alemán	deutsch	GERMAN	850
Hebreo	hebrew	HEBREW	862
Islandés	islensk	ICELANDIC	850
Italiano	italiano	ITALIAN	850
Bokmal de Noruega	norbook	BM_NORWEGIAN	850
Nynorsk de Noruega	norntn	NN_NORWEGIAN	850
Portugués	portugal	PORTUGUESE	850
Ruso	russian	RUSSIAN	866
Español	espana	SPANISH	850
Sueco	svensk	SWEDISH	850
Alemán de Suiza	dschweiz	SWISS_GERMAN	850
Tailandés	thai	THAI	850
Inglés del Reino Unido	uk	UK_ENGLISH	850
Inglés de EE.UU	us	US_ENGLISH	850

Los archivos se distinguen por su extensión.

Contenido	Extensión
Diccionario	DIC
Lista de palabras de parada	STW
Lista de abreviaturas	ABR

Modificación de los archivos de palabras de parada y de abreviaturas

Hay un archivo de palabras de parada y un archivo de abreviaturas por idioma. Para comprender las implicaciones de editar estos archivos, consulte el apartado “Por qué es necesario indexar los documentos” en la página 29.

Consejo

Antes de empezar la edición de uno de estos archivos, haga una copia de seguridad.

Los archivos de palabras de parada y de abreviaturas de los sistemas OS/2 y Windows están en:

unidad:\dmb\db2tx\dict

En sistemas AIX, HP-UX y SUN-Solaris, están en:

DB2TX_INSTOWNER /db2tx/dicts

Utilice su propio editor para editar estos archivos. Utilizan el CCSID 850, para asegurar que el CCSID de la aplicación también se establece en 850 antes de empezar.

Elimine las palabras y las abreviaturas que desea indexar. Añada las palabras que no desea que se indexen.

Capítulo 5. Preparación del texto para la búsqueda

El “Capítulo 3. Iniciación” en la página 25 le ayuda a familiarizarse en la preparación del texto para la búsqueda mediante DB2 Text Extender realizando un ejemplo sencillo. Este capítulo describe la preparación del texto para la búsqueda con más detalle y describe todos los aspectos que debe tener en cuenta antes de empezar.

Los pasos para preparar el texto para la búsqueda son:

1. Preparación meticulosa
2. Inicio del procesador de línea de mandatos de DB2 Text Extender
3. Conexión a una base de datos
4. Habilitación de una base de datos para la búsqueda de texto
5. Habilitación de una tabla de texto para la búsqueda de texto (no es necesario si crea un índice por columna de texto)
6. Habilitación de una columna de texto para la búsqueda de texto

Preparación anterior a la preparación del texto para la búsqueda

Consejo

Lea esta sección con cuidado. Lista las opciones que necesita conocer *antes* de preparar el texto para la búsqueda.

- **Para usuarios de UNIX, configure su perfil**
Si no es propietario de la instancia de DB2 Text Extender, consulte el apartado “Configuración” en la página 17 para averiguar cómo configurar el perfil.
- **¿Crear un índice para toda la tabla?**
Debe decidir si desea crear un índice para toda la tabla de texto o un índice separado para cada columna de texto. El apartado “Creación de uno o varios índices de texto para una tabla” en la página 44 le ayudará a tomar esta decisión.
- **Conozca los documentos**
Cuando prepare los documentos para la búsqueda debe especificar sus CCSID, los idiomas y los formatos del texto. Para obtener más información, consulte el apartado “Capítulo 4. Planificación para las necesidades de la búsqueda” en la página 29.
- **Decida el tipo de índices de texto que necesita**

Preparación anterior a la preparación del texto para la búsqueda

El tipo de índice que necesita se determina por la clase de búsquedas que desea realizar (exacta, similar, etcétera) y por si los documentos son SBCS o DBCS. Encontrará más información en el “Capítulo 4. Planificación para las necesidades de la búsqueda” en la página 29.

- **Decida dónde desea almacenar los índices**

Cuando se preparan los documentos para la búsqueda, DB2 Text Extender crea un índice de texto. Debe especificar el directorio en el que desea almacenar el índice. Asegúrese de que habrá suficiente espacio de disco (consulte el apartado “Cálculo del tamaño de un índice” en la página 45).

- **Compruebe el nombre de base de datos por omisión**

El nombre de base de datos por omisión en la variable de entorno DB2DBDFT de DB2 (consulte el apartado “Variables de entorno” en la página 17) es el nombre que DB2 Text Extender utiliza si no especifica ningún nombre de base de datos al preparar el texto para la búsqueda.

- **Establezca la configuración del texto**

La configuración del texto determina los valores por omisión para el CCSID del índice, el idioma y el formato de los documentos, el tipo de índice, la frecuencia de actualización del índice, el nombre de espacio de tabla y el directorio del índice.

Puede alterar temporalmente estos valores cuando prepare el texto para la búsqueda, pero es más conveniente establecer los valores por omisión correctamente de antemano. Los valores iniciales de configuración del texto al instalar DB2 Text Extender se describen en el apartado “Valores de la configuración del texto” en la página 19. Para cambiar los valores de instalación y configurar sus propios valores por omisión, consulte el apartado “CHANGE TEXT CONFIGURATION” en la página 131.

- **Configure el soporte de secciones**

Si debe restringir las búsquedas a una sección en particular de un documento, lea el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48 para informarse de cómo especificar los modelos en el archivo de modelos de documentos.

- **Modifique las listas de palabras de parada y de abreviaturas**

Consulte los apartados “Por qué es necesario indexar los documentos” en la página 29 y “Modificación de los archivos de palabras de parada y de abreviaturas” en la página 55 para comprender el concepto de listas de *palabras de parada* y listas de abreviaturas y decida si va a modificarlas antes de empezar la indexación.

Cuando haya reunido la información y tomado las decisiones descritas en el apartado “Preparación anterior a la preparación del texto para la búsqueda” en la página 57, estará listo para preparar el texto para la búsqueda.

Inicio del procesador de línea de mandatos de DB2 Text Extender

Resumen	
Cuándo	Opcional. Al principio de cada sesión.
Mandato	db2tx
Autorización	Ninguna

Puede entrar mandatos de cliente DB2 Text Extender en el indicador del sistema operativo. Estos son similares a los mandatos DB2, pero en lugar de anteponerles el prefijo db2, debe anteponer db2tx.

Como alternativa a anteponer el prefijo db2tx a todos los mandatos de DB2 Text Extender puede iniciar el procesador de línea de mandatos de DB2 Text Extender. La ventaja de esto es que no se interrumpe la conexión de base de datos después de cada mandato; el usuario permanece conectado. (Sin el procesador de línea de mandatos se volverá a conectar automáticamente cada vez que ejecute un mandato db2tx, pero debe asegurarse de que el nombre de la base de datos a la que desea conectarse esté establecido en la variable de entorno DB2DBDFT.)

1. (Sólo para Windows): Inicie el procesador de línea de mandatos DB2

En el indicador del sistema operativo, entre:

```
=>db2 cmd
```

2. Inicie el procesador de línea de mandatos de DB2 Text Extender

En sistemas UNIX, en el indicador del sistema operativo, entre:

```
=>db2tx
```

En sistemas Windows, en el indicador de DB2, entre:

```
db2=>db2tx
```

Se visualiza el indicador db2tx:

```
db2tx=>
```

y todos los mandatos posteriores se interpretan como mandatos de DB2 Text Extender.

Para abandonar esta modalidad, entre:

```
db2tx=>quit
```

Inicio del procesador de línea de mandatos de DB2 Text Extender

Para Windows 95 y Windows 98, puede utilizar el mandato `db2tx`, tal como se describe arriba, o puede pulsar el botón Inicio, apuntar a Programas, apuntar a Cliente DB2 para Windows 95/98 y, a continuación, pulsar Ventana de mandatos DB2.

Si se salta este paso, puede emitir los mandatos de DB2 Text Extender directamente desde el indicador del sistema operativo anteponiéndoles el prefijo `db2tx`. Este es un ejemplo de mandato emitido desde el indicador del sistema operativo:

```
=>db2tx enable base de datos
```

Consejo

Si trabaja desde un cliente, debe estar en el entorno de mandatos de DB2 Text Extender y utilizar el formato completo de la sentencia de conexión, especificando el ID de usuario y la contraseña.

```
db2tx => connect to nombre-basedatos user id-usuario using contraseña
```

Ayuda del procesador de línea de mandatos

Para visualizar una lista de mandatos, entre:

```
db2tx ?
```

Para visualizar la sintaxis de un mandato individual, entre:

```
db2tx ? mandato
```

Por ejemplo:

```
db2tx ? CHANGE TEXT CONFIGURATION
```

Conexión a una base de datos

Resumen

Cuándo	Opcional. Para iniciar una sesión con una base de datos que no sea la base de datos por omisión especificada en la variable de entorno DB2DBDFT.
Mandato	CONNECT
Autorización	CONNECT en la base de datos

Antes de poder emitir más mandatos en una sesión DB2 Text Extender, debe estar conectado a una base de datos. Si emite un mandato sin estar conectado a una base de datos, DB2 Text Extender le conecta automáticamente a la base

de datos por omisión especificada en la variable de entorno DB2DBDFT. Puede conectarse a una base de datos explícitamente utilizando el mandato de DB2 Text Extender CONNECT TO.

```
db2tx CONNECT TO basedatos-ejemplo
```

Se visualiza lo siguiente (tal como se ve en AIX):

```
>-----Información de conexión de base de datos-----<
Producto base datos = DB2/6000 05.02.0000
Usuario             = nombre-usuario
Alias base de datos = basedatos-ejemplo
```

Nota: Si después emite un mandato de DB2 Text Extender fuera del procesador de mandatos anteponiendo el prefijo db2tx, se pierde la conexión de base de datos actual y se establece una nueva conexión a la base de datos por omisión.

Si está utilizando un cliente OS/2 y es la primera vez que se conecta a una base de datos desde que ha iniciado el sistema, UPM (Gestión de perfiles de usuario) le solicita un ID de usuario y una contraseña. Si ya ha iniciado la sesión, UPM toma el ID de usuario y la contraseña de esta sesión e intenta conectarse a la base de datos remota del servidor.

Habilitación de una base de datos

Resumen

Cuándo	Una vez para cada base de datos que contenga columnas de texto en las que se haya de buscar.
Mandato	db2tx=>ENABLE DATABASE
Autorización	SYSADM o DBADM

Este mandato no toma ningún parámetro. Prepara la base de datos conectada para que DB2 Text Extender la utilice.

Este mandato también declara las funciones de DB2 Text Extender y los tipos diferenciados de DB2 Text Extender para DB2. Estas son las funciones SQL que se utilizan luego para la búsqueda de texto. Se describen en el “Capítulo 11. Funciones de búsqueda” en la página 197. Estas declaraciones se aplican a todas las sesiones posteriores.

Se crea una vista del catálogo, TEXTINDEXES, que hace un seguimiento de las columnas de texto habilitadas. Consulte el apartado “Trabajar con la vista del catálogo de DB2 Text Extender” en la página 111.

Habilitación de una base de datos

Este mandato crea la información de configuración del texto para la base de datos y contiene los valores por omisión para el índice, el texto y las características del proceso. Se describen en el “Valores de la configuración del texto” en la página 19.

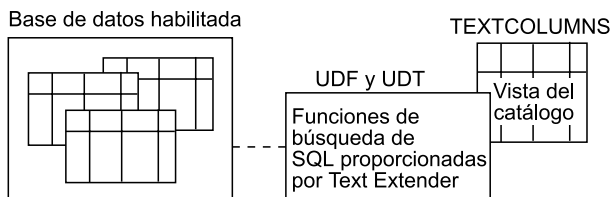


Figura 8. Habilitación de una base de datos

Una vez se habilita una base de datos, permanece habilitada hasta que el usuario la inhabilita. Para invertir los cambios realizados por `ENABLE DATABASE`, consulte el apartado “Inhabilitación de una base de datos” en la página 115.

Consejos

Si se utiliza la variable de entorno `DB2TX_INSTOWNER`, debe establecerse en el nombre del propietario de la instancia antes de que se habilite la base de datos. Esto es especialmente importante para los usuarios de UNIX porque, en UNIX, esta variable se establece por omisión.

Si posteriormente decide desechar una base de datos habilitada, primero debe inhabilitarla para asegurarse de que se eliminan las funciones de DB2 Text Extender declaradas, la vista del catálogo, etcétera.

Habilitación de una tabla de texto (opcional)

Resumen

Cuándo	Opcional. Utilice este mandato únicamente para crear un índice común para todas las columnas de texto de una tabla. Consulte el apartado “Creación de uno o varios índices de texto para una tabla” en la página 44.
Mandato	<code>db2tx=>ENABLE TEXT TABLE ...</code> (vea los ejemplos)
Autorización	<code>ALTER</code> , <code>SELECT</code> , <code>UPDATE</code> en la tabla

Habilitación de una tabla de texto (opcional)

Este paso determina si tiene un índice común para todas las columnas de texto de la tabla o varios índices, es decir, un índice separado para cada columna de texto.

Para tener un índice común, ejecute `ENABLE TEXT TABLE`, después ejecute `ENABLE TEXT COLUMN` para cada columna de texto. Para tener índices separados, sátese `ENABLE TEXT TABLE` y ejecute solamente `ENABLE TEXT COLUMN` para cada columna de texto. Esto se muestra en la Figura 9 y la Figura 10 en la página 66.

Durante este paso, DB2 Text Extender crea un índice de texto vacío que es común a todas las columnas de texto habilitadas con posterioridad. El usuario especifica el tipo de texto, la frecuencia con la que se ha de actualizar el índice y el directorio en el que se ha de almacenar el índice. Los valores por omisión para los parámetros que no especifique se toman de los valores de configuración del texto.

Consejo

Si un valor como, por ejemplo, el tipo de índice, debe ser igual para la mayoría de las columnas de texto, utilice la información de configuración del texto para especificar los valores por omisión. Consulte el apartado “Valores de la configuración del texto” en la página 19.

Este paso también crea una tabla de registro cronológico vacía para registrar los documentos de la tabla que se añaden, cambian o suprimen. Se crean desencadenantes para mantener actualizada la tabla de registro cronológico.

No puede ejecutar `ENABLE TEXT TABLE` para una tabla que ya contiene una columna de texto que se ha habilitado para DB2 Text Extender.

Para suprimir un índice creado por `ENABLE TEXT TABLE`, consulte el apartado “Inhabilitación de una tabla de texto” en la página 115.

Consejo

Si posteriormente decide desechar una tabla de texto habilitada, primero debe inhabilitarla para asegurarse de que se eliminan el índice, la tabla de registro cronológico, etcétera.

Ejemplos

El siguiente ejemplo habilita la tabla de texto `DB2TX.MYTABLE`:

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

Habilitación de una tabla de texto (opcional)

Los valores por omisión para las características del índice se toman de los valores de configuración del texto.

El siguiente ejemplo establece explícitamente las características del índice común que se crea para la tabla. Para un servidor UNIX:

```
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
```

Para un servidor OS/2, Windows NT o Windows 2000:

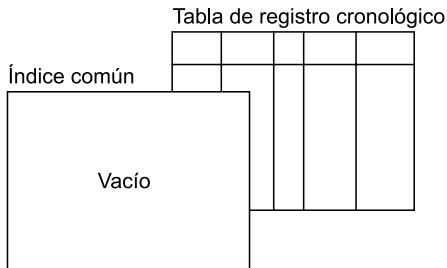
```
db2tx ENABLE TEXT TABLE    db2tx.mytable
                               INDEXTYPE    linguistic
                               UPDATEFREQ    min(100) d(1,2,3,4,5) h(12,15) m(00)
                               DIRECTORY     d:\dmb\db2tx\indexes
```

El ejemplo establece el tipo de índice y el directorio del índice y, después, establece la frecuencia de actualización del índices para que se actualice a las 12:00 o 15:00, de lunes a viernes, si hay un mínimo de 100 documentos de texto en cola.

Habilitación de una tabla de texto (opcional)

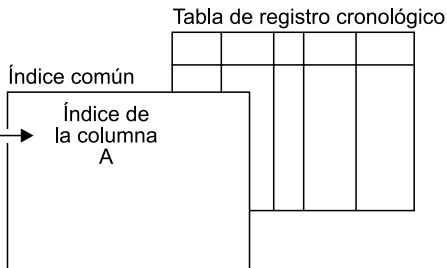
1. Habilitar tabla de texto

	A		B
	Texto		
			Texto



2. Habilitar columna de texto A

	A		B
	Texto		
			Texto



3. Habilitar columna de texto B

	A		B
	Texto		
			Texto

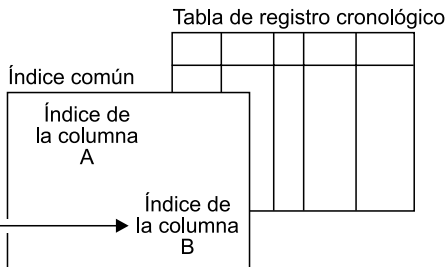
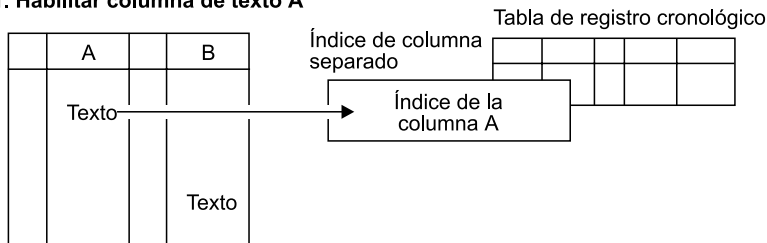


Figura 9. Creación de un índice común para todas las columnas de texto de una tabla

Habilitación de una columna de texto

1. Habilitar columna de texto A



2. Habilitar columna de texto B

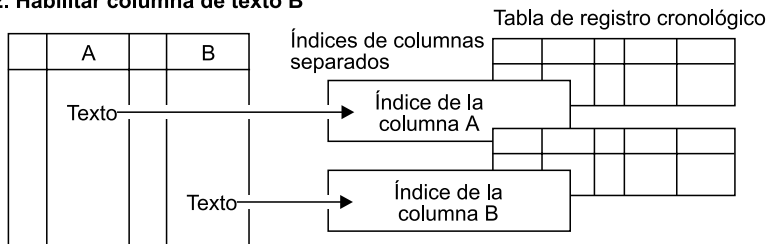


Figura 10. Creación de un índice separado para cada columna de texto

Habilitación de una columna de texto

Resumen

Cuándo	Una vez para cada columna que contenga texto para buscar.
Mandato	db2tx=>ENABLE TEXT COLUMN ... (vea los ejemplos)
Autorización	ALTER, SELECT, UPDATE en la tabla

Consejo

Si un valor como, por ejemplo, el tipo de índice, debe ser igual para la mayoría de las columnas de texto, utilice la información de configuración del texto para especificar los valores por omisión. Consulte el apartado “Valores de la configuración del texto” en la página 19.

Para anular los cambios realizados por ENABLE TEXT COLUMN, utilice el mandato DISABLE TEXT COLUMN. Para inhabilitar todas las columnas de texto habilitadas de una tabla, utilice el mandato DISABLE TEXT TABLE.

Habilitación de una columna de texto

Cuando habilite una columna de texto, se añade una columna de descriptores de contexto a la tabla, se establece la información del documento (formato, idioma, CCSID), se crea una tabla de registro cronológico y se crea un índice.

Se añade una columna de descriptores de contexto

Durante este paso, DB2 Text Extender añade una columna de descriptores de contexto VARCHAR de 60 bytes a la tabla – una columna que contiene los descriptores de contexto asociados a la columna de texto que se habilita. Los descriptores de contexto contienen información acerca del texto de la columna de texto asociada y de los archivos externos asociados. Esta información incluye un ID exclusivo de documento, el idioma, el formato y el CCSID del documento y el nombre de índice.

Consejo

Esta versión de DB2 Text Extender no puede utilizar índices creados con un release anterior. Debe inhabilitar y volver a crear todas las bases de datos y las tablas habilitadas para DB2 Text Extender.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
Datos	Datos	Datos	Datos	Texto

Figura 11. Estructura de la tabla DB2TX.MYTABLE—antes de la habilitación

La columna que contiene los bloques de texto es COMMENT. Antes de poder buscar en el texto de esta columna, debe preparar la base de datos y la columna COMMENT para que DB2 Text Extender las utilice.

Después de este paso de preparación, la tabla DB2TX.MYTABLE contiene una columna adicional para los descriptores de contexto.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT	COMMENTHANDLE
Datos	Datos	Datos	Datos	Texto	Descriptores de contexto

Figura 12. Estructura de la tabla DB2TX.MYTABLE—después de la habilitación

Nota: Cuando posteriormente busque texto, especifique la columna de descriptores de contexto, no la columna de texto, como la columna en que se ha de buscar.

Habilitación de una columna de texto

Se establece la información de los documentos

Especifique el tipo de documentos de texto que almacena normalmente en esta columna de texto: su formato (por ejemplo, ASCII), su idioma y su CCSID. Se pueden especificar valores por omisión para esta información en los valores de configuración del texto. Consulte el apartado “Valores de la configuración del texto” en la página 19.

Se crea una tabla de registro cronológico

Durante este paso, se crea una tabla de registro cronológico y una vista llamada LOGIXnnnnnn, donde *IXnnnnnn* es el nombre de índice (disponible en la vista del catálogo). Si se especifica un espacio de tabla por omisión en la configuración del texto, la tabla de registro cronológico se almacena allí; de lo contrario, se almacena en el espacio de tabla por omisión del sistema DB2. Para optimizar el rendimiento y la utilización del espacio de disco, puede especificar un espacio de tabla diferente para utilizarse para las tablas de registro cronológico.

También se crean desencadenantes que añaden información a la tabla de registro cronológico siempre que se añade o cambia un documento de la columna. Esta información hace que estos documentos se indexen la próxima vez que tenga lugar una indexación.

Si se añaden o cambian archivos externos, estos desencadenantes no se dan cuenta de los cambios. En estos casos, para que los desencadenantes añadan la información a la tabla de registro cronológico, utilice la sentencia UPDATE, tal como se muestra en el ejemplo del apartado “Actualización de un índice para archivos externos” en la página 100.

Si se producen errores durante la indexación como, por ejemplo, cuando no se ha podido encontrar un documento en cola para indexación, los llamados *sucesos de error* se añaden a la tabla de registro cronológico y pueden visualizarse de la forma que describe el apartado “Visualización de sucesos de error” en la página 108.

Consejo

Si se queda sin espacio de registro cronológico en este paso, consulte el apartado “Habilitación de una columna de texto en una tabla grande” en la página 70 para ver las posibles soluciones.

En bases de datos particionadas, cada tabla se asigna a un espacio de tabla y a un grupo de nodos. Es importante que la tabla de registro cronológico se asigne a un espacio de tabla que pertenezca al mismo grupo de nodos que la tabla de usuario habilitada. DB2 Text Extender lo comprueba durante el mandato ENABLE.

Se crea un índice

Si piensa tener un índice separado para cada columna de texto, es decir, se ha saltado el paso `ENABLE TEXT TABLE`, DB2 Text Extender crea un índice separado para la columna de texto durante este paso. El usuario especifica el tipo de texto, la frecuencia con la que se ha de actualizar el índice y el directorio en el que se ha de almacenar el índice. Por otro lado, si piensa tener un índice para toda la tabla, es que ya ha ejecutado `ENABLE TEXT TABLE` y ha especificado los parámetros de índice; se pasan por alto si los repite ahora.

Utilice la palabra clave `UPDATEINDEX` para determinar si la indexación de los documentos de texto de la columna de texto especificada empieza inmediatamente o para cuándo está planificada la indexación periódica. Si no utiliza esta palabra clave, se toma el valor especificado en los valores de configuración del texto.

Creación de índices de varios tipos para una columna de texto. Puede crear más de un índice para una columna de texto. Esto es útil si desea permitir, por ejemplo, la búsqueda lingüística y similar en la misma columna de texto, asociando la columna a tipos de índices diferentes como, por ejemplo, los índices lingüísticos y Ngram. Esto se hace ejecutando `ENABLE TEXT COLUMN` de nuevo, especificando no sólo el tipo de índice adicional que se ha de crear, sino también un nombre exclusivo de columna de descriptores de contexto.

Consejo

Si está utilizando una base de datos particionada y decide cambiar la configuración de un grupo de nodos e iniciar una redistribución de las filas de la tabla, debe volver a crear los índices de texto y las tablas de registro cronológico, es decir, debe inhabilitar y volver a habilitar todas las tablas habilitadas para DB2 Text Extender de ese grupo de nodos.

Ejemplos

El ejemplo siguiente habilita la columna de texto `COMMENT` de la tabla `DB2TX.MYTABLE` y asigna el nombre `COMMENTHANDLE` a la columna de descriptores de contexto que se crea:

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

Los valores por omisión para la información de texto y para las características de los índices se toman de los valores de configuración del texto.

El ejemplo siguiente establece explícitamente los valores para el tipo de documentos que están en la columna `COMMENT`. Los valores por omisión para las características de índice se toman de los valores de configuración del texto.

Habilitación de una columna de texto

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

El siguiente ejemplo establece explícitamente los valores para las características del índice que se crea para la columna COMMENT. El ejemplo establece el tipo de índice, el directorio del índice y la frecuencia de actualización del índice para que se actualice a las 12:00 o 15:00, de lunes a viernes, si hay un mínimo de 100 documentos de texto en cola. Los valores por omisión para la información del texto se toman de los valores de configuración del texto.

Para un servidor UNIX:

```
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
```

Para un servidor OS/2, Windows NT o Windows 2000:

```
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
```

Habilitación de una columna de texto en una tabla grande

Si está trabajando con una tabla que tiene una gran longitud de fila, recuerde que la habilitación de una columna de texto añade una columna de descriptores de contexto del tipo DB2TEXTH (VARCHAR 60). Esto podría ser importante si la tabla se acerca a la longitud máxima de fila, tal como la determina DB2.

También, cuando habilite una columna de texto en una tabla grande, utilice el programa de utilidad REORGANIZE de DB2 UDB para comprobar si es necesario reorganizar la tabla. Cuando se habilita una tabla grande por primera vez, los pasos siguientes agilizan la indexación:

1. Habilite la tabla utilizando la opción NOUPDATE. Esto crea los descriptores de contexto, pero no indexa los documentos.
2. Reorganice la tabla utilizando el programa de utilidad REORGANIZE de DB2 UDB .
3. Cree el índice ejecutando UPDATE INDEX.

Habilitación de una columna de texto

Cuando se habilita una columna de texto o archivos externos, DB2 Text Extender añade una columna de descriptores de contexto a la tabla e inicializa los valores de los descriptores de contexto, lo que provoca que se graben entradas en el registro cronológico de DB2 UDB . Si se ha de grabar un número de entradas de registro cronológico excepcionalmente grande, DB2 UDB puede quedarse sin espacio de registro cronológico.

Hay dos maneras de manejar esta situación; la primera es mejor por razones de rendimiento:

- Incremente el espacio de registro cronológico disponible utilizando el mandato UPDATE DB CFG de DB2 UDB para modificar los parámetros de configuración de base de datos para LOGPRIMARY, LOGSECOND y LOGFILSIZ. Los valores siguientes se toman de la experiencia; es posible que deba cambiarlos para adecuarlos a su instalación.

```
LOGSECOND    50
```

Asegúrese que la suma de LOGPRIMARY y LOGSECOND no sea superior a 128. También debe incrementar el tamaño de almacenamiento dinámico de la aplicación.

```
APPLHEAPSZ   512
```

- Fuerce a que DB2 UDB ejecute un COMMIT intermedio utilizando el parámetro de configuración COMMITCOUNT descrito en el apartado “Configuración” en la página 17. El valor que especifique indica después de cuántas sentencias de inserción o de actualización emite DB2 Text Extender una sentencia de compromiso DB2 UDB . Esto reduce el tamaño necesario para las tablas de registro cronológico, aunque también incrementa el tiempo necesario para el paso de habilitación.

Habilitación de columnas de texto de un tipo de datos no soportado

Las columnas de texto deben ser CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR o CLOB para que DB2 Text Extender las habilite. Si los documentos están en una columna de tipo diferente como, por ejemplo, un tipo diferenciado definido por el usuario (UDT), debe proporcionar una función que tome el tipo del usuario como entrada y proporcione un tipo de salida CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR o CLOB.

Utilice la palabra clave FUNCTION en ENABLE TEXT COLUMN para especificar el nombre de esta función.

Ejemplo: Ha intentado almacenar texto comprimido en una tabla.

1. Cree un tipo diferenciado definido por el usuario (UDT) para el texto:

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

2. Cree una tabla e inserte el texto en ella:

Habilitación de una columna de texto

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

Para habilitar la columna de texto para que DB2 Text Extender la utilice:

1. Cree una función definida por el usuario (UDF) denominada, por ejemplo, UNCOMPRESS, que reciba un valor de tipo COMPRESSED_TEXT y devuelva el texto correspondiente descomprimido como, por ejemplo, un valor CLOB(10M).
2. Habilite la columna de texto utilizando la palabra clave FUNCTION para identificar la UDF UNCOMPRESS:

```
db2tx ENABLE TEXT COLUMN MYTABLE texto
        FUNCTION uncompress
        HANDLE descriptor
...
```

Habilitación de columnas que contienen el tipo de datos DATALINK

DB2 Text Extender soporta el tipo de datos DATALINK. La tecnología Data Link de DB2 proporciona una integridad de referencia entre los datos almacenados en DB2 y los datos de los archivos externos. Lo hace proporcionando un nuevo tipo de datos DATALINK que contiene referencias a un objeto almacenado externamente a la base de datos. Un valor de DATALINK almacenado en una tabla DB2 codifica el nombre del servidor Data Links que contiene el archivo y el nombre del archivo en términos de una Uniform Resource Location (URL).

Si los datos DATALINK están en archivos externos en lugar de en una tabla DB2, siga la descripción del apartado “Habilitación de archivos de texto externos” en la página 73.

Antes de habilitar una columna de texto que contiene el tipo de datos DATALINK, es importante ejecutar el ejemplo de verificación de la instalación (consulte el apartado “Preparación de una base de datos de ejemplo para la verificación de la instalación” en la página 16).

Ejemplo AIX: La sentencia siguiente muestra cómo se inserta un valor de enlace de datos en la columna de texto utilizando el tipo de datos DATALINK en AIX:

```
db2 "insert into mytable values ('Peter',
                                dlvalue ('http://d1fs.boeblingen.ibm.com/test/documents/letter1',
                                           'URL', 'se trata de una carta de Peter'))"
```

En este ejemplo, *d1fs.boeblingen.ibm.com* es el nombre de servidor de archivos, *test* es el nombre del sistema de archivos y *documents/letter1* es la vía de acceso y el nombre de archivo.

Habilitación de una columna de texto

Asegúrese de que el sistema de archivos esté montado en el servidor DB2 con el mismo nombre especificado en la URL.

Ejemplo de Windows NT o Windows 2000: La sentencia siguiente muestra cómo se inserta el valor de enlace de datos en la columna de texto utilizando el tipo de datos DATALINK en Windows:

```
db2 "insert into mytable ('Peter',  
                           dvalue ('unc:\\dfs.boeblingen.ibm.com\cdrive\documents\letter1',  
                                   'URL', 'se trata de una carta de Peter'))"
```

En este ejemplo, *cdrive* es el nombre de compartición de la unidad en la que está almacenado el archivo.

Si surgen problemas al indexar los documentos almacenados en una columna de enlace de datos, asegúrese de que puede acceder a los documentos desde el servidor DB2 utilizando las API proporcionadas por DB2.

Habilitación de archivos de texto externos

Esta opción sólo se proporciona para mantener la compatibilidad con los releases anteriores. El método preferido para la indexación de archivos no almacenados en las tablas DB2 es utilizar una columna que tenga el tipo de datos DATALINK. Consulte el apartado “Habilitación de columnas que contienen el tipo de datos DATALINK” en la página 72.

Resumen

Cuándo	Opcional. Una vez para cada tabla asociada a los archivos externos en los que se ha de buscar.
Mandato	db2tx=>ENABLE TEXT FILES ... (vea los ejemplos)
Autorización	ALTER, SELECT, UPDATE en la tabla

DB2 Text Extender no sólo puede buscar en el texto almacenado en tablas DB2 UDB, sino que también puede buscar en documentos de texto almacenados en archivos. Este paso de preparación es necesario si piensa buscar texto en archivos externos. La tabla asociada a los archivos de texto externos no debe haberse habilitado por el mandato ENABLE TEXT TABLE.

Se crea un índice y una tabla de registro cronológico y se establece la información de documentos, tal como se describe en el apartado “Habilitación de una columna de texto” en la página 66.

Habilitación de archivos de texto externos

Consejos

1. Si se queda sin espacio de registro cronológico en este paso, consulte el apartado “Habilitación de una columna de texto en una tabla grande” en la página 70 para ver las posibles soluciones.
2. Si la tabla que está habilitando está particionada y utiliza un grupo de nodos que tiene múltiples nodos físicos, asegúrese de que los archivos externos a los que está haciendo referencia en las columnas de la tabla están ubicados en el nodo en el que reside la partición de la tabla respectiva.

Una columna de descriptores de contexto del tipo DB2TEXTFH se añade a una tabla DB2 UDB existente. La columna de descriptores de contexto contendrá las referencias a archivos externos, cada descriptor de contexto contendrá información de índices y de documentos (CCSID, formato e idioma).

Consulte el apartado “Descriptores de contexto para archivos externos” en la página 81 para ver una descripción.

Puede especificar parámetros adicionales como, por ejemplo, las características del índice por omisión, de la misma manera que para habilitar una columna de texto.

Después de haber creado el índice, puede mover o suprimir los archivos externos. Puede seguir buscando en los archivos. Puede insertar nuevas filas en la tabla y utilizar UPDATE INDEX para actualizar el índice con las nuevas referencias de archivo.

Si la tabla que está habilitando utiliza un grupo de nodos con múltiples nodos físicos, asegúrese de que los archivos externos a los que hace referencia en las columnas de la tabla están ubicados en el nodo en el que reside la partición de la tabla.

Ejemplos

1. Cree la tabla DB2TX.EXTFILE que tenga como mínimo una columna o utilice una tabla existente.
2. Añada una columna de descriptores de contexto FILEHANDLES a la tabla DB2TX.EXTFILE

```
db2tx ENABLE TEXT FILES db2tx.extfile
        HANDLE          filehandles
        INDEXTYPE       linguistic
```


Habilitación de archivos de texto externos

```
UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)
UPDATEINDEX NOUPDATE
DIRECTORY \any\db2tx\indices
```

3. Inicialice el descriptor de contexto

- Para cada fila de una nueva tabla:

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

- Para una tabla existente, actualice las columnas de descriptores de contexto para que reflejen la referencia de archivo externo, especificando el nombre del archivo externo:

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

Consejo

No utilice INIT_TEXT_HANDLE para actualizar las columnas de descriptores de contexto que hagan referencia a archivos externos.

4. Actualice el índice

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

Finalización de la sesión

Finalización de la sesión

Ahora ha completado los pasos para preparar los documentos de texto en los que se ha de buscar.

Si ha especificado NOUPDATE para la palabra clave UPDATEINDEX cuando ha habilitado la columna de texto, DB2 Text Extender no indexa el texto inmediatamente sino que espera a la siguiente indexación periódica. Para actualizar el índice ahora, consulte el apartado “Actualización de un índice” en la página 99.

Cuando haya finalizado la indexación de los documentos, puede empezar a recuperar la información tal como se describe en el “Capítulo 6. Cómo buscar” en la página 77.

Consejo

Utilice GET INDEX STATUS para determinar cuándo ha terminado la indexación.

Para finalizar el procesador de mandatos de DB2 Text Extender, entre:

```
db2tx=>quit
```

Capítulo 6. Cómo buscar

DB2 Text Extender proporciona funciones SQL que le permiten incluir subconsultas de búsqueda de texto en consultas de SQL. Se proporcionan estas funciones además de las que están disponibles normalmente en SQL y se hace referencia a ellas en este manual como funciones de DB2 Text Extender.

Consulte el “Capítulo 11. Funciones de búsqueda” en la página 197 para ver una descripción de la sintaxis de estas funciones.

Antes de buscar, lea el apartado “Tipos de búsqueda” en la página 40 y utilice también GET INDEX SETTINGS para averiguar qué tipo de índice está asociado al texto en el que está buscando. Una búsqueda puede producir resultados distintos según el tipo de índice.

El tipo de índice que se supone en los ejemplos de este capítulo es lingüístico.

Este capítulo describe:

- Las funciones de DB2 Text Extender de ejemplo
- La tabla de ejemplo DB2TX.SAMPLE
- Descriptores de contexto para archivos externos
- Establecimiento de la vía de acceso de función para dar a SQL el acceso a las funciones de DB2 Text Extender
- Búsqueda de texto, utilizando CONTAINS, NO_OF_MATCHES y RANK
- Especificación de argumentos de búsqueda en funciones de DB2 Text Extender, utilizando ejemplos de CONTAINS
- Redefinición de la búsqueda anterior, utilizando CONTAINS y REFINE
- Establecimiento y extracción de información en descriptores de contexto, utilizando INIT_TEXT_HANDLE, CCSID, FORMAT y LANGUAGE
- Mejora del rendimiento de la búsqueda, utilizando SEARCH_RESULT.

Dónde encontrar ejemplos de sintaxis de las funciones de búsqueda

DB2 Text Extender proporciona un archivo de entrada del procesador de línea de mandatos denominado txsample.udf. Contiene ejemplos de las funciones de búsqueda de DB2 Text Extender que se ejecutan en la tabla de ejemplo. Utilice este archivo para ver ejemplos de la sintaxis de la preparación del texto y la función de búsqueda de DB2 Text Extender y de la sintaxis utilizada en los argumentos de búsqueda.

Dónde encontrar ejemplos de sintaxis de las funciones de búsqueda

Para ejecutar este archivo, entre:

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

Tabla de ejemplo DB2TX.SAMPLE

La tabla de ejemplo DB2TX.SAMPLE se utiliza en muchos ejemplos de sintaxis de búsqueda. Puede intentar estos ejemplos utilizando DB2 Text Extender.

Para crear la tabla DB2TX.SAMPLE, consulte el apartado “Preparación de una base de datos de ejemplo para la verificación de la instalación” en la página 16.

En la Tabla 6 se muestra un extracto de la tabla de ejemplo.

Tabla 6. Un extracto de la tabla de ejemplo 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.

Tabla de ejemplo DB2TX.SAMPLE

Tabla 6. Un extracto de la tabla de ejemplo DB2TX.SAMPLE (continuación)

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>

Tabla de ejemplo DB2TX.SAMPLE

Tabla 6. Un extracto de la tabla de ejemplo DB2TX.SAMPLE (continuación)

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?

Esta es una parte de la estructura de la tabla que muestra la primera y la última columna:

DB2TX.SAMPLE

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

Figura 13. La estructura de la tabla DB2TX.SAMPLE

La columna que contiene el texto que se ha de buscar es COMMENT. Sin embargo, antes de poder buscar en el texto de esta columna debe preparar la columna COMMENT para que DB2 Text Extender la utilice mediante el mandato ENABLE TEXT COLUMN.

Después de este paso de preparación, la tabla DB2TX.SAMPLE tiene este aspecto:

DB2TX.SAMPLE

DOCID	COMMENT	COMMENTHANDLE
doc 1	Customer is ...	X'..descriptor de contexto..'
doc 2	After rechecking ...	X'..descriptor de contexto..'

Figura 14. La tabla de DB2TX.SAMPLE después de habilitarla

Ahora, la tabla tiene una columna adicional para los descriptores de contexto y cada objeto de texto tiene un descriptor de contexto exclusivo que lo representa.

Cuando luego inserte texto en una columna de texto habilitada, un desencadenante de inserción crea un descriptor de contexto para el mismo.

DB2TX.SAMPLE

DOCID	COMMENT	COMMENTHANDLE
doc 1	Customer is ...	X'..descriptor de contexto..'
doc 2	After rechecking ...	X'..descriptor de contexto..'

Descriptores de contextos creados por ENABLE TEXT COLUMN

Fila insertada:

doc 11	I have installed ...	X'..descriptor de contexto..'
--------	----------------------	-------------------------------

Descriptor de contexto creado por un desencadenante de inserción

Figura 15. El descriptor de contexto para una fila insertada se crea por un desencadenante

Un descriptor de contexto contiene la información siguiente:

- Un ID de documento

- El nombre y la ubicación del índice asociado

- La información de documentos: CCSID, formato e idioma.

Las funciones SQL proporcionadas por DB2 Text Extender toman un descriptor de contexto como parámetro y almacenan, acceden, buscan y manipulan el texto como parte del proceso SQL de la tabla.

Descriptores de contexto para archivos externos

DB2 Text Extender no sólo puede buscar en el texto almacenado en tablas DB2 UDB , sino que también puede buscar en archivos de texto almacenados en otro lugar. El apartado “Habilitación de archivos de texto externos” en la página 73 describe el paso de preparación que posibilita la búsqueda en documentos de texto que no están almacenados en tablas DB2 UDB . En este paso, el mandato ENABLE TEXT FILES crea una columna de descriptores de contexto del tipo DB2TEXTFH para descriptores de contexto de archivos externos. La columna de descriptores de contexto se añade a una tabla existente.

Puede, por ejemplo, crear una tabla que contenga columnas para el nombre del autor y para la fecha de creación del documento.

Los descriptores de contexto de los archivos se inicializan utilizando INIT_TEXT_HANDLE. Cada descriptor de contexto no sólo contiene un ID de documento, el nombre y la ubicación del índice asociado y la información de documentos (el CCSID, el formato y el idioma), sino que también contiene la referencia al archivo externo.

Establecimiento de la vía de acceso de la función actual

Establecimiento de la vía de acceso de la función actual

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

Utilice la sentencia de SQL SET CURRENT FUNCTION PATH para añadir DB2TX a los nombres de vías de acceso actuales para que SQL pueda encontrar las funciones de DB2 Text Extender. Si decide no hacerlo, puede calificar los nombres de función explícitamente escribiendo, por ejemplo, DB2TX.CONTAINS para la función CONTAINS.

Los ejemplos de este capítulo utilizan el formato calificado para las funciones de DB2 Text Extender. Puede utilizar las sentencias de ejemplo exactamente tal como están escritas sin tener que establecer la vía de acceso de la función actual.

Consejo

Recuerde que ha de establecer la vía de acceso de la función actual cada vez que se conecta a una base de datos.

Búsqueda de texto

► CONTAINS (—descriptor-contexto—, —argumento-búsqueda—) ◀
├── NO_OF_MATCHES ──┘
└── RANK ───────────┘

► SEARCH_RESULT ◀

► (—esquema—, —tabla—, —descriptor-contexto—, —argumento-búsqueda—) ◀

Esta sección describe cómo utilizar las funciones SQL proporcionadas con DB2 Text Extender para buscar en las bases de datos DB2 que contienen texto. Le indica cómo:

- Realizar una consulta
- Determinar cómo se han encontrado las coincidencias en un documento de texto
- Obtener el rango de un documento de texto encontrado.

La utilización de SEARCH_RESULT se describe en el apartado “Mejora del rendimiento de la búsqueda” en la página 97.

Cada una de estas funciones de DB2 Text Extender busca las apariciones del argumento de búsqueda en el índice de texto. Si hay 100 000 documentos de texto en la tabla, las funciones CONTAINS, RANK o NO_OF_MATCHES se llaman 100 000 veces. Pero no se busca en el índice de texto 100 000 veces. En su lugar, la primera vez que se llama a la función, se crea una lista interna de todos los documentos que contienen el término de búsqueda; las llamadas posteriores de la función determinan si el documento implicado está en la lista.

Consejo

Cuando utilice las funciones de DB2 Text Extender para buscar en una tabla, asegúrese de que pasa la columna de descriptores de contexto a la función, en lugar de la columna de texto. Si intenta buscar en una columna de texto, SQL responde con un mensaje que indica que el tipo de datos es incorrecto, por ejemplo:

No se ha encontrado ninguna función con el nombre "CONTAINS" que tenga argumentos compatibles en la vía de acceso de la función.

Si busca texto inmediatamente después de emitir los mandatos ENABLE TEXT TABLE o ENABLE TEXT COLUMN, puede producirse un error RC_SE_EMPTY_INDEX que indica que el índice de texto que crea el mandato no existe todavía. El tiempo que tarda en crearse un índice depende de factores como, por ejemplo, el número de documentos que se indexan y el rendimiento del sistema que realiza la indexación. Puede tardar varios minutos o varias horas, y debe realizarse cuando el sistema está poco cargado como, por ejemplo, por la noche.

Si se recibe este mensaje, intente realizar la búsqueda más tarde o utilice GET INDEX STATUS para comprobar si se han producido errores de indexación.

Realización de una consulta

Este ejemplo demuestra cómo la función CONTAINS busca texto en los documentos identificados por un descriptor de contexto. Devuelve 1 si el texto satisface el argumento de búsqueda, de lo contrario, devuelve 0.

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

En este ejemplo, se busca el término compress en el texto al que los descriptores de contexto hacen referencia en la columna

Búsqueda de texto

COMMENTHANDLE. Los descriptores de contexto de la columna COMMENTHANDLE indican dónde está indexado el texto COMMENT.

Consejo

Si ha creado identificadores en mayúsculas y minúsculas para las tablas o las columnas, recuerde que estos nombres deben ir entre comillas dobles. Por ejemplo:

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

Si especifica sentencias select de DB2 UDB desde la línea de mandatos, el analizador de la línea de mandatos del sistema operativo elimina los caracteres especiales como, por ejemplo, las comillas dobles de la serie de mandatos, por lo que debe utilizar una barra inclinada invertida para enmascarar estos símbolos especiales. Por ejemplo:

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

Búsqueda y devolución del número de coincidencias encontradas

Utilice la función NO_OF_MATCHES para determinar la frecuencia con que se encuentran los criterios de búsqueda en cada documento de texto.

```
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 devuelve un valor entero.

Búsqueda y devolución del rango de un documento de texto encontrado

RANK es un valor absoluto que indica el grado de coincidencia con los criterios de búsqueda del documento en relación a otros documentos encontrados. El valor indica el número de coincidencias encontradas en el documento en relación al tamaño del documento. Puede obtener el rango de un documento encontrado utilizando la función RANK.

Este es un ejemplo:

```
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 devuelve un valor DOUBLE entre 0 y 1.

Especificación de los argumentos de búsqueda

Los argumentos de búsqueda se utilizan en CONTAINS, NO_OF_MATCHES, RANK y HANDLE_LIST. Esta sección utiliza la función CONTAINS para mostrar diferentes ejemplos de argumentos de búsqueda en funciones DB2 Text Extender.

Búsqueda de varios términos

Puede haber más de un término en un argumento de búsqueda. Una forma de combinar varios términos de búsqueda es conectarlos entre sí utilizando comas, de la siguiente manera:

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

Este formato de argumento de búsqueda encuentra el texto que contiene cualquiera de los términos de búsqueda. En términos lógicos, los términos de búsqueda se conectan mediante un operador OR.

Búsqueda con los operadores booleanos AND y OR

(Consulte también el apartado “Búsqueda con el operador booleano NOT” en la página 90.)

Los términos de búsqueda pueden combinarse con otros términos de búsqueda utilizando los operadores booleanos “&” (AND) y “|” (OR). Por ejemplo:

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

Puede combinar varios términos de búsqueda utilizando los operadores booleanos:

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

Especificación de los argumentos de búsqueda

Si utiliza más de un operador booleano, DB2 Text Extender los evalúa de izquierda a derecha, pero el operador lógico AND (&) tiene un enlace superior al operador lógico OR (!). Por ejemplo, si no incluye los paréntesis,

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

se evalúa como:

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

Por lo que en el ejemplo siguiente, debe incluir los paréntesis:

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

Si combina operadores booleanos con términos de búsqueda encadenados mediante el separador coma, de esta manera:

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

la coma se interpreta como el operador booleano OR, de esta manera:

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

Búsqueda de variaciones de un término

Si está utilizando un índice **precise** (exacto), DB2 Text Extender busca los términos exactamente tal como los escribe. Por ejemplo, el término *media* sólo encuentra el texto que contiene “media”. El texto que contiene el singular “medium” no se encuentra.

Si está utilizando un índice **linguistic** (lingüístico), DB2 Text Extender busca también las variaciones de los términos como, por ejemplo, el plural de un nombre o un tiempo de verbo distinto.

Por ejemplo, el término *drive* encuentra el texto que contiene “drive”, “drives”, “driving”, “drove” y “driven”.

Búsqueda de fragmentos de un término (enmascaramiento de caracteres)

Los caracteres de máscara, conocidos como caracteres “comodines”, ofrecen una forma de flexibilizar la búsqueda. Representan caracteres opcionales al principio, en medio o al final de un término de búsqueda. Aumentan el número de documentos de texto que una búsqueda encuentra.

Consejo

Si utiliza caracteres de máscara, no puede utilizar la palabra clave SYNONYM FORM OF.

Especificación de los argumentos de búsqueda

Los caracteres de máscara son especialmente útiles para buscar variaciones de términos si tiene un índice exacto. Si tiene un índice lingüístico, muchas de las variaciones encontradas utilizando los caracteres de máscara también se encontrarían.

Tenga en cuenta que los fragmentos de palabras (palabras enmascaradas con caracteres comodín) no pueden reducirse a su formato base. Por lo tanto, si busca `pass%`, no encontrará las palabras “passes” ni “passed”, porque se reducen a su formato base “pass” en el índice. Para buscarlas, debe buscar `pass%`.

DB2 Text Extender utiliza dos caracteres de máscara: porcentaje (%) y subrayado (_):

- % representa **cualquier número de caracteres arbitrarios**. Este es un ejemplo de % utilizado como carácter de máscara al principio del término de búsqueda:

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

Este término de búsqueda encuentra los documentos de texto que contienen, por ejemplo, “username”, “file_name” y “table-name”.

% también puede representar una **palabra completa**: el ejemplo siguiente busca los documentos de texto que contienen frases como “graphic function” y “query function”.

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

- _ representa **un carácter** en un término de búsqueda: El ejemplo siguiente busca los documentos de texto que contienen “CLOB” y “BLOB”.

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

Búsqueda de términos que ya contienen un carácter de máscara

Si desea buscar un término que contiene el carácter “%” o el carácter “_”, debe anteponer al carácter el llamado carácter de *escape* y, después, identificar el carácter de escape utilizando la palabra clave ESCAPE.

Por ejemplo, para buscar “10% interest”:

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

Especificación de los argumentos de búsqueda

El carácter de escape en este ejemplo es "!".

Búsqueda de términos en cualquier orden

Si busca “hard disk” tal como se muestra en el ejemplo siguiente, sólo encontrará los dos términos si son adyacentes y aparecen en el orden mostrado, sin tener en cuenta el tipo de índice que esté utilizando.

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

Para buscar términos en cualquier orden, como “data disks and hard drives”, por ejemplo, utilice una coma para separar los términos:

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

Búsqueda de términos en la misma sentencia o párrafo

Este es un ejemplo de un argumento de búsqueda que busca documentos de texto en el que aparecen los términos de búsqueda en la misma sentencia:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              '"compress" IN SAME SENTENCE AS "decompress"') = 1
```

También puede buscar más de dos palabras que aparezcan juntas. En el siguiente ejemplo, se efectúa una búsqueda de varias palabras que aparecen en el mismo párrafo:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              '"compress" IN SAME PARAGRAPH AS "decompress"
                              AND "encryption"') = 1
```

Búsqueda de términos en secciones de documentos estructurados

Este es un ejemplo de un argumento de búsqueda para encontrar los documentos de texto en los que el término Williams aparece en la subsección author de la sección play de los documentos estructurados. La estructura del documento se especifica por el modelo play que se describe en un archivo de modelos de documentos. Consulte el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48 para obtener más información.

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'MODEL play SECTIONS (play/author) "williams"') = 1
```

Búsqueda de sinónimos de términos

Para un índice lingüístico, puede flexibilizar las búsquedas buscando además de los términos de búsqueda que especifique, las palabras que tienen un significado similar. Por ejemplo, cuando busca la palabra “book”, puede ser útil buscar también su sinónimo. Para hacerlo, especifique:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE, 'SYNONYM FORM OF "book"') = 1
```

Cuando utiliza SYNONYM FORM OF, se supone que los sinónimos del término están conectados por el operador lógico OR, es decir, el argumento de búsqueda se interpreta como:

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

Los sinónimos están en un diccionario que se proporciona con DB2 Text Extender. El diccionario por omisión utilizado para los sinónimos siempre es US_ENGLISH, no el idioma especificado en los valores de configuración del texto.

Puede cambiar el diccionario para una consulta en particular especificando un idioma diferente. Este es un ejemplo:

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'SYNONYM FORM OF UK_ENGLISH "programme"') = 1
```

Consejo

No puede utilizar la palabra clave SYNONYM si hay caracteres de máscara en el término de búsqueda o si se utiliza NOT con el argumento de búsqueda.

Realización de una búsqueda lingüística

DB2 Text Extender ofrece un poderoso proceso lingüístico para realizar una búsqueda basándose en los términos de búsqueda que el usuario proporciona. Las funciones lingüísticas se aplican cuando el índice es lingüístico. Las funciones lingüísticas se describen en el “Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto” en la página 255.

Un ejemplo es una búsqueda de un plural como, por ejemplo, “utilities”, y encontrar “utility”. El plural se reduce a su formato base utility, utilizando un diccionario inglés, antes de que empiece la búsqueda.

Especificación de los argumentos de búsqueda

Sin embargo, el diccionario inglés no contiene la información para reducir las variaciones de los términos en otros idiomas a su formato base. Para buscar el plural de un término en otro idioma debe utilizar el diccionario para ese idioma.

Si especifica GERMAN, por ejemplo, puede buscar “geflogen” (volado) y encontrar todas las variaciones de su formato base “fliegen” (volar)—no sólo “geflogen”, si no también “fliege”, “fliegt”, etcétera.

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'STEMMED FORM OF GERMAN "geflogen"') = 1
```

Consejo

Cuando busque en documentos que no estén en inglés de EE.UU., especifique el idioma en el argumento de búsqueda *sin tener en cuenta el idioma por omisión*.

Si especifica siempre el formato base de un término de búsqueda, en lugar de una variación del mismo, no es necesario que especifique un idioma.

Para entender porqué, considere lo que ocurre cuando se indexa el texto de la base de datos. Si está utilizando un índice lingüístico, se reducen todas las variaciones de un término a su formato base antes de que se almacenen los términos en el índice. Esto significa que, en la tabla DB2TX.SAMPLE, aunque el término “decompress” aparece en la primera entrada de la columna COMMENT, “decompression” aparece en la segunda entrada, el índice sólo contiene el formato base “decompress” e identifica este término (o sus variaciones) como si estuviese en las dos entradas.

Después, si busca el formato base “decompress”, encontrará todas las variaciones. Sin embargo, si busca una variación como “decompression”, no puede encontrarla directamente. Debe especificar un diccionario adecuado para la búsqueda, de modo que la variación pueda convertirse primero a su formato base.

Búsqueda con el operador booleano NOT

Puede utilizar el operador booleano NOT para excluir de la búsqueda unos documentos de texto en particular. Por ejemplo:

```
("compress", "compiler") & NOT "DB2"
```

Cualquier documento que contenga el término “DB2” se excluye de la búsqueda de “compress” o “compiler”.

Especificación de los argumentos de búsqueda

No puede utilizar el operador NOT en combinación con IN SAME SENTENCE AS o IN SAME PARAGRAPH AS descritas en el apartado “Búsqueda de términos en la misma sentencia o párrafo” en la página 88, tampoco puede utilizarlo con SYNONYM FORM OF descrito en el apartado “Búsqueda de sinónimos de términos” en la página 89.

Sólo puede utilizar el operador NOT con una búsqueda primaria, es decir, no puede combinar con libertad los operadores &, | y NOT (consulte el apartado “Sintaxis del argumento de búsqueda” en la página 217).

Ejemplo de utilización **no** permitida de NOT:

```
NOT("compress" & "compiler")
```

Lo permitido es:

```
NOT("compress" , "compiler")
```

Búsqueda similar

La búsqueda “similar” busca palabras que estén escritas de manera parecida al término de búsqueda. Está disponible para índices Ngram.

Por ejemplo:

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                          'FUZZY FORM OF 2 "compress"') = 1
```

Esta búsqueda encontraría una aparición de la palabra incorrectamente escrita conpress.

El nivel de coincidencia, en el ejemplo “2”, especifica el grado de exactitud. Se da soporte a cinco niveles, el nivel 1 proporciona la coincidencia inferior de aproximadamente el 20 por ciento y el nivel 5 proporciona la coincidencia más exacta de aproximadamente el 90 por ciento. Utilice una búsqueda similar cuando sea posible que haya palabras escritas incorrectamente en el documento, como es el caso de los documentos creados utilizando un dispositivo de Reconocimiento óptico de caracteres o entrada fonética.

Respeto de los límites de la palabra-frase

La búsqueda de “límites” se ha desarrollado para el idioma coreano. Asegura que DB2 Text Extender respete los límites de la palabra durante la búsqueda. Por ejemplo:

```
db2=>SELECT DATE, SUBJECT
      FROM DB2TX.SAMPLE
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                          'BOUND "expresión-coreana"') = 1
```

Especificación de los argumentos de búsqueda

Búsqueda de palabras con sonido similar

La búsqueda de “sonido” busca palabras que suenen igual que el argumento de búsqueda. Es útil en los documentos que pueden contener palabras que suenen igual, pero que se escriban de forma diferente. El nombre alemán que se pronuncia my-er, por ejemplo, se escribe de varias formas.

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

Esta búsqueda puede encontrar las apariciones de “Meyer”, “Mayer” y “Maier”.

Búsqueda en tesoro

La búsqueda en tesoro es otra de las poderosas funciones de ampliación del término de búsqueda de DB2 Text Extender. Los términos adicionales buscados se toman de un tesoro que el usuario crea, por lo que tiene control directo sobre los mismos. El usuario busca “base de datos”, por ejemplo, y puede encontrar términos como “depósito” y “DB2”.

Este tipo de búsqueda está pensado para áreas específicas de interés en las que se efectúan búsquedas frecuentes; un área en la que vale la pena invertir tiempo para crear un tesoro para producir resultados de búsqueda significativamente más eficaces.

Consulte el apartado “Conceptos del tesoro” en la página 268 para obtener más información y una descripción de cómo crear un tesoro. El ejemplo de la Figura 18 en la página 269 es un pequeño extracto de un tesoro sobre el tema de bases de datos. Se utiliza en los ejemplos siguientes para demostrar la sintaxis de la utilización de la ampliación del tesoro.

Este ejemplo toma el término “object relational database management system” y lo amplía, añadiendo todas las *instancias* de este término que se encuentran en el tesoro “myterms”. Aquí, “DB2” se añade a la búsqueda.

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                              'THESAURUS "myterms"
                              EXPAND "INST"
                              TERM OF "object relational database management system"') = 1
```

El siguiente ejemplo toma el término “document management system” y lo amplía, añadiendo todos sus *sinónimos*. Sólo hay un sinónimo – “library”.

