

DB2 Universal Database



# DB2 Extension Texte Administration et programmation

*Version 7*



DB2 Universal Database



# DB2 Extension Texte Administration et programmation

*Version 7*

**Important**

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à la section «Remarques» à la page 339.

**Première édition - juin 2000**

LE PRESENT DOCUMENT EST LIVRE "EN L'ETAT". IBM DECLINE TOUTE RESPONSABILITE, EXPRESSE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE QUALITE MARCHANDE OU D'ADAPTATION A VOS BESOINS. Certaines juridictions n'autorisent pas l'exclusion des garanties implicites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. Les informations qui y sont fournies sont susceptibles d'être modifiées avant que les produits décrits ne deviennent eux-mêmes disponibles. En outre, il peut contenir des informations ou des références concernant certains produits, logiciels ou services non annoncés dans ce pays. Cela ne signifie cependant pas qu'ils y seront annoncés.

Pour plus de détails, pour toute demande d'ordre technique, ou pour obtenir des exemplaires de documents IBM, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial.

Vous pouvez également consulter les serveurs Internet suivants :

- <http://www.fr.ibm.com> (serveur IBM en France)
- <http://www.can.ibm.com> (serveur IBM au Canada)
- <http://www.ibm.com> (serveur IBM aux Etats-Unis)

*Compagnie IBM France  
Direction Qualité  
Tour Descartes  
92066 Paris-La Défense Cedex 50*

© Copyright IBM France 2000. Tous droits réservés.

© Copyright International Business Machines Corporation 1995, 2000. All rights reserved.

---

# Table des matières

**Avis aux lecteurs canadiens . . . . . vii**

**Avant-propos . . . . . ix**

Lecteurs concernés . . . . . ix

Structure du manuel . . . . . ix

Lecture des diagrammes de syntaxe . . . . . x

Envoi de commentaires . . . . . xi

**Résumé des modifications . . . . . xiii**

---

## **Partie 1. Guide. . . . . 1**

**Chapitre 1. Présentation de DB2 Extension  
Texte . . . . . 3**

DB2 Extension Texte en environnement  
client/serveur DB2 . . . . . 4

DB2 Extension Texte en environnement de base  
de données partitionnée . . . . . 7

**Chapitre 2. Installation, configuration et  
maintenance . . . . . 9**

Installation sur systèmes AIX, Sun-Solaris et  
HP-UX . . . . . 9

Conditions d'installation . . . . . 9

Etape 1 pour AIX : Installation des  
composants du produit. . . . . 9

Etape 1 pour Sun-Solaris : Installation des  
composants du produit . . . . . 10

Etape 1 pour HP-UX : Installation des  
composants du produit . . . . . 11

Etape 2 : Mise en place de l'instance DB2  
Extension Texte . . . . . 13

Etape 3 : Démarrage du serveur DB2  
Extension Texte . . . . . 13

Installation sous Windows 95, Windows 98,  
Windows NT et Windows 2000 . . . . . 14

Serveur . . . . . 14

Client . . . . . 14

Installation de DB2 Extensions . . . . . 14

Création et gestion d'instances . . . . . 16

Préparation d'une base de données  
exemple pour vérifier l'installation . . . . . 16

Configuration . . . . . 17

Variables d'environnement . . . . . 17

Définitions de la configuration de texte . . . 19

Modification de la configuration du texte . 20

Configuration et gestion d'un serveur DB2

Extension Texte . . . . . 21

Création d'une instance DB2 Extension

Texte . . . . . 21

Démarrage et arrêt d'un serveur DB2

Extension Texte . . . . . 22

Sauvegarde et restauration des index et  
des bases de données activées . . . . . 22

Trace des incidents. . . . . 23

**Chapitre 3. Mise en route . . . . . 25**

Exemple simple de recherche dans un texte . 25

Modèle d'assistant pour l'indexation et la  
recherche . . . . . 27

**Chapitre 4. Planification de vos besoins en  
matière de recherche . . . . . 29**

Nécessité d'indexation des documents texte . 29

Documents pris en charge . . . . . 31

Particularités des documents HTML . . . . 32

Particularités des documents XML . . . . 32

Utilisation des formats de documents non  
pris en charge . . . . . 33

Langues . . . . . 35

CCSID . . . . . 35

EBCDIC . . . . . 35

ASCII . . . . . 36

DBCS . . . . . 37

UNICODE . . . . . 37

Le moyen d'éviter des incidents de pages  
de codes lors du stockage et de l'activation  
de texte . . . . . 37

Types de recherche. . . . . 40

Recherche linguistique . . . . . 41

Recherche exacte . . . . . 43

Recherche ou recherche de correspondance  
partielle dans des documents DBCS . . . . 43

Modification du type d'index . . . . . 44

Création d'un ou de plusieurs index de texte  
pour une table . . . . . 44

Calcul de la taille d'un index . . . . . 45

Mise à jour d'un index . . . . . 46

Gestion des index dans un environnement multi-noeud . . . . .	47	Recherche et renvoi du rang d'un document texte trouvé . . . . .	85
Utilisation de documents structurés (prise en charge de sections) . . . . .	48	Définition des arguments de recherche . . . . .	86
Sections avec attribut . . . . .	49	Recherche de plusieurs termes . . . . .	86
Fichiers à plat et documents HTML . . . . .	51	Recherche à l'aide des opérateurs booléens AND et OR . . . . .	86
Documents XML . . . . .	52	Recherche des variantes d'un terme . . . . .	87
Dictionnaires, listes de mots vides, listes d'abréviations et paramètres de langue . . . . .	54	Recherche à partir de fragments de termes (caractères génériques) . . . . .	87
Modification des fichiers de mots vides et d'abréviations . . . . .	55	Recherche de termes contenant un caractère générique . . . . .	88
<b>Chapitre 5. Préparation d'un texte pour la recherche . . . . .</b>	<b>57</b>	Recherche de termes conjoints ou disjoints . . . . .	89
Préparation préalable . . . . .	57	Recherche de termes d'une même phrase ou d'un même paragraphe . . . . .	89
Démarrage de l'interpréteur de commandes DB2 Extension Texte . . . . .	59	Recherche de termes dans des sections de documents structurés . . . . .	89
Aide sur l'interpréteur de commandes . . . . .	60	Recherche de synonymes. . . . .	90
Connexion à une base de données . . . . .	60	Recherche linguistique . . . . .	90
Activation d'une base de données . . . . .	61	Recherche à l'aide de l'opérateur booléen NOT . . . . .	91
Activation d'une table de texte (facultatif) . . . . .	63	Recherche de correspondance partielle . . . . .	92
Exemples . . . . .	64	Respect de la segmentation . . . . .	93
Activation d'une colonne de texte . . . . .	66	Recherche d'homophones . . . . .	93
Création d'une colonne de descripteurs . . . . .	67	Recherche par thésaurus . . . . .	93
Définition des informations relatives aux documents . . . . .	68	Recherche sur texte libre et recherche hybride . . . . .	94
Création d'une table de journalisation . . . . .	68	Affinement d'une recherche. . . . .	95
Création d'un index . . . . .	69	Définition et extraction d'informations dans les descripteurs . . . . .	97
Exemples . . . . .	69	Définition des informations de texte lors de l'insertion d'un nouveau texte . . . . .	97
Activation d'une colonne de texte dans une table volumineuse . . . . .	70	Extraction des informations d'un descripteur . . . . .	98
Activation de colonnes de texte d'un type de données non pris en charge. . . . .	71	Modification des informations d'un descripteur . . . . .	98
Activation de colonnes contenant le type de données DATALINK . . . . .	72	Amélioration des performances de recherche . . . . .	99
Activation de fichiers de texte externes . . . . .	73	<b>Chapitre 7. Administration . . . . .</b>	<b>101</b>
Exemples . . . . .	75	Gestion des index de texte . . . . .	101
Fin de la session . . . . .	75	Mise à jour d'un index . . . . .	101
<b>Chapitre 6. Mode de recherche . . . . .</b>	<b>77</b>	Mise à jour d'un index pour y intégrer les références de fichiers externes. . . . .	102
Localisation d'exemples de syntaxe pour les fonctions de recherche . . . . .	77	Modification des paramètres d'un index . . . . .	102
Modèle de table DB2TX.SAMPLE . . . . .	78	Réinitialisation de l'état d'un index . . . . .	103
Descripteurs pour fichiers externes . . . . .	82	Suppression des événements d'un index . . . . .	104
Définition du chemin de fonction en cours . . . . .	82	Réorganisation d'un index . . . . .	105
Recherche de texte . . . . .	83	Extraction d'informations utiles . . . . .	105
Création d'une requête . . . . .	84	Affichage des informations d'état . . . . .	106
Recherche et renvoi du nombre de correspondances trouvées . . . . .	85	Affichage des paramètres des variables d'environnement . . . . .	107

Affichage des définitions de la configuration de texte . . . . .	107
Affichage de l'état d'un index. . . . .	108
Affichage des événements d'erreur . . . . .	110
Affichage des paramètres d'index . . . . .	110
Affichage des paramètres de texte pour une colonne. . . . .	111
Utilisation de la vue de catalogue de DB2	
Extension Texte . . . . .	112
Annulation de la procédure de préparation de texte . . . . .	115
Désactivation d'une colonne de texte . . . . .	115
Désactivation de fichiers texte. . . . .	116
Désactivation d'une table de texte . . . . .	116
Désactivation d'une base de données . . . . .	117

**Chapitre 8. Utilisation des fonctions API pour la recherche et l'affichage . . . . . 119**

Configuration de l'application. . . . .	119
Liaison d'une application UNIX . . . . .	119
Liaison d'une application OS/2 ou Windows. . . . .	119
Présentation des fonctions API . . . . .	120
Recherche de texte . . . . .	121
Obtention d'une table de résultats de recherche (DesGetSearchResultTable) . . . . .	121
Affichage de texte. . . . .	122
Extraction d'informations d'affichage (DesGetBrowseInfo) . . . . .	123
Démarrage d'une session d'affichage (DesStartBrowseSession) . . . . .	123
Ouverture d'un document (DesOpenDocument). . . . .	123
Extraction de correspondances (DesGetMatches) . . . . .	124
Fermeture d'un document (DesCloseDocument). . . . .	124
Arrêt d'une session d'affichage (DesEndBrowseSession). . . . .	125
Libération des informations d'affichage (DesFreeBrowseInfo) . . . . .	125

**Partie 2. Référence . . . . . 127**

**Chapitre 9. Commandes de préparation de texte et d'administration du client . . . . . 129**

CHANGE INDEX SETTINGS . . . . .	131
CHANGE TEXT CONFIGURATION . . . . .	133
CONNECT . . . . .	137
DELETE INDEX EVENTS . . . . .	139

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

**Chapitre 10. Commandes d'administration du serveur. . . . . 175**

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

**Chapitre 11. Fonctions de recherche . . . . . 199**

Les types distincts dans DB2 Extension Texte	199
Récapitulatif des fonctions DB2 Extension Texte . . . . .	201
CCSID . . . . .	202
CONTAINS. . . . .	203
FILE . . . . .	204
FORMAT . . . . .	205
HANDLE . . . . .	206
HANDLE_LIST . . . . .	207
INIT_TEXT_HANDLE . . . . .	208

LANGUAGE . . . . .	209	Etape 2 : Correspondance étendue . . . . .	266
NO_OF_DOCUMENTS . . . . .	210	Thésaurus . . . . .	268
NO_OF_MATCHES . . . . .	211	Termes . . . . .	270
RANK . . . . .	212	Relations. . . . .	270
REFINE . . . . .	213	Relations dans un thésaurus Ngram . . . . .	272
SEARCH_RESULT . . . . .	214	Création d'un thésaurus . . . . .	273
		Création d'un thésaurus Ngram . . . . .	278
<b>Chapitre 12. Syntaxe des arguments de recherche . . . . .</b>	<b>215</b>	<b>Chapitre 16. Fichiers de configuration . . . . .</b>	<b>281</b>
Argument de recherche . . . . .	217	Fichier de configuration du client . . . . .	281
<b>Chapitre 13. Fonctions API de recherche et d'affichage. . . . .</b>	<b>229</b>	Fichier de configuration du serveur. . . . .	283
DesCloseDocument . . . . .	231	Fichier de configuration d'analyse de texte . . . . .	287
DesEndBrowseSession . . . . .	232	Options communes . . . . .	287
DesFreeBrowseInfo . . . . .	233	<b>Chapitre 17. Codes retour . . . . .</b>	<b>291</b>
DesGetBrowseInfo . . . . .	234	<b>Chapitre 18. Messages . . . . .</b>	<b>299</b>
DesGetMatches . . . . .	237	SQLSTATE renvoyés par les fonctions DB2	
DesGetSearchResultTable . . . . .	243	Extension Texte . . . . .	299
DesOpenDocument . . . . .	248	Messages de DB2 Extension Texte . . . . .	303
DesStartBrowseSession . . . . .	251	<b>Chapitre 19. Codes anomalie du moteur de recherche . . . . .</b>	<b>317</b>
<b>Chapitre 14. Modèle de programme API . . . . .</b>	<b>253</b>	<b>Chapitre 20. Codes anomalie dus à des événements d'erreur . . . . .</b>	<b>321</b>
<b>Chapitre 15. Traitement linguistique des index linguistiques et exacts . . . . .</b>	<b>255</b>	<hr/>	
Analyse linguistique lors de l'indexation . . . . .	255	<b>Partie 3. Annexes . . . . .</b>	<b>337</b>
Analyse de texte de base . . . . .	256	<b>Remarques . . . . .</b>	<b>339</b>
Réduction des termes à leur forme de base (lemmisation) . . . . .	262	Marques . . . . .	342
Filtrage des mots vides . . . . .	262	<b>Glossaire . . . . .</b>	<b>345</b>
Décomposition (éclatement des termes composés) . . . . .	262	<b>Index . . . . .</b>	<b>351</b>
Analyse linguistique pour l'extraction . . . . .	263	<b>Comment prendre contact avec IBM. . . . .</b>	<b>361</b>
Synonymes . . . . .	264	Infos produit . . . . .	361
Extension de thésaurus . . . . .	265		
Extension phonétique . . . . .	265		
Masquage de caractères et de mots . . . . .	265		
Analyse linguistique pour l'affichage . . . . .	266		
Etape 1 : Normalisation et extension des termes . . . . .	266		

---

## Avis aux lecteurs canadiens

Le présent document a été traduit en France. Voici les principales différences et particularités dont vous devez tenir compte.

### Illustrations

Les illustrations sont fournies à titre d'exemple. Certaines peuvent contenir des données propres à la France.

### Terminologie

La terminologie des titres IBM peut différer d'un pays à l'autre. Reportez-vous au tableau ci-dessous, au besoin.

IBM France	IBM Canada
ingénieur commercial	représentant
agence commerciale	succursale
ingénieur technico-commercial	informaticien
inspecteur	technicien du matériel

### Claviers

Les lettres sont disposées différemment : le clavier français est de type AZERTY, et le clavier français-canadien de type QWERTY.

### OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

- les pages de codes 850 (multilingue) et 863 (français-canadien),
- le code pays 002,
- le code clavier CF.

### Nomenclature

Les touches présentées dans le tableau d'équivalence suivant sont libellées différemment selon qu'il s'agit du clavier de la France, du clavier du Canada ou du clavier des États-Unis. Reportez-vous à ce tableau pour faire

correspondre les touches françaises figurant dans le présent document aux touches de votre clavier.

<b>France</b>	<b>Canada</b>	<b>Etats-Unis</b>
 (Pos1)		Home
Fin	Fin	End
 (PgAr)		PgUp
 (PgAv)		PgDn
Inser	Inser	Ins
Suppr	Suppr	Del
Echap	Echap	Esc
Attn	Intrp	Break
Impr écran	ImpEc	PrtSc
Verr num	Num	Num Lock
Arrêt défil	Défil	Scroll Lock
 (Verr maj)	FixMaj	Caps Lock
AltGr	AltCar	Alt (à droite)

## Brevets

Il est possible qu'IBM détienne des brevets ou qu'elle ait déposé des demandes de brevets portant sur certains sujets abordés dans ce document. Le fait qu'IBM vous fournisse le présent document ne signifie pas qu'elle vous accorde un permis d'utilisation de ces brevets. Vous pouvez envoyer, par écrit, vos demandes de renseignements relatives aux permis d'utilisation au directeur général des relations commerciales d'IBM, 3600 Steeles Avenue East, Markham, Ontario, L3R 9Z7.

## Assistance téléphonique

Si vous avez besoin d'assistance ou si vous voulez commander du matériel, des logiciels et des publications IBM, contactez IBM direct au 1 800 465-1234.

---

## Avant-propos

Le présent manuel explique comment utiliser DB2 Extension Texte pour préparer et gérer une base de données DB2 pour des données de type texte. Il décrit également comment utiliser les fonctions SQL fournies par DB2 Extension Texte et les interfaces API pour accéder et manipuler ces types de données. En incorporant les fonctions de DB2 Extension Texte aux instructions SQL de votre programme et en intégrant les API vous pouvez créer des programmes d'extraction de texte souples et puissants.

Dans le présent manuel, le terme "DB2" désigne DB2 UDB.

---

### Lecteurs concernés

Le présent manuel est destiné aux administrateurs de bases de données DB2 familiarisés avec les concepts, les outils et les techniques d'administration de DB2.

Il s'adresse également aux programmeurs d'applications DB2 familiarisés avec le langage SQL et un ou plusieurs langages de programmation pouvant être utilisés pour les programmes d'application DB2.

Le présent manuel est destiné aux utilisateurs de DB2 Extension Texte. Les utilisateurs des extensions DB2 Image, Audio et Vidéo doivent consulter le manuel *DB2 Extensions Image, Audio et Vidéo - Administration et programmation*.

---

### Structure du manuel

Ce manuel est organisé comme suit :

«Partie 1. Guide»

Cette partie présente DB2 Extension Texte, en décrit l'installation et la configuration, et traite de sa planification. Elle décrit également la préparation et la gestion d'une base de données DB2 afin de pouvoir rechercher du texte.

Prenez connaissance de cette partie si vous n'êtes pas encore familiarisé avec DB2 Extension Texte et que vous souhaitez apprendre à utiliser les fonctions de DB2 Extension Texte et les API pour rechercher du texte.

«Partie 2. Référence»

Cette partie présente des informations de référence sur les fonctions DB2 Extension Texte, les API, les commandes et des informations de diagnostic telles que messages et codes.

## Avant-propos

Prenez connaissance de cette partie si vous êtes déjà familiarisé avec les concepts et les tâches DB2 Extension Texte, mais que vous avez besoin d'informations sur une fonction DB2 Extension Texte, une API, une commande, un message ou un code particulier.

---

### Lecture des diagrammes de syntaxe

Tout au long de ce manuel, les données de syntaxe sont représentées de la manière suivante :

- Les diagrammes de syntaxe se lisent de la gauche vers la droite et du haut vers le bas. Ils figurent après le chemin d'accès. Le symbole ►— indique le début d'une instruction.

Le symbole —► indique que la fin de la syntaxe figure sur la ligne suivante.

Le symbole ►— indique que le début d'une instruction figure sur la ligne précédente.

Le symbole —◄ indique la fin d'une instruction.

- Les éléments obligatoires apparaissent sur la ligne horizontale (chemin principal).

►—élément obligatoire—◄

- Les éléments facultatifs apparaissent en-dessous du chemin principal.

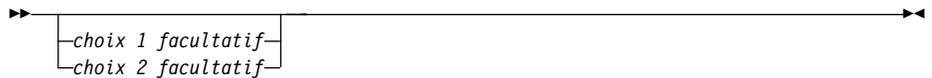
►—  
└élément facultatif┘—◄

- Si vous avez le choix entre plusieurs éléments, ceux-ci apparaissent les uns au-dessus des autres.

Si vous *devez* choisir l'un de ces éléments, un élément de la pile apparaît dans le chemin principal.

►—  
└choix 1 obligatoire┘  
└choix 2 obligatoire┘—◄

Si l'absence d'élément est admise, l'ensemble de la pile apparaît en-dessous du chemin principal.



Une flèche de répétition au-dessus d'une pile indique que vous pouvez choisir plusieurs éléments de la pile.



- Les mots clés apparaissent en majuscules. L'orthographe de ces mots clés doit être respectée. Les variables apparaissent en minuscules (par exemple, srcpath). Elles représentent les noms ou les valeurs fournis par l'utilisateur dans la syntaxe.
- Les marques de ponctuation, parenthèses, opérateurs arithmétiques et tout autre symbole présentés font également partie de la syntaxe.

---

### Envoi de commentaires

Vos commentaires nous permettent d'améliorer la qualité de nos documents. Veuillez nous faire parvenir vos commentaires sur ce manuel ou toute autre documentation relative à d'autres extensions DB2. Vous pouvez envoyer vos commentaires de l'une des manières suivante :

- Envoyez vos commentaires via le Web en utilisant le site Web suivant :

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

Ce site comporte une page de commentaires dans laquelle vous pouvez indiquer vos observations.

- Envoyez vos remarques par courrier électronique à l'adresse [swsdid@de.ibm.com](mailto:swsdid@de.ibm.com). Veillez à mentionner le nom et la référence du manuel, la version du produit et, le cas échéant, l'emplacement précis du texte concerné (par exemple, un numéro de page ou de tableau).

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie de ces informations que, de votre côté, vous pourrez évidemment continuer à exploiter.

## Avant-propos

---

## Résumé des modifications

Les modifications apportées à ce document pour la version 7.1 incluent :

- **La recherche de sections associées à un type particulier**

Les types date/heure et entier sont pris en charge. Ceci vous permet, par exemple, de rechercher des documents dont la date est postérieure au 13/12/99 dans la section Date de publication.

- **La prise en charge de Datalink**

Les documents auxquels il est fait référence via une colonne de type Datalink peuvent être indexés.

- **La prise en charge d'un plus grand nombre de CCSID**

Les restrictions sur les CCSID pour les index Ngram ont été levées.

- **Les index double effet ne sont plus acceptés.**

Utilisez plutôt un index linguistique ou un index exact normalisé.

- **La modification de la syntaxe de search\_result**

La syntaxe de la fonction search\_result a été modifiée. La syntaxe antérieure est toujours acceptée aux fins de compatibilité amont.

- **La prise en charge de plus de langues**

La liste des langues et CCSID pris en charge s'est étoffée.

- **Les possibilités de configuration ont été étendues**

Vous pouvez configurer le marqueur sémantique et personnaliser la configuration du serveur afin, par exemple, de réorganiser les seuils.

- **Une nouvelle commande de nettoyage du système pour UNIX**

Une nouvelle commande de serveur, TXIPCLEAN, pour plateformes UNIX a été développée pour faciliter le nettoyage du système si le démarrage du serveur DB2 Extension Texte échoue.

## Récapitulatif des modifications

---

## Partie 1. Guide



---

## Chapitre 1. Présentation de DB2 Extension Texte

DB2 Extension Texte fait partie de la famille DB2 Extensions. Elle permet aux programmeurs d'inclure dans leurs applications des requêtes SQL concernant des documents texte.

Les autres extensions permettent de rechercher des données image, audio et vidéo, et permettent de traiter des documents XML et des données spatiales.

DB2 Extension Texte permet aux requêtes SQL de tirer parti de l'extraction de texte intégral via des fonctions DB2 UDB permettant de stocker dans des bases de données des documents texte non structurés.

DB2 Extension Texte offre aux utilisateurs de DB2 UDB et aux programmeurs d'applications une méthode rapide, souple et intelligente de recherche dans les documents texte. Son principal atout réside dans sa capacité à effectuer des recherches à grande vitesse parmi des milliers de documents texte de grande taille et à trouver, non seulement les termes demandés, mais également leurs synonymes et variantes.

DB2 Extension Texte ne limite pas les recherches aux documents texte stockés dans des bases de données DB2 UDB. Il est possible d'élargir la recherche aux documents texte stockés dans des fichiers.

Le mode de fonctionnement de DB2 Extension Texte repose sur la technologie de recherche linguistique très évoluée d'IBM. Elle permet à vos applications d'accéder aux documents texte et d'en extraire des données de diverses manières. Vos applications peuvent :

- rechercher des documents contenant du texte spécifique, les synonymes d'un mot ou d'une expression ou des mots situés à proximité, c'est-à-dire, par exemple, dans une même phrase ou un même paragraphe ;
- effectuer des recherches globales à l'aide de caractères génériques placés en début, au milieu ou en fin de mot ;
- rechercher des documents rédigés en langues différentes dans des formats différents ;
- rechercher une correspondance "partielle" de mots dont l'orthographe ressemble au terme fourni. Ceci permet de retrouver des mots même s'ils sont mal orthographiés.
- rechercher un texte libre en utilisant un argument de recherche en langage courant.
- rechercher des mots dont la sonorité ressemble au terme fourni.

## Présentation

Votre recherche de texte peut être intégrée à des requêtes professionnelles. Par exemple, vous pouvez coder une requête SQL dans une application pour rechercher des documents texte créés par un auteur particulier, au cours d'une période précise, et contenant une expression ou un mot particulier. L'interface de programmation DB2 Extension Texte permet également de parcourir des documents.

En intégrant la recherche de texte intégral dans les requêtes SELECT de DB2, vous obtenez une puissante fonction d'extraction associant la recherche d'attributs et la recherche de texte intégral. Les instructions SQL suivantes en sont un exemple :

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

DB2TX.CONTAINS est l'une des nombreuses fonctions de recherche de DB2 Extension Texte. DB2BOOKS\_HANDLE est le nom d'une colonne de descripteurs faisant référence à la colonne DB2BOOKS contenant les documents texte dans lesquels s'effectue la recherche. Le restant de l'instruction est un exemple d'argument de recherche se composant des mots autorisation, table et supprimer, ou de l'un des synonymes de supprimer.

---

## DB2 Extension Texte en environnement client/serveur DB2

La figure 1 à la page 5 montre la façon dont DB2 Extension Texte est intégrée dans l'environnement client-serveur DB2.

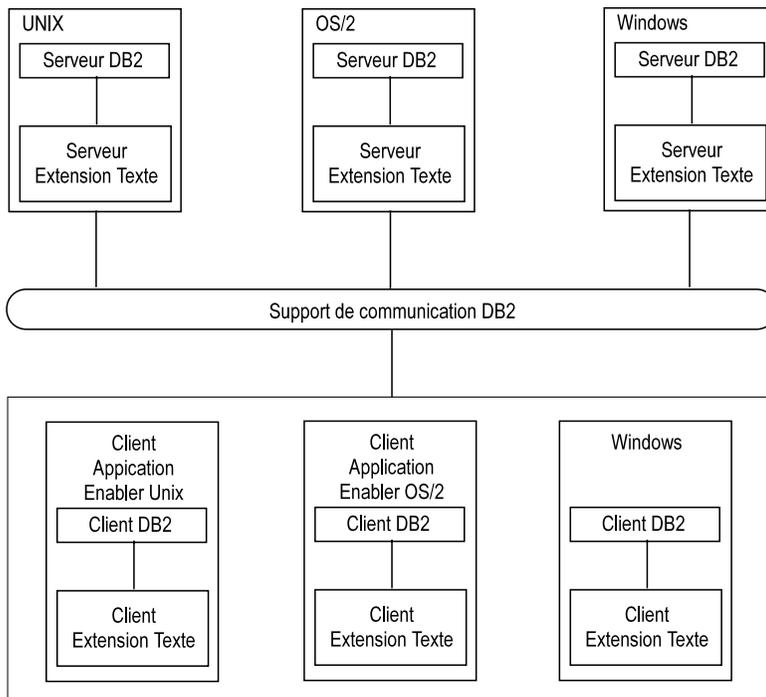


Figure 1. Intégration de DB2 Extension Texte dans l'environnement client-serveur DB2

Pour obtenir la liste des protocoles support de communication DB2 (tels que TCP/IP ou NETBIOS) pour un client, reportez-vous au manuel *DB2 - Mise en route* pour la plateforme appropriée.

La principale partie de DB2 Extension Texte est installée sur le même poste que le serveur DB2. Il n'est possible d'installer qu'une seule instance de serveur DB2 Extension Texte par instance de serveur DB2.

L'installation de DB2 Extension Texte est modulable et peut comprendre :

- un ou plusieurs des serveurs DB2 Extension Texte sur l'un des systèmes d'exploitation illustré par la figure 1, où UNIX désigne des postes de travail AIX(R), SUN-Solaris et HP-UX.
- des clients AIX, SUN-Solaris, HP-UX, OS/2, Windows NT, Windows 98, ou Windows 95 ayant accès à un ou plusieurs serveurs éloignés de DB2 Extension Texte.
- des clients AIX contenant un serveur local et ayant accès à des serveurs éloignés.

La figure 2 à la page 6 présente une configuration classique de DB2 Extension Texte. Pour exécuter DB2 Extension Texte à partir d'un client, vous devez tout

## Présentation

d'abord installer un client DB2 et certaines fonctions de DB2 Extension Texte. Ces fonctions constituent le "client" de DB2 Extension Texte, bien qu'elles ne soient pas un client à proprement parler. Le client communique avec le serveur via la connexion client DB2.

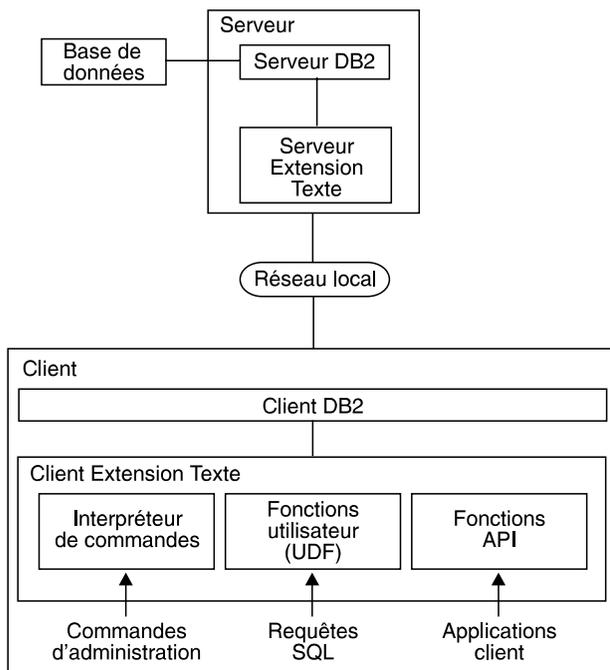


Figure 2. Exemple de configuration de DB2 Extension Texte

Les principaux composants de DB2 Extension Texte sont les suivants :

- **Un interpréteur de commandes.** Ces commandes vous permettent de disposer le texte en colonnes à des fins de recherche et de mise à jour des index de texte.
- **Des fonctions SQL.** Des fonctions peuvent être insérées dans des requêtes SQL pour la recherche dans un texte et, par exemple, pour connaître le nombre d'occurrences du terme recherché. Pour plus de clarté, la figure présente les fonctions SQL sur le poste client car elles peuvent être utilisées comme éléments d'une requête SQL. En fait, elles font partie de l'installation du serveur et s'exécutent sur ce dernier. Cependant, ces fonctions peuvent être utilisées à partir de n'importe quel poste client DB2 sans qu'il soit nécessaire d'installer le client DB2 Extension Texte.
- **Une interface de programmation d'applications (API)** se composant de fonctions pouvant être appelées dans des programmes en C pour la recherche dans un texte et pour l'affichage des résultats de la recherche.

**Conseil**

Les utilitaires client de DB2 Extension Texte offrent des fonction de préparation de texte, d'administration et une API. (Ces fonctions sont disponibles sur le serveur). Pour pouvoir les utiliser vous devez installer le client DB2 Extension Texte. Si vous n'effectuez la préparation de texte et l'administration de DB2 Extension Texte que sur le serveur de cette extension, vous ne devez installer les utilitaires client de DB2 Extension Texte que sur les clients utilisant les fonctions API.

Il est inutile d'installer le client DB2 Extension Texte si vous n'utilisez sur le client que la fonction de recherche par instructions SQL DB2 UDB. Toutes les communications sont gérées par DB2 UDB et le moteur de recherche DB2 Extension Texte ne fonctionne que sur le serveur.

**DB2 Extension Texte en environnement de base de données partitionnée**

La figure 3 montre la façon dont DB2 Extension Texte est intégrée dans l'environnement de base de données partitionnée DB2.

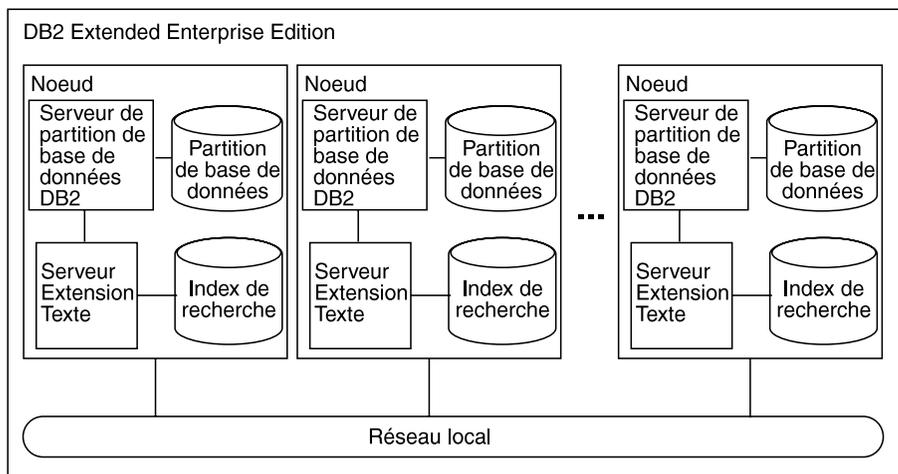


Figure 3. Intégration de DB2 Extension Texte dans l'environnement de base de données partitionnée DB2

Les serveurs de partitions de base de données, également appelés *noeuds*, peuvent être soit affectés chacun à une machine distincte, soit situés sur la même machine. Dans le premier cas, les serveurs sont considérés comme des *noeuds physiques*, tandis que ceux situés sur la même machine sont appelés *noeuds logiques*. DB2 Extension Texte prend en charge les deux configurations.

## Présentation

Un *groupe de noeuds* est un sous-ensemble nommé d'un ou de plusieurs serveurs de partition de base de données. Si une table dotée de DB2 Extension Texte est affectée à un groupe de noeuds multipartition, les index de texte créés par DB2 Extension Texte sont situés sur le même noeud que chaque partition de table correspondante. Ce procédé permet d'éviter que les données soient déplacées d'un noeud à l'autre au cours de l'indexation.

Un serveur DB2 Extension Texte s'exécute sur chaque partition d'un groupe de noeuds et gère les lignes indexées de la partition correspondante. Vous ne pouvez exploiter DB2 Extension Texte que par le biais d'une seule partition de base de données. Chaque partition peut servir à émettre des requêtes DB2 Extension Texte. Ces requêtes sont ensuite transmises à toutes les partitions de base de données concernées.

---

## Chapitre 2. Installation, configuration et maintenance

---

### Installation sur systèmes AIX, Sun-Solaris et HP-UX

DB2 Extension Texte Version 7 pour AIX, Sun-Solaris et HP-UX peut s'exécuter dans un environnement de base de données non partitionnée. Pour AIX et Sun-Solaris, elle peut également tirer parti de la prise en charge de base de données partitionnée qu'offre DB2 Extended Enterprise Edition.

Les étapes d'installation sont les suivantes :

1. Installez les composants du produit sur la machine cible.
2. Installez l'instance DB2 Extension Texte.
3. Démarrez le serveur DB2 Extension Texte.

### Conditions d'installation

Avant d'installer DB2 Extension Texte, vérifiez que vous disposez des éléments suivants pour les composants client et serveur.

#### Serveur

Les composants serveurs impliquent l'installation de DB2 Version 7.1 (ou suivante).

#### Client

Les composants client impliquent l'installation de DB2 Personal Edition, version 7.1 (ou suivante), de DB2 Client Application Enabler, version 7.1 (ou suivante) ou de DB2 Software Developer's Kit, version 7.1 (ou suivante).

### Etape 1 pour AIX : Installation des composants du produit

Lors de l'installation, un groupe appelé smadmin est créé automatiquement.

1. Connectez-vous en tant qu'utilisateur root à la machine serveur ou client cible.
2. Utilisez l'outil SMIT (System Management Interface Tool) pour exécuter `installp` à partir d'une interface appelée par invite, ou entrez la commande `installp`.
3. Entrez la commande `smit install_latest`. Le menu de l'outil SMIT s'affiche.
4. Entrez l'emplacement du module dans la zone INPUT device / directory for software.
5. Cliquez sur le bouton D0 ou appuyez sur Entrée. Cette opération confirme le répertoire d'installation.
6. Précisez dans la zone SOFTWARE TO INSTALL s'il s'agit d'installer les composants client ou serveur.

## Installation sur systèmes UNIX

7. Cliquez sur le bouton D0 ou appuyez sur Entrée. Une invite vous demande de confirmer les paramètres d'installation. Pour confirmer, appuyez sur Entrée.
8. Indiquez dans la zone SOFTWARE TO INSTALL le dictionnaire (s'il en existe) que vous souhaitez installer sur le client. Si vous n'en indiquez pas, le dictionnaire anglais est installé par défaut.

Le tableau 1 fournit la liste des dictionnaires DB2 Extension Texte. Pour installer tous les dictionnaires sur le poste client, entrez db2tx\_07\_01.dic dans la zone SOFTWARE TO INSTALL.

Tableau 1. Dictionnaires DB2 Extension Texte

Dictionnaire	Langue
db2tx_07_01.dic.Ar_AA	Arabe
db2tx_07_01.dic.Da_DK	Danois
db2tx_07_01.dic.NI_NL	Néerlandais
db2tx_07_01.dic.En_GB	Anglais (Royaume Uni)
db2tx_07_01.dic.En_US	Anglais (Etats-Unis)
db2tx_07_01.dic.Fi_FI	Finnois
db2tx_07_01.dic.Fr_FR	Français
db2tx_07_01.dic.Fr_CA	Français (Canada)
db2tx_07_01.dic.De_DE	Allemand
db2tx_07_01.dic.De_CH	Allemand (CH)
db2tx_07_01.dic.Iw_IL	Hébreu
db2tx_07_01.dic.Is_IS	Islandais
db2tx_07_01.dic.It_IT	Italien
db2tx_07_01.dic.No_NO	Norvégien
db2tx_07_01.dic.Pt_PT	Portugais
db2tx_07_01.dic.Ru_RU	Russe
db2tx_07_01.dic.Es_ES	Espagnol et Catalan
db2tx_07_01.dic.Sv_SE	Suédois
db2tx_07_01.dic.Th_TH	Thailandais

9. Déconnectez-vous.

### Etape 1 pour Sun-Solaris : Installation des composants du produit

Lors de l'installation, un groupe appelé smadmin est créé automatiquement.

1. Connectez-vous en tant qu'utilisateur root à la machine serveur ou client cible.

2. Exécutez la commande tar pour extraire le module installable du fichier fixpack :  
tar -xvf IBMdb2tx.tar
3. Exécutez la commande pkgadd pour transférer les fichiers de votre répertoire à la machine cible. Vous pouvez le faire de l'une des deux manières suivantes :
  - Installer le client et le serveur DB2 Extension Texte avec le dictionnaire anglais (Etats-Unis) comme suit :  
pkgadd -d . db2txclnt db2txsrv db2txcom db2txmsgs db2txmsc db2txdius
  - Préciser si vous souhaitez installer un client ou un serveur DB2 Extension Texte et les dictionnaires que vous voulez charger comme suit :  
./db2tx\_install
4. Déconnectez-vous.

### Etape 1 pour HP-UX : Installation des composants du produit

Lors de l'installation, un groupe appelé smadmin est créé automatiquement.

1. Connectez-vous en tant qu'utilisateur root à la machine serveur ou client cible.
2. Exécutez la commande tar pour extraire le module installable du fichier fixpack :  
tar -xvf b2tx.tar
3. Exécutez le programme swinstall :  
swinstall

Cette opération ouvre les fenêtres Software Selection et Specify Source.

4. Le cas échéant, modifiez la valeur indiquée pour Source Host Name.
5. Précisez votre répertoire tar en indiquant IBMdb2tx comme valeur pour la zone Source Depot Path.
6. Dans la zone Change Software View, sélectionnez Products.
7. Sélectionnez OK pour retourner à la fenêtre Software Selection, qui contient la liste des logiciels disponibles pour l'installation.
8. A présent, procédez comme suit :
  - Pour installer DB2 Extension Texte avec tous les dictionnaires et la documentation dans toutes les langues fournies, mettez DB2TX Product en évidence.
  - Pour installer des parties de DB2 Extension Texte:
    - a. Sélectionnez Open Item dans le menu Actions.

## Installation sur systèmes UNIX

- b. Pour chaque ensemble de fichiers que vous souhaitez installer, mettez-le en évidence et sélectionnez **Choose Mark for Install** dans le menu **Actions**. Vous pouvez choisir d'installer le serveur ou le client DB2 Extension Texte :

Serveur Extension Texte - DB2TX.db2txsrv  
Client Extension Texte - DB2TX.db2txclnt

- c. Choisissez les dictionnaires (s'ils existent) que vous souhaitez installer. Le tableau 2 fournit la liste des dictionnaires DB2 Extension Texte.

Tableau 2. Dictionnaires DB2 Extension Texte

Dictionnaire	Langue
db2txdiaa	Arabe
db2txdidk	Danois
db2txdinl	Néerlandais
db2txdigb	Anglais (Royaume Uni)
db2txdius	Anglais (Etats-Unis)
db2txdifi	Finnois
db2txdifr	Français
db2txdica	Français (Canada)
db2txdide	Allemand
db2txdich	Allemand (CH)
db2txdiil	Hébreu
db2txdiis	Islandais
db2txdiit	Italien
db2txdino	Norvégien
db2txdipt	Portugais
db2txdies	Russe
db2txdiru	Espagnol et Catalan
db2txdise	Suédois
db2txdith	Thailandais

9. Sélectionnez **Install** dans le menu **Actions** pour démarrer l'installation du produit et ouvrir la fenêtre **Install Analysis**.
10. Sélectionnez **OK** dans la fenêtre **Install Analysis** quand la zone **Status** affiche un message **Ready**.
11. Sélectionnez **Yes** dans la fenêtre **Confirmation** pour confirmer que vous voulez installer les logiciels DB2 pour HP-UX.

12. Affichez la fenêtre Install pour lire les données en cours de traitement pendant l'installation du logiciel. Celle-ci est terminée quand la zone Status indique Ready et que la fenêtre Note s'ouvre.

Le programme `swinstall` charge l'ensemble de fichiers et exécute les scripts de contrôle pour cet ensemble de fichiers.

Le système d'exploitation HP-UX fournit une aide détaillée sur `winstall`. Vous pouvez en disposer en entrant :

```
man swinstall
```

13. Déconnectez-vous.

### Etape 2 : Mise en place de l'instance DB2 Extension Texte

1. Vérifiez que vous êtes actif en tant qu'utilisateur root.
2. Sous AIX, entrez :

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

Sous Solaris et HP-UX, entrez :

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

3. Si vous disposez d'une instance existante, entrez :

```
./txiupdt instanceid db2instanceid
```

pour mettre à jour l'instance au moyen des nouveaux fichiers.

Sinon, pour créer une nouvelle instance DB2 Extension Texte, entrez :

```
./txicrt id-instance db2instanceid
```

où *id-instance* est un ID utilisateur d'instance DB2 existant ou un ID utilisateur à associer à un ID instance DB2.

L'exécution de `txicrt` crée le répertoire `/u/id-instance/db2tx`. Ne créez pas de répertoires ou fichiers supplémentaires sous `u/id-instance/db2tx` ; ils risquent d'être perdus en cas de suppression de l'instance.

Si vous utilisez un environnement de base de données partitionnée, des paramètres supplémentaires sont nécessaires. Pour plus de détails, reportez-vous à la section «TXICRT» à la page 176.

4. Déconnectez-vous.

### Etape 3 : Démarrage du serveur DB2 Extension Texte

1. Connectez-vous en tant qu'*id-instance*.
  2. Arrêtez et redémarrez DB2.
  3. Pour lancer le serveur DB2 Extension Texte, entrez :
- ```
txstart
```
4. Déconnectez-vous.

## Installation sur systèmes Windows

---

### Installation sous Windows 95, Windows 98, Windows NT et Windows 2000

DB2 Extension Texte est installé en tant que composant du module DB2 Extensions.

#### Serveur

Les composants serveur pour DB2 Extensions peuvent être installés dans un serveur Windows NT ou Windows 2000.

Les composants serveurs devant s'exécuter dans un environnement de base de données non partitionné impliquent l'installation de DB2 Version 7.1 (ou suivante).

Les composants serveurs devant s'exécuter dans un environnement de base de données partitionné impliquent l'installation de DB2 Extended Enterprise Edition version 7.1 (ou suivante).

#### Client

Les composants client pour DB2 Extensions peuvent être installés sous Windows 95, Windows 98, Windows NT ou Windows 2000.

Les composants client impliquent l'installation de DB2 Personal Edition, version 7.1 (ou suivante), de DB2 Client Application Enabler, version 7.1 (ou suivante) ou de DB2 Software Developer's Kit, version 7.1 (ou suivante) pour Windows.

#### Installation de DB2 Extensions

Pour installer DB2 Extensions sous Windows :

1. Insérez le CD-ROM dans le lecteur de CD-ROM de votre poste de travail client.
2. Si la fonction de démarrage automatique de CD est activée, le Tableau de bord DB2 Extensions démarre automatiquement. Sinon, cliquez sur le bouton Démarrer dans la barre des tâches. Sélectionnez Exécuter. Entrez :  
`d:\db2ext\winnt95\setup`  
dans la zone Ouvrir, où "d" est la lettre correspondant au lecteur de CD-ROM.
3. A partir du Tableau de bord, vous pouvez afficher le fichier README de DB2 Extension. Consultez les conditions préalables à l'installation ou installez le produit.
4. Sélectionne le bouton Installer sur la gauche pour commencer l'installation.
5. Dans la première fenêtre, sélectionnez la version de DB2 Extensions à installer Client/Serveur pour EEE, Client/Serveur pour non EEE ou Client seulement.

6. Dans la fenêtre suivante, choisissez le dossier dans lequel les extensions seront installées.
7. Sélectionnez ensuite le type d'installation que vous souhaitez : Standard, Sélective ou Minimale.

L'installation **Standard** englobe le client/serveur DB2 Extensions, les fichiers exemples, SDK, la documentation en ligne en anglais (Etats-Unis) et le dictionnaire DB2 Extension Texte en anglais (Etats Unis).

L'installation **Minimale** n'installe que le client/serveur et le dictionnaire DB2 Extension Texte en anglais (Etats Unis).

L'installation **Sélective** vous permet de choisir parmi les composants suivants :

- Sélectionnez **DB2 Extension - Serveur (EEE)** si vous installez les extensions DB2 dans un environnement de base de données partitionnée (DB2 Extended Enterprise Edition). Cette opération installe également le composant **DB2 Extensions - Client**.
- Sélectionnez **DB2 Extension - Serveur (non EEE)** si vous installez les extensions DB2 dans un environnement de base de données partitionnée (DB2 Extended Enterprise Edition). Cette opération installe également le composant **DB2 Extensions - Client**.
- Sélectionnez **DB2 Extensions - Client** si vous installez un environnement DB2 Extensions client.
- Sélectionnez **DB2 Extensions - SDK** si vous voulez installer des fichiers d'en-tête et des bibliothèques afin de développer des applications DB2 Extensions.
- Sélectionnez **Fichiers exemples** si vous voulez installer les exemples de programmes et de fichiers multi-média.
- Sélectionnez **Documentation en ligne** si vous voulez installer les publications DB2 Extensions au format électronique. Cliquez ensuite sur le bouton **Changer** dans la section **Description** pour visualiser la liste des publications disponibles selon les langues. Choisissez celles que vous souhaitez installer.
- Sélectionnez **Dictionnaires Extension Texte** si vous voulez installer le support linguistique associé à DB2 Extension Texte. Cliquez ensuite sur le bouton **Changer** dans la section **Description** pour visualiser la liste des dictionnaires disponibles selon les langues. Choisissez ceux qu'il convient d'installer. L'anglais (Etats Unis) est toujours obligatoire.

Si vous ne disposez pas de suffisamment d'espace disque pour les composants sélectionnés, cliquez sur **Précédent** pour sélectionner à nouveau le répertoire d'installation. Sino, cliquez sur **Suivant**.

8. Si vous installez le serveur Extensions (EEE) dans un environnement de base de données partitionnée et que votre machine est propriétaire d'instances, vous êtes invités à indiquer une plage de ports TCP/IP. Le même nom de partage doit exister sur tous les noeuds associés à l'instance DB2 en cours.

## Installation sur systèmes Windows

Si vous installez DB2 Extensions sur une machine ne possédant pas d'instance, vous êtes invité à indiquer le nom d'une machine propriétaire d'instances. Les plages de ports TCP/IP et les informations relatives au chemin de données seront obtenus via la machine propriétaire d'instances.

9. Cliquez sur le bouton *Suivant* pour installer les composants sélectionnés. Au démarrage, le programme d'installation vérifie que les logiciels nécessaires à DB2 Extensions sont installés sur votre poste de travail.
10. Arrêtez le poste de travail et redémarrez-le.

### Création et gestion d'instances

Vous pouvez créer et utiliser de multiples instances du serveur DB2 Extensions. Vous devez créer de multiples instances si vous avez créé plusieurs instances du serveur DB2 pour Windows. Chaque instance du serveur DB2 Extensions est associée à une instance de DB2 serveur pour Windows.

Par défaut, une instance de DB2 Extensions, nommée DB2, est créée lorsque vous installez un programme DB2 Extensions sous Windows dans un environnement de base de données non partitionnée.

Par défaut, une instance de DB2 Extensions, nommée DB2MPP, est créée lorsque vous installez un programme DB2 Extensions sous Windows dans un environnement de base de données partitionnée. Dans un environnement de base de données partitionnée, aucune instance par défaut n'est créée pour DB2 Extension Texte. Utilisez la commande `txicrt` pour créer une instance. Pour plus d'informations, reportez-vous à la section «Création d'une instance DB2 Extension Texte» à la page 21.

### Préparation d'une base de données exemple pour vérifier l'installation

DB2 Extension Texte inclut des utilitaires servant à créer une base de données exemple. Ils permettent de préparer rapidement du texte pour tester les fonctions de recherche de DB2 Extension Texte, immédiatement après l'installation, et plus généralement, pour tout test ultérieur.

Ces utilitaires se trouvent dans les répertoires suivants :

- Sous OS/2 (R), Windows NT et Windows 2000 : `DMBMMPPATH\samples` où `DMBMMPPATH` correspond au chemin d'installation.
- Sous AIX, HP-UX et SUN-Solaris : `$HOME/db2tx/samples`

Pour préparer un exemple de base de données, que ce soit sur le serveur ou sur un client :

1. Entrez la commande suivante à l'invite du système d'exploitation :  
`TXVERIFY nom_bdd [ID-utilisateur]  
[mot-de-passe]`

Cette commande crée et active un modèle de base de données.

- Entrez la commande suivante à l'invite du système d'exploitation :  
TXSAMPLE nom-bdd [ID-utilisateur] [mot-de-passe]

L'ID utilisateur et le mot de passe ne sont requis que si vous travaillez à partir d'un poste client. Cette commande effectue les opérations suivantes :

- Connexion à la base de données spécifiée

- Création de la table `db2tx.sample`

- Importation des modèles de document en anglais pour peupler la table

- Activation de la colonne de texte de la table, avec les types d'index spécifiés ci-dessous.

  - exact

  - précis normalisé

  - linguistique

  - linguistique avec activation de sections

  - Ngram

  - Ngram avec distinction entre les majuscules/minuscules

- Attente de la génération de l'index de texte.

---

## Configuration

La présente section décrit les variables d'environnement de DB2 Extension Texte ainsi que les informations de configuration. Dans les deux cas, vous avez la possibilité de spécifier des valeurs par défaut pour de nombreux paramètres requis par DB2 Extension Texte.

### Variables d'environnement

Les variables d'environnement définissent les valeurs par défaut des paramètres d'environnement. Pour afficher la définition des variables d'environnement, utilisez la commande `GET ENVIRONMENT` décrite à la section «Affichage des paramètres des variables d'environnement» à la page 107.

|                        |                                                                                                                                                                                                                   |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>DB2INSTANCE</b>     | Nom de l'instance DB2 UDB.                                                                                                                                                                                        |
| <b>DB2DBDFT</b>        | Nom de la base de données par défaut, c'est-à-dire le nom de la base de données DB2 UDB prise en compte si aucun nom de base de données n'est spécifié.                                                           |
| <b>DB2TX_INSTOWNER</b> | Nom de l'instance DB2 Extension Texte. Il s'agit du nom de l'utilisateur propriétaire de l'instance au moment de la connexion. Il est obligatoire si un client DB2 Extension Texte se connecte à un serveur UNIX. |

## Configuration

### DB2TX\_INSTOWNERHOMEDIR

Répertoire principal du propriétaire de l'instance (UNIX uniquement).

**En environnement Windows**, les variables d'environnement de DB2 Extension Texte sont spécifiées comme variables système dans la page Propriétés système. Elles sont ajoutées à votre système lors de l'installation.

**En environnement OS/2**, les variables d'environnement de DB2 Extension Texte sont ajoutées au fichier CONFIG.SYS lors de l'installation.

**En environnement UNIX**, les variables d'environnement de DB2 Extension Texte, à l'exception de DB2DBDFT, sont spécifiées dans les scripts de profils de shell suivants :

DB2TXCSHRC pour le shell C

DB2TXPROFILE pour les shells Korn et Bourne

Lorsque vous créez une instance, le fichier db2txprofile ou db2txcshrc est copié dans .profile.

Lorsque vous créez une instance, db2txprofile est ajouté au fichier .profile et db2profile du propriétaire de l'instance. Le script shell est stocké dans \$DB2TX\_INSTOWNERHOMEDIR/db2tx.

Vous avez également besoin des variables d'environnement suivantes :

**LANG** Spécifie les paramètres de langue.

**NLSPATH** Spécifie le chemin d'accès au fichier de messages de DB2 Extension Texte.

**DB2ENVLIST** Liste des variables d'environnement qui sont transmises par DB2 aux fonctions DB2 Extension Texte.

**RAHUSER** Spécifie le nom du propriétaire d'instance utilisé pour la commande DB2 db2a11. Variable uniquement utilisée dans l'environnement DB2 EEE.

Si vous êtes le propriétaire de l'instance, ces variables d'environnement sont automatiquement définies lors de la création de l'instance. Si vous n'êtes pas le propriétaire de l'instance, incluez le profil db2tx du propriétaire de l'instance, plutôt que d'essayer de définir ces variables d'environnement manuellement.

## Conseil

Chaque fois que vous modifiez des variables d'environnement, vous devez arrêter et redémarrer DB2 et DB2 Extension Texte pour que les modifications s'appliquent.

## Définitions de la configuration de texte

Chaque base de données comporte des paramètres de configuration de texte formés de :

- Caractéristiques de texte
- Caractéristiques d'index
- Caractéristiques de traitement

Ces paramètres sont définis lorsque vous autorisez l'utilisation de la base de données par DB2 Extension Texte. La commande `ENABLE DATABASE` utilise soit les paramètres que vous y définissez, soit les paramètres initiaux décrits ici. Vous avez la possibilité d'afficher et de modifier ces paramètres. Pour plus de détails à ce sujet, reportez-vous aux sections «Affichage des définitions de la configuration de texte» à la page 107, et «Modification de la configuration du texte» à la page 20.

### Caractéristiques de texte

Le «Chapitre 4. Planification de vos besoins en matière de recherche» à la page 29, décrit les formats de document, les langues et les CCSID pris en charge par DB2 Extension Texte. Les valeurs par défaut de ces éléments sont nécessaires dans de nombreuses commandes.

**FORMAT** Paramétrage initial : TDS

**LANGUAGE** Paramétrage initial : Langue (LANGUAGE) définie pour la base de données

**CCSID** Paramétrage initial : CCSID défini pour la base de données

### Caractéristiques d'index

**DIRECTORY** Répertoire utilisé pour stocker l'index.

Paramétrage initial pour les systèmes UNIX :

`DB2TX_INSTOWNERHOMEDIR/db2tx/indexes`

Paramétrage initial sous OS/2, Windows NT et Windows 2000 : `DMBMMPATH\instance\nom-instance \db2tx\index` où `DMBMMPATH` représente le chemin d'installation de DB2 Extension Texte. Pour les clients Windows, `DIRECTORY` doit correspondre à un nom de répertoire sur le poste serveur.