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
```

```
'THESAURUS "myterms"  
EXPAND "SYN"  
TERM OF "document management system") = 1
```

Búsqueda de texto libre e híbrida

La “búsqueda de texto libre” es una búsqueda en la cual el término de búsqueda se expresa como texto de formato libre. Una frase o una sentencia describe en idioma natural el tema que se ha de buscar. La secuencia de la palabras en una consulta de texto libre no es importante. Además, se da soporte a las llamadas *afinidades léxicas*. En la recuperación, son pares de palabras determinadas que aparecen en un término de consulta de texto libre y que aparecen en la colección de documentos, con una frecuencia mínima y a una distancia mínima. La distancia para los documentos en inglés es de cinco palabras.

Tenga en cuenta que no se da soporte al enmascaramiento de caracteres o de palabras para las series de búsqueda en un argumento de texto libre.

Por ejemplo:

```
db2=>SELECT DATE, SUBJECT  
      FROM DB2TX.SAMPLE  
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,  
        'IS ABOUT "everything related to AIX installation"') = 1
```

La búsqueda híbrida es una combinación de búsqueda booleana y de búsqueda de texto libre. Por ejemplo:

```
db2=>SELECT DATE, SUBJECT  
      FROM DB2TX.SAMPLE  
      WHERE DB2TX.CONTAINS (COMMENTHANDLE,  
        '"DB2" & IS ABOUT "everything related to AIX installation"') = 1
```

Redefinición de la búsqueda anterior

Cuando un argumento de búsqueda encuentra demasiadas apariciones, a menudo es útil reducir o *redefinir* la búsqueda combinando el argumento de búsqueda inicial con un segundo argumento de búsqueda en una relación AND booleana.

Puede redefinir los resultados de la búsqueda sin utilizar la función REFINE, almacenando los resultados en una tabla y realizando una nueva búsqueda en esa tabla. Sin embargo, dependiendo del número de términos calificados, este método es menos eficaz que almacenar el último argumento de búsqueda y utilizar REFINE.

Los pasos siguientes muestran cómo se realiza una búsqueda y, después, se redefine utilizando la función REFINE. La función REFINE devuelve un

Redefinición de la búsqueda anterior

argumento de búsqueda que es una combinación de AND booleano y sus dos parámetros de entrada. El argumento de búsqueda combinada devuelto por REFINE es un valor de tipo LONG VARCHAR.

1. Cree una tabla para los anteriores argumentos de búsqueda.

Cree la tabla PREVIOUS_SEARCHES para que contenga los argumentos de las búsquedas que ya se han realizado.

```
db2=>CREATE TABLE PREVIOUS_SEARCHES (step INT,  
                                       searchargument LONG VARCHAR)
```

PREVIOUS_SEARCHES

STEP	SEARCHARGUMENT
------	----------------

2. Busque el primer argumento de búsqueda.

Busque la palabra “compress” en la tabla de ejemplo.

```
db2=>SELECT COMMENT  
      FROM DB2TX.SAMPLE  
      WHERE DB2TX.CONTAINS (COMMENTHANDLE, "compress") = 1
```

Inserte el argumento de búsqueda en la tabla PREVIOUS_SEARCHES para utilizarlo en pasos posteriores.

```
db2=>INSERT INTO PREVIOUS_SEARCHES  
      VALUES (1, "compress")
```

PREVIOUS_SEARCHES

STEP	SEARCHARGUMENT
1	"compress"

3. Redefina la búsqueda.

Suponiendo que la búsqueda devuelva demasiados documentos de texto, redefina la búsqueda combinando el término que ha buscado anteriormente con la palabra “compiler” utilizando la función REFINE.

```
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  
          FROM DB2TX.SAMPLE, LAST_SEARCH  
          WHERE DB2TX.CONTAINS(COMMENTHANDLE,  
                                DB2TX.REFINE(LAST_SEARCH, "compiler")) = 1
```

Redefinición de la búsqueda anterior

Inserte el argumento de búsqueda redefinido en la tabla PREVIOUS_SEARCHES para utilizarlo en pasos posteriores.

```
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"

Puede repetir este paso hasta que el número de documentos de texto encontrados sea lo suficientemente pequeño.

Establecimiento y extracción de información en descriptores de contexto

Los descriptores de contexto contienen el CCSID, el formato y el idioma de los documentos de texto. Los descriptores de contexto para archivos externos contienen, además, un puntero para el archivo externo. Estos descriptores de contexto se crean al habilitar una columna de texto o los archivos externos.

Las funciones de DB2 Text Extender descritas aquí le permiten establecer o cambiar la información de texto de los descriptores de texto.

Establecimiento de información de texto al insertar nuevo texto

►►—INIT_TEXT_HANDLE—┌(—formato—, —idioma—)—————►
└(—CCSID—, —formato—, —idioma—, —nombrearchivo—)——┘

Cuando ejecute el mandato ENABLE TEXT COLUMN para habilitar una columna de texto que ya contiene texto, puede establecer implícitamente el formato y el idioma del texto en los valores especificados en los valores de configuración del texto. Estos valores de formato y de idioma se almacenan entonces en el descriptor de contexto. Si desea establecer valores de formato y de idioma diferentes, puede especificarlos explícitamente en el mandato ENABLE TEXT COLUMN.

Cuando ejecute el mandato ENABLE TEXT FILES, también puede establecer el CCSID y la ubicación del documento.

Establecimiento y extracción de información en descriptores de contexto

Luego, al insertar una fila que contenga texto, un desencadenante de inserción crea un descriptor de contexto y establece el formato y el idioma del texto en los valores que se han utilizado cuando se ha habilitado la columna de texto.

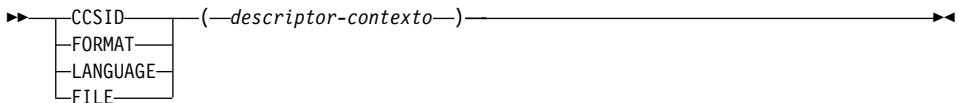
Para establecer el formato y el idioma en valores diferentes a estos, utilice la función `INIT_TEXT_HANDLE` en el mandato `INSERT`. Mientras se inserta la fila, la función `INIT_TEXT_HANDLE` crea un descriptor de contexto inicializado parcialmente que contiene los valores de idioma y de formato que especifique. Entonces el desencadenante de inserción rellena los demás valores del descriptor de contexto.

En el ejemplo siguiente, `INIT_TEXT_HANDLE` preestablece el idioma y el formato en el descriptor de contexto inicializado. El mandato `INSERT` coloca este descriptor de contexto en la columna `COMMENTHANDLE`.

```
db2=>INSERT INTO DB2TX.SAMPLE (DOCID, COMMENT, COMMENTHANDLE)
      VALUES ('doc 101',
              'I have installed...',
              DB2TX.INIT_TEXT_HANDLE('AMI', 'GERMAN') )
```

El valor que `INIT_TEXT_HANDLE` devuelve es de tipo `DB2TEXTH` o `DB2TEXTFH`.

Extracción de información de descriptores de contexto



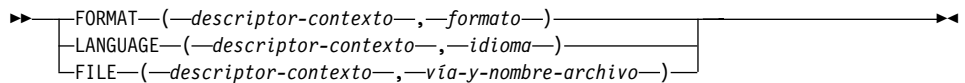
Este es un ejemplo de extracción de un `CCSID` de un descriptor de contexto:

```
db2=>SELECT DISTINCT DB2TX.CCSID(COMMENTHANDLE)
      FROM DB2TX.SAMPLE
```

De la misma manera, puede extraer el formato o el idioma de un documento de texto o la ubicación de los archivos externos. Este es un ejemplo que ilustra la utilización de la función `FORMAT`. Devuelve el número de documentos ASCII (TDS) de la tabla de ejemplo.

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

Cambio de la información de los descriptores de contexto



Las funciones `FORMAT` y `LANGUAGE` también pueden cambiar la especificación correspondiente en un descriptor de contexto. Estas funciones devuelven el descriptor de contexto cambiado como un valor de tipo `DB2TEXTH` o `DB2TEXTFH`.

El ejemplo siguiente muestra cómo cambiar el valor del idioma de un documento de texto.

```
db2=>UPDATE DB2TX.SAMPLE
        SET COMMENTHANDLE = DB2TX.LANGUAGE(COMMENTHANDLE, 'FRENCH')
        WHERE ...
```

Si vuelve a utilizar la función `LANGUAGE` de nuevo, puede ver que se ha producido el cambio:

```
db2=>SELECT DISTINCT DB2TX.LANGUAGE(COMMENTHANDLE)
        FROM DB2TX.SAMPLE
```

Mejora del rendimiento de la búsqueda

La función `SEARCH_RESULT` utiliza el concepto DB2 de las funciones de tabla con valores. La función se utiliza en la cláusula `FROM` de una sentencia de SQL y devuelve una tabla intermedia con resultado de la búsqueda de la serie de búsqueda especificada. La sintaxis de la serie de búsqueda es la misma que la descrita en el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215. La ventaja de esta función con respecto a `CONTAINS` o `RANK` es una mejora significativa del rendimiento cuando se emplean tablas grandes.

La tabla devuelta tiene la estructura siguiente:

Nombre de columna	Tipo de datos
<code>HANDLE</code>	<code>DB2TX.DB2TEXTH</code> , <code>DB2TX.DB2TEXTFH</code>
<code>NUMBER_OF_MATCHES</code>	<code>INTEGER</code>
<code>RANK</code>	<code>DOUBLE</code>

Ejemplo:

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

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

Mejora del rendimiento de la búsqueda

donde `T1.handle = T2.mydochandle`

Si no ha establecido la vía de acceso actual, la última línea del ejemplo es:

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

`SELECT NUMBER_OF_MATCHES, RANK, HANDLE` hace que se devuelvan los tres elementos, pero puede especificarlos en cualquier combinación. Por ejemplo, puede que desee omitir `RANK` para evitar el proceso intensivo asociado al mismo.

Si sólo necesita el valor de `HANDLE`, puede utilizar `SELECT COUNT(*)` simplemente.

Capítulo 7. Administración

Este capítulo describe cómo mantener los índices de texto, cómo obtener información útil y cómo anular el proceso de preparación de texto.

Mantenimiento de los índices de texto

Estas son las tareas de mantenimiento:

- Actualización de un índice
- Cambio de los valores de un índice
- Restablecimiento del estado de un índice
- Supresión de los sucesos de índice
- Reorganización de un índice.

Puede ejecutar estas tareas en cualquier momento y en cualquier orden.

Actualización de un índice

Resumen

Cuándo

Cuando un índice debe actualizarse inmediatamente sin esperar a que se produzca la indexación periódica. (Consulte el apartado “Habilitación de una columna de texto” en la página 66 para obtener información acerca de la indexación periódica.)

Mandato

UPDATE INDEX

Autorización

ALTER, SELECT, UPDATE en la tabla

Este ejemplo actualiza el índice para una tabla de índice común:

```
db2tx UPDATE INDEX db2tx.mytable
```

Este ejemplo actualiza el índice para una columna de una tabla de múltiples índices:

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

Mantenimiento de los índices de texto

Utilice este mandato para actualizar el índice inmediatamente, sin esperar a que se produzca automáticamente la siguiente indexación periódica. Es útil cuando se han añadido varios documentos de texto a una base de datos y desea buscar en ellos inmediatamente.

DB2 Text Extender indexa los documentos de texto de esta columna (o de todas las columnas de la tabla) que se han insertado o cambiado y elimina del índice los términos de los documentos que se han suprimido. La tabla de registro cronológico asociada al índice contiene información acerca de los documentos que se han insertado, actualizado y suprimido.

Actualización de un índice para archivos externos

Una tabla de registro cronológico no contiene automáticamente la información de los cambios de cualquiera de los archivos externos que pueda haber indexado (consulte el apartado “Habilitación de archivos de texto externos” en la página 73) como, por ejemplo, la sustitución de un documento por una versión más reciente que tiene el mismo nombre absoluto de vía de acceso. DB2 Text Extender no puede supervisar las actualizaciones que se producen en estos archivos, en las tablas de registro cronológico porque las actualizaciones no se producen dentro del ámbito de DB2 UDB .

Para que se reflejen las actualizaciones de archivos externos en un índice de DB2 Text Extender, puede hacer lo siguiente:

1. Fuerce que una entrada de “cambio” se coloque en la tabla de registro cronológico emitiendo una sentencia de actualización en la columna de descriptores de contexto correspondiente que no haga nada efectivamente:

```
UPDATE tabla
SET   coldescontarch = coldescontarch
WHERE DB2TX.FILE(coldescontarch) = nombreachivo
```

donde *nombreachivo* es el nombre absoluto de la vía de acceso del archivo externo que se ha actualizado.

2. Ejecute UPDATE INDEX para actualizar el índice, incluyendo el cambio realizado en el archivo externo.

Cambio de los valores de un índice

Resumen**Cuándo**

Cuando se haya de cambiar la frecuencia de actualización de un índice.

Mandato

CHANGE INDEX SETTINGS

Autorización

ALTER, SELECT, UPDATE en la tabla

Utilice este mandato para cambiar la frecuencia de actualización de un índice.

Frecuencia de actualización

Consulte el apartado “Actualización de un índice” en la página 46 para obtener más información. Si no especifica ninguna frecuencia de actualización, no se cambian los valores actuales.

Utilice la palabra clave UPDATEINDEX para determinar si la indexación de los documentos empieza inmediatamente o cuando esté planificada la próxima indexación periódica. Si no utiliza esta palabra clave, no se cambia el valor actual.

Ejemplos

Para cambiar la frecuencia de actualización para el índice de modo que se actualice a las 12:00 o 15:00, de lunes a viernes, si hay un mínimo de 100 documentos de texto en cola:

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

Para detener la actualización periódica de un índice:

```
db2tx CHANGE INDEX SETTINGS db2tx.mytable
      HANDLE      commenthandle
      UPDATEFREQ none
```

Mantenimiento de los índices de texto

Restablecimiento del estado del índice

Resumen

Cuándo

Cuando ya no se busque ni se actualice un índice.

Mandato

RESET INDEX STATUS

Autorización

Ninguna

Pueden darse algunas situaciones que le impidan buscar en un índice o actualizarlo. El apartado “Visualización del estado de un índice” en la página 106 describe cómo determinar si se ha producido uno de estos sucesos. RESET INDEX STATUS reactiva el índice para que pueda volver a buscar en él.

Este ejemplo restablece el estado del índice de una tabla de índice común:

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

La sintaxis le permite restablecer el estado del índice para una columna de texto en particular. Este ejemplo restablece el estado de índice de una columna de tabla con múltiples índices:

```
db2tx RESET INDEX STATUS db2tx.mytable HANDLE commenthandle
```

Supresión de sucesos de índice

Resumen

Cuándo

Cuando ya no necesite los mensajes de una tabla de registro cronológico de un índice.

Mandato

DELETE INDEX EVENTS

Autorización

Ninguna

Si algo le impide buscar en un índice o actualizarlo, o si un documento no puede indexarse, se conoce como *suceso* de indexación. La información acerca de los sucesos de indexación se almacenan en una tabla de registro

cronológico del índice. Puede ayudarle a determinar la causa del problema. Cuando ya no necesite estos mensajes, puede suprimirlos.

Este ejemplo suprime los mensajes del índice de una tabla de índice común:

```
db2tx DELETE INDEX EVENTS db2tx.mytable
```

La sintaxis también le permite suprimir los sucesos de indexación para una columna de texto en particular. Este ejemplo suprime los mensajes para el índice de una columna de tabla de múltiples índices:

```
db2tx DELETE INDEX EVENTS db2tx.mytable HANDLE commenthandle
```

Reorganización de un índice

Resumen

Cuándo

Quando GET INDEX STATUS indique que un índice debe reorganizarse manualmente.

Mandato

REORGANIZE INDEX

Autorización

Ninguna

Si una columna de texto se actualiza con frecuencia, la búsqueda en el índice puede ser ineficaz. Para que la búsqueda pueda volver a ser eficaz, se tiene que *reorganizar* el índice. Aunque DB2 Text Extender reconoce cuándo es necesario reorganizar un índice y lo hace automáticamente como proceso de fondo, pueden presentarse situaciones en las que sea necesario reorganizar manualmente un índice utilizando REORGANIZE INDEX. Puede utilizar el mandato GET INDEX STATUS para averiguar si es necesario reorganizar un índice.

Aunque se pueden realizar búsquedas en el índice mientras se ejecuta REORGANIZE INDEX, no se pueden efectuar actualizaciones.

En este ejemplo se reorganiza el índice de una tabla de índice común:

```
db2tx REORGANIZE INDEX db2tx.mytable
```

En este ejemplo, se reorganiza el índice de una columna de tabla de múltiples índices:

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

Obtención de información útil

Esta sección describe los mandatos para visualizar información sobre:

- El estado habilitado de las bases de datos, las tablas, las columnas y los archivos
- Los valores de las variables de entorno
- Los valores de la configuración del texto
- El estado de los índices
- Los sucesos de error
- Los valores de índice
- Los valores de texto para una columna.

Visualización de información del estado habilitado

Resumen

Cuándo

Cuando necesite información sobre el estado habilitado de las bases de datos, las tablas, las columnas de texto o los archivos externos.

Mandato

GET STATUS

Autorización

Ninguna

Entre:

```
db2tx GET STATUS
```

Este es un ejemplo de la salida visualizada por GET STATUS. Muestra el estado habilitado de la base de datos y de cualquier tabla, columna de texto o archivo de texto habilitados que contiene.

```
La base de datos está habilitada para Text Extender
```

```
La tabla DB2TX.MYTABLE está habilitada como tabla de índice común
```

```
La tabla DB2TX.MYTABLE está habilitada como tabla de índice común
```

```
NombreColTexto      NombreColDesCont
-----
COMMENT              COMMENTHANDLE
```

La tabla DB2TX.TEST está habilitada como una tabla de múltiples índices

NombreColTexto	NombreColDesCont
-----	-----
ABSTRACT1	ABSTRACT1HANDLE
ABSTRACT2	ABSTRACT2HANDLE

Visualización de los valores de las variables de entorno

Resumen

Cuándo

Cuando necesite información acerca de los valores de las variables de entorno.

Mandato

GET ENVIRONMENT

Autorización

Ninguna

Entre:

```
db2tx GET ENVIRONMENT
```

Este es un ejemplo de la salida visualizada por GET ENVIRONMENT.

Muestra los valores actuales de las variables de entorno de DB2 Text Extender.

```
Instancia gestor base datos actual (DB2INSTANCE) = db2
Nombre de base de datos por omisión (DB2DBDFT) = my_database
Nombre de instancia (DB2TX_INSTOWNER) = user1
Directorio instancias (DB2TX_INSTOWNERHOMEDIR) = /usr/instance1
```

Visualización de los valores de configuración del texto

Resumen

Cuándo

Cuando necesite información de los valores por omisión para el texto, el índice y el proceso.

Mandato

GET TEXT CONFIGURATION

Autorización

Ninguna

Obtención de información útil

Estos valores se describen en el apartado “Valores de la configuración del texto” en la página 19. Para cambiarlos, consulte el apartado “Cambio de configuración del texto” en la página 20.

Para visualizar la configuración del texto, entre:

```
db2tx GET TEXT CFG
```

Este es un ejemplo de la salida visualizada por GET TEXT CONFIGURATION. Muestra los valores actuales de la configuración del texto.

```
ID juego caract. codificad.(CCSID) = 850
Idioma (LANGUAGE) = US_ENGLISH
Formato (FORMAT) = TDS

Tipo de índice (INDEXTYPE) = LINGUISTIC
Frecuencia actualiz. (UPDATEFREQ) = NONE
Directorio índices (DIRECTORY) = user1/db2tx/indexes

Opción actual. índic.(UPDATEINDEX) = UPDATE
Cuenta compromiso (COMMITCOUNT) = 10 000
Nombre espacio tabla (TABLESPACE) = TXLOG
```

Visualización del estado de un índice

Resumen

Cuándo

Cuando necesite determinar si se puede buscar en un índice o si se puede actualizar el mismo.

Mandato

GET INDEX STATUS

Autorización

Ninguna

Pueden darse algunas situaciones que le impidan buscar en un índice o actualizarlo. En estas situaciones, se almacenan mensajes en la tabla de registro cronológico del índice que pueden ayudarle a determinar la causa. Por lo tanto, puede ser útil comprobar el estado de un índice y si hay algún mensaje disponible.

Este ejemplo visualiza el estado del índice de una tabla de índice común:

```
db2tx GET INDEX STATUS db2tx.mytable
```

La sintaxis le permite visualizar el estado del índice para una columna de texto en particular. Este ejemplo obtiene el estado del índice de una columna de tabla de múltiples índices:


```
db2tx GET INDEX STATUS db2tx.mytable HANDLE commenthandle
```

Este es un ejemplo de la salida visualizada por GET INDEX STATUS.

```
Nodo 1
Estado de búsqueda           = Búsqueda disponible
Estado de actualización      = Actualización disponible
Estado de reorganización     = iniciada 13.55
Documentos planificados      = 0
Documentos indexados         = 187000
Documentos índice primario  = 130000
Documentos índice secund.   = 57000
Sucesos de error            = Ningún suceso de error
```

Si el índice se divide entre varios nodos, la información se visualiza por nodo.

Estado de búsqueda

Indica si puede utilizar la columna de descriptores de contexto especificada para buscar en el índice. Si la búsqueda no está disponible, compruebe el código de razón indicado para obtener más información acerca de porqué se ha producido la situación y después utilice RESET INDEX STATUS para poder volver a trabajar con el índice. Consulte el apartado “Capítulo 20. Códigos de razón de sucesos de error” en la página 321.

Estado de actualización

Indica si puede actualizar el índice para la tabla o la columna especificada. Si la función de actualización del índice no está disponible, compruebe el código de razón indicado para obtener más información sobre porqué se ha producido la situación y, después, utilice RESET INDEX STATUS para poder volver a trabajar con el índice.

Estado de reorganización

Indica si puede reorganizar el índice para la tabla o la columna especificada. Si la función de reorganización no está disponible, compruebe el código de razón indicado para obtener más información acerca de porqué se ha producido la situación. Una razón normal para no poder utilizar la reorganización es porque se está actualizando el índice en este momento.

Documentos planificados

Muestra el número de documentos que se listan en la cola para indexación (o para suprimirlos del índice).

Documentos indexados

Muestra el número de documentos de la cola de documentos planificados que ya se han indexado.

Documentos índice primario

Muestra el número de documentos del índice primario.

Obtención de información útil

Documentos índice secund.

Muestra el número de documentos del índice secundario.

Sucesos de error

Muestra el número de sucesos que están disponibles en la tabla de registro cronológico del índice. Puede visualizar esta información tal como se describe en el apartado “Visualización de sucesos de error”. Cuando ya no necesite esta información, puede suprimirla tal como se describe en el apartado “Supresión de sucesos de índice” en la página 102.

Visualización de sucesos de error

Cuando se producen problemas durante la indexación como, por ejemplo, no se ha podido encontrar un documento planificado para la indexación, se graban los llamados *sucesos de error* en la tabla de registro cronológico del índice.

Los códigos de retorno de sucesos se describen en el “Capítulo 20. Códigos de razón de sucesos de error” en la página 321.

Puede acceder a los sucesos de error en una vista de la tabla de registro cronológico llamada `db2tx.LOGIXnnnnnn`, donde `IXnnnnnn` es el nombre del índice, que se puede obtener en la vista del catálogo.

Para obtener el nombre del índice:

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

La vista de sucesos de error tiene el diseño siguiente:

```
UPDATESTATUS  SMALLINT  
EVENTREASON   INTEGER  
EVENTMESSAGE  VARCHAR(1024)  
UPDATETIME    TIMESTAMP  
HANDLE        DB2TEXTX o DB2TEXTFH  
NODENUMBER    INTEGER
```

Este es un ejemplo para mostrar cómo puede acceder a la información del registro cronológico del índice:

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

Visualización de los valores de índice

Resumen**Cuándo**

Cuando necesite información acerca de los valores de un índice.

Mandato

GET INDEX SETTINGS

Autorización

Ninguna

Este ejemplo obtiene los valores del índice de una tabla de índice común:

```
db2tx GET INDEX SETTINGS db2tx.mytable
```

Este ejemplo obtiene los valores del índice de una columna de tabla de múltiples índices:

```
db2tx GET INDEX SETTINGS db2tx.mytable
        HANDLE commenthandle
```

Si la tabla se ha habilitado como tabla de múltiples índices, este mandato visualiza los valores de índice de todas las columnas de texto habilitadas en la tabla.

Este es un ejemplo de la salida visualizada por GET INDEX SETTINGS para una tabla de índice común. La salida para una tabla de múltiples índices muestra información similar para cada índice. La sintaxis le permite pedir los valores de índice para una columna de texto en particular.

Valores de índice actuales:

```
Tipo de índice      (INDEXTYPE) = LINGUISTIC
Opción actual. indic. (UPDATEINDEX) = UPDATE
Frecuencia actualiz. (UPDATEFREQ) = NONE
Nodo 1
Directorio de índices (DIRECTORY) = /home/user1/db2tx/indices
```

Si el índice se divide entre varios nodos, se visualiza la información de nodo para el directorio de índices.

Visualización de los valores de texto para una columna

Resumen

Cuándo

Cuando necesite información acerca de los valores de texto para una columna.

Mandato

GET TEXT INFO

Autorización

Ninguna

Este ejemplo obtiene la información de texto para el índice de una tabla de índice común:

```
db2tx GET TEXT INFO db2tx.mytable
```

Este ejemplo obtiene la información de texto para el índice de una columna de tabla de múltiples índices:

```
db2tx GET TEXT INFO db2tx.mytable HANDLE commenthandle
```

La sintaxis le permite especificar un nombre de tabla y el nombre de una columna de descriptores de contexto.

Si sólo especifica un nombre de tabla en el mandato, se visualiza la información de texto para cada columna habilitada en esta tabla. Si también especifica un nombre de columna de descriptores de contexto, sólo se visualiza la información de texto para esa columna.

Este es un ejemplo de lo que visualiza este mandato para una tabla de múltiples índices:

```
Información de texto para la columna ABSTRACT1
con la columna de descriptores de contexto ABSTRACT1HANDLE:
ID juego caract. codificad.(CCSID) = 850
Idioma (LANGUAGE) = US_ENGLISH
Formato (FORMAT) = TDS
```

```
Información de texto para la columna ABSTRACT2
con la columna de descriptores de contexto ABSTRACT2HANDLE:
ID juego caract. codificad.(CCSID) = 850
Idioma (LANGUAGE) = US_ENGLISH
Formato (FORMAT) = TDS
```

Trabajar con la vista del catálogo de DB2 Text Extender

DB2 Text Extender crea y mantiene una vista del catálogo llamada DB2TX.TEXTINDEXES para cada base de datos. Se crea cuando se ejecuta ENABLE DATABASE. Contiene información acerca de las tablas y las columnas que están habilitadas para DB2 Text Extender.

Se crean nuevas entradas en DB2TX.TEXTINDEXES siempre que se habilita una tabla, una columna o archivos externos. Las entradas se actualizan siempre que se modifican los valores de índice utilizando el mandato CHANGE INDEX SETTINGS. Se suprimen las entradas si se inhabilitan las columnas o las tablas.

Los datos de la vista del catálogo están disponibles mediante los recursos normales de consulta de SQL. Sin embargo, no puede modificar la vista del catálogo utilizando mandatos normales de manipulación de datos SQL. No puede crear ni desechar explícitamente la vista del catálogo. La Tabla 7 muestra el contenido de la vista del catálogo.

Tabla 7. Vista del catálogo de DB2 Text Extender

Nombre de columna	Tipo de datos	Anulable	Descripción
TABLESCHEMA	CHAR(8)	No	Esquema de la tabla a la que se aplica esta entrada.
TABLERNAME	VARCHAR(18)	No	Nombre de la tabla a la que se aplica esta entrada.
COLUMNNAME	VARCHAR(18)	Sí	Nombre de una columna que se ha habilitado en esta tabla. Este valor es nulo si se ha habilitado la tabla, pero no se ha habilitado ninguna columna.
HANDLENAME	VARCHAR(18)	Sí	Nombre de una columna de descriptores de contexto. Este valor es nulo si no hay ninguna columna habilitada en la tabla TABLESCHEMA.TABLERNAME.
INDEXNAME	CHAR(8)	No	Nombre del índice de texto creado durante la habilitación de la tabla de texto o una columna de texto.
LOGTABLE	VARCHAR(18)	No	Nombre de la tabla de registro cronológico para el índice INDEXNAME. La tabla DB2TX.LOGTABLE contiene información acerca de los documentos de texto que están planificados para la siguiente actualización del índice de texto y sucesos de error.
INDEXTYPE	VARCHAR(30)	No	Tipo de índice: LINGUISTIC, PRECISE, NGRAM.

Trabajar con la vista del catálogo de DB2 Text Extender

Tabla 7. Vista del catálogo de DB2 Text Extender (continuación)

Nombre de columna	Tipo de datos	Anulable	Descripción
MINIMUM	INTEGER	Sí	El número mínimo de peticiones de actualización de índice necesarias para que se realice una actualización de índice. Consulte el apartado “Actualización de un índice” en la página 46. Este valor es nulo si se establece la frecuencia de actualización en NONE.
DAYS	VARCHAR(15)	Sí	Los días en los que se ha de planificar una actualización. Consulte el apartado “Actualización de un índice” en la página 46. Este valor es nulo si se establece la frecuencia de actualización en NONE.
HOURS	VARCHAR(75)	Sí	Las horas en las que se ha de planificar la actualización de un índice. Consulte el apartado “Actualización de un índice” en la página 46. Este valor es nulo si se establece la frecuencia de actualización en NONE.
MINUTES	VARCHAR(185)	Sí	Los minutos en los que se ha planificado una actualización. Consulte el apartado “Actualización de un índice” en la página 46. Este valor es nulo si se establece la frecuencia de actualización en NONE.
INDEXDIRECTORY	VARCHAR(254)	No	Nombre del directorio donde se almacena el índice de texto dentro en el sistema de archivos.
UPDATEONCREATE	VARCHAR(10)	No	Los valores “update” o “noupdate”, el que se ha especificado con la opción UPDATEINDEX en ENABLE TEXT TABLE o ENABLE TEXT COLUMN o en la última CHANGE INDEX SETTINGS.
COMMONINDEX	VARCHAR(4)	No	“sí” si la tabla TABLESCHEMA.TABLENAME es una tabla de índice común. “no” si la tabla TABLESCHEMA.TABLENAME es una tabla de múltiples índices.
CCSID	SMALLINT	Sí	CCSID para la columna de texto TEXTCOLUMN especificada con el mandato de habilitación de columna de texto. Este valor es nulo si TEXTCOLUMN es nulo.
LANGUAGE	VARCHAR(30)	Sí	Nombre del diccionario utilizado al procesar la columna de texto TEXTCOLUMN. Este valor es nulo si TEXTCOLUMN es nulo.
FORMAT	VARCHAR(30)	Sí	Formato especificado para la columna de texto TEXTCOLUMN. Este valor es nulo si TEXTCOLUMN es nulo.

Tabla 7. Vista del catálogo de DB2 Text Extender (continuación)

Nombre de columna	Tipo de datos	Anulable	Descripción
FUNCTIONSCHEMA	CHAR(8)	Sí	Esquema de la función de acceso especificada en el mandato ENABLE TEXT COLUMN utilizando la opción FUNCTION. Este valor es nulo si no se especifica ninguna opción FUNCTION.
FUNCTIONNAME	VARCHAR(18)	Sí	Nombre de la función de acceso especificada en el mandato ENABLE TEXT COLUMN utilizando la opción FUNCTION. Este valor es nulo si no se especifica ninguna opción FUNCTION.
PROTOTYPEHANDLE	VARCHAR(60)	Sí	Descriptor de contexto para utilizarlo en las funciones de rendimiento. Sólo contiene el nombre de índice que es común para toda la columna de texto.
INDEXOPTION	VARCHAR(30)	Sí	Opción utilizada al crear el índice: CASE_ENABLED.
INDEXPROPERTY	VARCHAR(30)	Sí	Propiedad utilizada al crear el índice: SECTIONS_ENABLED.
NODENUMBER	INTEGER	No	Número de nodo de la partición de tabla.

Anulación del proceso de preparación de texto

Cuando se prepara texto para que DB2 Text Extender lo utilice, se realizan algunos cambios administrativos. Esta sección describe las funciones que le ayudan a anular este proceso.

Inhabilitación de una columna de texto

Resumen

Cuándo

Cuando ya no piense realizar búsquedas de texto en una columna de texto.

Mandato

DISABLE TEXT COLUMN

Autorización

ALTER, SELECT, UPDATE en la tabla

Ejemplo:

```
db2tx DISABLE TEXT COLUMN db2tx.mytable
      HANDLE commenthandle
```

Anulación del proceso de preparación de texto

Cuando inhabilita una columna de texto se produce lo siguiente:

- Si se trata de una tabla de múltiples índices, es decir, la columna tiene un índice de texto y un tabla de registro cronológico propios, se suprimen el índice, la tabla de registro cronológico y los desencadenantes de la tabla de registro cronológico.
- Si se trata de una tabla de índice común, es decir, hay un solo índice compartido por todas las columnas de texto, los términos para los documentos de esta columna se eliminan del índice común. Si se trata de la única columna de texto habilitada que queda en la tabla, se suprimen el índice, la tabla de registro cronológico y los desencadenantes de la tabla de registro cronológico.

Inhabilitación de los archivos de texto

Resumen

Cuándo

Cuando ya no piense realizar búsquedas de texto en un conjunto de archivos de texto externos.

Mandato

DISABLE TEXT FILES

Autorización

ALTER, SELECT, UPDATE en la tabla

Ejemplo:

```
db2tx DISABLE TEXT FILES db2tx.mytable
      HANDLE commenthandle
```

Cuando inhabilita los archivos de texto externos, se produce lo siguiente:

- Se suprime el índice para esta columna de descriptores de contexto.
- Se suprimen la tabla de registro cronológico y los desencadenantes.

Inhabilitación de una tabla de texto

Resumen

Cuándo

Cuando ya no piense realizar búsquedas de texto en una tabla de texto.

Mandato

DISABLE TEXT TABLE

Autorización

ALTER, SELECT, UPDATE en la tabla

Ejemplo:

```
db2tx DISABLE TEXT TABLE db2tx.mytable
```

Cuando inhabilita una tabla de texto, se produce lo siguiente:

- Si hay un índice común para las columnas de texto de la tabla, se suprime este índice. Si, en su lugar, hay índices individuales para cada columna de texto, se suprimen *todos* los índices para las columnas de texto.
- Se suprime la tabla de registro cronológico común utilizada para registrar automáticamente los documentos de texto que se han de indexar. Si, en su lugar, hay tablas de registro cronológico individuales para cada columna de texto, se suprimen todas las tablas de registro cronológico.
- Se suprimen los desencadenantes utilizados para mantener las tablas de registro cronológico.
- El contenido de las columnas de descriptores de contexto se establece en nulo.

Inhabilitación de una base de datos

Resumen

Cuándo

Cuando ya no piense realizar búsquedas de texto en esa base de datos.

Mandato

DISABLE DATABASE

Autorización

SYSADM o DBADM en la base de datos

Anulación del proceso de preparación de texto

Para inhabilitar la base de datos conectada, entre:

```
db2tx DISABLE DATABASE
```

Cuando inhabilita una base de datos, se suprimen los siguientes objetos:

- La vista del catálogo de DB2 Text Extender que se ha creado cuando se ha habilitado la base de datos
- La declaración de las funciones SQL (UDF) de DB2 Text Extender y los tipos diferenciados (UDT) de DB2 Text Extender para esta base de datos
- Todos los índices relacionados con cualquiera de estas tablas de texto o columnas de texto de la base de datos
- Las tablas de registro cronológico utilizadas para registrar automáticamente los documentos de texto que se han de indexar y los desencadenantes utilizados para mantenerlos.

Puesto que no se pueden suprimir las columnas de descriptores de contexto, y la columna de descriptores de contexto tiene un tipo diferenciado, algunos tipos diferenciados no se suprimen.

Capítulo 8. Utilización de las funciones API para buscar y examinar

Este capítulo describe cómo utilizar las funciones de búsqueda y de examen de la API de DB2 Text Extender. Para ver una descripción detallada de estas funciones, consulte el “Capítulo 13. Funciones API para la búsqueda y examen” en la página 229. Los ejemplos de los programas que utilizan las funciones API se proporcionan en el “Capítulo 14. Programa API de ejemplo” en la página 253. El mismo capítulo describe la función de examen de ejemplo `DesBrowseDocument`.

Consejo

Antes de buscar, debe leer el apartado “Tipos de búsqueda” en la página 40. Una búsqueda puede producir resultados diferentes dependiendo del tipo de índice utilizado. Utilice `GET INDEX SETTINGS` para averiguar el tipo de índice de texto que está asociado al texto que está buscando.

Configuración de la aplicación

Un programa de aplicación que utiliza la API de DB2 Text Extender es una aplicación DB2 CLI, porque algunas funciones API necesitan un descriptor de contexto de conexión de base de datos como entrada. Esto significa que las normas que se han de tener en cuenta para las aplicaciones DB2 CLI también se aplican a las aplicaciones que utilizan la API de DB2 Text Extender.

En la aplicación, incluya `des_ext.h` que se proporciona en el subdirectorio `include` del directorio de instalación de DB2 Text Extender.

Para utilizar el programa de aplicación con la API de DB2 Text Extender, enlace el programa a la API.

Enlace de una aplicación UNIX

Debe enlazar la biblioteca `libdesc1.a` a la aplicación. Esta biblioteca está en el subdirectorio `lib` del directorio de instalación de DB2 Text Extender.

Enlace de una aplicación OS/2 o Windows

Todas las funciones API de DB2 Text Extender están contenidas en la biblioteca de enlace dinámico `desclapi.dll`. Las llamadas a función API

Configuración de la aplicación

externa se resuelven en la biblioteca de importación `desclapi.lib`. Estas se proporcionan en los subdirectorios `dll` y `lib`, respectivamente del directorio de instalación de DB2 Text Extender.

Para enlazar la aplicación a la biblioteca de enlace dinámico API `desclapi.dll`, utilice `desclapi.lib` como archivo de biblioteca.

Visión general de las funciones API

Estas son las funciones de búsqueda y de examen; la primera es una función de búsqueda, las demás son funciones de examen:

- DesGetSearchResultTable
- DesGetBrowseInfo
- DesStartBrowseSession
- DesOpenDocument
- DesGetMatches
- DesCloseDocument
- DesEndBrowseSession
- DesFreeBrowseInfo.

Consejo

Muchas funciones API necesitan un descriptor de contexto de conexión (`hdbc`). El usuario debe proporcionar este descriptor de contexto utilizando la función `SQLConnect`, pero esto no impide que llame a DB2 Text Extender desde programas de SQL incorporado. La publicación *DB2 Call Level Interface Guide and Reference* describe cómo mezclar sentencias CLI con sentencias de SQL incorporado.

Puede utilizar las funciones API para:

- **Buscar texto**

En este escenario, sólo se necesita la función `DesGetSearchResultTable`. Toma como entrada un argumento de búsqueda y un nombre de columna de descriptores de contexto. Busca el texto y pone la información acerca de los documentos encontrados en una tabla del resultado que el usuario ha preparado con anterioridad.

Esta función se describe con más detalle en el apartado “Búsqueda de texto” en la página 119. Consulte también el “Capítulo 14. Programa API de ejemplo” en la página 253.

- **Examinar texto**

Utilice las funciones siguientes en el orden mostrado:

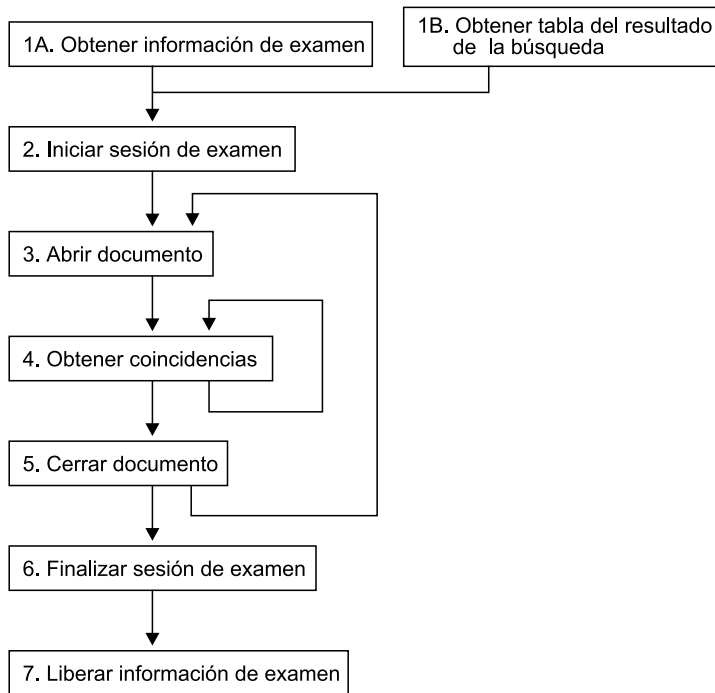


Figura 16. Orden para la utilización de las funciones API

Estas funciones obtienen información de resaltado, después inician una sesión de examen para visualizar un documento de texto y resaltar los términos encontrados.

En una sesión de examen, puede abrir y visualizar más documentos utilizando la misma información de resaltado. Estas funciones se describen con más detalle en el apartado “Examen de texto” en la página 120.

Búsqueda de texto

Hay una función API para buscar texto: la función `DesGetSearchResultTable`.

Obtención de una tabla del resultado (`DesGetSearchResultTable`)

La función `DesGetSearchResultTable` recibe un argumento de búsqueda para buscar en los documentos de una columna de texto en particular y almacena el resultado en una tabla. La tabla del resultado contiene el descriptor de contexto de cada documento encontrado. También puede contener información del rango y el número de coincidencias, dependiendo de la opción de búsqueda especificada.

Búsqueda de texto

También puede obtener esta información utilizando las funciones RANK y NO_OF_MATCHES. Este es un ejemplo:

```
INSERT INTO RESULT
SELECT COMMENTHANDLE,
       RANK(COMMENTHANDLE, '"stored procedures"'),
       NO_OF_MATCHES(COMMENTHANDLE, '"stored procedures"')
FROM DB2TX.SAMPLE
WHERE CONTAINS(COMMENTHANDLE, '"stored procedures") = 1
```

DesGetSearchResultTable sólo se puede utilizar en tablas base, pero puede ser más rápida que utilizar las funciones de DB2 Text Extender si la consulta es de sólo texto; va directamente al servidor DB2 Text Extender para obtener el rango y el número de coincidencias y sólo repite en bucle el número de documentos coincidentes encontrados. Por otro lado, en el ejemplo UDF, la función CONTAINS se llama una vez para cada fila de la tabla; después, para cada fila calificada, se llama a las funciones RANK y NO_OF_MATCHES. Para cada documento encontrado, se realizan tres búsquedas separadas.

Entrada

Los parámetros de entrada son:

- El descriptor de contexto para la conexión de base de datos
- La tabla que se ha de buscar
- El nombre de la columna descriptores de contexto que se asocia a la columna de texto en la que se ha de buscar
- Un argumento de búsqueda
- Opciones de búsqueda
- Una opción de examen (para devolver la información de examen)
- El nombre de la tabla en la que se ha de almacenar el resultado.

Salida

Si se especifica una opción de examen, esta función devuelve un puntero a una información de examen.

Examen de texto

Este grupo de funciones de la Figura 16 en la página 119 averigua los términos que se han de resaltar. Después, inicia una sesión de examen, abre un documento y obtiene la información coincidente en el formato de una corriente de datos que un programa de aplicación que llama al examinador puede analizar.

Obtención de información de examen (DesGetBrowseInfo)

La función DesGetBrowseInfo recibe un argumento de búsqueda y un descriptor de contexto. Devuelve un puntero a la información de examen

necesario para `DesStartBrowseSession`. La información de examen incluye una lista de todos los términos que se han de resaltar.

Otro método de obtener información de examen es especificar la opción de examinar en la función `DesGetSearchResultTable`.

Entrada

Los parámetros de entrada son:

- El descriptor de contexto para la conexión de base de datos
- Un descriptor de contexto
- Un argumento de búsqueda.

Salida

Esta función devuelve un puntero a la información de examen.

Inicio de una sesión de examen (`DesStartBrowseSession`)

La función `DesStartBrowseSession` inicia una sesión de examen, estableciendo el entorno necesario para examinar un documento de texto y resaltando sus coincidencias. Recibe un puntero para la información de examen, de `DesGetBrowseInfo` o de `DesGetSearchResultTable`, y devuelve un descriptor de contexto de sesión de examen para que lo utilicen las demás funciones de examen.

Entrada

El parámetro de entrada es:

- Un puntero para la información de examen de `DesGetBrowseInfo` o `DesGetSearchResultTable`
- Un ID de usuario
- Una contraseña.

Salida

Esta función devuelve un descriptor de contexto de sesión de examen.

Apertura de un documento (`DesOpenDocument`)

La función `DesOpenDocument` recibe un puntero de sesión de examen, un descriptor de contexto y una opción `DES_FAST` o `DES_EXTENDED` que indica el tipo de proceso lingüístico que se ha de utilizar para resaltar los términos encontrados. Consulte la “Etapa 2: Comparación ampliada” en la página 266. `DES_FAST` utiliza el análisis básico de texto, sin utilizar ningún diccionario para determinar los términos que se han de resaltar. `DES_EXTENDED` utiliza la comparación ampliada.

`DesOpenDocument` prepara el documento de texto que corresponde al descriptor de contexto para obtener el texto del documento y la información de resaltado y devuelve un descriptor de contexto del documento que se utiliza para llamar iterativamente a `DesGetMatches`.

Examen de texto

Entrada

Los parámetros de entrada son:

- Un descriptor de contexto de sesión de examen de DesStartBrowseSession
- Un descriptor de contexto de texto
- Una opción de coincidencia: DES_FAST o DES_EXTENDED.

Salida

Esta función devuelve un descriptor de contexto de documento que DesGetMatches y DesCloseDocument utilizan.

Obtención de coincidencias (DesGetMatches)

La función DesGetMatches devuelve un puntero a la información de resaltado para el documento de texto descrito por un descriptor de contexto de documento. La información de resaltado es una corriente de datos. Comprende el contexto del texto (un párrafo como mínimo) e información para el resaltado de texto en ese contexto. La corriente de datos se describe en el apartado “Sintaxis de la corriente de datos” en la página 238. Un programa de aplicación puede analizar la corriente de datos y procesarla utilizando el examinador propio del usuario.

DesGetMatches sólo devuelve una parte de la corriente de datos, indicando la longitud de la parte en la estructura de salida.

Una secuencia de llamadas a DesGetMatches obtiene todo el contenido del documento de texto. Cuando se alcanza el final del documento de texto, se devuelve un indicador.

Entrada

Los parámetros de entrada son:

- Un descriptor de contexto de sesión de examen
- Un descriptor de contexto de documento de DesOpenDocument.

Salida

Esta función devuelve un puntero a una estructura que contiene la parte de la corriente de datos y su longitud.

Cierre de un documento (DesCloseDocument)

La función DesCloseDocument cierra un documento de texto abierto por DesOpenDocument, y libera el almacenamiento asignado durante la devolución del texto del documento y la información de resaltado.

Entrada

Los parámetros de entrada son:

- Un descriptor de contexto de sesión de examen
- Un descriptor de contexto de documento de DesOpenDocument.

Salida

Ninguna.

Finalización de una sesión de examen (DesEndBrowseSession)

La función DesEndBrowseSession finaliza una sesión de examen iniciada por DesStartBrowseSession, y libera el almacenamiento asignado para la sesión de examen.

Entrada

El parámetro de entrada es:

Un descriptor de contexto de sesión de examen.

Salida

Ninguna.

Liberalización de la información de examen (DesFreeBrowseInfo)

La función DesFreeBrowseInfo libera el almacenamiento asignado para la información de examen por DesGetBrowseInfo.

Entrada

El parámetro de entrada es:

Un puntero para la información de examen.

Salida

Ninguna.

Parte 2. Consulta

Capítulo 9. Mandatos de preparación y administración de texto para el cliente

Mandato	Finalidad	Página
CHANGE INDEX SETTINGS	Cambia las características de un índice	129
CHANGE TEXT CONFIGURATION	Cambia los valores de configuración del texto	131
CONNECT	Le conecta a una base de datos	135
DELETE INDEX EVENTS	Suprime los sucesos de índice de una tabla de registro cronológico	137
DISABLE DATABASE	Inhabilita una base de datos para que DB2 Text Extender no la utilice	138
DISABLE TEXT COLUMN	Inhabilita una columna de texto para que DB2 Text Extender no la utilice y suprime su índice asociado	139
DISABLE TEXT FILES	Inhabilita archivos de texto para que DB2 Text Extender no los utilice y suprime su índice asociado	140
DISABLE TEXT TABLE	Inhabilita una tabla para que DB2 Text Extender no la utilice y suprime los índices asociados a la tabla	141
ENABLE DATABASE	Prepara una base de datos para que DB2 Text Extender la utilice	142
ENABLE TEXT COLUMN	Prepara una columna de texto para que DB2 Text Extender la utilice y crea un índice de texto individual para la columna	143
ENABLE TEXT FILES	Prepara archivos de texto para que DB2 Text Extender los utilice y crea un índice de texto individual para los archivos	152
ENABLE TEXT TABLE	Crea un índice de texto común para una tabla	155
GET ENVIRONMENT	Visualiza los valores actuales de las variables de entorno	160
GET INDEX SETTINGS	Visualiza las características de un índice	161
GET INDEX STATUS	Visualiza la información de estado de un índice	163
GET STATUS	Visualiza el estado habilitado de las bases de datos, las tablas y las columnas	164
GET TEXT CONFIGURATION	Visualiza los valores de configuración del texto	165
GET TEXT INFO	Visualiza la información de texto para una columna de texto	166
QUIT	Salida de la modalidad de procesador de línea de mandatos de DB2 Text Extender	167

Mandato	Finalidad	Página
REORGANIZE INDEX	Reorganiza un índice para mejorar la eficacia de la búsqueda	168
RESET INDEX STATUS	Restablece el estado de un índice para permitir que se pueda volver a utilizar	169
UPDATE INDEX	Actualiza un índice de texto	170

Este capítulo describe la sintaxis de los mandatos de preparación y administración de texto para el cliente. El “Capítulo 5. Preparación del texto para la búsqueda” en la página 57 y el “Capítulo 7. Administración” en la página 99 describen cómo utilizar estos mandatos.

Antes de utilizar estos mandatos, inicie el procesador de línea de mandatos de DB2 Text Extender entrando el mandato `db2tx`. Le pone en una modalidad de entrada interactiva en la que todos los mandatos posteriores se interpretan como mandatos DB2 Text Extender. Normalmente, iniciaría el procesador de mandatos a la vez que inicia DB2.

Para abandonar esta modalidad, entre `QUIT`.

Como en DB2, puede emitir mandatos directamente desde el indicador del sistema operativo anteponiéndoles el prefijo `db2tx` y puede incluir peticiones de procesador de línea de mandatos en un archivo de mandatos de script de shell.

Consejo

Los mandatos con el prefijo `db2tx` hacen que se establezca una conexión con la base de datos por omisión especificada en la variable de entorno `DB2DBDFT`. La siguiente secuencia de mandatos no habilita la base de datos `MYDATABASE`, sino que habilita la base de datos por omisión.

```
db2tx CONNECT TO MYDATABASE
db2tx ENABLE DATABASE
```

CHANGE INDEX SETTINGS

Este mandato cambia las características de un índice **después** de haber habilitado la base de datos.

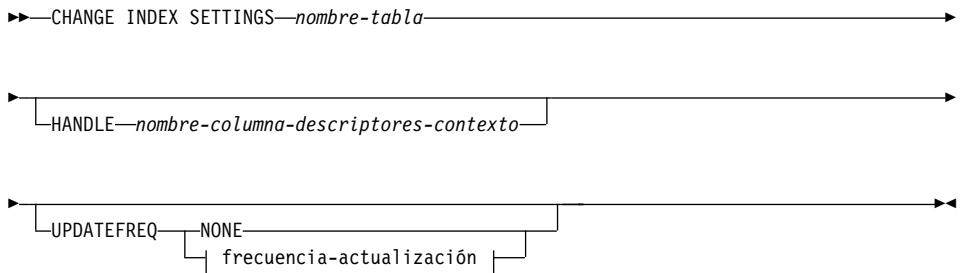
Para cambiar los valores por omisión que se utilizan al habilitar una base de datos por primera vez, consulte el apartado “CHANGE TEXT CONFIGURATION” en la página 131.

Autorización

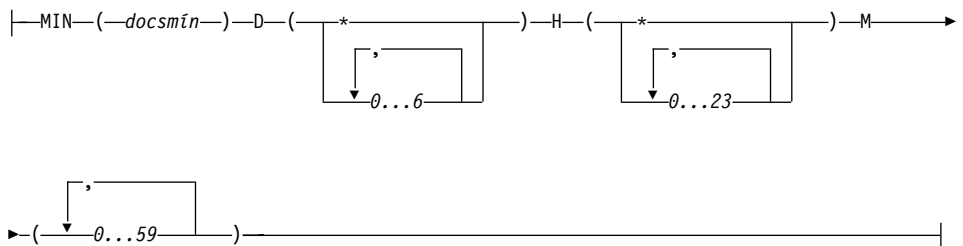
Debe tener como mínimo uno de los siguientes privilegios para la tabla:

- Privilegio ALTER
- Privilegio SELECT
- Privilegio UPDATE.

Sintaxis del mandato



frecuencia-actualización:



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna de texto cuya frecuencia de actualización del índice se ha de cambiar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

Mandato CHANGE INDEX SETTINGS

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuya frecuencia de actualización del índice se ha de cambiar. Es necesario si la columna de texto tiene su propio índice, es decir, si el índice se ha creado utilizando el mandato ENABLE TEXT COLUMN.

Sin embargo, si el índice se ha creado utilizando el mandato ENABLE TEXT TABLE, es decir, la tabla tiene un índice e texto para todas las columnas de texto, se pasa por alto esta palabra clave.

UPDATEFREQ frecuencia-actualización

La frecuencia de actualización del índice en función de cuándo se ha de realizar la actualización y del número mínimo de documentos de texto que deben estar en cola en la tabla de registro cronológico. Si no hay suficientes documentos de texto en la tabla de registro cronológico en el día y la hora dadas, el índice no se actualiza.

La sintaxis se describe en el apartado “Actualización de un índice” en la página 46.

NONE

No se realizan más actualizaciones de índice. Está pensado para una columna de texto en la que no habrán más cambios.

Si no especifica la palabra clave UPDATEFREQ, los valores de frecuencia se dejan sin cambios.

CHANGE TEXT CONFIGURATION

Este mandato cambia los valores por omisión de la configuración del texto que se utilizan al habilitar una base de datos. Se tratan de los valores de *configuración del texto*. Los valores de la configuración del texto inicial cuando se instala DB2 Text Extender se describen en el apartado “Valores de la configuración del texto” en la página 19.

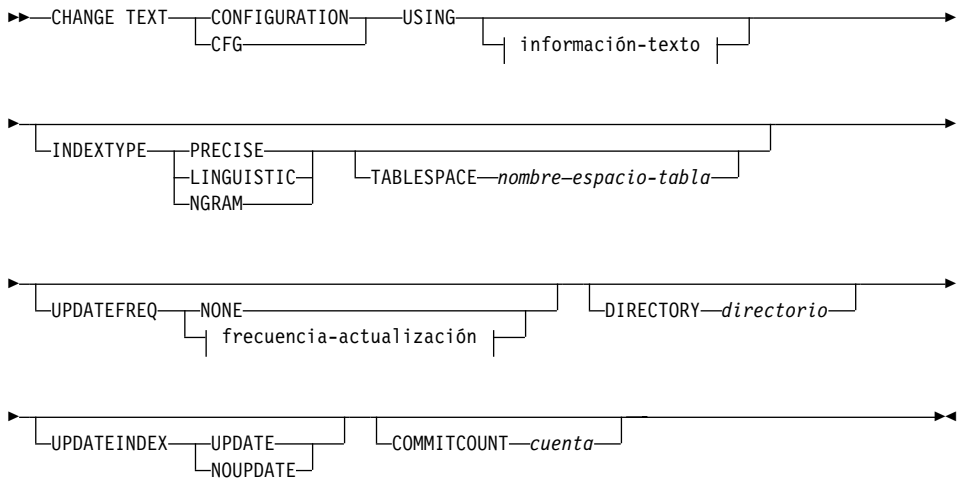
Para cambiar estos valores para una base de datos en particular **después** de haber habilitado la misma, consulte el apartado “CHANGE INDEX SETTINGS” en la página 129.

Autorización

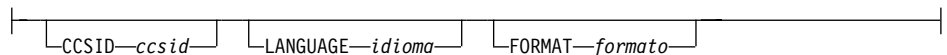
El usuario debe tener:

- Autorización SYSADM
- Autorización DBADM.

Sintaxis del mandato

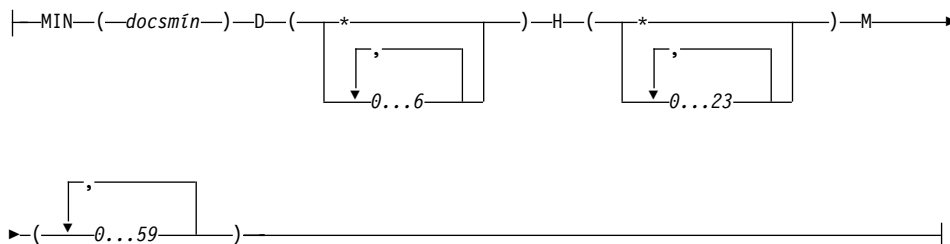


información-texto:



frecuencia-actualización:

Mandato CHANGE TEXT CONFIGURATION



Parámetros del mandato

INDEXTYPE

Para cambiar el tipo de índice por omisión, elija uno de los siguientes. Para obtener más información, consulte el apartado “Tipos de búsqueda” en la página 40.

PRECISE

Los términos se indexan y se buscan exactamente tal como aparecen en los documentos de texto.

LINGUISTIC

Los términos se procesan lingüísticamente antes de indexarse. También se procesan lingüísticamente los términos de búsqueda antes de que empiece la búsqueda.