## Configuration

**INDEXTYPE** Type d'index à utiliser. Pour obtenir une description, reportez-vous à la section «Types de recherche» à la page 40.  
Paramétrage initial : NGRAM

**UPDATEFREQ** Fréquence de mise à jour périodique de l'index. Pour obtenir une description, reportez-vous à la section «Mise à jour d'un index» à la page 46.  
Paramétrage initial : NONE

### Caractéristiques de traitement

**UPDATEINDEX**  
Paramètre définissant à quel moment intervient la première mise à jour de l'index : immédiatement, lors de l'étape d'activation, ou plus tard, suite à une commande UPDATE INDEX explicite ou en fonction des paramètres de fréquence de mise à jour de l'index (NOUPDATE).  
Paramétrage initial : UPDATE

**COMMITCOUNT**  
Paramètre définissant le nombre d'instructions de mise à jour ou d'insertion à l'issue desquelles DB2 Extension Texte émet une instruction de validation DB2 UDB. Pour plus de détails, reportez-vous à la section «Activation d'une colonne de texte dans une table volumineuse» à la page 70.  
Paramétrage initial : 0

### Modification de la configuration du texte

Lors de l'installation initiale de DB2 Extension Texte, les valeurs par défaut sont définies pour la configuration de texte. Pour afficher les valeurs en cours, reportez-vous à la section «Affichage des définitions de la configuration de texte» à la page 107.

Pour modifier les définitions par défaut de la configuration de texte servant à la création des index, entrez :

```
db2tx "CHANGE TEXT CFG USING paramètres"
```

#### Exemples :

Pour modifier le type d'index et le répertoire par défaut pour les futurs index :

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

Pour modifier la fréquence de mise à jour par défaut des index, afin qu'ils soient actualisés à 12 ou 15 heures, du lundi au vendredi, et s'il existe au moins 100 documents texte en file d'attente :

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

Pour arrêter la mise à jour périodique d'un index :

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

---

### Configuration et gestion d'un serveur DB2 Extension Texte

La configuration d'un serveur DB2 Extension Texte pour utilisation par des clients DB2 Extension Texte constitue l'étape suivante pour un DB2 Extension Texte nouvellement installé. Cette opération est normalement effectuée par un administrateur de serveur DB2 Extension Texte et comporte :

1. La création d'une instance DB2 Extension Texte
2. Le démarrage d'un serveur DB2 Extension Texte

Parmi les autres tâches nécessaires à la la gestion d'un serveur DB2 Extension Texte :

1. Sauvegarde et restauration des index DB2 Extension Texte et des bases de données activées
2. Trace des incidents

Reportez-vous au «Chapitre 10. Commandes d'administration du serveur» à la page 175 pour connaître la syntaxe des commandes.

#### Création d'une instance DB2 Extension Texte

Avant de commencer à utiliser DB2 Extension Texte, vous devez créer une instance DB2 Extension Texte. Chaque instance créée constitue un environnement indépendant dans lequel vous pouvez gérer des index, stockés dans des répertoires différents.

**Pour créer une instance dans un environnement mono-noeud UNIX**, entrez :

```
txicrt nom-instance nom-instance-db2
```

où *nom-instance* est l'ID utilisateur du propriétaire de l'instance et *nom-instance-db2* est le nom de l'instance DB2 UDB associée.

**Sous Windows NT et Windows 2000** une instance par défaut est créée lors de l'installation. Pour créer une instance supplémentaire, entrez :

```
txicrt nom-instance-db2
```

## Configuration et gestion d'un serveur

**Pour savoir comment créer des instances dans un environnement multi-noeud**, reportez-vous à la section «TXICRT» à la page 176.

**Pour supprimer une instance sur des systèmes UNIX**, entrez :

```
txidrop nom-instance nom-instance-db2
```

**La commande équivalente sous Windows NT et Windows 2000 est :**

```
txidrop nom-instance-db2
```

**Pour déterminer l'instance avec laquelle vous souhaitez travailler :**

```
set DB2INSTANCE = nom-instance
```

puis arrêtez et redémarrez DB2 et DB2 Extension Texte.

### Démarrage et arrêt d'un serveur DB2 Extension Texte

**Pour démarrer le serveur DB2 Extension Texte**, connectez-vous sous l'ID utilisateur du propriétaire de l'instance (systèmes UNIX seulement), puis entrez :

```
txstart
```

**Pour afficher l'état du serveur DB2 Extension Texte**, entrez la commande suivante :

```
txstatus
```

**Pour arrêter le serveur DB2 Extension Texte**, entrez la commande suivante :

```
txstop
```

#### Conseil pour les utilisateurs UNIX

Si nous ne parvenez pas à démarrer un serveur DB2 Extension Texte, essayez de nettoyer les ressources partagées. Reportez-vous à la section «Commande TXIPCLEAN (sous UNIX uniquement)» à la page 180.

### Sauvegarde et restauration des index et des bases de données activées

Vous pouvez sauvegarder et restaurer les bases de données activées ainsi que les index de texte créés par DB2 Extension Texte.

Pour effectuer une **sauvegarde** :

1. Déterminez les tables qui ont été activées par DB2 Extension Texte. Pour ce faire, entrez la commande :  

```
db2tx GET STATUS
```
2. Déterminez le nom des répertoires d'index utilisés par la base de données. Pour ce faire, entrez la commande :  

```
db2tx GET INDEX SETTINGS nom-table
```

3. Arrêtez le serveur DB2 Extension Texte Pour ce faire, entrez la commande :  
TXSTOP
4. Sauvegardez les répertoires d'index avec leurs sous-répertoires index et work.
5. Sauvegardez le fichier desmastr.dat, qui est situé dans :  
    rèp-personnel-propriétaire-instance/db2tx/txins000 (UNIX)  
    ou  
    DMBMPPATH\db2tx\txins000 (OS/2, Windows NT et Windows 2000)

Notez que, pour les bases de données partitionnées, il existe plusieurs répertoires d'instances, nommés txins000, txins001, ... Il est donc plus pratique de sauvegarder la totalité du répertoire .../db2tx, plutôt que les fichiers se trouvant dans chaque sous-répertoire.

6. Redémarrez le serveur DB2 Extension Texte par :  
TXSTART

Pour effectuer une **restauration** :

1. Arrêtez le serveur DB2 Extension Textepar :  
TXSTOP
2. Sauvegardez le fichier desmastr.dat existant.
3. Restaurez la copie de sauvegarde du fichier desmastr.dat.
4. Restaurez les copies de sauvegarde des répertoires d'index, avec le même chemin d'accès qu'auparavant.
5. Redémarrez le serveur DB2 Extension Texte par :  
TXSTART

### Trace des incidents

Si vous avez besoin de signaler une erreur à un technicien de maintenance d'IBM, celui-ci peut vous demander d'activer la fonction trace de manière à ce que les informations soient écrites dans un fichier qui servira à localiser l'erreur. N'utilisez la fonction trace qu'à la demande du technicien IBM ou de votre propre support technique.

Les performances du système étant ralenties lorsque la fonction trace est activée, n'utilisez celle-ci que dans le cas d'erreurs.

Pour activer la fonction trace, entrez :

```
txtrace on [options]
```

Vous trouverez la syntaxe ainsi que la liste des événements et des composants à la section «TXTRACE» à la page 192. D'autres options y sont également décrites.

## Configuration et gestion d'un serveur

Vous pouvez également filtrer la fonction trace en paramétrant un "masque" qui, pour chaque enregistrement de la fonction trace, déterminera en fonction de son ID s'il doit être accepté ou rejeté. Par défaut, il n'y a aucun filtre.

Un masque comporte quatre segments séparés par des points, par exemple : 2.2-6.1,3.\*; où :

- 2 Désigne DB2 Extension Texte de DB2 UDB.
- 2-6 N'inclut que les événements dont l'ID est compris entre 2 et 6.
- 1,3 N'inclut que les événements signalés par les composants 1 et 3.
- \* Inclut toutes les fonctions de ces composants.

Vous avez la possibilité d'exclure les erreurs système en-deçà d'un certain seuil de gravité, et de spécifier, au cas où la mémoire tampon de la fonction trace serait saturée, lesquels, des premiers ou des derniers enregistrements, doivent être conservés.

Pour reproduire l'erreur et générer un fichier d'image mémoire où les informations de la trace seront écrites en binaire, entrez :

```
txtrace dump fich-image-mémoire
```

Pour générer une version formatée du fichier d'image mémoire, entrez :

```
txtrace format fich-image-mémoire fich-formaté
```

Lorsque la trace est active, vous pouvez également demander que les informations qu'elle délivre soient directement écrites, à partir de la mémoire partagée, dans un fichier formaté. Pour ce faire, entrez :

```
txtrace format > fichier-formaté
```

Une fois les informations de trace écrites dans un fichier, désactivez la fonction trace par :

```
txtrace off
```

---

## Chapitre 3. Mise en route

Ce chapitre vous permet de vous familiariser avec les principales fonctions permettant des recherches dans du texte. Vous êtes censé utiliser un système DB2 Extension Texte actif, qui a été installé et configuré et sur lequel une instance DB2 Extension Texte a été créée et démarrée.

### Conseil

Ce chapitre ne traite que des fonctions de base permettant la recherche dans un texte. Avant de préparer votre propre texte pour la recherche, reportez-vous à la section «Préparation préalable» à la page 57.

Deux méthodes vous permettent de vous familiariser avec la recherche dans un texte :

- Parcourir un exemple simple en entrant manuellement les commandes, ou
- Exécuter un modèle d'indexation et d'assistant de recherche.

---

### Exemple simple de recherche dans un texte

#### 1. Sous Windows, l'interpréteur de commandes DB2

Sous UNIX, entrez la commande suivante à l'invite du système d'exploitation :

```
=>db2 cmd
```

#### 2. Démarrez l'interpréteur de commandes DB2 Extension Texte

Sous UNIX, entrez la commande suivante à l'invite du système d'exploitation :

```
=>db2tx
```

Sous Windows, entrez la commande suivante à l'invite du système d'exploitation :

```
db2=>db2tx
```

#### 3. Connexion à une base de données

Choisissez une base de données qui contient du texte que vous voulez utiliser pour des recherches. Pour vous connecter à la base de données, entrez la commande suivante :

```
db2tx=>CONNECT TO ma-bdd
```

### 4. Activation d'une base de données pour la recherche de texte

Pour activer la base de données connectée, entrez la commande suivante :

```
db2tx=>ENABLE DATABASE
```

### 5. Activation d'une table de texte pour la recherche de texte

Vous devez activer une table de texte si vous souhaitez créer un seul index pour toute la table.

#### Un ou plusieurs index ?

La section «Création d'un ou de plusieurs index de texte pour une table» à la page 44 explique comment vous pouvez permettre la recherche dans une table soit en créant un seul index de texte pour la table toute entière, soit en créant plusieurs index, à raison d'un par colonne de texte.

- Pour créer un seul index pour toute la table, exécutez à ce stade la commande `ENABLE TEXT TABLE` pour créer un index vide, puis exécutez la commande `ENABLE TEXT COLUMN` (voir étape suivante) à plusieurs reprises, une fois par colonne de texte, afin de remplir l'index unique.
- Sinon, pour créer un index distinct pour chaque colonne de texte, vous ignorez l'étape de création d'un index unique pour la table entière, autrement dit, vous n'exécutez pas `ENABLE TEXT TABLE`, mais uniquement `ENABLE TEXT COLUMN` autant de fois qu'il y a d'index à créer et à remplir, à raison d'un par colonne de texte.

Pour cet exemple, vous n'allez pas créer d'index pour la table entière. Vous allez passer à l'étape suivante pour créer un index par colonne de texte.

### 6. Activation d'une colonne de texte pour la recherche de texte

Entrez la commande suivante pour permettre à DB2 Extension Texte de rechercher dans la colonne de texte `ma-colonne` dans la table `db2tx.sample` et d'affecter le nom `mon-descripteur` à la colonne de descripteurs que crée cette commande.

```
db2tx=>ENABLE TEXT COLUMN db2tx.sample ma-colonne HANDLE mon-descripteur
```

Cette commande crée un index de texte. Des valeurs par défaut sont utilisées pour le type de documents en cours d'indexation et pour les caractéristiques d'index.

### 7. Vérification de l'état de l'index que vous créez

Entrez :

```
db2tx=>GET INDEX STATUS db2tx.sample HANDLE mon-descripteur
```

### 8. Sortie de l'interpréteur de commandes DB2 Extension Texte

Entrez :  
db2tx=>QUIT

### 9. Démarrage de l'interpréteur de commandes DB2

A l'invite du système d'exploitation, entrez la commande suivante :  
=>db2

### 10. Recherche de texte

La recherche peut commencer dans vos documents. Essayez la commande SELECT suivante qui trouve toutes les occurrences de searchterm dans le texte que vous venez d'indexer :

```
db2=>SELECT COUNT (*)
        FROM sample
        WHERE DB2TX.CONTAINS (mon-descripteur,"searchterm") = 1
```

---

## Modèle d'assistant pour l'indexation et la recherche

Un assistant d'indexation et de recherche fournit un exemple de création d'index sur des tables et des colonnes sélectionnées et de recherche sur ces index. Cet assistant vous guide dans les différentes étapes à parcourir pour effectuer une recherche sur les colonnes sélectionnées dans une table à l'aide des fonctions de recherche DB2 Extension Texte.

**Pour démarrer cet assistant**, entrez la commande suivante :

```
txwizard
```

Vous pouvez utiliser le code source Java fourni avec le modèle d'assistant pour réaliser votre propre application de préparation et de recherche de texte ou simplement pour avoir une idée sur la manière d'utiliser et d'appeler les fonctions DB2 Extension Texte dans un environnement Java.

Sous AIX, le code source Java de l'assistant est stocké dans le répertoire  
/usr/lpp/db2tx\_06\_01/samples/wizard

Sous Windows NT et Windows 2000, il est stocké dans le répertoire  
%DMBMPATH%\samples\wizard

## Mise en route

---

## Chapitre 4. Planification de vos besoins en matière de recherche

Avant d'exécuter les étapes décrites dans le «Chapitre 5. Préparation d'un texte pour la recherche» à la page 57, vous devez déterminer :

- Le format et la page de codes de vos documents ainsi que la langue dans laquelle ils sont rédigés
- Le moyen d'éviter des incidents de pages de codes
- Le type de fonctionnalité de recherche dont vous aurez besoin
- La quantité d'espace disque nécessaire
- Le concept d'index et le choix entre un index commun pour une table DB2 ou un index distinct pour chacune des colonnes de la table

Le présent chapitre précise comment recueillir cette information.

Vous avez le choix entre plusieurs types d'index : linguistique, exact, double effet et Ngram. Le choix du type d'index est important. Par exemple, si vous choisissez l'index de type *linguistique*, vous pouvez rechercher les variantes et les synonymes d'un mot. Le type d'index a également des conséquences sur les performances d'indexation et la taille de l'index. Vous pouvez également combiner les capacités de recherche en créant plusieurs index, chacun ayant un type différent, par colonne de texte.

---

### Nécessité d'indexation des documents texte

Un système d'extraction d'informations rapide n'effectue pas de recherche séquentielle dans les documents car cela prendrait trop de temps. Il s'applique à un index de texte généré préalablement. Un index de texte se compose de termes significatifs extraits de documents texte, chaque terme étant stocké avec des informations concernant le document dont il provient.

Un index de texte ne contient que les informations significatives. Les termes qui ne le sont pas, tels que "et", "de" ou "qui", ne sont pas indexés. (Aucun filtrage de mots vides n'est effectué pour les index Ngram.) DB2 Extension Texte se réfère à une liste de ces mots, appelés *mots vides*, pour éviter de les indexer. Le système d'extraction recherche dans l'index les termes demandés afin de déterminer les documents dans lesquels ils apparaissent.

#### Conseil

Si vous devez modifier la liste des mots vides, ne le faites qu'une fois, au moment de l'installation.

## Nécessité d'indexation des documents texte

Il existe un fichier contenant la liste des mots vides par langue, qu'il est possible de modifier (consultez la section «Modification des fichiers de mots vides et d'abréviations» à la page 55). Cependant, comme il n'y a qu'un fichier pour tout le système, ne le modifiez qu'une fois, lorsque vous installez DB2 Extension Texte pour la première fois. Les modifications que vous y apportez plus tard ne sont pas répercutées dans les index existants.

Par exemple, supposons que certains documents contiennent le nom d'un magazine, "Demain". Si vous retirez ce mot de la liste des mots vides, il sera indexé et les recherches suivantes le trouveront. Mais les index créés avant le retrait ne contiennent pas le mot "Demain" et la recherche de ce terme est infructueuse.

Si vous décidez de modifier la liste des mots vides et désirez refléter ce changement partout, il vous faut recréer tous vos index.

L'indexation se fait en deux étapes. La première consiste à enregistrer dans une *table de journalisation* les documents pour lesquels un index doit être créé. Cette opération s'effectue automatiquement via les *déclencheurs* DB2 chaque fois que vous insérez, mettez à jour ou supprimez un document dans une colonne.

La seconde étape consiste à indexer les documents texte figurant dans la table de journalisation. Elle peut être effectuée périodiquement. Les termes des documents insérés ou modifiés dans la colonne sont ajoutés à l'index. Par contre, les termes des documents supprimés de la colonne sont supprimés de l'index.

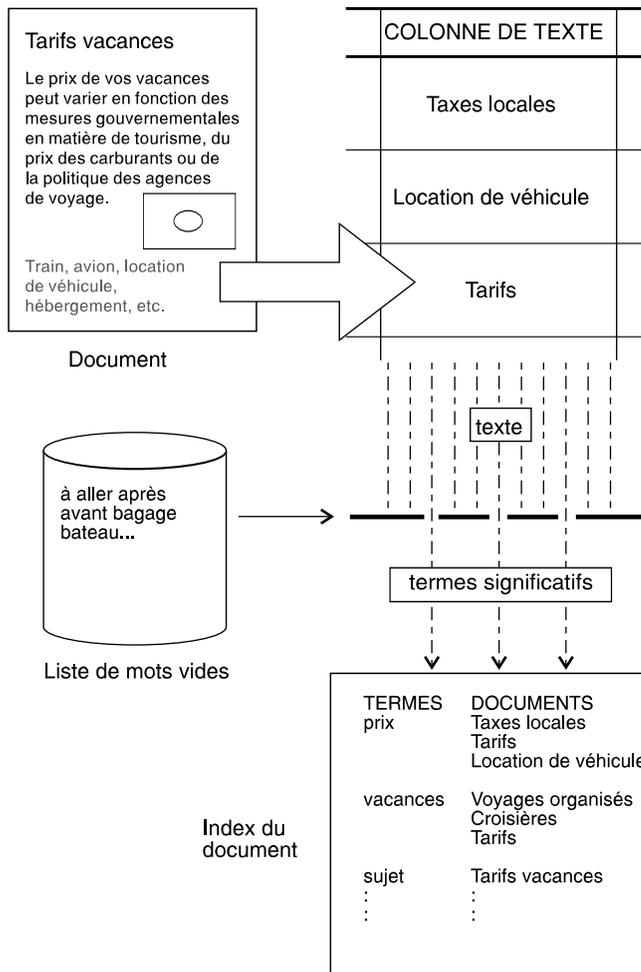


Figure 4. Indexation des termes significatifs

## Documents pris en charge

DB2 Extension Texte doit avoir connaissance du format (ou type) des documents texte devant fait l'objet d'une recherche (par exemple, HTML ou ASCII). Ces informations sont nécessaires lors de l'indexation des documents texte.

Les formats de documents texte acceptés sont :

|                       |                                        |
|-----------------------|----------------------------------------|
| <b>HTML</b>           | Hypertext Markup Language              |
| <b>XML</b>            | Extended Markup Language               |
| <b>ASCII_SECTIONS</b> | ASCII structuré contenant des sections |

## Formats de documents

|               |                                                         |
|---------------|---------------------------------------------------------|
| <b>TDS</b>    | ASCII à plat                                            |
| <b>AMI</b>    | AmiPro Architecture version 4                           |
| <b>FFT</b>    | IBM FFT : DCA                                           |
| <b>MSWORD</b> | Microsoft Word versions 5.0 et 5.5                      |
| <b>RFT</b>    | IBM RFT : DCA                                           |
| <b>RTF</b>    | Microsoft Rich Text Format (RTF) version 1              |
| <b>WP5</b>    | WordPerfect (OS/2 et Windows), versions 5.0, 5.1 et 5.2 |

### Particularités des documents HTML

La prise en compte des trémas et caractères diacritiques dans les documents HTML dépend de la page de codes du document :

- Pour les pages de codes 37, 273, 277, 278, 280, 284, 297, 437, 500, 819, 850, 858, 860, 863, 865, 871, 923, 924 et 1252, les règles suivantes s'appliquent :
  - Notation par entité pour les trémas et les caractères spéciaux, par exemple, `&auml;` pour ä.
  - Seuls les caractères qui comportent un point de code pour la page de codes 819 (ASCII) ou 500 (EBCDIC) sont corrects.
  - Lors de l'indexation, les documents comportant des caractères propres à certaines langues, ä, par exemple, sont générateurs d'erreurs lors de la reconnaissance des mots si la page de codes du document n'est ni 819 ni 500.
  - Si vous ajoutez des documents à un index Ngram, il doit avoir été créé avec la page de codes 819, 500 ou UTF8.
- Pour toutes les autres pages de codes SBCS (jeu de caractères simple octet), les règles suivantes s'appliquent :
  - Les entités ne sont pas résolues.
  - Les caractères spéciaux doivent être écrits en points de code propres à la langue.

### Particularités des documents XML

Les documents XML doivent comporter une chaîne de codage XML qui sera utilisée lors de l'analyse syntaxique. Les codages suivants sont acceptés :

UTF8, utf8, utf-8, ibm-1208, utf\_8,  
UTF16\_BigEndian, utf-16be, utf16  
UTF16\_LittleEndian, utf-16le  
LATIN\_1, latin1, latin-1, ascii, ibm-819, iso-8859-1, Latin-1  
ibm-912, iso-8859-2  
ibm-913, iso-8859-3  
ibm-914, iso-8859-4  
ibm-915, iso-8859-5  
ibm-1089, iso-8859-6

ibm-813, iso-8859-7  
ibm-916, iso-8859-8i  
ibm-920, iso-8859-9  
ibm-0037, ebcdic-cp-us, ebcdic-cp-ca, ebcdic-cp-nl, ebcdic-cp-dk, ebcdic-cp-no,  
ebcdic-cp-fi, ebcdic-cp-se, ebcdic-cp-it, ebcdic-cp-es, ebcdic-cp-gb  
ibm-297, ebcdic-cp-fr, ebcdic-cp-ar1, ebcdic-cp-he, ebcdic-cp-ch, ebcdic-cp-roece,  
ebcdic-cp-yu, ebcdic-cp-is, ebcdic-cp-ar2  
ibm-954, euc-jp eucjis  
ibm-943, shift\_jis, sjis, shiftjis, shift-jis  
ibm-950 , big-5, big5  
ibm-949, iso2022kr, euc-kr  
ibm-878, koi8-r

La page de codes par défaut est ignorée pour les documents XML.

Si vous ajoutez des documents XML à un index Ngram, celui-ci doit avoir été créé avec la page de codes UTF8. Si l'index prend en charge les sections, seul l'identificateur de section de niveau feuille est utilisé lors de l'indexation ; la résolution du texte dans les sections imbriquées n'est pas prise en charge.

### Utilisation des formats de documents non pris en charge

Pour les formats de documents non pris en charge, définissez un ID numérique, compris entre 1 et 100. Cette valeur est transmise comme format source à l'exit utilisateur qui convertit le format d'origine en TDS (ASCII à plat).

Si, au cours de l'indexation, DB2 Extension Texte détecte un document dont le type n'est pas pris en charge, elle envoie un exit qui écrit le document sur un disque et appelle un programme chargé de convertir le texte dans l'un des formats pris en charge. L'exit utilisateur doit être enregistré dans le fichier de configuration du serveur (DESSRV.INI) et dans le fichier de configuration du client (DESCL.INI). Modifiez l'option USEREXIT dans la section [DOCUMENTFORMAT] en y indiquant le nom de l'exit utilisateur.

Pour activer l'exit utilisateur, modifiez les fichiers ASCII suivants :

Windows NT et Windows 2000 :  
%DMBMPATH%\instance%\DB2INSTANCE%\db2tx\descl.ini  
%DMBMPATH%\instance%\DB2INSTANCE%\db2tx\txinsnnn\dessrv.ini

UNIX :  
\$DB2TX\_INSTOWNERHOMEDIR/db2tx/descl.ini  
\$DB2TX\_INSTOWNERHOMEDIR/db2tx/txinsnnn/dessrv.ini

Ajoutez-y les instructions suivantes :

```
[DOCUMENTFORMAT]  
USEREXIT=nom-exécutable
```

## Formats de documents

où <nom-exécutable> représente le nom de l'exécutable utilisateur. Vous pouvez spécifier un chemin d'accès complet ou, si l'exécutable utilisateur est stocké dans un répertoire mentionné dans l'instruction PATH, seulement le nom du fichier.

Pour appeler l'exécutable utilisateur, utilisez la syntaxe suivante :

```
<nom-exit-utilisateur>
        -sourcefile      <nom-fichier-source>
        -targetfile     <nom-fichier-cible>
        -sourceccsid    <ccsid-source>
        -targetccsid    <ccsid-cible>
        -sourceformat   <format-source>
        -targetformat   <format-cible>
```

### **nom-fichier-source**

Nom du fichier à convertir par l'exécutable utilisateur. Ce nom de fichier est qualifié complet et se situe dans le répertoire de travail défini soit dans le profil client soit dans l'instance du serveur.

### **nom-fichier-cible**

Nom du fichier contenant le résultat de l'exécutable utilisateur. Ce fichier est ensuite utilisé lors du traitement par DB2 Extension Texte. Ce nom de fichier est qualifié complet et pointe sur le répertoire de travail défini soit dans le profil client, soit dans l'instance du serveur. Les entrées dans le profil client sont utilisées pour l'appel de l'API

EhwGetMatches et celle de l'instance du serveur pour l'appel de l'API EhwUpdateIndex.

### **ccsid-source**

Page de codes du fichier source. Il s'agit de la page de codes par défaut.

### **ccsid-cible**

Page de codes attendue par DB2 Extension Texte. La page de codes est 850.

### **format-source**

Format du fichier source. Il s'agit du format par défaut.

### **format-cible**

Format du fichier attendu par DB2 Extension Texte. Actuellement, seul le format fichier à plat (TDS) ou, pour les index utilisant des sections, ASCIISECTION sont acceptés.

L'exécutable utilisateur doit pouvoir renvoyer les valeurs suivantes :

- 0 La conversion de format a réussi.
- >0 La conversion de format a échoué. Lors de l'indexation, les messages d'erreur sont écrits dans la table des erreurs du document. Vous pouvez afficher les messages d'erreur par la commande `desmsgix`.

---

**Langues**

DB2 Extension Texte a également besoin de savoir dans quelle langue un document est rédigé, afin d'utiliser le dictionnaire approprié pour l'analyse linguistique. Le tableau 5 à la page 54 fournir la liste des paramètres de langue que vous pouvez indiquer lors de l'activation d'une colonne de texte ou de documents externes.

---

**CCSID**

<p><b>Conseil</b></p> <p>Avent d'indiquer un CCSID lors de l'activation d'une colonne de texte, prenez connaissance de la section «Le moyen d'éviter des incidents de pages de codes lors du stockage et de l'activation de texte» à la page 37</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Les documents peuvent être indexés s'ils utilisent l'un des CCSID ci-après.

**Remarque :** Les CCSID 861, 865 et 4946 ne sont pas pris en charge par DB2 UDB. Pour indexer des documents associés à ces CCSID, stockez les documents dans une colonne de type binaire (BLOB ou FOR BIT DATA).