NGRAM

Los términos se indexan analizando los juegos de caracteres en lugar de utilizar un diccionario. Este tipo de diccionario es obligatorio si los documentos que se indexan contienen caracteres DBCS, aunque también pueda utilizarse un índice Ngram para documentos SBCS.

TABLESPACE nombre-espacio-tabla

Especifique el nombre de un espacio de tabla existente. El espacio de tabla se utiliza para contener las tablas específicas de índice que DB2 Text Extender crea (por ejemplo, las tablas de registro cronológico). Para tablas grandes, utilice un espacio de tabla separado. Si no especifica un espacio de tabla, las tablas se crean en el espacio de tabla por omisión de DB2.

UPDATEFREQ frecuencia-actualización

La frecuencia de actualización del índice en función de cuándo se ha de realizar la actualización y del número mínimo de documentos de texto que deben estar en cola en la tabla de registro cronológico. Si no hay suficientes documentos de texto en la tabla de registro cronológico en el día y la hora dadas, el índice no se actualiza.

La sintaxis se describe en el apartado “Actualización de un índice” en la página 46.

NONE

No se realizan más actualizaciones de índice. Está pensado para una columna de texto en la que no habrán más cambios.

DIRECTORY directorio

El directorio en el que se ha de almacenar el índice de texto.

UPDATEINDEX

Una palabra clave que determina si los documentos de texto se indexan inmediatamente después de completarse el mandato que utiliza esta opción, sin esperar la próxima indexación periódica establecida por UPDATEFREQ. Estos mandatos son ENABLE TEXT COLUMN y ENABLE TEXT FILES.

UPDATE

La indexación de documentos de texto se produce inmediatamente después de que se haya completado el mandato.

NOUPDATE

La indexación se produce en un momento establecido por los valores de frecuencia de actualización especificados en este mandato por UPDATEFREQ o por el valor de configuración del texto.

COMMITCOUNT cuenta

Un valor de 500 a 1 000 000 que indica el número de inserciones o actualizaciones después de las cuales se emite una sentencia de compromiso intermedia de DB2 UDB . Esto puede evitar una situación en la que no haya suficiente espacio de registro cronológico al habilitar tablas o columnas grandes o un gran número de archivos externos.

CCSID ccsid

El identificador de juego de caracteres codificado que se ha de utilizar al indexar documentos de texto.

Para obtener información acerca de los CCSID que pueden soportarse consulte el apartado “CCSID” en la página 35.

LANGUAGE idioma

El idioma en el que está escrito el texto. Esto determina el diccionario que se va a utilizar al indexar los documentos de texto y al buscar en los documentos de texto. El “Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto” en la página 255 describe cómo se utilizan los diccionarios.

Los idiomas soportados se listan en la Tabla 5 en la página 54.

FORMAT formato

El tipo de documento de texto almacenado, como, por ejemplo

Mandato CHANGE TEXT CONFIGURATION

WordPerfect o ASCII. DB2 Text Extender necesita esta información para indexar los documentos. Los formatos de documento soportados se listan en el apartado “Formatos de documentos que se soportan” en la página 31.

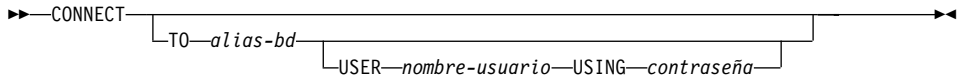
CONNECT

Este mandato conecta DB2 Text Extender a una base de datos .

Autorización

Autorización CONNECT para la base de datos .

Sintaxis del mandato



Parámetros del mandato

TO alias-bd

La base de datos a la que se ha de conectar. Debe ser el nombre de un alias de base de datos, no un nombre de base de datos.

USER nombre-usuario

Si no se ha especificado ningún nombre de usuario, se recupera del sistema operativo.

USING contraseña

Sólo es necesaria una contraseña si se especifica un nombre de usuario.

CONNECT ningún operando

Si no especifica ningún operando y no hay ninguna base de datos conectada, CONNECT establece una conexión implícita con la base de datos especificada en la variable de entorno DB2DBDFT. Si no especifica ningún operando y hay una base de datos conectada, CONNECT visualiza la información acerca de la base de datos actual.

Notas del uso

Cuando emita un mandato DB2TX sin estar todavía conectado a una base de datos , DB2 Text Extender le conecta a la base de datos especificada en la variable de entorno DB2DBDFT.

Para conectarse explícitamente a una base de datos en particular, emita el mandato CONNECT TO.

Sólo puede estar conectado a una base de datos cada vez; se denomina la base de datos actual. En modalidad interactiva, una conexión dura hasta que otra sentencia CONNECT TO cambia la base de datos o hasta que se emite un mandato QUIT. En modalidad de línea de mandatos, un mandato CONNECT no surte ningún efecto.

Mandato CONNECT

Si está utilizando un cliente OS/2 y es la primera vez que se conecta a una base de datos después de haber iniciado el sistema, UPM (Gestión de perfiles de usuario) le solicita un ID de usuario y una contraseña. Si ya ha iniciado la sesión, UPM toma el ID de usuario y la contraseña de esta sesión e intenta conectarse al servidor remoto. Consulte el apartado “Inicio del procesador de línea de mandatos de DB2 Text Extender” en la página 59.

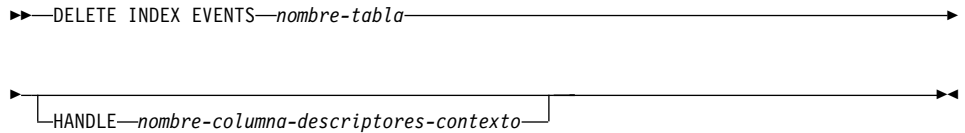
DELETE INDEX EVENTS

Este mandato suprime los sucesos de indexación de la tabla de registro cronológico de un índice para una columna de descriptores de contexto o tabla determinadas.

Autorización

Ninguna.

Sintaxis del mandato



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada cuyos sucesos de error se han de suprimir de la tabla de registro cronológico. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuyos mensajes se han de suprimir de la tabla de registro cronológico.

Notas del uso

Si se proporciona un nombre de columna de descriptores de contexto, sólo se suprimen los sucesos de indexación para esta columna.

El usuario limpia las tablas de registro cronológico, debe suprimir los sucesos de índice después de haber comprobado la razón del suceso y, posiblemente, haber eliminado la fuente del error.

Mandato DISABLE DATABASE

DISABLE DATABASE

Este mandato restablece cualquier trabajo de preparación realizado por DB2 Text Extender para una base de datos e inhabilita todas las tablas de texto para que DB2 Text Extender no las utilice.

Autorización

El usuario debe tener:

Autorización SYSADM

Autorización DBADM.

Sintaxis del mandato

►—DISABLE DATABASE—◄

Parámetros del mandato

Ninguna.

Notas del uso

Este mandato restablece la base de datos conectada para que DB2 Text Extender ya no realice ninguna búsqueda en ella; es decir, inhabilita todas las tablas de texto y las columnas de texto de DB2 Text Extender de la base de datos. Se restablecen todas las modificaciones que se han realizado en la base de datos para habilitar las tablas de texto, las columnas de texto y los archivos externos de DB2 Text Extender: se suprimen todos los índices de texto relacionados, la vista del catálogo TEXTCOLUMNS de DB2 Text Extender se suprime de la base de datos y se suprimen todos los desencadenantes de DB2 Text Extender.

DISABLE TEXT COLUMN

Este mandato inhabilita una columna de texto para que DB2 Text Extender no la utilice.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

- Privilegio ALTER
- Privilegio SELECT
- Privilegio UPDATE.

Sintaxis del mandato

►►—DISABLE TEXT COLUMN—*nombre-tabla*—HANDLE—*nombre-columna-descriptores-contexto*—►►

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna que se ha de inhabilitar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto que se ha de inhabilitar para que DB2 Text Extender no la utilice.

Notas del uso

Se suprime el índice.

Se suprime la tabla de registro cronológico utilizada para registrar los cambios de la columna de descriptores de contexto (inserciones, actualizaciones y supresiones).

Se suprimen los desencadenantes que graban entradas en la tabla de registro cronológico.

La columna de descriptores de contexto no se cambia.

Mandato DISABLE TEXT FILES

DISABLE TEXT FILES

Este mandato inhabilita un conjunto de archivos de texto externos para que DB2 Text Extender no los utilice.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio ALTER

Privilegio SELECT

Privilegio UPDATE.

Sintaxis del mandato

►►—DISABLE TEXT FILES—*nombre-tabla*—HANDLE—*nombre-columna-descriptores-contexto*—►►

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna de descriptores de contexto para los archivos de texto externos que se ha de inhabilitar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto que se ha de inhabilitar para que DB2 Text Extender no la utilice.

Notas del uso

Se suprime el índice.

Se suprime la tabla de registro cronológico utilizada para registrar los cambios de la columna de descriptores de contexto (inserciones, actualizaciones y supresiones). También se suprimen los desencadenantes que graban entradas en la tabla de registro cronológico.

DISABLE TEXT TABLE

Este mandato inhabilita todas las columnas de texto de una tabla para que DB2 Text Extender no las utilice.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio ALTER

Privilegio SELECT

Privilegio UPDATE.

Sintaxis del mandato

►►—DISABLE TEXT TABLE—*nombre-tabla*—◄◄

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna que se ha de inhabilitar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

Notas del uso

Este mandato hace que DB2 Text Extender no pueda utilizar ninguna de las columnas de texto de la tabla.

Si ENABLE TEXT COLUMN inhabilitase individualmente las columnas de texto de esta tabla, se suprimirían todos sus índices de texto asociados. (Para inhabilitar las columnas de texto y suprimir individualmente sus índices de texto asociados, utilice el mandato DISABLE TEXT COLUMN.) Si las columnas de texto de esta tabla se hubiesen habilitado a la vez por ENABLE TEXT TABLE, sólo habría un índice común para todas las columnas de texto. Este mandato suprime el índice común.

Se suprimen las tablas de registro cronológico utilizadas para registrar los cambios en las columnas de texto (inserciones, actualizaciones y supresiones). Se suprimen los desencadenantes que graban entradas en la tabla de registro cronológico.

ENABLE DATABASE

Este mandato habilita la base de datos actual para que almacene los datos de texto.

Autorización

El usuario debe tener:

Autorización SYSADM

Autorización DBADM.

Sintaxis del mandato

►—ENABLE DATABASE—◄

Parámetros del mandato

Ninguno.

Notas del uso

Debe estar conectado a una base de datos , explícita o implícitamente antes de emitir este mandato (consulte “CONNECT” en la página 135).

Este mandato prepara la base de datos conectada para que DB2 Text Extender la utilice. Es un paso obligatorio para poder habilitar una tabla de texto o una columna de texto de DB2 Text Extender en la base de datos .

ENABLE DATABASE crea una vista del catálogo de DB2 Text Extender denominada DB2TX.TEXTINDEXES, que se describe en el apartado “Trabajar con la vista del catálogo de DB2 Text Extender” en la página 111 y una vista del catálogo denominada DB2TX.TEXTCOLUMNS que se utiliza para las consultas de “rendimiento”.

También crea los valores de configuración del texto, descritos en el capítulo “Valores de la configuración del texto” en la página 19.

También se realiza algún otro trabajo como, por ejemplo, la declaración de tipos diferenciados de DB2 Text Extender y las funciones de DB2 Text Extender.

ENABLE TEXT COLUMN

Este mandato habilita una columna de texto para que DB2 Text Extender la utilice.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio ALTER

Privilegio SELECT

Privilegio UPDATE.

Sintaxis del mandato

►—ENABLE TEXT COLUMN—*nombre-tabla*—*nombre-columna-texto*—►

►—HANDLE—*nombre-columna-descriptores-contexto*—►
 └─FUNCTION—*nombre-función*—┘

└─información-texto ┘ └─características-índice ┘

└─UPDATEINDEX—┘ └─UPDATE—┘ └─NOUPDATE—┘ └─TABLESPACE—*nombre-espacio-tabla*—┘

└─COMMITCOUNT—*cuenta*—┘

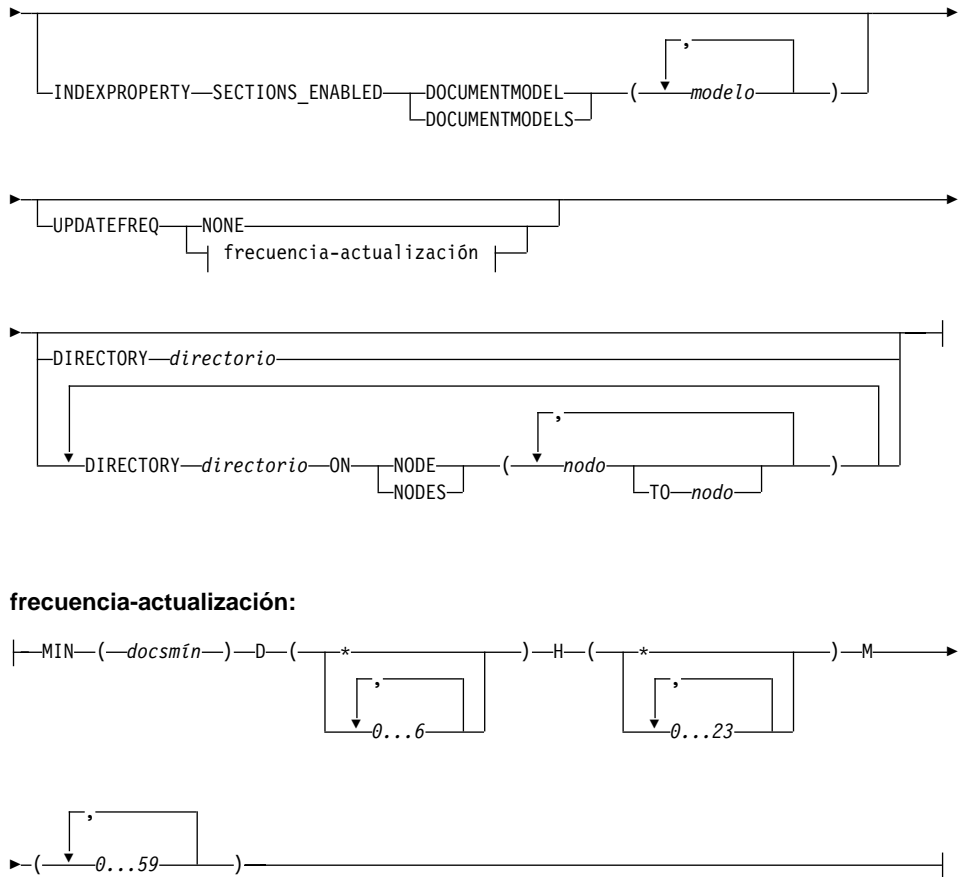
información-texto:

└─CCSID—*ccsid*—┘ └─LANGUAGE—*idioma*—┘ └─FORMAT—*formato*—┘

características-índice:

└─INDEXTYPE—┘ └─PRECISE—┘ └─LINGUISTIC—┘ └─NGRAM—┘ └─INDEXOPTION—┘ └─CASE_ENABLED—┘ └─NORMALIZED—┘

Mandato ENABLE TEXT COLUMN



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna que se ha de habilitar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

nombre-columna-texto

El nombre de la columna que se ha de habilitar para que DB2 Text Extender la utilice. Esta columna debe ser de tipo CHAR, VARCHAR, LONG VARCHAR, CLOB, DBCLOB, GRAPHIC, VARGRAPHIC o LONG VARGRAPHIC. Si el tipo de documento no es ninguno de estos, utilice FUNCTION para convertir el tipo de documento.

FUNCTION nombre-función

El nombre de una función definida por el usuario que DB2 Text Extender utilizará para acceder a los documentos de texto que están

en la columna que no es de tipo CHAR, VARCHAR, LONG VARCHAR, CLOB, DBCLOB, GRAPHIC, VARGRAPHIC ni LONG VARGRAPHIC. Consulte el apartado “Habilitación de columnas de texto de un tipo de datos no soportado” en la página 71 para obtener más información.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto que se ha de añadir a la tabla para que la utilicen las funciones de DB2 Text Extender.

CCSID ccsid

El identificador de juego de caracteres codificado que se ha de utilizar al indexar documentos de texto.

Si especifica un CCSID al habilitar una columna de texto para un índice Ngram, el CCSID debe ser el mismo CCSID de la base de datos y el CCSID utilizado durante la búsqueda (el CCSID de la base de datos) debe coincidir con este CCSID.

Para buscar el CCSID por omisión, utilice:

```
db2tx get text cfg
```

El valor por omisión de instalación es el CCSID de la base de datos .

Si no se especifica esta palabra clave, se utiliza el CCSID especificado en los valores de configuración del texto. Los cambios posteriores en los valores de configuración del texto se pasan por alto; el valor utilizado es el que existía en el momento en que se ha habilitado la columna, no el que existe cuando se indexan los documentos de texto.

Para obtener información acerca de otros CCSID que pueden soportarse, consulte el apartado “CCSID” en la página 35.

LANGUAGE idioma

El idioma en el que está escrito el texto. Esto determina el diccionario que se va a utilizar al indexar los documentos de texto y al buscar en los documentos de texto. El “Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto” en la página 255 describe cómo se utilizan los diccionarios.

Esta palabra clave especifica el idioma una vez para toda la columna. Puede alterar temporalmente este valor para los documentos de texto insertados individualmente utilizando la función INIT_TEXT_HANDLE en una sentencia INSERT.

Si no se especifica esta palabra clave, se utiliza el idioma especificado en los valores de configuración de texto. Los cambios posteriores en los valores de configuración del texto se pasan por alto; el valor

Mandato ENABLE TEXT COLUMN

utilizado es el que existía en el momento en que se ha habilitado la columna, no el que existe cuando se indexan los documentos de texto.

Los idiomas soportados se listan en la Tabla 5 en la página 54.

FORMAT formato

El tipo de documento de texto almacenado, como, por ejemplo WordPerfect o ASCII. DB2 Text Extender necesita esta información para indexar los documentos. Los formatos de documento soportados se listan en el apartado “Formatos de documentos que se soportan” en la página 31.

Los formatos de documentos soportados para los documentos estructurados son:

- **ASCII_SECTIONS**

Los documentos que tienen el formato ASCII_SECTIONS no pueden contener secciones anidadas. (Para obtener información acerca de las secciones anidadas, consulte el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48.)

El código inicial para una sección finaliza en el siguiente código inicial.

- **HTML**

Se proporciona un archivo de modelos de documentos para los documentos HTML. Contiene un subconjunto de las definiciones HTML estándar, que el usuario puede modificar si es necesario. Los documentos HTML contienen secciones anidadas.

- **XML**

El proceso de documentos XML incluye la evaluación de la Definición de tipo de documento (DTD). El modelo asignado al documento se comprueba con la DTD. Si los códigos definidos en el archivo de modelos de documentos no están definidos en la DTD, no se indexa el documento. Si no se ha definido ningún modelo para una DTD reconocida, el documento no se indexará. Los documentos XML pueden contener secciones anidadas.

Para estos formatos, debe especificar la información de estructura en un archivo de modelos de documentos. Consulte el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48. Si se especifican los formatos TDS e INDEXPROPERTY SECTION_ENABLED, se supone que el formato del documento es ASCII_SECTIONS.

Los identificadores que no están definidos en el archivo de modelos se indexan de manera normal, de acuerdo al tipo de índice.

Esta palabra clave especifica el formato una vez para toda la columna. Puede alterar temporalmente este valor para los documentos de texto insertados individualmente utilizando la función INIT_TEXT_HANDLE en una sentencia INSERT.

Si no se especifica esta palabra clave, se utiliza el formato especificado en los valores de configuración del texto. Los cambios posteriores en los valores de configuración del texto se pasan por alto; el valor utilizado es el que existía en el momento en que se ha habilitado la columna, no el que existe cuando se indexan los documentos de texto.

INDEXTYPE

El tipo de índice que se ha de crear. Para obtener más información, consulte el apartado “Tipos de búsqueda” en la página 40.

PRECISE

Los términos se indexan y se buscan exactamente tal como aparecen en los documentos de texto.

LINGUISTIC

Los términos se procesan lingüísticamente antes de indexarse. También se procesan lingüísticamente los términos de búsqueda antes de que empiece la búsqueda.

NGRAM

Los términos se indexan analizando los juegos de caracteres en lugar de utilizar un diccionario. Este tipo de índice es obligatorio si los documentos que está indexando contienen caracteres DBCS, aunque también se puede utilizar un índice Ngram para documentos SBCS.

Si no especifica la palabra clave INDEXTYPE, se utiliza el valor de los valores de configuración del texto.

No se da soporte a los documentos en formato XML para índices Ngram.

INDEXOPTION

Las opciones que se han de utilizar al crear el índice.

CASE_ENABLED

Esta opción está disponible **únicamente para índices Ngram**. Normalmente, los índices Ngram no permiten una búsqueda sensible a las mayúsculas y minúsculas. Con la especificación de CASE_ENABLED, asegura que los documentos se indexen de modo que sea posible una búsqueda sensible a las mayúsculas y minúsculas. Para obtener más información, consulte el apartado “Efectuar una búsqueda similar o una búsqueda en documentos DBCS” en la página 43.

Mandato ENABLE TEXT COLUMN

NORMALIZED

Esta opción está disponible **únicamente para índices exactos**. Un índice exacto normalizado difiere de un índice exacto en que:

- No es sensible a las mayúsculas y minúsculas; todas las palabras, excepto las que están completamente en mayúsculas, se convierten a minúsculas.
- Las palabras escritas con todas las letras en mayúsculas no son objeto de filtro de palabra de parada; por ejemplo, se indexa la abreviatura UK.
- Los términos de búsqueda del idioma inglés se pueden ampliar para que incluyan formatos de lemas que utilizan un algoritmo heurístico, de modo que la búsqueda de house también busque houses.

INDEXPROPERTY SECTIONS_ENABLED DOCUMENTMODEL(S) *modelo*

Propiedades de un tipo de índice seleccionado.

SECTIONS_ENABLED especifica que el tipo de índice seleccionado puede contener información acerca de la estructura de documentos.

DOCUMENTMODEL/DOCUMENTMODELS *modelo* especifica el modelo o modelos que se han de asociar como valor por omisión para los documentos que se han de indexar. Debe especificarse un nombre de modelo si se utiliza la propiedad de índice SECTIONS_ENABLED. Si se especifica una lista de modelos, se utiliza el primer modelo como modelo por omisión para el índice. El modelo por omisión se utiliza durante la indexación si el documento no hace ninguna referencia a un modelo o si no se especifica ningún modelo durante la búsqueda.

Los caracteres que pueden utilizarse para el nombre de modelo son a-z, A-Z y 0-9.

El nombre de modelo especificado debe corresponder a una definición de modelo del archivo de definiciones de modelos `desmodel.ini`.

Para cambiar el modelo o modelos asociados a un índice,

1. Utilice DISABLE TEXT COLUMN para inhabilitar el índice
2. Utilice ENABLE TEXT COLUMN para volver a indexar los documentos, especificando nombres de modelos de documentos diferentes.

UPDATEFREQ *frecuencia-actualización*

La frecuencia de actualización del índice en función de cuándo se ha de realizar la actualización y del número mínimo de documentos de texto que deben estar en cola en la tabla de registro cronológico. Si no hay suficientes documentos de texto en la tabla de registro cronológico en el día y la hora dadas, el índice no se actualiza.

La sintaxis se describe en el apartado “Actualización de un índice” en la página 46 .

Si no especifica UPDATEFREQ, se utiliza la frecuencia por omisión especificada en los valores de configuración del texto.

Consejo

Si tiene muchas tablas, considere la posibilidad de evitar la utilización de los valores por omisión. Creando valores de frecuencia de actualización individuales para las tablas puede evitar la indexación simultánea de todas las tablas que provoca una carga prolongada innecesaria en los recursos del sistema.

NONE

No se realizan más actualizaciones de índice. Está pensado para una columna de texto en la que no habrán más cambios.

Estos valores de frecuencia de actualización se pasan por alto si ya se han establecido para toda la tabla por ENABLE TEXT TABLE.

DIRECTORY directorio

La vía de acceso del directorio en el que se ha de almacenar el índice de texto. La vía de acceso especificada se concatena con “txinsnnn” donde *nnn* es el número de nodo.

Se trata de un directorio del sistema en el que se ejecuta el servidor DB2 Text Extender. Si todavía no existe el directorio, se crea. Si ya existe, asegúrese en las plataformas UNIX que el directorio tenga permisos de lectura/grabación para el propietario de la instancia y compruebe el 'ID de usuario con barreras' si se ha especificado uno diferente del especificado cuando se ha creado la instancia de DB2.

Si no especifica la palabra clave DIRECTORY, se utiliza el valor de DIRECTORY de los valores de configuración del texto.

Este valor se pasa por alto si ya se ha establecido para toda la tabla por ENABLE TEXT TABLE.

Si está utilizando bases de datos particionadas, puede especificar un nombre de vía de acceso del índice común para todos los nodos o para un rango de nodos, o vías de acceso de directorios de índices diferentes para cada nodo.

ON NODE **nodo** [TO **nodo**]

El número del nodo o del rango de nodos al que se está asignando el nombre de vía de acceso del directorio.

UPDATEINDEX

Una palabra clave que determina si los documentos de texto asociados

Mandato ENABLE TEXT COLUMN

a esta columna de descriptores de contexto se indexan inmediatamente después de que se haya completado este mandato, sin esperar a la próxima indexación periódica establecida por UPDATEFREQ.

UPDATE

La indexación de los documentos de texto se produce inmediatamente después de que se haya completado este mandato.

NOUPDATE

La indexación se produce en un momento establecido por los valores de frecuencia de actualización especificados en este mandato por UPDATEFREQ o por el valor de configuración del texto.

Si no especifica esta palabra clave, se toma el valor de los valores de configuración del texto.

TABLESPACE nombre-espacio-tabla

El nombre del espacio de tabla para el índice que se crea internamente en la columna de descriptores de contexto. El espacio de tabla debe haberse creado previamente.

COMMITCOUNT cuenta

Un valor de 500 a 1 000 000 que indica el número de inserciones o actualizaciones después de las cuales DB2 UDB debe emitir una sentencia de compromiso intermedia. Esto puede evitar una situación en la que no haya suficiente espacio de registro cronológico al habilitar tablas o columnas grandes o un gran número de archivos externos.

Notas del uso

Este mandato añade una columna de descriptores de contexto a la tabla DB2 especificada. Cada columna de descriptores de contexto se asocia a una columna de texto y la utilizan las funciones de DB2 Text Extender.

Si aún no se ha habilitado esta tabla para crear un índice común, se crea un índice que está asociado a esta columna de texto.

También, se crea una tabla de registro cronológico en la base de datos. La tabla de registro cronológico se utiliza para registrar los cambios en la columna de texto, es decir, las inserciones, las actualizaciones y las supresiones. Se definen desencadenantes de inserción, actualización y supresión para la columna de texto para mantener actualizada automáticamente la tabla de registro cronológico.

Mandato ENABLE TEXT COLUMN

Si la columna de texto que está habilitando pertenece a una tabla que forma parte de un grupo de nodos de múltiples nodos, el directorio de índices que especifique debe estar disponible en todos los nodos físicos. Si utiliza el directorio por omisión especificado en la configuración del texto, asegúrese de que la vía de acceso esté disponible en todos los nodos del grupo de nodos. Si no es conveniente, puede especificar una vía de acceso específica para cada nodo en el mandato ENABLE TEXT COLUMN.

Si cambia la configuración de los nodos de un grupo de nodos que contiene una tabla que está habilitada para DB2 Text Extender, debe volver a indexar la tabla.

Consejo

Si se queda sin espacio de registro cronológico en este paso, consulte el apartado “Habilitación de una columna de texto en una tabla grande” en la página 70 para ver las posibles soluciones.

Mandato ENABLE TEXT FILES

ENABLE TEXT FILES

Este mandato permite que DB2 Text Extender busque en los archivos de texto que no están en una base de datos DB2 UDB .

Consejo

Este mandato no se puede utilizar si las columnas de texto de la tabla comparten un índice común, tal como se describe en el apartado “Habilitación de una tabla de texto (opcional)” en la página 62.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio ALTER

Privilegio SELECT

Privilegio UPDATE.

Sintaxis del mandato

►—ENABLE TEXT FILES—*nombre-tabla*—HANDLE—*nombre-columna-descriptores-contexto*—►

┌──┴──┐ ┌──┴──┐
| información-texto | | características-índice |
└──┬──┘ └──┬──┘

┌──┴──┐ ┌──┴──┐ ┌──┴──┐
| UPDATEINDEX | | UPDATE | | TABLESPACE—*nombre-espacio-tabla* |
└──┬──┘ └──┬──┘ └──┬──┘
| NOUPDATE |

┌──┴──┐
| COMMITCOUNT—*cuenta* |
└──┬──┘

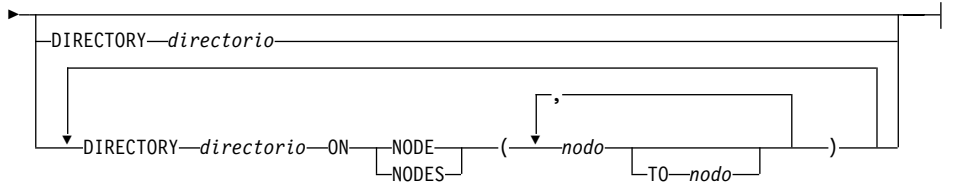
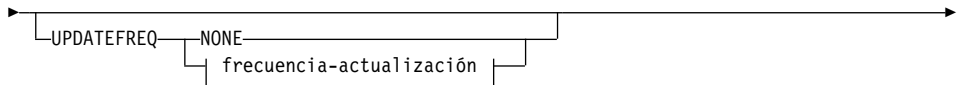
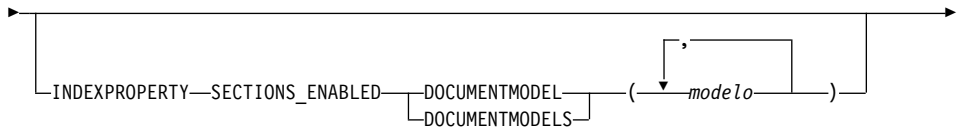
información-texto:

┌──┴──┐ ┌──┴──┐ ┌──┴──┐
| CCSID—*ccsid* | | LANGUAGE—*idioma* | | FORMAT—*formato* |
└──┬──┘ └──┬──┘ └──┬──┘

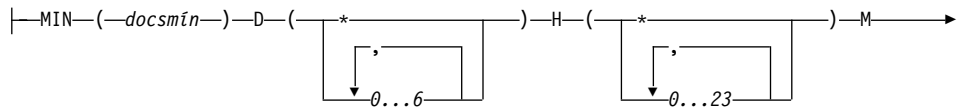
características-índice:

┌──┴──┐ ┌──┴──┐ ┌──┴──┐
| INDEXTYPE | | PRECISE | | INDEXOPTION | | CASE_ENABLED |
└──┬──┘ └──┬──┘ └──┬──┘ └──┬──┘
| LINGUISTIC | | NGRAM | | NORMALIZED |

Mandato ENABLE TEXT FILES



frecuencia-actualización:



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que se ha de asociar a los archivos de texto externos que se han de indexar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

Para ver una descripción de los parámetros del mandato, consulte “ENABLE TEXT COLUMN” en la página 143.

Notas del uso

Este mandato añade una columna de descriptores de contexto a la tabla DB2 especificada. Cada columna de descriptores de contexto se asocia a una colección de archivos de texto externos y la utilizan las funciones de DB2 Text Extender. Se crea un índice que se asocia a estos archivos.

Mandato ENABLE TEXT FILES

Después de haber habilitado los archivos de texto, inicialice los descriptores de contexto de la columna de descriptores de contexto utilizando INIT_TEXT_HANDLE. Después, rellene el índice utilizando UPDATE INDEX.

No puede volver a utilizar un nombre de columna de contexto si ese nombre ya se ha utilizado antes en ENABLE TEXT FILES para identificar una columna de descriptores de contexto de una columna de texto.

Se crea una tabla de registro cronológico para registrar los cambios en los archivos, pero debe activar los desencadenantes manualmente para registrar esos cambios porque DB2 UDB no tiene los archivos bajo su control y, por lo tanto, no conoce esos cambios. Consulte el apartado “Actualización de un índice para archivos externos” en la página 100 para ver una descripción de cómo hacer esto.

Si se queda sin espacio de registro cronológico en este paso, consulte el apartado “Habilitación de una columna de texto en una tabla grande” en la página 70 para ver las posibles soluciones.

ENABLE TEXT TABLE

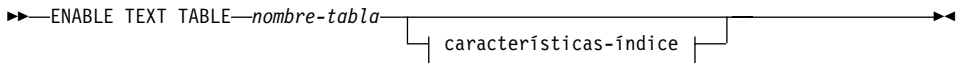
Crea un índice común para que lo utilice cualquiera de las columnas de texto de la tabla que se habilite con posterioridad. Entonces, se trata de una tabla de índice común. Una tabla que no se habilita de esta manera, en que las columnas de texto que luego se habilitan crean sus propios índices individuales, es una tabla de múltiples índices.

Autorización

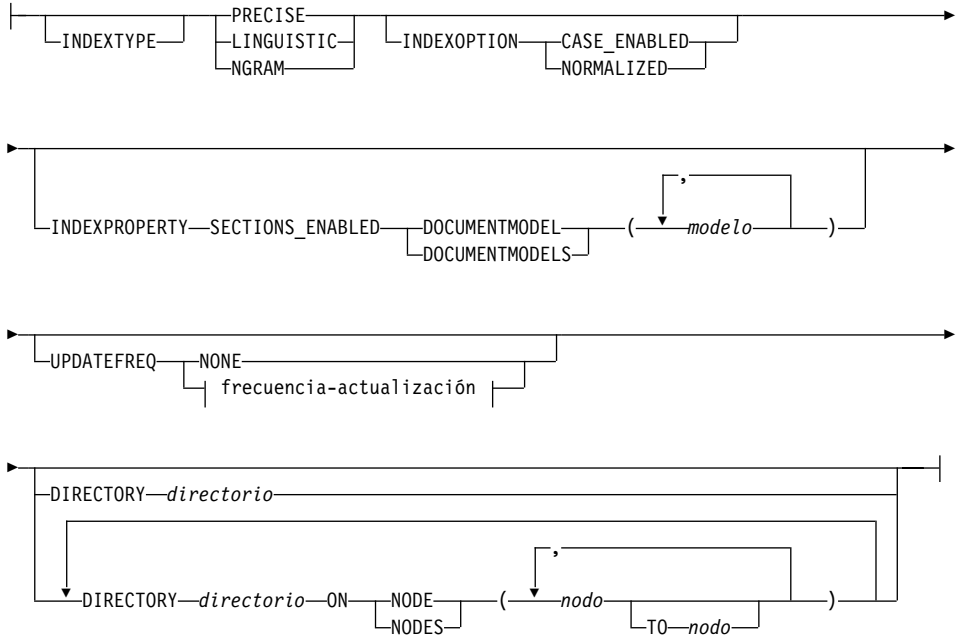
Debe tener como mínimo uno de los siguientes privilegios para la tabla:

- Privilegio ALTER
- Privilegio SELECT
- Privilegio UPDATE.

Sintaxis del mandato

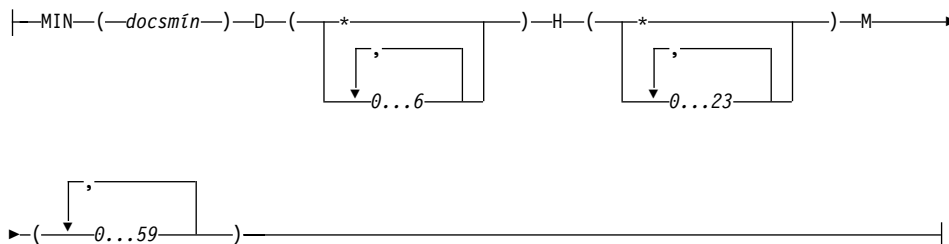


características-índice:



Mandato ENABLE TEXT TABLE

frecuencia-actualización:



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto que se ha de habilitar en la base de datos conectada. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

INDEXTYPE

El tipo de índice que se ha de crear. Para obtener más información, consulte el apartado “Tipos de búsqueda” en la página 40.

PRECISE

Los términos se indexan y se buscan exactamente tal como aparecen en los documentos de texto.

LINGUISTIC

Los términos se procesan lingüísticamente antes de indexarse. También se procesan lingüísticamente los términos de búsqueda antes de que empiece la búsqueda.

NGRAM

Los términos se indexan analizando los juegos de caracteres en lugar de utilizar un diccionario. Este tipo de diccionario es obligatorio si los documentos que se indexan contienen caracteres DBCS, aunque también pueda utilizarse un índice Ngram para documentos SBCS.

Si no especifica la palabra clave INDEXTYPE, se utiliza la configuración del texto.

INDEXOPTION

Las opciones que se han de utilizar al crear el índice.

CASE_ENABLED

Esta opción está disponible **únicamente para índices Ngram**. Normalmente, los índices Ngram no permiten una búsqueda sensible a las mayúsculas y minúsculas. Con la especificación

de CASE_ENABLED, asegura que los documentos se indexen de modo que sea posible una búsqueda sensible a las mayúsculas y minúsculas. Para obtener más información, consulte el apartado “Efectuar una búsqueda similar o una búsqueda en documentos DBCS” en la página 43.

INDEXPROPERTY SECTIONS_ENABLED DOCUMENTMODEL(S) modelo

Propiedades de un tipo de índice seleccionado.

SECTIONS_ENABLED especifica que el tipo de índice seleccionado puede contener información acerca de la estructura de documentos.

DOCUMENTMODEL/DOCUMENTMODELS *modelo* especifica el modelo o modelos que se han de asociar como valor por omisión para los documentos que se han de indexar. Debe especificarse un nombre de modelo si se utiliza la propiedad de índice SECTIONS_ENABLED. Si se especifica una lista de modelos, se utiliza el primer modelo como modelo por omisión para el índice. El modelo por omisión se utiliza durante la indexación si el documento no hace ninguna referencia a un modelo o si no se especifica ningún modelo durante la búsqueda.

Los caracteres que pueden utilizarse para el nombre de modelo son a-z, A-Z y 0-9.

El nombre de modelo especificado debe corresponder a una definición de modelo del archivo de definiciones de modelos `desmodel.ini`.

Para cambiar el modelo o modelos asociados a un índice,

1. Utilice DISABLE TEXT TABLE para inhabilitar el índice
2. Utilice ENABLE TEXT TABLE para volver a indexar los documentos, especificando nombres de modelos de documentos diferentes.

UPDATEFREQ frecuencia-actualización

La frecuencia de actualización del índice en función de *cuándo* se ha de realizar la actualización y *cuántos documentos de texto deben ponerse en cola* en la tabla de registro cronológico. Si no hay suficientes documentos de texto en la tabla de registro cronológico en el día y la hora dadas, el índice no se actualiza.

La sintaxis se describe en el apartado “Actualización de un índice” en la página 46.

Si no especifica UPDATEFREQ, se utiliza la frecuencia por omisión especificada en los valores de configuración del texto.

NONE

No se realizan más actualizaciones de índice. Está pensado para una columna de texto en la que no habrán más cambios.

Consejo

Si tiene muchas tablas, considere la posibilidad de evitar la utilización de los valores por omisión. Creando valores de frecuencia de actualización individuales para las tablas puede evitar la indexación simultánea de todas las tablas que provoca una carga prolongada innecesaria en los recursos del sistema.

DIRECTORY directorio

La vía de acceso del directorio en el que se ha de almacenar el índice de texto. La vía de acceso especificada se concatena con “txinsnnn” donde *nnn* es el número de nodo.

Se trata de un directorio del sistema en el que se ejecuta el servidor DB2 Text Extender. Si todavía no existe el directorio, se crea. Si ya existe, asegúrese en las plataformas UNIX que el directorio tenga permisos de lectura/grabación para el propietario de la instancia y compruebe el 'ID de usuario con barreras' si se ha especificado uno diferente del especificado cuando se ha creado la instancia de DB2.

Si no especifica la palabra clave DIRECTORY, se utiliza el valor de DIRECTORY de los valores de configuración del texto.

Si está utilizando bases de datos particionadas, puede especificar un nombre de vía de acceso de índice común para todos los nodos o para un rango de nodos, o diferentes vías de acceso de directorio de índices para cada nodo.

ON NODE nodo [TO nodo]

El número del nodo o del rango de nodos al que se está asignando el nombre de vía de acceso del directorio.

Notas del uso

Se crea un nuevo índice de texto que se asocia a todas las columnas de texto de esta tabla. Esto se hace cuando se desea tener un índice común para todas las columnas de texto de una tabla, en lugar de un índice separado para cada columna de texto.

Cuando ha habilitado una tabla, debe ejecutar ENABLE TEXT COLUMN para cada una de las columnas de texto en las que desea buscar.

Se crea una tabla de registro cronológico en la base de datos. La tabla se utiliza para registrar los cambios, es decir, las inserciones, las actualizaciones y las supresiones de las columnas de texto que se habilitan luego.

Mandato ENABLE TEXT TABLE

Cuando se habilita una columna de texto, se crean desencadenantes que supervisan los cambios en el texto y que mantienen automáticamente un registro en la tabla de registro cronológico de los documentos que es necesario indexar.

DB2 Text Extender indexa los documentos de texto listados en la tabla de registro cronológico periódicamente, tal como se especifica por la palabra clave UPDATEFREQ.

Mandato GET ENVIRONMENT

GET ENVIRONMENT

Este mandato visualiza los valores de las variables de entorno.

Autorización

Ninguna.

Sintaxis del mandato

▶—GET ENVIRONMENT—▶

Parámetros del mandato

Ninguno.

Notas del uso

Estas son las variables de entorno visualizadas:

DB2INSTANCE

Nombre de instancia de DB2 UDB .

DB2DBDFT

Nombre de base de datos por omisión.

DB2TX_INSTOWNER

Nombre de instancia de DB2 Text Extender. Sólo es necesaria si se conecta a un servidor UNIX.

DB2TX_INSTOWNERHOMEDIR

Directorio inicial del propietario de la instancia. Sólo es necesaria si se conecta a un servidor UNIX.

GET INDEX SETTINGS

Este mandato visualiza los valores de un índice, mostrando lo siguiente:

- Tipo de índice
- Opción de índice (opcional)
- Actualizar opción de índice
- Directorio de índices
- Frecuencia de actualización
- Modelo por omisión.

Autorización

Ninguna.

Sintaxis del mandato

```
►►—GET INDEX SETTINGS—nombre-tabla—————►
```



```
└─┬—HANDLE—nombre-columna-descriptores-contexto—┘
```

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada cuyos valores de índice se han de visualizar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuyos valores de índice se han de visualizar.

Notas del uso

Si la tabla se habilita como una tabla de múltiples índices, este mandato visualiza los valores de índice de todas las columnas de texto habilitadas en la tabla. Si se proporciona un *nombre-columna-descriptores-contexto*, este mandato visualiza los valores de índice de la columna especificada.

Si la tabla es una tabla de índice común, se visualizan los valores del índice común. Si se proporciona un *nombre-columna-descriptores-contexto*, se pasa por alto.

Si la tabla o la columna se habilita con la propiedad de índice `SECTIONS_ENABLED`, el mandato `GET INDEX SETTINGS` visualiza el modelo por omisión para el índice. El modelo por omisión es el nombre del

Mandato GET INDEX SETTINGS

modelo que ha especificado durante la habilitación o el primer nombre de modelo de una lista de nombres de modelo. Este es un ejemplo:

Valores de índice actuales:

Tipo de índice	(INDEXTYPE)	=	LINGUISTIC
Modelo por omisión	(DOCUMENTMODEL)	=	mimodelo
Opción actual. índice	(UPDATEINDEX)	=	UPDATE
Frecuencia actualización	(UPDATEFREQ)	=	NONE
Nodo 1			
Directorio de índices	(DIRECTORY)	=	/home/user1/db2tx/indices

GET INDEX STATUS

Este mandato visualiza la siguiente información del estado del índice para una columna de descriptores de contexto o una tabla determinadas:

- Si está disponible la función de búsqueda
- Si está disponible la función de actualización de índice
- Si está disponible la función de reorganización
- El número de documentos planificados
- El número de documentos indexados
- El número de documentos indexados en el índice primario
- El número de documentos indexados en el índice secundario
- Sucesos de error.

Autorización

Ninguna.

Sintaxis del mandato



Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene las columnas de texto cuyo estado se ha de visualizar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuyo estado se ha de visualizar.

Notas del uso

Para una tabla de múltiples índices, debe especificar el nombre de la columna de descriptores de contexto.

Mandato GET STATUS

GET STATUS

Este mandato visualiza la información acerca del estado habilitado de las tablas o columnas de texto de la base de datos .

Autorización

Ninguna.

Sintaxis del mandato

▶—GET STATUS—◀

Parámetros del mandato

Ninguno.

Notas del uso

Este mandato visualiza si la base de datos está habilitada, los nombres de las tablas de texto habilitadas en la base de datos , los nombres de las columnas de texto habilitadas y sus columnas de descriptores de contexto asociadas y los nombres de las columnas de descriptores de contexto del archivo externo.

GET TEXT CONFIGURATION

Este mandato visualiza los valores por omisión para la configuración del texto para la base de datos conectada.

Para cambiar estos valores por omisión, utilice “CHANGE TEXT CONFIGURATION” en la página 131.

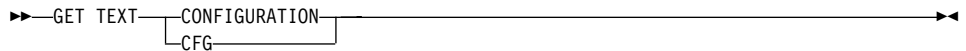
Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio SELECT

Sintaxis del mandato

```
▶▶ GET TEXT CONFIGURATION
```



Parámetros del mandato

Ninguno.

Notas del uso

Para ver un ejemplo de la información de configuración del texto, consulte el apartado “Visualización de los valores de configuración del texto” en la página 105.

Mandato GET TEXT INFO

GET TEXT INFO

Este mandato visualiza los valores de información del texto para las columnas de texto:

- CCSID
- Idioma
- Formato.

Autorización

Ninguno.

Sintaxis del mandato

```
▶▶ GET TEXT INFO nombre-tabla ───────────────────────────────────▶▶  
└─ HANDLE nombre-columna-descriptores-contexto ───────────┘
```

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene las columnas de texto cuyos valores de información de texto se han de visualizar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuyos valores de información de texto se han de visualizar.

Notas del uso

Si se proporciona un nombre de columna de descriptores de contexto, se visualiza la información de texto sólo para esta columna.

Si no se proporciona ningún nombre de columna de descriptores de contexto, se visualiza la información de texto para cada columna habilitada en esta tabla.

QUIT

Este mandato detiene el procesador de línea de mandatos de DB2 Text Extender y devuelve el control al sistema operativo.

Autorización

Ninguna.

Sintaxis del mandato

»—QUIT—«

Parámetros del mandato

Ninguno.

Notas del uso

Se termina la conexión con la base de datos.

REORGANIZE INDEX

Si una columna de texto se actualiza con frecuencia, la búsqueda en el índice puede ser ineficaz. Para que la búsqueda vuelva a ser eficaz, se ha de *reorganizar* el índice. Aunque DB2 Text Extender reconoce cuándo se ha de reorganizar un índice y lo hace en el fondo automáticamente, pueden presentarse situaciones en las que sea necesario reorganizar manualmente un índice utilizando REORGANIZE INDEX. Puede utilizar el mandato GET INDEX STATUS para averiguar si es necesario reorganizar un índice.

Autorización

Ninguna.

Sintaxis del mandato

►—REORGANIZE INDEX—*nombre-tabla*—►

└─┬─HANDLE—*nombre-columna-descriptores-contexto*—┘

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada cuyo índice se ha de reorganizar. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

El nombre de la columna de descriptores de contexto cuyo índice se ha de reorganizar.

Notas del uso

Para una tabla de múltiples índices, debe especificar un nombre de columna de descriptores de contexto.

Aunque se pueden realizar búsquedas en el índice mientras se ejecuta REORGANIZE INDEX, no se pueden efectuar actualizaciones.

UPDATE INDEX

Este mandato inicia la indexación inmediatamente. Actualiza el índice para que refleje el contenido actual de la columna o columnas de texto a las que está asociado este índice.

Para que las actualizaciones de archivos externos se reflejen en el índice, debe forzar a que se coloque una entrada de “cambio” en la tabla de registro cronológico emitiendo una sentencia de actualización en la columna de descriptores de contexto correspondiente. Consulte el apartado “Actualización de un índice para archivos externos” en la página 100 para ver un ejemplo.

Autorización

Debe tener como mínimo uno de los siguientes privilegios para la tabla:

Privilegio ALTER

Privilegio SELECT

Privilegio UPDATE.

Sintaxis del mandato

```
► UPDATE INDEX nombre-tabla ───────────────────────────────────►
    └── HANDLE nombre-columna-descriptores-contexto ───┘
└──────────────────────────────────────────────────────────►
  └── COMMITCOUNT cuenta ───┘
```

Parámetros del mandato

nombre-tabla

El nombre de la tabla de texto de la base de datos conectada que contiene la columna de texto cuyo índice se ha de actualizar. También puede ser el nombre de una tabla de índice común. El nombre debe incluir un nombre de esquema explícito (calificador) a menos que el nombre de esquema sea el mismo que el ID de usuario.

HANDLE nombre-columna-descriptores-contexto

Si se trata de un índice común, el *nombre-columna-descriptores-contexto* no es necesaria y se pasa por alto. El índice que se ha de actualizar se asocia a toda la tabla y no a una columna de texto individual.

Si se trata de una tabla de múltiples índices, el *nombre-columna-descriptores-contexto* es el nombre de la columna de descriptores de contexto cuyo índice se ha de actualizar.

COMMITCOUNT cuenta

Un valor de 500 a 1 000 000 que indica el número de inserciones o actualizaciones después de las cuales DB2 UDB debe emitir una

sentencia de compromiso intermedia. Esto puede evitar una situación en la que no haya suficiente espacio de registro cronológico al habilitar tablas o columnas grandes o un gran número de archivos externos.

Mandato UPDATE INDEX

Capítulo 10. Mandatos de administración para el servidor

Este capítulo describe la sintaxis de los mandatos de administración para el servidor. La administración del servidor consiste en las tareas que el usuario puede realizar para iniciar, detener y comprobar el estado del servidor DB2 Text Extender y para crear una base de datos de ejemplo y una tabla de ejemplo. El apartado “Configuración y mantenimiento de un servidor DB2 Text Extender” en la página 21 describe cómo utilizar estos mandatos.

Mandato	Finalidad	Página
TXICRT	Crea una instancia de DB2 Text Extender	174
TXIDROP	Desecha una instancia de DB2 Text Extender	176
TXILIST (sólo para UNIX)	Lista las instancias de DB2 Text Extender (sólo para UNIX)	177
TXIPCLEAN (sólo para UNIX)	Limpia los recursos del sistema después de una anomalía del sistema (sólo para UNIX)	178
TXNADD	Añade servidores DB2 Text Extender	179
TXNCHECK	Comprueba la integridad de la base de datos de DB2 Text Extender	180
TXNDROP	Desecha los servidores DB2 Text Extender	181
TXSAMPLE	Crea y habilita una tabla de ejemplo	182
TXSTART	Inicia los servicios de DB2 Text Extender	183
TXSTATUS	Visualiza el estado del servicio de búsqueda	184
TXSTOP	Detiene los servicios de DB2 Text Extender	185
TXTHESC	Compila un archivo de definiciones de tesaurus	186
TXTHESN	Compila un archivo de definiciones de tesaurus Ngram	188
TXTRACE	Produce información de rastreo	190
TXVERIFY	Crea y habilita una base de datos de ejemplo	195
TXWIZARD	Un programa de ejemplo para indexar y buscar (sólo para AIX, Windows NT y Windows 2000)	196

Mandato TXICRT

TXICRT

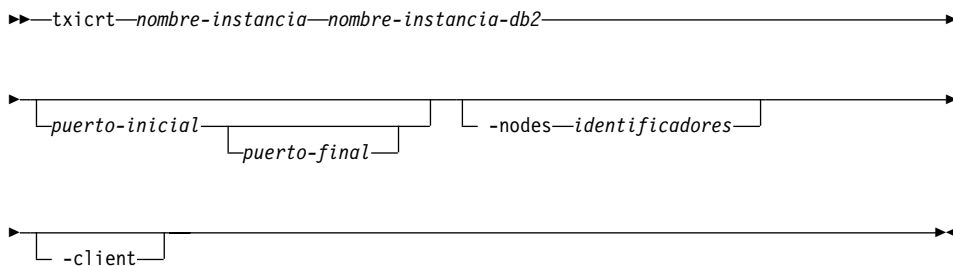
Este mandato crea una instancia de DB2 Text Extender.

Autorización

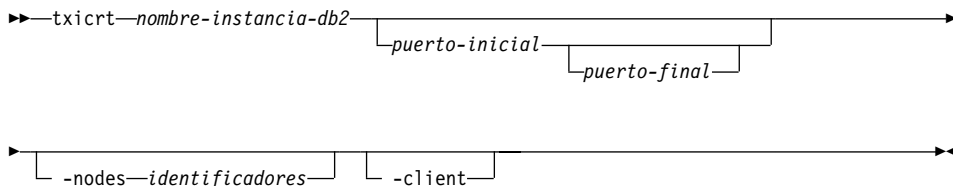
(Sólo para UNIX) Root

Sintaxis del mandato

Para UNIX:



Para Windows NT y Windows 2000:



Parámetros del mandato

nombre-instancia

(Sólo para UNIX) El ID de usuario del propietario de la instancia.

nombre-instancia-db2

El nombre de la instancia de DB2 que se ha de asociar a la instancia que se crea.

puerto-inicial/puerto-final

El rango de puertos TCP/IP que se han de utilizar si está trabajando con más de un nodo físico. No es necesario si sólo utiliza nodos lógicos.

El rango de puertos debe ser igual al número de nodos en los que desea trabajar. Si está trabajando con un nodo, especifique solamente el parámetro puerto-inicial. Para dar soporte a un subconjunto de nodos, utilice el parámetro puerto-final.

No especifique ningún rango de puertos si la instancia se crea en una base de datos de una sola partición.

Utilice únicamente números de puerto que no se listen en el archivo TCP/IP etc/services. No especifique números de puertos para las instalaciones de sólo cliente.

-nodes *identificadores*

Una lista de identificadores de nodos. Ejemplo: -nodes 0 1 2.

Utilice este parámetro si la tabla pertenece a nodos específicos de un grupo de nodos.

Si no se especifica este parámetro, se utilizan todos los nodos, es decir, todos los nodos definidos en db2nodes.cfg.

No especifique identificadores de nodos para instalaciones de sólo cliente.

-client

Especifique este parámetro si tiene una instalación de sólo cliente DB2 Text Extender. Si no especifica este parámetro, TXICRT configura una instancia de cliente/servidor de DB2 Text Extender.

Notas del uso

La habilitación de tablas o columnas de texto sólo es posible en los nodos en los que ha creado una instancia de DB2 Text Extender.

Ya debe haberse creado una instancia de DB2 para el ID de usuario.

En instalaciones UNIX, este mandato específico de instancia está en el directorio /usr/lpp/db2tx_07_01/instance.

En sistemas Windows NT o Windows 2000, se crea una instancia por omisión en el momento de la instalación en un entorno de bases de datos no particionadas.

Mandato TXIDROP

TXIDROP

Este mandato desecha una instancia de DB2 Text Extender junto con todos sus índices.

Autorización

(Sólo para UNIX) Root

Sintaxis del mandato

Para UNIX:

►►txidrop—*nombre-instancia*—*nombre-instancia-db2*—►►

Para Windows NT y Windows 2000:

►►txidrop—*nombre-instancia*—►►

Parámetros del mandato

nombre-instancia

El nombre de la instancia que se ha de desechar.

nombre-instancia-db2

El nombre de la instancia de DB2 asociada a la instancia que se ha de desechar.

Notas del uso

Antes de desechar una instancia, inhabilite las bases de datos que se hayan habilitado para la misma.

En instalaciones UNIX, este mandato específico de instancia está en el directorio /usr/lpp/db2tx_07_01/instance.

TXILIST (sólo para UNIX)

Este mandato lista las instancias de la versión de DB2 Text Extender actual.

Autorización

SYSADM, DBADM

Sintaxis del mandato

▶▶—txilist—▶▶

Parámetros del mandato

Ninguno.

Notas del uso

Si hay más de una versión de DB2 Text Extender instalada, este mandato sólo lista las instancias para la versión de DB2 Text Extender con la que está trabajando actualmente, es decir, la versión determinada por el directorio actual.

Mandato TXIPCLEAN (sólo para UNIX)

TXIPCLEAN (sólo para UNIX)

Este mandato limpia los recursos del sistema para el usuario que ha iniciado la sesión, es decir mata el proceso desssdem, limpia la memoria compartida, segmenta las colas de mensajes y semáforos.

Autorización

Miembro del grupo smadmin

Sintaxis del mandato

►—txipclean—◄

Parámetros del mandato

Ninguno.

Notas del uso

Utilice este mandato con cuidado; no sólo limpia los recursos de DB2 Text Extender, si no todos los recursos compartidos del usuario que ha iniciado la sesión. Inicie siempre la sesión como propietario de la instancia para limitar la limpieza a los recursos de la instancia.

Pueden haber situaciones en las que no pueda iniciar o detener DB2 Text Extender de forma adecuada. En este caso, intente detener DB2 y DB2 Text Extender primero, antes de ejecutar TXIPCLEAN.

TXNADD

Este mandato crea un servidor DB2 Text Extender adicional en el nodo especificado. El mandato TXSTART posteriormente inicia el servidor.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

►►—txnadd—*número-nodo* *puerto* ◀◀

Parámetros del mandato

número-nodo

El número del nodo que se ha de añadir. Especifique el número de nodo exactamente tal como se especifica en el archivo `db2nodes.cfg`.

puerto El número de puerto que se ha de utilizar para el nodo añadido. Debe especificar un número de puerto si está añadiendo un nodo físico o si la configuración de nodos contiene nodos físicos.

Notas del uso

Utilice TXNCHECK para comprobar los índices que debe volver a crear para mantenerlos síncronos con la base de datos.

Consejo

Si añade un nodo físico a una configuración de nodos que tiene nodos lógicos, primero debe desechar la instancia y, después, volver a crearla, especificando los números de puerto para todos los nodos.

Mandato TXNCHECK

TXNCHECK

Este mandato comprueba la integridad de DB2 Text Extender para una base de datos determinada.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

► txncheck nombre-base-datos ◄

Parámetros del mandato

nombre-base-datos

El nombre de la base de datos que se ha de comprobar. Si no se especifica ningún nombre de base de datos, se utiliza el nombre de la base de datos por omisión de DB2.