**EBCDIC**

37	Anglais américain, canadien
273	Autrichien, allemand
277	Danois, norvégien
278	Finnois, suédois
280	Italien
284	Espagnol, latino-américain
285	Anglais britannique
297	Français
420	Arabe
424	Hébreu
437	Anglais américain
500	Latin-1 International
871	Islandais
875	Grec

## CCSID

1025	Russe
1112	Lituanien
1122	Estonien
1123	Ukrainien

## ASCII

813 AIX, HP, SUN	Grec
819 AIX, HP, SUN	Latin-1
850 AIX, OS/2	Latin-1
855 OS/2	Bulgare
860 OS/2	Portugais
861, voir la remarque	Islandais
862 OS/2	Hébreu
864 OS/2	Arabe
863 OS/2	Canadien
865, voir la remarque	Danois, norvégien
866 OS/2	Russe
869 OS/2	Grec
915 AIX, OS/2, HP	Russe
916 AIX	Hébreu
921 AIX, OS/2, WIN	Lituanien
922 AIX, OS/2, WIN	Estonien
1064 AIX	Arabe
1089 AIX, HP	Arabe
1124 AIX	Ukrainien
1125 OS/2	Ukrainien
1129	Vietnamien
1130	Vietnamien
1131	Vietnamien
1250 WIN	Croate, Biélorusse
1251 WIN	Russe
1252 WIN	Latin-1

1253 WIN	Tchèque
1255 WIN	Hébreu
1256 WIN	Arabe
1257 WIN	Grec
1258 WIN	Vietnamien

**DBCS**

932 AIX, OS/2	Japonais, combinaison SBCS/DBCS
942 OS/2	Japonais, combinaison SBCS/DBCS
943 OS/2, WIN	Japonais, combinaison SBCS/DBCS
5039 HP	Japonais, combinaison SBCS/DBCS
954 AIX, HP, SUN	Japonais
949 OS/2	Coréen
970 AIX, HP, SUN	Coréen
1363 WIN	Coréen
948 OS/2	Chinois traditionnel, combinaison SBCS/DBCS
950 AIX, HP, OS/2, SUN, WIN	Chinois traditionnel, combinaison SBCS/DBCS
964 AIX, HP, SUN	Chinois traditionnel, combinaison SBCS/DBCS
1381 OS/2, WIN	Chinois simplifié, combinaison SBCS/DBCS
1383 AIX, HP, SUN	Chinois simplifié, combinaison SBCS/DBCS
1386 AIX, OS/2, WIN	Chinois simplifié, combinaison SBCS/DBCS
494, voir la remarque	Latin-1 (CP850)
5039 HP	Japonais

**UNICODE**

1208	UTF8
13488	UCS2

**Le moyen d'éviter des incidents de pages de codes lors du stockage et de l'activation de texte**

Des paramètres de page de codes sont définis dans les éléments suivants :

- L'environnement d'application actif
- Chacun des documents
- Chacune des bases de données DB2
- Chaque index DB2 Extension Texte

## CCSID

Lorsque vous stockez des documents dans une colonne de base de données DB2 de type caractère, VARCHAR et CLOB par exemple, DB2 suppose que tous les documents ont la même page de codes que l'environnement d'application actif et convertit le document de cette page de codes à celle de la base de données. La page de codes de la base de données est soit déjà celle de l'environnement d'application actif (auquel cas il n'y a pas de conversion), soit la page de codes définie lors de la création de la base de données et qui est différente de celle de la page de codes de l'application (et dans ce cas, il y a conversion).

Lorsque vous stockez des données dans une colonne de base de données DB2 de type binaire, BLOB ou FOR BIT DATA, par exemple, DB2 ne convertit pas les données et les documents conservent leur CCSID d'origine.

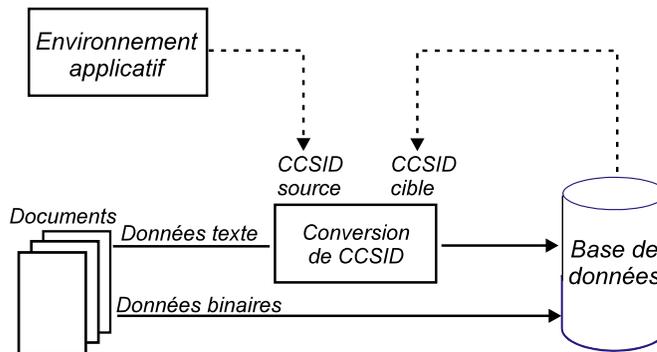


Figure 5. Définition du CCSID d'un document de base de données par DB2

Lorsque vous activez une colonne de texte en vue de son utilisation par DB2 Extension Texte, autrement dit, lorsque vous exécutez la commande `ENABLE TEXT COLUMN` pour créer un index pour la recherche, la page de codes est définie soit par celle de la base de données (par défaut), soit à la valeur par défaut en cours qui peut être définie par la commande `CHANGE TEXT CONFIGURATION`, soit à la page de code que vous avez défini dans la commande `ENABLE TEXT COLUMN`.

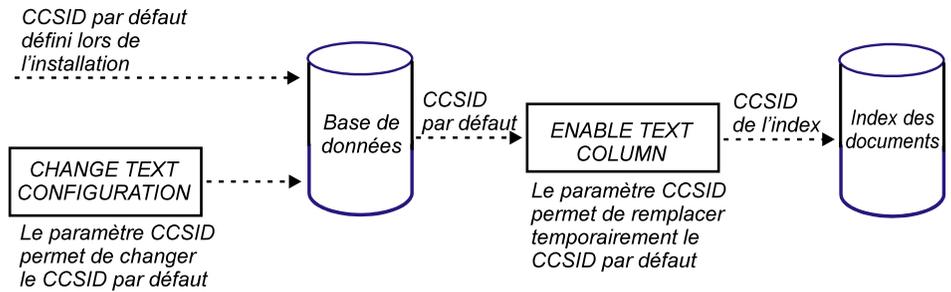


Figure 6. Définition du CCSID d'un index de texte par DB2 Extension Texte

Lors de la recherche, le CCSID de la base de données est utilisé pour interpréter le CCSID de chaque chaîne de recherche.

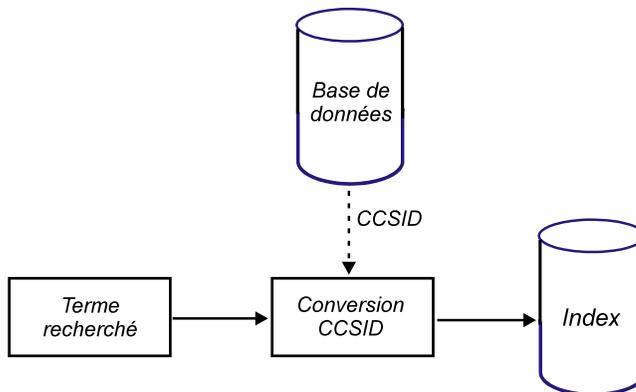


Figure 7. Recherche utilisant le CCSID de la base de données

Le paramétrage de la page de codes peut être vérifié comme suit :

- Pour vérifier la page de codes de la base de données, exécutez la commande DB2 :  
db2 get db cfg
- Pour vérifier la page de codes d'index par défaut, exécutez la commande DB2 Extension Texte :  
db2tx get text cfg

Pour éviter des incidents de pages de codes, il est important d'en coordonner correctement le paramétrage :

- **Exemple 1** : Environnement 850, Document 850, Base de données 850, Index 850

**Correct.** DB2 suppose, à juste titre, que le document a la même page de codes que l'environnement et, comme la base de données a la même page de codes que l'environnement, DB2 n'effectue aucune conversion et stocke

le document avec la page de codes 858 dans la base de données. Lorsque vous activez la colonne de texte en prenant le paramétrage de page de codes par défaut, (850, la page de codes de la base de données), le document est indexé correctement avec l'index de la page de codes 850.

- **Exemple 2** : Environnement 1252, Document 1252, Base de données 850, Index 850

**Correct.** DB2 suppose, à juste titre, que le document a la même page de codes que l'environnement et effectue une conversion des pages de code 1252 à 850, lors du stockage du document dans la base de données. Lorsque vous activez la colonne de texte en prenant le paramétrage de page de codes par défaut, (850, la page de codes de la base de données), le document est indexé correctement avec l'index de la page de codes 850.

- **Exemple 3** : Environnement 1252, Document 850, Base de données 850, (Index ANY)

**Erreur.** DB2 suppose indûment que le document a la même page de codes que l'environnement (page de codes 1252) et effectue une conversion inappropriée entre pages de codes 1252 et 850 lors du stockage du document dans la base de données.

- **Exemple 4** : Environnement 1252, Document 850, Base de données 1252, Index 850

**Correct.** DB2 suppose indûment que le document a la même page de codes que l'environnement mais, du fait que cette dernière est identique à la page de codes de la base de données, DB2 n'effectue pas de conversion et stocke le document dans la page de codes 850 de la base de données. Lorsque vous activez la colonne de texte, vous devez toutefois indiquer la page de codes 850 pour le document de sorte que celui-ci soit correctement indexé avec l'index de la page de codes 850.

- **Exemple 5** : Environnement 1252, Document 1252, Base de données 850, Index 1252

**Erreur potentielle.** DB2 suppose, à juste titre, que le document a la même page de codes, 1252, que l'environnement et convertit le document en page de codes 850 lorsqu'il le stocke dans la base de données. L'erreur potentielle se produit si vous indiquez la page de codes d'origine (1252) du document lorsque vous activez la colonne de texte avec l'index de la page de codes 850. L'opération correcte consisterait à utiliser le paramétrage de page de codes par défaut 850 (page de codes de la base de données).

---

## Types de recherche

Vous pouvez affecter l'un des types d'index et diverses options à une colonne contenant du texte à rechercher : *linguistique*, *exact* et *Ngram*. Vous devez décider du type d'index à créer avant de préparer ce type de colonne pour DB2 Extension Texte. Pour obtenir une description plus détaillée de la façon

dont chaque type d'index affecte l'analyse linguistique, consultez le «Chapitre 15. Traitement linguistique des index linguistiques et exacts» à la page 255.

### Résumé

- Pour **rechercher des variantes linguistiques d'un mot**, utilisez un **index linguistique**.  
Recherche les variantes d'un mot en se fondant sur la normalisation et la lemmisation ainsi que sur l'utilisation d'un dictionnaire ; c'est la méthode qui utilise le moins d'espace disque.
- Pour **rechercher des correspondances totales**, utilisez un **index exact**.  
Retrouve le terme exactement comme il a été entré ; l'indexation et la recherche sont plus rapides ; cette méthode utilise plus d'espace disque. Si l'option NORMALIZED est indiquée, les recherches font la distinction entre majuscules et minuscules.
- Pour **rechercher des variantes de caractères**, utilisez un **index Ngram**.  
Retrouve les mots même en cas de fautes d'orthographe ; si l'option CASE\_ENABLED est indiquée pour permettre une recherche faisant la distinction entre majuscules et minuscules, l'index a besoin de plus d'espace disque et les recherches peuvent durer plus longtemps.
- Pour **rechercher dans des documents DBCS**, utilisez un **index Ngram**.  
C'est la seule possibilité pour des documents DBCS, mais elle peut également être utilisée pour des documents SBCS de type TDS.

DB2 Extension Texte propose de nombreuses options de recherche, qui dépendent du type d'index utilisé. Consultez le tableau 8 à la page 223, et le tableau 9 à la page 223, avant de décider du type d'index à utiliser.

## Recherche linguistique

Dans le cas d'un index linguistique, les opérations linguistiques s'exécutent lors de l'analyse de chaque texte de document en vue de son indexation. Cela signifie que les mots sont réduits à leur forme de base avant d'être stockés dans un index. Par exemple, le terme "chevaux" est stocké sous cheval.

Lors d'une requête portant sur un index linguistique, c'est la même analyse linguistique qui s'applique aux termes recherchés avant de lancer la recherche dans l'index de texte. Ainsi, si vous recherchez le terme "chevaux", il est réduit à sa forme de base, cheval, avant le début de la recherche. Le tableau 18 à la page 256, résume la façon dont les termes sont extraits en vue de leur indexation lors de l'utilisation d'un index linguistique.

## Types de recherche

L'avantage de ce type d'index est que toute variante du terme recherché retrouvée dans l'un des documents texte indexés est prise en compte. Le terme cheval permettra de retenir les termes "cheval", "chevaux", "CHEVAUX" (majuscules), etc. De même, le terme Chevaux correspondra à ces mêmes termes.

Ce type d'index requiert un espace disque minimal. Cependant, l'indexation et la recherche peuvent prendre plus de temps que pour un index de type exact.

Les types d'analyse linguistique disponibles dépendent de la langue du document. La liste des types existants est la suivante :

- Séparation des mots et des phrases.
- Analyse des débuts de phrase.
- Suppression des tirets.
- Normalisation des termes sous une forme standard sans majuscules dans laquelle les lettres accentuées telles que "ü" sont remplacées par des lettres non accentuées. Par exemple, le mot allemand "Tür" (porte) est indexé sous la forme tuer.
- Réduction des termes à leur forme de base. Par exemple, "acheté" est indexé sous la forme acheter et "chevaux" sous la forme cheval.

### Conseil

Les fragments de mots (mots masqués comprenant des caractères génériques) ne peuvent pas être réduits à leur forme de base. Ainsi, si vous effectuez une recherche sur cheva%, vous ne trouverez pas le mot "chevaux" car il est réduit à sa forme de base dans l'index. Pour le trouver, vous devez effectuer la recherche sur cheva%.

Les variantes de mots mal orthographiés ne peuvent être ramenées à leur forme de base. Prenons, par exemple, le mot allemand röstete qui est indexé correctement dans sa forme de base, le verbe rösten. Un terme röstete ou rösteten recherché est ramené correctement à la forme de base rösten, et le terme est localisé dans l'index. Un terme rostete ne peut ramené à rösten, et le terme n'est pas localisé dans l'index.

- Décomposition des mots au cours de laquelle les mots composés tels que le mot allemand "Wetterbericht" (bulletin météo) sont indexés sous la forme wetterbericht, ainsi que sous les formes wetter et bericht.
- Filtrage des mots vides par lequel les termes non significatifs ne sont pas indexés. La phrase "Un rapport sur tous les animaux" est indexée sur les termes rapport et animaux.

- Filtrage partiel, qui est similaire au filtrage des mots vides, au cours duquel seuls les noms, les verbes et les adjectifs sont indexés. La phrase “Je roule rapidement avec ma voiture” est indexée sur les termes rouler et voiture. Les mots “Je”, “avec” et “ma” sont supprimés en tant que mots vides et l’adverbe “rapidement” est supprimé par le filtrage partiel.

### Recherche exacte

Dans un index exact, les termes des documents texte sont indexés sous la forme exacte sous laquelle ils figurent dans le document. Par exemple, le terme cheval permet de trouver “cheval” mais pas “chevaux”, ni “Cheval” car la recherche dans un index exact prend en compte les majuscules et les minuscules.

Dans une requête, la même analyse s’applique aux termes de la requête, qui sont ensuite comparés aux termes trouvés dans l’index. Cela signifie que les termes trouvés sont strictement identiques aux termes recherchés. Vous pouvez utiliser des caractères génériques pour élargir la recherche ; par exemple, la recherche du terme index% permet de trouver “indexation”, “indexé”, etc.

Le tableau 19 à la page 256, illustre la façon dont des termes sont extraits du texte pour être insérés dans un index exact.

L’avantage de ce type d’index est que la recherche est plus précise, et que l’indexation, ainsi que l’extraction, sont plus rapides. Etant donné que toutes les formes et orthographes d’un terme sont indexées, vous devez disposer de plus d’espace disque que pour un index linguistique.

Les analyses linguistiques effectuées pour indexer des documents texte dans un index exact sont les suivantes :

Séparation des mots et des phrases

Filtrage des mots vides.

### Recherche ou recherche de correspondance partielle dans des documents DBCS

Un index Ngram analyse des ensembles de caractères. L’analyse ne se base pas sur un dictionnaire.

Si le texte contient des caractères DBCS, utilisez un index Ngram. Aucun autre type d’index ne prend en charge les caractères DBCS.

Ce type d’index prend en charge les recherches de correspondance “ partielle ”, ce qui signifie qu’il permet de trouver des chaînes de caractères similaires au terme fourni. Par exemple, rechercher Extension permet de trouver un mot similaire mal orthographié, comme Extensino. Il est aussi possible de préciser le degré de similarité requis.

## Types de recherche

**Remarque :** Même dans le cas d'une recherche de correspondance partielle, au moins les trois premiers caractères doivent correspondre.

Dans un index Ngram, pour effectuer une recherche faisant la distinction entre majuscules et minuscules, il ne suffit pas de préciser dans la requête le mot-clé `PRECISE FORM OF`. Car, normalement, un index Ngram ne fait pas de distinction entre les majuscules et les minuscules sur les caractères indexés. Cependant, pour le forcer à faire cette distinction, utilisez l'option `CASE_ENABLED`, au moment de la création de l'index. Ensuite, dans votre requête, précisez le mot-clé `PRECISE FORM OF`.

Avec l'option `CASE_ENABLED`, l'index occupe plus de place et les recherches prennent quelquefois plus de temps.

Veuillez vous reporter à la section «CCSID» à la page 35 pour connaître la liste des CCSID pris en charge par les index Ngram. Un index Ngram accepte une liste de CCSID natifs. Pour tous les autres CCSID, les données sont mappées de ce CCSID à UTF8.

Bien que les index de type Ngram aient été conçus pour l'indexation de documents DBCS, ils peuvent également être utilisés dans le cas de documents SBCS. Cependant, ils ne prennent en charge que les documents TDS.

Les différentes syntaxes de recherche possibles ne sont pas toutes prises en charge. Reportez-vous au récapitulatif des règles et des restrictions dans le «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

### Modification du type d'index

Si vous constatez que le type d'index utilisé ne vous convient pas, vous devez d'abord supprimer l'index en désactivant la colonne de texte ou la table de texte ; vous devez ensuite recréer l'index en réactivant la colonne de texte ou la table de texte.

---

## Création d'un ou de plusieurs index de texte pour une table

Le «Chapitre 5. Préparation d'un texte pour la recherche» à la page 57, décrit comment préparer des tables afin de pouvoir y rechercher du texte. Avant d'effectuer cette préparation, vous devez toutefois choisir entre la création d'un index de texte commun à toutes les colonnes de type texte indexées de la table ou la création de plusieurs index de textes, un par colonne de texte indexée. Une table comportant un index distinct pour chaque colonne de texte est appelée *table à index multiples*.

### Conseil

Si vous avez l'intention d'indexer des fichiers externes (reportez-vous à la section «Activation de fichiers de texte externes» à la page 73), la table associée doit être une table à index multiples.

L'utilisation d'index multiples présente les avantages suivants :

- Création d'un type d'index distinct pour chaque colonne de texte  
Ceci vous offre une grande souplesse quant aux autres caractéristiques associées à une colonne de type texte (moment de la mise à jour périodique de l'index et répertoire dans lequel il est stocké). Pour obtenir une description de ces caractéristiques, reportez-vous à la section «ENABLE TEXT COLUMN» à la page 145.
- Indexation de colonnes à différents moments  
L'indexation est une activité pouvant consommer du temps et des ressources. Le fait de disposer d'une table à index multiples permet d'étaler cette activité dans le temps en indexant les colonnes à des moments différents.

Si la souplesse offerte par une table à index multiples s'avère inutile, utilisez un index commun qui facilitera la gestion de DB2 Extension Texte ; lorsque vous activez une table de texte, vous définissez les paramètres d'indexation qui seront utilisés comme valeurs par défaut pour toutes ses colonnes de texte. De même, si vous voulez désactiver les colonnes, vous pouvez le faire par une seule commande qui désactive la table de texte.

---

### Calcul de la taille d'un index

L'espace disque nécessaire pour un index dépend de la taille et du type de données à indexer, ainsi que du type d'index. Les documents texte écrits au moyen d'un logiciel de traitement de texte ont besoin de moins de place, une grande partie de leur contenu étant occupée par des caractères de contrôle. Nous vous conseillons de réserver un espace disque équivalant à 0,7 fois la taille des documents à indexer, puis de multiplier cette valeur par 2 pour réserver un espace temporaire permettant de réorganiser l'index. Pour un index Ngram, vous aurez besoin de deux fois plus d'espace disque.

Si vous avez plusieurs index volumineux, mémorisez-les sur des unités disques différentes, spécialement si vous avez des accès concurrents aux index au moment de leur mise à jour ou des recherches.

## Mise à jour d'un index

### Mise à jour d'un index

Lorsqu'un document texte est ajouté à une base de données, ou modifié dans une base de données, il doit être indexé afin que le contenu de l'index soit synchronisé avec celui de la base de données. De même, lorsqu'un document texte est supprimé d'une base de données, ses termes doivent être supprimés de l'index.

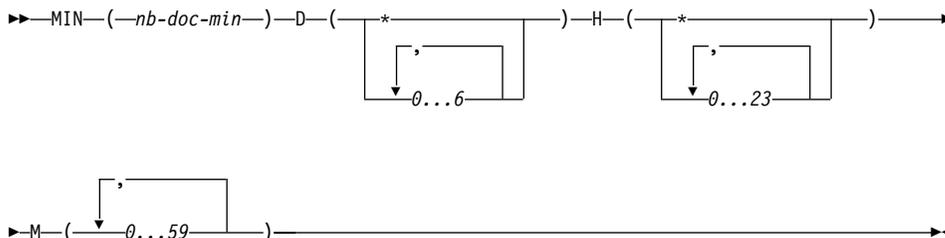
Les informations relatives à l'ajout, la modification ou la suppression de documents sont automatiquement stockées par les déclencheurs dans une table de journalisation. Les documents figurant dans cette table sont indexés lors de la mise à jour suivante d'un index.

La commande UPDATE INDEX permet de mettre à jour à jour un index immédiatement, à la demande. Cependant, en règle générale, vous mettre automatiquement à jour un index à des intervalles définis dans la variable d'environnement DB2TXUPDATEFREQ. C'est la variable d'environnement qui détermine les valeurs par défaut. Ces dernières peuvent être ignorées lorsque vous créez un index à l'aide des commandes ENABLE TEXT COLUMN ou ENABLE TEXT TABLE. La fréquence de mise à jour d'un index existant peut être modifiée via la commande CHANGE INDEX SETTINGS.

Vous devez définir la fréquence de mise à jour d'index en indiquant le moment auquel cette mise à jour doit être effectuée et le nombre minimal de documents texte devant être mis en file d'attente. Si la table de journalisation ne contient pas suffisamment de documents au jour et à l'heure prévus pour la mise à jour, l'index n'est pas mis à jour.

Vous devez planifier soigneusement l'indexation périodique car cette opération consomme du temps et des ressources. Le temps nécessaire dépend de nombreux facteurs, tels que le nombre de documents texte ajoutés ou modifiés depuis la précédente mise à jour d'index, la taille des documents et la puissance du processeur.

### Syntaxe



### **MIN nb-doc-min**

Nombre minimal de documents texte devant être mis en file d'attente pour que l'index soit mis à jour.

**D** Jour(s) de la semaine choisi(s) pour la mise à jour de l'index.

\* Tous les jours

0 Dimanche

1 Lundi

2 Mardi

3 Mercredi

4 Jeudi

5 Vendredi

6 Samedi

**H** Heure(s) de mise à jour de l'index.

\* Chaque heure

0...23 A l'heure indiquée

**M** Minute(s) s'ajoutant à l'heure de mise à jour de l'index.

0...59 A la minute indiquée

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

Si, à 12h00 ou 15h00, du lundi au vendredi, il y a eu au moins 100 documents texte en attente, l'index est mis à jour.

Vous pouvez combiner plusieurs spécifications de fréquence.

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

La mise à jour de l'index est prévue du lundi au vendredi, à 12h00 et 15h00 comme précédemment, mais, chaque jour à 22h00, l'index est mis à jour même s'il n'y a qu'un seul document texte dans la table de journalisation.

---

## Gestion des index dans un environnement multi-noeud

Si la colonne de texte que vous activez se trouve dans une table appartenant à un groupe de noeuds, le répertoire d'index spécifié doit être disponible sur tous les noeuds physiques. Si vous utilisez le répertoire par défaut spécifié dans la configuration de texte, assurez-vous que le chemin d'accès est

## Gestion des index dans un environnement multi-noeud

accessible à partir de tous les noeuds du groupe. Si ce n'est pas le cas, vous pouvez définir un chemin spécifique pour chaque noeud dans la commande `ENABLE TEXT COLUMN`.

Si vous modifiez la configuration d'un groupe de noeuds qui contient une table activée pour DB2 Extension Texte, vous devrez réindexer la table. Pour vérifier si la configuration des noeuds a été modifiée, utilisez la commande `TXNCHECK`.

### Conseil

Si vous utilisez DB2 UDB Extended Enterprise Edition et disposez de tables partitionnées sur plusieurs noeuds, il est possible qu'une différence d'heure système existe entre les noeuds (si l'un est en horaire d'été, par exemple). Pour éviter les incidents, vérifiez que ces différences restent dans l'intervalle défini par le paramètre `MAX_TIME_DIFF` de DB2. Sous Windows NT et Windows 2000, utilisez la commande `NET TIME` pour synchroniser les paramètres d'heures système.

---

## Utilisation de documents structurés (prise en charge de sections)

La prise en charge de sections vous permet d'indexer et de rechercher des sections particulières dans un document structuré, par exemple dans le titre, l'auteur ou la description. Les documents peuvent être au format XML ou HTML ou au format fichier à plat avec des marques de type HTML. Vous définissez les marques de mise en forme et les noms de section correspondants dans un *modèle de document*. Celui-ci définit les sections des documents qui sont indexées, dont utilisables pour la recherche. Une section porte un nom descriptif utilisé dans les requêtes portant sur cette section.

Pour que la prise en charge des sections soit disponible, vous devez spécifier `INDEXPROPERTY SECTIONS_ENABLED` au moment où vous activez la colonne de texte contenant le document.

Un *fichier de modèles de documents* contient tous les modèles de documents définis pour l'instance serveur. Lors de la création d'une instance serveur, un exemple de fichier de modèles de documents `DESMODEL.INI` est créé dans le répertoire de l'instance serveur. Ce fichier est en page de codes ASCII .

Les informations relatives au modèle de document sont copiées dans le répertoire d'index. Si vous modifiez le fichier de modèles de documents pour l'instance serveur après avoir créé l'index, cela n'affecte pas la prise en charge de sections pour l'index créé.

## Documents structurés (prise en charge de sections)

Une recherche sur un index avec prise en charge de sections, portant sur McDaniel dans la section Author, pourrait se présenter comme ci-dessous. La section, ici Author, est toujours précédée du nom du modèle.

```
db2 "select count (*)
      from db2tx.htmltable
      where db2tx.contains(mon-descripteur,'MODEL mon-modèle-html SECTION (author) "McDaniel") = 1
```

### Sections avec attribut

Les sections peuvent être de types différents. Celles qui ne comportent que du texte ne sont associées à aucun type. Les sections ayant un type déclaré sont appelées *sections avec attribut*. Les types d'attributs admis sont :

- DATE
- TIME
- FLOAT
- INTEGER

Vous pouvez utiliser des plages de valeurs pour effectuer des recherches dans des documents contenant des sections avec attribut. La section doit être associée au type approprié. Dans l'exemple qui suit, la section ABC doit être de type TIME pour que la requête puisse être prise en charge :

extraction de tous les documents contenant des valeurs comprises entre 20:00:00 et 22:00:00 dans la section ABC

Plusieurs valeurs d'attribut peuvent être associés à une même session. Séparez les valeurs par des espaces ou de retour ligne. Les sections avec attribut ne peuvent pas être imbriquées.

### Formats de date pris en charge

Vous pouvez inclure des dates, par exemple la date de création ou la date de publication, dans la liste des sections prises en charge dans un modèle de document. Les formats de date suivants sont acceptés :

Tableau 3. Formats de date pris en charge

Nom format	Format	Exemple
ISO	YYYY-MM-DD	1986-12-25 100-12-25 10123-2-17
	DD-MM-YYYY	01-10-1999 5-8-1990
Standard IBM USA (USA)	MM/DD/YYYY	08/18/1999 5/13/1977
	YYYY/MM/DD	1999/08/18
Standard IBM Européen (EUR)	DD.MM.YYYY	20.10.1999 1.5.1980

## Documents structurés (prise en charge de sections)

La recherche sur un index avec prise en charge des sections pour tous les documents publiés à une date déterminée pourrait se présenter comme suit :  
dessrch -s SERVER -x TESTIX -section sample/Date-publication -term 1999-09-09

### Formats d'heure pris en charge

Les formats d'heure suivants sont acceptés :

Tableau 4. Formats d'heure pris en charge

Nom format	Format	Exemple
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

Pour les heures, vous pouvez ne pas indiquer les zéros non significatifs. Les valeurs se situant hors de plages suivantes ne sont pas admises :

Heures	Minutes	Secondes
0–23	0–59	0–59

### Formats en virgule flottante pris en charge

Les formats suivants sont acceptés :

- Une suite de chiffres comprise entre 1.175494351e–38F et 3.402823466e+38F
- Une notation scientifique utilisant e ou E.

Séparez les valeurs en virgule flottante par des espaces ou de retour ligne.