Notas del uso

La tabla habilitada por DB2 Text Extender debe volverse a indexar si se cambia el grupo de nodos al que pertenece. Utilice TXNCHECK para visualizar las tablas habilitadas por DB2 Text Extender que es necesario volver a indexar debido a que se ha redistribuido un grupo de nodos.

TXNDROP

Este mandato elimina un servidor DB2 Text Extender para un nodo determinado.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

►►—txndrop—*número-nodo*—◄◄

Parámetros del mandato**número-nodo**

El número de nodo que se ha de desechar. Especifique el número de nodo exactamente tal como se especifica en el archivo `db2nodes.cfg`.

Notas del uso

Utilice TXNCHECK para comprobar los índices que debe volver a crear para mantenerlos síncronos con la base de datos.

Mandato TXSAMPLE

TXSAMPLE

Este mandato crea una tabla de ejemplo, carga los documentos de ejemplo en inglés en la columna COMMENT y habilita la columna de texto. Consulte el apartado “Preparación de una base de datos de ejemplo para la verificación de la instalación” en la página 16.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

► `txsample—nombre-base-datos` id-usuario—contraseña ◀

Parámetros del mandato

nombre-base de datos

El nombre de la base de datos , que ya ha habilitado DB2 Text Extender, en la que se ha de crear la tabla de ejemplo.

id-usuario

Sólo es necesario si está trabajando desde una estación de trabajo cliente.

contraseña

Sólo es necesaria si está trabajando desde una estación de trabajo cliente.

Notas del uso

Si el CCSID de base de datos no es 850, consulte el paso de habilitación de ejemplo para el tipo de índice Ngram y, allí, cambie el CCSID para que coincida con el CCSID de la base de datos.

Consejo

Este mandato también puede utilizarse en una estación de trabajo cliente.

TXSTART

Este mandato inicia los servicios de DB2 Text Extender.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

```
► txstart [NODENUM—nódono] ◀
```

Parámetros del mandato

NODENUM nódono

El número de nodo del servidor que se ha de iniciar. Si no se especifica ningún número de nodo, se inician todos los servidores.

Notas del uso

Ejecute este mandato:

- Mientras está conectado con un ID de usuario en el grupo de administración SM
- Siempre que detenga y reinicie el sistema servidor

Si tiene problemas para iniciar los servicios de DB2 Text Extender, detenga DB2 y ejecute el mandato TXIPCLEAN. Consulte el apartado “TXIPCLEAN (sólo para UNIX)” en la página 178.

Mandato TXSTATUS

TXSTATUS

Este mandato visualiza si DB2 Text Extender está activo y en ejecución.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

► txstatus NODENUM—*númodo* ◄

Parámetros del mandato

NODENUM *númodo*

El número de nodo del servidor cuyo estado se ha de comprobar. Si no se especifica un número de nodo, se visualiza el estado de todos los servidores.

TXSTOP

Este mandato detiene los servicios de DB2 Text Extender.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

►► txstop [NODENUM—*númnodo*] ◄◄

Parámetros del mandato**NODENUM *númnodo***

El número de nodo del servidor cuyo estado se ha de comprobar. Si no se especifica ningún número de nodo, se detienen todos los servidores.

Notas del uso

Este mandato no detiene DB2.

TXTHESC

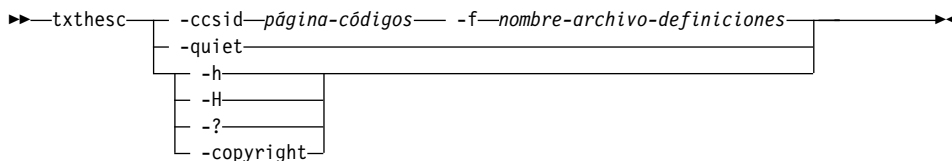
Este mandato compila un archivo de definiciones de tesauo. Este tesauo sólo puede utilizarse para búsquedas en índices exactos o lingüísticos.

Para compilar un archivo de definiciones de tesauo Ngram, consulte "TXTHESN" en la página 188.

Autorización

(Sólo para UNIX) Ninguna.

Sintaxis del mandato



Parámetros del mandato

-f *nombre-archivo-definiciones*

El nombre del archivo SGML que contiene la definición de tesauo. El nombre de archivo debe contener la vía de acceso absoluta de la vía de acceso relativa para el archivo.

El diccionario tesauo se genera en el mismo directorio que el archivo de definición. Tiene el mismo nombre que el archivo de definición pero con las extensiones th1 a th6.

Consejo

Puesto que los archivos de tesauo se sobregaban cuando tienen los mismos nombres, utilice un directorio separado para cada tesauo.

-ccsid *página-códigos*

La página de códigos en la que se graba el archivo de definición de tesauo. Actualmente, sólo se da soporte a la página de códigos 850.

-quiet No se visualiza información de salida.

-copyright

Devuelve el número interno de creación del producto. Utilice este número al informar de problemas.

-h, -H o -?

Visualiza información de ayuda.

Notas del uso

Utilice este mandato para compilar un archivo de definición de tesoro estándar en un formato de definición de tesoro binario. El archivo de definición debe estar en formato SGML.

Para utilizar un archivo de tesoro compilado, muévelo al directorio de diccionarios de la instancia del servidor y, después, especifique la ubicación de los archivos durante la búsqueda.

El directorio de diccionarios de los sistemas OS/2 y Windows es:

unidad:\dmb\db2tx\dict

En sistemas AIX, HP-UX y SUN-Solaris, el directorio de diccionarios es:

DB2TX_INSTOWNER /db2tx/dicts

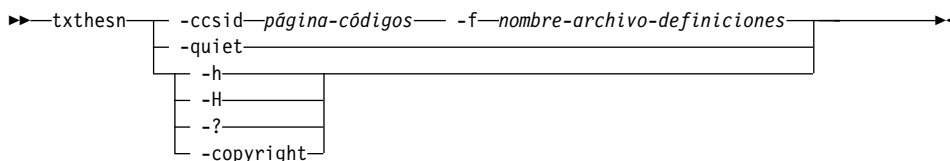
TXTHESN

Este mandato compila un archivo de definición de tesoro Ngram. Este tesoro sólo puede utilizarse para búsquedas en índices Ngram.

Autorización

(Sólo para UNIX) Ninguna.

Sintaxis del mandato



Parámetros del mandato

-f nombre-archivo-definiciones

El nombre del archivo que contiene la definición de tesoro. El nombre de archivo debe contener la vía de acceso absoluta o la vía de acceso relativa para el archivo. El nombre de archivo está restringido a 8+3 caracteres. La extensión es opcional.

El diccionario tesoro se genera en el mismo directorio que el archivo de definición. Tiene el mismo nombre que el archivo de definición pero con las extensiones wdf, wdv, grf, grv, MEY, ROS, NEY, SOS y lkn, donde *n* es un dígito.

Consejo

Puesto que los archivos de tesoro se sobregaban cuando tienen los mismos nombres, debe utilizar un directorio separado para cada tesoro.

-ccsid página-códigos

La página de códigos en la que se graba el archivo de definición de tesoro. Para ver una lista de páginas de códigos soportados, consulte el apartado "CCSID" en la página 35.

-quiet No se visualiza información de salida.

-copyright

Devuelve el número interno de creación del producto. Utilice este número al informar de problemas.

-h, -H o -?

Visualiza información de ayuda.

Notas del uso

Utilice este mandato para compilar un archivo de definición de tesoro en un formato de definición de tesoro binario. El archivo de definición debe estar en el formato descrito en “Creación de un tesoro Ngram” en la página 278.

Para utilizar un archivo de tesoro compilado, muévelo al directorio de diccionarios de la instancia del servidor y, después, especifique la ubicación de los archivos durante la búsqueda.

El directorio de diccionarios de los sistemas OS/2 y Windows es:

unidad:\dmb\db2tx\dict

En sistemas AIX, HP-UX y SUN-Solaris, el directorio de diccionarios es:

DB2TX_INSTOWNER /db2tx/dicts

Mandato TXTRACE

TXTRACE

Este mandato graba la información del proceso en un almacenamiento intermedio de rastreo en la memoria compartida. Esta información se puede utilizar para rastrear anomalías. Puede grabarse en binario desde el almacenamiento intermedio de rastreo en un archivo para formatearlo más tarde cuando se haya desactivado el rastreo o puede formatearse y grabarse en un archivo mientras el rastreo esté activo todavía.

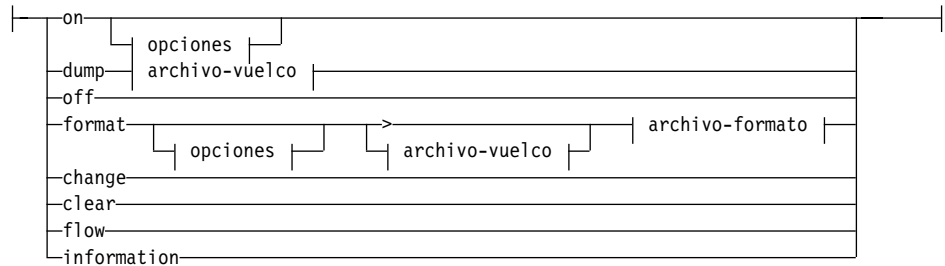
Autorización

(Sólo para UNIX) SYSADM, DBADM

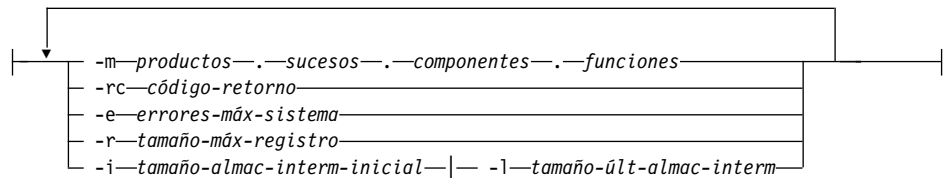
Sintaxis del mandato

► txtrace | parámetros |

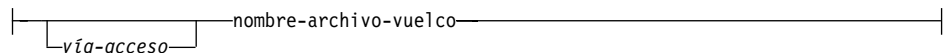
parámetros:



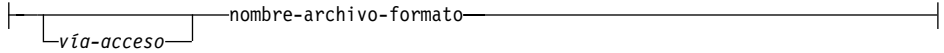
opciones:



archivo-vuelco:



archivo-formato:



Parámetros del mandato

Nota:

También está disponible la opción `-u` con todos los parámetros de TXTRACE para visualizar la información acerca del parámetro.

on Para iniciar el recurso de rastreo.

dump | **dmp**

Para grabar la información de rastreo en binario en un archivo.

off Para detener el recurso de rastreo.

format | **fmt**

Para formatear la información de rastreo en binario. Puede formatear el archivo de vuelco, cuando el rastreo esté desactivado, especificando el nombre del archivo de vuelco y el nombre del archivo que ha de contener la información de rastreo formateada. Para formatear la información de rastreo directamente desde el almacenamiento intermedio de rastreo mientras el rastreo todavía está activado, entre: `destrc format > nombarch.tmp`.

change | **chg**

Para cambiar la máscara de rastreo, `maxSysErrors` o `maxRecordSize`.

clear | **clr**

Para borrar el rastreo.

flow | **flw**

Para mostrar el flujo de control del rastreo.

information | **info** | **inf**

Para obtener información acerca del rastreo.

opciones

Para filtrar la información de rastreo cuando el rastreo está activado (esto reduce la utilización de memoria compartida) o cuando se formatea la información de rastreo. A menos que el rastreo sea muy grande, normalmente resulta más fácil grabar toda la información de rastreo y después filtrarla durante el formateo.

-m Para añadir una "máscara" para especificar los sucesos, los componentes y las funciones que se han de incluir en el rastreo. El valor por omisión es rastrearlo todo. La máscara

Mandato TXTRACE

consta de cuatro partes, separadas por puntos; por ejemplo: 2.2-6.1,3.* Puede especificar un rango utilizando “-” como carácter separador o una lista utilizando “,” como carácter separador. Por ejemplo: 2-6 sólo incluye los sucesos cuyos ID están en el rango de 2 a 6. Para incluir solamente los componentes 2 y 6, especifique 2,6

productos

El ID de producto. El ID de producto para DB2 Text Extender es “2”. El ID de producto para TextMiner es “3”.

sucesos El conjunto de tipos de sucesos que se ha de incluir en el rastreo:

0	error_sistema
1	error_sistema
2	error_sistema
3	error_no-grave
4	error_no-grave
5	códderror_api
7	códderror_fnc
8	error de ruptura
10	entrada_api
11	salida_api
13	códret_api
15	datos_api
30	entrada_fnc
31	salida_fnc
33	códret_fnc
35	datos_fnc

componentes

Los componentes que se han de rastrear.

Los ID de componente para DB2 Text Extender son:

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

Los ID de componente para TextMiner son:

1	automachine
----------	-------------

2	bgproc (proceso de fondo)
3	cluster
4	common
5	commsrvc (servicios comunes)
6	communic (comunicación)
7	daemon
8	dsclient
9	environ (entorno)
10	glue
11	idxcomm (creación índice, parte común)
12	libsrv (servicios de biblioteca)
13	search
14	trace
15	guru
16	indexbld (creación de índice, sólo tm)
17	indexeng (aplicación de índice, sólo tm)
18	smsearch
19	(aplicación de índice, sólo tm)
20	tmsearch
21	gtrcm (gtr, parte común)
22	gtrsrch (búsqueda, sólo gtr)
23	gtridx (creación de índice, sólo gtr)

funciones

Asterisco (*). El conjunto de funciones que se ha de rastrear. Utilice un asterisco (*) para rastrear todas las funciones a menos que el Centro de soporte de IBM le indique lo contrario.

-rc *código-retorno*

Trata el *código-retorno* como un error del sistema.

-e *errores-máx-sistema*

Entero. Para detener el rastreo después de este número de errores. El valor por omisión es 1 que especifica que cuando se produce el primer error del sistema, se suprime el rastreo de todos los sucesos de gravedad inferior. Esto es aceptable si sólo está interesado en el primer error principal, pero debe especificar un número superior (por ejemplo, -e 50) si desea ver todo el rastreo después del error del sistema inicial. El destino del rastreo es la memoria compartida.

-r *tamaño-máx-registro*

Entero. Para detener el rastreo después de que se haya grabado este número de registros en el archivo de rastreo. El valor por omisión es 16 KB.

Mandato TXTRACE

-i *tamaño-alm-interm-inicial*

Entero. Para mantener este número de registros desde el principio del rastreo. Si se especifica -i, el valor por omisión es 16 KB. En un sistema UNIX, el tamaño de almacenamiento intermedio recomendado es de 2 MB.

Si no se especifican -i ni -l, -l es el valor por omisión.

Si especifica -i, no se produce el reinicio; no se graban más entradas de rastreo si el volumen de registros sobrepasa el *tamaño-máx-registro*, incluso si borra todas las entradas de rastreo. Para grabar las nuevas entradas de rastreo, aumente el tamaño de almacenamiento intermedio, desactive el rastreo y, después, vuelva a activarlo.

-l *tamaño-últ-almac-interm*

Entero. Para mantener este número de registros del final del rastreo. El valor por omisión es 16 KB. En un sistema UNIX, el tamaño de almacenamiento intermedio recomendado es de 2 MB.

vía-acceso

El directorio en el que se almacena el archivo correspondiente.

nombre-archivo-vuelco

El nombre del archivo que contiene la información de rastreo en binario.

nombre-archivo-formateado

El nombre del archivo que contiene la información de rastreo formateada.

Ejemplos

Consulte el apartado “Rastreo de anomalías” en la página 23.

TXVERIFY

Este mandato crea una base de datos de ejemplo y la habilita.

Autorización

(Sólo para UNIX) SYSADM, DBADM

Sintaxis del mandato

```
▶—txverify—nombre-base-datos—┬──────────────────────────────────────────┐──▶  
                                └──id-usuario—contraseña──┘
```

Parámetros del mandato**nombre-base-datos**

El nombre de la base de datos en la que se ha de crear la tabla de ejemplo.

id-usuario

Sólo es necesario si está trabajando desde una estación de trabajo cliente.

contraseña

Sólo es necesaria si está trabajando desde una estación de trabajo cliente.

Notas del uso

Este mandato también se puede utilizar en una estación de trabajo cliente.

TXWIZARD

Este mandato sólo está disponible en sistemas AIX, Windows NT y Windows 2000. Se proporciona como un ejemplo de cómo crear índices en las tablas o columnas seleccionadas y cómo buscar en estos índices.

Sintaxis del mandato

►—txwizard—◄

Parámetros del mandato

Ninguno.

Notas del uso

Este asistente de ejemplo le conduce por los pasos necesarios para posibilitar la búsqueda en las columnas seleccionadas de una tabla utilizando las funciones de búsqueda de DB2 Text Extender. Puede utilizar el código fuente Java que se proporciona con el asistente de ejemplo para crear su propia aplicación de preparación y búsqueda de texto o aprender cómo se pueden utilizar y llamar las funciones de DB2 Text Extender en un entorno Java.

En AIX, el código fuente Java del asistente se almacena en el directorio
`/usr/lpp/db2tx_06_01/samples/wizard`

En Windows NT y Windows 2000, se almacena en el directorio
`%DMBMPATH%\samples\wizard`

Capítulo 11. Funciones de búsqueda

DB2 Text Extender proporciona funciones SQL para buscar en documentos de texto almacenados en DB2 UDB y para trabajar con los resultados de una búsqueda. Algunos de los parámetros de las funciones son tipos de datos conocidos como *tipos diferenciados* que se proporcionan con DB2 Text Extender.

Este capítulo describe las funciones SQL de DB2 Text Extender y los tipos diferenciados.

DB2 Text Extender proporciona un archivo de entrada del procesador de línea de mandatos de DB2 UDB denominado `txsample.udf`. Contiene ejemplos de las funciones de DB2 Text Extender que se ejecutan en la tabla de ejemplo descrita en el apartado “Preparación de una base de datos de ejemplo para la verificación de la instalación” en la página 16. Utilice este archivo para ver los ejemplos de las funciones de preparación y búsqueda de texto.

Los tipos diferenciados de DB2 Text Extender

Tipo diferenciado	Tipo de datos fuente	Comentarios
DB2TEXTH	VARCHAR(60) FOR BIT DATA	<p>Descriptor de contexto de texto. Una serie de longitud variable que contiene la información necesaria para la indexación de un documento de texto almacenado en una columna de texto. La información de un descriptor de contexto incluye un ID de documento, el nombre del servidor en el que se ha de indexar el texto, el nombre del índice y la información acerca del documento de texto.</p> <p>Los descriptores de contexto se almacenan en columnas que DB2 Text Extender crea y asocia a cada columna de texto.</p> <p>Las funciones <code>HANDLE</code> e <code>INIT_TEXT_HANDLE</code> devuelven este tipo de datos.</p>

Tipos diferenciados de DB2 Text Extender

Tipo diferenciado	Tipo de datos fuente	Comentarios
DB2TEXTFH	VARCHAR(210) FOR BIT DATA	<p>Descriptor de contexto de archivo. Una serie de longitud variable que contiene la información necesaria para indexar un archivo de texto externo – un archivo almacenado fuera del control de DB2 UDB . La información de un descriptor de contexto de texto incluye un ID de documento, el nombre del servidor en el que se ha de indexar el texto, el nombre del índice, la información acerca del archivo de texto y la información acerca de la ubicación del archivo.</p> <p>Los descriptores de contexto de archivo se almacenan en columnas que DB2 Text Extender crea y asocia a cada grupo de archivos externos.</p> <p>Las funciones FILE e INIT_TEXT_HANDLE devuelven este tipo de datos.</p>
DB2TEXTHLISTP	VARCHAR(16) FOR BIT DATA	<p>Puntero de lista de descriptores de contexto. Un puntero para una lista de descriptores de contexto asociados a los documentos de texto encontrados por una búsqueda.</p> <p>La función HANDLE_LIST devuelve este tipo de datos.</p>
DB2TEXTFHLISTP	VARCHAR(16) FOR BIT DATA	<p>Puntero de lista de descriptores de contexto. Un puntero para una lista de descriptores de contexto asociados a los archivos externos encontrados por una búsqueda.</p> <p>La función HANDLE_LIST devuelve este tipo de datos.</p>

Un resumen de las funciones de DB2 Text Extender

Función de búsqueda	Finalidad	Página
CCSID	Devuelve el CCSID de un descriptor de contexto	200
CONTAINS	Realiza una búsqueda de texto en un documento en particular	201

Resumen de las funciones de DB2 Text Extender

Función de búsqueda	Finalidad	Página
FILE	Devuelve o cambia la vía de acceso y el nombre de un archivo de un descriptor de contexto existente	202
FORMAT	Devuelve o cambia el valor de formato del documento de un descriptor de contexto	203
HANDLE ¹	Devuelve un descriptor de contexto de una lista de descriptores de contexto	204
HANDLE_LIST ¹	Busca y devuelve una lista de descriptores de contexto	205
INIT_TEXT_HANDLE	Devuelve un descriptor de contexto inicializado parcialmente que contiene información como, por ejemplo, los valores de formato y de idioma	206
LANGUAGE	Devuelve o cambia el valor del idioma en un descriptor de contexto	208
NO_OF_DOCUMENTS ¹	Devuelve el número de documentos listados en una lista de descriptores de contexto	209
NO_OF_MATCHES	Busca y devuelve el número de coincidencias encontradas	210
RANK	Busca y devuelve el valor de rango de un documento de texto encontrado	211
REFINE	Toma un argumento de búsqueda y un argumento de redefinición de búsqueda y devuelve un argumento de búsqueda combinado	212
SEARCH_RESULT	Devuelve una tabla intermedia con el resultado de la búsqueda de la serie de búsqueda especificada	213

Se proporcionan ejemplos de la utilización de las funciones de DB2 Text Extender en el “Capítulo 6. Cómo buscar” en la página 77.

1. Estas funciones de búsqueda son características de un release anterior de DB2 Text Extender. Se siguen soportando por razones de compatibilidad. Su funcionalidad se ha sustituido por la función de búsqueda SEARCH_RESULT.

Función CCSID

CCSID

La función CCSID devuelve el CCSID (tipo de datos SMALLINT) de un descriptor de contexto. Este es el parámetro CCSID utilizado para indexar el documento de texto correspondiente. Se describe en el apartado “CCSID” en la página 35. Se establece para cada columna de texto por el mandato ENABLE TEXT COLUMN.

Sintaxis de la función

►—CCSID—(*—descriptor-contexto—*)—►◄

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH o DB2TEXTH. Normalmente, es el nombre de una columna de descriptores de contexto desde la cual se ha de devolver el valor de CCSID.

CONTAINS

La función CONTAINS busca el texto en un documento de texto en particular. Devuelve un valor INTEGER 1 si el documento contiene el texto. De lo contrario, devuelve 0.

Sintaxis de la función

►►—CONTAINS—(—*descriptor-contexto*—,—*argumento-búsqueda*—)—————►◄

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH o DB2TEXTH. Normalmente es el nombre de una columna que contiene los descriptores de contexto de los documentos de texto en los que se ha de buscar.

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Función FILE

FILE

La función FILE realiza una de las acciones siguientes:

- Devuelve la vía de acceso y el nombre de archivo de un descriptor de contexto
- Cambia la vía de acceso y el nombre de archivo de un descriptor de contexto y devuelve la vía de acceso y el nombre de archivo.

El descriptor de contexto devuelto es un valor de tipo DB2TEXTFH.

Sintaxis de la función

►—FILE—(*—descriptor-contexto—*)—►

►—FILE—(*—descriptor-contexto—*, *—nombre-archivo—*)—►

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH. Normalmente, es el nombre de una columna de descriptores de contexto desde la que se ha de devolver el nombre de archivo.

nombre-archivo

Una serie de tipo VARCHAR(150) que especifica la nueva vía de acceso absoluta y el nombre del archivo externo que se ha de asociar al descriptor de contexto. La vía de acceso podría ser, por ejemplo, una unidad de la LAN o una unidad montada en NFS. Los permisos de acceso del archivo deben permitir que el propietario de la instancia de DB2 UDB acceda al archivo.

FORMAT

La función FORMAT realiza una de las acciones siguientes:

- Devuelve el formato del documento especificado en un descriptor de contexto
- Cambia la especificación de formato del descriptor de contexto de un documento y devuelve el descriptor de contexto cambiado.

El formato de documento devuelto es una serie de tipo VARCHAR(30). El descriptor de contexto devuelto es de tipo DB2TEXTFH oDB2TEXTH.

Este es el parámetro de formato utilizado para indexar el documento de texto correspondiente. Los formatos de documento soportados se listan en el apartado “Formatos de documentos que se soportan” en la página 31.

Sintaxis de la función

►►—FORMAT⁽¹⁾—(—*descriptor-contexto*—)—————►►

Notas:

- 1 Devuelve un valor de formato, tipo VARCHAR(30).

►►—FORMAT⁽¹⁾—(—*descriptor-contexto*—,—*formato*—)—————►►

Notas:

- 1 Devuelve un descriptor de contexto, tipo DB2TEXTFH oDB2TEXTH.

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH oDB2TEXTH. Normalmente, es el nombre de una columna de descriptores de contexto desde la cual se ha de devolver o establecer el valor de formato.

formato

El nuevo valor de formato de documento de tipo de datos VARCHAR(30).

Si se especifica *formato*, el formato de este documento se establece en el descriptor de contexto; en este caso, se devuelve el descriptor de contexto en lugar del valor de formato.

Función HANDLE

HANDLE

La función HANDLE devuelve un descriptor de contexto de tipo DB2TEXTFH o DB2TEXTH, seleccionado por el número de secuencia en una lista de descriptores de contexto.

Utilice la función HANDLE únicamente en una sentencia INSERT.

Consejo

Esta función es una característica de un release anterior de DB2 Text Extender. Se sigue soportando por razones de compatibilidad. Su funcionalidad se ha sustituido por la función de búsqueda SEARCH_RESULT.

Sintaxis de la función

►—HANDLE—(—*lista-descriptores-contexto*—,—*entero*—)—————►◀

Parámetros de la función

lista-descriptores-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTHLISTP o DB2TEXTFHLISTP. Se devuelve por la función HANDLE_LIST.

Se trata de un puntero para una lista de descriptores de contexto de los documentos encontrados por una búsqueda.

entero Un valor INTEGER que indica el descriptor de contexto de la lista que se ha de devolver.

HANDLE_LIST

La función HANDLE_LIST busca documentos de texto utilizando un argumento de búsqueda y devuelve un valor de tipo DB2TEXTHLISTP o DB2TEXTFHLISTP que apunta a una lista de descriptores de contexto para los documentos encontrados.

Nota: Sólo se puede acceder al puntero para esta lista desde el ámbito de la sentencia de SQL que utiliza esta función.

Si no se encuentran documentos que contengan el término o los términos de búsqueda, la lista de descriptores de contexto está vacía. Utilice la función NO_OF_DOCUMENTS para averiguar si la lista está vacía.

Si busca texto almacenado en una columna de una tabla base, utilice esta función sólo una vez por sentencia. Si busca texto almacenado en una columna de una vista y la vista se ha creado a partir de más de una tabla base (es decir, es una vista de uniones), utilice un descriptor de contexto para cada columna de descriptores de contexto como argumento de entrada de *descriptor-contexto* para la función HANDLE_LIST.

Consejo

Esta función es una característica de un release anterior de DB2 Text Extender. Se sigue soportando por razones de compatibilidad. Su funcionalidad se ha sustituido por la función de búsqueda SEARCH_RESULT.

Sintaxis de la función

►►—HANDLE_LIST—(—*descriptor-contexto*—,—*argumento-búsqueda*—)—————►►

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTH o DB2TEXTFH. Normalmente, es la columna de descriptores de contexto que contiene los descriptores de contexto de los documentos de texto que se han de buscar. También puede ser un descriptor de contexto prototipo extraído de la vista del catálogo TEXTCOLUMNS.

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Función INIT_TEXT_HANDLE

INIT_TEXT_HANDLE

La función INIT_TEXT_HANDLE devuelve un descriptor de contexto inicializado parcialmente que contiene valores preestablecidos para el formato o el idioma del texto. Puede insertarse en una columna de descriptores de contexto. Es útil cuando se añade una fila que contiene texto cuyo formato e idioma no son iguales a los especificados en los valores de configuración del texto.

El descriptor de contexto devuelto es de tipo DB2TEXTH.

Si piensa buscar en texto almacenado en *archivos externos* en lugar de en una tabla DB2 UDB , puede utilizar la función INIT_TEXT_HANDLE para devolver un descriptor de contexto completamente inicializado que contenga los valores preestablecidos para el CCSID del texto, el formato, el idioma y la ubicación del archivo.

El descriptor de contexto devuelto es un valor de tipo DB2TEXTFH.

Utilice la función INIT_TEXT_HANDLE para insertar o actualizar los valores de descriptor de contexto.

Sintaxis de la función

►►INIT_TEXT_HANDLE(—*formato*—,—*idioma*—)—————►►

►►INIT_TEXT_HANDLE(—*CCSID*—,—*formato*—,—*idioma*—,—*nombre-archivo*—)—————►►

Parámetros de la función

formato

Una serie de tipo VARCHAR(30) que especifica el nuevo valor de formato del documento. Los formatos soportados se listan en el apartado “Formatos de documentos que se soportan” en la página 31.

idioma

Una serie de tipo VARCHAR(30) que especifica el nuevo valor de idioma del documento. Los idiomas soportados se listan en la Tabla 5 en la página 54.

nombre-archivo

Una serie de tipo VARCHAR(150) que especifica la vía de acceso absoluta y el nombre del archivo externo que se ha de asociar al descriptor de contexto. Para tener acceso al archivo en UNIX, el propietario de la instancia de DB2 UDB debe incluirse en los permisos

Función INIT_TEXT_HANDLE

de acceso del archivo. Para usuarios de OS/2 y Windows, los permisos de acceso del archivo deben incluir los ID de usuario de inicio de sesión.

LANGUAGE

La función LANGUAGE efectúa una de las acciones siguientes:

- Devuelve el idioma del documento especificado en un descriptor de contexto
- Cambia la especificación del idioma del descriptor de contexto de un documento y devuelve el descriptor de contexto cambiado.

El idioma del documento devuelto es una serie de tipo VARCHAR(30). El descriptor de contexto devuelto es de tipo DB2TEXTFH oDB2TEXTH.

Este es el parámetro de idioma utilizado para indexar el documento de texto correspondiente. Los idiomas soportados se listan en la Tabla 5 en la página 54.

Sintaxis de la función

►►LANGUAGE⁽¹⁾(—*descriptor-contexto*—)►►

Notas:

- 1 Devuelve un valor de idioma, tipo VARCHAR(30).

►►LANGUAGE⁽¹⁾(—*descriptor-contexto*—,—*idioma*—)►►

Notas:

- 1 Devuelve un descriptor de contexto, tipo DB2TEXTFH oDB2TEXTH.

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH oDB2TEXTH. Normalmente, es el nombre de una columna de descriptores de contexto desde la que se ha de devolver o establecer el valor del idioma.

idioma

El nuevo valor de idioma del documento de tipo de datos VARCHAR(30).

Si se especifica *idioma*, el idioma de este documento se establece en el descriptor de contexto; se devuelve el descriptor de contexto en lugar del valor de idioma.

NO_OF_DOCUMENTS

La función NO_OF_DOCUMENTS devuelve un valor INTEGER que indica el número de elementos de una lista de documentos de texto que una búsqueda ha encontrado. El valor devuelto es el número de entradas encontradas en una lista de descriptores de contexto.

Consejo

Esta función es una característica de un release anterior de DB2 Text Extender. Se sigue soportando por razones de compatibilidad. Su funcionalidad se ha sustituido por la función de búsqueda SEARCH_RESULT.

Sintaxis de la función

►►—NO_OF_DOCUMENTS—(*—lista-descriptores-contexto—*)—————►

Parámetros de la función**lista-descriptores-contexto**

Una expresión cuyo resultado es un valor de tipo DB2TEXTHLISTP o DB2TEXTFHLISTP. Se devuelve por la función HANDLE_LIST.

Se trata de un puntero para una lista de descriptores de contexto de los documentos encontrados por una búsqueda.

Las funciones HANDLE_LIST y NO_OF_DOCUMENTS deben estar en la misma sentencia de SQL porque la lista sólo existe en el ámbito de la sentencia.

Función NO_OF_MATCHES

NO_OF_MATCHES

NO_OF_MATCHES puede buscar en documentos de texto y devolver un valor INTEGER que indique cuántas coincidencias se han producido por documento.

Sintaxis de la función

►►NO_OF_MATCHES(—*descriptor-contexto*—,—*argumento-búsqueda*—)◀◀

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH oDB2TEXTH. Normalmente es el nombre de una columna que contiene los descriptores de contexto de los documentos de texto en los que se ha de buscar.

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

RANK

RANK puede buscar en documentos de texto y devolver un valor de rango para cada documento encontrado, indicando el grado de exactitud con que el argumento de búsqueda describe el documento encontrado.

RANK devuelve un valor DOUBLE entre 0 y 1. El valor del rango es absoluto e indica el grado con que el documento encontrado satisface los criterios de búsqueda en relación a otros documentos encontrados. El valor indica el número de coincidencias encontradas en el documento en relación al tamaño del documento.

Sintaxis de la función

►►—RANK—(—*descriptor-contexto*—,—*argumento-búsqueda*—)—►►

Parámetros de la función

descriptor-contexto

Una expresión cuyo resultado es un valor de tipo DB2TEXTFH o DB2TEXTH. Normalmente es el nombre de una columna que contiene los descriptores de contexto de los documentos de texto en los que se ha de buscar.

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Función REFINE

REFINE

La función REFINE toma dos argumentos de búsqueda y devuelve un argumento de búsqueda combinado de tipo LONG VARCHAR, que consta de los dos argumentos de búsqueda originales conectados por el operador booleano AND.

Sintaxis de la función

►►REFINE(—argumento-búsqueda—,—argumento-búsqueda—)◄◄

Parámetros de la función

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

El argumento de búsqueda no debe contener los parámetros de búsqueda IS ABOUT, THESAURUS ni EXPAND.

SEARCH_RESULT

La función SEARCH_RESULT devuelve el resultado de una búsqueda en una tabla intermedia. Esta función puede utilizarse en la cláusula FROM de una sentencia de SQL.

La tabla devuelta tiene la estructura siguiente:

Nombre de columna	Tipo de datos
HANDLE	DB2TX.DB2TEXTH, DB2TX.DB2TEXTFH
NUMBER_OF_MATCHES	INTEGER
RANK	DOUBLE

Los valores sólo se generan para las columnas seleccionadas de la tabla intermedia. La selección de count(*) sólo genera la columna HANDLE. Puesto que el cálculo de los valores de rango consume muchos recursos del sistema, no debe seleccionar el valor de rango de la tabla intermedia si no es necesario.

Esta función es más rápida que CONTAINS o RANK cuando se procesan tablas grandes.

Sintaxis de la función

►SEARCH_RESULT—(*—esquema—*, *—tabla—*, *—descriptor-contexto—*, *—argumento-búsqueda—*)►

Parámetros de la función

esquema

El nombre de esquema de la tabla en la que está buscando.

tabla El nombre de la tabla en la que está buscando.

descriptor-contexto

El nombre de una columna de descriptores de contexto que corresponde a la columna que contiene los documentos que se han de buscar.

argumento-búsqueda

Una serie de tipo LONG VARCHAR que contiene los términos que se han de buscar. Consulte el apartado “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Ejemplos

Para ver un ejemplo, consulte el apartado “Mejora del rendimiento de la búsqueda” en la página 97 o vea las funciones de DB2 Text Extender de ejemplo proporcionadas en el archivo descrito en el apartado “Dónde encontrar ejemplos de sintaxis de las funciones de búsqueda” en la página 77.

Función SEARCH_RESULT

Capítulo 12. Sintaxis de los argumentos de búsqueda

Un argumento de búsqueda es la condición que el usuario especifica al buscar términos en documentos de texto. Consta de uno o varios términos de búsqueda y parámetros de búsqueda.

En el apartado “Especificación de los argumentos de búsqueda” en la página 85 se ofrecen ejemplos de argumentos de búsqueda y un archivo denominado `txsample.udf`. Contiene ejemplos de las funciones de DB2 Text Extender que se ejecutan en la tabla de ejemplo descrita en el apartado “Preparación de una base de datos de ejemplo para la verificación de la instalación” en la página 16.

Las funciones de DB2 Text Extender que utilizan argumentos de búsqueda son:

- **CONTAINS**. Esta función utiliza un argumento de búsqueda para buscar texto en un documento de texto en particular. Devuelve un valor **INTEGER** 1 si el documento contiene el texto. De lo contrario, devuelve 0.
- **NO_OF_MATCHES**. Esta función utiliza un argumento de búsqueda para buscar en documentos de texto. Devuelve un valor **INTEGER** que indica cuántas coincidencias se han producido por documento.
- **RANK**. Esta función utiliza un argumento de búsqueda para buscar en documentos de texto. Devuelve un valor para cada documento encontrado, que indica el grado de exactitud con que el argumento de búsqueda describe el documento encontrado.
- **REFINE**. Esta función toma dos argumentos de búsqueda y devuelve un argumento de búsqueda combinado de tipo **LONG VARCHAR**, que consta de los dos argumentos de búsqueda originales conectados por el operador booleano **AND**.
- **HANDLE_LIST**. Esta función utiliza un argumento de búsqueda para buscar en documentos de texto. Devuelve un valor de tipo **DB2TEXTHLISTP** o **DB2TEXTFHLISTP** que apunta a una lista de descriptores de contexto para los documentos encontrados.
- **SEARCH_RESULT**. Esta función devuelve una tabla que contiene la información solicitada, es decir, el rango, el número de coincidencias y el descriptor de contexto.

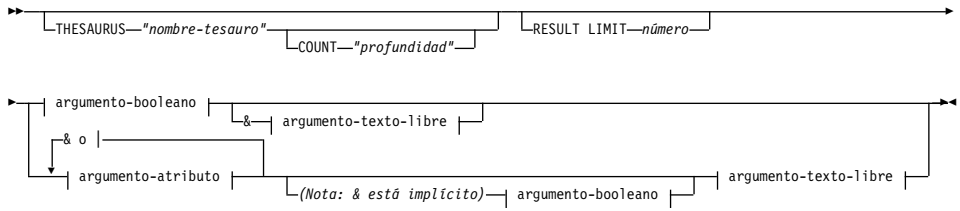
Sintaxis de los argumentos de búsqueda

Las funciones de API que utilizan argumentos de búsqueda son:

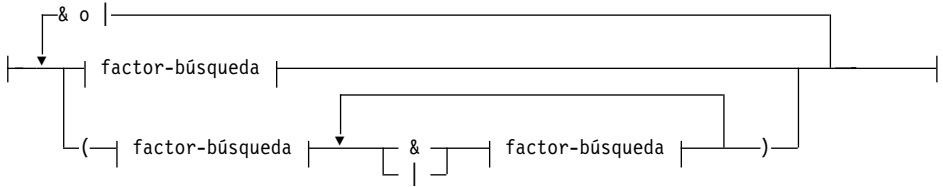
- `DesGetBrowseInfo`. Esta función utiliza un argumento de búsqueda para buscar en el texto identificado por un descriptor de contexto. Devuelve un puntero para la información de examen que `DesStartBrowseSession` necesita para el resaltado de términos.
- `DesGetSearchResultTable`. Esta función utiliza un argumento de búsqueda para buscar en documentos de texto identificados por una columna de texto. Los datos del descriptor de contexto de los elementos de texto encontrados se graban en una tabla de resultados. La información de examen acerca del rango y del número de coincidencias también puede grabarse en la tabla de resultados.

Argumento de búsqueda

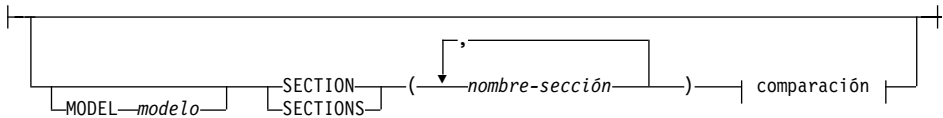
Sintaxis del argumento de búsqueda



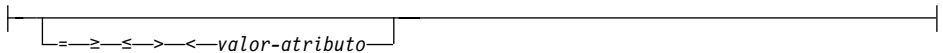
argumento-booleano:



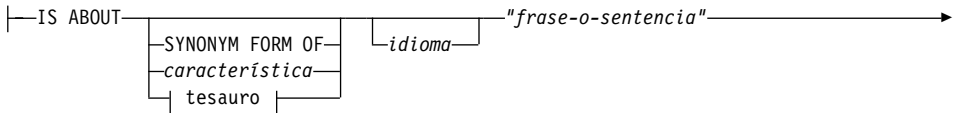
argumento-atributo:



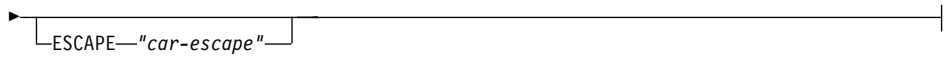
Comparación:



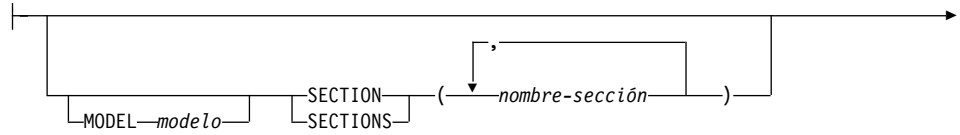
argumento-texto-libre:



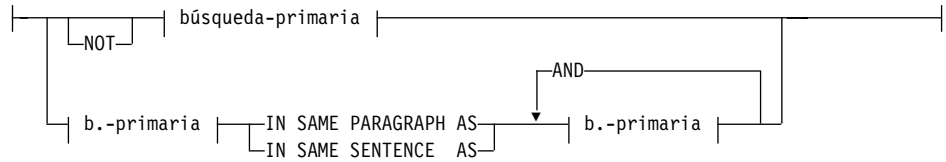
Sintaxis de los argumentos de búsqueda



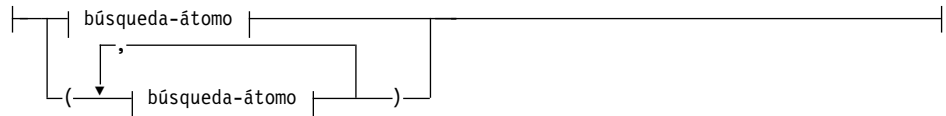
factor-búsqueda:



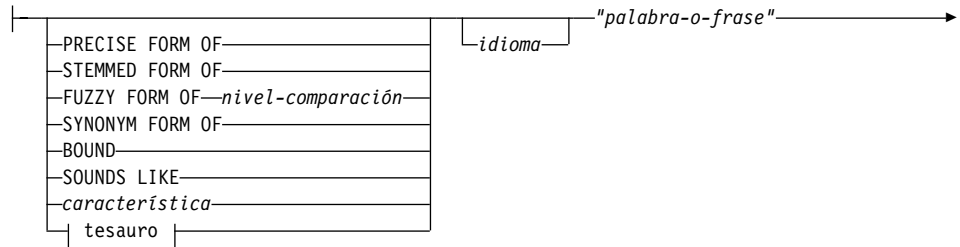
elemento-búsqueda:



búsqueda-primaria:



búsqueda-átomo:




```
└─ESCAPE─"carácter-escape"─┘
```

tesauro (si se especifica THESAURUS):

```
└─EXPAND─"relación"─┘ TERM OF
```

Ejemplos

Se proporcionan ejemplos en el apartado “Especificación de los argumentos de búsqueda” en la página 85.

Parámetros de búsqueda

IS ABOUT

Una opción que le permite especificar un argumento de búsqueda de texto libre, es decir, una frase o sentencia en un idioma natural que describe el concepto que se ha de buscar. Consulte el apartado “Búsqueda de texto libre e híbrida” en la página 93.

MODEL *modelo*

Una palabra clave utilizada para especificar el nombre del modelo de documento que se ha de utilizar en el término de búsqueda. El modelo de documento describe la estructura de los documentos que contienen secciones identificables para que pueda buscarse individualmente en el contenido de estas secciones.

El nombre de modelo debe especificarse en un archivo de modelos de documentos descrito en el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48. El nombre de modelo puede enmascarse con caracteres comodín.

Si no especifica ningún modelo, se utiliza el modelo por omisión especificado al crear el índice.

SECTION(S) *nombre-sección*

Una palabra clave utilizada para especificar una o varias secciones a las que se restringe la búsqueda. Debe especificarse el nombre de sección en un modelo de un archivo de modelos de documentos, descrito en el apartado “Trabajar con documentos estructurados (soporte de secciones)” en la página 48. Un nombre de sección puede enmascarse utilizando los caracteres comodín % y _.

Las secciones pueden anidarse dentro de otras secciones, por ejemplo:

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

Sintaxis de los argumentos de búsqueda

Restricciones: La búsqueda en secciones anidadas sólo es posible para los documentos almacenados en columnas habilitadas con el formato XML. Para los índices Ngram, sólo puede buscarse en un nombre de sección y el formato XML no está soportado.

<valor-atributo

Un valor utilizado junto con un operador de comparación precedente para los atributos listados en la lista de secciones precedente. Una consulta que pide que un valor de atributo esté dentro de un rango determinado puede utilizar dos operadores de comparación en una condición de atributo.

Una combinación de operadores que utiliza la misma clase de comparación (como \geq en el primero y $>$ en el segundo) de la misma condición no es válida. La especificación de dos comparaciones con el operador $=$ tampoco es válida.

- $=$ Pide una comparación de igualdad de un atributo en el documento indexado con el valor de atributo siguiente.
- \geq Pide una comparación "mayor que o igual a" del atributo del documento indexado con el valor del atributo siguiente.
- $>$ Pide una comparación "mayor que" del atributo del documento indexado con el valor del atributo siguiente.
- \leq Pide una comparación "menor que o igual a" del atributo del documento indexado con el valor del atributo siguiente.
- $<$ Pide una comparación "menor que" del atributo del documento indexado con el valor del atributo siguiente.

Las secciones pueden anidarse dentro de otras secciones, por ejemplo:
`play/Act/Title=play/act/title`

Restricciones: La búsqueda en secciones anidadas sólo es posible para los documentos almacenados en columnas habilitadas con el formato XML. Para los índices Ngram, sólo puede buscarse en un nombre de sección y el formato XML no está soportado.

THESAURUS *nombre-tesauro*

Una palabra clave utilizada para especificar el nombre del tesauro utilizado para ampliar el término de búsqueda. El nombre de tesauro es el nombre de archivo (sin su extensión) de un tesauro que se ha compilado utilizando el compilador de tesauro TXTHESC o TXTHESN. Existen los tesauros por omisión `desthes` y `desnthes`, almacenados en el directorio de ejemplo, donde `desnthes` es un

tesauro Ngram. También puede especificar el nombre de vía de acceso del archivo. El nombre de vía de acceso por omisión es la vía de acceso del diccionario.

COUNT *profundidad*

Una palabra clave utilizada para especificar el número de niveles (la profundidad) de los términos del tesauro que se han de utilizar para ampliar el término de búsqueda para una relación determinada. Si no especifica esta palabra clave, se supone una cuenta de 1.

RESULT LIMIT *número*

Una palabra clave utilizada para especificar el número máximo de entradas que se han de devolver en la lista de resultados. *número* es un valor de 1 a 32767. Si se utiliza una búsqueda de texto libre, sólo se establece el rango de la lista de resultados de la búsqueda con respecto a la lista de resultados de búsqueda completa. De lo contrario, sólo se establece el rango del resultado limitado de la búsqueda a partir de las entradas de la lista.

EXPAND *relación*

Se utiliza una palabra clave para especificar la relación como, por ejemplo, INSTANCE, entre el término de búsqueda especificado en TERM OF y los términos del tesauro que se han de utilizar para ampliar el término de búsqueda. El nombre de la relación debe corresponder a la relación utilizada en el tesauro. Consulte el apartado "Conceptos del tesauro" en la página 268.

Para un tesauro Ngram, utilice el nombre de relación de miembros descrito en el apartado "Creación de un tesauro Ngram" en la página 278. Para las relaciones de miembros definidas por el usuario, utilice :RELATION *n* donde *n* es el número de relación de miembros especificado en :RELATED (*número*).

TERM OF *"palabra-o-frase"*

El término de búsqueda, o término de búsqueda de múltiples palabras, al que los demás términos de búsqueda se han de añadir desde el tesauro.

factor-búsqueda

Un operando que puede combinarse con otros operandos para formar un argumento de búsqueda. El orden de la evaluación es de izquierda a derecha. El operador lógico AND (&) tiene una vinculación superior al operador lógico OR (|). Ejemplo:

```
"pasajero" & "vehículo" | "transporte" & "público"
```

se evalúa como:

```
("pasajero" & "vehículo") | ("transporte" & "público")
```

Sintaxis de los argumentos de búsqueda

Para buscar:

```
"pasajero" & ("vehículo" | "transporte") & "público"
```

debe incluir los paréntesis que aparecen aquí.

NOT búsqueda-primaria

Un operador que le permite excluir los documentos de texto de la búsqueda que contienen un término en particular.

Cuando se utiliza NOT en un factor de búsqueda, no puede utilizar la palabra clave SYNONYM FORM OF.

búsqueda-primaria IN SAME PARAGRAPH AS búsqueda-primaria

Una palabra clave que le permite buscar una combinación de términos que aparecen en el mismo párrafo.

El siguiente argumento de búsqueda busca documentos de texto que contienen el término “tráfico” sólo si el término “aéreo” está en el mismo párrafo.

```
"tráfico" IN SAME PARAGRAPH AS "aéreo"
```

No puede utilizar la palabra clave IN SAME PARAGRAPH AS cuando se utiliza NOT como factor de búsqueda.

búsqueda-primaria IN SAME SENTENCE AS búsqueda-primaria

Una palabra clave que le permite buscar una combinación de términos que aparecen en la misma sentencia. Similar a IN SAME PARAGRAPH AS.

AND búsqueda-primaria

Una palabra clave que le permite combinar varias búsquedas primarias para buscar en la misma sentencia o el mismo párrafo.

El siguiente argumento de búsqueda busca “bosque”, “lluvia”, “erosión” y “tierra” en la misma sentencia.

```
"bosque" IN SAME SENTENCE AS "lluvia" AND "erosión" AND "tierra"
```

búsqueda-átomo

Si conecta una serie de átomos de búsqueda mediante comas, una búsqueda será satisfactoria si se encuentra un término de cualquiera de los átomos de búsqueda. Cada átomo de búsqueda debe contener como mínimo una palabra o una frase.

La sentencia siguiente es verdadera si se encuentran uno o varios argumentos de búsqueda.

```
CONTAINS (mytexthandle, '( "texto",  
"gráfico",  
"audio",  
"vídeo" )') = 1
```

PRECISE FORM OF, STEMMED FORM OF, FUZZY FORM OF, SYNONYM FORM OF, BOUND

La Tabla 8 muestra las opciones que corresponden a los diferentes tipos de índice. Por ejemplo, para un índice lingüístico, cualquiera de las opciones son correctas excepto PRECISE FORM OF. Si especifica PRECISE FORM OF, se pasa por alto y se toma el valor por omisión.

El proceso del término de búsqueda se describe con más detalle en la Tabla 9.

Tabla 8. Opciones lingüísticas

Palabra clave de átomo de búsqueda	Tipo de índice				
	Lingüístico	Exacto	Exacto normalizado	Ngram	Ngram sensible a las mayúsculas y minúsculas
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=valor por omisión O=función disponible

Tabla 9. Opciones de término de búsqueda para índices Ngram

Palabra clave de átomo de búsqueda	Proceso de término de búsqueda				
	Mayúsculas/minúsculas		Reducción a la raíz	Coincidencia	
	Sensible	No sensible		Exacta	Similar
PRECISE FORM OF	cuando se habilitan las mayúsculas y minúsculas	X		X	
STEMMED FORM OF		X	X		

Sintaxis de los argumentos de búsqueda

Tabla 9. Opciones de término de búsqueda para índices Ngram (continuación)

Palabra clave de átomo de búsqueda	Proceso de término de búsqueda				
	Mayúsculas/minúsculas		Reducción a la raíz	Coincidencia	
	Sensible	No sensible		Exacta	Similar
FUZZY FORM OF		X			X
X=valor por omisión					

Si utiliza una palabra clave que no está disponible para este tipo de índice, se pasa por alto y se utiliza la palabra clave por omisión en su lugar o se devuelve un mensaje.

PRECISE FORM OF

Una palabra clave que hace que se busque la palabra (o cada palabra de la frase) que sigue a PRECISE FORM OF exactamente tal como se ha escrito, en lugar de reducirse primero a su formato raíz. Para los índices exactos, este formato de búsqueda es sensible a las mayúsculas y minúsculas; es decir, el uso de letras en mayúsculas o en minúsculas es significativo. Por ejemplo, si busca ratón no encontrará "Ratón".

Esta es la opción por omisión para los índices exactos. Para un índice exacto normalizado, el formato por omisión de la búsqueda no es sensible a las mayúsculas y minúsculas. Si especifica esta palabra clave para un índice lingüístico, se pasa por alto y se asume STEMMED FORM OF.

STEMMED FORM OF

Una palabra clave que hace que la palabra (o cada palabra de la frase) que sigue a STEMMED FORM OF se reduzca a su raíz antes de que se lleve a cabo la búsqueda. Este formato de búsqueda no es sensible a las mayúsculas y minúsculas. Por ejemplo si busca ratón encontrará "Ratón".

La manera en que estas palabras se reducen a su raíz depende del idioma.

Ejemplo: programming computer systems se sustituye por program compute system cuando se utiliza un diccionario inglés de EE.UU y por programme compute system cuando se utiliza un diccionario inglés del Reino Unido.

Esta frase de búsqueda puede encontrar "programmer computes system", "program computing systems", "programming computer system", etcétera.

Esta es la opción por omisión para los índices lingüísticos. Si especifica esta palabra clave para un índice exacto, se pasa por alto y en su lugar se asume PRECISE FORM OF.

FUZZY FORM OF

Una palabra clave para realizar una búsqueda “similar” de términos que se escriben de forma similar al término de búsqueda. Esto es especialmente útil al buscar en documentos que se han creado por un programa de reconocimiento óptico de caracteres (OCR). Estos documentos incluyen con frecuencia palabras escritas incorrectamente. Por ejemplo, un programa OCR podría reconocer la palabra *economía* como *economía*. *nivel-comparación*: Un entero de 1 a 5 que especifica el grado de similitud, donde 5 es más similar que 1.

SYNONYM FORM OF

Una palabra clave que hace que se busque la palabra o frase que sigue a SYNONYM FORM OF junto con sus sinónimos. Los sinónimos los proporciona el diccionario especificado por *idioma* o sino por el diccionario por omisión.

Los sinónimos para una frase son frases alternativas que contienen todas las combinaciones posibles de sinónimos que pueden obtenerse sustituyendo cada palabra de la frase original por uno de sus sinónimos. El orden de las palabras permanece igual que en la frase original.

Si especifica esta palabra clave para un índice exacto, se pasa por alto y en su lugar se asume PRECISE FORM OF.

No puede especificar esta palabra clave cuando se utiliza NOT en el factor de búsqueda o cuando la palabra o frase que se ha de buscar contiene caracteres de máscara.

BOUND

Una palabra clave para buscar en documentos que utilizan el CCSID coreano. Hace que la búsqueda respete los límites de la frase o palabra. Si se especifica *idioma*, se pasa por alto; se asume el coreano.

idioma Una variable que determina el diccionario que se utiliza en el proceso lingüístico de los documentos de texto durante la indexación y recuperación. Esto no sólo se aplica a los índices lingüísticos, sino también a los índices exactos porque utilizan un diccionario para procesar las palabras de parada.

El proceso lingüístico incluye el proceso de sinónimos y el proceso de palabras raíz.

Los idiomas soportados se listan en la Tabla 5 en la página 54.

Sintaxis de los argumentos de búsqueda

Nota: Cuando se busca en documentos que no están en inglés de EE.UU, debe especificar el idioma del argumento de búsqueda *sin tener en cuenta el idioma por omisión*.

"palabra-o-frase"

Una palabra o frase que se ha de buscar. Los caracteres que se pueden utilizar en una palabra dependen del idioma. También depende del idioma si se deben separar las palabras por caracteres separadores. Para el inglés y la mayoría de los demás idiomas, cada palabra de una frase debe ir separada por un carácter en blanco.

Búsqueda exacta o lingüística. DB2 Text Extender puede buscar utilizando el formato exacto de la palabra o frase, o una variación de la misma. Si no especifica ninguna de las opciones de la Tabla 8 en la página 223, se utilizan las opciones de lingüísticas por omisión de acuerdo al tipo de índice que se utiliza.

Para buscar una serie de caracteres que contiene comillas dobles, escriba las comillas dobles dos veces. Por ejemplo, para buscar el texto carácter "comodín", utilice:

```
"carácter ""comodín"""
```

Enmascaramiento de caracteres. Una palabra puede contener los siguientes caracteres de máscara:

_ (subrayado)

Representa un carácter cualquiera.

% (porcentaje)

Representa cualquier número arbitrario de caracteres. Si una palabra consiste en un solo %, representa una palabra opcional de cualquier longitud.

Una palabra no puede estar compuesta exclusivamente de caracteres de máscara, excepto cuando se utiliza un solo % para representar una palabra opcional.

Si utiliza un carácter de máscara, no puede utilizar SYNONYM OF, *característica* ni THESAURUS.

ESCAPE *carácter-escape*

Un carácter que identifica el siguiente carácter como uno que se ha de buscar y no uno utilizado como carácter de máscara.

Ejemplo: Si *carácter-escape* es \$, entonces \$%, \$_ y \$\$ representan %, _ y \$ respectivamente. Cualquier carácter % y _ no precedido por \$ representan caracteres de máscara.

Resumen de las normas y restricciones:

Operaciones booleanas

NOT no está permitido después de OR.

FUZZY FORM OF

Los 3 primeros caracteres deben coincidir. No se puede utilizar si una palabra del átomo de búsqueda contiene un carácter de máscara. No se puede utilizar en combinación con NOT. Sólo puede utilizarse con un índice Ngram.

IN SAME PARAGRAPH AS

No puede utilizarse si se utiliza NOT en un factor de búsqueda.

IN SAME SENTENCE AS

No puede utilizarse si se utiliza NOT en un factor de búsqueda.

Índice lingüístico

Impide la utilización de PRECISE FORM OF. Toma por omisión STEMMED FORM OF. Pueden utilizarse caracteres de máscara. Las búsquedas no son sensibles a las mayúsculas y minúsculas.

Carácter de máscara

Impide la utilización de SYNONYM FORM OF y THESAURUS.