Voici quelques exemples de formats corrects et incorrects pour des valeurs en virgule flottante :

Format correct	Format incorrect
1000	1 000
1	1,000
1.0e3	e3
1.0E3	
1E3	
.1E4	
0.1E4	

## Documents structurés (prise en charge de sections)

### Formats de nombres entiers pris en charge

Les nombres entiers compris entre -2147483648 et 2147483647 sont pris en charge. Séparez les entiers par des espaces. Dans une section de type INTEGER, la chaîne "42 234" est résolue en valeurs "42" et "234" et non pas en "42234".

Voici quelques exemples de formats corrects et incorrects pour des valeurs de type INTEGER :

Format correct	Format incorrect
1000	1.000 1,000 1 000
1000000	100.000 100,00 100 000

#### Remarque

Les séparateurs propres à chaque langue ainsi que les formats monétaires ne sont pas pris en charge.

### Fichiers à plat et documents HTML

Pour les fichiers à plat, les sections sont marquées par des marques de type HTML, telles que <title> et <subject>. Un document avec des sections marquées peut se présenter comme suit :

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

Un fichier de modèles de documents pour des fichiers à plat ou des documents HTML peut se présenter comme ci-dessous. Les noms de modèles et les noms de section ne peuvent comporter que les lettres A à Z, a à z et les chiffres 0 à 9. Dans les noms de modèles, la distinction est faite entre majuscules et minuscules. Dans les noms de section, cette distinction peut ou non être faite en fonction du paramètre que vous avez défini lors de la création du modèle.

```
;liste des modèles de documents
;un modèle commence toujours par 'modelname' et le nom du modèle
[MODELS]
modelname=sample
modelname=sample2
modelname=sample3
```

```
; définition de modèle de document 'sample'
; à gauche - identificateur du nom de section
```

## Documents structurés (prise en charge de sections)

```
; à droite - marque du nom de section
[sample]
Title = title
Author = author
Subject = subject
Abstract = abstract
Content = content
PublishingDate=date

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

Si un document contient une section marquée qui n'est pas définie dans le modèle de document, le contenu de la section est intégré à la section précédente définie pour l'indexation et la recherche. Un document, par exemple, contient les sections marquées suivantes :

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

Le modèle de document, book, est défini par :

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

La section <subject> ne figure pas dans le modèle de document book. Lors de l'indexation du document, le contenu de la section subject est indexé au moyen du contenu de la section title. Il peut également être utilisé pour la recherche dans la section title.

Si lors de la création de l'index, vous avez défini une liste de modèles, le modèle par défaut est le premier de la liste. Vous pouvez changer de modèle par défaut par la commande `desmodix`.

## Documents XML

Pour les index à activation de section, les documents XML doivent être correctement structurés et contenir un élément racine. Le nom de cet élément

## Documents structurés (prise en charge de sections)

doit être identique à l'un des noms de modèles défini et être comme lui en majuscules et/ou minuscules. La définition donnée dans le fichier de modèles de documents doit être un sous-ensemble du modèle de document défini dans le fichier DTD (définition de type de document) pour ce document.

La description de l'élément doit commencer par l'élément racine. Pour chaque élément XML que vous voulez utiliser en tant que section, vous devez en indiquer l'arborescence complète dans la description du modèle. S'il s'agit d'une section de type date, elle doit se situer au niveau feuille dans l'arborescence du modèle de document. L'imbrication des sections d'attributs n'est pas prise en charge.

Une description de modèle pour des documents XML peut se présenter comme suit :

```
; liste de modèles de documents
[MODELS]
modelname = LETTER
; exemple de définition de modèle de document XML
; à gauche = identificateur de nom de section codant la totalité du chemin d'accès
; à droite = marques de nom de section indiquant une marque pour chaque
;           élément du chemin d'accès dans l'arborescence jusqu'au
;           noeud indiqué. Le délimiteur de marque est /.
[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 document XML peut se présenter comme indiqué ci-après. Cet exemple montre également comment sont indexées les sections qui ne sont pas définies dans le modèle.

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

<LETTER>
  <HEADER>Cette marque n'ayant pas été indiquée dans la définition, ce texte sera
    ajouté à la section nommée LETTER

</HEADER>
  <DATE>
    01.01.2000 03.02.2000
  </DATE>

  <ADDRESS>
    Du texte sera ajouté à la section nommée LETTER/address.
    <CITY>
      Du texte sera ajouté à la section nommée LETTER/address/City.
    </CITY>
```

## Documents structurés (prise en charge de sections)

```
</ADDRESS>

<CONTENT>
    Du texte sera ajouté à la section nommée LETTER/Content.

<NOSECTION> Du texte sera ajouté à la section LETTER/Content
    car l'élément NOSECTION n'a pas été défini.
</NOSECTION>
<GREETINGS>
    Du texte sera ajouté à la section nommée LETTER/Content/Greetings.
</GREETINGS>
</CONTENT>
</LETTER>
```

---

## Dictionnaires, listes de mots vides, listes d'abréviations et paramètres de langue

Le tableau 5 indique les langues prises en compte et les noms des fichiers fournis en tant que dictionnaires, listes de mots vides et listes d'abréviations. Les fichiers dictionnaires sont au format binaire et ne peuvent pas être changés. Les fichiers de mots vides et les fichiers d'abréviations (éventuels) ont un format plat et peuvent être modifiés. Si vous modifiez l'un de ces fichiers, veuillez à utiliser la page de codes correspondant à la langue.

Ce tableau indique également les paramètres de langue que vous devez indiquer lors de l'activation d'une colonne de texte ou de documents externes. Il précise à DB2 Extension Texte dans quelle langue sont écrits les documents, afin d'utiliser le dictionnaire approprié pour l'analyse linguistique.

Tableau 5. Fonctions linguistiques utilisées pour les langues prises en charge

Langue	Nom du fichier	Paramètre LANGUAGE	Page de codes
Arabe	arabic	ARABIC	864
Portugais (Brésil)	brazil	BRAZILIAN	850
Français canadien	canadien	CAN_FRENCH	850
Catalan	catala	CATALAN	850
Danois	dansk	DANISH	850
Néerlandais	nederlnd	DUTCH	850
Finnois	suomi	FINNISH	850
Français	français	FRENCH	850
Allemand	deutsch	GERMAN	850
Hébreu	hebrew	HEBREW	862
Islandais	islensk	ICELANDIC	850
Italien	italiano	ITALIAN	850
Norvégien (Bokmal)	norbook	BM_NORWEGIAN	850

## Dictionnaires, listes de mots vides et listes d'abréviations

Tableau 5. Fonctions linguistiques utilisées pour les langues prises en charge (suite)

Langue	Nom du fichier	Paramètre LANGUAGE	Page de codes
Norvégien (Nynorsk)	norntn	NN_NORWEGIAN	850
Portugais	portugal	PORTUGUESE	850
Russe	russian	RUSSIAN	866
Espagnol	espana	SPANISH	850
Suédois	svensk	SWEDISH	850
Suisse allemand	dschweiz	SWISS_GERMAN	850
Thaïlandais	thai	THAI	850
Anglais britannique	uk	UK_ENGLISH	850
Anglais américain	us	US_ENGLISH	850

Les fichiers se différencient par leur extension :

Contenu	Extension
Dictionnaire	DIC
Liste de mots vides	STW
Liste d'abréviations	ABR

### Modification des fichiers de mots vides et d'abréviations

Il existe un fichier de mots vides et un fichier d'abréviations par langue prise en charge. Pour connaître les répercussions engendrées par la modification de ces fichiers, reportez-vous à la section «Nécessité d'indexation des documents texte» à la page 29.

#### Conseil

Avant de modifier l'un de ces fichiers, faites une copie de sauvegarde.

Sous OS/2 et Windows, les fichiers de mots vides et d'abréviations se trouvent dans :

*unité*:\dmb\db2tx\dict

Sous AIX, HP-UX et SUN-Solaris, ils sont dans :

*DB2TX\_INSTOWNER*/db2tx/dicts

## **Dictionnaires, listes de mots vides et listes d'abréviations**

Utilisez votre éditeur de texte pour les modifier. Ces fichiers utilisent le CCSID 850 ; assurez-vous, avant de commencer, que le paramètre CCSID de votre application est bien défini avec la valeur 850.

Retirez les mots et les abréviations que vous voulez indexer. A l'inverse, ajoutez les mots que vous voulez soustraire à l'indexation.

---

## Chapitre 5. Préparation d'un texte pour la recherche

Le «Chapitre 3. Mise en route» à la page 25 vous aide à vous familiariser avec la préparation d'un texte à la recherche via DB2 Extension Texte en déroulant un exemple simple. Ce chapitre décrit plus en détail la préparation d'un texte à la recherche et tous les points qu'il faut prendre en compte avant de commencer.

Les étapes de préparation d'un texte à la recherche sont :

1. Une préparation minutieuse
2. Démarrage de l'interpréteur de commandes DB2 Extension Texte
3. Connexion à une base de données
4. Activation d'une base ed données pour la recherche de texte
5. Activation d'une table de texte pour la recherche de texte (inutile si vous créez un index par colonne de texte)
6. Activation d'une colonne de texte pour la recherche de texte

---

### Préparation préalable

#### Conseil

Lisez attentivement cette section. Elle détaille les informations que vous devez connaître *avant* de préparer votre texte pour la recherche.

- **Pour les utilisateurs UNIX, définition de leur profil**

Si vous n'êtes pas propriétaire d'une instance DB2 Extension Texte, reportez-vous à la section «Configuration» à la page 17, qui explique comment configurer votre profil.

- **Création d'un seul index pour toute la table ?**

Vous devez déterminer si vous souhaitez créer un seul index pour une table toute entière, ou un index distinct pour chaque colonne de texte. La section «Création d'un ou de plusieurs index de texte pour une table» à la page 44 vous aidera à en décider.

- **Bonne connaissance des documents**

Lorsque vous préparez des documents pour la recherche, vous devez préciser le CCSID, la langue et le format du texte. Pour plus d'informations, reportez-vous à la section «Chapitre 4. Planification de vos besoins en matière de recherche» à la page 29.

## Préparation préalable

- **Choix du type d'index de texte nécessaire**

Le type d'index dont vous avez besoin est fonction de la nature de recherches que vous faites (correspondance, totale, partielle, etc) et du type des documents, SBCS ou DBCS. Pour plus de détails, reportez-vous au «Chapitre 4. Planification de vos besoins en matière de recherche» à la page 29.

- **Lieu de stockage des index**

Lors de la préparation de documents pour la recherche, DB2 Extension Texte crée un index de texte. Vous devez indiquer le répertoire de stockage de cet index. Vérifiez que vous disposez bien de l'espace disque nécessaire (reportez-vous à la section «Calcul de la taille d'un index» à la page 45).

- **Vérification du nom de base de données par défaut**

Le nom par défaut de la base de données dans la variable d'environnement DB2 DB2DBDFT (reportez-vous à la section «Variables d'environnement» à la page 17) est celui qu'utilise DB2 Extension Texte si vous n'en indiquez pas lors de la préparation du texte à la recherche.

- **Définition de la configuration du texte**

La configuration du texte détermine le paramétrage par défaut du CCSID de l'index, de la langue et du format du document, du type d'index, la fréquence de mise à jour de l'index, le nom d'espace table et le répertoire de l'index.

Vous pouvez ignorer ces paramètres lors de la préparation du texte pour la recherche, mais il est plus commode de disposer de valeurs par défaut définies correctement préalablement. Les paramètres de configuration de texte initiaux, définis lors de l'installation de DB2 Extension Texte, sont décrits à la section «Définitions de la configuration de texte» à la page 19. Pour modifier les paramètres de l'installation et définir vos propres valeurs par défaut, reportez-vous à la commande «CHANGE TEXT CONFIGURATION» à la page 133.

- **Activation de la prise en charge de section**

Si vous devez limiter les recherches à une section déterminée d'un document, prenez connaissance de la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48 pour savoir comment définir les modèles dans le fichier de modèles de documents.

- **Modifications des listes de mots vides et d'abréviations**

Reportez-vous aux sections «Nécessité d'indexation des documents texte» à la page 29 et «Modification des fichiers de mots vides et d'abréviations» à la page 55 pour appréhender le concept de listes de *mots vides* et de listes d'abréviations et déterminer si vous souhaitez les modifier avant de lancer l'indexation.

Après avoir recueilli les informations et pris les décisions décrites dans la présente section, vous êtes en mesure de préparer votre texte à la recherche.

### Démarrage de l'interpréteur de commandes DB2 Extension Texte

Résumé	
<b>Utilisation</b>	Facultatif. Au début de chaque session
<b>Commande</b>	db2tx
<b>Droits</b>	Indéterminés

Vous pouvez entrer les commandes client DB2 Extension Texte à l'invite du système d'exploitation. Celles-ci sont similaires aux commandes DB2, à ceci près qu'elles débutent par le préfixe db2tx et non db2.

Vous pouvez également, plutôt que de faire précéder chaque commande DB2 Extension Texte de db2tx, lancer l'interpréteur de commandes DB2 Extension Texte. Cette solution présente l'avantage de ne pas interrompre votre connexion à la base de données après chaque commande ; vous restez connecté. (En l'absence d'interpréteur de commandes, vous êtes automatiquement reconnecté lors de chaque exécution d'une commande db2tx, mais vous devez alors vérifier que le nom de la base de données à laquelle vous voulez vous connecter est bien définie dans la variable d'environnement DB2DBDFT.)

#### 1. Windows seulement : Lancez l'interpréteur de commandes DB2

Entrez la commande suivante à l'invite du système d'exploitation :

```
=>db2 cmd
```

#### 2. Lancez l'interpréteur de commandes DB2 Extension Texte

Sous UNIX, entrez la commande suivante à l'invite du système d'exploitation :

```
=>db2tx
```

Sous Windows, entrez la commande suivante à l'invite de DB2 :

```
db2=>db2tx
```

L'invite db2tx s'affiche :

```
db2tx=>
```

et toutes les commandes qui suivent sont interprétées en tant que commandes DB2 Extension Texte.

Pour quitter ce mode, entrez :

```
db2tx=>quit
```

## Démarrage de l'interpréteur de commandes DB2 Extension Texte

Pour Windows 95 et Windows 98, vous pouvez utiliser la commande db2tx comme indiqué plus haut ou cliquer sur le bouton Démarrer, pointer sur Programmes, puis sur Client DB2 pour Windows 95/98, et enfin cliquer sur l'icône Fenêtre de commande DB2.

Si vous omettez cette étape, vous pouvez exécuter des commandes de DB2 Extension Texte directement à partir de l'invite du système d'exploitation en les faisant précéder de db2tx. Voici un exemple de commande lancée à partir de l'invite du système d'exploitation :

```
=>db2tx enable bdd
```

### Conseil

Si vous travaillez à partir d'un client, vous devez être dans l'environnement de commandes de DB2 Extension Texte et entrer l'instruction complète de connexion en précisant votre ID utilisateur et votre mot de passe.

```
db2tx => connect to nom-bdd user ID-utilisateur using mot-de-passe
```

## Aide sur l'interpréteur de commandes

Pour afficher une liste de commandes, entrez :

```
db2tx ?
```

Pour afficher la syntaxe d'une commande spécifique, entrez :

```
db2tx ? commande
```

Par exemple :

```
db2tx ? CHANGE TEXT CONFIGURATION
```

---

## Connexion à une base de données

### Résumé

<b>Utilisation</b>	Facultatif. Permet de démarrer une session avec une base de données autre que celle par défaut indiquée dans la variable d'environnement DB2DBDFT.
<b>Commande</b>	CONNECT
<b>Droits</b>	CONNECT sur la base de données

Avant de pouvoir lancer d'autres commandes dans une session DB2 Extension Texte, vous devez être connecté à une base de données. Si vous lancez une

commande sans être connecté à une base de données, DB2 Extension Texte vous connecte à la base de données par défaut spécifiée dans la variable d'environnement DB2DBDFT. Vous pouvez vous connecter explicitement à une base de données par la commande CONNECT TO de DB2 Extension Texte.

```
db2tx CONNECT TO modèle-bdd
```

Voici ce qui s'affiche sous AIX :

```
>-----Database Connection Information -----<
Database Product      = DB2/6000 05.02.0000
User                  = nom-util
Database alias        = modèle-bd
```

**Remarque :** Si vous lancez ultérieurement une commande de DB2 Extension Texte hors de l'interpréteur de commandes, en la faisant précéder de db2tx, la connexion en cours à la base de données est interrompue et une nouvelle connexion est établie à la base de données par défaut.

Si vous utilisez un poste client OS/2 et qu'il s'agit de votre première connexion à une base de données depuis le démarrage du système, les fonctions Gestion des profils utilisateur (UPM) vous invitent à entrer un ID utilisateur et un mot de passe. Si vous êtes déjà connecté, les fonctions UPM tentent d'effectuer une connexion à la base de données éloignée sur le serveur sous l'ID utilisateur et du mot de passe de la session en cours.

---

### Activation d'une base de données

#### Résumé

<b>Utilisation</b>	Une fois pour chaque base de données contenant des colonnes de texte dans lesquelles doit s'effectuer la recherche.
<b>Commande</b>	db2tx=>ENABLE DATABASE
<b>Droits</b>	SYSADM ou DBADM

Cette commande ne comporte pas de paramètres. Elle prépare la base de données connectée en vue de son utilisation par DB2 Extension Texte.

Cette commande déclare également les fonctions DB2 Extension Texte et les types distincts DB2 Extension Texte à DB2. Il s'agit des fonctions SQL que vous utiliserez ultérieurement pour rechercher du texte. Elles sont décrites

## Activation d'une base de données

dans le «Chapitre 11. Fonctions de recherche» à la page 199. Ces déclarations s'appliquent à toutes les sessions à venir.

Une vue de catalogue TEXTINDEXES est créée et conservera une trace des colonnes de texte activées. Pour plus de détails, reportez-vous à la section «Utilisation de la vue de catalogue de DB2 Extension Texte» à la page 112.

Cette commande crée, pour la base de données, des informations sur la configuration de texte incluant les valeurs par défaut associées aux caractéristiques de texte, d'index et de traitement. Ces caractéristiques sont décrites en détail à la section «Définitions de la configuration de texte» à la page 19.

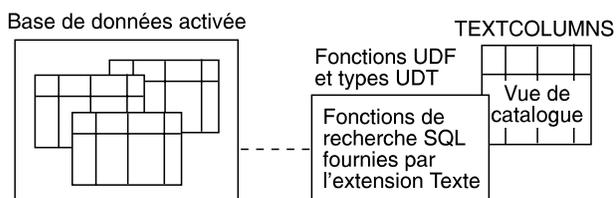


Figure 8. Activation d'une base de données

Lorsqu'une base de données est activée, elle le reste jusqu'à sa désactivation. Pour inverser les modifications qu'entraîne `ENABLE DATABASE`, reportez-vous à la section «Désactivation d'une base de données» à la page 117.

### Conseils

Si vous utilisez la variable d'environnement `DB2TX_INSTOWNER`, elle doit être définie par le nom du propriétaire de l'instance avant que la base de données ne soit activée. Ce point est particulièrement important pour les utilisateurs d'UNIX car, sous UNIX, cette variable est activée par défaut.

Si vous décidez par la suite de supprimer une base de données activée, vous devrez d'abord la désactiver afin que les fonctions DB2 Extension Texte déclarées, la vue de catalogue, etc., soient bien supprimées.

### Activation d'une table de texte (facultatif)

Résumé	
<b>Utilisation</b>	Facultatif. Création d'un index commun à toutes les colonnes de texte d'une table Reportez-vous à la section «Création d'un ou de plusieurs index de texte pour une table» à la page 44.
<b>Commande</b>	db2tx=>ENABLE TEXT TABLE ... (voir Exemples)
<b>Droits</b>	ALTER, SELECT, UPDATE sur la table

Cette étape détermine si vous disposez d'un index commun pour toutes les colonnes de texte de la table, ou de plusieurs index, à savoir un index distinct pour chaque colonne de texte.

Pour obtenir un index commun, exécutez `ENABLE TEXT TABLE`, puis `ENABLE TEXT COLUMN` pour chaque colonne de texte. Pour obtenir des index distincts, n'utilisez pas `ENABLE TEXT TABLE`. Exécutez uniquement `ENABLE TEXT COLUMN` pour chaque colonne de texte, comme illustré par la figure 9 et la figure 10 à la page 66.

Au cours de cette étape, DB2 Extension Texte crée un index de texte vide, qui sera commun à toutes les colonnes de texte activées ultérieurement. Vous indiquez le type d'index, sa fréquence de mise à jour et le répertoire dans lequel il doit être stocké. Les valeurs par défaut utilisées pour les paramètres que vous ne spécifiez pas proviennent du paramétrage de la configuration de texte.

Conseil	
	Si un paramètre tel que le type d'index doit être identique pour la plupart des colonnes de texte, il est plus simple d'utiliser les informations de configuration de texte pour spécifier les paramètres par défaut. Pour plus de détails, reportez-vous à la section «Définitions de la configuration de texte» à la page 19.

Cette étape crée également une table de journalisation vide pour y consigner les documents qui sont ajoutés, modifiés ou supprimés dans une table. Des déclencheurs sont créés afin de mettre à jour la table de journalisation.

Vous ne pouvez pas exécuter `ENABLE TEXT TABLE` pour une table contenant déjà une colonne de texte activée pour DB2 Extension Texte.

## Activation d'une table de texte (facultatif)

Pour supprimer un index créé par ENABLE TEXT TABLE, reportez-vous à la section «Désactivation d'une table de texte» à la page 116.

### Conseil

Si vous décidez par la suite de supprimer une table de texte, vous devrez d'abord la désactiver afin que l'index, la table de journalisation, etc., soient bien eux aussi supprimés.

## Exemples

L'exemple suivant active la table de texte DB2TX.MYTABLE :

```
db2tx ENABLE TEXT TABLE db2tx.ma-table
```

Les valeurs par défaut associées aux caractéristiques de l'index proviennent des définitions de la configuration de texte.

L'exemple suivant définit explicitement les caractéristiques de l'index commun créé pour la table. Pour un serveur UNIX :

```
db2tx ENABLE TEXT TABLE    db2tx.ma-table
                           INDEXTYPE    linguistic
                           UPDATEFREQ   min(100) d(1,2,3,4,5) h(12,15) m(00)
                           DIRECTORY    DB2TX_INSTOWNERHOMEDIR/db2tx/indexes
```

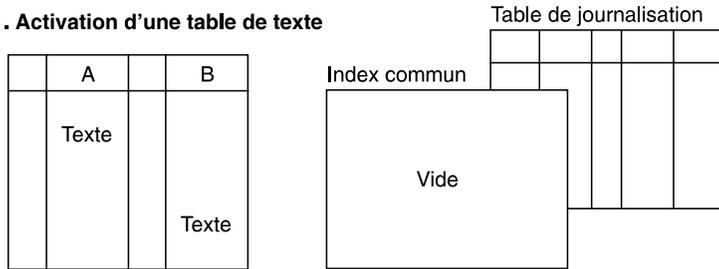
Pour un serveur OS/2 ou Windows NT :

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

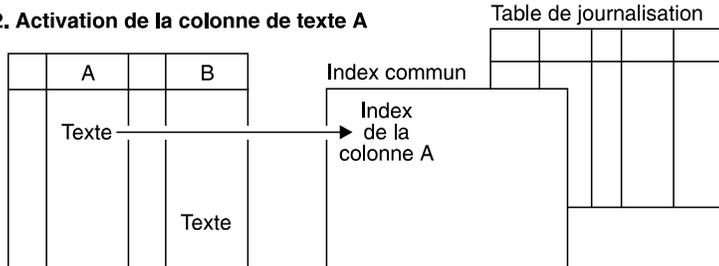
Cet exemple définit le type d'index et son répertoire, puis sa fréquence de mise à jour. Ainsi, l'index est mis à jour à 12 ou à 15 heures, du lundi au vendredi, lorsqu'il y a au moins 100 documents texte en attente.

## Activation d'une table de texte (facultatif)

### 1. Activation d'une table de texte



### 2. Activation de la colonne de texte A



### 3. Activation de la colonne de texte B

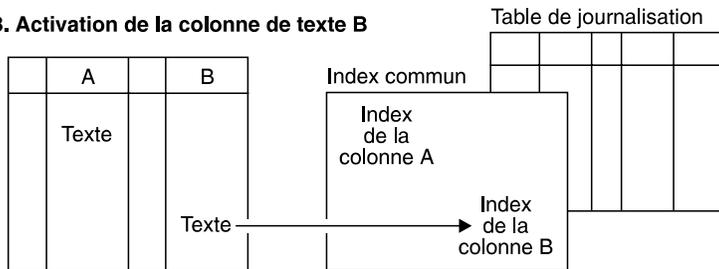
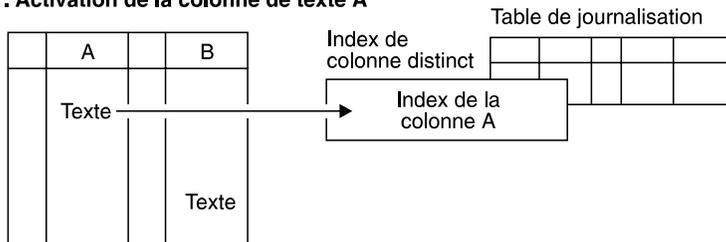


Figure 9. Création d'un index commun à toutes les colonnes de texte d'une table

## Activation d'une colonne de texte

### 1. Activation de la colonne de texte A



### 2. Activation de la colonne de texte B

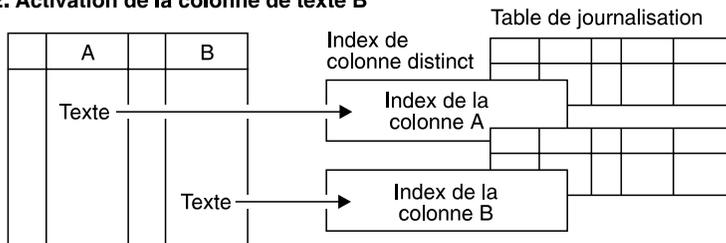


Figure 10. Création d'un index distinct pour chaque colonne de texte

---

## Activation d'une colonne de texte

### Résumé

**Utilisation** Une fois pour chaque colonne contenant du texte dans lequel doit s'effectuer la recherche.

**Commande** `db2tx=>ENABLE TEXT COLUMN ...` (voir Exemples)

**Droits** ALTER, SELECT, UPDATE sur la table

### Conseil

Si un paramètre tel que le type d'index doit être identique pour la plupart des colonnes de texte, il est plus simple d'utiliser les informations de configuration de texte pour spécifier les paramètres par défaut. Pour plus de détails, reportez-vous à la section «Définitions de la configuration de texte» à la page 19.

Pour inverser les modifications qu'entraîne `ENABLE TEXT COLUMN`, utilisez la commande `DISABLE TEXT COLUMN`. Pour désactiver toutes les colonnes de texte activées dans une table, utilisez la commande `DISABLE TEXT TABLE`.

## Activation d'une colonne de texte

Lorsque vous activez une colonne de texte, une colonne de descripteurs est ajoutée à la table, les informations relatives aux documents (format, langue, CCSID) sont définies, et une table de journalisation et un index sont créés.

### Création d'une colonne de descripteurs

Au cours de cette étape, DB2 Extension Texte ajoute une colonne de descripteurs VARCHAR de 60 octets à la table : il s'agit d'une colonne qui contient des descripteurs associés à la colonne de texte en cours d'activation. Ces descripteurs comportent des informations sur le texte de la colonne et des fichiers externes associés. Il s'agit d'un ID de document unique, de la langue, du format et du CCSID du document, ainsi que du nom de l'index.

#### Conseil

Cette version de DB2 Extension Texte ne prend pas en charge les index créés à l'aide d'une version précédente. Vous devez désactiver puis réactiver toutes les bases de données et tables activées pour DB2 Extension Texte.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
Données	Données	Données	Données	Texte

Figure 11. Structure de la table DB2TX.MYTABLE—avant activation

La colonne COMMENT contient des blocs de texte. Avant de pouvoir effectuer une recherche dans le texte de cette colonne, vous devez préparer la base de données et la colonne COMMENT en vue de son utilisation par DB2 Extension Texte.

Après cette étape de préparation, la table DB2TX.MYTABLE contient une colonne supplémentaire pour les descripteurs de texte.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT	COMMENTHANDLE
Données	Données	Données	Données	Texte	Descripteurs de texte

Figure 12. Structure de la table DB2TX.MYTABLE—après activation

**Remarque :** Par la suite, lorsque vous effectuerez une recherche, celle-ci se fera dans la colonne de descripteurs, pas dans la colonne de texte.

## Activation d'une colonne de texte

### Définition des informations relatives aux documents

Vous indiquez le type de documents texte généralement stockés dans cette colonne de texte : leur format (ASCII par exemple), leur langue et leur CCSID. Les valeurs par défaut relatives à ces informations peuvent être indiquées dans les définitions de la configuration de texte. Pour plus de détails, reportez-vous à la section «Définitions de la configuration de texte» à la page 19.

### Création d'une table de journalisation

Au cours de cette étape, une table de journalisation et une vue appelée LOGIXnnnnnn sont créées, où IXnnnnnnn représente le nom de l'index (que vous pouvez trouver dans la vue de catalogue). Si un espace table par défaut est spécifié dans la configuration du texte, la table de journalisation y est stockée, sinon elle est placée dans l'espace table par défaut du système DB2. Pour optimiser les performances et l'utilisation de l'espace disque, vous pouvez spécifier un espace table différent qui servira aux tables de journalisation.

Des déclencheurs sont également créés pour que, lorsqu'un document est ajouté à la colonne ou y est modifié, l'information soit ajoutée à la table de journalisation. De ce fait, ces documents seront indexés lors de la prochaine indexation.

En revanche, si des fichiers externes sont modifiés ou ajoutés, les déclencheurs n'en sont pas directement informés. Dans ces cas, pour que les déclencheurs ajoutent l'information dans la table de journalisation, vous devez utiliser l'instruction UPDATE, comme indiqué dans l'exemple de la section «Mise à jour d'un index pour y intégrer les références de fichiers externes» à la page 102.

Si des erreurs surviennent au cours de l'indexation, comme lorsqu'un document en attente d'indexation n'a pu être trouvé, des *événements d'erreur* sont ajoutés à la table de journalisation et peuvent être consultés, comme indiqué à la section «Affichage des événements d'erreur» à la page 110.

#### Conseil

A ce stade, si vous manquez d'espace de journalisation, reportez-vous à la section «Activation d'une colonne de texte dans une table volumineuse» à la page 70, pour prendre connaissance des solutions possibles.

Dans les bases de données partitionnées, chaque table est affectée à un espace table et à un groupe de noeuds. Il est important que la table de journalisation

soit associée à un espace table appartenant au même groupe de noeuds que la table utilisateur activée. DB2 Extension Texte le vérifie lors de l'exécution de la commande ENABLE.

### Création d'un index

Si vous souhaitez que chaque colonne de texte ait son propre index (vous n'avez donc pas effectué l'étape ENABLE TEXT TABLE), DB2 Extension Texte crée un index distinct pour la colonne de texte au cours de cette étape. Vous indiquez le type d'index, sa fréquence de mise à jour et le répertoire dans lequel il doit être stocké. Par contre, si vous préférez créer un index commun à toutes les colonnes de la table, vous avez déjà exécuté ENABLE TEXT TABLE et indiqué les paramètres de l'index. Par conséquent, si vous les indiquez à nouveau ici, ils sont ignorés.

Utilisez le mot clé UPDATEINDEX pour déterminer si l'indexation des documents texte de la colonne de texte spécifiée débute immédiatement, ou seulement lors de la prochaine indexation périodique. Si vous n'utilisez pas ce mot clé, le système utilisera la valeur spécifiée dans les définitions de la configuration de texte.

**Création d'index de types différents pour une colonne de texte.** Vous pouvez créer plusieurs index pour une colonne de texte. Ceci peut être utile pour permettre, par exemple, une recherche linguistique et une recherche de correspondance partielle sur la même colonne de texte en utilisant différents types d'index, tels qu'un index linguistique et un index Ngram. Pour ce faire, exécutez de nouveau ENABLE TEXT COLUMN, en spécifiant non seulement le type d'index supplémentaire à créer, mais aussi un nom de colonne de descripteurs qui lui sera propre.

#### Conseil

Si vous utilisez une base de données partitionnée et que vous avez décidé de modifier la configuration d'un groupe de noeuds puis de lancer une redistribution des lignes d'une table, vous devez recréer les index de texte et les tables de journalisation. Pour ce faire, désactivez puis réactivez toutes les tables activées pour DB2 Extension Texte dans ce groupe de noeuds.

### Exemples

L'exemple suivant active la colonne de texte COMMENT de la table DB2TX.MYTABLE, et attribue le nom COMMENTHANDLE à la colonne de descripteurs créée :

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

## Activation d'une colonne de texte

Les valeurs associées aux informations relatives au texte et aux caractéristiques de l'index proviennent des paramètres de configuration de texte.

L'exemple suivant définit explicitement les valeurs pour le type de documents se trouvant dans la colonne COMMENT. Les valeurs par défaut associées aux caractéristiques de l'index proviennent des paramètres de configuration de texte.

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

L'exemple suivant définit explicitement les valeurs pour les caractéristiques de l'index créé pour la colonne COMMENT. Cet exemple définit le type d'index, son répertoire et sa fréquence de mise à jour. Ainsi, l'index est mis à jour à 12 ou à 15 heures, du lundi au vendredi, lorsqu'il y a au moins 100 documents texte en attente. Les valeurs par défaut pour les informations relatives au texte proviennent des définitions de la configuration de texte.

Pour un serveur 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
```