Índice Ngram

Pueden utilizarse caracteres de máscara, aunque no después de un carácter no alfanumérico. Las búsquedas no son sensibles a las mayúsculas y minúsculas a menos que se habiliten las mayúsculas/minúsculas del índice y se utilice PRECISE FORM OF.

NOT Impide la utilización de SYNONYM FORM OF, IN SAME PARAGRAPH AS e IN SAME SENTENCE AS.

PRECISE FORM OF

Se pasa por alto para un índice lingüístico.

Índice exacto

Impide la utilización de STEMMED FORM OF y SYNONYM FORM OF. Toma por omisión PRECISE FORM OF. Pueden utilizarse caracteres de máscara. Las búsquedas son sensibles a las mayúsculas y minúsculas.

STEMMED FORM OF

Se pasa por alto para un índice exacto, pero está disponible para un índice exacto normalizado que contiene documentos en inglés.

SYNONYM FORM OF

No se puede utilizar si una palabra del átomo de búsqueda contiene un carácter de máscara. No se puede utilizar en combinación con NOT. No se puede utilizar con un índice exacto.

Sintaxis de los argumentos de búsqueda

Capítulo 13. Funciones API para la búsqueda y examen

DB2 Text Extender proporciona funciones C para buscar documentos de texto y para examinar (visualizar) los documentos encontrados. Estas funciones constituyen la interfaz de programas de aplicación (API) de DB2 Text Extender. Este capítulo describe las funciones API por orden alfabético.

El “Capítulo 8. Utilización de las funciones API para buscar y examinar” en la página 117 proporciona una introducción a las funciones y describe cómo pueden utilizarse a la vez.

Función	Finalidad	Página
DesCloseDocument	Libera el almacenamiento asignado por DesOpenDocument.	231
DesEndBrowseSession	Cierra una sesión de examen y libera el almacenamiento asignado por DesStartBrowseSession.	232
DesFreeBrowseInfo	Libera el almacenamiento asignado por DesGetBrowseInfo.	233
DesGetBrowseInfo	Busca el texto en el documento utilizando un argumento de búsqueda y crea información de examen.	234
DesGetMatches	Devuelve un puntero de la información de resaltado para el documento de texto descrito por un descriptor de contexto de documento. La información de resaltado es una corriente de datos. Comprende el contexto del texto (un párrafo como mínimo) e información para el resaltar el texto en ese contexto.	237
DesGetSearchResultTable	Toma un argumento de búsqueda para buscar documentos de texto de una columna de texto determinada y almacena el resultado en una tabla proporcionada por el usuario. También puede devolver información de examen.	243
DesOpenDocument	Recibe un puntero para la sesión de examen, un descriptor de contexto y una opción DES_FAST o DES_EXTENDED que indica el tipo de proceso lingüístico que se ha de utilizar para resaltar los términos encontrados. Prepara el documento de texto que corresponde al descriptor de contexto para obtener el texto del documento y la información de resaltado y devuelve un descriptor de contexto de documento que se utiliza para llamar iterativamente a DesGetMatches.	248
DesStartBrowseSession	Abre una sesión de examen utilizando la información de examen de DesGetBrowseInfo y devuelve un descriptor de contexto de sesión de examen para que otras funciones de examen lo utilicen.	251

Consejo

Muchas funciones API necesitan un descriptor de contexto de conexión (hdbc). El usuario debe proporcionar este descriptor de contexto utilizando la función `SQLConnect`, pero esto no impide que llame a DB2 Text Extender desde programas de SQL incorporado. La publicación *DB2 Call Level Interface Guide and Reference* describe cómo mezclar sentencias CLI con sentencias de SQL incorporado.

DesCloseDocument

Finalidad

Cierra un documento de texto abierto por DesOpenDocument y libera el almacenamiento asignado durante la devolución del texto del documento y la información de resaltado.

Sintaxis

```

DESRETURN
DesCloseDocument
    (DESBROWSESESSION    Sesiónexamen,
    DESHANDLE             Descriptorcontexto);
    
```

Argumentos de la función

Tabla 10. Argumentos de DesCloseDocument

Tipo de datos	Argumento	Uso	Descripción
DESBROWSESESSION	<i>BrowseSession</i>	entrada	Descriptor de contexto de sesión de examen.
DESHANDLE	<i>DocumentHandle</i>	entrada	Descriptor de contexto devuelto por DesOpenDocument que identifica un documento de texto abierto.

Códigos de retorno

```

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
    
```

Restricciones

Sólo puede llamarse a esta función después de haber abierto un documento de texto llamando a DesOpenDocument.

Función API DesEndBrowseSession

DesEndBrowseSession

Finalidad

Finaliza una sesión iniciada por DesStartBrowseSession y libera el almacenamiento asignado para la sesión de examen.

Sintaxis

```
DESRETURN  
DesEndBrowseSession  
(DESBROWSESESSION BrowseSession);
```

Argumentos de la función

Tabla 11. Argumentos de DesEndBrowseSession

Tipo de datos	Argumento	Uso	Descripción
DESBROWSESESSION	<i>BrowseSession</i>	entrada	Descriptor de contexto de sesión de examen.

Uso

Esta función no libera el almacenamiento asignado para la sesión de examen por DesGetBrowseInfo. Este almacenamiento contiene la información de examen que otra sesión de examen todavía puede utilizar. Para liberar este almacenamiento, llame a DesFreeBrowseInfo.

Códigos de retorno

```
RC_SUCCESS  
  
RC_INVALID_SESSION  
RC_INVALID_PARAMETER  
RC_SE_UNEXPECTED_ERROR
```

Restricciones

Sólo puede llamarse a esta función después de haber iniciado una sesión de examen llamando a DesStartBrowseSession.

DesFreeBrowseInfo

Finalidad

Libera el almacenamiento asignado para la información de examen por DesGetBrowseInfo.

Sintaxis

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

Argumentos de la función

Tabla 12. Argumentos de DesFreeBrowseInfo

Tipo de datos	Argumento	Uso	Descripción
DESBROWSEINFO	<i>BrowseInfo</i>	entrada	Información de examen

Códigos de retorno

RC_SUCCESS

RC_INVALID_PARAMETER

Restricciones

Sólo puede llamarse a esta función después de haber asignado el almacenamiento para la información de examen llamando a DesGetBrowseInfo.

Función API DesGetBrowseInfo

DesGetBrowseInfo

Finalidad

Recibe un argumento de búsqueda para buscar en el texto identificado por un descriptor de contexto. Devuelve un puntero para la información de examen necesaria para que DesStartBrowseSession resalte los términos encontrados.

Sintaxis

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

Argumentos de la función

Tabla 13. Argumentos de DesGetBrowseInfo

Tipo de datos	Argumento	Uso	Descripción
SQLHDBC	<i>hdbc</i>	entrada	Un descriptor de contexto de conexión de base de datos.
SQLCHAR *	<i>pHandle</i>	entrada	Puntero para un descriptor de contexto que se ha extraído de la base de datos.
DESUSHORT	<i>HandleLength</i>	entrada	Longitud de pHandle. DES_NTS no puede utilizarse aquí.
char *	<i>pSearchArgument</i>	entrada	Puntero para el argumento de búsqueda de texto que especifica la información que desea buscar.
DESSMALLINT	<i>ArgumentLength</i>	entrada	La longitud de pSearchArgument (sin incluir un terminador de byte nulo) o DES_NTS.
DESBROWSEINFO *	<i>pBrowseInfo</i>	salida	Puntero para la información de examen que contiene los datos necesarios para examinar un documento. Este puntero se pasa a DesStartBrowseSession.
DESMESSAGE *	<i>pErrorMessage</i>	salida	Texto de mensaje definido para la implementación. Si se produce un error, DB2 Text Extender devuelve un código de error y un mensaje de error. El programa de aplicación asigna el tamaño de almacenamiento intermedio DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> es el puntero nulo, no se devuelve ningún mensaje de error.

Uso

El programa de aplicación debe establecer una conexión con la base de datos antes de que llame a DesGetBrowseInfo.

Para el puntero para el argumento de búsqueda, se utiliza *char** en lugar de *SQLCHAR**. Esto es porque es posible que el valor del parámetro no venga de la base de datos.

Para la correlación entre los tipos de datos SQL y los tipos de datos C, debe utilizar el nombre simbólico SQL *SQL_VARBINARY* para un descriptor de contexto. El tipo de variables de sistema principal que apunta a la representación C de los valores de *Descriptor de contexto* es *SQLCHAR**.

DB2 Text Extender asigna el almacenamiento para la información de examen. El programa de aplicación debe liberar el almacenamiento y los recursos relacionados llamando a DesFreeBrowseInfo.

Puesto que los valores de *Descriptor de contexto* son datos de bits y contienen varios caracteres '\0', debe especificar la longitud de *pHandle*.

El argumento de búsqueda de *pSearchArgument* se describe en el "Capítulo 12. Sintaxis de los argumentos de búsqueda" en la página 215.

Códigos de retorno

```

RC_SUCCESS
RC_NO_BROWSE_INFO

RC_ALLOCATION_ERROR
RC_FILE_IO_PROBLEM
RC_INTERNAL_ERROR
RC_INVALID_PARAMETER
RC_PARSER_INVALID_ESCAPE_CHARACTER
RC_PARSER_INVALID_USE_OF_ESCAPE_CHAR
RC_PARSER_SYNTAX_ERROR
RC_SE_COMMUNICATION_PROBLEM
RC_SE_EMPTY_INDEX
RC_SE_EMPTY_QUERY
RC_SE_FUNCTION_DISABLED
RC_SE_FUNCTION_IN_ERROR
RC_SE_INCORRECT_HANDLE
RC_SE_INDEX_DELETED
RC_SE_INDEX_NOT_ACCESSIBLE
RC_SE_INDEX_SUSPENDED
RC_SE_INSTALLATION_PROBLEM
RC_SE_IO_PROBLEM
    
```

Función API DesGetBrowseInfo

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

Avisos: Los siguientes códigos de retorno indican que la función ha devuelto un resultado, pero puede que no sea tal como se esperaba.

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

Restricciones

Sólo puede llamarse a esta función después de haber realizado una conexión con la base de datos y utilizado una función de DB2 Text Extender para extraer un descriptor de contexto de la base de datos.

DesGetMatches

Finalidad

Devuelve una corriente de datos que contiene información de resaltado para el documento de texto descrito por un descriptor de contexto de documento. Consulte el apartado “Sintaxis de la corriente de datos” en la página 238. La información de resaltado comprende el contexto del texto (un párrafo como mínimo) y la información para resaltar el texto en ese contexto.

DesGetMatches sólo devuelve una parte de la corriente de datos, indicando la longitud de la parte en la estructura de salida.

Una secuencia de llamadas a DesGetMatches obtiene todo el contenido del documento de texto. Cuando se alcanza el final del documento de texto, se devuelve RC_SE_END_OF_INFORMATION.

Sintaxis

```

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

Argumentos de la función

Tabla 14. Argumentos DesGetMatches

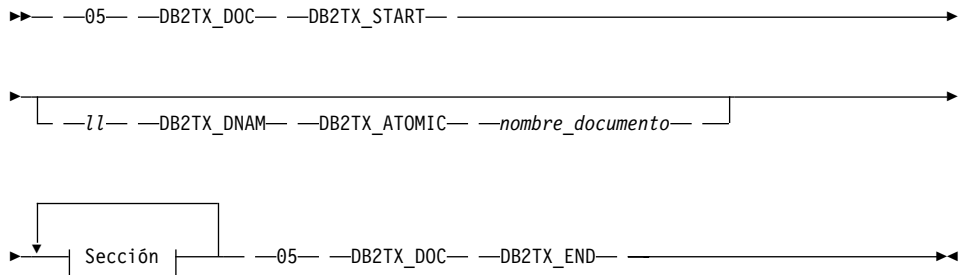
Tipo de datos	Argumento	Uso	Descripción
DESBROWSESESSION	<i>BrowseSession</i>	entrada	Descriptor de contexto de sesión de examen.
DESHANDLE	<i>DocumentHandle</i>	entrada	Descriptor de contexto de documento devuelto por DesOpenDocument.
DESMATCHINFO *	<i>pMatchInfo</i>	salida	Puntero para un almacenamiento intermedio que contiene la parte de la corriente de datos recibida. DesGetMatches asigna ese almacenamiento intermedio.
DESULONG *	<i>pMatchInfoLength</i>	salida	La longitud de la parte de la corriente de datos a la que apunta pMatchInfo.

Función API DesGetMatches

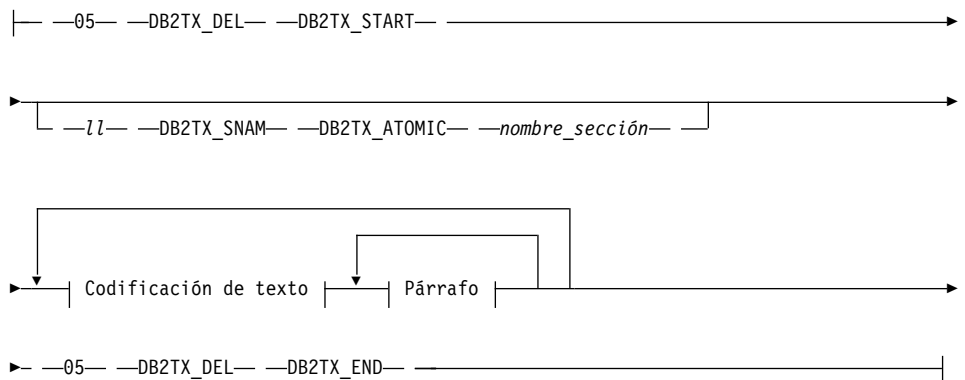
Tabla 14. Argumentos DesGetMatches (continuación)

Tipo de datos	Argumento	Uso	Descripción
DESMESSAGE *	<i>pErrorMessage</i>	salida	Texto de mensaje definido para la implementación. Si se produce un error, DB2 Text Extender devuelve un código de error y un mensaje de error. El programa de aplicación asigna el tamaño de almacenamiento intermedio DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> es el puntero nulo, no se devuelve ningún mensaje de error.

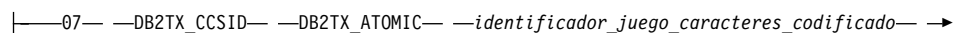
Sintaxis de la corriente de datos

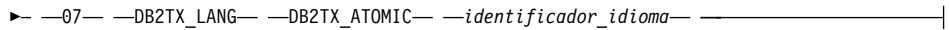


Sección:

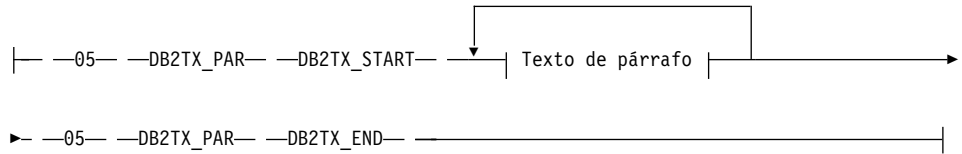


Codificación de texto:

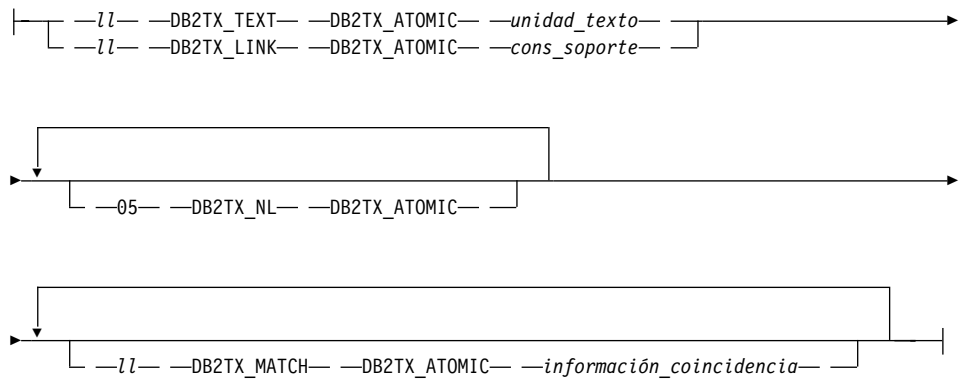




Párrafo:



Texto de párrafo:



Cada segmento de un diagrama de sintaxis como, por ejemplo, 05 DB2TX_DOC DB2TX_START, empieza por un campo de longitud de tipo entero, que en el diagrama es un número explícito como 05 o una variable *ll*. La longitud del segmento incluye el campo de longitud de 2 bytes.

Nota: La longitud está en formato big-endian.

Cada segmento incluye uno de los siguientes identificadores de tipo de 1 byte:

DB2TX_START

Indica el inicio de un segmento como, por ejemplo, un documento o un párrafo.

DB2TX_END

Indica el final de un segmento.

Función API DesGetMatches

DB2TX_ATOMIC

Indica que el elemento que sigue es de átomo como, por ejemplo, un nombre de documento o un identificador de idioma.

Los elementos de corriente de datos tienen cada uno dos bytes de longitud. Son:

DB2TX_DOC

Indica el inicio y el final de un documento.

DB2TX_DNAM

Un nombre de documento. Si no se especifica ningún nombre, se utiliza el identificador del documento.

DB2TX_DEL

Indica el inicio y el final de un elemento de documento. El único tipo de elemento de documento actualmente soportado es una *sección de texto*.

DB2TX_SNAM

Especifica el nombre de una sección de texto. Actualmente, DB2 Text Extender soporta una sola sección de texto y suministra automáticamente un nombre por omisión. Si especifica un nombre de sección, se pasa por alto.

DB2TX_PAR

Indica el inicio y el final de un párrafo de texto dentro de la sección actual.

DB2TX_TEXT

Especifica una parte de texto dentro del párrafo actual. Normalmente, la *unidad de texto* contiene una línea de texto y el elemento TEXT va seguido por un elemento DB2TX_NL; pero las líneas de texto también se pueden dividir en varias partes, cada una especificada en su propio elemento DB2TX_TEXT.

El texto utiliza el CCSID y el idioma asociado al párrafo actual.

DB2TX_LINK

Especifica una consulta hipertexto de DB2 Text Extender. Utiliza el CCSID del párrafo actual.

DB2TX_NL

Indica el inicio de una nueva línea del párrafo actual.

DB2TX_MATCH

Contiene la información de apariciones de las coincidencias en la parte de texto actual. La información se suministra como una secuencia de pares de números binarios. El primer número de cada par es el desplazamiento de una coincidencia en la parte de texto actual, el segundo número es la longitud en caracteres de esa coincidencia. La longitud dada podría

exceder de la parte de texto dada. Tanto el desplazamiento como la longitud son valores de dos bytes especificados en formato big-endian.

DB2TX_CCSID

El CCSID para el texto de los párrafos posteriores hasta que un párrafo vaya precedido por un nuevo elemento DB2TX_CCSID. Se devuelven los siguientes CCSID:

DB2TX_CCSID_00500

para el texto en la página de códigos 500 de EBCDIC Latin-1.

DB2TX_CCSID_04946

para el texto en la página de códigos 850 de ASCII Latin-1.

DB2TX_CCSID_00819

para el texto en la página de códigos 819 de ASCII.

Esos nombres simbólicos para los CCSID se definen en el archivo DES_EXT.H proporcionado por DB2 Text Extender. Los valores binarios de dos bytes se especifican en formato big-endian.

DB2TX_LANG

El identificador del idioma para el texto de los párrafos posteriores hasta que un párrafo vaya precedido por un nuevo elemento DB2TX_LANG. El archivo DES_EXT.H proporcionado con DB2 Text Extender define nombres simbólicos para todos los identificadores soportados por DB2 Text Extender. Los valores binarios de dos bytes se especifican en formato big-endian.

Uso

DesGetMatches devuelve RC_SE_END_OF_INFORMATION cuando se alcanza el final del documento de texto.

Códigos de retorno

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

Avisos: Los siguientes códigos de retorno indican que la función ha devuelto un resultado, pero puede que no sea tal como se esperaba.

Función API DesGetMatches

RC_SE_DICTIONARY_NOT_FOUND

Restricciones

Sólo puede llamarse a esta función después de haber abierto un documento de texto llamando a DesOpenDocument.

DesGetSearchResultTable

Finalidad

Utiliza un argumento de búsqueda para buscar en documentos de texto identificados por una columna de texto. Los datos del descriptor de contexto de los elementos de texto encontrados se graban en una tabla de resultados. La información de examen acerca del rango y del número de coincidencias también puede grabarse en la tabla de resultados.

Sintaxis

```

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);
    
```

Argumentos de la función

Tabla 15. Argumentos de DesGetSearchResultTable

Tipo de datos	Argumento	Uso	Descripción
SQLHDBC	hdbc	entrada	Un descriptor de contexto de conexión de base de datos.
char *	pTableSchema	entrada	El esquema de la tabla base en la que se ha de buscar.
DESSMALLINT	TableSchemaLength	entrada	La longitud de pTableSchema (sin incluir un terminador de byte nulo) o DES_NTS.
char *	pTableName	entrada	Puntero para el nombre de la tabla base en la que se ha de buscar.
DESSMALLINT	TableNameLength	entrada	La longitud de pTableName (sin incluir el terminador de byte nulo) o DES_NTS.

Función API DesGetSearchResultTable

Tabla 15. Argumentos de DesGetSearchResultTable (continuación)

Tipo de datos	Argumento	Uso	Descripción
char *	pColumnName	entrada	Puntero para el nombre de la columna a la que se ha de direccionar la búsqueda de texto que se intenta realizar. La columna debe ser de tipo DESTEXTH.
DESSMALLINT	ColumnNameLength	entrada	La longitud de pColumnName (sin incluir el terminador de byte nulo) o DES_NTS.
char *	pSearchArgument	entrada	Puntero para el argumento de búsqueda de texto.
DESSMALLINT	ArgumentLength	entrada	La longitud de pSearchArgument (sin incluir el terminador de byte nulo) o DES_NTS.
char *	pResultSchema	entrada	Puntero para el esquema que contiene la tabla de resultados.
DESSMALLINT	ResultSchemaLength	entrada	La longitud de pSchemaName (sin incluir el terminador de byte nulo) o DES_NTS.
char *	pResultTableName	entrada	Puntero para el nombre de la tabla de resultados que ha creado previamente en la que se ha de almacenar el resultado de la búsqueda. Consulte la Figura 17 en la página 245 para ver la estructura de esta tabla.
DESSMALLINT	ResultTableNameLength	entrada	La longitud de pResultTableName (sin incluir el terminador de byte nulo) o DES_NTS.
DESSEARCHOPTION	SearchOption	entrada	<p>Una opción que determina si está pidiendo información de rango, para el número de coincidencias o sólo para los descriptores de contexto de los documentos de texto coincidentes.</p> <p>DES_RANK DES_MATCH DES_RANKANDMATCH DES_TEXTHANDLEONLY</p> <p>Esta opción determina el contenido de la tabla de resultados, tal como se describe en el apartado “Uso” en la página 245.</p>
DESBROWSEOPTION	BrowseOption	entrada	Reservado.

Tabla 15. Argumentos de DesGetSearchResultTable (continuación)

Tipo de datos	Argumento	Uso	Descripción
DESBROWSEINFO *	pBrowseInfo	salida	Puntero para la información de examen o un puntero para nulo, dependiendo del valor de BrowseOption.
DESMESSAGE *	pErrorMessage	salida	Texto de mensaje definido para la implementación. Si se produce un error, DB2 Text Extender devuelve un código de error y un mensaje de error. El programa de aplicación asigna el tamaño de almacenamiento intermedio DES_MAX_MESSAGE_LENGTH. Si pErrorMessage es el puntero nulo, no se devuelve ningún mensaje de error.

Uso

La conexión con la base de datos debe establecerla el programa de aplicación llamando a DesGetSearchResultTable.

El nombre *pResultTableName* hace referencia a una tabla de resultados que se ha creado con anterioridad. El programa de utilidad DESRESTB del directorio de ejemplos crea una tabla de resultados para los descriptores de contexto de texto. Después de llamar a esta función, la tabla de resultados contiene información que identifica los valores de texto que coinciden con el argumento de búsqueda. Esta es la estructura de la tabla de resultados:

RESULT TABLE

TEXTHANDLE	RANK	MATCHES

Figura 17. Estructura de la tabla de resultados

El tipo de datos de TEXTHANDLE es DB2TEXTH o DB2TEXTFH. El tipo de datos de RANK es DOUBLE. El tipo de datos de MATCHES es INTEGER.

El argumento de búsqueda de *pSearchArgument* se describe en el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215.

Si el valor de *BrowseOption* es *BROWSE*, DB2 Text Extender devuelve información de examen desde la aplicación de búsqueda de DB2 Text Extender ubicada en el servidor. *pBrowseInfo* apunta a la información de examen que es la entrada para DesStartBrowseSession. Si el valor de *BrowseOption* es *NO_BROWSE* *pBrowseInfo* apunta a nulo.

Función API DesGetSearchResultTable

Códigos de retorno

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

Avisos: Los siguientes códigos de retorno indican que la función ha devuelto un resultado, pero puede que no sea tal como se esperaba.

RC_SE_CONFLICT_WITH_INDEX_TYPE
RC_SE_DICTIONARY_NOT_FOUND

Función API DesGetSearchResultTable

RC_SE_STOPWORD_IGNORED
RC_SE_UNKNOWN_SECTION_NAME
RC_SE_DOCMOD_READ_PROBLEM

Función API DesOpenDocument

DesOpenDocument

Finalidad

Recibe un puntero de sesión de examen, un descriptor de contexto y una opción (DES_EXTENDED o DES_FAST) que indica si el documento de texto debe analizarse con o sin la utilización de un diccionario. Prepara el documento de texto que corresponde al descriptor de contexto para obtener el texto del documento y la información de resaltado y devuelve un descriptor de contexto de documento que se utiliza para llamar iterativamente a DesGetMatches.

Sintaxis

```
DESRETURN  
DesOpenDocument  
(DESBROWSESESSION BrowseSession,  
SQLCHAR *pHandle,  
DESUSHORT HandleLength,  
DESMATCHMODE MatchMode,  
DESHANDLE *pDocumentHandle,  
DESMESSAGE *pErrorMessage);
```

Argumentos de la función

Tabla 16. Argumentos de DesOpenDocument

Tipo de datos	Argumento	Uso	Descripción
DESBROWSESESSION	<i>BrowseSession</i>	entrada	Descriptor de contexto de sesión de examen.
SQLCHAR *	<i>pHandle</i>	entrada	Puntero para el descriptor de contexto extraído de la base de datos.
DESUSHORT	<i>HandleLength</i>	entrada	Longitud de pHandle (no se puede utilizar DES_NTS).
DESMATCHMODE	<i>MatchMode</i>	entrada	Modalidad para determinar si se utiliza un diccionario para buscar la información de resaltado. DES_FAST No utilizar un diccionario DES_EXTENDED Utilizar un diccionario
DESHANDLE *	<i>pDocumentHandle</i>	salida	Un descriptor de contexto de documento para llamar iterativamente a DesGetMatches.

Tabla 16. Argumentos de DesOpenDocument (continuación)

Tipo de datos	Argumento	Uso	Descripción
DESMESSAGE *	<i>pErrorMessage</i>	salida	Texto de mensaje definido para la implementación. Si se produce un error, DB2 Text Extender devuelve un código de error y un mensaje de error. El programa de aplicación asigna el tamaño de almacenamiento intermedio DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> es el puntero nulo, no se devuelve ningún mensaje de error.

Uso

DES_FAST y DES_EXTENDED hacen referencia al uso del proceso lingüístico para buscar los términos que se han de resaltar en el texto examinado. Consulte el apartado “Proceso lingüístico para examen” en la página 266 para obtener más información. Especifique DES_FAST para utilizar el análisis de texto básico y DES_EXTENDED para utilizar la comparación ampliada.

Para la correlación entre los tipos de datos SQL y los tipos de datos C, debe utilizar el nombre simbólico SQL SQL_VARBINARY para un descriptor de contexto. El tipo de variables de sistema principal que apuntan a la representación C de valores *TextHandle* es SQLCHAR*.

DB2 Text Extender asigna el almacenamiento para la información de examen. El programa de aplicación debe liberar este almacenamiento y los recursos relacionados llamando a DesFreeBrowseInfo.

Puesto que los valores de *TextHandle* son datos de bits y contienen varios caracteres '\0', debe especificar la longitud de *pHandle*.

El llamante debe tener acceso de lectura para la tabla que contiene el documento de texto al que *pHandle* hace referencia.

Códigos de retorno

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

Función API DesOpenDocument

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

Restricciones

Sólo puede llamarse a esta función después de haber iniciado una sesión de examen llamando a DesStartBrowseSession.

DesStartBrowseSession

Finalidad

Inicia una sesión de examen, estableciendo el entorno necesario para examinar un documento de texto y resaltar sus coincidencias. Recibe un puntero para la información de examen, de DesGetBrowseInfo o de DesGetSearchResultTable, y devuelve un descriptor de contexto de sesión de examen para que lo utilicen las demás funciones de examen.

Sintaxis

```

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

Argumentos de la función

Tabla 17. Argumentos de DesStartBrowseSession

Tipo de datos	Argumento	Uso	Descripción
DESBROWSEINFO	<i>BrowseInfo</i>	entrada	Puntero para la información necesario para examinar y resaltar las coincidencias de un documento de texto. El puntero lo devuelve DesGetSearchResultTable o DesGetBrowseInfo.
char *	<i>pUserId</i>	entrada	ID de usuario para la base de datos
DESSMALLINT	<i>UserIdLength</i>	entrada	Longitud del ID de usuario para la base de datos
char *	<i>pPassword</i>	entrada	Contraseña para la base de datos
DESSMALLINT	<i>PasswordLength</i>	entrada	Longitud de la contraseña para la base de datos
DESBROWSESESSION *	<i>pBrowseSession</i>	salida	Un descriptor de contexto para una sesión de examen para que la utilicen otras funciones de examen.

Función API DesStartBrowseSession

Tabla 17. Argumentos de DesStartBrowseSession (continuación)

Tipo de datos	Argumento	Uso	Descripción
DESMESSAGE *	<i>pErrorMessage</i>	salida	Texto de mensaje definido para la implementación. Si se produce un error, DB2 Text Extender devuelve un código de error y un mensaje de error. El programa de aplicación asigna el tamaño de almacenamiento intermedio DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> es el puntero nulo, no se devuelve ningún mensaje de error.

Uso

Esta función abre una sesión de examen para examinar documentos de texto. Se le solicita el ID de usuario y la contraseña para comprobar la autorización para acceder a la base de datos.

La sesión de examen se cierra llamando a DesEndBrowseSession.

BrowseInfo depende del argumento de búsqueda y de la columna de texto base utilizada para crear la información de examen.

Códigos de retorno

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

Restricciones

Debe llamar a DesGetBrowseInfo o DesGetSearchResultTable con la opción de examen adecuada antes de llamar a esta función.

Capítulo 14. Programa API de ejemplo

DB2 Text Extender proporciona un programa de ejemplo, DESSAMP1.C, ubicado en el directorio SAMPLES.

DESSAMP1.C es un ejemplo de programa que utiliza la función DesGetSearchResultTable y su propio programa examinador. Sigue la secuencia de las llamadas a funciones API que aparecen en la Figura 16 en la página 119.

El usuario debe acceder a una base de datos habilitada y a una columna de texto habilitada. Para ejecutar este programa, haga lo siguiente:

1. Opcional. Copie el archivo fuente DESSAMP1.C en un directorio local de la máquina cliente.
2. Utilice el archivo de creación suministrado (dessamp.mak para OS/2, Windows NT y Windows 2000; dessamp para sistemas UNIX) para compilar y enlazar los archivos de ejemplo.
3. Ejecute el programa de utilidad DESRESTB para crear una tabla de resultados en la base de datos que piensa utilizar con el código de ejemplo:

```
DESRESTB nombre-base-datos
```

Esta tabla se utiliza para almacenar información como, por ejemplo, los resultados de la búsqueda; tiene la estructura siguiente:

Columna	Tipo de datos
HANDLE	DB2TX.DB2TEXTH o DB2TX.DB2TEXTFH
RANK	DOUBLE
MATCHES	INTEGER

Programa API de ejemplo

Capítulo 15. Proceso lingüístico para los índices lingüístico y exacto

DB2 Text Extender ofrece el proceso lingüístico en estas áreas de recuperación:

- **Indexación.** Cuando DB2 Text Extender analiza documentos para extraer los términos que se han de almacenar en el índice de texto, el texto se procesa lingüísticamente para extraer los términos correctos para el índice. Esto se hace para que la recuperación sea lo más sencilla y rápida posible.
- **Recuperación.** Cuando DB2 Text Extender busca los documentos que contienen apariciones de los términos de búsqueda que ha especificado, en el índice de documentos, los términos de búsqueda también se procesan lingüísticamente para compararlos con los términos indexados.
- **Examen.** Cuando examina un documento que se ha encontrado después de una búsqueda, se utiliza el proceso lingüístico para resaltar los términos encontrados en el documento.

Proceso lingüístico al indexar

Cuando DB2 Text Extender indexa y recupera documentos, realiza un análisis lingüístico del texto. Como puede ver en la tabla siguiente, la cantidad de proceso lingüístico depende del tipo de índice. Para índices Ngram, no se aplica ningún proceso lingüístico.

El proceso lingüístico utilizado para la indexación de documentos consta de:

- Análisis básico de texto
 - Reconocimiento de términos (señalización)
 - Normalización de términos a un formato estándar
 - Reconocimiento de sentencias
- Reducción de términos a su formato base
- Filtro de palabras de parada
- Descomposición (división de términos compuestos).

La Tabla 18 en la página 256 muestra un resumen de cómo se indexan los términos cuando el tipo de índice es **lingüístico** y no se han pedido propiedades de índice adicionales.

Proceso lingüístico al indexar

Tabla 18. Extracción de términos para un índice lingüístico

Texto de documento	Término en índice	Proceso lingüístico
Ratón Käfer	ratón kaefer	Análisis básico de texto (normalización)
ratones nadado	ratón nadar	Reducción al formato base
coche-cama Wetterbericht	coche-cama, coche cama wetterbericht, wetter bericht	Descomposición
un informe sobre animales	informe animal	Filtro de palabras de parada. Las palabras de parada son: un, sobre

Por comparación, la Tabla 19 muestra un resumen de cómo se indexan los términos cuando el tipo de índice es **exacto**.

Tabla 19. Extracción de términos para un índice exacto

Texto de documento	Término en índice	Proceso lingüístico
Ratón Käfer	Ratón Käfer	No hay normalización
ratones nadado	ratones nadado	No hay reducción al formato base
un informe sobre animales	informe animales	Filtro de palabras de parada. Las palabras de parada son: un, sobre
coche-cama Wetterbericht	coche-cama Wetterbericht	No hay descomposición

Análisis básico de texto

DB2 Text Extender procesa el análisis básico de texto sin utilizar ningún diccionario electrónico.

Reconocimiento de términos que contienen caracteres no alfanuméricos

Cuando se indexan los documentos, se reconocen los términos incluso si contienen caracteres no alfanuméricos por ejemplo: "\$14.225,23", "sofá-cama" y "22/10/90".

Los siguientes elementos se toman como parte del término:

Acentos

Signos monetarios

Caracteres separadores de números (por ejemplo, “/” o “.”)

El carácter “@” de una dirección de correo electrónico (sólo en inglés)

El signo “+”.

También se utilizan las normas específicas del idioma para reconocer los términos que contienen:

- Apóstrofes en idiomas románicos como por ejemplo, *l'aventure* en francés.
- Formatos nacionales para fechas, horas y números.
- Alternativas como, por ejemplo, *mission/responsibility*, indicadas en inglés mediante la utilización del carácter “/”.
- Los apóstrofes de cola en las palabras italianas como, por ejemplo, *securita'*. Es normal que en los textos escritos en italiano, cuando el juego de caracteres no incluye los caracteres acentuados, se escriba el acento *después* del carácter; por ejemplo, “à” se escribe “a”.

Normalización de términos a un formato estándar

La normalización reduce los términos escritos en mayúsculas y minúsculas y los términos que contienen caracteres especiales acentuados a un formato estándar. Esto se hace por omisión cuando el tipo de índice es lingüístico. (En un índice exacto las letras mayúsculas y minúsculas no se cambian—las búsquedas son sensibles a las mayúsculas y minúsculas.)

Por ejemplo, el término *Computer* se indexa como *computer*, la letra en mayúscula se cambia a minúscula. Una búsqueda del término *computer* encuentra las apariciones de *computer* y también de *Computer*. El efecto de la normalización durante la indexación es que los términos se indexan de la misma manera, sin tener en cuenta las mayúsculas del documento.

La normalización no sólo se aplica durante la indexación, sino también durante la recuperación. Los caracteres en mayúsculas de un término de búsqueda se cambian a minúsculas cuando se realiza la búsqueda. Cuando el término de búsqueda es, por ejemplo, *Computer*, el término utilizado en la búsqueda es *computer*.

Los caracteres acentuados y especiales se normalizan de manera similar. Cualquier variación de *écolé* como, por ejemplo, *Écolé*, encuentra *écolé*, *Ecolé*, etcétera. *Bürger* encuentra *buerger*, *Maße* encuentra *masse*.

Si el término de búsqueda incluye caracteres de máscara (comodín), la normalización se realiza antes de procesar los caracteres de máscara. Ejemplo: *Bür_er* se convierte en *buer_er*.

Proceso lingüístico al indexar

Reconocimiento de sentencias

Puede buscar términos que aparecen en la misma sentencia. Para que esto sea posible, se analiza cada documento durante la indexación para averiguar donde finaliza cada sentencia.

DB2 Text Extender ofrece dos tipos de reconocimiento de final de sentencia:

- Universal Unicode Tokenizer para los idiomas que no sean el árabe y el hebreo.

Este es el método más sencillo, pero más rápido. El señalizador busca un punto, el signo de admiración o de interrogación, precedido por un carácter de señal como, por ejemplo, una letra y seguido de un carácter blanco, tabulador o de nueva línea. Para comprobar que es realmente el final de una sentencia y no sólo una abreviatura que finaliza con un punto, se comprueba la lista de abreviaturas específica del idioma.

- El señalizador basado en POE para el árabe y el hebreo

Este señalizador es lingüísticamente más avanzado, pero necesita más poder de proceso. El señalizador encuentra el final de las sentencias principalmente a través de la comparación de las puntuaciones, pero también de las entradas de tipo especial y el número de palabras.

Señalizador basado en POE para el reconocimiento de sentencias

El señalizador basado en POE determina los límites de la sentencia (o fragmento de sentencia) utilizando las reglas de puntuación y el proceso específico del idioma que implica el proceso de abreviaturas, aunque el nivel de funcionamiento varía ampliamente según el idioma. La mayoría de idiomas que utilizan páginas de códigos de un solo byte tienen un diccionario suplementario de abreviaturas que se proporciona con POE. Puesto que normalmente los idiomas de doble byte no emplean abreviaturas con períodos, no hay diccionarios suplementarios de abreviaturas disponibles para estos idiomas.

La determinación del final de una sentencia se realiza principalmente a través de la comparación de la puntuación. La tabla siguiente lista los caracteres de puntuación de terminación y sus GCGID (identificadores globales de caracteres gráficos).

GCGID de caracteres SBCS	GCGID de caracteres DBCS	Descripción
SP110000	SP110080	Punto
SP020000	SP020080	Signo de admiración
SP150000	SP150080	Signo de interrogación
SP140000	N/D	Punto y coma (signo de interrogación griego)

GCGID de caracteres SBCS	GCGID de caracteres DBCS	Descripción
N/D	JQ730080	Período de círculo de doble byte

Se supone que un carácter de puntuación de terminación como, por ejemplo, un punto, un signo de admiración o un signo de interrogación, marca el final de una sentencia a menos que se produzca alguna de las situaciones siguientes:

- El carácter de puntuación de terminación va seguido de un carácter de puntuación de cierre de la tabla siguiente como, por ejemplo, cerrar paréntesis.

GCGID de caracteres SBCS	GCGID de caracteres DBCS	Descripción
SP070000	SP070080	Cerrar paréntesis
SP040000	SP040080	Comillas dobles
SP050000	SP050080	Comillas simples
SP180000	SP070083	Comillas dobles en ángulo
N/D	SM140080	Cerrar llaves
N/D	SM080080	Cerrar corchetes
N/D	JQ720080	Comillas simples cuadradas
N/D	JQ720081	Comillas dobles cuadradas
N/D	SP200080	Cerrar comillas simples de corchete
N/D	SP220080	Cerrar comillas dobles de corchete
N/D	SP070081	Cerrar corchetes de caparazón
N/D	SP070082	Cerrar comillas simples en ángulo
N/D	SP070084	Cerrar paréntesis en pico
N/D	SP370080	Cierre vertical de comillas simples cuadradas
N/D	SP370081	Cierre vertical de comillas dobles cuadradas
N/D	SP250084	Cierre vertical de paréntesis cuadrados

Proceso lingüístico al indexar

GCGID de caracteres SBCS	GCGID de caracteres DBCS	Descripción
N/D	SP250080	Cierre vertical de paréntesis
N/D	SP350080	Cierre vertical de llaves
N/D	SP250081	Cierre vertical de corchetes de caparazón
N/D	SP250083	Cierre vertical de comillas dobles en ángulo
N/D	SP250082	Cierre vertical de comillas simples en ángulo

Nota: POE no considera los elementos marcados con N/D como caracteres de puntuación de cierre y los caracteres de cierre vertical sólo están soportados en chino.

Ejemplo:

...esta frase finaliza con dos paréntesis.))

En el ejemplo, el segundo paréntesis se detecta como final de la sentencia. Sin embargo, en alemán nacional, un signo de cerrar comillas no se considera como marca de final de sentencia si va seguido de una coma.

- El carácter de puntuación de terminación va seguido de otro carácter de terminación. Ejemplo:

iiiEsta es una gran admiración!!!

El signo de admiración final se detecta como final de la sentencia.

- El carácter de puntuación de terminación va precedido de un carácter numérico o de puntuación y seguido de un carácter numérico. Esto impide que series como, por ejemplo, '1,25' y '0,314' finalicen una sentencia.
- El carácter de puntuación de terminación es un punto y forma parte de una abreviatura que no está permitida al final de una sentencia. El proceso de abreviaturas limitadas se realiza para cada idioma.
- El carácter de puntuación de terminación es un punto y no va seguido de un carácter de espacio en blanco como, por ejemplo, un blanco o un carácter de nueva línea. Esto es para evitar que las cabeceras como, por ejemplo, 'III.IV' se detecten como final de una sentencia.

El señalizador basado en POE realiza un proceso de abreviaturas para determinar si un punto forma parte de una abreviatura o si marca el final de una sentencia. Puede añadir abreviaturas a un diccionario suplementario de abreviaturas. Si no se pasa ningún diccionario al señalizador basado en POE, todas las letras solas seguidas de puntos se marcan como abreviaturas; no tiene lugar ningún otro proceso de abreviaturas.

Decidir si un fragmento de texto es una abreviatura o no es con frecuencia ambiguo, porque puede tomarse una palabra normal seguida de un punto como una abreviatura. Por ejemplo, tome en consideración los caracteres "no." en las sentencias siguientes:

Entre el no. de exenciones que está reclamando.
Conteste a la pregunta con un sí o un no.

Pero incluso cuando se sabe que un fragmento de texto es una abreviatura, todavía sigue existiendo la ambigüedad de si finaliza una sentencia. Algunas abreviaturas nunca finalizan una sentencia, mientras otras sí que lo hacen a veces. Por ejemplo, considere la utilización de la abreviatura "Hwy." en las sentencias en inglés siguientes:

The drive along Hwy. 1 to Santa Cruz was beautiful.
Many people speak highly of the Pacific Coast Hwy.

Puesto que las abreviaturas pueden ser ambiguas y que algunas abreviaturas no pueden aparecer al final de una sentencia, POE intenta clasificar las abreviaturas encontradas. Si se encuentra un punto formando parte de una abreviatura que a veces finaliza una sentencia, se realiza un proceso más profundo. Si POE determina que la abreviatura no está al final de la sentencia, la señal que representa el punto se une a la señal para el texto de la abreviatura. De lo contrario, la señal que representa el punto permanece en una señal separada.

El proceso de abreviaturas basado en POE utiliza tres conjuntos de criterios para determinar si un punto forma parte de una abreviatura:

- Todas las letras solas seguidas de un punto se consideran abreviaturas. Las abreviaturas de una sola letra se clasifican como posibles finales de sentencias.
- Las palabras contenidas en el diccionario suplementario de abreviaturas siempre se consideran abreviaturas. La determinación de si una abreviatura que se encuentra en el diccionario suplementario puede finalizar una sentencia se toma por la información asociada con esa palabra en el diccionario suplementario. Por ejemplo, "Mr." se marca en el diccionario suplementario de abreviaturas inglés de EE.UU. como una abreviatura que no puede finalizar una sentencia, mientras que "etc." es una abreviatura que algunas veces puede finalizar una sentencia.
- Cualquier palabra de dos a seis caracteres de longitud seguida de un punto y que no se encuentra en ninguno de los diccionarios ni en el diccionario suplementario de abreviaturas también pueden considerarse abreviaturas. Esto es para manejar casos como "Jrnl. Comp. Ling.". Las abreviaturas determinadas por la búsqueda en diccionario siempre se tratan como posibles finalizaciones de sentencias.

Proceso lingüístico al indexar

Si se identifica una abreviatura como posible final de sentencia, POE examina el texto que sigue a la abreviatura para determinar si está al final de la sentencia actual comprobando si la siguiente palabra empieza por una letra en mayúsculas.

Si una abreviatura va seguida por dos o más elementos de datos de nueva línea, nueva sentencia o nuevo párrafo, POE supone que se ha alcanzado un final de sentencia. También, si el texto posterior es un signo de abrir interrogación, abrir admiración, se inserta un marcador de final de sentencia en la salida.

Si POE determina que el punto forma parte de una abreviatura que no finaliza una sentencia, continúa su búsqueda de un delimitador de sentencia. De lo contrario, comprueba las demás condiciones de excepción de carácter de puntuación de terminación (la siguiente puntuación de terminación o puntuación de cierre) antes de marcar un final de sentencia.

Reducción de términos a su formato base (lematización)

En un índice lingüístico, puede buscar ratón, por ejemplo y encontrar ratones. Los términos se reducen a su formato base para la indexación; el término ratones se indexa como ratón. Después, cuando se utiliza el término de búsqueda ratón, se encuentra el documento. También se encuentra el documento si busca ratones.

El efecto es que se encuentran los documentos que contienen información sobre los ratones sin tener en cuenta la variación del término ratón que aparece en el documento ni si ese utiliza como término de búsqueda.

De la misma manera, los verbos conjugados se reducen a su infinitivo; comprado, por ejemplo, pasa a comprar.

Filtro de palabras de parada

Las palabras de parada como, por ejemplo, preposiciones y pronombres, que aparecen muy frecuentemente en los documentos y que, por lo tanto, no son adecuados como términos de búsqueda. Dichas palabras están en una lista de palabras de parada asociada a cada diccionario y se excluyen del proceso de indexación.

El proceso de palabras de parada no es sensible a las mayúsculas y las minúsculas. Por lo tanto la palabra de parada acerca excluye también la primera palabra de una sentencia Acerca. Las listas de palabras de parada, suministradas en varios idiomas, pueden modificarse.

Un índice Ngram no tiene una lista de palabras de parada.

Descomposición (división de términos compuestos)

Los idiomas germánicos, como el alemán u holandés, son ricos en términos compuestos como, por ejemplo, Versandetiketten, que significa etiquetas (Etiketten) de correos (Versand). Estos términos compuestos pueden dividirse en sus componentes.

Para un índice exacto, los términos compuestos se indexan sin cambiarlos como una palabra. Para un índice lingüístico, los términos compuestos se dividen durante la indexación. Cuando se realiza una búsqueda, los términos compuestos se dividen si tiene un índice lingüístico.

Se encuentran los componentes si aparecen en cualquier orden en un documento siempre que estén contenidos en una sola sentencia. Por ejemplo, cuando se busca la palabra en alemán Wetterbericht (informe del tiempo), también se encontraría un documento que contenga la frase Bericht über das Wetter (informe acerca del tiempo).

Se efectúa un intento de dividir un término si:

- El idioma del término utiliza términos compuestos
- El término tiene una longitud mínima determinada
- El término no es en sí mismo una entrada en el diccionario electrónico—las palabras compuestas que se utilizan comúnmente como la palabra alemana Geschäftsbericht (informe comercial) *están* en el diccionario alemán.

Si se encuentra que es posible una división, las partes componentes del término también se reducen a su formato base. Estos son algunos ejemplos del danés, alemán y holandés:

Término compuesto	Partes componentes
børsmæglerselskab	børsmæglerselskab børs mægler selskab
Kindersprachen	kindersprache kind sprache
probleemkinderen	probleemkinderen probleemkind kind probleem

Proceso lingüístico para la recuperación

El proceso lingüístico aspira a debilitar los términos de búsqueda para que se incremente la velocidad de llamada de las búsquedas, es decir, se encuentren más documentos relevantes. Hay dos operaciones básicas sobre términos de consulta para conseguir este objetivo; las expansiones y las reducciones. Además, algunas operaciones de términos de búsqueda implican tanto la ampliación como la reducción.

- Las ampliaciones toman un término de una palabra o de múltiples palabras de un término de búsqueda y lo asocian a un conjunto de términos de búsqueda alternativos, cada uno de los cuales puede ser un término de múltiples palabras a su vez. La expresión fuente y el conjunto de expresiones de destino forman una expresión OR booleana en el lenguaje de consulta de DB2 Text Extender. Puesto que las ampliaciones no cambian el término fuente, hasta cierto límite son independientes del tipo de índice. Las siguientes operaciones son ampliaciones:

Ampliación de sinónimos

Ampliación de tesoro

- Las reducciones cambian el término de búsqueda a un formato que es más general que el especificado por el usuario. Puesto que cambia el término de búsqueda, las reducciones dependen del tipo de índice para asegurar que el término cambiado coincide. Por lo tanto, DB2 Text Extender deriva la información de reducción del tipo de estos índices o índice a los que se dirige la consulta. Las siguientes operaciones son reducciones:

Lematización (consulte el apartado “Reducción de términos a su formato base (lematización)” en la página 262)

Normalización (consulte el apartado “Normalización de términos a un formato estándar” en la página 257).

Palabras de parada (consulte el apartado “Filtro de palabras de parada” en la página 262).

- Algunas operaciones cambian el término de búsqueda y lo amplían con un conjunto de términos alternativos. Debido a la reducción inherente, éstos dependen de nuevo de la información contenida en el índice. Las siguientes operaciones pertenecen a esta clase:

Enmascaramiento de caracteres y de palabras

Ampliación de sonido.

Sinónimos

Los sinónimos son palabras relacionadas semánticamente. Normalmente, estas palabras tienen la misma clase o clases de palabra (por ejemplo, un nombre, un verbo, etcétera) como término origen. Los sinónimos se obtienen de un archivo separado para cada idioma. Siempre se devuelven en formato base y, menos unas cuantas excepciones, no son términos de múltiples palabras. Las

palabras de búsqueda siempre se reducen a su formato base cuando se buscan sinónimos. Estos son algunos ejemplos de sinónimos de una palabra en tres idiomas:

- Inglés

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

- Francés

mot:

expression parole terme vocable lettre billet missive épître
plaisanterie

- Alemán

Wort:

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

Ampliación de tesoro

Un término de búsqueda puede ampliarse utilizando términos del tesoro que pueden obtenerse a través de una relación específica. Estas relaciones pueden ser jerárquicas (por ejemplo la relación “Narrower term”), asociativas (por ejemplo, una relación de “Related term”) o puede ser una relación de sinónimos. Un término del tesoro puede ser, y a menudo es, un término de múltiples palabras.

El apartado “Conceptos del tesoro” en la página 268 describe la ampliación de tesoro con más detalles.

El término de búsqueda (término inicial) no se normaliza cuando se efectúa una búsqueda en el tesoro. Las palabras resultantes de la búsqueda en el tesoro se reducen a su formato base de acuerdo al tipo de índice.

Ampliación de sonido

La ampliación de sonido amplía palabras aisladas a un conjunto de palabras que suenan de forma similar. Es especialmente útil cuando no se sabe cómo se escribe exactamente el término que se ha de buscar.

Enmascaramiento de caracteres y palabras

El enmascaramiento es una técnica de ampliación no lingüística, en la que se sustituye una expresión regular por una disyunción de todas las palabras que la satisfagan. Ni una expresión enmascarada ni ninguna de sus ampliaciones son objeto de lematización, extracción de palabras de parada ni ninguna de las demás técnicas de ampliación. Esto puede tener el efecto que, por ejemplo,

Proceso lingüístico para la recuperación

un formato de verbo irregular como juegan, cuando se busca con el término enmascarado jueg*, coincida en un índice exacto, pero no en un índice lingüístico, donde este formato se ha lematizado y convertido en jugar.

Si utiliza el enmascaramiento de palabras, el rendimiento puede ser lento, especialmente al buscar en índices grandes.

Proceso lingüístico para examen

El proceso lingüístico también se utiliza al examinar documentos que se han encontrado después de una búsqueda. Se realiza en dos etapas:

1. Análisis básico del texto: normalización y ampliación de términos
2. Comparación ampliada.

Etapas 1: Normalización y ampliación de términos

La primera etapa se realiza sin utilizar ningún diccionario electrónico.

Normalización

La normalización se describe en el apartado “Análisis básico de texto” en la página 256.

Ampliación de términos

La ampliación de términos es lo contrario a la reducción de un término a su formato base. Si el índice es lingüístico, los términos de búsqueda se reducen a su formato base antes de que empiece la búsqueda.

De manera similar, si tiene un índice lingüístico, los términos de los documentos se reducen a su formato base antes de añadirlos al índice. Por lo tanto, los documentos se encuentran en base a un formato base del término.

Sin embargo, cuando examina un documento encontrado, espera ver resaltadas todas las variantes del formato base. Para resaltar estas variantes, se expande el término base encontrado.

Se pueden producir todas las variantes (inflexiones) para el término de búsqueda encontrado en los diccionarios. Estas son las inflexiones producidas para la palabra en alemán gehen (ir):

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

Etapas 2: Comparación ampliada

La segunda etapa es la comparación ampliada, que puede utilizarse en raras ocasiones cuando el análisis básico de texto y la normalización no pueden resaltar un término encontrado. La comparación ampliada encuentra las coincidencias más difíciles.

Puede elegir la comparación ampliada especificando DES_EXTENDED como parámetro de la función API DesOpenDocument.

La comparación ampliada utiliza el mismo proceso lingüístico que se realiza durante la indexación lingüística.

Estas son las ocasiones en las que la comparación ampliada puede encontrar coincidencias adicionales:

- El término de búsqueda incluye caracteres de máscara y es una inflexión. Los caracteres de máscara se procesan y se realiza una reducción a la raíz para el término de búsqueda y se encuentran los términos correspondientes. Sin la comparación ampliada, no se resaltaría el texto que coincide con los criterios de búsqueda especificados.
Ejemplo: Un documento contiene el término inflexado *juegan*.
 - Durante la indexación, este término se reduce a *jug*.
 - Si el término de búsqueda es *jug%*, se encuentra el documento anterior porque la reducción a la raíz es *jug*.
 - Sin la comparación ampliada, sólo se resaltan aquellas palabras que coinciden con el término *jug%*. Con la comparación ampliada, el término inflexado *juegan* también se resalta.
- Si se han indexado palabras compuestas. Cuando un documento en un idioma germánico contiene una palabra compuesta y se indexa utilizando un índice lingüístico, el índice del documento conserva las partes de la palabra compuesta y la palabra compuesta en sí. Cuando busca una parte de una palabra compuesta, se encuentran los documentos que contienen la palabra compuesta, pero sin la comparación ampliada la palabra no se resalta.
Ejemplo: Un documento contiene la palabra en alemán *Apfelbaum* (literalmente, árbol de manzanas).
 - Durante la indexación lingüística la palabra se reduce a *apfel* y *baum*.
 - Cuando se busca en el índice el término *baum*, se encuentra el término *Baum* y el documento que lo contiene en el índice.
 - Sin la comparación ampliada, no se resalta ningún término porque el documento contiene *Apfelbaum*, pero no *Baum*. Con la comparación ampliada, la palabra compuesta *Apfelbaum* se divide y el fragmento *Baum* se encuentra y se resalta.
- Si las palabras llevan guiones al final de una línea. Si el procesador de texto inserta automáticamente los guiones, se puede encontrar y resaltar la palabra con el guión. Sin embargo, si el guión lo escribe el usuario, se encuentran los documentos que contienen la palabra, pero sin la comparación ampliada la palabra no se resalta.

Proceso lingüístico para examen

Ejemplo: Un documento contiene la palabra con guión contenedor, separada al final de una línea de la siguiente manera:

Otro nombre
para una carpeta es un con-
tenedor.

- Durante la indexación la palabra se normaliza a contenedor.
- Cuando se busca el término contenedor en el índice, se encuentran la palabra y el documento que la contiene.
- Se intenta resaltar cualquier palabra del documento que coincida con contenedor. Sin la comparación ampliada, sólo se encuentra una coincidencia si el guión de con-tenedor lo ha insertado el procesador de texto y no lo ha escrito un usuario.

Conceptos del tesoro

Un tesoro es un vocabulario controlado de términos relacionados semánticamente que, normalmente, cubre el área de un tema específico. Puede visualizarse como una red semántica en la que cada término se representa por un nodo. Si dos términos están relacionados entre sí, sus nodos se conectan mediante un enlace etiquetado con el nombre de la relación. Todos los términos se relacionan directamente con un término determinado que puede obtenerse siguiendo todas las conexiones que salen de su nodo. Se pueden obtener más términos relacionados siguiendo iterativamente todas las conexiones que salen de los nodos obtenidos en el paso anterior. La Figura 18 en la página 269 muestra un ejemplo de la estructura de un tesoro muy pequeño.

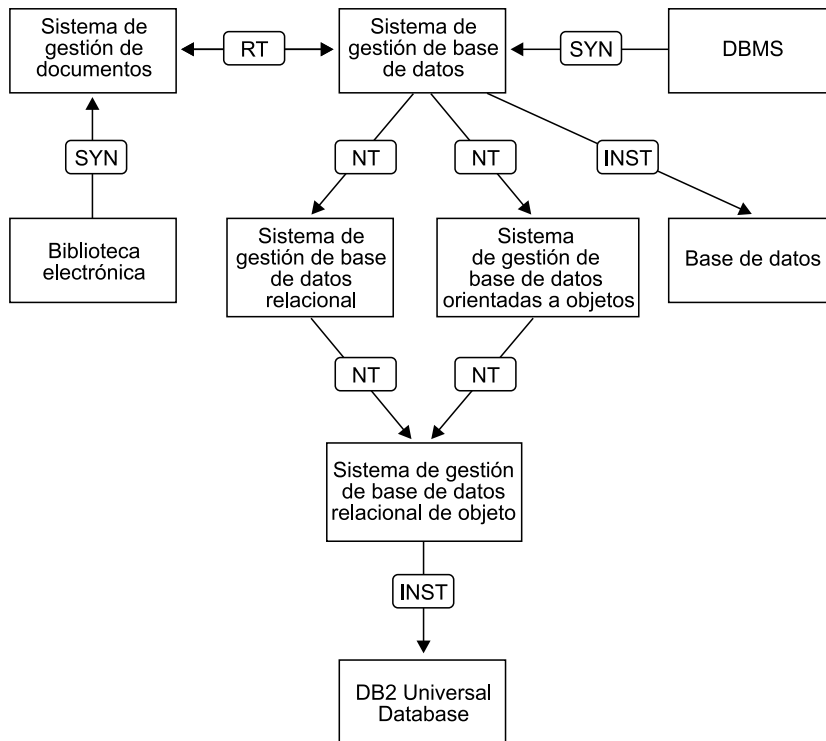


Figura 18. Tesoro visualizado como una red

DB2 Text Extender le permite ampliar un término de búsqueda añadiendo términos adicionales de un tesoro que ha creado previamente. Consulte el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215 para averiguar cómo utilizar la ampliación del tesoro en una consulta.

Para crear un tesoro para utilizarlo en una aplicación de búsqueda es necesario un archivo de definiciones de tesoro que se ha de compilar en un formato interno, el diccionario tesoro.

El formato del diccionario utilizado por un índice lingüístico y uno exacto difieren del utilizado por un índice Ngram. Así, se proporcionan dos compiladores de tesoro diferentes con el producto. No sólo son levemente diferentes en los conceptos en los que se basan, sino que necesitan formatos fuente diferentes. Por lo tanto, primero debe decidir el tipo de índice que utilizará antes de empezar una definición del tesoro para la aplicación de búsqueda.

Los componentes básicos de un tesoro son los “términos” y las “relaciones”.

Conceptos del tesoro

Términos

Un término es una palabra o una expresión que indica un concepto dentro del dominio de temas del tesoro. Por ejemplo, los siguientes términos podrían estar en uno o varios tesauros:

proceso de datos
helicóptero
producto nacional bruto

En un tesoro DB2 Text Extender, los términos se clasifican como descriptores o no descriptores. Un *descriptor* es un término de una clase de sinónimos que es el preferido para la indexación y la búsqueda. Los otros términos de la clase se llaman *no descriptores*. Por ejemplo, contorno y forma son sinónimos, donde forma puede ser el descriptor y contorno un no descriptor.

Un tesoro Ngram no distingue entre descriptores y no descriptores.

Relaciones

Una relación es una expresión de una asociación entre dos términos. Las relaciones tienen las propiedades siguientes:

- La *profundidad* de una relación es el número de niveles a los que se amplía la relación. Esto se especifica en la sintaxis de búsqueda utilizando la palabra clave THESDEPTH. Consulte el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215 para averiguar cómo utilizar la ampliación del tesoro en una consulta.
- La *direccionalidad* de una relación especifica si la relación es igual desde un término al otro (bidireccional) o sólo en una dirección (unidireccional).

La ampliación del tesoro puede utilizar todas las relaciones definidas en el tesoro. El usuario también puede especificar la profundidad de la ampliación. Este es el número máximo de transiciones desde un término fuente a un término de destino. Sin embargo, tenga en cuenta que el conjunto de términos puede incrementarse exponencialmente a medida que aumenta la profundidad.

El ejemplo siguiente muestra los términos que se van añadiendo a medida que aumenta la profundidad.

sanitario

servicio sanitario, paramédicos, medicina, enfermedad

alergología, virología, medicina veterinaria, toxicología, cirugía, estomatología, reumatología, radioterapia, siquiatria, medicina preventiva, patología, odontología, nutrición, medicina nuclear, neurología, nefrología, chequeo médico, medicina de empresa, hematología, medicina general, epidemiología, reclamaciones clínicas, cardiología, oncología

Relaciones del tesoro DB2 Text Extender

Estos son los tipos de relaciones proporcionadas por un tesoro de DB2 Text Extender

- Asociativa
- Sinónima
- Jerárquica
- Otras

En un tesoro DB2 Text Extender no hay relaciones predefinidas. Puede dar a cada relación un nombre como, por ejemplo, BROADER TERM, que puede ser una abreviatura mnemotécnica como BT. Las relaciones comunes utilizadas en el diseño del tesoro son:

- BT o BROADER TERM
- NT o NARROWER TERM
- RT o RELATED TERM
- SYN o SYNONYM
- USE
- UF o USE FOR

Asociativa: Una relación asociativa es una relación bidireccional entre descriptores, que se amplía a cualquier profundidad. Enlaza dos términos que no son equivalentes ni jerárquicos, aunque están semánticamente asociados en tal medida que el enlace entre ellos puede sugerir términos adicionales en la indexación o en la recuperación.

Las relaciones asociativas se designan normalmente como RT (término relacionado). Son ejemplos:

perro RT seguridad
mascota RT veterinario

Sinónimo: Cuando se distingue entre descriptores y no descriptores, como se hace en un tesoro DB2 Text Extender, la relación sinónima es unidireccional entre dos términos que tienen un significado igual o similar. En una clase de sinónimos, uno de los términos se designa como el descriptor. Los demás términos se llaman no descriptores. Consulte el apartado “Relaciones de tesoro Ngram” en la página 272 para ver una definición de la relación sinónima cuando no se distingue entre descriptores y no descriptores.

La designación común USE conduce desde un no descriptor dado a su descriptor. La designación común USE FOR conduce desde el descriptor a cada no descriptor. Por ejemplo:

felino USE gato
abogado UF defensor

Conceptos del tesoro

Jerárquica: Una relación jerárquica es una relación unilateral entre descriptores que indica que uno de los términos es más específico, o menos genérico, que el otro. Esta diferencia conduce a la representación de los términos como una jerarquía, donde un término representa una clase y los términos subordinados hacen referencia a sus partes miembro. Por ejemplo, el término “ratón” pertenece a la clase “roedor”.

BROADER TERM y NARROWER TERM son relaciones jerárquicas. Por ejemplo:

coche NT limusina
equino BT caballo

Otras: Una relación de tipo *otra* es la más general. Representa una asociación que no recae con facilidad en ninguna de las demás categorías. Una relación de tipo *otra* puede ser bidireccional o unidireccional, no hay restricción en la profundidad y pueden existir relaciones entre descriptores y no descriptores.

Esta relación se utiliza con frecuencia para nuevos términos de un tesoro hasta que se determina la relación adecuada con otros términos.

Claro que puede definir su propia relación sinónima bidireccional utilizando el tipo de relación *asociativa* para una relación sinónima entre descriptores o incluso con el tipo de relación *otra* para una relación sinónima entre términos arbitrarios.

Relaciones de tesoro Ngram

Un tesoro Ngram da soporte a los dos tipos siguientes:

- Asociativa
- Sinónima

Hay dos relaciones predefinidas, cada una de ellas basada en estos dos tipos. Puede definir sus propias relaciones en base al tipo *asociativo*. Para ver los detalles, consulte el apartado “Creación de un tesoro Ngram” en la página 278.

Asociativa

Una relación asociativa es una relación bidireccional entre dos términos que no expresan el mismo concepto pero que están relacionados el uno con el otro. La relación predefinida *RELATED_TO* y todas las relaciones definidas por el usuario se basan en este tipo de relación.

Son ejemplos:

tenis RELATED_TO raqueta
alemán RELATED_TO salchicha

Sinónimo

Una relación sinónima es una relación bidireccional entre dos términos que tienen un significado igual o similar y que pueden utilizarse como alternativos entre sí. Esta relación puede, por ejemplo, utilizarse para un término y su abreviatura. La relación predefinida SYNONYM_OF es la única relación basada en este tipo.

Son ejemplos:

mancha SYNONYM_OF tacha
EE.UU. SYNONYM_OF Estados Unidos

Creación de un tesoro

Consulte también el apartado “Creación de un tesoro Ngram” en la página 278.

Hay un archivo de entrada de ejemplo de compilador de tesoro en inglés `desthes.sgm` almacenado en el directorio de ejemplos de la vía de acceso de la instalación. El directorio de diccionarios de los sistemas OS/2 y Windows es:

```
drive:\dmb\db2tx\samples
```

En sistemas AIX, HP-UX y SUN-Solaris, el directorio es:

```
DB2TX_INSTOWNER /db2tx/samples
```

Una versión compilada de este tesoro y su archivo de entrada SGML se almacena en el directorio de diccionarios.

```
unidad:\dmb\db2tx\dict  
o  
DB2TX_INSTOWNER /db2tx/dicts
```

Los archivos que pertenecen a este tesoro se denominan `desthes.th1`, `desthes.th2`, ... y `desthes.th6..`

Para crear un tesoro, defina primero su contenido en un archivo. Se recomienda utilizar un directorio plano para cada tesoro que defina. El archivo puede tener cualquier extensión excepto `th1` a `th6`, que se utilizan para el diccionario tesoro. Si utiliza el mismo directorio para el tesoro Ngram, consulte el apartado “Creación de un tesoro Ngram” en la página 278 para ver más extensiones excluidas.

Después compile el archivo ejecutando:

```
txthesc -f nombarch -c ccsid
```

donde `nombarch` sólo puede contener los caracteres a-z, A-Z y 0-9.

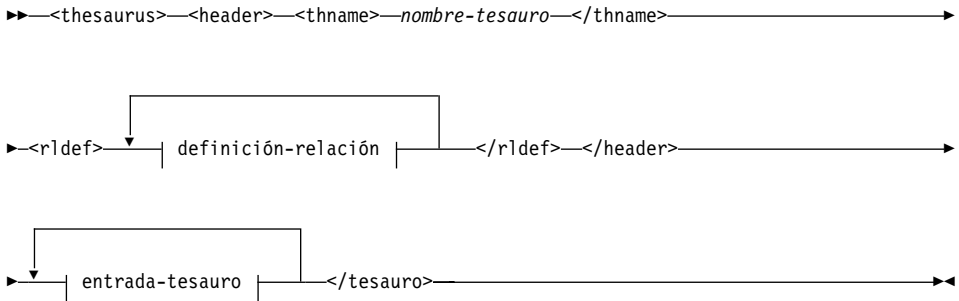
Actualmente, sólo se da soporte al CCSID 850.

Conceptos del tesoro

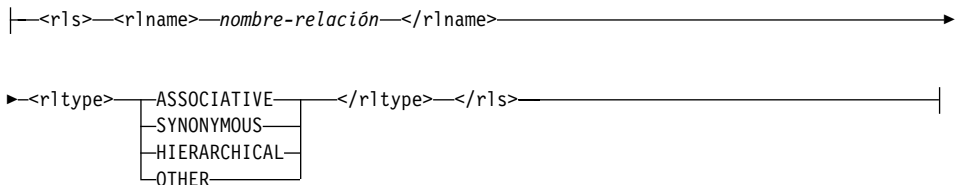
txthesc produce archivos de tesoro que se denominan *nombarch* sin ninguna extensión y con las extensiones th1 a th6, en el mismo directorio en el que está ubicado el archivo de definiciones. Si ya hay un tesoro con el mismo nombre, se sobregaba sin avisar.

Consulte el “Capítulo 12. Sintaxis de los argumentos de búsqueda” en la página 215 para averiguar cómo se utiliza un tesoro en una consulta.

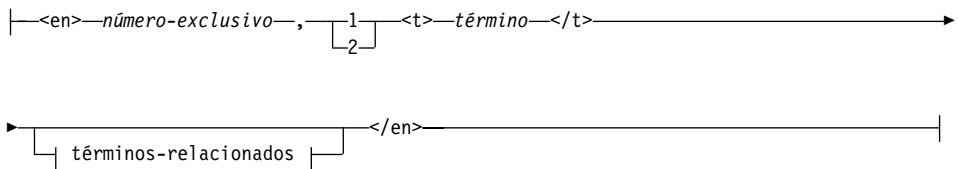
Especifique el contenido de un tesoro utilizando Standard Generalized Markup Language (SGML). El diagrama siguiente muestra las reglas de sintaxis que se han de seguir al crear un tesoro.



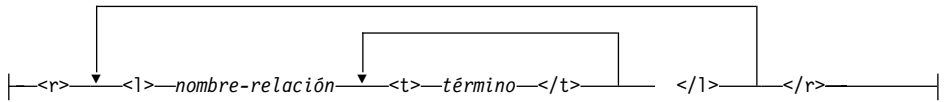
definición-relación:



entrada-tesauro:



términos-relacionados:



nombre-relación sólo puede contener los caracteres a-z, A-Z y 0-9.

La Figura 19 en la página 276 muestra la definición SGML del tesoro mostrado en la Figura 18 en la página 269.

Conceptos del tesoro

```
<thesaurus>
<header>
<thname>tesauro de ejemplo tesac</thname>
<rldef>

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

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

<rls>
<rlname>Instancia</rlname>
<rltype>hierarchical</rltype>
</rls>

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

<en> 2, 1
<t> sistema de gestión de bases de datos</t>
<r>
  <l>Narrower Term
  <t>oo sistema de gestión de bases de datos</t>
  <t> sistema de gestión de bases de datos relacionales</t>
  </l>

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

  <l>Related Term
  <t> sistema de gestión de documentos</t>
  </l>

  <l>Instancia
  <t> base de datos</t>
  </l>
</r>
</en>
```

Figura 19. Definición de un tesoro sencillo (Pieza 1 de 2)

```
<en> 5, 1
<t> sistema de gestión de bases de datos relacionales</t>
<r>
  <l>Narrower Term
  <t>sistema de gestión de bases de datos relacionales de objetos</t>
  </l>
</r>
</en>

<en> 3, 1
<t>sistema de gestión de bases de datos relacionales de objetos</t>
<r>
  <l>Instancia
  <t>DB2 Universal Database</t>
  </l>
</r>
</en>

<en> 6, 1
<t>sistema de gestión de bases de datos orientada a objetos</t>
<r>
  <l>Narrower Term
  <t>sistema de gestión de bases de datos relacionales de objetos</t>
  </l>
</r>
</en>

<en> 4, 1
<t>sistema de gestión de documentos</t>
<r>
  <l>Synonym
  <t>biblioteca</t>
  </l>
</r>
</en>

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

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

<en> 11, 1
<t>base de datos</t>
</en>
</tesaurus>
```

Figura 19. Definición de un tesaurus sencillo (Pieza 2 de 2)

Conceptos del tesoro

Creación de un tesoro Ngram

Hay un archivo de entrada de ejemplo de compilador de tesoro Ngram en inglés `desnthes.def` almacenado en el directorio de diccionarios de la vía de acceso de instalación. El directorio de diccionarios de los sistemas OS/2 y Windows es:

```
unidad:\dmb\db2tx\dict
```

En sistemas AIX, HP-UX y SUN-Solaris, el directorio de diccionarios es:

```
DB2TX_INSTOWNER /db2tx/dicts
```

También se almacena allí una versión compilada del tesoro de ejemplo. Los archivos que pertenecen a este tesoro se denominan `desnthes.<extension>` con la siguiente extensión, donde `n` es un dígito:

- Para archivos de diccionario: `wdf`, `wdv`, `grf`, `grv`, `MEY`, `ROS`, `NEY`, `SOS`, `lkn`
- Para archivos temporales: `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` y `$21`

Para crear un tesoro Ngram, primero defina su contenido en un archivo de definiciones. Puede tener varios tesauros en el mismo directorio, pero se recomienda tener un directorio para cada tesoro. La longitud del nombre de archivo sin la extensión no debe exceder de 8 caracteres. La extensión es opcional pero se restringe a 3 caracteres y deben ser diferentes de cualquiera de las extensiones listadas más arriba.

Si utiliza el mismo directorio para otros tesauros DB2 Text Extender, no utilice las extensiones listadas en el apartado “Creación de un tesoro” en la página 273 .

Después compile el archivo ejecutando:

```
txthesn -f nombre-archivo-definiciones -ccsid página-códigos
```

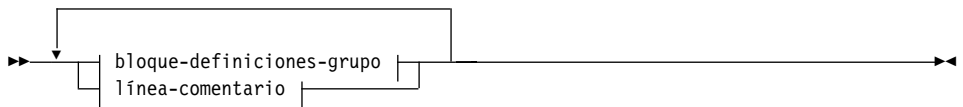
Esta es una lista de las páginas de códigos que soporta un tesoro Ngram:

932	AIX, OS/2	Japonés
942	OS/2	Japonés
943	OS/2, Windows	Japonés
949	OS/2	Coreano
950	AIX, HP-UX, OS/2, SUN-Solaris, Windows	Chino tradicional
970	AIX, HP-UX, SUN-Solaris	Coreano
1381	OS/2, Windows	Chino simplificado
1363	Windows	Coreano
1383	AIX, HP-UX, SUN-Solaris	Chino simplificado

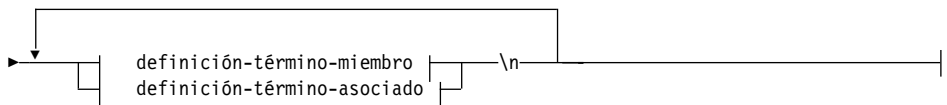
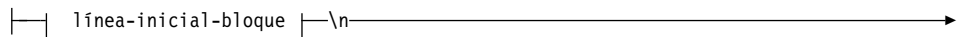
850	AIX, OS/2	Latin-1
1252	Windows	Latin-1

txthesn produce archivos de tesoro que tienen el mismo nombre que el *nombre-archivo-definiciones* con las extensiones mencionadas más arriba. Los archivos se crean en el mismo directorio que el archivo de definiciones. Si ya existe un tesoro con el mismo nombre en este directorio se sobregraba sin avisar.

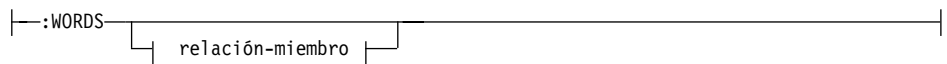
Especifique el contenido del tesoro utilizando el diagrama de sintaxis siguiente:



bloque-definiciones-grupo:



línea-inicial-bloque:



relación-miembro:

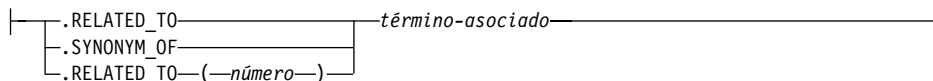


definición-término-miembro:



Conceptos del tesoro

definición-término-asociado:



línea-comentario:



Cada término miembro debe estar escrito en una sola línea. Cada término asociado debe ir precedido por el nombre de relación. Si los términos miembro están relacionados entre sí, especifique una relación de miembros.

La longitud de los términos miembros y los términos asociados se restringe a 164 caracteres. Los caracteres de un solo byte y los caracteres de doble byte de la misma letra se contemplan como iguales. No se distingue entre letras en mayúsculas y en minúsculas. Un término puede contener un carácter en blanco pero se puede utilizar el carácter de punto de un solo byte "." o los dos puntos ":",

Todas las relaciones definidas por el usuario se basan en el tipo *asociativo*. Se identifican por números exclusivos entre 1 y 128.

Si una aplicación desea utilizar nombres simbólicos para sus relaciones de tesoro en lugar del nombre y número de la relación, debe administrar la correlación ella misma. Por ejemplo, si la relación `OPPOSITE_OF` se ha definido como `RELATED_TO(1)`, la aplicación tiene que correlacionar este nombre con el nombre de relación interna `RELATED_TO(1)`. Consulte el "Capítulo 12. Sintaxis de los argumentos de búsqueda" en la página 215 para averiguar cómo utilizar la ampliación del tesoro en una consulta.

Capítulo 16. Archivos de configuración

Este capítulo describe los archivos de configuración. Estos archivos se generan automáticamente al crear una instancia de DB2 Text Extender con el mandato `txicrt` o cuando se crea el perfil de cliente con el mandato `descrc1`. Estos archivos se generan en la página de códigos 819.

Puede editar estos archivos para ajustar el sistema, sin embargo, asegúrese de que utiliza la página de códigos correcta al hacerlo. Los nombres de secciones y los nombres de opciones son independientes de las mayúsculas y minúsculas. Se utiliza un punto y coma como delimitador de comentarios.

Cuando la opción es un valor booleano, los valores TRUE, YES, ON y 1 se consideran verdaderos sin tener en cuenta si están escritos en mayúsculas o minúsculas; todos los demás valores se consideran falsos.

Archivo de configuración de cliente

Nombre de archivo

DESCL.INI

Ubicación

UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx

Windows NT y Windows 2000:

%DMBMPATH%\instancia%\DB2INSTANCE%\db2tx

Las opciones actualizadas se activan en la siguiente función `StartSession`.

Sección	Opción	Valor por omisión	Descripción
[INSTANCE]	DESWORKCL		Apunta a un directorio de trabajo que se utiliza para archivos temporales.
	DESNLPCL		Apunta al directorio de recursos.
	DESDTDPATHSRV		Apunta al directorio en el que se almacenan las DTD para los documentos XML.
	DESDTDPATHCL		Apunta a la vía de acceso donde deben almacenarse los archivos DTD. Se utiliza para archivos XML.

Archivo de configuración de cliente

Sección	Opción	Valor por omisión	Descripción
[BUFFER]	BUFFERSEGMENTSIZ	32 000	El tamaño de los segmentos de bloque, en bytes, utilizados para el almacenamiento intermedio.
	BUFFERSEGMENTCOUNT	3	El número máximo de segmentos utilizados antes de intercambiar los almacenamientos intermedios a archivos temporales. Un segmento de almacenamiento intermedio se define por BUFFERSEGMENTSIZ.
[DOCUMENTFORMAT]	USEREXIT		<p>El nombre de la salida de usuario utilizada para trabajar con formatos de documentos no soportados. Especifique un nombre de archivo si la salida de usuario se almacena en un directorio que forma parte de la sentencia PATH o un nombre de archivo totalmente calificado.</p> <p>Para obtener más información sobre la salida de usuario para la conversión de formato, consulte el apartado “Utilización de formatos de documentos no soportados” en la página 33.</p>
	FORMATRECOGNITION	TRUE	<p>Activa el reconocimiento de los formatos de documentos.</p> <p>TRUE: el reconocimiento de formato está activado</p> <p>FALSE: el reconocimiento de formato está desactivado</p>

Archivo de configuración de cliente

Sección	Opción	Valor por omisión	Descripción
	UseExitForAllFormats	FALSE	<p>Determina cuándo se llama a la salida de usuario para trabajar con formatos de documentos no listada en deslsdef.h. Debe establecer un valor para USEREXIT.</p> <p>TRUE: siempre se llama a la salida de usuario. Si se establece este valor, FORMATRECOGNITION se pasa por alto.</p> <p>FALSE: llama a la salida de usuario para todos los formatos por encima del valor de EHW_USER_FORMATS.</p>

Archivo de configuración del servidor

Nombre de archivo

DESSRV.INI

Ubicación

UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx/TXINSnnn

Windows NT y Windows 2000:

%DMBMMPPATH%\instancia%\DB2INSTANCE%\db2tx\txinsnnn

Este archivo sólo existe una vez por instancia del servidor.

Las opciones actualizadas se activan la próxima vez que se inicia la instancia del servidor.

Tabla 20. Opciones de archivo de configuración del servidor

Sección	Opción	Valor por omisión	Descripción
[INSTANCE]	DESWORKSRV		Apunta a un directorio de trabajo que se utiliza para archivos temporales.
	DESNLPSSRV		Apunta al directorio de recursos.

Archivo de configuración del servidor

Tabla 20. Opciones de archivo de configuración del servidor (continuación)

Sección	Opción	Valor por omisión	Descripción
[DAEMON]	MaxMtEntries	30 en AIX, Windows, OS/390 10 en Solaris	El número máximo de índices utilizados en paralelo en cualquier momento dado. Disminuya este número si no tiene muchos recursos como, por ejemplo, semáforos y memorias compartidas. Los recursos disponibles dependen de la plataforma y, por lo tanto, los valores por omisión también dependen de la plataforma.
	MaxIndexEntries	1000	El número máximo de índices utilizados. Disminuya este número si tiene pocos recursos como, por ejemplo, memorias compartidas.
[BUFFER]	BUFFERSEGMENTSIZ	32 000	El tamaño de los segmentos de bloque, en bytes, utilizados para el almacenamiento intermedio. Se utiliza por EhwUpdate.
	BUFFERSEGMENTCOUNT	3	El número máximo de segmentos utilizados durante el proceso de actualización del índice antes de que los almacenamientos intermedios se intercambien a archivos temporales. Aumente este número si las colecciones de documentos contienen documentos grandes.
	BUFFERSORTSIZE	20 000 000	El tamaño del almacenamiento intermedio, en bytes, utilizado para clasificar los archivos de trabajo temporales.

Tabla 20. Opciones de archivo de configuración del servidor (continuación)

Sección	Opción	Valor por omisión	Descripción
[DOCUMENTFORMAT]	USEREXIT		<p>El nombre de la salida de usuario utilizada para trabajar con los formatos de documentos listados en deslsdef.h. Especifique un nombre de archivo si la salida de usuario está almacenada en un directorio que forma parte de la sentencia PATH o un nombre de archivo totalmente calificado.</p> <p>Para obtener más información sobre la salida de usuario para la conversión de formato, consulte el apartado “Utilización de formatos de documentos no soportados” en la página 33.</p>
	FORMATRECOGNITION	TRUE	<p>Activa el reconocimiento de los formatos de documento listados en deslsdef.h.</p> <p>TRUE: el reconocimiento de formato está activado</p> <p>FALSE: el reconocimiento de formato está desactivado</p>

Archivo de configuración del servidor

Tabla 20. Opciones de archivo de configuración del servidor (continuación)

Sección	Opción	Valor por omisión	Descripción
	UseExitForAllFormats	FALSE	<p>Determina cuándo se llama a la salida de usuario para trabajar con formatos de documentos no listada en deslsdef.h. Debe establecer un valor para USEREXIT.</p> <p>TRUE: siempre se llama a la salida de usuario. Si se establece este valor, FORMATRECOGNITION se pasa por alto.</p> <p>FALSE: llama a la salida de usuario para todos los formatos por encima del valor de EHW_USER_FORMATS.</p>
[LINGPREC] Para todos los índices con un tipo base lingüístico o exacto.	UPDATETHRESHOLD	4 000 000	Un proceso de actualización de índice se divide internamente en varias ejecuciones de actualización y de reorganización. Este valor especifica el número de palabras que se han de reunir en un paso de actualización.
	UPDATESLICE	1	El número de ejecuciones de actualización que tienen lugar antes de iniciarse un proceso de reorganización interna. Una ejecución de actualización la define UPDATETHRESHOLD.
[NGRAM] Para todos los índices con soporte de DBCS.	UPDATETHRESHOLD	10 000 000	El tamaño total, en bytes, de los documentos añadidos a un índice durante una ejecución de actualización. Si se ha excedido el umbral, se inicia automáticamente un proceso de reorganización.

Tabla 20. Opciones de archivo de configuración del servidor (continuación)

Sección	Opción	Valor por omisión	Descripción
	UPDATESLICE	10 000	El número máximo de documentos de un índice secundario. Este número se comprueba después de cada ejecución de actualización. Si el número de documentos es mayor que este valor, se inicia automáticamente un proceso de reorganización.

Archivo de configuración de análisis de prueba

Nombre de archivo

DESTAF.TCF

Ubicación

UNIX: DB2TX_INSTOWNERHOMEDIR/db2tx/TXINSnnn

Windows NT y Windows 2000:

%DMBMMPPATH%*instancia*\%DB2INSTANCE%\db2tx\txinsnnn

Este archivo sólo existe una vez por instancia del servidor.

El archivo de configuración de análisis de prueba contiene secciones para el proceso de actualización y de consulta de cada tipo de índice, excepto los índices ngram. Sólo debe cambiar las opciones descritas aquí. Si cambia cualquiera de las demás entradas, puede conducir a errores durante el proceso de actualización o de consulta.

Opciones comunes

Las opciones siguientes se aplican a:

- [TAF_Application:UPDATE:PREC]) para el tipo de índice PRECISE
- [TAF_Application:UPDATE:LING]) para el tipo de índice LINGUISTIC
- [TAF_Application:UPDATE:NORM]) para el tipo de índice PRECISE NORMALIZED
- [TAF_Application:QUERY:PREC])
- [TAF_Application:QUERY:LING])
- [TAF_Application:QUERY:NORM])

Archivo de configuración de análisis de prueba

Opción	Valor por omisión	Descripción
FilterSet	Ninguno, obligatorio	Nombre de la sección del archivo de configuración que contiene las normas para las palabras de parada, por ejemplo, StwUseAllUpCs, StwIgnAllUpCs.
Normalize	No	<p>Especifica si se lleva a cabo la normalización en el índice.</p> <p>Si se establece en sí, la normalización convierte todos los caracteres a minúsculas y sustituye los caracteres diacríticos por caracteres base.</p> <p>Si se establece en No, se pasan por alto las entradas para NormalizeUmlauts y ConvertAllUpperToLower.</p>
NormalizeUmlauts	Sí	Especifica si se han de sustituir las diéresis, por ejemplo, ä por ae. Esto sólo se aplica a los documentos en alemán, danés, sueco o noruego.
ConvertAllUpperToLower	No	Especifica que las palabras con todas las letras en mayúsculas se convierten a minúsculas.

[TAF_Filter:STWUSEALLUPCS] [TAF_Filter:STWIGNALLUPCS]

Las opciones de estas secciones describen los conjuntos de filtros de palabras de parada. La sección [TAF_Filter:STWUSEALLUPCS] trata las palabras con todas las letras escritas en mayúsculas como palabras en mayúsculas y minúsculas [TAF_Filter:STWIGNALLUPCS] define un trato especial para las palabras con todas las letras escritas en mayúsculas.

Opción	Valor por omisión	Descripción
FilterType	TAF_FILTER_STW	La sección del archivo de configuración que contiene las normas para el filtro de palabras de parada.
MinTokenLength	3	El número mínimo de caracteres necesario para una señal. Si la longitud de la señal es menor que este valor, no se indexará. Es un valor entero ≥ 1 .
MinUpperCaseTokenLength	2	El número mínimo de caracteres necesario para una señal con todas las letras en mayúsculas. Si la longitud de la señal es menor que este valor, no se indexará. Es un valor entero ≥ 1 .

Archivo de configuración de análisis de prueba

Opción	Valor por omisión	Descripción
UseAlternateTerritories	No	<p>Especifica si se ha de utilizar un diccionario alternativo para el proceso de palabras de parada si no se puede encontrar el diccionario específico del idioma.</p> <p>Si se establece en sí, se utiliza el primer diccionario de palabras de parada que coincide con el idioma. Por ejemplo, si el idioma actual es en-gb pero no hay ningún archivo de palabras de parada en-gb.tsw, se utiliza el archivo de palabras de parada en-us.tsw.</p> <p>Si se establece en no, no se efectúa el filtro de palabras de parada.</p>

[TAF_AnnotationType:TAF_ThesaurusEntry]

Esta sección establece los valores por omisión para las consultas que necesitan la ampliación del tesoro.

Palabra clave	Valor por omisión	Descripción
ThesBaseName	desthes	Establece el nombre por omisión del tesoro utilizado para las consultas que necesitan la ampliación de la consulta. El nombre es una serie de caracteres.
Depth	1	Establece la vía de acceso por omisión para atravesar el tesoro utilizado para consultas que necesitan la ampliación del tesoro. Es un valor entero ≥ 1 .

[TAF_Plugin:TAF_POE]

Para cambiar el valor por omisión para el entorno de idioma que está precargado cuando se inicia Aplicación Text Search, puede definir un idioma en la opción PreloadLanguages.

Opción	Valor por omisión	Descripción
PreloadLanguages	EN-US	Especifica un conjunto de recursos específicos de idioma que se han de precargar en el arranque. Por ejemplo, si una aplicación se utiliza en documentos en italiano, se cargan los recursos para el italiano. El valor del campo toma el formato XX-YY donde XX es el idioma e YY es el territorio.

Archivo de configuración de análisis de prueba

Capítulo 17. Códigos de retorno

Este capítulo lista los códigos que la API de DB2 Text Extender devuelve en respuesta a una llamada de función. Se listan por orden alfabético.

Todas las llamadas a API de DB2 Text Extender devuelven un código de retorno numérico como valor de función C. Los códigos de retorno se definen en el archivo include DES_EXT.H proporcionado con DB2 Text Extender.

La API de DB2 Text Extender intercepta situaciones de error e informa de las condiciones de error utilizando un código de retorno.

Las aplicaciones que llaman a las funciones API de DB2 Text Extender deben comprobar siempre el código de retorno antes de intentar procesar cualquier otro parámetro de salida. Los códigos de retorno posibles para cada llamada se listan, junto con sus parámetros, en el “Capítulo 13. Funciones API para la búsqueda y examen” en la página 229.

En algunos casos, una entrada incorrecta como, por ejemplo, un puntero de sesión obsoleta puede provocar una condición de fin anómalo en los servicios API que DB2 Text Extender no puede interceptar.

RC_ALLOCATION_ERROR

Explicación: No se puede asignar almacenamiento para uso interno.

Lo que se ha de hacer: Asegúrese de que hay suficiente memoria disponible.

RC_FILE_IO_PROBLEM

Explicación: DB2 Text Extender no ha podido leer o grabar un archivo.

Lo que se ha de hacer: Compruebe que hay suficiente espacio de disco y memoria disponible en el servidor. Compruebe que las variables de entorno y los valores de configuración del texto son correctos.

RC_INVALID_BROWSE_INFO

Explicación: La información de examen devuelta por DesGetSearchResultTable o por DesGetBrowseInfo y utilizada como entrada para DesStartBrowseSession no es válida.

Lo que se ha de hacer: Compruebe si un error de programación prevalece sobre la información de rastreo.

RC_INVALID_BROWSE_OPTION

Explicación: La opción de examen de DesGetSearchResultTable no es válida.

Lo que se ha de hacer: Asegúrese de que la opción es BROWSE o NO_BROWSE.

Códigos de retorno

RC_INVALID_MATCH_OPTION

Explicación: Las opciones de coincidencia utilizadas en DesOpenDocument no son válidas.

Lo que se ha de hacer: Compruebe que la opción es FAST o EXTENDED.

RC_INVALID_PARAMETER

Explicación: Uno de los parámetros de entrada es incorrecto.

Lo que se ha de hacer: Lea el mensaje de error devuelto por DB2 Text Extender para determinar la causa.

RC_INVALID_SEARCH_OPTION

Explicación: La opción de búsqueda de DesGetSearchResultTable no es válida.

Lo que se ha de hacer: Asegúrese de que la opción es DES_TEXTHANDLEONLY, DES_RANK, DES_MATCH o DES_RANKANDMATCH.

RC_INVALID_SESSION

Explicación: El puntero de sesión especificado en la llamada al servicio actual no es correcta o es obsoleta.

Lo que se ha de hacer: Guarde la información que sirva de ayuda para encontrar la causa del error, después finalice la aplicación.

RC_NO_BROWSE_INFO

Explicación: DB2 Text Extender no ha devuelto ninguna información de examen. Esto es debido a que el argumento de búsqueda ha generado un resultado de búsqueda vacío. No se trata de ningún error.

Lo que se ha de hacer: No es necesaria ninguna acción.

RC_PARSER_INVALID_ESCAPE_CHARACTER

Explicación: Los criterios de búsqueda contienen un carácter de escape incorrecto. Se informa de este error si se utiliza un blanco como carácter de escape o si se especifica más de un carácter de escape para una palabra o una frase en los criterios de búsqueda. Ejemplo: ESCAPE " " o ESCAPE "\$".

Lo que se ha de hacer: Compruebe la sintaxis del argumento de búsqueda y vuelva a intentarlo.

RC_PARSER_INVALID_USE_OF_ESCAPE_CHAR

Explicación: La sintaxis del carácter de escape de los criterios de búsqueda no puede interpretarse.

Lo que se ha de hacer: Compruebe la sintaxis del carácter de escape. Por ejemplo, si \$ es el carácter de escape especificado, la palabra o la frase sólo puede contener \$\$, \$_ o \$%, donde _ y % son los símbolos de máscara.

RC_PARSER_SYNTAX_ERROR

Explicación: La sintaxis de los criterios de búsqueda no se puede interpretar.

Lo que se ha de hacer: Compruebe la sintaxis del argumento de búsqueda, consultando el "Capítulo 12. Sintaxis de los argumentos de búsqueda" en la página 215.

RC_RESULT_TABLE_NOT_EXIST

Explicación: Está intentando almacenar el resultado de una búsqueda en una tabla que no existe.

Lo que se ha de hacer: Cree una tabla de resultados tal como se muestra en la Figura 17 en la página 245.

RC_SE_BROWSER_TIME_OUT

Explicación: Se ha iniciado el proceso de búsqueda pero no ha respondido en un tiempo aceptable. DB2 Text Extender ha cancelado el proceso pendiente.

Este error puede producirse cuando el sistema no tiene suficiente espacio de almacenamiento o tiene una carga excesiva.

Lo que se ha de hacer: Termine la sesión de examen llamando a `DesEndBrowseSession`, libere el almacenamiento asignado llamando a `DesFreeBrowseInfo` y vuelva a intentarlo.

RC_SE_CAPACITY_LIMIT_EXCEEDED

Explicación: La función pedida no puede procesarse. No hay suficiente memoria o espacio de disco.

Lo que se ha de hacer: Finalice el programa y compruebe los recursos del sistema.

RC_SE_COMMUNICATION_PROBLEM

Explicación: Ha fallado la comunicación con el servidor DB2 Text Extender. El error puede haberse producido por una falta de espacio de almacenamiento o por una instalación incorrecta de DB2 Text Extender.

Lo que se ha de hacer: Guarde la información que sirva de ayuda para encontrar el error, después finalice la aplicación.

RC_SE_CONFLICT_WITH_INDEX_TYPE

Explicación: La especificación del término de búsqueda lingüístico de la consulta con corresponde al tipo de índice. Por ejemplo, `PRECISE FORM OF` no se puede utilizar con un índice lingüístico. La especificación de lingüístico por omisión se utiliza tal como se muestra en la Tabla 8 en la página 223.

Lo que se ha de hacer: Adapte la aplicación

para impedir la especificación de opciones de consulta que entren en conflicto con el tipo de índice.

RC_SE_DICTIONARY_NOT_FOUND

Explicación: Los servicios lingüísticos de DB2 Text Extender no pueden encontrar los archivos de diccionario. La consulta se procesa sin el soporte lingüístico. Los archivos de diccionarios correspondientes al código o códigos del idioma especificado no están en la vía de acceso esperada.

Lo que se ha de hacer: Puede continuar llamando a API. Para UNIX, compruebe que el diccionario necesario está en la vía de acceso `{DB2TX_INSTOWNERHOMEDIR}/db2tx/dicts`. Para OS/2, compruebe que el diccionario necesario está en la vía de acceso especificada en el valor de configuración del texto `DB2TX_DATA`. Si es necesario, instale el diccionario necesario.

RC_SE_DOCMOD_READ_PROBLEM

Explicación: Cuando se crea la instancia de DB2 Text Extender, se pone un archivo de modelos de documentos denominado `desmodel.ini` en el directorio de instancia. Cuando se crea un índice, también se pone un archivo `desmodel.ini` en el directorio de índices `IXnnnnnn`. Este archivo de modelos de documentos no se ha podido leer.

Lo que se ha de hacer: Compruebe que existe el archivo de modelos de documentos y que está en el directorio correcto.

RC_SE_DOCUMENT_NOT_ACCESSIBLE

Explicación: Se ha encontrado el documento de texto pedido, pero no se puede acceder en este momento.

Lo que se ha de hacer: Compruebe si otra tarea o usuario está accediendo de forma exclusiva al documento.

Códigos de retorno

RC_SE_DOCUMENT_NOT_FOUND

Explicación: El documento de texto pedido no se ha encontrado. La causa más probable es que se ha suprimido un documento de texto del almacenamiento, pero todavía no se ha eliminado del índice de DB2 Text Extender. Esto también puede ocurrir si intenta examinar un documento identificado por un descriptor de contexto dañado.

Lo que se ha de hacer: En la mayoría de casos, puede pasar por alto este código de retorno. No se volverá a visualizar después de la próxima actualización del índice.

Si persiste, compruebe que el programa de aplicación está pasando el descriptor de contexto encontrado correctamente para el examen.

RC_SE_EMPTY_INDEX

Explicación: El índice de DB2 Text Extender correspondiente a la columna de descriptores de contexto direccionada por la petición de búsqueda está vacío. No se han añadido documentos de texto a este índice o se han eliminado del mismo todos los documentos de texto.

Esto puede producirse cuando se ha habilitado una columna de texto, pero los documentos de la columna no se han indexado todavía. Es decir, ha especificado en el mandato ENABLE TEXT COLUMN que el índice se ha de crear más tarde, en el momento determinado por los valores de la indexación periódica.

También puede producirse cuando se ha habilitado una tabla de texto para crear un índice común vacío para todas las columnas de texto, pero no se ha habilitado ninguna de las columnas de texto.

Lo que se ha de hacer: Si se ha utilizado ENABLE TEXT TABLE para crear un índice común vacío para todas las columnas de texto, ejecute ENABLE TEXT COLUMN para una de las columnas de texto como mínimo que contenga el texto que se ha de buscar. En este

mandato, puede determinar si el índice se crea inmediatamente o en el momento determinado por los valores de indexación periódica.

Ejecute GET INDEX STATUS para comprobar que el índice se ha creado satisfactoriamente.

RC_SE_EMPTY_QUERY

Explicación: DB2 Text Extender ha analizado y procesado lingüísticamente los criterios de búsqueda especificados. Un error de programación ha provocado que se efectúe una consulta que no contenía ningún término de búsqueda o todos los términos de búsqueda eran palabras de parada (palabras que DB2 Text Extender no indexa) que se han eliminado de la consulta. El resultado ha sido ningún término de búsqueda.

Lo que se ha de hacer: Vuelva a escribir la consulta. Si el problema persiste, compruebe si existe un error de programación.

RC_SE_END_OF_INFORMATION

Explicación: No se trata de ningún error. Se ha alcanzado el final del documento. No hay más información disponible para DesGetMatches.

Lo que se ha de hacer: Utilice este código de retorno para finalizar el proceso iterativo del documento con DesGetMatches.

RC_SE_FUNCTION_DISABLED

Explicación: La función pedida ha llamado a una función de DB2 Text Extender que el administrador ha impedido.

Lo que se ha de hacer: Pida ayuda al administrador. Puede que sea necesario detener y reiniciar DB2 Text Extender (txstop/txstart).

RC_SE_FUNCTION_IN_ERROR

Explicación: La función pedida se ha bloqueado debido a una situación de error en el servidor

DB2 Text Extender. La llamada a la API no puede procesarse.

Lo que se ha de hacer: Compruebe el estado del índice. Compruebe el espacio disponible en el directorio de índices. Restablezca el estado del índice y reintente el mandato.

RC_SE_INCORRECT_HANDLE

Explicación: Un descriptor de contexto especificado en un parámetro de entrada como, por ejemplo, *descriptor de contexto de sesión de examen* no es válido. Debe ser un descriptor de contexto que una llamada anterior haya devuelto y que no esté obsoleto.

Lo que se ha de hacer: Guarde la información que pueda servir de ayuda para encontrar la causa del error y, a continuación, termine la sesión llamando a `DesEndBrowseSession`.

Compruebe si un error de programación ha producido un descriptor de contexto incorrecto.

RC_SE_INDEX_DELETED

Explicación: El índice de DB2 Text Extender al que se está accediendo se ha suprimido.

Lo que se ha de hacer: Póngase en contacto con el administrador de DB2 Text Extender para volver a crear el índice.

RC_SE_INDEX_NOT_ACCESSIBLE

Explicación: No se puede acceder al índice de DB2 Text Extender y la llamada actual no puede procesarse.

Lo que se ha de hacer: Pida al administrador de DB2 Text Extender que compruebe la posibilidad de acceso del índice.

RC_SE_INDEX_SUSPENDED

Explicación: DB2 Text Extender ha recibido una petición relativa a un índice que otra sesión o la

sesión actual ha suspendido.

Lo que se ha de hacer: Pida al administrador de DB2 Text Extender que compruebe el estado del índice.

RC_SE_INSTALLATION_PROBLEM

Explicación: DB2 Text Extender ha encontrado un problema de instalación.

Lo que se ha de hacer: Compruebe el valor actual de las variables de entorno `DB2INSTANCE`, `DB2TX_INSTOWNER`, `DB2TXINSTOWNERHOMEDIR`. Utilice `descfgc1 -d` y `descfgsv -d -i txinsnnn` para comprobar la configuración del servicio de búsqueda.

RC_SE_IO_PROBLEM

Explicación: Se ha producido un error cuando el servidor intentaba abrir o leer uno de sus archivos de índice. Esto puede ser debido a una de las siguientes situaciones:

Una acción no intencionada del administrador como, por ejemplo, la supresión de un archivo de índice de DB2 Text Extender

Un valor incorrecto en la configuración del texto.

Lo que se ha de hacer: Termine la aplicación. Compruebe con el administrador que:

- Existen todos los archivos del índice de DB2 Text Extender actual
 - Los valores de configuración del texto son correctos.
-

RC_SE_LS_FUNCTION_FAILED

Explicación: Una función que accedía a la base de datos para recuperar documentos de texto para el examen ha fallado. El usuario ya no puede acceder a la base de datos o el usuario no está autorizado para la tabla de texto.

Lo que se ha de hacer: Compruebe que la entrada para la función como, por ejemplo, el ID de usuario es correcta. Compruebe que se puede

Códigos de retorno

acceder a la base de datos y que el usuario está autorizado para la tarea.

RC_SE_LS_NOT_EXECUTABLE

Explicación: Una función que intenta acceder a la base de datos para recuperar documentos de texto para examen no se puede ejecutar.

Lo que se ha de hacer: Compruebe que DB2 Text Extender está instalado correctamente. Si el problema persiste, póngase en contacto con el representante de IBM.

RC_SE_MAX_OUTPUT_SIZE_EXCEEDED

Explicación: Se ha encontrado un número inusualmente grande de coincidencias. El tamaño de la información de examen ha excedido el máximo que se puede manejar. La petición no puede procesarse.

Lo que se ha de hacer: Especifique más la consulta o asegúrese de que está disponible más memoria del sistema.

RC_SE_MAX_NUMBER_OF_BUSY_INDEXES

Explicación: El servicio de búsqueda ha impedido la función pedida porque el número máximo de índices está activo actualmente.

Lo que se ha de hacer: Vuelva a emitir la función después de un corto período de tiempo. En general, el problema es sólo temporal.

RC_SE_NO_DATA

Explicación: No se trata de ningún error. Ningún documento de texto coincide con los criterios de búsqueda. Si pide información de examen, no se devuelve ninguna. No se asigna almacenamiento para la información de examen.

Lo que se ha de hacer: No es necesaria ninguna acción.

RC_SE_NOT_ENOUGH_MEMORY

Explicación: No hay suficiente espacio de almacenamiento en el sistema cliente o servidor. No puede procesarse la petición actual.

Lo que se ha de hacer: Libere espacio de almacenamiento y finalice la aplicación.

RC_SE_PROCESSING_LIMIT_EXCEEDED

Explicación: La petición de búsqueda actual ha excedido el tamaño máximo del resultado o el tiempo máximo de proceso especificado para el entorno de cliente/servidor. Se ha cancelado la petición.

Lo que se ha de hacer: Especifique más la petición de búsqueda. Tome en consideración la posibilidad de aumentar el tiempo máximo de proceso.

RC_SE_QUERY_TOO_COMPLEX

Explicación: La consulta especificada es demasiado compleja.

Lo que se ha de hacer: Adapte la aplicación para evitar el uso excesivo de caracteres de máscara y de sinónimos.

El uso excesivo de símbolos de máscara o el uso excesivo de la opción SYNONYM puede ampliar la consulta a un tamaño que DB2 Text Extender no pueda gestionar.

RC_SE_REQUEST_IN_PROGRESS

Explicación: Se ha llamado a un servicio API de examen de DB2 Text Extender mientras estaba activada otra petición API de examen para la misma sesión.

Lo que se ha de hacer: Finalice la sesión llamando a DesEndBrowseSession y libere almacenamiento llamando a DesFreeBrowseInfo.

La API de examen de DB2 Text Extender no

soporta el acceso simultáneo a la misma sesión de examen.

Todas las aplicaciones que se ejecutan simultáneamente en el mismo proceso deben manejar sus propias sesiones de examen.

RC_SE_SERVER_BUSY

Explicación: El cliente DB2 Text Extender no puede establecer actualmente una sesión con el servidor DB2 Text Extender pedido o se ha interrumpido el enlace de comunicaciones del servidor DB2 Text Extender y no se puede volver a establecer.

El servidor DB2 Text Extender se ha iniciado correctamente, pero se ha alcanzado el número máximo de procesos paralelos del servidor.

Lo que se ha de hacer: Si no se trata de un problema temporal, cambie la configuración de las comunicaciones en el servidor.

RC_SE_SERVER_CONNECTION_LOST

Explicación: Se ha interrumpido la comunicación entre el cliente y el servidor y no se puede volver a establecer.

Puede que un administrador haya detenido la tarea del servidor DB2 Text Extender o puede que se haya concluido la estación de trabajo servidor.

Lo que se ha de hacer: Compruebe si se ha producido alguna de estas condiciones y corríjalas.

RC_SE_SERVER_NOT_AVAILABLE

Explicación: Los servicios API de DB2 Text Extender no han podido establecer una sesión con el servidor DB2 Text Extender pedido.

Puede que no se haya iniciado el servidor DB2 Text Extender.

Lo que se ha de hacer: Compruebe que el servidor DB2 Text Extender se haya iniciado

correctamente. Si persiste el error, es que hay un problema de instalación.

RC_SE_STOPWORD_IGNORED

Explicación: Este código informativo se devuelve cuando la consulta especificada contiene como mínimo un término de búsqueda que sólo consta de palabras de parada. Se ha pasado por alto el término de búsqueda al procesar la consulta.

Lo que se ha de hacer: Puede continuar emitiendo llamadas a API. Evite la utilización de palabras de parada en consultas DB2 Text Extender.

RC_SE_UNEXPECTED_ERROR

Explicación: Se ha producido un error que podría estar provocado por una instalación incorrecta de DB2 Text Extender.

Lo que se ha de hacer: Finalice la aplicación, guardando la información que pueda ayudar a averiguar la causa del error.

RC_SE_UNKNOWN_INDEX_NAME

Explicación: El nombre del índice de texto asociado a la columna de texto forma parte del descriptor de contexto.

Lo que se ha de hacer: Asegúrese de que el descriptor de contexto que utiliza como entrada para DesGetBrowseInfo sea correcto.

RC_SE_UNKNOWN_SECTION_NAME

Explicación: Un nombre de sección especificado no forma parte de ningún modelo especificado en un archivo de modelos de documentos o del modelo por omisión utilizado.

Lo que se ha de hacer: Especifique un nombre de sección que forme parte del modelo especificado o del modelo por omisión.

Códigos de retorno

RC_SE_UNKNOWN_SERVER_NAME

Explicación: El nombre del servidor DB2 Text Extender forma parte del descriptor de contexto.

Lo que se ha de hacer: Asegúrese de que el descriptor de contexto que utiliza como entrada para DesGetBrowseInfo sea correcto.

RC_SE_WRITE_TO_DISK_ERROR

Explicación: Se ha producido un error de grabación que podría estar provocado por un disco lleno en la estación de trabajo del servidor DB2 Text Extender o por una instalación incorrecta de DB2 Text Extender.

Lo que se ha de hacer: Finalice la aplicación, guardando la información que pueda ayudar a averiguar la causa del error. Compruebe que haya suficiente espacio de disco disponible en el servidor.

RC_SQL_ERROR_WITH_INFO

Explicación: Se ha producido un error SQL. Se devuelve un mensaje de error.

Lo que se ha de hacer: Compruebe el mensaje de error devuelto por DB2 Text Extender para obtener más información como, por ejemplo, el mensaje de error SQL, el SQLState y el código de error SQL nativo.

RC_SQL_ERROR_NO_INFO

Explicación: Se ha producido un error SQL. No se ha devuelto ningún mensaje.

RC_TEXT_COLUMN_NOT_ENABLED

Explicación: La columna de descriptores de contexto especificada no es una columna de la tabla que ha especificado.

Lo que se ha de hacer: Compruebe si el nombre de la columna de descriptores de contexto que

ha especificado es correcto. Asegúrese de que la columna de texto de esa tabla se ha habilitado.

Capítulo 18. Mensajes

Este capítulo describe lo siguiente:

- **Indicaciones SQL devueltas por funciones de DB2 Text Extender:** Estos mensajes se pueden visualizar cuando se utilizan funciones de DB2 Text Extender.
- **Mensajes del procesador de línea de mandatos DB2TX:** Estos mensajes pueden visualizarse cuando se entran mandatos utilizando el procesador de línea de mandatos DB2TX. Cada número de mensaje lleva el prefijo DES.

Indicaciones SQL devueltas por funciones de DB2 Text Extender

Las funciones SQL proporcionadas por DB2 Text Extender pueden devolver indicaciones de error. Ejemplo:

```
SQL0443N La función definida por el usuario
"DB2TX.CONTAINS" (nombre específico "DES5A")
ha devuelto un error SQLSTATE con el
texto de diagnóstico "No se puede abrir el archivo de mensajes".
SQLSTATE=38702
```

Los mensajes de esta sección están ordenados por número de SQLSTATE.

01H10 El archivo *nombre-archivo* no se puede abrir.

Lo que se ha de hacer: Asegúrese de que existe el archivo y que el nombre de instancia de DB2 tiene los permisos necesarios para abrirlo.

01H11 El descriptor de contexto está incompleto

Explicación: Se ha intentado utilizar un descriptor de contexto que se ha inicializado, pero no se ha completado. Se ha creado un descriptor de contexto parcial utilizando INIT_TEXT_HANDLE que contiene valores preestablecidos para el idioma y el formato del documento. Sin embargo, un desencadenante no ha completado el descriptor de contexto.

Lo que se ha de hacer: Utilice solamente los descriptores de contexto que se hayan completado. Si el descriptor de contexto implicado se ha almacenado en una columna de

descriptores de contexto, habilite o vuelva a habilitar su columna de texto correspondiente.

01H12 Argumentos de búsqueda demasiado largos. Se ha pasado por alto el segundo argumento.

Explicación: Se ha utilizado la función REFINE para combinar dos argumentos de búsqueda, pero la longitud combinada de los argumentos de búsqueda es superior al máximo permitido para LONG VARCHAR. La función REFINE devuelve el primer argumento de búsqueda en lugar de uno combinado.

Lo que se ha de hacer: Reduzca la longitud de uno o de ambos argumentos de búsqueda, después repita la consulta.

Indicaciones SQL devueltas por funciones DB2 Text Extender

01H13 El argumento de búsqueda contiene una palabra de parada.

Explicación: La consulta especificada contiene como mínimo un término de búsqueda que sólo consta de palabras de parada. Se ha pasado por alto el término de búsqueda al procesar la consulta.

Lo que se ha de hacer: Evite la utilización de palabras de parada en consultas DB2 Text Extender.

01H14 Falta un diccionario de idioma para el proceso lingüístico.

Explicación: Los servicios lingüísticos de DB2 Text Extender no pueden encontrar los archivos de diccionario. La consulta se procesa sin el soporte lingüístico. Los archivos de diccionario correspondientes al código o códigos del idioma especificado no están en la vía de acceso esperada.

Lo que se ha de hacer: Para UNIX, compruebe que el diccionario necesario está en la vía de acceso {DB2TX_INSTOWNERHOMEDIR}/db2tx/dicts. Para OS/2, compruebe que el diccionario necesario está en la vía de acceso especificada en los valores de configuración del texto. Si es necesario, instale el diccionario necesario.

01H15 Una especificación de término de búsqueda lingüística no coincide con el tipo de índice.

Explicación: La especificación del término de búsqueda lingüístico de la consulta no corresponde al tipo de índice. Por ejemplo, PRECISE FORM OF no debe utilizarse con un índice lingüístico. La especificación de lingüístico por omisión se utiliza tal como se muestra en la Tabla 8 en la página 223.

Lo que se ha de hacer: Adapte la aplicación para impedir la especificación de opciones de consulta que entren en conflicto con el tipo de índice.

38700 La biblioteca de Text Extender no está actualizada.

Explicación: Se ha intentado utilizar un descriptor de contexto que sólo puede interpretarse por una versión posterior de DB2 Text Extender.

Lo que se ha de hacer: Asegúrese de que la vía de acceso para la versión actual de la biblioteca esté establecida correctamente y que tiene los permisos necesarios para acceder a ella.

Consulte la vista del catálogo de DB2 SYSCAT.FUNCTIONS, en la columna IMPLEMENTATION, para la función que ha causado el problema.

38701 *archivorastreo* No se puede abrir este archivo de rastreo.

Explicación: Se ha intentado utilizar una función de rastreo que graba en el archivo DB2TX_TRACEFILE del directorio DB2TX_TRACEDIR. El archivo no existe, no se puede encontrar o no están disponibles los permisos necesarios para el archivo.

38702 No se puede abrir el archivo de mensajes *archivo-mensajes*.

Explicación: Se ha generado una situación que ha provocado que DB2 Text Extender intentase devolver un mensaje. El archivo que contiene los mensajes no existe o no se ha podido encontrar o los permisos necesarios para el archivo no están disponibles.

Lo que se ha de hacer: Asegúrese de que el archivo existe, que la vía de acceso está establecida correctamente y que tiene los permisos necesarios para abrir el archivo.

38704 El formato del descriptor de contexto es incorrecto.

Explicación: Se ha utilizado un descriptor de contexto que tiene un formato incorrecto como argumento para una función de DB2 Text Extender.

Lo que se ha de hacer: Asegúrese de que el

Indicaciones SQL devueltas por funciones DB2 Text Extender

descriptor de contexto no lo ha producido
INIT_TEXT_HANDLE.

38705 *nombreudf* **Declaración de UDF incorrecta.**

Explicación: El nombre específico de una función de DB2 Text Extender se ha cambiado en el script en el que se declaran las funciones. Los nombres de funciones de DB2 Text Extender pueden cambiarse, pero no sus nombres específicos.

Lo que se ha de hacer: Compruebe el script DESCVDF.DDL que contiene las declaraciones de funciones de DB2 Text Extender, para asegurarse de que se siguen utilizando los nombres correctos. Compruebe los nombres con los del soporte de almacenamiento de distribución original.