Pour un serveur OS/2, Windows NT ou 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  mon-espace-table
                STOGROUP     mon-groupe-stockage
```

## Activation d'une colonne de texte dans une table volumineuse

Si vous gérez une table définie avec une longueur de ligne élevée, n'oubliez pas que l'activation d'une colonne de texte ajoute une colonne de descripteurs de type DB2TEXTH (VARCHAR 60). De même, l'activation d'un fichier externe ajoute une colonne de descripteurs de type DB2TEXTFH (VARCHAR 210). Cette opération peut avoir des répercussions si la longueur des lignes de la table approche de la valeur maximale déterminée par DB2.

En outre, avant d'activer une colonne de texte dans une table volumineuse, vérifiez par l'utilitaire REORGANIZE de DB2 UDB si la table a besoin d'être réorganisée. Lorsque vous activez une telle table pour la première fois, suivez la procédure ci-dessous pour que l'indexation soit plus rapide :

1. Activez la table avec l'option NOUPDATE, ce qui a pour effet de créer les descripteurs, sans pour autant indexer les documents.
2. Réorganisez la table à l'aide de l'utilitaire REORGANIZE de DB2 UDB.
3. Puis, créez l'index en lançant la commande UPDATE INDEX.

Lorsque vous activez une colonne de texte ou des fichiers externes, DB2 Extension Texte ajoute une colonne de descripteurs à la table et initialise les valeurs associées. Cela a pour effet de générer des entrées dans la table de journalisation de DB2 UDB. Si le nombre de ces entrées est très élevé, il peut arriver que l'espace de journalisation de DB2 UDB soit insuffisant.

Il existe deux méthodes pour remédier à cette situation, la première étant préférable pour des raisons de performances :

- Augmentez l'espace de journalisation disponible par la commande UPDATE DB CFG de DB2 UDB, de manière à modifier les paramètres de configuration pour LOGPRIMARY, LOGSECOND et LOGFILSIZ. Les valeurs suivantes sont issues de la pratique ; vous aurez peut-être besoin de les modifier en fonction de votre installation.

```
LOGSECOND    50
```

Vérifiez que la somme des valeurs de LOGPRIMARY et de LOGSECOND ne dépasse 128. Augmentez également la taille de la mémoire dynamique allouée à l'application.

```
APPLHEAPSZ   512
```

- Forcez des validations DB2 UDB intermédiaires par le paramètre de configuration COMMITCOUNT, comme indiqué à la section «Configuration» à la page 17. La valeur à entrer correspond au nombre d'instructions de mise à jour ou d'insertion à l'issue desquelles DB2 Extension Texte émet une instruction de validation DB2 UDB. Cette procédure réduit l'espace requis pour les tables de journalisation, mais augmente, en revanche, la durée nécessaire à l'étape d'activation.

### Activation de colonnes de texte d'un type de données non pris en charge

Les colonnes de texte doivent être de type CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR ou CLOB pour être activées par DB2 Extension Texte. Si les documents sont dans une colonne de type différent, un type UDT, par exemple, vous devez fournir une fonction qui convertit le type UDT en entrée en type de sortie CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR ou CLOB.

## Activation d'une colonne de texte

Utilisez le mot clé `FUNCTION` dans `ENABLE TEXT COLUMN` pour indiquer le nom de cette fonction.

Exemple : vous souhaitez stocker du texte condensé dans une table.

1. Créez un type UDT pour le texte :

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

2. Créez une table et insérez-y le texte :

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

Pour activer la colonne de texte en vue de son utilisation par DB2 Extension Texte :

1. Créez une fonction UDF, appelée `UNCOMPRESS`, par exemple, qui reçoit une valeur de type `COMPRESSED_TEXT` et renvoie le texte non condensé correspondant sous la forme d'une valeur `CLOB (10M)`, par exemple.
2. Activez la colonne de texte en utilisant le mot clé `FUNCTION` pour identifier la fonction UDF `UNCOMPRESS` :

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

## Activation de colonnes contenant le type de données DATALINK

DB2 Extension Texte prend en charge le type de données `DATALINK`. La technologie Data Link de DB2 garantit l'intégrité référentielle entre les données stockées dans DB2 et celles des fichiers externes. Elle y parvient via un nouveau type de données, `DATALINK`, qui contient des références à un objet stocké hors de la base de données. Une valeur `DATALINK` stockée dans une table DB2 code le nom du serveur Data Link contenant le fichier, ainsi que le nom du fichier, sous forme d'une URL.

Si les données `DATALINK` se trouvent dans des fichiers externes et non dans une table DB2, suivez les indications de la section «Activation de fichiers de texte externes» à la page 73.

Avant d'activer une colonne de texte contenant le type de données `DATALINK`, exécutez le modèle de vérification d'installation (reportez-vous à la section «Préparation d'une base de données exemple pour vérifier l'installation» à la page 16).

## Activation d'une colonne de texte

**Exemple sous AIX :** L'instruction suivante illustre l'insertion d'une valeur Data Link dans votre colonne de texte au moyen du type de données DATALINK sous AIX :

```
db2 "insert into ma-table valeurs ('Peter',
    valeur-dl ('http://d1fs.boeblingen.ibm.com/test/documents/lettre1',
        'URL', 'ceci est une lettre de Peter'))"
```

Dans cet exemple, *d1fs.boeblingen.ibm.com* est le nom du serveur de fichiers, *test* est le nom du système de fichiers et *documents/lettre1* est le chemin et le nom du fichier.

Vérifiez que le système de fichiers est monté sur le serveur DB2 sous le même nom que celui défini dans l'URL.

**Exemple sous Windows NT ou Windows 2000 :** L'instruction suivante illustre l'insertion d'une valeur Data Link dans une colonne de texte au moyen du type de données DATALINK sous Windows :

```
db2 "insert into ma-table ('Peter',
    valeur-dl ('unc:\\d1fs.boeblingen.ibm.com\lecteur-cd\documents\lettre1',
        'URL', 'ceci est une lettre de Peter'))"
```

Dans cet exemple, *lecteur-cd* est le nom de partage de l'unité sur laquelle est stocké le fichier.

Si vous avez des difficultés lors de l'indexation de documents stockés dans une colonne datalink, vérifiez que vous pouvez accéder aux documents à partir de votre serveur DB2 au moyen des API fournies par DB2.

---

## Activation de fichiers de texte externes

Cette option n'est fournie qu'à des fins de compatibilité avec les versions antérieures. L'utilisation d'une colonne avec type de données DATALINK constitue la meilleure méthode pour indexer des fichiers non stockés dans des tables DB2. Pour plus de détails, reportez-vous à la section «Activation de colonnes contenant le type de données DATALINK» à la page 72.

### Résumé

<b>Utilisation</b>	Facultatif. Une fois pour chaque table associée à des fichiers externes sur lesquels porteront les recherches.
<b>Commande</b>	db2tx=>ENABLE TEXT FILES ... (voir Exemples)
<b>Droits</b>	ALTER, SELECT, UPDATE sur la table

## activation de fichiers de texte externes

DB2 Extension Texte peut effectuer des recherches non seulement dans le texte stocké dans les tables DB2 UDB, mais également dans les documents texte stockés dans des fichiers. Cette étape de préparation est nécessaire pour pouvoir élargir la recherche au texte contenu dans des fichiers externes. La table associée aux fichiers de texte externes ne doit pas avoir été activée par la commande `ENABLE TEXT TABLE`.

Un index et une table de journalisation sont créés, et les informations relatives aux documents sont définies, selon la même procédure que celle décrite à la section «Activation d'une colonne de texte» à la page 66.

### Conseils

1. A ce stade, si vous manquez d'espace de journalisation, reportez-vous à la section «Activation d'une colonne de texte dans une table volumineuse» à la page 70, pour prendre connaissance des solutions possibles.
2. Si la table que vous activez est partitionnée et qu'elle utilise un groupe de noeuds comportant plusieurs noeuds physiques, vérifiez que les fichiers externes référencés dans les colonnes de la table sont situés sur le noeud qui héberge leur partition de table respective.

Une colonne de descripteurs de type `DB2TEXTFH` est ajoutée à la table `DB2 UDB` existante. Son rôle est de conserver les références des fichiers externes, tous les descripteurs contenant un index et les informations spécifiques au document (`CCSID`, format et langue).

Pour obtenir une description, reportez-vous à la section «Descripteurs pour fichiers externes» à la page 82.

Vous pouvez spécifier des paramètres supplémentaires, tels que les caractéristiques par défaut de l'index, en procédant de la même manière que pour activer une colonne de texte.

Une fois l'index créé, vous pouvez déplacer ou supprimer les fichiers externes. Il est toujours possible d'effectuer une recherche dans les fichiers. Vous pouvez également insérer de nouvelles lignes dans la table, et utiliser la commande `UPDATE INDEX` pour actualiser l'index en y intégrant les références du nouveau fichier.

Si la table que vous activez utilise un groupe de noeuds contenant plusieurs noeuds physiques, vérifiez que les fichiers externes référencés dans les colonnes de la table sont situés sur le noeud qui héberge la partition de table.

### Exemples

1. Créez une table DB2TX.EXTFILE contenant au moins une colonne, ou utilisez une table existante.
2. Ajoutez la colonne de descripteurs FILEHANDLES à la table DB2TX.EXTFILE :

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

3. Initialisez le descripteur.

- Pour chaque ligne d'une nouvelle table :

```
db2 INSERT INTO db2tx.EXTFILE (FILEHANDLES)
      VALUES (db2tx.INIT_TEXT_HANDLE
              (850, 'TDS'
```

```
'US_ENGLISH',
```

```
          'd:\dmb\samples\tx')
```

- Dans le cas d'une table préexistante, actualisez les colonnes de descripteurs de manière à ce qu'elles intègrent les références du fichier externe ; pour ce faire, spécifiez le nom de ce fichier dans la commande :

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

#### Conseil

N'utilisez pas INIT\_TEXT\_HANDLE pour mettre à jour des colonnes de descripteurs qui font référence à des fichiers externes.

4. Actualisez l'index :

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

---

### Fin de la session

Vous avez effectué toutes les étapes de préparation de vos documents texte en vue d'une recherche.

Si vous avez indiqué NOUPDATE pour le mot clé UPDATEINDEX lorsque vous avez activé la colonne de texte, DB2 Extension Texte n'indexe pas immédiatement le texte, mais attend la prochaine indexation périodique. Pour mettre à jour l'index immédiatement, reportez-vous à la section «Mise à jour d'un index» à la page 101.

## Fin de la session

Lorsque l'indexation des documents est terminée, vous pouvez commencer à extraire des informations comme indiqué au «Chapitre 6. Mode de recherche» à la page 77.

### **Conseil**

Utilisez la commande `GET INDEX STATUS` pour être averti de la fin de l'indexation.

Pour quitter l'interpréteur de commandes DB2 Extension Texte, entrez :  
`db2tx=>quit`

---

## Chapitre 6. Mode de recherche

DB2 Extension Texte fournit des fonctions SQL vous permettant d'inclure des sous-requêtes de recherche de texte dans des requêtes SQL. Ces fonctions sont fournies en supplément de celles normalement disponibles dans SQL et il y fait référence ici en tant que fonctions DB2 Extension Texte.

Pour obtenir une description de la syntaxe de ces fonctions, reportez-vous au «Chapitre 11. Fonctions de recherche» à la page 199.

Avant de commencer la recherche, consultez la section «Types de recherche» à la page 40, et utilisez GET INDEX SETTINGS pour déterminer le type d'index associé au texte dans lequel vous effectuez la recherche. En effet, les résultats d'une recherche varient en fonction du type d'index.

Dans les exemples de ce chapitre, l'index est de type linguistique.

Ce chapitre décrit :

- Les modèles de fonctions DB2 Extension Texte
- Le modèle de table DB2TX.SAMPLE
- Les descripteurs pour les fichiers externes.
- La définition du chemin de fonctions permettant à SQL d'accéder aux fonctions DB2 Extension Texte
- La recherche de texte par CONTAINS, NO\_OF\_MATCHES et RANK
- La définition d'arguments de recherche dans les fonctions DB2 Extension Texte, en utilisant des exemples de la fonction CONTAINS.
- L'affinement d'une recherche précédente par les fonctions CONTAINS et REFINE
- La définition et l'extraction d'informations dans les descripteurs de texte par les fonctions INIT\_TEXT\_HANDLE, CCSID, FORMAT et LANGUAGE
- L'amélioration des performances de recherche via SEARCH\_RESULT

---

### Localisation d'exemples de syntaxe pour les fonctions de recherche

DB2 Extension Texte fournit un fichier d'entrée pour l'interpréteur de commandes, appelé txsample.udf. Il contient des exemples de fonctions de recherche DB2 Extension Texte qui s'exécutent sur la table modèle. Utilisez ce fichier pour visualiser des exemples de la syntaxe de la préparation de texte et des fonctions de recherche DB2 Extension Texte, ainsi que de la syntaxe utilisée dans les arguments de recherche.

## Localisation d'exemples de syntaxe pour les fonctions de recherche

Pour exécuter ce fichier, entrez :

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

---

### Modèle de table DB2TX.SAMPLE

Le modèle de table DB2TX.SAMPLE est utilisé dans de nombreux exemples de syntaxes de recherche. Vous pouvez mettre en oeuvre ces exemples dans DB2 Extension Texte.

Pour la création de la table DB2TX.SAMPLE, reportez-vous à la section «Préparation d'une base de données exemple pour vérifier l'installation» à la page 16.

Le tableau 6, présente un extrait de la table DB2TX.SAMPLE.

Tableau 6. Extrait du modèle de table DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 5	RSSHERM at CHGVMIC1	erreur LIBDB2E.A	1995-07-25 -20.13.59	Le message 'No such file or directory' apparaît sur le poste du client sur LIBDB2E.A. Il semble qu'il s'agisse d'un message d'erreur autre que celui associé au pilote d'entrées-sorties asynchrones. Il utilise la version bêta 4 sur 3.2.5. J'ai demandé au client de comparer ses droits d'accès et la propriété des fichiers /usr/lpp/db2_02_01/lib aux miens ; ils sont à présent identiques. Son fichier .profile et son ENV semblent corrects. Malheureusement, il a validé (par la commande COMMIT) l'installation. Je ne vois pas d'où pourrait venir l'erreur.
doc 6	EDWARDSC at SYDVM1	ID utilisateur et mot de passe reçus en minuscules de DDCS/2	1995-07-25 -20.15.20	Après avoir vérifié à nouveau, l'instance dans laquelle j'avais rencontré des incidents relatifs aux caractères majuscules-minuscules utilisait une passerelle DB2/2 vers MVS. L'ID utilisateur en minuscules que j'ai entré n'a pas été accepté (le mot de passe n'a pas été pris en compte). La connexion n'a abouti que lorsque j'ai entré l'ID utilisateur en majuscules. J'imagine que cela ne vous est d'aucune aide. Désolé.

Tableau 6. Extrait du modèle de table DB2TX.SAMPLE (suite)

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 7	SKY at TOROLAB4	ODBC & procédures mémorisées	1995-07-25 -20.42.27	<p>Il existe deux ensembles de modèles de programme expliquant l'utilisation des procédures mémorisées par CLI (ODBC).</p> <p>Le fichier en C inpsrv2.c (sur le serveur) et inpli2.c (sur le poste client) constituent le modèle présentant l'utilisation des procédures mémorisées en entrée. Les fichiers outsrv2.c et outcli2.c constituent le modèle présentant l'utilisation des procédures mémorisées en sortie.</p> <p>Ces fichiers font partie des fichiers .../sqllib/samples/cli. Le fichier MAKE les crée automatiquement et transfère le fichier du serveur vers le sous-répertoire adéquat.</p>
doc 8	ADAMACHE at TOROLAB2	Violation d'accès DB2SYS.DLL	1995-07-25 -21.13.22	<p>Avez-vous fait installer une version bêta précédente ? Si c'est le cas, l'avez-vous retirée par Software Installer ?</p> <p>Avez-vous supprimé les répertoires de bases de données (SQLDBDIR et SQL00001, etc.) des pilotes bêta précédents ?</p>

## Modèle de table DB2TX.SAMPLE

Tableau 6. Extrait du modèle de table DB2TX.SAMPLE (suite)

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 9	ADAMACHE at TOROLAB2	CREATE DB = SYS3175: db2sysc.exe dans db2eng.dll	1995-07-25 -21.40.09	<p>De nombreux utilisateurs de versions bêta DB2/2 suppriment une bêta précédente par Software Installer, installent la bêta 5 (ou code confirmé), créent une base de données et obtiennent : SYS3175: db2sysc.exe in db2eng.dll</p> <p>Cela se produit car le format du répertoire a changé entre la version bêta 4 et bêta 5. Notre installation DB2/2 ne transfère pas le répertoire sqlldbidr entre les pilotes de bêta. Vous devez supprimer toutes les occurrences des répertoires sqlldbidr, sql000x et le répertoire \sqllib\db2\sqlldbidr.</p> <p>Vous devez commencer par supprimer la version bêta précédente par Software Installer, puis supprimer toutes les occurrences des répertoires sqlldbidr, sql000x et le répertoire \sqllib\db2\sqlldbidr, et ensuite seulement installer le nouveau code.</p>
doc 10	RSSHERM at CHGVMIC1	DB2/NT - support SNA	1995-07-25 -22.10.15	<p>DB2/NT va-t-il pouvoir faire office à la fois de serveur vers les clients CAE/WIN et de poste client (hopping) vers DB2/6000 et/ou vers DB2/MVS sur un réseau SNA ?</p> <p>L'autre possibilité serait DRDA à partir de DB2/NT vers DB2/6000 et/ou vers DB2/MVS - à nouveau via le réseau SNA, qui est pris en charge, je suppose ?</p>

Voici la première et la dernière colonne de la structure de la table :

**DB2TX.SAMPLE**

DOCID	COMMENT
doc 1	Le client est ...
doc 2	Après vérification ...

Figure 13. Structure de la table DB2TX.SAMPLE

La colonne COMMENT contient le texte dans lequel doit s'effectuer la recherche. Vous devez préalablement préparer la colonne COMMENT en vue de son utilisation par DB2 Extension Texte à l'aide de la commande ENABLE TEXT COMLUMN.

Suite à cette étape de préparation, la table DB2TX.SAMPLE se présente comme suit :

**DB2TX.SAMPLE**

DOCID	COMMENT	COMMENTHANDLE
doc 1	Le client est ...	X'..descr..'
doc 2	Après vérification ...	X'..descr..'

Figure 14. Structure de la table DB2TX.SAMPLE après activation

A présent, la table comporte une colonne supplémentaire pour les descripteurs et chaque objet texte est représenté par un descripteur unique.

Lorsque vous insérez du texte dans une colonne de texte activée, un déclencheur d'insertion crée un descripteur pour celui-ci.

**DB2TX.SAMPLE**

DOCID	COMMENT	COMMENTHANDLE
doc 1	Le client est ...	X'..descr..'
doc 2	Après vérification ...	X'..descr..'

Descripteurs créés par  
ENABLE TEXT COLUMN

Ligne insérée :

doc 11	J'ai installé ...	X'..descr..'
--------	-------------------	--------------

Descripteur crée par  
un déclencheur d'insertion

Figure 15. Un déclencheur crée le descripteur pour une ligne insérée.

Un descripteur contient les informations suivantes :

- un ID document
- le nom et l'emplacement de l'index associé

## Modèle de table DB2TX.SAMPLE

les informations relatives au document : le CCSID, le format et la langue.

Les fonctions SQL fournies par DB2 Extension Texte utilisent un descripteur comme paramètre, stockent, accèdent, recherchent et manipulent le texte dans le cadre du traitement SQL de la table.

---

### Descripteurs pour fichiers externes

DB2 Extension Texte peut effectuer des recherches non seulement dans le texte stocké dans les tables DB2 UDB, mais également dans des fichiers de texte stockés dans des fichiers. La section «Activation de fichiers de texte externes» à la page 73, décrit l'étape de préparation nécessaire à la recherche dans des documents texte extérieurs à des tables DB2 UDB. Lors de cette étape, la commande `ENABLE TEXT FILES` permet de créer une colonne de descripteurs de type `DB2TEXTFH` pour les fichiers externes. Cette colonne est ajoutée à la table existante.

Vous pouvez, par exemple, créer une table avec des colonnes correspondant au nom de l'auteur et à la date de création du document.

Pour initialiser les descripteurs de fichiers, utilisez la commande `INIT_TEXT_HANDLE`. Chaque descripteur contient, outre l'ID du document, le nom et l'emplacement de l'index associé, les informations spécifiques au document (CCSID, format et langue), la référence du fichier externe.

---

### Définition du chemin de fonction en cours

►—`SET—CURRENT FUNCTION PATH`——`DB2TX, ...`—►

Utilisez l'instruction SQL `SET CURRENT FUNCTION PATH` pour ajouter `DB2TX` aux noms de chemins en cours afin que SQL puisse trouver les fonctions DB2 Extension Texte. Si vous décidez de ne pas le faire, vous pouvez qualifier explicitement les noms des fonctions en tapant, par exemple, `DB2TX.CONTAINS` pour la fonction `CONTAINS`.

Les exemples du présent chapitre utilisent la forme qualifiée pour les fonctions de DB2 Extension Texte. Vous pouvez utiliser les exemples d'instructions tels qu'ils sont écrits, sans avoir à définir le chemin de fonction en cours.

### Conseil

N'oubliez pas de définir le chemin de fonction en cours à chaque connexion à une base de données.

## Recherche de texte

► CONTAINS (—descripteur—, —argument-recherche—) →  
    | NO\_OF\_MATCHES |  
    | RANK |

►► SEARCH\_RESULT (—schéma—, —table—, —descripteur—, —argument-recherche—) →

Cette section explique l'utilisation des fonctions SQL fournies avec DB2 Extension Texte pour effectuer une recherche dans des bases de données DB2 contenant du texte. Elle traite des sujets suivants :

- Création d'une requête
- Recherche du nombre de correspondances trouvées dans un document texte
- Recherche du rang d'un document texte trouvé

L'utilisation de la fonction SEARCH\_RESULT est décrite à la section «Amélioration des performances de recherche» à la page 99.

Chacune de ces fonctions DB2 Extension Texte recherche dans l'index de texte les occurrences de l'argument de recherche. S'il existe 100 000 documents texte dans la table (par exemple), la fonction CONTAINS, RANK ou NO\_OF\_MATCHES est appelée 100 000 fois. Mais l'index de texte ne fait pas l'objet de 100 000 recherches. Au lieu de cela, lors du premier appel de la fonction, une liste interne de tous les documents contenant le terme de recherche est créée ; les appels ultérieurs de la fonction déterminent si le document concerné se trouve dans la liste.

### Conseil

Lorsque vous utilisez les fonctions DB2 Extension Texte pour effectuer une recherche dans une table, transmettez la colonne de descripteurs, à la fonction et non la colonne de texte. En effet, si vous tentez d'effectuer une recherche dans une colonne de texte, SQL affiche un message indiquant que le type de données est incorrect, par exemple :

Aucune fonction nommée "CONTAINS" et ayant des arguments compatibles n'a été trouvée dans le chemin de fonctions.

Si vous recherchez du texte immédiatement après avoir émis la commande `ENABLE TEXT TABLE` ou `ENABLE TEXT COLUMN`, une erreur `RC_SE_EMPTY_INDEX` peut se produire. Elle indique que l'index en cours de création par la commande n'existe pas encore. Le temps nécessaire à la création d'un index dépend de facteurs tels que le nombre de documents indexés et les performances du système effectuant l'indexation. Cette opération peut prendre quelques minutes ou plusieurs heures, et doit être effectuée de préférence lorsque le système est peu sollicité (la nuit, par exemple).

Si ce message apparaît, faites une nouvelle tentative de recherche ultérieurement, ou utilisez la commande `GET INDEX STATUS` pour vérifier s'il y a eu des erreurs d'indexation.

### Création d'une requête

Cet exemple explique comment la fonction `CONTAINS` recherche du texte dans des documents identifiés par un descripteur. Elle renvoie la valeur 1 si le texte correspond à l'argument de recherche. Sinon, elle renvoie la valeur 0.

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

Dans cet exemple, le terme `compresser` est recherché dans le texte auquel les descripteurs font référence dans la colonne `COMMENTHANDLE`. Les descripteurs de la colonne `COMMENTHANDLE` indiquent où est indexé le texte `COMMENT`.

**Conseil**

Si vous avez créé des identificateurs constitués d'une combinaison de majuscules et de minuscules pour des tables ou des colonnes, n'oubliez pas de les mettre entre guillemets. Par exemple :

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

Si vous utilisez des instructions SELECT de DB2 UDB dans la ligne de commande, l'analyseur syntaxique du système d'exploitation retirera les caractères spéciaux, tels que les guillemets, de la chaîne de commande ; pour y remédier, faites précéder ces caractères d'une barre oblique inversée afin de les masquer. Par exemple :

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

**Recherche et renvoi du nombre de correspondances trouvées**

Utilisez la fonction NO\_OF\_MATCHES pour déterminer la fréquence de localisation des critères de recherche dans chaque document texte.