38706 *atributo* **No se puede reconocer el valor de este atributo.**

Explicación: Se ha intentado establecer un CCSID, un formato o un idioma en un valor desconocido.

Lo que se ha de hacer: Consulte el “Capítulo 4. Planificación para las necesidades de la búsqueda” en la página 29 para ver los valores correctos.

38707 **La función solicitada ya no se soporta.**

Explicación: La función especificada ya no se soporta.

Lo que se ha de hacer: Compruebe la función especificada.

38708 *código de retorno*

Explicación: Se ha producido un error al procesar la petición de búsqueda.

Lo que se ha de hacer: Consulte la descripción del código de retorno en el “Capítulo 17. Códigos de retorno” en la página 291.

38709 **No hay suficiente memoria disponible.**

Explicación: No hay suficiente memoria disponible para ejecutar la función de DB2 Text Extender.

Lo que se ha de hacer: Cierre cualquier aplicación no necesaria para liberar memoria, después vuelva a intentarlo.

38710 *númeroerror* **No se puede acceder a los resultados de búsqueda.**

Explicación: Se ha producido un error al intentar leer la lista de documentos encontrados (lista de resultados) devueltos por el servicio de búsqueda.

Lo que se ha de hacer: Intente repetir la búsqueda. Si no es satisfactoria, reinicie el servicio de búsqueda. Si el problema persiste, informe del mismo al representante local de IBM, indicando el número del error.

38711 **Error interno grave.**

Explicación: Se ha producido un error grave.

Lo que se ha de hacer: Informe del error al representante local de IBM, indicando las circunstancias bajo las cuales se ha producido.

38712 *nombreíndice* **Descriptor de contexto incorrecto en este índice de texto.**

Explicación: Se ha dañado un descriptor de contexto.

Lo que se ha de hacer: Utilice UPDATE INDEX para volver a crear el índice.

38714 **Reduzca la variable de entorno DB2TX_INSTOWNERHOMEDIR.**

Explicación: El nombre del directorio inicial del propietario de la instancia no puede sobrepasar los 256 caracteres.

Lo que se ha de hacer: Utilice enlaces para reducir la longitud del nombre de directorio.

Indicaciones SQL devueltas por funciones DB2 Text Extender

38717 **No se ha podido encontrar el tesoro especificado.**

Explicación: No se ha podido encontrar el tesoro especificado.

Lo que se ha de hacer: Compruebe el nombre de tesoro especificado.

38718 **No se ha podido encontrar en el tesoro el nombre de relación especificado.**

Explicación: La relación especificada no existe en el tesoro especificado.

Lo que se ha de hacer: Asegúrese de que la relación especificada existe.

38719 **Se ha producido un error de proceso de búsqueda. Código de razón: rc.**

Explicación: No se ha podido efectuar la búsqueda debido a la razón especificada.

Lo que se ha de hacer: Intente resolver el problema informado por el código de razón. Si la razón especificada no sirve de ayuda y no se encuentra más información en el archivo `desdiag.log`, cree un rastreo e informe de la información al representante local de IBM.

38720 **Se ha producido un error de conexión de memoria compartida.**

Explicación: El sistema no puede obtener acceso a la memoria compartida.

Lo que se ha de hacer: Compruebe la configuración del sistema y aumente los recursos compartidos o consulte el uso actual del recurso compartido (ipcs) y limpie los recursos que ya no se necesitan.

38721 **Se ha producido un error de creación/acceso de semáforo.**

Explicación: El sistema no puede crear o no puede obtener acceso a un semáforo.

Lo que se ha de hacer: Compruebe la configuración del sistema y aumente los recursos compartidos o consulte el uso actual del recurso compartido (ipcs) y limpie los recursos que ya no se necesitan.

38722 **No ha vuelto un proceso de búsqueda.**

Explicación: Se ha producido un error al procesar la petición de búsqueda.

Lo que se ha de hacer: Verifique la configuración del sistema `descfgc1` y compruebe si todos los nodos están activados y en ejecución.

38723 **El CCSID de la consulta y del índice no coinciden.**

Explicación: El CCSID de base de datos utilizado para la serie de consulta no es el mismo que el CCSID del índice de texto.

Lo que se ha de hacer: Inhabilite el índice de texto y vuelva a crearlo utilizando el CCSID de la base de datos.

38724 **El nombre de sección o de modelo es incorrecto.**

Explicación: El nombre de sección o de modelo especificado en la consulta es incorrecto.

Lo que se ha de hacer: Compruebe el nombre de sección o de modelo.

38726 **Se ha producido un error de lectura de archivo de modelos.**

Explicación: No se ha encontrado un archivo de definiciones de modelos o no se puede abrir.

Lo que se ha de hacer: Compruebe que el archivo de definiciones de modelos existe y que está en el directorio de índices.

Mensajes de DB2 Text Extender

Cada mensaje tiene un identificador de mensaje que consta de un prefijo (DES), el número del mensaje y un sufijo de una letra. El sufijo de una letra indica la gravedad de la aparición que ha producido el mensaje:

- I** Mensaje informativo
- W** Mensaje de aviso
- N** Mensaje de error (o “negativo”)
- C** Mensaje de error grave.

DES0001N Número incorrecto de argumentos para el mandato db2txinstance.

Explicación: El mandato db2txinstance necesita dos argumentos.

Lo que se ha de hacer: Entre el mandato de nuevo con estos argumentos:

```
db2txinstance nombreInstancia nombreInstanciaDb2
```

donde *nombreInstancia* es el nombre de inicio de sesión de un usuario UNIX existente que se está asignando como el propietario de esta instancia y *nombreInstanciaDb2* es el nombre de inicio de sesión del propietario de la instancia DB2 correspondiente.

DES0002N Nombre de instancia no válido.

Explicación: El nombre de instancia especificado debe ser el nombre de inicio de sesión de un usuario UNIX existente.

Lo que se ha de hacer: Corrija el nombre de instancia o seleccione un usuario UNIX existente o cree un usuario UNIX para que sea el propietario de la instancia.

Entre el mandato db2txinstance de nuevo de la siguiente manera:

```
db2txinstance nombreInstancia
```

donde *nombreInstancia* es el nombre de inicio de sesión del usuario UNIX seleccionado.

DES0004N La instancia especificada ya existe. No se puede procesar el mandato.

Explicación: El *nombreInstancia* especifica el nombre de inicio de sesión de un usuario UNIX que es el propietario de la instancia. Este propietario de la instancia ya tiene un directorio db2tx en el directorio inicial.

Lo que se ha de hacer: Para crear la instancia, suprime la instancia existente y, a continuación, vuelva a intentar el mandato.

DES0005N No se puede encontrar el catálogo de mensajes de instalación.

Explicación: Falta en el sistema, el catálogo de mensajes necesario para los scripts de instalación; puede que se haya suprimido o puede que se hayan cargado incorrectamente los productos de bases de datos.

Lo que se ha de hacer: Verifique que la opción del producto db2tx_01_01_0000.client está instalada correctamente. Si hay errores de verificación, vuelva a instalar la opción.

DES0015W Una especificación de término de búsqueda lingüístico no coincide con el tipo de índice.

Explicación: La especificación del término de búsqueda lingüístico de la consulta con corresponde al tipo de índice. Por ejemplo, PRECISE FORM OF no debe utilizarse con un índice lingüístico. La especificación de lingüístico por omisión se utiliza tal como se muestra en la Tabla 8 en la página 223.

Mensajes de DB2 Text Extender

Lo que se ha de hacer: Adapte la aplicación para impedir la especificación de opciones de consulta que entren en conflicto con el tipo de índice.

DES0016W Una especificación de idioma no está soportada para el tipo de índice actual.

Explicación: El idioma que ha especificado no se soporta para el tipo de índice especificado.

Lo que se ha de hacer: Consulte la documentación para ver una lista de idiomas soportados para el tipo de índice.

DES0017W La extracción de características no se ha habilitado.

Explicación: Ha utilizado un argumento de búsqueda de características en la consulta pero el índice no se ha creado con la opción de índice FEATURE_EXTRACTION.

Lo que se ha de hacer: Cambie la opción de índice a FEATURE_EXTRACTION.

DES0018W opción no se soporta para el tipo de índice actual.

Explicación: Ha pedido una opción de búsqueda que no se soporta para el tipo de índice y la opción de índice actuales.

Lo que se ha de hacer: Compruebe qué tipo de índice u opción de índice soporta la opción de búsqueda pedida. Consulte el apartado Tabla 8 en la página 223.

DES0121N No se ha podido asignar memoria (malloc ha fallado).

Explicación: No se ha podido reservar almacenamiento para la aplicación.

Lo que se ha de hacer: Aumente el espacio de paginación.

DES0333N Se ha producido un problema de E/S de archivo de índice de texto.

Explicación: El cliente DB2 Text Extender no puede establecer una sesión con el servidor pedido.

Lo que se ha de hacer: Compruebe que el servidor DB2 Text Extender se haya iniciado. Si no, ejecute TXSTART.

DES0709W El diccionario del idioma especificado no está instalado.

Explicación: DB2 Text Extender no puede encontrar los archivos del diccionario.

Lo que se ha de hacer: Instale o vuelva a instalar el diccionario para el idioma especificado.

DES0377N Se ha producido un problema de E/S de archivo de índice de texto.

Explicación: DB2 Text Extender no puede acceder al índice de texto. Esto puede ocurrir si el valor de DIRECTORY en la configuración del texto apunta a un directorio no válido.

Lo que se ha de hacer: Compruebe los valores de configuración del texto.

DES0700N El valor de número de nodo 'nodo' no está contenido en la definición del grupo de nodos.

Explicación: El número de nodo especificado no es válido.

Lo que se ha de hacer: Compruebe el número de nodo DB2.

DES0701N El valor del número de nodo 'nodo' está fuera de rango.

Explicación: El número de nodo especificado no es válido.

Lo que se ha de hacer: Compruebe el número de nodo DB2.

DES0704N El formato '*formato*' necesita la especificación de una propiedad de índice.

Explicación: El formato del documento no es compatible con la información de tipo de índice.

Lo que se ha de hacer: Especifique una propiedad de índice que sea compatible con el formato de documento.

DES0705N No se ha encontrado el nombre de modelo de documento '*modelo*' en el archivo de definición de modelos.

Explicación: El nombre de modelo de documento no se ha encontrado en el archivo de definición de modelos. Tenga en cuenta que el nombre de modelo es sensible a las mayúsculas y minúsculas.

Lo que se ha de hacer: Utilice un nombre de modelo que esté especificado en el archivo de definición de modelos.

DES0706N No se puede acceder al archivo de definición de modelos en el servidor DB2 Text Extender.

Explicación: No se ha encontrado el archivo de definición de modelos o no se puede abrir.

Lo que se ha de hacer: Compruebe que existe el archivo de definición de modelos.

DES0707N El formato '*formato*' no soporta la propiedad de índice especificada.

Explicación: El formato del documento no soporta la propiedad de índice.

Lo que se ha de hacer: Especifique una propiedad de índice que sea compatible con el formato de documento.

DES0710N No se permite un puntero nulo para el parámetro '*parámetro*'.

Explicación: No se ha especificado ningún valor para *parámetro*.

Lo que se ha de hacer: Especifique un valor para el parámetro.

DES0711N Se ha producido un error interno de Text Extender. Información de diagnóstico: *mensaje*.

Explicación: Se ha producido un error de proceso interno.

Lo que se ha de hacer: Compruebe el mensaje de diagnóstico para resolver el problema. Si el error interno no lo ha provocado un problema de instalación, puede haber información adicional en el archivo *desdiag.log* o en el archivo de rastreo creado. Si esto no sirve de ayuda, reúna la información disponible y llame al servicio técnico de IBM.

DES0712N El parámetro '*parámetro*' es demasiado largo.

Explicación: El parámetro especificado está fuera de rango.

Lo que se ha de hacer: Especifique el parámetro utilizando una longitud válida.

DES0713N La longitud del parámetro '*parámetro*' es incorrecta: %d1.

Explicación: El parámetro especificado está fuera de rango.

Lo que se ha de hacer: Especifique el parámetro utilizando una longitud válida.

DES0714N Especifique el parámetro *parámetro* directamente o en de la tabla de configuración.

Explicación: No se ha especificado el CCSID, el formato o el idioma y no hay ningún valor de configuración del texto para este valor.

Lo que se ha de hacer: Especifique el parámetro que falta directamente en el mandato `ENABLE TEXT COLUMN` o establezca un valor en los valores de configuración del texto.

Mensajes de DB2 Text Extender

DES0715N El tipo de datos *esquema.tipo* no está soportado para datos de texto.

Explicación: *esquema.tipo* es el nombre de esquema y el nombre de tipo de la columna de texto o el resultado de una función de acceso. El tipo de datos para una columna de texto no está soportado por DB2 Text Extender. Debe ser CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR o CLOB. Si este no es el caso, debe proporcionar una función de acceso cuya entrada sea del tipo de datos de la columna de texto y cuya salida sea VARCHAR, LONG VARCHAR o LOB.

Lo que se ha de hacer: Si *esquema.tipo* es un tipo de columna de texto, debe registrar una función de acceso con un resultado de tipo VARCHAR, LONG VARCHAR o LOB. Si *esquema.tipo* es el resultado de una función de acceso, no se puede utilizar. Proporcione una función de acceso con un resultado del tipo necesario.

DES0716N El formato *formato* no está soportado.

Explicación: *formato* es un formato que no está soportado por Text Extender.

Lo que se ha de hacer: Compruebe la lista de formatos soportados en el apartado “Formatos de documentos que se soportan” en la página 31.

DES0717N El idioma *idioma* no está soportado.

Explicación: *idioma* es un idioma que DB2 Text Extender no soporta o el idioma seleccionado no está soportado por el tipo de índice especificado.

Lo que se ha de hacer: Compruebe la lista de idiomas soportados en la Tabla 5 en la página 54.

DES0718N El CCSID *valor_ccsid* no está soportado.

Explicación: Ha especificado un valor de CCSID no válido.

Lo que se ha de hacer: Consulte la

documentación para ver una lista de los CCSID soportados.

DES0719N Una llamada al programa DB2 Text Extender *programa* ha fallado con el código de retorno *rc*.

Explicación: Puede haberse producido un error durante la instalación. Los códigos de retorno se listan en el archivo DES_EXT.H.

Lo que se ha de hacer: Compruebe si la instalación ha sido satisfactoria. Compruebe que las variables de entorno como, por ejemplo, DB2TX_INSTOWNER y DB2TX_INSTOWNERHOMEDIR estén establecidas correctamente.

DES0720N La función de acceso *esquema.función* no está registrada en la base de datos.

Explicación: El nombre de la función es incorrecto o no se ha registrado en la base de datos.

Lo que se ha de hacer: Compruebe el nombre de la función de acceso. Si es correcto, compruebe que se conozca la función en el sistema de bases de datos. Utilice CREATE FUNCTION para registrar la función de acceso en la base de datos.

DES0721N La base de datos es incoherente; falta una vista del catálogo de DB2 Text Extender.

Explicación: Una de las vistas del catálogo de DB2 Text Extender no está en la base de datos.

Lo que se ha de hacer: Utilice el mandato DISABLE DATABASE para eliminar las vistas del catálogo que quedan y, a continuación, vuelva a entrar ENABLE DATABASE. Los datos del índice se han perdido; vuelva a indexar los documentos de texto.

DES0722N La tabla *esquema.tabla* no es una tabla base de la base de datos.

Explicación: La tabla no existe en la base de datos o es una tabla de resultados o una vista. Debe haber una columna de texto en una tabla base para que pueda habilitarse para DB2 Text Extender.

Lo que se ha de hacer: Asegúrese de que el nombre de tabla sea correcto y de que sea una tabla base.

DES0723N La creación de un índice para la columna de descriptores de contexto *nombrecolumnadescriptorescontexto* en la tabla *esquema.tabla* ha fallado.

Explicación: No se ha podido crear el índice de texto para la columna de descriptores de contexto.

Lo que se ha de hacer: Utilice `txstatus` para comprobar el estado del servidor. Si los servicios del servidor están ejecutándose correctamente, utilice `DISABLE TEXT COLUMN` o `DISABLE TEXT TABLE` para obtener un estado coherente de nuevo. Después habilite la columna de texto de nuevo utilizando `ENABLE TEXT COLUMN` o `ENABLE TEXT TABLE`.

DES0724N Falta una entrada en la vista del catálogo `TextIndices` para la columna de descriptores de contexto *columnadescriptorescontexto* en la tabla *esquema.tabla*.

Explicación: La vista del catálogo `TextIndices` está dañada.

Lo que se ha de hacer: Utilice `DISABLE TEXT COLUMN` o `DISABLE TEXT TABLE` para obtener un estado coherente de nuevo. Después habilite la columna de texto utilizando `ENABLE TEXT COLUMN` o `ENABLE TEXT TABLE`.

DES0727N La columna *columna* de la tabla *esquema.tabla* ya está habilitada.

Explicación: Este mensaje puede aparecer si la tabla se ha desechado y, después, se vuelve a crear utilizando la misma columna de texto, sin primero inhabilitar la columna.

Lo que se ha de hacer: Inhabilite la columna y, después, vuelva a intentarlo.

DES0728N La columna *columna* no existe en la tabla *esquema.tabla*.

Explicación: Está intentando habilitar una columna de texto que no existe.

Lo que se ha de hacer: Cambie el nombre de tabla o el nombre de columna y vuelva a intentarlo.

DES0729N La columna de descriptores de contexto *columnadescriptorescontexto* no existe en la tabla *esquema.tabla*.

Explicación: Está intentando utilizar una columna de descriptores de contexto que no existe.

Lo que se ha de hacer: Utilice el mandato `GET STATUS` para comprobar si existe la columna de descriptores de contexto y que su nombre se haya especificado correctamente.

DES0730N La tabla *esquema.tabla* ya está habilitada como tabla de índice común.

Explicación: Está intentando habilitar una tabla que ya se ha habilitado como tabla de índice común.

Lo que se ha de hacer: Continúe sin habilitar la tabla o ejecute el mandato `DISABLE TEXT TABLE` para inhabilitar la tabla antes de volverla a habilitar.

Mensajes de DB2 Text Extender

DES0731N La tabla *esquema.tabla* no está habilitada para Text Extender; no puede inhabilitarse.

Explicación: Está intentando inhabilitar una tabla que no se ha habilitado.

Lo que se ha de hacer: Compruebe el nombre de tabla.

DES0732N La frecuencia de actualización es incorrecta cerca de la *ubicación*; la esperada era *parámetro*.

Explicación: La especificación de *parámetro* para la Frecuencia de actualización no era correcta.

Lo que se ha de hacer: Compruebe el parámetro de frecuencia de actualización y vuelva a entrar el mandato.

DES0733N La tabla *esquema.tabla* contiene una columna habilitada; no puede habilitarse como tabla de índice común.

Explicación: Esta tabla contiene una columna de texto que ya tiene su propio índice. No puede crear un índice común para todas las columnas de texto mientras existe este índice individual.

Lo que se ha de hacer: Utilice DISABLE TEXT COLUMN para inhabilitar las columnas habilitadas, después entre el mandato ENABLE TEXT TABLE de nuevo.

DES0734N La columna de descriptores de contexto *columnadescriptorescontexto* pertenece a la tabla de texto parcial *esquema.tabla*; no puede inhabilitarse por separado.

Explicación: No puede inhabilitar una única columna de texto de una tabla que se ha habilitado como tabla de texto parcial.

Lo que se ha de hacer: Inhabilite toda la tabla de texto parcial.

DES0736N *columnadescriptorescontexto* ya es una columna de descriptores de contexto en la tabla *esquema.tabla*.

Explicación: Está intentando utilizar un nombre de columna de descriptores de contexto existente.

Lo que se ha de hacer: Vuelva a entrar el mandato, utilizando un nombre diferente para la columna de descriptores de contexto.

DES0737N La tabla *esquema.nombretabla* se ha habilitado como una tabla de índice común con la opción **STORAGEopción_almacenamiento**.

Explicación: No es posible habilitar una tabla de índice común para archivos externos.

Lo que se ha de hacer: Si desea habilitar una tabla para archivos externos, utilice una tabla de múltiples índices.

DES0738N La función de acceso *esquema.función* tiene parámetros incorrectos.

Explicación: Los parámetros de entrada o salida de *esquema.función* son incorrectos.

- Sólo puede haber un parámetro de entrada y debe ser del tipo de datos de la columna de texto que se ha de habilitar.
 - El parámetro de salida debe ser de tipo CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR o CLOB.
-

DES0739W El programa de actualización de índice para la tabla *esquema.tabla*, nombre de columna de descriptores de contexto *columnadescriptorescontexto*, no se ha podido iniciar.

Explicación: El programa que actualiza los índices no se ha podido iniciar. Puede haberse producido un error durante la instalación.

Lo que se ha de hacer: Compruebe si la instalación ha sido satisfactoria. Compruebe que las variables de entorno como, por ejemplo,

DB2TX_INSTOWNER y DB2TX_INSTOWNERHOMEDIR estén establecidas correctamente.

DES0740N La columna de descriptores de contexto '*columna*' sólo puede volverse a utilizar para una columna de texto de tipo *tipocolumna*.

Explicación: La columna de descriptores de contexto ya se ha utilizado para otro tipo de columna de descriptores de contexto.

Lo que se ha de hacer: Especifique un nuevo nombre de columna de descriptores de contexto.

DES0741N No se ha encontrado el *parámetro de programa o archivo, o no se ha podido iniciar*.

Explicación: El mandato ENABLE DATABASE o DISABLE DATABASE no ha podido abrir el *parámetro* de archivo. Puede haberse producido un error durante la instalación.

Lo que se ha de hacer: Compruebe si la instalación ha sido satisfactoria.

DES0742N Esta tabla utiliza índices de columna. Especifique una columna de descriptores de contexto en el mandato.

Explicación: La tabla especificada se ha habilitado como tabla de múltiples índices. Para trabajar con una columna específica especifique la columna de descriptores de contexto relacionada.

Lo que se ha de hacer: Especifique una columna de descriptores de contexto.

DES0745N El propietario de la instancia de DB2TX *propietario-instancia* no es un ID de usuario válido.

Explicación: La variable de entorno DB2TX_INSTOWNER no contiene un ID de usuario válido.

Lo que se ha de hacer: Corrija la variable de entorno.

DES0747N El CCSID actual no está soportado para el tipo de índice *tipo_índice*.

Explicación: Ha especificado un CCSID que no está soportado para el tipo de índice pedido.

Lo que se ha de hacer: Consulte la documentación para ver una lista de los CCSID soportados.

DES0748N El valor de cuenta de compromiso '*cuentacompromiso*' no lo soporta DB2 Text Extender.

Explicación: El valor de cuenta de compromiso especificado no está soportado.

Lo que se ha de hacer: Especifique un valor de cuenta de compromiso válido.

DES0749N El valor de índice de actualización '*valoríndice*' no lo conoce DB2 Text Extender.

Explicación: El valor de 'índice de actualización' no es válido.

Lo que se ha de hacer: Especifique un valor de índice de actualización válido.

DES0750N El valor de tipo de índice '*tipóíndice*' no lo conoce DB2 Text Extender.

Explicación: El valor de tipo de índice especificado no es válido.

Lo que se ha de hacer: Especifique un valor de tipo de índice de válido

DES0751N No tiene la autorización para realizar la operación especificada.

Explicación: No tiene la autorización necesaria de administrador de bases de datos para realizar esta operación.

Lo que se ha de hacer: Pida a un administrador de bases de datos que realice esta operación.

Mensajes de DB2 Text Extender

DES0756N La base de datos no está habilitada para Text Extender.

Explicación: La base de datos debe estar habilitada para que se pueda ejecutar este mandato.

Lo que se ha de hacer: Ejecute ENABLE DATABASE y, a continuación, vuelva a someter el mandato.

DES0763N La frecuencia de actualización no es correcta.

Explicación: El valor de 'frecuencia de actualización' especificado no es válido.

Lo que se ha de hacer: Especifique una 'frecuencia de actualización' válida.

DES0765N La base de datos ya está habilitada para Text Extender.

Explicación: Está intentando habilitar una base de datos que ya está habilitada.

Lo que se ha de hacer: Continúe sin habilitar la base de datos o utilice DISABLE DATABASE para inhabilitar la base de datos antes de volverla a habilitar.

DES0766N Una acción ha provocado que se sobrepasase el tamaño máximo de fila de la tabla o una tabla temporal.

Explicación: El mandato ENABLE TEXT COLUMN añade una columna de descriptores de contexto a la tabla. Si la tabla ya es grande, puede provocar que el tamaño de fila de la tabla exceda el valor máximo de 4005.

El mandato ENABLE TEXT COLUMN también crea una tabla temporal cuyo tamaño es proporcional al número de columnas de texto que ya se han habilitado. Si ya hay muchas columnas de texto habilitadas, el tamaño de la tabla temporal puede sobrepasar el valor máximo.

Lo que se ha de hacer: Utilice ENABLE TEXT COLUMN solamente en tablas que no provocan

que se sobrepase este límite.

DES0769W Aviso: Se han especificado características de índice pero se pasarán por alto. La tabla 'nombretabla' es una tabla de índice común.

Explicación: La tabla especificada es una tabla de índice común y, por lo tanto, no se puede especificar ninguna característica de índice.

Lo que se ha de hacer: No es necesaria ninguna acción.

DES0770N La variable de entorno *variable-ent* no está definida.

Explicación: No se ha especificado un parámetro para un mandato y el sistema ha intentado leer el valor por omisión de la variable de entorno *variable-ent*, pero esta variable de entorno no está definida.

Lo que se ha de hacer: Defina la variable de entorno necesaria.

DES0774N La longitud *longitud* del valor para la variable '*variable*' está fuera de rango.

Explicación: La longitud del valor del parámetro está fuera de rango.

Lo que se ha de hacer: Especifique el parámetro utilizando una longitud válida.

DES0775N El valor de directorio de índices '*directorio*' no es correcto.

Explicación: El valor de directorio de índice no es correcto, puede que la longitud del directorio sea incorrecta.

Lo que se ha de hacer: Especifique un valor de directorio de índices válido.

DES0776N El nombre de espacio de tabla '*espaciotabla*' no es correcto.

Explicación: El nombre de espacio de tabla especificado no es correcto, puede que la

longitud del valor de espacio de tabla sea incorrecta.

Lo que se ha de hacer: Especifique un valor de espacio de tabla válido.

DES0777N El sistema de gestión de bases de datos no conoce el espacio de tabla *espaciotabla*.

Explicación: El sistema de bases de datos no conoce el espacio de tabla especificado.

Lo que se ha de hacer: Compruebe que el espacio de tabla especificado existe en la base de datos.

DES0778N '*espaciotabla*' no es un espacio de tabla REGULAR. Se ha creado con la palabra clave '*palabra clave*'.

Explicación: El tipo de datos para el espacio de tabla especificado no está soportado.

Lo que se ha de hacer: Especifique un espacio de tabla regular.

DES0779I La indexación se ha iniciado satisfactoriamente. Para comprobar el estado de la indexación utilice 'GET INDEX STATUS'.

Explicación: El programa de indexación se ha iniciado. Puede utilizar el mandato 'GET INDEX STATUS' para comprobar el estado del proceso de indexación.

Lo que se ha de hacer: Compruebe la salida del mandato GET INDEX STATUS.

DES0780I La reorganización del índice se ha iniciado satisfactoriamente. Para comprobar el estado del índice utilice 'GET INDEX STATUS'.

Explicación: La reorganización del programa ha iniciado. Puede utilizar el mandato 'GET INDEX STATUS' para comprobar el estado del proceso de reorganización.

Lo que se ha de hacer: Compruebe la salida del

mandato GET INDEX STATUS.

DES0789W Aviso: La correlación de partición del grupo de nodos actual se ha actualizado, por favor llame al programa de utilidad TXNCHECK.

Explicación: La correlación de partición del grupo de nodos actual se ha actualizado.

Lo que se ha de hacer: Utilice el mandato TXNCHECK.

DES0800I El mandato *mandato* se ha completado satisfactoriamente.

Explicación: El mandato especificado se ha completado satisfactoriamente.

Lo que se ha de hacer: No es necesaria ninguna acción.

DES0810N Falta un cerrar comillas.

Explicación: Se han encontrado unas comillas, pero faltan las segundas comillas.

Lo que se ha de hacer: Compruebe la sintaxis del mandato e inténtelo de nuevo.

DES0811N "*señal*" es inesperada. Compruebe el índice de características o la información del texto.

Explicación: El índice de características o la información del texto es incorrecto.

Lo que se ha de hacer: Compruebe la sintaxis y vuelva a intentarlo.

DES0812N La tabla *esquema.tabla* no existe o no está habilitada para DB2 Text Extender.

Explicación: Al ejecutar el mandato GET, el nombre de una tabla de bases de datos es incorrecto o la tabla no existe o no se ha habilitado todavía.

Lo que se ha de hacer: Si el nombre de tabla es correcto, utilice GET STATUS para comprobar

Mensajes de DB2 Text Extender

que se ha habilitado. Habilite la tabla y vuelva a intentarlo.

DES0813N La tabla *esquema.tabla* no existe o no está habilitada para DB2 Text Extender o no contiene una columna de descriptores de contexto *columna*.

Explicación: Al ejecutar el mandato GET, no se han encontrado entradas para la columna de descriptores de contexto en la tabla. Si la tabla existe, no está habilitada o no contiene ninguna columna de descriptores de contexto.

Lo que se ha de hacer: Si el nombre de tabla es correcto, utilice GET STATUS para comprobar que se ha habilitado. Habilite la tabla y vuelva a intentarlo.

DES0814N La tabla *nombretabla* no existe o no está habilitada para DB2 Text Extender o no hay ninguna columna de texto habilitada en esta tabla.

Explicación: La tabla especificada no está habilitada para DB2 Text Extender.

Lo que se ha de hacer: Habilite la tabla para DB2 Text Extender.

DES0815N Se han encontrado unas comillas vacías *''*. Se esperaba un nombre dentro de las comillas.

Explicación: Se han encontrado dos comillas consecutivas sin texto entre ellas.

Lo que se ha de hacer: Compruebe la sintaxis y vuelva a intentarlo.

DES0816N La palabra *"señal"* es inesperada. Utilice una de las palabras clave *palabra clave* o *palabra clave*.

Explicación: Se ha encontrado una señal inesperada.

Lo que se ha de hacer: En el mandato, utilice una de las palabras clave proporcionadas en el mensaje.

DES0817N *"señal"* es inesperada. Utilice la palabra clave *palabra clave*.

Explicación: Se ha encontrado una señal inesperada.

Lo que se ha de hacer: En el mandato, utilice la palabra clave proporcionada en el mensaje.

DES0818N Fin de mandato no esperado. Se espera la palabra clave *palabra clave*.

Explicación: Falta una palabra clave.

Lo que se ha de hacer: En el mandato, utilice la palabra clave proporcionada en el mensaje.

DES0819N Fin de mandato no esperado. Se esperaba una de las palabras clave siguientes: *palabra clave* o *palabra clave*.

Explicación: Falta una palabra clave.

Lo que se ha de hacer: En el mandato, utilice una de las palabras clave proporcionadas en el mensaje.

DES0820N La opción de índice *opción índice* no está soportada para el tipo de índice *tipo índice*.

Explicación: Ha especificado una opción de índice que no está soportada para el tipo de índice dado.

Lo que se ha de hacer: Consulte la documentación para las opciones de índice soportadas para el tipo de índice dado.

DES0821N El nombre *"señal"* es demasiado largo. Sólo están permitidos *nn* caracteres para nombres de *variable*.

Explicación: Un nombre es demasiado largo.

Lo que se ha de hacer: Especifique un nombre que tenga una longitud aceptable.

DES0822N El mandato contiene una señal no reconocida "*señal*". Se espera el fin del mandato.

Explicación: Se ha encontrado el fin del mandato, pero se espera una palabra clave.

Lo que se ha de hacer: Compruebe la sintaxis del mandato e inténtelo de nuevo.

DES0823N Se espera un nombre de tabla después de "*esquema*".

Explicación: Falta un nombre de tabla o de función después del ".".

Lo que se ha de hacer: Compruebe la sintaxis del mandato e inténtelo de nuevo.

DES0824N Fin de mandato no esperado; se necesita una *palabra clave*.

Explicación: Falta la palabra clave del mensaje en la sintaxis.

Lo que se ha de hacer: Compruebe la sintaxis del mandato e inténtelo de nuevo.

DES0825N '*alias*' no es un nombre de alias de base de datos conocido. Sólo se permiten *nn* caracteres.

Explicación: El nombre de alias de base de datos especificado no se conoce en el sistema de bases de datos.

Lo que se ha de hacer: Compruebe que el nombre de alias especificado sea válido.

DES0826N El alias de base de datos *alias* no debe ir entrecomillado.

Explicación: El nombre del mensaje se ha interpretado como un alias de base de datos. No debe ir entrecomillado.

Lo que se ha de hacer: Compruebe la sintaxis del mandato e inténtelo de nuevo.

DES0827N El CCSID "*ccsid*" no está soportado.

Explicación: El CCSID no es uno de los que DB2 Text Extender soporta.

Lo que se ha de hacer: Consulte la documentación para ver una lista de CCSID soportados.

DES0829N El nombre de usuario *idusuario* no debe estar entrecomillado.

Explicación: Ha entrado un nombre de usuario entrecomillado.

Lo que se ha de hacer: Suprima las comillas.

DES0830N El parámetro "*parámetro*" del mandato *enable/disable DATABASE* no se reconoce. Se espera el fin del mandato.

Explicación: Los mandatos ENABLE DATABASE y DISABLE DATABASE no toman parámetros.

Lo que se ha de hacer: Entre el mandato de nuevo sin parámetros.

DES0831N Fin de mandato no esperado. Falta el nombre de tabla.

Explicación: El mandato necesita un nombre de tabla.

Lo que se ha de hacer: Entre el nombre de tabla adecuado.

DES0832N Fin del mandato no esperado. Falta el nombre de base de datos.

Explicación: El mandato necesita un nombre de base de datos.

Lo que se ha de hacer: Entre el nombre de base de datos adecuado.

Mensajes de DB2 Text Extender

DES0833N **Fin del mandato no esperado.
Falta el nombre de columna.**

Explicación: El mandato necesita un nombre de columna.

Lo que se ha de hacer: Entre el nombre de columna adecuado.

DES0899N **Mandato DB2TX desconocido:
mandato.**

Explicación: El mandato especificado no está soportado por DB2 Text Extender

Lo que se ha de hacer: Escriba `db2tx ?` para obtener una lista de mandatos.

DES0971N **El directorio de índices puede especificarse una vez sin la especificación de nodo o varias veces con la especificación de nodo.**

Explicación: La especificación del directorio/s de índices no es correcta.

Lo que se ha de hacer: Compruebe la especificación del directorio de índices. Puede especificar un directorio de índices sin una especificación de nodo o múltiples directorios de índices con la especificación de nodo.

DES0972N **La especificación de nodo no es correcta. Se espera un valor de dígito sin signo.**

Explicación: Se ha especificado un valor que no es un dígito para el número de nodo.

Lo que se ha de hacer: Especifique un valor de dígito sin signo para el número de nodo.

DES0973N **La especificación de nodo no es correcta. Se han encontrado uno o varios caracteres antes del cerrar paréntesis.**

Explicación: La especificación de nodo es sintácticamente incorrecta.

Lo que se ha de hacer: Compruebe la sintaxis

de la especificación de nodo y vuelva a intentarlo.

DES0974N **La secuencia de los números de nodo de una cláusula TO no es correcta (el segundo nodo menor que el primero).**

Explicación: La especificación de nodo es sintácticamente incorrecta.

Lo que se ha de hacer: Compruebe la sintaxis de la especificación de nodo y vuelva a intentarlo.

DES0975N **La sintaxis para la especificación de la información de nodo no es correcta.**

Explicación: La especificación de nodo es sintácticamente incorrecta.

Lo que se ha de hacer: Compruebe la sintaxis de la especificación de nodo y vuelva a intentarlo.

DES0976N **La información de nodo está incompleta.**

Explicación: La especificación de nodo está incompleta - falta parte de la información.

Lo que se ha de hacer: Compruebe la sintaxis de la especificación de nodo y vuelva a intentarlo.

DES0977N **La información de nodo está incompleta. Falta el abrir paréntesis.**

Explicación: La especificación de nodo está incompleta - falta el abrir paréntesis.

Lo que se ha de hacer: Corrija la especificación de nodo y vuelva a intentarlo.

DES0998N **La longitud del nombre de modelo de documento es incorrecta.**

Explicación: La longitud del valor 'nombre de modelo de documento' es incorrecta.

Lo que se ha de hacer: Compruebe el valor de nombre de modelo y vuelva a intentarlo.

DES0999N **La sintaxis para la especificación del modelo o modelos de documentos no es correcta.**

Explicación: La especificación del nombre o nombres de modelos es sintácticamente incorrecta.

Lo que se ha de hacer: Compruebe la sintaxis de la especificación del modelo o modelos y vuelva a intentarlo.

DES9994N **Se ha producido un error de la aplicación de búsqueda de texto. Código de razón: código-razón**

Explicación: La aplicación de búsqueda de texto utilizada por DB2 Text Extender ha generado un error.

Lo que se ha de hacer: Compruebe el código de razón de la aplicación de búsqueda en el “Capítulo 19. Códigos de razón de la aplicación de búsqueda” en la página 317. Si la razón informada no ayuda a resolver el problema, puede haber información adicional en el archivo desdiag.log o en un archivo de rastreo creado. Si no sirve de ayuda, reúna la información disponible y llame al servicio técnico de IBM.

DES9995N **Se ha producido un error de Text Extender. Texto del mensaje: texto-mensaje**

Explicación: Se ha producido un error de DB2 Text Extender.

Lo que se ha de hacer: Utilice el mensaje proporcionado por DB2 Text Extender para resolver el problema. Si el mensaje informado no ayuda a resolver el problema, puede haber información adicional en el archivo desdiag.log o en un archivo de rastreo creado. Si no sirve de ayuda, reúna la información disponible y llame al servicio técnico de IBM.

DES9996N **Se ha producido un error interno de DB2 Text Extender. Código de razón: código_razón**

Explicación: Se ha producido un error de proceso interno.

Lo que se ha de hacer: Compruebe que la instalación de DB2 Text Extender se haya completado satisfactoriamente. Si es así, anote el código de razón y llame al servicio técnico de IBM.

DES9997N **Se ha producido un error SQL. SqlState: estado Código de error QL: rc; Mensaje de error SQL: mensaje**

Explicación: Se ha producido un error SQL.

Lo que se ha de hacer: Realice la acción del mensaje de error SQL que se visualiza con el mensaje.

DES9998N **Se ha producido un error SQL. No hay más información disponible.**

DES9999N **No hay ningún mensaje de error correspondiente.**

Explicación: Se ha producido un error de proceso interno.

Lo que se ha de hacer: Compruebe el mensaje de diagnóstico para resolver el problema. Si ningún problema de instalación provoca el error interno, puede haber información adicional en el archivo desdiag.log o en un archivo de rastreo creado. Si no sirve de ayuda, reúna la información disponible y llame al servicio técnico de IBM.

Capítulo 19. Códigos de razón de la aplicación de búsqueda

Este capítulo lista los códigos de razón que devuelve la aplicación de búsqueda de DB2 Text Extender.

Tabla 21. Códigos de razón de la aplicación de búsqueda

Códigos de razón	Valores
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

Códigos de razón de la aplicación de búsqueda

Tabla 21. Códigos de razón de la aplicación de búsqueda (continuación)

Códigos de razón	Valores
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

Códigos de razón de la aplicación de búsqueda

Tabla 21. Códigos de razón de la aplicación de búsqueda (continuación)

Códigos de razón	Valores
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

Códigos de razón de la aplicación de búsqueda

Capítulo 20. Códigos de razón de sucesos de error

Este capítulo lista los sucesos de error que pueden producirse cuando DB2 Text Extender indexa documentos. Esto puede producirse, por ejemplo, cuando:

- No se pueden indexar los documentos
- Los documentos están indexados, pero se produce un problema
- No se puede encontrar un diccionario de idioma.

Consejo

Si un código de razón no está documentado:

1. Compruebe que hay suficiente espacio de disco.
2. Reúna toda la información de error que está disponible:
 - Archivo desdiag.log
 - Mensaje de suceso
3. Llame al servicio técnico de IBM.

- 1 No hay almacenamiento. El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

116

Error de sintaxis de corriente de datos

280

El documento no se ha indexado. No se ha podido abrir uno de los archivos de índice.

281

No se ha indexado el documento. No se ha podido leer uno de los archivos de índice.

441

No se ha indexado el documento. Este mensaje sólo aparece para índices Ngram. La página de códigos del documento es diferente de la página de códigos con la que se ha creado el índice. Esto puede ocurrir para documentos HTML y XML si no se ha creado el índice en UTF8.

500

No se ha indexado el documento. No se han podido cargar los Servicios de biblioteca. Compruebe que el DLL esté disponible y que la vía de acceso del recurso sea válida.

Códigos de razón de sucesos de error

501

No se ha indexado el documento. Ha fallado Lib_Init de los Servicios de biblioteca En sistemas de archivos planos: no se ha encontrado el archivo DIT o no es un directorio válido o el contenido de DIT no es correcto.

502

No se ha indexado el documento. Se ha producido un error al leer el contenido del documento en el servicio de biblioteca LIB_read_doc_content.

503

No se ha indexado el documento. Se ha producido un error en el servicio de biblioteca LIB_access_doc.

504

No se ha indexado el documento. El servicio de biblioteca LIB_doc_index_status ha devuelto un error.

505

Ha fallado el cierre del documento. El servicio de biblioteca LIB_close_doc ha devuelto un error.

506

Ha fallado la finalización de los Servicios de biblioteca El servicio de biblioteca LIB_end ha devuelto un error.

507

Ha fallado la llamada al servicio de biblioteca LIB_read_doc_content con un código de retorno no esperado.

508

La llamada al servicio de biblioteca LIB_close_doc ha devuelto un error RC_TERMINATION.

545

No se ha indexado el documento. No se ha podido abrir uno de los archivos de índice temporal.

546

No se ha indexado el documento. No se ha podido cerrar uno de los archivos de índice temporal.

548

Error interno. Envíe la información del registro cronológico de diagnóstico al representante de IBM.

549, 550

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

551-564

No se ha indexado el documento. No se ha podido abrir, leer, grabar o

cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

565

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

566-587

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice.

588-590

Error interno. Envíe la información del registro cronológico de diagnóstico al representante de IBM.

591-604

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice.

605

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

606-623

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

624

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

625-631

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

632

Uno de los archivos temporales creados durante la indexación no se ha podido abrir con acceso de grabación. Compruebe los derechos de acceso.

633

Uno de los archivos temporales creados durante la indexación no se ha podido cerrar.

Códigos de razón de sucesos de error

634

Uno de los archivos temporales creados durante la indexación no se ha podido grabar. Compruebe que el directorio de trabajo del índice tiene suficiente espacio de disco.

635

Uno de los archivos temporales creados durante la indexación no se ha podido leer.

636

Uno de los archivos temporales creados durante la indexación no se ha podido abrir con acceso de lectura. Compruebe los derechos de acceso.

659

Uno de los archivos temporales creados durante la indexación no se ha podido abrir.

660

Uno de los archivos temporales creados durante la indexación no se ha podido grabar.

661

Uno de los archivos temporales creados durante la indexación no se ha podido cerrar.

662

Uno de los archivos temporales creados durante la indexación no se ha podido abrir.

663

Uno de los archivos temporales creados durante la indexación no se ha podido grabar.

664

Uno de los archivos temporales creados durante la indexación no se ha podido cerrar.

665

Uno de los archivos temporales creados durante la indexación no se ha podido abrir.

667

Uno de los archivos temporales creados durante la indexación no se ha podido grabar.

668-669

No se ha indexado el documento. Se ha producido un problema de comparación con los códigos de sección encontrados en el documento con respecto a los definidos en el archivo de modelos de documentos.

670-672

No se ha indexado el documento. No se ha podido abrir, leer, grabar o

cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

673

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo.

674

Error interno, envíe la información del registro cronológico de diagnóstico al representante de IBM.

675-687

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

688, 690

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo. Intente utilizar valores inferiores en el archivo de configuración.

689

Error interno. Envíe la información del registro cronológico de diagnóstico al representante de IBM.

691-695

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

696-707

Error interno, envíe la información del registro cronológico de diagnóstico al representante de IBM.

708

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo. Intente utilizar valores inferiores en el archivo de configuración.

709-718

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

719-721

Error interno, envíe la información del registro cronológico de diagnóstico al representante de IBM.

Códigos de razón de sucesos de error

722, 729

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo. Intente utilizar valores inferiores en el archivo de configuración.

730, 732, 733, 735-738

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco para el índice y que los derechos de acceso son correctos.

731, 739-742, 744-746, 749, 755-758, 760-761, 767

Error interno, envíe la información del registro cronológico de diagnóstico al representante de IBM.

743, 748, 750-754, 759, 765-766, 768-770

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

747, 763, 764

No hay almacenamiento (alloc ha fallado). El servidor se ha quedado sin memoria. Reduzca la carga de trabajo. Intente utilizar valores inferiores en el archivo de configuración.

815

Hay dos causas posibles:

- No se ha encontrado uno de los archivos de recursos necesario para soportar el idioma utilizado para el documento que provoca la anomalía.
- El idioma solicitado por ese documento no está soportado por DB2 Text Extender

831

No se ha indexado el documento. No se ha encontrado texto. La longitud del documento es de 0 bytes.

860

Error al abrir el archivo. No se han podido encontrar algunos diccionarios o los archivos de tesoro. Compruebe la vía de acceso de recursos para los archivos de diccionario. Si ha especificado información de vía de acceso para los archivos de tesoro durante la búsqueda, compruebe la ubicación y el nombre de archivo.

954-956

Error interno. Envíe la información del registro cronológico de diagnóstico al representante de IBM.

957-967

No se ha indexado el documento. No se ha podido abrir, leer, grabar o cerrar uno de los archivos de índice. Compruebe que hay suficiente espacio en el disco utilizado para el índice y que los derechos de acceso son correctos.

1000

Se ha producido un error durante la apertura del archivo. Compruebe los derechos de acceso.

1001

Se ha producido un error durante la adición al archivo. Compruebe los derechos de acceso.

1002

Se ha producido un error durante la lectura del archivo. Puede que el archivo esté dañado.

1003

Se ha producido un error durante la grabación del archivo. Compruebe el espacio de disco y los derechos de acceso.

1005

Se ha producido un error durante la lectura del archivo (posicionamiento en el archivo). Puede que el archivo esté dañado.

1006

Se ha producido un error durante el cambio de nombre del archivo temporal. Compruebe los derechos de acceso.

1007

Se ha producido un error durante la creación de un archivo. Compruebe los derechos de acceso.

1008

Se ha producido un error durante la compresión de un archivo. Compruebe los derechos de acceso.

1009

Se ha producido un error durante el cierre del archivo. Puede que el archivo esté dañado.

1010

El nombre de índice especificado ya se está utilizando. Utilice otro nombre de índice.

1011

La vía de acceso especificada ya se está utilizando. Utilice otra ubicación.

1012

Se utiliza la misma vía de acceso para el directorio de datos y de trabajo. Utilice otra ubicación.

Códigos de razón de sucesos de error

1013

El nombre de índice especificado no es válido. Los nombres de índice deben estar en mayúsculas o ser dígitos y no pueden tener más de 8 caracteres.

1014

Se ha producido un error durante la copia del archivo. Compruebe los derechos de acceso y el espacio de disco.

1017

El nombre de índice es desconocido. Compruebe si se ha escrito correctamente.

1019

Se ha producido un error durante la supresión del archivo. Compruebe los derechos de acceso. Este mensaje de error puede aparecer como un "error secundario" - consulte el archivo de diagnósticos para ver si una entrada de error anterior proporciona más información.

1020

Error general de archivo. Compruebe los derechos de acceso.

1070-1074

No se ha indexado el documento. La página de códigos especificada no es válida en general o no es válida para el índice que se está accediendo.

1085

No se ha indexado el documento. Se ha producido un error al leer la cola de índices.

1086

No se ha indexado el documento. La cola de índices está vacía.

1116-1117

No se ha indexado el documento. No se ha podido procesar la información del archivo de inicialización de la instancia del servidor. Asegúrese de que las entradas del archivo de inicialización son válidas y que la aplicación puede acceder al archivo.

1129

No se ha indexado ningún documento. Ha fallado el inicio del proceso de fondo.

1158

Se ha producido un error al cambiar el nombre de un archivo. Compruebe los derechos de acceso y el espacio de disco.

1162

Los archivos de un índice Ngram pueden estar dañados.

1163, 1164

El documento no se ha indexado debido a un error no esperado.

1165

El documento no se ha indexado debido a una condición de fin de archivo no esperado.

1176

No se pueden indexar más documentos para el índice Ngram. Hay una condición de desbordamiento para los números de documento (desbordamiento de longitud). Si han habido muchas supresiones o actualizaciones repetidas del mismo documento, intente llamar a EhwrReorg para resolver el problema. Si no, considere la posibilidad de utilizar un segundo índice para los nuevos documentos.

1177

No se ha indexado el documento. El indexador Ngram lo ha considerado demasiado grande.

1189

No se ha indexado el documento. Ha habido un problema con la secuencia de límites (específico del idioma coreano).

1198-1200

No se ha indexado el documento. Ha habido un problema con el acceso de índice. Puede que el índice esté dañado.

1201

No se ha indexado el documento. No se ha podido convertir la página de códigos del documento a la página de códigos específica del índice. Este error sólo es para los índices Ngram en la página de códigos UTF8.

1202

No se ha indexado el documento. No se ha podido convertir la página de códigos del documento a la página de códigos específica del índice debido a que hay datos no válidos en el documento. Este error sólo es para los índices Ngram en la página de códigos UTF8.

1500-1505

El componente de análisis de documentos tiene problemas. Puede que no se haya inicializado (compruebe LIBPATH y el contenido del archivo de configuración) o ha fallado debido a problemas internos. Consulte el archivo de diagnósticos para obtener más información.

1904

No se ha indexado el documento. Existe un problema para acceder al modelo de documento para un índice con secciones habilitadas. Compruebe los derechos de acceso y la existencia del archivo.

2000

No se ha indexado el documento. El tipo de documento no está soportado. El servicio de biblioteca Lib_access_doc ha devuelto un tipo de documento no válido.

Códigos de razón de sucesos de error

2001

No se ha indexado el documento. Se ha detectado una secuencia incorrecta de campos en la corriente de datos del documento.

2002

No se ha indexado el documento. Se ha detectado un campo estructurado incorrectamente en la corriente de datos del documento.

2003

No se ha indexado el documento. Sólo se permite una sección de texto para un documento en el formato de texto de DB2 Text Extender.

2005

No se ha indexado el documento. Un idioma especificado en la corriente de datos del documento no está soportado.

2006

No se ha indexado el documento. Un CCSID especificado en la corriente de datos del documento no está soportado.

2007

El formato de documento esperado proporcionado por la biblioteca o por la norma por omisión no es correcto. La cabecera del documento es incorrecta para el formato. Compruebe si la norma por omisión es un documento con una cabecera de documento especial y cambie si la norma no es correcta.

2008

No se ha indexado el documento porque no se ha podido acceder al mismo.

2009

No se ha indexado el documento porque se estaba utilizando y no se ha podido acceder al mismo.

2010

No se ha indexado el documento. El CCSID especificado no es correcto.

2011

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. El fin de página debe ser el último control en el cuerpo de texto del documento.

2012

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. Un campo estructurado contiene una especificación incorrecta de longitud.

2013

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. Se ha detectado un control incorrecto en el documento.

2014

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. Se ha detectado un control de múltiples bytes o un campo estructurado incorrecto en el documento.

2015

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. Se han encontrado parámetros duplicados en el documento.

2016

No se ha indexado el documento porque no es un documento IBM DCA RFT o FFT válido. Se ha encontrado una unidad de texto vacía.

2018

El documento está en un formato no soportado o existe una entrada "excluye" en el DIT para la extensión del documento. Compruebe que el documento tenga una extensión que le permita indexarlo.

2020

No se ha indexado el documento. No es un documento WordPerfect ni un archivo WordPerfect.

2021

No se ha indexado el documento. Es un archivo WordPerfect pero no un documento WordPerfect.

2022

No se ha indexado el documento. Esta versión de WordPerfect no está soportada.

2023

No se ha indexado el documento. Se trata de un archivo WordPerfect cifrado. Almacene el documento sin cifrarlo.

2026

Se ha producido END_TXT en una nota al pie o una nota final. Compruebe el archivo WordPerfect, puede que esté dañado.

2028

El analizador ha devuelto un texto que no es de documento. Compruebe el contenido del archivo, especialmente con respecto a las palabras específicas de formato. Compruebe si el formato del documento está soportado. Si falla el reconocimiento automático del formato, asegúrese de que se llama al analizador correcto.

2030

No se ha indexado el documento. Se trata de un archivo que no es Microsoft Word o que su versión de Word no está soportada.

Códigos de razón de sucesos de error

2031

No se ha indexado el documento. Se ha detectado un fin de archivo no esperado en un documento Microsoft Word.

2032

No se ha indexado el documento. Se ha detectado un control incorrecto en un documento Microsoft Word.

2033

No se ha indexado el documento. Se ha guardado en formato *complex* con la opción *fastsave*. Guárdelo con la opción *fastsave* desactivada.

2034

No se ha indexado el documento. Falta una marca de fin de archivo obligatoria en un documento Microsoft Word.

2035

El documento está cifrado. Almacene el documento en Microsoft Word sin cifrar.

2036

Se trata de un documento Word para Macintosh; no se puede procesar. Almacene el documento en formato Word para Windows.

2037

Este documento Word contiene objetos OLE incorporados.

2040

El documento no se ha indexado porque no es un archivo ECTF válido.

2041

No se ha indexado el documento. Contiene un control .SO LEN que no va seguido de un número.

2042

No se ha indexado el documento. Contiene un control .SO LEN que va seguido de un número incorrecto. El número debe estar entre 1 y 79.

2043

No se ha indexado el documento. Sólo está permitido un control .SO DOC. Guarde cada documento ECTF en un archivo separado.

2044

No se ha indexado el documento. Un control .SO HDE debe ir seguido de códigos de inicio y de fin.

2046

No se ha indexado el documento. El documento contiene texto antes del control .SO DOC.

2047

No se ha indexado el documento. El documento contiene texto antes de un control .SO PID.

2048

No se ha indexado el documento. Falta un código de fin después de un código de inicio.

2050

No se ha indexado el documento. Se han detectado códigos incorrectos después de un control .SO HDE.

2051

No se ha indexado el documento. Se ha detectado un fin de la línea después de un control .SO.

2052

No se ha indexado el documento. Se ha detectado un fin de texto no esperado.

2060

No se ha indexado el documento. No es un documento AmiPro o su versión de AmiPro no está soportada.

2061

No se ha indexado el documento. La longitud de un control de un documento AmiPro es demasiado larga.

2062

No se ha indexado el documento. Esta versión de AmiPro no está soportada. Sólo se soporta AmiPro Architecture Versión 4.

2063

AmiPro Style Sheets no se han indexado.

2064

No se ha indexado el documento. Se ha detectado un juego de caracteres incorrecto. Sólo se soporta el juego de caracteres Lotus 82 (Windows ANSI).

2065

No se ha indexado el documento. Se ha detectado un fin de archivo no esperado en un documento AmiPro.

2072

No se puede explorar el documento porque está cifrado.

2073

El formato del documento no es coherente.

2074

El documento tiene establecido el bit de distintivo “archivo incorrecto”.

Códigos de razón de sucesos de error

2080

No se ha indexado el documento. No es un documento RTF o es de una versión de RTF que no está soportada.

2081

No se ha indexado el documento. Se ha detectado una palabra de control RTF que es demasiado larga.

2083

No se ha indexado el documento. La página de códigos Macintosh no está soportada.

2084

No se ha indexado el documento. Se trata de un documento RTF, pero no se soporta esta versión de RTF. Sólo se soporta RTF Versión 1.

2090

No se ha indexado el documento. Se trata de un documento HTML, que contiene un código considerado demasiado largo por el analizador.

2093

No se ha indexado el documento. Se trata de un documento XML, que se ha rechazado por el analizador XML.

2100

El documento está dañado o no se puede leer por otra razón. Un nuevo analizador común podría corregir el problema.

2101

El documento no puede indexarse porque está vacío o no contiene texto. Compruebe si el documento sólo contiene gráficos.

2102

El documento no puede indexarse porque está protegido con contraseña o está cifrado.

2105

El tipo de documento es conocido, pero el filtro no está disponible.

2106

El documento no puede indexarse porque está vacío.

2107

El documento no puede indexarse porque no puede abrirse. Compruebe el acceso de documento.

2112

El documento no puede indexarse porque es un archivo ejecutable.

2113

El documento no puede indexarse porque está comprimido.

2114

El documento no puede indexarse porque es un gráfico. Si el formato del documento gráfico devuelve un fragmento aceptable de texto, pida que se incluya este formato de documento en el proceso de indexación.

2120

El archivo de salida de la salida del usuario no existe o no se puede acceder. Una nueva versión de analizador común podría corregir el problema.

2121

No se puede abrir el archivo de salida para lectura o está vacío. Una nueva versión de analizador común podría corregir el problema.

2122

Se intenta utilizar un archivo de salida de usuario, pero no se ha proporcionado ni establecido ningún nombre de archivo en el objeto.

2130

No se ha podido ejecutar ningún programa de salida de usuario. Compruebe si se puede encontrar el ejecutable en la vía de acceso establecida por la variable de entorno PATH. Cree un rastreo y un vuelco para obtener información adicional acerca de los códigos de retorno (errno) de entorno.

2131

El programa de salida de usuario ha fallado con un código de retorno de error. Cree un rastreo y un vuelco para obtener información adicional acerca de los códigos de retorno (errno) de entorno.

Parte 3. Apéndices

Notas

Puede que IBM no comercialice los productos, servicios o características que se explican en este documento en todos los países. Solicite al representante local de IBM información sobre los productos y servicios disponibles actualmente en el área. Las referencias a un producto, programa o servicio de IBM no indican ni implican que sólo pueda utilizarse ese producto, programa o servicio de IBM. En su lugar, puede utilizarse cualquier producto, programa o servicio funcionalmente equivalente que no infrinja el derecho de propiedad intelectual de IBM. Sin embargo, es responsabilidad del usuario el evaluar y verificar el funcionamiento de cualquier producto, programa o servicio que no sea de IBM.

IBM puede tener aplicaciones patentadas o pendientes de patente que cubran un tema descrito en este documento. Con este documento no se otorga la licencia de ninguna de estas patentes. Puede remitir la solicitud de licencia, por escrito, a:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
Estados Unidos

Para la solicitud de licencias relativas a la información de doble byte (DBCS), póngase en contacto con el departamento de propiedad intelectual de IBM en su país o remita las solicitudes, por escrito, a:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japón

El siguiente párrafo no se aplica al Reino Unido ni a ningún otro país en el que tales disposiciones sean incoherentes con la legislación local:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN "TAL CUAL" SIN GARANTÍA DE NINGUNA CLASE, NI EXPRESA NI IMPLÍCITA, INCLUYENDO, PERO SIN LIMITARSE A, LAS GARANTÍAS IMPLÍCITAS DEL NO INCUMPLIMIENTO, COMERCIALIZACIÓN O IDONEIDAD PARA UN PROPÓSITO DETERMINADO. Algunos países no permiten la declaración de limitación de responsabilidad de garantías expresas ni implícitas para determinadas transacciones, por lo tanto, puede que esta declaración no se aplique en su caso.

Esta información podría incluir inexactitudes técnicas o errores tipográficos. Se efectúan cambios periódicos en la información que contiene; estos cambios se incorporarán a nuevas ediciones de la publicación. IBM puede realizar mejoras y/o cambios en el producto o productos y/o programa o programas descritos en esta publicación en cualquier momento sin previo aviso.

Cualquier referencia de esta información a sitios Web que no son de IBM se efectúa solamente por conveniencia y no son de ninguna manera una aprobación de estos sitios. Los materiales de esos sitios Web no forman parte de los materiales de este producto de IBM y el riesgo de la utilización de estos sitios Web la asume el usuario.

IBM puede utilizar o distribuir cualquier información que suministre en la manera que encuentre adecuada sin incurrir en ninguna obligación con el remitente.

Las personas con licencia de este programa que deseen información acerca del mismo para permitir: (i) el intercambio de información entre programas creados independientemente y otros programas (incluyendo este) y (ii) el uso mutuo de la información que se ha intercambiado, deben ponerse en contacto con:

IBM Canada Limited
Office of the Lab Director
1150 Eglinton Ave. East
North York, Ontario
M3C 1H7
CANADÁ

Es posible que esta información esté disponible, sujeta a los términos y condiciones adecuadas, incluyendo en algunos casos el pago de una cuota.

El programa bajo licencia descrito en esta información y todo el material bajo licencia disponible para el mismo se proporcionan por IBM bajo los términos de IBM Customer Agreement, IBM International Program License Agreement u otro contrato equivalente entre las partes.

Cualquier dato de rendimiento que contenga esta publicación se ha determinado en un entorno controlado. Por lo tanto, los resultados obtenidos en otros entornos operativos puede variar significativamente. Algunas mediciones se pueden haber realizado en un sistema en nivel de desarrollo y no existen garantías de que sean iguales en los sistemas disponibles en general. Además, algunas mediciones pueden haberse estimado mediante extrapolación. Los resultados reales pueden variar. Los usuarios de este documento deben verificar los datos aplicables para su entorno específico.

La información relativa a los productos que no son de IBM se han obtenido de los distribuidores de estos productos, los anuncios publicados u otras fuentes públicamente disponibles. IBM no ha probado esos productos y no puede confirmar la exactitud del rendimiento, la compatibilidad ni otras afirmaciones relacionadas con productos que no son IBM. Las preguntas sobre las posibilidades de los productos que no son IBM deben dirigirse a los distribuidores de esos productos.

Todas las declaraciones relativas a las orientaciones o intenciones futuras de IBM están sujetas a cambios o a su abandono sin previo aviso y sólo representan metas y objetivos.

Esta información puede contener ejemplos de datos e informes utilizados en las operaciones comerciales diarias. Para ilustrarlas lo mejor posible, los ejemplos incluyen nombres de personas, compañías, marcas y productos. Todos estos nombres son ficticios y cualquier similitud a nombres y direcciones utilizadas por empresas reales es pura coincidencia.

LICENCIA DE COPYRIGHT:

Esta información puede contener programas de aplicación de ejemplo en lenguaje fuente, que ilustren las técnicas de programación en varias plataformas operativas. Puede copiar, modificar y distribuir estos programas de ejemplo en cualquier forma sin ningún pago a IBM, para desarrollo, utilización, comercialización o distribución de programas de aplicación que conforman la interfaz de programas de aplicación para la plataforma operativa para la que se han escrito los programas de ejemplo. Estos ejemplos no se han probado en profundidad bajo todas las condiciones. Por lo tanto, IBM no puede garantizar ni suponer la fiabilidad, el servicio o la función de estos programas.

Cada copia o cualquier porción de estos programas de ejemplo o cualquier trabajo derivado debe incluir el aviso de copyright siguiente:

© (el nombre de su compañía) (año). Partes de este código se derivan de programas de ejemplo de IBM Corp. © Copyright IBM Corp. _entre el año o años_. Reservados todos los derechos.

Marcas registradas

Los términos siguientes son marcas registradas de International Business Machines Corporation en Estados Unidos y/o en otros países.

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	RISC System/6000
DataRefresher	RS/6000
DB2	S/370
DB2 Connect	SP
DB2 Extenders	SQL/400
DB2 OLAP Server	System/370
DB2 Universal Database	System/390
Distributed Relational Database Architecture	SystemView
DRDA	VisualAge
Extended Services	VM/ESA
FFST	VSE/ESA
First Failure Support Technology	VTAM
	WebExplorer
	WIN-OS/2

Los términos siguientes son nombres comerciales o marcas registradas de otras compañías:

Microsoft, Windows, Windows NT y el logotipo de Windows son marcas registradas de Microsoft Corporation en Estados Unidos y/o en otros países.

Java y los nombres comerciales basados en Java son marcas registradas de Sun Microsystems, Inc. en Estados Unidos y/o en otros países.

UNIX es una marca registrada de Open Group en Estados Unidos y en otros países.

Puede haber otros nombres de compañías, productos o servicios que sean marcas registradas o de servicio de otras empresas.

Glosario

Este glosario define muchos de los términos y las abreviaturas utilizados en este manual. Si no encuentra el término que busca, consulte el índice o el *Dictionary of Computing*, New York: McGraw-Hill, 1994.

A

ampliar. Acción de añadir a un término de búsqueda términos adicionales derivados de un tesoro.

aparición. Sinónimo de *coincidencia*.

API. Interfaz de programas de aplicación.

archivo externo. Documento de búsqueda en el formato de un archivo almacenado en el sistema de archivos del sistema operativo, en lugar de en el formato de una celda de una tabla bajo el control de DB2.

argumento de búsqueda. Condiciones especificadas cuando se realiza una búsqueda que constan de uno o varios términos de búsqueda y parámetros de búsqueda.

B

base de datos particionada. Base de datos que consta de varias partes, cada una de las cuales se mantiene por un servidor separado de particiones de bases de datos.

búsqueda booleana. Búsqueda en la que se combinan uno o varios términos de búsqueda utilizando operadores booleanos.

búsqueda de características. Búsqueda de términos como, por ejemplo, nombres de personas, lugares u organizaciones, realizada en un índice lingüístico creado utilizando la opción de indexación `FEATURE_EXTRACTION`.

búsqueda de límites. Búsqueda en documentos en coreano que respeta los límites de las palabras.

búsqueda de texto libre. Búsqueda en la cual el término de búsqueda se expresa como texto de formato libre – una frase o una sentencia que describe en idioma natural el tema que se ha de buscar.

búsqueda híbrida. Una *búsqueda booleana* y una *búsqueda de texto libre* combinadas.

búsqueda similar. Búsqueda que puede encontrar palabras que se escriben de manera similar al término de búsqueda.

C

carácter comodín. Ver *carácter de máscara*.

carácter de escape. Carácter que indica que el carácter subsiguiente no se interpreta como un *carácter de máscara*.

carácter de máscara. Carácter utilizado para representar caracteres opcionales al principio, en medio o al final de un término de búsqueda. Normalmente, los caracteres de máscara se utilizan para buscar variaciones de un término en un índice exacto.

características de índice. Propiedades de un *índice de texto* determinando:

El directorio en el que está almacenado el índice

El tipo de índice

La frecuencia con que se actualiza el índice

Cuándo se ha de producir la primera actualización del índice.

CCSID. Identificador de juego de caracteres codificado.

coincidencia. Aparición de un término de búsqueda en un documento de texto.

columna de texto. Columna de texto que contiene *documentos de texto*.

comparación ampliada. Proceso que implica la utilización de un *diccionario* para resaltar los términos que no son coincidencias obvias del término de búsqueda.

configuración del texto. Valores por omisión para los valores de índice, texto y proceso.

corriente de datos. Información devuelta por una función API, que comprende el texto (un párrafo como mínimo) que contiene el término que se busca y la información para resaltar el término encontrado en ese texto.

cuenta. Palabra clave utilizada para especificar el número de niveles (la profundidad) de los términos en el tesoro que se han de utilizar para ampliar el término de búsqueda para una relación dada.

D

DBCS. Soporte de caracteres de doble byte.

DB2 Extender. Programa del grupo de programas que le permite almacenar y recuperar tipos de datos más allá de los datos tradicionales numéricos y de caracteres como, por ejemplo, datos de imagen, de audio y de vídeo y documentos complejos.

descriptor de contexto. Valor binario que identifica un documento de texto. Incluye:

- Un ID de documento

- El nombre y ubicación del índice asociado

- La *información de texto* del documento

- Si el documento está ubicado en un archivo externo que no está bajo el control de DB2, la vía de acceso y el nombre del archivo.

Se crea un descriptor de contexto para cada documento de texto de una columna de texto cuando esa columna está *habilitada* para que DB2 Text Extender la utilice.

descriptor de contexto de archivo. Ver *descriptor de contexto*.

descriptor de contexto de documento. Ver *descriptor de contexto*.

descriptor de contexto inicializado. Un *descriptor de contexto*, preparado con anterioridad, que sólo contiene el formato de texto o el idioma del texto o ambos.

desencadenante. Mecanismo que añade automáticamente información acerca de los documentos que es necesario indexar en una *tabla de registro cronológico* siempre que se añade, cambia o suprime un documento de una columna de texto.

diccionario. Conjunto de información lingüística relacionada con el idioma que DB2 Text Extender utiliza durante el análisis de texto, la indexación, la recuperación y el resaltado de documentos en un idioma en particular.

documento. Ver *documento de texto*.

documento de texto. Texto de tipo CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR o CLOB, almacenado en una tabla DB2.

E

examinador. Una función de DB2 Text Extender que le permite visualizar texto en un monitor del sistema.

examinar. Ver el texto visualizado en un monitor del sistema.

expansor. Ver *DB2 Extender*.

F

formato. Tipo de un documento como, por ejemplo, ASCII o WordPerfect.

frecuencia de actualización. Frecuencia con la que se actualiza un índice de texto, expresada en función del día, hora y minuto y el número

mínimo de nombres de documentos que deben listarse en la *tabla de registro cronológico* para la indexación, antes de que tenga lugar la misma.

función. Ver *función de acceso*.

función de acceso. Función proporcionada por el usuario que convierte el tipo de los datos de texto almacenados en una columna en un tipo que DB2 Text Extender puede procesar.

función definida por el usuario (UDF).

Función SQL creada por un usuario de DB2, en contraposición a la función SQL proporcionada por DB2. DB2 Text Extender proporciona funciones de administración y de búsqueda como, por ejemplo, CONTAINS, en el formato de UDF.

grupo de nodos. Subconjunto con nombre de uno o varios servidores de particiones de bases de datos. *Nodo* asignado a una máquina separada físicamente. Ver también *nodo lógico*.

H

habilitar. Preparar una base de datos , una tabla de texto o una columna de texto para que DB2 Text Extender las utilice.

I

idioma. Nombre de un *diccionario* que se ha de utilizar al *indexar*, buscar y *examinar*.

indexación periódica. Indexación a intervalos de tiempo predeterminados, expresados en función del día, hora y minuto y el número mínimo de nombres de documentos que deben listarse en la *tabla de registro cronológico* para indexación, antes de que ésta tenga lugar.

indexar. Extraer los términos significativos del texto y almacenarlos en un *índice de texto*.

índice de texto. Colección de términos significativos extraídos de documentos de texto. Cada término se asocia al documento del que se ha extraído. Se consigue una mejora significativa en el tiempo de búsqueda buscando en el índice

en lugar de en los documentos en sí. Ver también *índice exacto* e *índice lingüístico*.

índice exacto. *Índice de texto* que contiene los términos exactamente tal como aparecen en el documento de texto del que se han extraído. Ver también *índice lingüístico* e *índice Ngram*.

índice lingüístico. *Índice de texto* que contiene términos que se han reducido a su formato base por el proceso lingüístico. “Ratones”, por ejemplo, se indexaría como “ratón”. Ver también *índice exacto* e *índice Ngram*.

índice Ngram. *Índice de texto* que soporta documentos DBCS y la búsqueda similar de documentos SBCS. Ver también *índice lingüístico* e *índice exacto*.

información de resaltado. Ver *corriente de datos*.

información de texto. Propiedades de un *documento de texto* que describen:

El *CCSID*

El *formato*

El *idioma*.

inhabilitar. Restaurar una base de datos , una tabla de texto o una columna de texto a la condición previa en la que estaba antes de que DB2 Text Extender la habilite eliminando los elementos creados durante el proceso de habilitación.

instancia. Entorno lógico de DB2 Text Extender. Pueden haber varias instancias de DB2 Text Extender en la misma estación de trabajo, pero sólo una instancia para cada instancia de DB2. Puede utilizar estas instancias para:

Separar el entorno de desarrollo del entorno de producción.

Restringir la información sensible a un grupo en particular de personas.

interfaz de programas de aplicación (API).

Interfaz de propósito general entre programas de aplicación y los servicios de recuperación de información de DB2 Text Extender.

L

M

modelo de documento. Definición de la estructura de un documento en función de las secciones que contiene. Un modelo de documento hace que DB2 Text Extender conozca las secciones de los documentos al indexarlos. Un modelo de documento lista los códigos de marcación que identifican las secciones. Para cada código puede especificar un nombre de sección descriptivo para utilizarlo en consultas en esa sección. Puede especificar uno o varios modelos de documento en un archivo de modelos de documentos.

N

nodo. Servidor de un entorno de *bases de datos particionadas*. Ver también *nodo lógico*, *nodo físico* y *grupo de nodos*.

nodo físico. *Nodo* asignado a una máquina separada físicamente. Ver también *nodo lógico*.

nodo lógico. *Nodo* asignado con otros nodos a la misma máquina física. Ver también *nodo físico*.

O

P

página de códigos. Asignación de caracteres gráficos y significados de funciones de control para todos los elementos de código. Por ejemplo, la asignación de los caracteres y los significados para los elementos de códigos 256 para un código de 8 bits.

palabra de parada. Palabra común como, por ejemplo, “antes”, de un *documento de texto* que se ha de excluir del *índice de texto* y se pasa por alto si se incluye en un *argumento de búsqueda*.

perfil. Ver *perfil de entorno*.

perfil de entorno. Script proporcionado con DB2 Text Extender que contiene los valores para las *variables de entorno*.

perfil de shell. Ver *perfil de entorno*.

procesador de línea de mandatos. Programa llamado DB2TX que:

- Permite al usuario entrar mandatos DB2 Text Extender

- Procesa los mandatos

- Visualiza el resultado.

R

rango. Valor absoluto de tipo DOUBLE entre 0 y 1 que indica el grado con qué un documento coincide con los criterios de búsqueda en relación a los demás documentos encontrados. El valor indica el número de coincidencias encontradas en el documento en relación al tamaño del documento.

rastreo. Acción de almacenamiento de información en un archivo que puede utilizarse luego para buscar la causa de un error.

recuperar. Buscar un documento de texto utilizando un argumento de búsqueda en una de las funciones de búsqueda de DB2 Text Extender.

redefinir. Añadir los criterios de búsqueda de una búsqueda anterior a otros criterios de búsqueda para reducir el número de *coincidencias*.

S

SBCS. Soporte de caracteres de un solo byte.

T

tabla de índice común. Tabla DB2 cuyas columnas de texto comparten un índice de texto común. Ver también *tabla de múltiples índices*.

tabla de múltiples índices. Tabla DB2 cuyas columnas de texto tienen *índices de texto* individuales. Ver también *tabla de índice común*.

tabla de registro cronológico. Tabla creada por DB2 Text Extender que contiene información acerca de los documentos de texto que se han de indexar. Se utilizan *desencadenantes* para almacenar esta información en una tabla de registro cronológico siempre que se añade, cambia o suprime un documento de una columna de texto habilitada.

tabla de texto. Tabla DB2 que contiene *columnas de texto*.

tipo de índice. Característica de un *índice de texto* que determina si contiene los formatos exacto o lingüístico de los términos del documento. Ver *índice exacto*, *índice lingüístico* e *índice Ngram*.

tipo diferenciado. Ver *tipo diferenciado definido por el usuario*.

tipo diferenciado definido por el usuario (UDT). Tipo de datos creado por un usuario de DB2, en contraposición al tipo de datos proporcionado por DB2 como, por ejemplo, LONG VARCHAR.

U

UDF. Función definida por el usuario.

UDT. Tipo diferenciado definido por el usuario.

variable de entorno. Variable utilizada para proporcionar valores por omisión para los valores del entorno de DB2 Text Extender.

variable de instancia. Variable utilizada para proporcionar un valor por omisión para el nombre del propietario de la *instancia* o el nombre del directorio inicial del propietario de la instancia.

vista del catálogo. Vista de una tabla del sistema creada por DB2 Text Extender con finalidad de administración. Una vista del catálogo contiene información acerca de las tablas y las columnas que se han habilitado para que DB2 Text Extender las utilice.

W

Índice

A

abreviaturas

- edición de un archivo de abreviaturas 55
- listas de 54

abrir documento, función API

- descripción 248
- utilización 121

actualización de un índice de texto

- cambio de la frecuencia 101
- periódicamente 46
- UPDATEFREQ en CHANGE INDEX SETTINGS 130
- UPDATEFREQ en CHANGE TEXT CONFIGURATION 132

administración

- actualización de un índice inmediatamente 99
- actualización de un índice para archivos externos 100
- archivo de abreviaturas, edición 55
- archivo de palabras de parada, modificación 55
- base de datos de ejemplo, programa de utilidad para crear 16
- cambio de configuración del texto 20
- cambio de valores de índice 101
- compilación de archivo de definiciones de tesauro 186
- compilación de un archivo de definición de tesauro Ngram 188
- creación de una base de datos de ejemplo 195
- creación de una instancia de DB2 Text Extender 21
- DELETE INDEX EVENTS 137
- desechar una instancia de DB2 Text Extender 21
- detención del servidor DB2 Text Extender 185
- hacer copia de seguridad y restaurar 22
- información de estado, obtención 104

administración (continuación)

- inhabilitación de archivos de texto 114
- inhabilitación de una base de datos 115
- inhabilitación de una columna de texto 113
- inhabilitación de una tabla de texto 115
- inicio del servidor DB2 Text Extender 183
- mandato CONNECT 135
- mandato CHANGE INDEX SETTINGS 129
- mandato CHANGE TEXT CONFIGURATION 131
- mandato DB2TX 128
- mandato DELETE INDEX EVENTS 137
- mandato DISABLE DATABASE 138
- mandato DISABLE TEXT COLUMN 139
- mandato DISABLE TEXT FILES 140
- mandato DISABLE TEXT TABLE 141
- mandato GET ENVIRONMENT 160
- mandato GET INDEX SETTINGS 161
- mandato GET INDEX STATUS 163
- mandato GET STATUS 164
- mandato GET TEXT CONFIGURATION 165
- mandato GET TEXT INFO 166
- mandato QUIT 167
- mandato REORGANIZE INDEX 168
- mandato RESET INDEX STATUS 169
- mandato TXICRT 174
- mandato TXIDROP 176
- mandato TXILIST 177
- mandato TXIPCLEAN 178
- mandato TXNADD 179
- mandato TXNCHECK 180
- mandato TXNDROP 181

administración (continuación)

- mandato TXSAMPLE 182
- mandato TXSTART 183
- mandato TXSTATUS 184
- mandato TXSTOP 185
- mandato TXTHESC 186
- mandato TXTHESN 188
- mandato TXTRACE 190
- mandato TXVERIFY 195
- mandato TXWIZARD 196
- mandato UPDATE INDEX 170
- mantenimiento de los índices de texto 99
- modificación de archivos de palabras de parada y de abreviaturas 55
- procesador de línea de mandatos 128
- rastreo de anomalías 23
- reorganización de un índice 103
- restablecimiento del estado del índice 102
- resumen de mandatos, cliente 127
- resumen de mandatos, servidor 173
- supresión de sucesos de índice 102
- variables de entorno 17
- visualización de estado de índice 106
- visualización de los valores de índice 109
- visualización de los valores de información de texto 110
- visualización del estado de bases de datos, tablas y columnas 104
- visualización del estado del servidor 22
- AmiPro, formato de documento 31
- ampliación de términos para resultado 266
- análisis de texto
 - para examen 266
 - para indexación 256

- análisis de texto básico
 - de términos que contienen caracteres no alfanuméricos 256
 - normalización 257
 - para indexar términos 256
 - para resaltado 266
 - reconocimiento de sentencias 258
- AND
 - operador booleano 85
 - palabra clave en argumento de búsqueda 222
- apariciones de un término de búsqueda 210
- archivo de cabecera des_ext.h 117
- archivo de cabecera DES_EXT.H 117
- archivo de modelos de documentos, contenido 52
- archivo include des_ext.h 117
- archivos de configuración 281
- archivos externos
 - cambio de la vía de acceso/nombre en el descriptor de contexto 97
 - consideraciones de la actualización de índices 100
 - descriptores de contexto para 81
 - extracción de vía de acceso/nombre de un descriptor de contexto 96
 - función FILE 202
 - habilitación 73
 - inhabilitación 114
 - inicialización de descriptores de contexto 206
 - mandato DISABLE TEXT FILES 140
 - mandato ENABLE TEXT FILES 152
 - obtención o cambio de un nombre de archivo en un descriptor de contexto 202
- argumento de búsqueda
 - búsqueda con & y | 85
 - búsqueda con NOT 90
 - búsqueda de fragmentos de un término 86
 - búsqueda de límites 91
 - búsqueda de palabras con sonido similar 92
 - búsqueda de sinónimos 89
 - búsqueda de términos en cualquier orden 88

- argumento de búsqueda (*continuación*)
 - búsqueda de términos en el mismo párrafo 88
 - búsqueda de términos en la misma sentencia 88
 - búsqueda de términos en secciones de documentos 88
 - búsqueda de términos en varios idiomas 89
 - búsqueda de texto libre 93
 - búsqueda de variaciones de un término 86
 - búsqueda de varios términos 85
 - búsqueda híbrida 93
 - búsqueda similar 91, 223
 - búsquedas en tesaurus 92
 - descripción 215
 - especificación 85
 - IN SAME PARAGRAPH AS 222
 - IN SAME SENTENCE AS 222
 - operador & (AND) 221
 - operador | (OR) 221
 - palabra clave AND 222
 - palabra clave BOUND 223
 - palabra clave COUNT 221
 - palabra clave EXPAND 221
 - palabra clave FUZZY FORM OF 223
 - palabra clave MODEL 219
 - palabra clave NOT 222
 - palabra clave PRECISE FORM OF 223
 - palabra clave RESULT LIMIT 221
 - palabra clave SECTION 219
 - palabra clave STEMMED FORM OF 223
 - palabra clave SYNONYM FORM OF 223
 - palabra clave TERM OF 221
 - palabra clave THESAURUS 220
 - resumen de las normas y las restricciones 226
 - sintaxis 217
 - utilización de caracteres de máscara 86
 - valor de atributo 220
- argumento de búsqueda booleano 221
- ASCCI plano, formato de documento 31
- ASCII, formato de documento 31
- asistente de ejemplo 27
- ayuda para mandatos 60

B

- base de datos
 - conexión a 60
 - habilitación 61
 - hacer copia de seguridad y restaurar 22
 - información de estado, visualización 104
 - inhabilitación 115
 - mandato CONNECT 135
 - mandato DISABLE DATABASE 138
 - mandato ENABLE DATABASE 142
 - mandato GET STATUS 164
- base de datos de ejemplo
 - creación (TXVERIFY) 16
- búsqueda de anomalías 23
- búsqueda de límites, ejemplo 91
- búsqueda de texto
 - creación de una lista de documentos encontrados 205
 - ejemplo de programa 253
 - función REFINE 212
 - función SEARCH_RESULT 213
 - mejora del rendimiento 97
 - obtención del número de coincidencias encontradas 84
 - obtención del rango de un documento encontrado 84
 - realización de una consulta 83
 - redefinición de una búsqueda previa 93
 - sintaxis 217
 - utilización de la API 119
 - visión general 82
- búsqueda de texto libre
 - ejemplo 93
- búsqueda exacta 43
- búsqueda híbrida, ejemplo 93
- búsqueda similar 41
- búsqueda similar, ejemplo 91
- búsquedas en tesaurus
 - compilación de archivo de definiciones de tesaurus 186
 - compilación de un archivo de definición de tesaurus Ngram 188
 - conceptos 268
 - creación de un tesaurus 273
 - creación de un tesaurus Ngram 278
 - ejemplo 92
 - mandato TXTHESC 186
 - mandato TXTHESN 188

búsquedas en tesoro (*continuación*)
palabra clave THESAURUS 220
sintaxis 220

C

carácter de escape
sintaxis 226
utilización 87

caracteres comodín
en un término de búsqueda 86
enmascaramiento de
palabras 265

características de índice
en ENABLE TEXT
COLUMN 143
en ENABLE TEXT FILES 152
en ENABLE TEXT TABLE 155
valores por omisión en los
valores de la configuración del
texto 19
visualización 109

características del proceso
valores por omisión en los
valores de la configuración del
texto 20

características del texto
CCSID 35
en ENABLE TEXT
COLUMN 143
en ENABLE TEXT FILES 152
formato 31
idioma 35
valores por omisión en la
configuración del texto 19

CCSID
descripción 35
en CHANGE TEXT
CONFIGURATION 133
en ENABLE TEXT
COLUMN 145
evitar problemas de página de
códigos 37
extracción de un descriptor de
contexto 96
función 200
inicialización en descriptores de
contexto 206
lista de 35
mandato GET TEXT INFO 166
valor por omisión en la
configuración del texto 19

cerrar documento, función API
descripción 231
utilización 122

cliente/servidor, entorno 4

códigos de razón de la aplicación de
búsqueda 317

códigos de razón de sucesos 321

códigos de retorno 291

coincidencia
de DesGetSearchResultTable 119
en un resultado de búsqueda 84
función DesGetMatches 237
función
NUMBER_OF_MATCHES 210

columna
habilitación 66
habilitación de una tabla
grande 70
habilitación de varios tipos de
índice 69
inhabilitación 113
mandato DISABLE TEXT
COLUMN 139
mandato ENABLE TEXT
COLUMN 143

comparación ampliada 266

compilación de archivo de
definiciones de tesoro 186

compilación de un archivo de
definición de tesoro Ngram 188

conceptos de paralelismo 7

conexión a una base de datos
cómo 60
mandato CONNECT 135

configuración 17

creación de una base de datos de
ejemplo 195

creación de una instancia de DB2
Text Extender
mandato TXICRT 174

creación de una tabla de ejemplo
mandato TXSAMPLE 182

D

DB2 Extenders
ejemplo de uso 4

DB2DBDFT, variable de entorno 17

DB2INSTANCE, variable de
entorno 17

DB2TX, procesador de línea de
mandatos
sintaxis 128
utilización 59

DB2TX_ variables de entorno
descripción 17
visualización 105

DB2TX_INSTOWNER, variable de
entorno 17

DB2TX_INSTOWNERHOMEDIR,
variable de entorno 18

DES_BROWSE, opción en
DesGetSearchResultTable 244

DES_EXTENDED, opción en
DesOpenDocument 248

DES_FAST, opción en
DesOpenDocument 248

DES_MATCH, opción en
DesGetSearchResultTable 244

DES_NOBROWSE, opción en
DesGetSearchResultTable 244

DES_RANK, opción en
DesGetSearchResultTable 244

DES_RANKANDMATCH, opción en
DesGetSearchResultTable 244

DES_TEXTHANDLEONLY, opción
en DesGetSearchResultTable 244

DESK.INI 281

descomposición de términos
compuestos 263

descriptor de contexto
cambio de formato y de
idioma 97
descripción 80
establecimiento y extracción de
información en 95
extracción de CCSID, formato e
idioma 96
función CCSID 200
función FORMAT 203
función LANGUAGE 208
inicialización 95
obtención de una lista de
descriptores de contexto 204
para archivos externos 81
tipo diferenciado
DB2TEXTFH 197
tipo diferenciado
DB2TEXTH 197
utilización de listas para mejorar
el rendimiento 97

desechar una instancia
cómo 21
mandato TXIDROP 176

desencadenantes
creación 68
descripción 30

DESMODEL.INI 48

DESRESTB, para crear una tabla de
resultados 245

DESSAMP1, programa de
ejemplo 253

DESSRV.INI 283

- detención del servidor DB2 Text Extender 185
 - directorio para índices
 - mandato GET INDEX SETTINGS 161
 - documento
 - CCSID 35
 - conversión de formato 33
 - conversión de tipos de datos 71
 - estructura 48
 - formato, descripción 31
 - formato en CHANGE TEXT CONFIGURATION 133
 - formato en ENABLE TEXT COLUMN 146
 - formatos soportados 31
 - idioma 35
 - indexación 29
 - información acerca de 110
 - mandato GET TEXT INFO 166
 - tipos de datos soportados 144
 - visualización de los valores 110
 - documentos DBCS, búsqueda en 43
 - documentos de archivo plano, soporte de secciones 51
 - documentos estructurados HTML 48
 - documentos estructurados XML 48
 - documentos HTML, soporte de secciones 51
 - documentos XML, soporte de secciones 53
- E**
- ejemplo de asistente
 - descripción 27
 - mandato TXWIZARD 196
 - ejemplo de GUI 27
 - ejemplo de interfaz 27
 - enmascaramiento de caracteres 265
 - entorno de cliente/servidor 4
 - entorno de múltiples nodos 48
 - espacio de disco para índices 45
 - espacio de registro cronológico, falta de 70
 - espacio de tabla 68
 - espacios de tabla y grupos de nodos 68
 - estado de actualización, restablecimiento
 - ejemplo 102
 - sintaxis 169
 - estado de actualización, visualización
 - ejemplo y salida 106
 - sintaxis 163
 - estado de búsqueda, restablecimiento
 - ejemplo 102
 - sintaxis 169
 - estado de búsqueda, visualización
 - ejemplo y salida 106
 - sintaxis 163
 - estado de índice, restablecimiento
 - ejemplo 102
 - sintaxis 169
 - estado de índice, visualización
 - ejemplo y salida 106
 - sintaxis 163
 - estado de un índice
 - restablecimiento 102
 - restablecimiento después de un error 102
 - visualización 106
 - visualización del estado actual 163
 - estado del servidor de DB2 Text Extender 184
 - estructura de documentos
 - ejemplo de búsqueda 88
 - habilitación de soporte de secciones 48
 - palabra clave MODEL en la sintaxis de búsqueda 219
 - palabra clave SECTION en la sintaxis de búsqueda 219
 - valor de atributo en la sintaxis de búsqueda 220
 - examen
 - ejemplo de programa 253
 - proceso lingüístico para 266
 - utilización de examinador propio 120
 - expansión de sonidos
 - descripción 265
 - ejemplo 92
 - Extenders
 - ejemplo de uso 4
- F**
- FFT, formato de documento 31
 - finalizar examen de sesión, función API
 - descripción 232
 - utilización 123
 - formato base, reducción de términos a 262
 - formato de documentos de texto 31
 - cambio en el descriptor de contexto 97
 - conversión de no soportado 33
 - descripción 31
 - formato de documentos de texto 31 (*continuación*)
 - en CHANGE TEXT CONFIGURATION 133
 - en ENABLE TEXT COLUMN 146
 - extracción de un descriptor de contexto 96
 - FORMAT, palabra clave 133
 - función FORMAT 203
 - inicialización en descriptores de contexto 206
 - lista de soportados 31
 - mandato GET TEXT INFO 166
 - palabra clave FORMAT 146
 - valor por omisión en la configuración del texto 19
 - formatos de enteros, soporte de secciones 51
 - formatos de fecha, soporte de secciones 49
 - formatos de hora, soporte de secciones 50
 - formatos flotantes soportados 50
 - frecuencia de la actualización
 - cambio 101
 - descripción 46
 - mandato GET INDEX SETTINGS 161
 - sintaxis 46
 - UPDATEFREQ en CHANGE INDEX SETTINGS 130
 - UPDATEFREQ en CHANGE TEXT CONFIGURATION 132
 - valor por omisión en valores de configuración del texto 19
 - función
 - establecimiento de la vía de acceso para las funciones de DB2 Text Extender 82
 - funciones API 117
 - funciones de búsqueda 77
 - para convertir tipos de datos 71
 - sentencia SET CURRENT FUNCTION PATH 82
 - función CONTAINS
 - ejemplo 83
 - sintaxis 201
 - función de acceso
 - descripción 71
 - en ENABLE TEXT COLUMN 144
 - función DesCloseDocument
 - descripción 231
 - utilización 122

función DesEndBrowseSession
 descripción 232
 utilización 123

función DesFreeBrowseInfo
 descripción 233
 utilización 123

función DesGetBrowseInfo
 descripción 234
 utilización 120

función DesGetMatches
 descripción 237
 utilización 122

función DesGetSearchResultTable
 descripción 243
 utilización 119

función DesOpenDocument
 descripción 248
 utilización 121

función DesStartBrowseSession
 descripción 251
 utilización 121

función FILE
 ejemplo 96
 sintaxis 202

función FORMAT
 ejemplo 96
 sintaxis 203

función HANDLE
 sintaxis 204
 utilización 97

función HANDLE_LIST
 sintaxis 205
 utilización 97

función INIT_TEXT_HANDLE
 ejemplo 95
 sintaxis 206

función LANGUAGE
 ejemplo 96
 sintaxis 208

función NO_OF_DOCUMENTS
 sintaxis 209

función NUMBER_OF_MATCHES,
 sintaxis 210

función RANK
 ejemplo 84
 sintaxis 211

función REFINE
 ejemplo 93
 sintaxis 212

función SEARCH_RESULT
 ejemplo 97
 sintaxis 213

funciones
 búsqueda de texto 82
 CCSID 200

funciones (*continuación*)
 consulta 197
 CONTAINS 201
 descripción 77
 especificación de los argumentos
 de búsqueda 85
 establecimiento y extracción de
 información en descriptores de
 contexto 95
 FILE 202
 FORMAT 203
 HANDLE 204
 HANDLE_LIST 205
 indicaciones SQL devueltas
 por 299
 INIT_TEXT_HANDLE 206
 LANGUAGE 208
 mejora del rendimiento de
 búsqueda 97
 NO_OF_DOCUMENTS 209
 NUMBER_OF_MATCHES 210
 RANK 211
 redefinición de una búsqueda
 previa 93
 REFINE 212
 SEARCH_RESULT 213
 vía de acceso de función 82
 visión general 198

funciones de búsqueda 117

funciones de DB2 Text Extender de
 ejemplo
 ejecución 77

funciones de examen 117

G

grupo de nodos de múltiples
 particiones 7

H

hacer copia de seguridad y
 restaurar 22

HTML, formato de documento 31

I

idioma de documentos de texto
 cambio en el descriptor de
 contexto 97
 descripción 35
 en un argumento de
 búsqueda 89
 extracción de un descriptor de
 contexto 96
 función LANGUAGE 208
 inicialización en descriptores de
 contexto 206
 lista de 35

idioma de documentos de texto
 (*continuación*)
 mandato GET TEXT INFO 166
 valor por omisión en la
 configuración del texto 19

indexación, proceso lingüístico 255

indicaciones SQL devueltas por
 funciones DB2 Text Extender 299

índice
 actualización de un índice
 inmediatamente 99
 actualización para archivos
 externos 100
 actualización periódica de
 índices 46
 cálculo del tamaño 45
 cambio de configuración del
 texto 20
 cambio de la frecuencia de
 actualización 101
 cambio de los valores
 actuales 101
 cambio del tipo de índice 44
 creación de varios tipos para una
 columna de texto 69
 exacta 43
 frecuencia de la actualización 46
 hacer copia de seguridad y
 restaurar 22
 INDEXOPTION en ENABLE
 TEXT COLUMN 147
 INDEXOPTION en ENABLE
 TEXT TABLE 156
 INDEXTYPE en CHANGE TEXT
 CONFIGURATION 132
 INDEXTYPE en ENABLE TEXT
 COLUMN 147
 INDEXTYPE en ENABLE TEXT
 TABLE 156
 lingüístico 41
 mandato CHANGE INDEX
 SETTINGS 129
 mandato CHANGE TEXT
 CONFIGURATION 131
 mandato GET INDEX
 SETTINGS 161
 Mandato GET INDEX
 STATUS 163
 mandato GET TEXT
 CONFIGURATION 165
 mandato UPDATE INDEX 170
 mantenimiento 99
 múltiples, utilización 44
 Ngram 43
 opción CASE_ENABLED 44

- índice *(continuación)*
 - planificación 29
 - reorganización 103
 - tabla de índice común 44
 - TABLESPACE en CHANGE TEXT CONFIGURATION 132
 - tipo de valor por omisión en los valores de configuración del texto 19
 - tipos de 40
 - visión general 29
 - visualización de los valores actuales 109
 - índice exacto
 - descripción 43
 - valores por omisión de opciones de búsqueda 223
 - índice lingüístico
 - descripción 41
 - valores por omisión de opciones de búsqueda 223
 - índice Ngram
 - descripción 43
 - opción CASE_ENABLED 44
 - valores por omisión de opciones de búsqueda 223
 - información acerca de documentos de texto
 - CCSID 35
 - formato 31
 - idioma 35
 - mandato GET TEXT INFO 166
 - tipos de 31
 - visualización del valor actual 110
 - información de resaltado
 - corriente de datos 122
 - sintaxis de corriente de datos 237
 - iniciación 25
 - inicialización de un descriptor de contexto
 - cómo 95
 - función
 - INIT_TEXT_HANDLE 206
 - iniciar sesión de examen, función API
 - descripción 251
 - utilización 121
 - inicio del servidor DB2 Text Extender 183
 - instancias
 - creación 21
 - desechar 21
 - instancias *(continuación)*
 - mandato para la lista de, TXLIST 177
 - variables de entorno 17
 - interfaz de programación de aplicaciones (API)
 - apertura de un documento para examen 121
 - búsqueda de texto 119
 - cierre de un documento 122
 - códigos de retorno 291
 - consulta 229
 - ejemplo de programa 253
 - finalización de una sesión de examen 123
 - función DesCloseDocument 231
 - función
 - DesEndBrowseSession 232
 - función DesFreeBrowseInfo 233
 - función DesGetBrowseInfo 234
 - función DesGetMatches 237
 - función
 - DesGetSearchResultTable 243
 - función DesOpenDocument 248
 - función
 - DesStartBrowseSession 251
 - funciones de búsqueda 117
 - funciones de examen 117
 - información de resaltado 237
 - inicio de una sesión de examen 121
 - liberalización del almacenamiento de información de examen 123
 - mensajes 299
 - obtención de coincidencias 122
 - obtención de información de examen 120
 - obtención de puntero para la información de resaltado 237
 - obtención de una tabla del resultado 119
 - resumen 229
 - visión general 118
- L**
- liberar almacenamiento para información de examen, función API
 - descripción 233
 - utilización 123
- M**
- mandato CONNECT
 - sintaxis 135
 - utilización 60
 - mandato CHANGE INDEX SETTINGS
 - sintaxis 129
 - utilización 101
 - mandato CHANGE TEXT CONFIGURATION
 - sintaxis 131
 - utilización 20
 - mandato DELETE INDEX EVENTS
 - ejemplo 102
 - sintaxis 137
 - mandato DISABLE DATABASE
 - sintaxis 138
 - utilización 115
 - mandato DISABLE TEXT COLUMN
 - sintaxis 139
 - utilización 113
 - mandato DISABLE TEXT FILES
 - sintaxis 140
 - utilización 114
 - mandato DISABLE TEXT TABLE
 - sintaxis 141
 - utilización 115
 - mandato ENABLE DATABASE
 - sintaxis 142
 - utilización 61
 - mandato ENABLE TEXT COLUMN
 - sintaxis 143
 - utilización 66
 - mandato ENABLE TEXT FILES
 - sintaxis 152
 - utilización 73
 - mandato ENABLE TEXT TABLE
 - sintaxis 155
 - utilización 62
 - mandato GET ENVIRONMENT
 - ejemplo y salida 105
 - sintaxis 160
 - mandato GET INDEX SETTINGS
 - ejemplo y salida 109
 - sintaxis 161
 - mandato GET INDEX STATUS
 - ejemplo y salida 106
 - sintaxis 163
 - mandato GET STATUS
 - ejemplo y salida 104
 - sintaxis 164
 - mandato GET TEXT CONFIGURATION
 - ejemplo y salida 105
 - sintaxis 165
 - mandato GET TEXT INFO
 - ejemplo y salida 110
 - sintaxis 166

- mandato QUIT
 - sintaxis 167
 - utilización 76
 - mandato REORGANIZE INDEX
 - ejemplo 103
 - sintaxis 168
 - mandato RESET INDEX STATUS
 - ejemplo 102
 - sintaxis 169
 - mandato TXICRT
 - creación de una instancia de DB2
 - Text Extender 21
 - sintaxis 174
 - mandato TXIDROP
 - sintaxis 176
 - mandato TXILIST
 - sintaxis 177
 - mandato TXIPCLEAN
 - sintaxis 178
 - mandato TXNADD
 - sintaxis 179
 - mandato TXNCHECK
 - sintaxis 180
 - mandato TXNDROP
 - sintaxis 181
 - mandato TXSAMPLE
 - sintaxis 182
 - utilización 16
 - mandato TXSTART
 - sintaxis 183
 - utilización 22
 - mandato TXSTATUS
 - sintaxis 184
 - utilización 22
 - mandato TXSTOP
 - sintaxis 185
 - utilización 22
 - mandato TXTHESC
 - sintaxis 186
 - mandato TXTHESN
 - sintaxis 188
 - mandato TXTRACE
 - sintaxis 190
 - utilización 23
 - mandato TXVERIFY
 - sintaxis 195
 - mandato TXWIZARD
 - sintaxis 196
 - mandato UPDATE INDEX
 - ejemplo 99
 - sintaxis 170
 - mandatos
 - CONNECT 135
 - CHANGE INDEX SETTINGS 129
 - mandatos (*continuación*)
 - CHANGE TEXT
 - CONFIGURATION 131
 - DB2TX 128
 - DELETE INDEX EVENTS 137
 - DISABLE DATABASE 138
 - DISABLE TEXT COLUMN 139
 - DISABLE TEXT FILES 140
 - DISABLE TEXT TABLE 141
 - ENABLE DATABASE 142
 - ENABLE TEXT COLUMN 143
 - ENABLE TEXT FILES 152
 - ENABLE TEXT TABLE 155
 - GET ENVIRONMENT 160
 - GET INDEX SETTINGS 161
 - GET INDEX STATUS 163
 - GET STATUS 164
 - GET TEXT
 - CONFIGURATION 165
 - GET TEXT INFO 166
 - QUIT 167
 - REORGANIZE INDEX 168
 - RESET INDEX STATUS 169
 - resumen, mandatos de
 - cliente 127
 - resumen, mandatos del
 - servidor 173
 - TXICRT 174
 - TXIDROP 176
 - TXILIST 177
 - TXIPCLEAN 178
 - TXNADD 179
 - TXNCHECK 180
 - TXNDROP 181
 - TXSAMPLE 182
 - TXSTART 183
 - TXSTATUS 184
 - TXSTOP 185
 - TXTHESC 186
 - TXTHESN 188
 - TXTRACE 190
 - TXVERIFY 195
 - TXWIZARD 196
 - UPDATE INDEX 170
 - máscara
 - en un término de búsqueda 86
 - proceso lingüístico 265
 - mensajes 299
 - Microsoft, formato de documento 31
 - modelo de documento
 - descripción 48
 - modificación del archivo de modelos de documentos 48
 - modelo de documento (*continuación*)
 - palabra clave MODEL en la sintaxis de búsqueda 219
 - palabra clave SECTION en la sintaxis de búsqueda 219
 - valor de atributo en la sintaxis de búsqueda 220
 - múltiples índices, utilización 44
- ## N
- nodo
 - asignación 7
 - entorno de múltiples nodos 48
 - grupos de nodos y espacios de tabla 68
 - mandato añadir nodo
 - TXNADD 179
 - mandato comprobar nodo
 - TXNCHECK 180
 - mandato desechar nodo
 - TXNDROP 181
 - nodo físico, asignación 7
 - nodo lógico, asignación 7
 - nombres de archivos de diccionarios 54
 - normalización de términos 257
 - normas para los argumentos de búsqueda 226
 - NOT
 - operador booleano 90
 - palabra clave en argumento de búsqueda 222
- ## O
- obtener coincidencias, función API
 - descripción 237
 - utilización 122
 - obtener información de examen, función API
 - descripción 234
 - utilización 120
 - obtener tabla de resultados de búsquedas, función API
 - descripción 243
 - utilización 119
 - operador & (AND) en argumento de búsqueda
 - cómo utilizar 85
 - sintaxis del argumento de búsqueda 221
 - operador | (OR) en argumento de búsqueda
 - cómo utilizar 85
 - sintaxis del argumento de búsqueda 221
 - operador booleano OR 85

operadores booleanos
 & (AND) y | (OR) 85
 NOT 90

P

palabra clave BOUND 223
palabra clave CASE_ENABLED
 en ENABLE TEXT
 COLUMN 147
 en ENABLE TEXT TABLE 156
palabra clave COUNT 221
palabra clave DIRECTORY
 en CHANGE TEXT
 CONFIGURATION 133
 en ENABLE TEXT
 COLUMN 149
 en ENABLE TEXT TABLE 158
 valor por omisión en valores de
 configuración del texto 19
 visualización del valor
 actual 109
palabra clave EXPAND 221
palabra clave FUNCTION
 descripción 71
 en ENABLE TEXT
 COLUMN 144
palabra clave FUZZY FORM
 OF 223
palabra clave IN SAME
 PARAGRAPH AS 222
palabra clave IN SAME SENTENCE
 AS 222
palabra clave INDEXOPTION
 en CHANGE TEXT
 CONFIGURATION 132
 en ENABLE TEXT
 COLUMN 147
 en ENABLE TEXT TABLE 156
palabra clave INDEXPROPERTY
 en ENABLE TEXT
 COLUMN 148
 en ENABLE TEXT TABLE 157
palabra clave INDEXTYPE
 en CHANGE TEXT
 CONFIGURATION 132
 en ENABLE TEXT
 COLUMN 147
 en ENABLE TEXT TABLE 156
palabra clave LANGUAGE 133, 145
palabra clave NODE
 en ENABLE TEXT
 COLUMN 149
 en ENABLE TEXT TABLE 158

palabra clave NORMALIZED
 en ENABLE TEXT
 COLUMN 148
palabra clave ON NODE
 en ENABLE TEXT
 COLUMN 149
 en ENABLE TEXT TABLE 158
palabra clave PRECISE FORM
 OF 223
palabra clave RESULT LIMIT 221
palabra clave STEMMED FORM
 OF 223
palabra clave TERM OF 221
palabra clave UPDATEFREQ
 en CHANGE INDEX
 SETTINGS 130
 en CHANGE TEXT
 CONFIGURATION 132
 en ENABLE TEXT
 COLUMN 148
 en ENABLE TEXT TABLE 157
palabra clave UPDATEINDEX
 en CHANGE TEXT
 CONFIGURATION 133
 en ENABLE TEXT
 COLUMN 149
 mandato GET INDEX
 SETTINGS 161
 valor por omisión en valores de
 configuración del texto 20
 visualización del valor
 actual 109
palabras de parada
 como parte del análisis básico del
 texto 262
 descripción 29
 edición de un archivo de
 palabras de parada 55
 listas de 54
parámetro de configuración
 COMMITCOUNT
 conservación de espacio de
 registro cronológico 71
 descripción 70
 en CHANGE TEXT
 CONFIGURATION 133
 en ENABLE TEXT
 COLUMN 150
 en ENABLE TEXT TABLE 170
 valor por omisión en valores de
 configuración del texto 20
parámetros de idiomas, lista de 54
parámetros LOGPRIMARY,
 LOGSECOND y LOGFILSIZ en
 DB2 UDB 70

perfiles de shell 17
procesador de línea de mandatos
 ayuda de 60
 inicio 59
 mandato DB2TX 128
 mandato QUIT 167
proceso lingüístico
 al indexar 255
 ampliación de sonido 265
 ampliación de términos 266
 análisis básico de texto 256
 comparación ampliada 266
 descripción 255
 división de términos
 compuestos 263
 enmascaramiento de
 caracteres 265
 enmascaramiento de
 palabras 265
 filtro de palabras de parada 262
 máscara 265
 para examen 266
 para la recuperación 264
 reducción de términos a su
 formato base 262
 sinónimos 264
profundidad de términos en un
 tesauro, especificación 221
programa API de ejemplo 253
puntero de lista de descriptores de
 contexto (tipo diferenciado
 DB2TEXTFHLISTP) 198
puntero de lista de descriptores de
 contexto (tipo diferenciado
 DB2TEXTHLISTP) 198

R

rango
 de DesGetSearchResultTable 119
 en un resultado de búsqueda 84
rastreo de anomalías
 configuración 23
 mandato TXTRACE 190
reconocimiento de sentencias 258
recuperación, proceso lingüístico
 para 264
redefinición de una búsqueda
 previa 93
reducción de términos a su formato
 base 262
rendimiento, mejora 97
requisitos de espacio para
 índices 45
restricciones para los argumentos de
 búsqueda 226

S

salida de usuario, conversión de formato de documento 33

secciones en documentos

- archivo de modelos de documentos, contenido 52
- DESMODEL.INI 48
- documentos de archivo plano 51
- documentos HTML 51
- documentos XML 53
- ejemplo de búsqueda 88
- formatos de enteros soportados 51
- formatos de fecha soportados 49
- formatos de hora soportados 50
- formatos flotantes soportados 50
- habilitación de soporte de secciones 48
- palabra clave MODEL en la sintaxis de búsqueda 219
- palabra clave SECTION en la sintaxis de búsqueda 219
- secciones de atributos 49
- valor de atributo en la sintaxis de búsqueda 220

sentencia SET CURRENT FUNCTION PATH 82

separación de frases 43

separación de palabras 43

servidor

- configuración y mantenimiento 21
- detención 185
- inicio 183
- mandato TXICRT 174
- mandato TXIDROP 176
- mandato TXILIST 177
- mandato TXIPCLEAN 178
- mandato TXNADD 179
- mandato TXNCHECK 180
- mandato TXNDROP 181
- mandato TXSAMPLE 182
- mandato TXSTART 183
- mandato TXSTATUS 184
- mandato TXSTOP 185
- mandato TXTRACE 190
- mandato TXVERIFY 195
- mandato TXWIZARD 196
- rastreo de anomalías 23
- visualización del estado 184

sinónimos

- descripción 264

sinónimos (*continuación*)

- en un argumento de búsqueda 89
- palabra clave SYNONYM FORM OF 223

sintaxis de corriente de datos 238

soporte de bases de datos particionadas 7

soporte de idioma nacional 37

sucesos de error

- códigos de razón 321
- DELETE INDEX EVENTS 137
- mandato GET INDEX STATUS 163
- registro 68
- supresión 102
- visualización 108

sucesos de indexación

- códigos de razón 321

sucesos de índice, supresión

- ejemplo 102
- sintaxis 137

T

tabla DB2TX.SAMPLE

- descripción 78
- programa de utilidad para la creación 16
- supresión 116

tabla de configuración

- creación 62
- mandato CHANGE TEXT CONFIGURATION 131
- mandato GET TEXT CONFIGURATION 165
- visualización 105

tabla de ejemplo

- descripción 78
- mandato TXSAMPLE 182
- supresión 116

tabla de índice común

- creación 62
- descripción 44, 45
- mandato ENABLE TEXT TABLE 155

tabla de registro cronológico

- asignación a un espacio de tabla 68
- creación 68
- descripción 30
- extracción de sucesos de error 108

tabla de resultados 245

tabla de texto

- habilitación de una columna en una tabla grande 70
- hacer copia de seguridad y restaurar 22
- mandato DISABLE TEXT TABLE 141
- mandato ENABLE TEXT TABLE 155

tablas grandes, habilitación 70

términos compuestos, división 263

tipo de datos DATALINK 72

tipo de índice, cambio

- cambio 44
- creación de varios tipos para una columna de texto 69

tipo diferenciado DB2TEXTFH 197

tipo diferenciado DB2TEXTFHLISTP 198

tipo diferenciado DB2TEXTH 197

tipo diferenciado DB2TEXTHLISTP 198

tipos de datos de documentos de texto

- función para la conversión 71, 72
- habilitación de las columnas que contienen el tipo de datos DATALINK 72
- soportado 144

tipos de índice de texto

- exacta 43
- INDEXTYPE en CHANGE TEXT CONFIGURATION 132
- INDEXTYPE en ENABLE TEXT COLUMN 147
- INDEXTYPE en ENABLE TEXT TABLE 156
- lingüístico 41
- mandato CHANGE TEXT CONFIGURATION 132
- mandato GET INDEX SETTINGS 161
- Ngram 43
- opción CASE_ENABLED 44
- valor por omisión en valores de configuración del texto 19
- valores por omisión de opciones de búsqueda 223

tipos diferenciados 197

TXSAMPLE.UDF

- ejecución 77

TXVERIFY

- creación de una base de datos de ejemplo 16

U

UDT 197

V

valores de configuración del texto

- cambio 20
- valores por omisión de la
instalación 19
- visualización 105

variables

- descripción de variables de
entorno 17
- mandato GET
ENVIRONMENT 160
- visualización de variables de
entorno 105

variables de entorno 17

- descripción 17
- mandato GET
ENVIRONMENT 160
- visualización 105

verificación de la instalación 16

visión general de DB2 Text
Extender 3

vista del catálogo

- contenido 111
- creación 61
- supresión 116

vista del catálogo TEXTINDEXES

- contenido 111
- creación 61
- supresión 116

W

WordPerfect, formato de
documento 31

X

XML, formato de documento 31

Ponerse en contacto con IBM

Si se le presenta un problema técnico, revise y efectúe las acciones sugeridas en la *Guía para la resolución de problemas* antes de ponerse en contacto con el Soporte al cliente DB2. Esta guía sugiere información que puede reunir para ayudar a que el Soporte al cliente DB2 le preste un mejor servicio.

Para obtener información o pedir cualquier producto DB2 Universal Database póngase en contacto con un representante de IBM en una sucursal local o póngase en contacto con cualquier comercializador de software de IBM autorizado.

Si vive en EE.UU., puede llamar a uno de los números siguientes:

- 1-800-237-5511 para soporte al cliente
- 1-888-426-4343 para conocer las opciones de servicio disponibles

Información sobre el producto

Si vive en EE.UU., puede llamar a uno de los números siguientes:

- 1-800-IBM-CALL (1-800-426-2255) ó 1-800-3IBM-OS2 (1-800-342-6672) para pedir productos u obtener información general.
- 1-800-879-2755 para pedir publicaciones.

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

Las páginas de la World Wide Web de DB2 proporcionan información actual de DB2 acerca de las noticias, las descripciones de productos, las planificaciones de formación y otras.

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

DB2 Product and Service Technical Library proporciona acceso a las preguntas formuladas con frecuencia, los arreglos, los manuales e la información técnica actualizada de DB2.

Nota: Puede que esta información sólo esté en inglés.

<http://www.elink.ibm.com/pbl/pbl/>

El sitio Web de pedido de publicaciones internacionales proporciona información sobre cómo pedir los manuales.

<http://www.ibm.com/education/certify/>

Professional Certification Program del sitio Web de IBM proporciona información de la prueba de certificado para diversos productos de IBM, incluyendo DB2.

ftp.software.ibm.com

Inicie la sesión como anónimo. En el directorio /ps/products/db2, puede encontrar demostraciones, arreglos, información y herramientas relacionadas con DB2 y muchos otros productos.

comp.databases.ibm-db2, bit.listserv.db2-l

Estos grupos de noticias de Internet están disponibles para que los usuarios expliquen sus experiencias con los productos DB2.

En Compuserve: GO IBMDB2

Entre este mandato para acceder a los fórums de la familia IBM DB2. En estos fórums se soportan todos los productos DB2.

Para obtener información sobre cómo ponerse en contacto con IBM fuera de Estados Unidos, consulte el Apéndice A de la publicación *IBM Software Support Handbook*. Para acceder a este documento, vaya a la página Web siguiente: <http://www.ibm.com/support/> y seleccione el enlace IBM Software Support Handbook situado en la parte inferior de la página.

Nota: En algunos países, los concesionarios autorizados de IBM deben ponerse en contacto con su estructura de soporte a concesionarios, en lugar de con el Centro de soporte de IBM.

Hoja de Comentarios

DB2[®] Universal Database
DB2 Text Extender
Administración y programación
Versión 7

Número de Publicación SC10-3527-00

Por favor, sírvase facilitarnos su opinión sobre esta publicación, tanto a nivel general (organización, contenido, utilidad, facilidad de lectura,...) como a nivel específico (errores u omisiones concretos). Tenga en cuenta que los comentarios que nos envíe deben estar relacionados exclusivamente con la información contenida en este manual y a la forma de presentación de ésta.

Para realizar consultas técnicas o solicitar información acerca de productos y precios, por favor dirijase a su sucursal de IBM, business partner de IBM o concesionario autorizado.

Para preguntas de tipo general, llame a "IBM Responde" (número de teléfono 901 300 000).

Al enviar comentarios a IBM, se garantiza a IBM el derecho no exclusivo de utilizar o distribuir dichos comentarios en la forma que considere apropiada sin incurrir por ello en ninguna obligación con el remitente.

Comentarios:

Gracias por su colaboración.

Para enviar sus comentarios:

- Envíelos por correo a la dirección indicada en el reverso.

Si desea obtener respuesta de IBM, rellene la información siguiente:

Nombre

Dirección

Compañía

Número de teléfono

Dirección de e-mail



IBM S.A.
National Language Solutions Center
Av. Diagonal 571, Edif. "L'Illa"
08029 Barcelona
España



Printed in Denmark by IBM Danmark A/S.

SC10-3527-00



Spine information:



DB2[®] Universal Database

**DB2 Text Extender Administración y
programación**

Versión 7