```
db2=>WITH TEMPTABLE (DATE, SUBJECT, MATCHES)
      AS (SELECT DATE, SUBJECT,
                DB2TX.NO_OF_MATCHES (COMMENTHANDLE, "compresser")
          FROM DB2TX.SAMPLE)
SELECT *
FROM TEMPTABLE
WHERE MATCHES > 0
```

NO\_OF\_MATCHES renvoie un entier.

**Recherche et renvoi du rang d'un document texte trouvé**

RANK est une valeur absolue indiquant le degré de pertinence du document par rapport aux autres document trouvés. Elle indique le nombre de correspondances trouvées dans le document par rapport à sa taille. Vous pouvez obtenir le rang d'un document trouvé en utilisant la fonction RANK.

Voici un exemple :

```
db2=>WITH TEMPTABLE (DATE, SUBJECT, RANK)
      AS (SELECT DATE, SUBJECT,
                DB2TX.RANK (COMMENTHANDLE, "compresser")
          FROM DB2TX.SAMPLE)
```

## Recherche de texte

```
SELECT *
  FROM TEMPTABLE
 WHERE RANK > 0
 ORDER BY RANK DESC
```

RANK renvoie une valeur de type DOUBLE comprise entre 0 et 1.

---

## Définition des arguments de recherche

Les arguments de recherche sont utilisés dans CONTAINS, NO\_OF\_MATCHES, RANK et HANDLE\_LIST. Cette section utilise la fonction CONTAINS pour présenter différents exemples d'arguments de recherche dans des fonctions DB2 Extension Texte.

### Recherche de plusieurs termes

Un argument de recherche peut être constitué de plusieurs termes. Pour les associer, reliez-les par des virgules, comme suit :

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

Cette forme d'argument recherche du texte contenant l'un des termes indiqués. En terme de logique, ils sont reliés par un opérateur OR.

### Recherche à l'aide des opérateurs booléens AND et OR

(Reportez-vous également à la section «Recherche à l'aide de l'opérateur booléen NOT» à la page 91.)

Des termes de recherche peuvent être combinés à d'autres à l'aide des opérateurs booléens "&" (AND) et "|" (OR). Par exemple :

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

Vous pouvez combiner plusieurs termes à l'aide des opérateurs booléens :

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

Si vous utilisez plusieurs opérateurs booléens, DB2 Extension Texte les évalue de gauche à droite mais l'opérateur logique AND (&) est prioritaire par rapport à l'opérateur OR (|).. Par exemple, si vous n'indiquez pas de parenthèses,

```
"DB2" & "compiler" | "support" & "compresser"
```

## Définition des arguments de recherche

est évalué comme suit :

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

Si vous souhaitez une autre interprétation, vous devez donc inclure des parenthèses, par exemple :

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

Si vous combinez des opérateurs booléens à des termes de recherche reliés par des virgules, comme suit :

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

la virgule est interprétée comme étant un opérateur booléen OR :

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

### Recherche des variantes d'un terme

Si vous utilisez un index **exact**, DB2 Extension Texte recherche les termes tels que vous les avez entrés. Par exemple, pour le terme *supports*, l'extension ne propose que du texte contenant "supports". Du texte contenant le singulier "support" n'est pas proposé.

Si vous utilisez un index **linguistique**, DB2 Extension Texte recherche également les variantes des termes, comme le pluriel d'un nom ou un verbe sous sa forme conjuguée.

Par exemple, si vous tapez le terme *pilote*, DB2 Extension Texte propose du texte contenant les termes "pilote", "pilotes", "pilotage", "pilotait" et "pilote".

### Recherche à partir de fragments de termes (caractères génériques)

Les caractères "génériques" permettent une recherche plus souple. Ils représentent des caractères facultatifs se trouvant au début, au milieu ou à la fin d'un terme de recherche. Ils permettent d'augmenter le nombre de documents texte localisés par une recherche.

#### Conseil

Lorsque vous utilisez des caractères génériques, vous ne pouvez pas employer le mot clé SYNONYM FORM OF.

Les caractères génériques sont particulièrement utiles pour rechercher les variantes d'un terme dans un index exact. Dans un index linguistique, leur emploi se révèle superflu, puisque les variantes des termes sont toujours proposées.

## Définition des arguments de recherche

Notez que les fragments de mots (mots contenant des caractères génériques) ne peuvent pas être réduits à leur forme de base. Ainsi, si vous effectuez une recherche sur `cheva%`, vous ne trouverez pas le mot “chevaux” car il est réduit à sa forme de base “cheval” dans l’index. Pour le trouver, vous devez effectuer la recherche sur `cheva%`.

DB2 Extension Texte utilise deux caractères génériques : le trait de soulignement (`_`) et le symbole pour-cent (`%`) :

- Le symbole `%` représente **un nombre indéfini de caractères quelconques**. Dans l’exemple ci-dessous, le symbole `%` fait office de caractère générique et est placé au début du terme de recherche :

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

Suite à cette recherche, les documents texte contenant, par exemple, “username”, “file\_name” et “table-name” sont proposés.

Le symbole `%` peut également représenter un **mot** : l’exemple suivant recherche des documents texte contenant des expressions telles que “graphic function” ou “query function”.

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

- Le trait de soulignement (`_`) représente **un caractère unique** d’un terme de recherche : l’exemple suivant recherche des documents texte contenant les termes “CLOB” et “BLOB”.

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

## Recherche de termes contenant un caractère générique

Si vous souhaitez rechercher un terme qui contient le caractère “%” ou “\_”, faites précéder celui-ci d’un caractère *d’échappement*, puis identifiez ce dernier par le mot clé `ESCAPE`.

Par exemple, pour rechercher “intérêt 10%” :

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

Dans cet exemple, le caractère d’échappement est `!"`.

### Recherche de termes conjoints ou disjoints

Si vous indiquez les termes conjoints "hard disk" comme dans l'exemple ci-dessous, la recherche ne porte que sur cette expression, et non sur les deux termes qui la composent, quel que soit l'index utilisé.

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

Pour rechercher toute occurrence de ces deux termes comme dans "data disks and hard drives", par exemple, séparez-les par une virgule :

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

### Recherche de termes d'une même phrase ou d'un même paragraphe

Voici un exemple d'argument de recherche portant sur des documents texte dans lesquels les termes de recherche se trouvent dans une même phrase :

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

Vous pouvez également rechercher plusieurs mots apparaissant conjointement. Dans l'exemple suivant, la recherche porte sur plusieurs mots apparaissant dans le même paragraphe.

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

### Recherche de termes dans des sections de documents structurés

Voici un exemple d'argument de recherche portant sur des documents texte dans lesquels le terme de recherche, Williams, apparaît dans la sous-section author dans la section play de documents structurés. La structure du document est spécifiée par le modèle play décrit dans le fichier des modèles de documents. Pour plus de détails, reportez-vous à la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48.

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

## Définition des arguments de recherche

### Recherche de synonymes

Pour un index linguistique, vous pouvez élargir votre recherche aux mots ayant la même signification que les termes de recherche. Par exemple, lorsque vous recherchez le mot "brochure", il peut s'avérer utile de rechercher également ses synonymes. Pour cela, indiquez ce qui suit :

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

Lorsque vous utilisez SYNONYM FORM OF, les synonymes du terme sont reliés par l'opérateur logique OR et l'argument de recherche est interprété comme suit :

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

Les synonymes sont définis dans un dictionnaire fourni avec DB2 Extension Texte. Le dictionnaire par défaut utilisé pour les synonymes est toujours US\_ENGLISH, et non celui de la langue spécifiée dans les définitions de la configuration de texte.

Vous pouvez changer de dictionnaire pour une requête spécifique en indiquant une autre langue. Voici un exemple :

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

#### Conseil

Vous ne pouvez pas utiliser le mot clé SYNONYM si le terme de recherche contient des caractères génériques, ou si NOT est utilisé avec l'argument de recherche.

### Recherche linguistique

DB2 Extension Texte offre un large éventail de fonctions linguistiques puissantes permettant d'effectuer une recherche basée sur les termes de recherche fournis. Ces fonctions sont mises en oeuvre pour les index linguistiques. Les fonctions linguistiques sont décrites dans le «Chapitre 15. Traitement linguistique des index linguistiques et exacts» à la page 255.

Ce type de fonction permet, par exemple, de rechercher un terme au pluriel, tel que "utilities" et de trouver son singulier, "utility". Le pluriel de ce terme est réduit à sa forme de base, utility, à l'aide d'un dictionnaire anglais, avant le début de la recherche.

## Définition des arguments de recherche

Cependant, ce dictionnaire ne dispose pas d'informations suffisantes pour réduire les variantes de termes d'autres langues à leur forme de base. Pour rechercher le pluriel d'un terme dans une autre langue, vous devez utiliser le dictionnaire correspondant.

Si vous indiquez GERMAN, par exemple, vous pouvez rechercher "geflogen" (volé) et trouver toutes les variantes de sa forme de base "fliegen" (voler)—, mais également d'autres formes conjuguées comme "fliege" ou "fliegt", et ainsi de suite.

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

### Conseil

Lorsque vous effectuez des recherches dans des documents rédigés dans d'autres langues, indiquez la langue correspondant dans l'argument de recherche *sans tenir compte de la langue par défaut*.

Si vous n'indiquez que la forme de base d'un terme de recherche, et jamais ses variantes, il est inutile d'indiquer une langue.

Pour en comprendre la raison, pensez à ce qui se produit lors de l'indexation du texte de votre base de données. Si vous utilisez un index linguistique, toutes les variantes d'un terme sont réduites à leur forme de base avant que les termes ne soient stockés dans l'index. Cela signifie que dans la table DB2TX.SAMPLE, bien que le terme "décompresser" apparaisse dans la première entrée de la colonne COMMENT, et "décompression" dans la seconde entrée, l'index ne contient que la forme de base "décompresser" et identifie ce terme (ou ses variantes) comme figurant dans ces deux entrées.

Par la suite, si vous recherchez la forme de base "décompresser", vous obtiendrez toutes ses variantes. Par contre, si vous recherchez une variante telle que "décompression", vous ne la trouverez pas directement. Vous devez indiquer un dictionnaire approprié pour la recherche, afin que la variante puisse d'abord être convertie à sa forme de base.

### Recherche à l'aide de l'opérateur booléen NOT

Vous pouvez utiliser l'opérateur booléen NOT pour exclure certains documents texte d'une recherche. Par exemple :

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

Tout document texte contenant le terme "DB2" est exclu de la recherche relative au terme "compresser" ou "compiler".

## Définition des arguments de recherche

Vous ne pouvez pas combiner l'opérateur NOT avec IN SAME SENTENCE AS ou IN SAME PARAGRAPH AS, décrit à la section «Recherche de termes d'une même phrase ou d'un même paragraphe» à la page 89, ni avec SYNONYM FORM OF, décrit à la section «Recherche de synonymes» à la page 90.

Vous pouvez utiliser l'opérateur NOT uniquement avec une recherche primaire, autrement dit vous ne pouvez pas combiner les opérateurs &, | et NOT (reportez-vous à la section «Syntaxe de l'argument de recherche» à la page 217).

Exemple d'utilisation de NOT **non** autorisée :

```
NOT("compresser" & "compiler")
```

Par contre, vous pouvez indiquer :

```
NOT("compresser" , "compiler")
```

## Recherche de correspondance partielle

La recherche de correspondance “partielle” permet d'effectuer des recherches sur des mots orthographiés de manière similaire au terme de base. Elle peut être utilisée dans le cadre de l'index Ngram.

Par exemple :

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

Une telle recherche permet de trouver une occurrence du mot mal orthographié `compresser`.

Le niveau de correspondance, “2” dans l'exemple précédent, indique le degré de précision. Il existe cinq niveaux de correspondance, le niveau 1 faisant référence à la correspondance la plus faible (environ 20 %) et le niveau 5 indiquant une quasi-concordance des termes (similitude à environ 90 %). Il est intéressant d'utiliser la recherche de correspondance partielle lorsqu'il y a des possibilités d'erreurs orthographiques dans un document, comme le cas se présente souvent dans des textes saisis à l'aide d'un logiciel de reconnaissance optique ou vocale.

### Respect de la segmentation

La recherche “segmentée” a été développée spécialement pour la langue coréenne. Elle garantit que DB2 Extension Texte respecte la segmentation des mots au cours de la recherche. Par exemple :

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

### Recherche d’homophones

La recherche “homophonique” permet de trouver des mots dont la consonance ressemble à celle de l’argument de recherche. Elle s’avère utile lorsque des documents peuvent contenir des termes à consonance identique, mais s’orthographiant différemment. Le terme allemand prononcé MEYER, par exemple, connaît différents homophones.

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

La recherche intégrera alors les occurrences suivantes : “Meyer”, “Mayer” et “Maier”.

### Recherche par thésaurus

La recherche par thésaurus fait également partie des puissantes fonctions d’extension de recherche offertes par DB2 Extension Texte. Les termes supplémentaires recherchés proviennent d’un thésaurus que vous aurez vous-même constitué, si bien que vous pouvez les contrôler directement. En effectuant une recherche sur “base de données”, par exemple, vous pourrez trouver des termes comme “référentiel” ou “DB2”.

Ce type de recherche s’applique à des domaines spécifiques faisant l’objet de fréquentes recherches, et pour lesquels il vaut mieux prendre le temps de constituer un thésaurus afin d’améliorer l’efficacité des recherches.

Reportez-vous à la section «Thésaurus» à la page 268, pour plus de détails sur la constitution d’un thésaurus. L’exemple de la figure 18 à la page 269, est un bref extrait d’un thésaurus sur les bases de données. Il est utilisé dans les exemples suivants pour illustrer la syntaxe applicable aux recherches par thésaurus.

Cet exemple étend la recherche portant sur le terme “système de gestion d’une base de données relationnelle objets” en ajoutant toutes les *instances* de ce terme trouvées dans le thésaurus “mestermes”. Ici, “DB2” est ajouté à la recherche.

## Définition des arguments de recherche

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS(COMMENTHANDLE,
                              'THESAURUS "mestermes"
                              EXPAND "INST"
                              TERM OF "système de gestion d'une base de données relationnelle objet") = 1
```

L'exemple suivant étend la recherche portant sur le terme "système de gestion de documents" en ajoutant tous les *synonymes* de ce terme. Il y a un synonyme : "bibliothèque".

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS(COMMENTHANDLE,
                              'THESAURUS "mestermes"
                              EXPAND "SYN"
                              TERM OF "système de gestion de documents") = 1
```

## Recherche sur texte libre et recherche hybride

La "recherche sur texte libre" est une procédure où le terme recherché est exprimé sous forme de texte libre, c'est-à-dire sous la forme d'une expression ou d'une phrase décrivant en langage naturel l'objet de la recherche. L'ordre des mots n'a pas d'importance dans une interrogation en texte libre. De plus, les *affinités lexicales* sont prises en charge. Dans le processus d'extraction, il s'agit de certaines combinaisons de mots contenus dans un terme de requête indiqué en texte libre, et apparaissant dans l'ensemble de textes selon une fréquence et une distance minimales déterminées. Pour les documents en anglais, cette longueur est fixée à cinq mots.

Notez que le masquage de caractères ou de mots n'est pas admis dans les chaînes de recherche d'un argument écrit en texte libre.

Par exemple :

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS(COMMENTHANDLE,
                              'IS ABOUT "toute information associée à l'installation d'AIX"') = 1
```

La recherche hybride combine les caractéristiques de la recherche booléenne et de la recherche sur texte libre. Par exemple :

```
db2=>SELECT DATE, SUBJECT
        FROM DB2TX.SAMPLE
        WHERE DB2TX.CONTAINS(COMMENTHANDLE,
                              '"DB2" & IS ABOUT "toute information associée à l'installation d'AIX"') = 1
```

## Affinement d'une recherche

Lorsque de trop nombreuses occurrences d'un argument de recherche sont proposées, il est souvent nécessaire de réduire ou d'*affiner* la recherche en combinant l'argument de recherche initial avec un second argument dans une relation booléenne de type AND.

Vous pouvez affiner les résultats de recherche sans utiliser la fonction REFINE, en stockant les résultats dans une table et en y effectuant la recherche suivante. Cependant, en fonction du nombre de termes qualifiants, cette méthode est moins efficace que celle qui consiste à stocker le dernier argument de recherche et à utiliser REFINE.

Les étapes suivantes indiquent comment effectuer une recherche, puis comment l'affiner à l'aide de la fonction REFINE. Cette fonction renvoie un argument de recherche qui est une combinaison booléenne de type AND de ses deux paramètres d'entrée. L'argument de recherche combiné renvoyé par REFINE est une valeur de type LONG VARCHAR.

1. Création d'une table pour les arguments de recherche précédents.

Créez une table PREVIOUS\_SEARCHES afin d'y placer les arguments des recherches qui ont déjà été effectuées.

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

### PREVIOUS\_SEARCHES

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

2. Recherche du premier argument de recherche.

Recherchez le mot "compresseur" dans le modèle de table.

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

Insérez l'argument de recherche dans la table PREVIOUS\_SEARCHES en vue de son utilisation dans les étapes suivantes.

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

## Affinement d'une recherche

### PREVIOUS\_SEARCHES

STEP	SEARCHARGUMENT
1	"compresser"

### 3. Affinement de la recherche.

Le nombre de documents texte renvoyés étant trop nombreux, vous devez affiner la recherche en combinant le terme de recherche précédent avec le mot "compiler" à l'aide de la fonction 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
```

Insérez l'argument de recherche affiné dans la table PREVIOUS\_SEARCHES en vue de son utilisation dans les étapes suivantes.

```
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	"compresser"
2	"compresser" & "compilateur"

Vous pouvez répéter cette étape jusqu'à ce que le nombre de documents texte soit suffisamment réduit.

### Définition et extraction d'informations dans les descripteurs

Les descripteurs contiennent le CCSID, le format et la langue de leurs documents texte. Les descripteurs associés aux fichiers externes contiennent, en outre, un pointeur désignant le fichier concerné. Les descripteurs sont créés lorsque vous activez une colonne de texte ou des fichiers externes.

Les fonctions DB2 Extension Texte décrites dans la présente section permettent de définir ou de modifier les informations de texte dans les descripteurs.

### Définition des informations de texte lors de l'insertion d'un nouveau texte

►►—INIT\_TEXT\_HANDLE—┌(—format—,—langue—)—————►  
└(—CCSID—,—format—,—langue—,—nom-fichier—)┘

Lorsque vous exécutez la commande `ENABLE TEXT COLUMN` pour activer une colonne de texte qui contient déjà du texte, vous pouvez définir implicitement le format et la langue de ce dernier par les valeurs indiquées dans les définitions de la configuration de texte. Ces paramètres de format et de langue sont ensuite stockés dans le descripteur. Pour les modifier, indiquez explicitement les valeurs souhaitées dans la commande `ENABLE TEXT COLUMN`.

Lorsque vous lancez la commande `ENABLE TEXT FILES`, vous pouvez également définir le CCSID et l'emplacement du document.

Lorsque, par la suite, vous insérez une ligne contenant du texte, un déclencheur d'insertion crée un descripteur et définit le format et la langue du texte par les valeurs utilisées lors de l'activation de la colonne de texte.

Pour indiquer d'autres valeurs, utilisez la fonction `INIT_TEXT_HANDLE` dans la commande `INSERT`. Lorsque la ligne est en cours d'insertion, la fonction `INIT_TEXT_HANDLE` crée un descripteur partiellement initialisé contenant les valeurs relatives au format et à la langue, indiquées par vos soins. Ensuite, les autres valeurs du descripteur sont indiquées par le déclencheur d'insertion.

Dans l'exemple suivant, `INIT_TEXT_HANDLE` prédéfinit la langue et le format dans un descripteur initialisé. La commande `INSERT` place ce descripteur dans la colonne `COMMENTHANDLE`.

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

## Définition et extraction d'informations dans les descripteurs

La valeur renvoyée par INIT\_TEXT\_HANDLE est de type DB2TEXTH ou DB2TEXTFH.

### Extraction des informations d'un descripteur



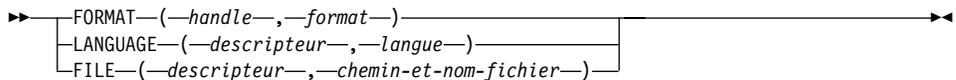
Voici un exemple d'extraction de CCSID dans un descripteur :

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

De la même manière, vous pouvez extraire le format ou la langue d'un document texte, ou l'emplacement des fichiers externes. L'exemple ci-dessous illustre l'utilisation de la fonction FORMAT. Elle renvoie le nombre de documents ASCII (TDS) contenus dans le modèle de table.

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

### Modification des informations d'un descripteur



Les fonctions FORMAT et LANGUAGE peuvent également modifier les spécifications correspondantes dans un descripteur. Ces fonctions renvoient le descripteur modifié comme une valeur de type DB2TEXTH ou DB2TEXTFH.

L'exemple suivant indique comment modifier le paramètre de langue d'un document de texte.

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

En utilisant à nouveau la fonction LANGUAGE, vous pouvez constater que la modification a été prise en compte :

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

### Amélioration des performances de recherche

La fonction `SEARCH_RESULT` tire parti du concept DB2 de fonctions s'appliquant à des valeurs de table. Cette fonction, utilisée dans la clause `FROM` d'une instruction SQL, renvoie une table intermédiaire contenant le résultat de la recherche sur la chaîne spécifiée. La syntaxe de la chaîne de recherche est identique à celle décrite au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215. Cette fonction présente l'avantage d'être nettement plus rapide que `CONTAINS` ou `RANK` en présence de tables volumineuses.

La table renvoyée présente la structure suivante :

Column Name	Datatype
<code>HANDLE</code>	<code>DB2TX.DB2TEXTH, DB2TX.DB2TEXTFH</code>
<code>NUMBER_OF_MATCHES</code>	<code>INTEGER</code>
<code>RANK</code>	<code>DOUBLE</code>

Exemple :

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

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

où `T1.handle = T2.mydochandle`

Si vous n'avez pas défini le chemin d'accès en cours, la dernière ligne de l'exemple est comme suit :

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

La commande `SELECT NUMBER_OF_MATCHES,RANK,HANDLE` renvoie la totalité des trois éléments mais vous pouvez préciser la combinaison de votre choix. Par exemple, vous pouvez choisir d'omettre `RANK` pour éviter le traitement intensif associé.

Si vous n'avez besoin que de la valeur `HANDLE`, vous pouvez utiliser simplement `SELECT COUNT(*)`.

## Récapitulatif des modifications

---

## Chapitre 7. Administration

Le présent chapitre décrit le mode de gestion des index, l'extraction d'informations utiles et l'inversion du processus de préparation de texte.

---

### Gestion des index de texte

Les tâches relatives à la gestion d'un index sont les suivantes :

- Mise à jour d'un index
- Modification des paramètres d'un index
- Réinitialisation de l'état d'un index
- Suppression des événements d'un index
- Réorganisation d'un index.

Vous pouvez exécuter ces tâches à tout moment sans ordre imposé.

### Mise à jour d'un index

#### Résumé

##### Utilisation

Lorsqu'un index doit être immédiatement mis à jour sans attendre l'indexation périodique. (Pour plus d'informations sur l'indexation périodique, reportez-vous à la section «Activation d'une colonne de texte» à la page 66.)

##### Commande

UPDATE INDEX

**Droits** ALTER, SELECT, UPDATE sur la table

Cet exemple met à jour l'index d'une table à index commun :

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

Cet exemple met à jour l'index pour une colonne d'une table à index multiples :

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

Utilisez cette commande pour mettre immédiatement à jour l'index sans attendre que la prochaine indexation périodique se produise

## Gestion des index de texte

automatiquement. Ceci s'avère utile si vous avez ajouté plusieurs documents texte à une base de données et que vous souhaitez y faire des recherches immédiatement.

DB2 Extension Texte indexe les documents texte de cette colonne (ou de toutes les colonnes de la table) qui ont été insérés ou modifiés, et supprime de l'index les termes provenant des documents qui ont été supprimés. La table de journalisation associée à l'index contient des informations sur les documents qui ont été insérés, mis à jour et supprimés.

### Mise à jour d'un index pour y intégrer les références de fichiers externes

Une table de journalisation ne contient pas automatiquement les modifications apportées aux fichiers externes que vous avez pu indexer (reportez-vous à la section «Activation de fichiers de texte externes» à la page 73), telles que le remplacement d'un document par une version plus récente de celui-ci, ayant le même nom et le même chemin absolu. Les mises à jour opérées sur ces fichiers ne peuvent être contrôlées par DB2 Extension Texte dans les tables de journalisation, étant donné qu'elles ne se produisent pas dans le cadre de DB2 UDB.

Pour que les mises à jour des fichiers externes soient répercutées dans un index DB2 Extension Texte, procédez comme suit :

1. Forcez l'insertion d'une entrée de type "modification" dans la table de journalisation en exécutant une instruction UPDATE sur la colonne de descripteurs concernée qui, de fait, n'est plus opérationnelle :

```
UPDATE table
SET   filehandlecol = filehandlecol
WHERE DB2TX.FILE(filehandlecol) = nom-fichier
```

où *nom-fichier* correspond au nom et au chemin complet du fichier externe qui a été modifié.

2. Lancez UPDATE INDEX pour mettre l'index à jour, y compris les modifications apportées au fichier externe.

### Modification des paramètres d'un index

#### Résumé

#### Utilisation

Lorsque la fréquence de mise à jour d'un index doit être modifiée.

#### Commande

CHANGE INDEX SETTINGS

**Droits** ALTER, SELECT, UPDATE sur la table

Utilisez cette commande pour modifier la fréquence de mise à jour d'un index.

### Fréquence de mise à jour

Pour plus d'informations, reportez-vous à la section «Mise à jour d'un index» à la page 46. Si vous ne spécifiez pas de fréquence de mise à jour, les paramètres en cours ne sont pas modifiés.

Utilisez le mot clé `UPDATEINDEX` pour déterminer si l'indexation des documents texte débute immédiatement, ou seulement lors de la prochaine indexation périodique. Si vous n'utilisez pas ce mot clé, le paramètre en cours n'est pas modifié.

### Exemples

Pour modifier la fréquence de mise à jour d'un index, afin qu'il soit mis à jour à 12 ou à 15 heures, du lundi au vendredi, s'il existe au moins 100 documents texte en attente :

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

Pour arrêter la mise à jour périodique d'un index :

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

## Réinitialisation de l'état d'un index

### Résumé

#### Utilisation

Lorsqu'il est impossible d'effectuer une recherche sur un index ou de le modifier.

#### Commande

```
RESET INDEX STATUS
```

**Droits** Aucun

Il peut arriver, dans certaines situations, que vous ne puissiez ni effectuer de recherche sur un index, ni le mettre à jour. La section «Affichage de l'état d'un index» à la page 108, vous explique comment déterminer si l'un de ces événements s'est produit. La commande `RESET INDEX STATUS` réactive l'index pour que vous puissiez à nouveau l'utiliser.

## Gestion des index de texte

Cet exemple réinitialise l'état d'un index dans le cas d'une table à index commun :

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

Vous pouvez également réinitialiser l'état de l'index pour une colonne de texte spécifique. L'exemple ci-dessous s'applique à l'index d'une colonne de table à index multiples :

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

### Suppression des événements d'un index

#### Résumé

##### Utilisation

Lorsque vous n'avez plus besoin des messages consignés dans la table de journalisation de l'index.

##### Commande

```
DELETE INDEX EVENTS
```

**Droits** Aucun

S'il vous est impossible d'effectuer une recherche sur un index, de le mettre à jour ou d'indexer un document, cet incident sera considéré par le système comme un *événement* d'indexation. Les informations relatives à ce type d'événements sont consignées dans la table de journalisation de l'index. Elles peuvent vous aider à déterminer l'origine du problème. Lorsque vous n'avez plus besoin de ces messages, vous pouvez les supprimer.

Cet exemple montre comment supprimer les messages d'indexation dans le cas d'une table à index commun :

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

Vous pouvez également supprimer les événements d'indexation relatifs à une colonne de texte spécifique. L'exemple ci-dessous s'applique à l'index d'une colonne de table à index multiples :

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

## Réorganisation d'un index

### Résumé

#### Utilisation

Quand la commande GET INDEX STATUS indique qu'il est nécessaire de réorganiser manuellement un index.

#### Commande

REORGANIZE INDEX

#### Droits

Aucun

Si une colonne de texte est mise à jour fréquemment, une recherche avec index peut devenir inefficace. Pour rendre son efficacité à la recherche, l'index doit être *réorganisé*. DB2 Extension Texte est capable d'estimer le moment où un index a besoin d'être réorganisé et effectue cette opération automatiquement, en arrière-plan. Cependant, dans certains cas, il peut être nécessaire de réorganiser un index manuellement par la commande REORGANIZE INDEX. Vous pouvez utiliser la commande GET INDEX STATUS pour savoir si un index a besoin d'une réorganisation.

Des recherches peuvent être effectuées pendant l'exécution de REORGANIZE INDEX, mais non des mises à jour d'index.

Cet exemple réorganise l'index dans le cas d'une table à index commun :

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

Cet exemple réorganise l'index dans le cas d'une colonne de table à index multiples :

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

---

## Extraction d'informations utiles

Cette section décrit les commandes permettant l'affichage des informations suivantes :

- Précisions sur l'état Activé (ou non) des bases de données, tables, colonnes et fichiers
- Paramètres des variables d'environnement
- Définitions de la configuration de texte
- Etat de l'index
- Evénements d'erreur
- Paramètres d'index

## Extraction d'informations utiles

- Paramètres de texte pour une colonne

### Affichage des informations d'état

#### Résumé

##### Utilisation

Lorsque vous avez besoin d'informations sur l'état d'activation de bases de données, de tables, de colonnes de texte ou de fichiers externes.

##### Commande

GET STATUS

**Droits** Aucun

Entrez :

```
db2tx GET STATUS
```

Voici un exemple de données de sortie affichées par GET STATUS. Elles indiquent l'état Activé (ou non) de la base de données et de l'ensemble des tables, colonnes de texte et les fichiers qu'elle contient.

```
Database is enabled for Text Extender
```

```
Table DB2TX.MYTABLE is enabled as a common-index table
```

```
Table DB2TX.MYTABLE is enabled as a common-index table
```

TextColumnName	HandleColumnName
-----	-----
COMMENT	COMMENTHANDLE

```
Table DB2TX.TEST is enabled as a multi-index table
```

TextColumnName	HandleColumnName
-----	-----
ABSTRACT1	ABSTRACT1HANDLE
ABSTRACT2	ABSTRACT2HANDLE

## Affichage des paramètres des variables d'environnement

### Résumé

#### Utilisation

Lorsque vous avez besoin d'informations relatives aux paramètres des variables d'environnement.

#### Commande

GET ENVIRONMENT

**Droits** Aucun

Entrez :

```
db2tx GET ENVIRONMENT
```

Voici un exemple de données de sortie affichées par GET ENVIRONMENT. Elles indiquent les valeurs en cours des variables d'environnement de DB2 Extension Texte.

```
Current database manager instance (DB2INSTANCE) = db2
Default database name             (DB2DBDFT) = my_database
Instance name                      (DB2TX_INSTOWNER) = user1
Instance directory                 (DB2TX_INSTOWNERHOMEDIR) = /usr/instance1
```

## Affichage des définitions de la configuration de texte

### Résumé

#### Utilisation

Lorsque vous avez besoin de connaître les paramètres par défaut des caractéristiques de texte, d'index et de traitement.

#### Commande

GET TEXT CONFIGURATION

**Droits** Aucun

Ces paramètres sont décrits à la section «Définitions de la configuration de texte» à la page 19. Pour les modifier, reportez-vous à la section «Modification de la configuration du texte» à la page 20.

Pour visualiser la configuration de texte en cours, entrez :

```
db2tx GET TEXT CFG
```

## Extraction d'informations utiles

Voici un exemple de données de sortie affichées par GET TEXT CONFIGURATION. Elles indiquent les définitions de la configuration de texte.

```
Coded character set ID      (CCSID) = 850
Language                   (LANGUAGE) = US_ENGLISH
Format                     (FORMAT) = TDS

Index type                 (INDEXTYPE) = LINGUISTIC
Update frequency          (UPDATEFREQ) = NONE
Index directory            (DIRECTORY) = user1/db2tx/index

Update index option       (UPDATEINDEX) = UPDATE
Commit count              (COMMITCOUNT) = 10 000
Tablespace name           (TABLESPACE) = TXLOG
```

## Affichage de l'état d'un index

### Résumé

#### Utilisation

Lorsque vous avez besoin de savoir si une recherche peut être effectuée sur un index ou s'il peut être mis à jour.

#### Commande

```
GET INDEX STATUS
```

**Droits** Aucun

Il peut arriver, dans certaines situations, que vous ne puissiez ni effectuer de recherche sur un index, ni le mettre à jour. Dans de tels cas, des messages susceptibles de vous aider à déterminer l'origine de l'événement sont stockés dans la table de journalisation de l'index. Il peut donc être utile de vérifier l'état d'un index et de savoir si de tels messages sont disponibles.

Cet exemple permet de visualiser l'état d'un index dans le cas d'une table à index commun :

```
db2tx GET INDEX STATUS db2tx.mytable
```

Vous pouvez également visualiser l'état de l'index pour une colonne de texte spécifique. L'exemple ci-dessous s'applique à l'index d'une colonne de table à index multiples :

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

Voici un exemple de données de sortie affichées par GET INDEX STATUS.

```
Node 1
Search status           = Search available
Update status           = Update available
Reorg status            = started 13.55
```

Scheduled documents	= 0
Indexed documents	= 187000
Primary index documents	= 130000
Secondary index documents	= 57000
Error events	= No error events

Si l'index est réparti sur plusieurs noeuds, les informations sont affichées pour chaque noeud.

### Search status

Indique si vous pouvez utiliser la colonne de descripteurs spécifiée pour effectuer une recherche sur l'index. Si celle-ci est impossible, vérifiez le code anomalie indiqué pour avoir plus d'informations sur l'incident, puis utilisez la commande `RESET INDEX STATUS` pour réactiver l'index. Pour plus de détails, reportez-vous au «Chapitre 20. Codes anomalie dus à des événements d'erreur» à la page 321.

### Update status

Indique si vous pouvez mettre à jour l'index d'une table ou colonne spécifique. Si cette opération est impossible, vérifiez le code anomalie indiqué pour avoir plus d'informations sur l'incident, puis utilisez la commande `RESET INDEX STATUS` pour réactiver l'index.

### Reorg status

Indique si vous pouvez réorganiser l'index d'une table ou colonne spécifique. Si cette opération n'est pas possible, vérifiez le code anomalie indiqué pour avoir plus d'informations sur l'incident. Une mise à jour en cours de l'index est une raison habituelle de l'indisponibilité de la réorganisation.

### Scheduled documents

Indique le nombre de documents en attente d'indexation (ou de retrait de l'index).

### Indexed documents

Indique le nombre de documents déjà indexés parmi ceux figurant dans la file d'attente des documents programmés.

### Primary index documents

Indique le nombre de documents dans l'index primaire.

### Secondary index documents

Indique le nombre de documents dans l'index secondaire.

### Error events

Indique le nombre d'événements consignés dans la table de journalisation de l'index. Pour visualiser ces informations, reportez-vous à la section «Affichage des événements d'erreur» à la page 110. Lorsque vous n'en avez plus besoin, vous pouvez les supprimer, comme indiqué à la section «Suppression des événements d'un index» à la page 104.

## Extraction d'informations utiles

### Affichage des événements d'erreur

Lorsque des incidents surviennent au cours de l'indexation, par exemple, lorsqu'il est impossible de trouver un document planifié pour cette opération, des messages appelés *événements d'erreur* sont consignés dans la table de journalisation de l'index.

Pour plus de détails sur les codes retour des événements, reportez-vous au «Chapitre 20. Codes anomalie dus à des événements d'erreur» à la page 321.

Vous pouvez consulter les événements d'erreur dans une vue de la table de journalisation dont le nom est `db2tx.LOGIXnnnnnn`, `IXnnnnnnn` représentant le nom de l'index, accessible à partir de la vue de catalogue.

Pour obtenir le nom de l'index, entrez :

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

L'écran des événements d'erreur se présente sous la forme suivante :

```
UPDATESTATUS  SMALLINT  
EVENTREASON   INTEGER  
EVENTMESSAGE  VARCHAR(1024)  
UPDATETIME    TIMESTAMP  
HANDLE        DB2TEXTX or DB2TEXTFH  
NODENUMBER    INTEGER
```

L'exemple ci-dessous illustre l'accès aux informations du journal de l'index :

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

### Affichage des paramètres d'index

#### Résumé

#### Utilisation

Lorsque vous avez besoin d'informations relatives aux paramètres d'un index.

#### Commande

```
GET INDEX SETTINGS
```

**Droits** Aucun

Cet exemple extrait les paramètres de l'index d'une table à index commun :

```
db2tx GET INDEX SETTINGS db2tx.mytable
```

L'exemple ci-dessous s'applique à l'index d'une colonne de table à index multiples :

```
db2tx GET INDEX SETTINGS db2tx.mytable  
                        HANDLE      commenthandle
```

Si la table est activée en tant que table à index multiples, cette commande affiche les paramètres d'index de toutes les colonnes de texte activées dans la table.

Voici un exemple de données de sortie affichées par GET INDEX SETTINGS dans le cas d'une table à index commun. Dans le cas d'une table à index multiples, l'information sera répétée autant de fois qu'il y a d'index. Vous pouvez également demander les paramètres d'index d'une colonne de texte spécifique.

Current index settings:

```
Index type           (INDEXTYPE) = LINGUISTIC  
Update index option (UPDATEINDEX) = UPDATE  
Update frequency    (UPDATEFREQ) = NONE  
Node 1  
Index directory      (DIRECTORY) = /home/user1/db2tx/index
```

Si l'index est réparti sur plusieurs noeuds, les informations de noeud sont affichées pour le répertoire d'index.

### Affichage des paramètres de texte pour une colonne

#### Résumé

#### Utilisation

Lorsque vous avez besoin d'informations relatives aux paramètres de texte d'une colonne.

#### Commande

```
GET TEXT INFO
```

**Droits** Aucun

Cet exemple permet d'obtenir les caractéristiques de texte d'un index dans le cas d'une table à index commun :

```
db2tx GET TEXT INFO db2tx.mytable
```

## Extraction d'informations utiles

Cet exemple permet d'obtenir les caractéristiques de texte d'un index dans le cas d'une table à index multiples :

```
db2tx GET TEXT INFO db2tx.mytable HANDLE commenthandle
```

Vous pouvez indiquer un nom de table et le nom d'une colonne de descripteurs.

Si vous n'indiquez qu'un nom de table dans la commande, les informations de texte pour chaque colonne activée dans cette table sont affichées. Si vous indiquez également un nom de colonne de descripteurs, seules sont affichées les informations de texte relatives à cette colonne.

Voici un exemple des informations affichées par cette commande dans le cas d'une table à index multiples :

```
Text information for column ABSTRACT1
      with handle column ABSTRACT1HANDLE:
Coded character set ID      (CCSID) = 850
Language                   (LANGUAGE) = US_ENGLISH
Format                     (FORMAT) = TDS

Text information for column ABSTRACT2
      with handle column ABSTRACT2HANDLE:
Coded character set ID      (CCSID) = 850
Language                   (LANGUAGE) = US_ENGLISH
Format                     (FORMAT) = TDS
```

---

## Utilisation de la vue de catalogue de DB2 Extension Texte

DB2 Extension Texte crée et tient à jour une vue de catalogue appelée DB2TX.TEXTINDEXES pour chaque base de données . Elle est générée lorsque vous exécutez la routine ENABLE DATABASE. Elle contient des informations relatives aux tables et aux colonnes activées en vue de leur utilisation par DB2 Extension Texte.

De nouvelles entrées sont créées dans DB2TX.TEXTINDEXES lors de l'activation d'une table, d'une colonne ou de fichiers externes. Les entrées sont mises à jour lors de la modification des paramètres d'index par la commande CHANGE INDEX SETTINGS. Elles sont supprimées si les colonnes ou les tables sont désactivées.

Vous pouvez accéder aux données de la vue de catalogue par le biais des fonctions de requête SQL normales. Cependant, vous ne pouvez pas modifier la vue de catalogue par les commandes de manipulation de données SQL normales. Vous ne pouvez pas non plus créer ou supprimer explicitement la vue de catalogue. Le tableau 7 à la page 113, présente le contenu de la vue de catalogue.

## Utilisation de la vue de catalogue de DB2 Extension Texte

Tableau 7. Vue de catalogue de DB2 Extension Texte

Nom de colonne	Type de données	Valeur NULL admise	Description
TABLESCHEMA	CHAR(8)	Non	Schéma de la table concernée par cette entrée.
TABLENAME	VARCHAR(18)	Non	Nom de la table concernée par cette entrée.
COLUMNNAME	VARCHAR(18)	Oui	Nom d'une colonne qui a été activée dans cette table. Cette valeur est définie par NULL si la table a été activée, mais qu'aucune colonne ne l'a été.
HANDLENAME	VARCHAR(18)	Oui	Nom d'une colonne de descripteurs. Cette valeur est définie par NULL si aucune colonne n'est activée dans la table TABLESCHEMA.TABLENAME.
INDEXNAME	CHAR(8)	Non	Nom de l'index de texte créé pendant l'activation de la table de texte ou d'une colonne de texte.
LOGTABLE	VARCHAR(18)	Non	Nom de la table de journalisation pour l'index INDEXNAME. La table DB2TX.LOGTABLE contient des informations sur les documents texte concernés par la prochaine mise à jour de l'index de texte et sur les événements d'erreur.
INDEXTYPE	VARCHAR(30)	Non	Type de l'index : LINGUISTIC, PRECISE, NGRAM.
MINIMUM	INTEGER	Oui	Nombre minimal de demandes de mise à jour d'index requises avant la mise à jour d'un index. Pour plus de détails, reportez-vous à la section «Mise à jour d'un index» à la page 46. Cette valeur est définie par NULL si la fréquence de mise à jour est définie par NONE.
DAYS	VARCHAR(15)	Oui	Jour de la mise à jour. Pour plus de détails, reportez-vous à la section «Mise à jour d'un index» à la page 46. Cette valeur est définie par NULL si la fréquence de mise à jour est définie par NONE.
HOURS	VARCHAR(75)	Oui	Heure de la mise à jour d'un index. Pour plus de détails, reportez-vous à la section «Mise à jour d'un index» à la page 46. Cette valeur est définie par NULL si la fréquence de mise à jour est définie par NONE.

## Utilisation de la vue de catalogue de DB2 Extension Texte

Tableau 7. Vue de catalogue de DB2 Extension Texte (suite)

Nom de colonne	Type de données	Valeur NULL admise	Description
MINUTES	VARCHAR(185)	Oui	Minute de la mise à jour. Pour plus de détails, reportez-vous à la section «Mise à jour d'un index» à la page 46. Cette valeur est définie par NULL si la fréquence de mise à jour est définie par NONE.
INDEXDIRECTORY	VARCHAR(254)	Non	Nom du répertoire, dans le système de fichiers, dans lequel est stocké l'index de texte.
UPDATEONCREATE	VARCHAR(10)	Non	Valeur "update" ou "noupdate", quelle que soit la valeur indiquée par l'option UPDATEINDEX dans ENABLE TEXT TABLE ou ENABLE TEXT COLUMN, ou dans la dernière commande CHANGE INDEX SETTINGS.
COMMONINDEX	VARCHAR(4)	Non	"oui" si la table TABLESCHEMA.TABLENAME est à index commun. "non" si la table TABLESCHEMA.TABLENAME est à index multiples.
CCSID	SMALLINT	Oui	CCSID de la colonne de texte TEXTCOLUMN définie par la commande d'activation de la colonne de texte. Cette valeur est définie par NULL si TEXTCOLUMN l'est aussi.
LANGUAGE	VARCHAR(30)	Oui	Nom du dictionnaire utilisé lors du traitement de la colonne de texte TEXTCOLUMN. Cette valeur est définie par NULL si TEXTCOLUMN l'est aussi.
FORMAT	VARCHAR(30)	Oui	Format indiqué pour la colonne de texte TEXTCOLUMN. Cette valeur est définie par NULL si TEXTCOLUMN l'est aussi.
FUNCTIONSCHEMA	CHAR(8)	Oui	Schéma de la fonction d'accès indiquée dans la commande ENABLE TEXT COLUMN par l'option FUNCTION. Cette valeur est définie par NULL si aucune option FUNCTION n'est indiquée.
FUNCTIONNAME	VARCHAR(18)	Oui	Nom de la fonction d'accès indiquée dans la commande ENABLE TEXT COLUMN par l'option FUNCTION. Cette valeur est définie par NULL si aucune option FUNCTION n'est indiquée.
PROTOTYPEHANDLE	VARCHAR(60)	Oui	Descripteur utilisé dans les fonctions de performances. Il ne contient que le nom de l'index commun à la colonne de texte complète.

## Utilisation de la vue de catalogue de DB2 Extension Texte

Tableau 7. Vue de catalogue de DB2 Extension Texte (suite)

Nom de colonne	Type de données	Valeur NULL admise	Description
INDEXOPTION	VARCHAR(30)	Oui	Option utilisée à la création de l'index : CASE_ENABLED.
INDEXPROPERTY	VARCHAR(30)	Oui	Propriété utilisée à la création de l'index : SECTIONS_ENABLED
NODENUMBER	INTEGER	Non	Numéro de noeud de la partition de table.

### Annulation de la procédure de préparation de texte

Lorsque du texte est préparé en vue de son utilisation par DB2 Extension Texte, certaines modifications lui sont apportées. La présente section décrit des fonctions permettant d'inverser ce processus.

#### Désactivation d'une colonne de texte

##### Résumé

##### Utilisation

Lorsque vous ne souhaitez plus effectuer de recherches de texte dans une colonne de texte.

##### Commande

DISABLE TEXT COLUMN

**Droits** ALTER, SELECT, UPDATE sur la table

Exemple:

```
db2tx DISABLE TEXT COLUMN db2tx.mytable
      HANDLE      commenthandle
```

Les événements suivants se produisent lorsque vous désactivez une colonne de texte :

- S'il s'agit d'une table à index multiples, c'est-à-dire que la colonne possède son propre index de texte et sa propre table de journalisation, ces derniers sont supprimés ainsi que les déclencheurs associés à la table de journalisation.
- S'il s'agit d'une table à index commun, c'est-à-dire où toutes les colonnes de texte partagent un même index, seuls les termes relatifs aux documents de cette colonne seront retirés de l'index commun. L'index, la table de journalisation et les déclencheurs associés ne sont supprimés que s'il s'agit de la dernière colonne de texte encore active dans la table.

## Annulation de la procédure de préparation de texte

### Désactivation de fichiers texte

#### Résumé

##### Utilisation

Lorsque vous ne souhaitez plus effectuer de recherches de texte dans un groupe de fichiers de texte externes.

##### Commande

DISABLE TEXT FILES

**Droits** ALTER, SELECT, UPDATE sur la table

Exemple:

```
db2tx DISABLE TEXT FILES db2tx.mytable
      HANDLE      commenthandle
```

Lorsque vous désactivez des fichiers de texte externes, voici ce qui se produit :

- L'index associé à cette colonne de descripteurs est supprimé.
- La table de journalisation et les déclencheurs sont supprimés.

### Désactivation d'une table de texte

#### Résumé

##### Utilisation

Lorsque vous ne souhaitez plus effectuer de recherches dans une table de texte.

##### Commande

DISABLE TEXT TABLE

**Droits** ALTER, SELECT, UPDATE sur la table

Exemple:

```
db2tx DISABLE TEXT TABLE db2tx.mytable
```

Lorsque vous désactivez une table de texte, voici ce qui se produit :

- S'il existe un index commun à toutes les colonnes de texte de la table, il est supprimé. Si chaque colonne de texte dispose de son propre index, *tous* les index sont supprimés.

## Annulation de la procédure de préparation de texte

- La table de journalisation commune utilisée pour enregistrer automatiquement les documents texte à indexer est supprimée. Si chaque colonne de texte dispose de sa propre table de journalisation, elles sont toutes supprimées.
- Les déclencheurs utilisés pour la gestion de la table de journalisation sont supprimés.
- Le contenu des colonnes de descripteurs est défini par NULL.

### Désactivation d'une base de données

#### Résumé

##### Utilisation

Lorsque vous ne souhaitez plus effectuer de recherches de texte dans cette base de données.

##### Commande

DISABLE DATABASE

**Droits** SYSADM ou DBADM sur la base de données

Pour désactiver la base de données connectée, entrez :

```
db2tx DISABLE DATABASE
```

Lorsque vous désactivez une base de données, les objets suivants sont supprimés :

- La vue de catalogue de DB2 Extension Texte créée lors de l'activation de la base de données
- La déclaration des fonctions SQL (UDF) de DB2 Extension Texte' et des types UDT DB2 Extension Texte pour cette base de données
- Tous les index relatifs aux tables et aux colonnes de texte de cette base de données
- Les tables de journalisation utilisées pour enregistrer automatiquement les documents texte à indexer, ainsi que les déclencheurs utilisés pour leur gestion

Les colonnes de descripteurs ne pouvant être supprimées et étant de types distincts, certains types distincts ne sont pas supprimés.

## Récapitulatif des modifications

---

## Chapitre 8. Utilisation des fonctions API pour la recherche et l'affichage

Le présent chapitre explique comment utiliser les fonctions de recherche et d'affichage de l'API DB2 Extension Texte. Pour obtenir une description détaillée de ces fonctions, reportez-vous au «Chapitre 13. Fonctions API de recherche et d'affichage» à la page 229. Des exemples de programmes utilisant les fonctions API sont fournis dans le «Chapitre 14. Modèle de programme API» à la page 253. Ce même chapitre décrit un modèle de fonction d'affichage `DesBrowseDocument`.

### Conseil

Avant d'effectuer une recherche, lisez la section «Types de recherche» à la page 40. Une recherche peut conduire à des résultats différents selon le type d'index utilisé. Utilisez l'option `GET INDEX SETTINGS` pour connaître le type d'index de texte associé au texte dans lequel vous effectuez une recherche.

---

### Configuration de l'application

Un programme d'application qui utilise l'API DB2 Extension Texte est une application DB2 CLI, car certaines des fonctions API nécessitent en entrée un descripteur de connexion à la base de données. Cela signifie que les règles devant être prises en compte pour les applications DB2 CLI s'appliquent également aux applications utilisant l'API DB2 Extension Texte.

Incluez dans l'application le fichier `des_ext.h` qui se trouve dans le sous-répertoire `include` du répertoire d'installation de DB2 Extension Texte.

Pour utiliser votre programme d'application avec l'API DB2 Extension Texte, liez-le à l'API.

#### Liaison d'une application UNIX

Vous devez lier la bibliothèque `libdesc1.a` à votre application. Cette bibliothèque se trouve dans le sous-répertoire `lib` du répertoire d'installation de DB2 Extension Texte.

#### Liaison d'une application OS/2 ou Windows

Toutes les fonctions API de DB2 Extension Texte sont contenues dans la bibliothèque dynamique `desclapi.dll`. Les appels de la fonction API externe sont résolus dans la bibliothèque d'importation `desclapi.lib`. Ces

## Configuration de l'application

bibliothèques se trouvent respectivement dans les sous-répertoires `dll` et `lib` du répertoire d'installation de DB2 Extension Texte.

Pour lier votre application à la bibliothèque dynamique d'API `desc1api.dll`, utilisez `desc1api.lib` comme fichier bibliothèque.

---

## Présentation des fonctions API

Vous trouverez ci-après la liste des fonctions de recherche (la première) et d'affichage (toutes les autres).

- DesGetSearchResultTable
- DesGetBrowseInfo
- DesStartBrowseSession
- DesOpenDocument
- DesGetMatches
- DesCloseDocument
- DesEndBrowseSession
- DesFreeBrowseInfo.

### Conseil

La plupart des fonctions API ont besoin d'un descripteur de connexion (hdbc). Vous devez le fournir via la fonction `SQLConnect`, mais cela ne vous interdit pas de lancer DB2 Extension Texte à partir de programmes en SQL intégré. Le manuel *DB2 Call Level Interface Guide and Reference* décrit comment associer des instructions CLI à des instructions en SQL intégré.

Vous pouvez utiliser les fonctions API pour les opérations suivantes :

- **Recherche de texte**

Dans ce scénario, seule la fonction de recherche `DesGetSearchResultTable` est utilisée. Elle nécessite en entrée un argument de recherche et un nom de colonne de descripteurs. Elle permet de rechercher du texte et de placer des informations relatives aux documents trouvés dans une table de résultats préparée préalablement.

Cette fonction est décrite plus en détail à la section «Recherche de texte» à la page 121. Reportez-vous également au «Chapitre 14. Modèle de programme API» à la page 253.

- **Affichage de texte**

Utilisez les fonctions suivantes dans l'ordre indiqué :

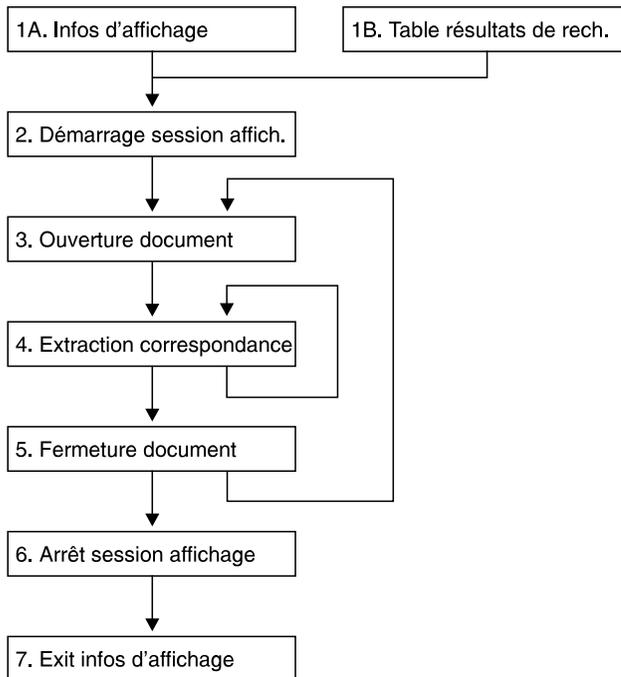


Figure 16. Séquence d'utilisation des fonctions API

Ces fonctions extraient les informations de mise en évidence, puis lancent une session d'affichage en vue d'afficher un document texte et de mettre en évidence les termes trouvés.

Dans une session d'affichage, vous pouvez ouvrir et afficher d'autres documents en utilisant ces mêmes informations de mise en évidence. Ces fonctions sont décrites plus en détail à la section «Affichage de texte» à la page 122.

---

## Recherche de texte

Il existe une API de recherche de texte : la fonction `DesGetSearchResultTable`.

### Obtention d'une table de résultats de recherche (`DesGetSearchResultTable`)

La fonction `DesGetSearchResultTable` reçoit un argument de recherche destiné à effectuer une recherche dans des documents texte dans une colonne de type texte particulière, et stocke les résultats dans une table. La table de résultats

## Recherche de texte

contient le descripteur de chaque document trouvé. Elle peut aussi classer les informations et le nombre de correspondances, selon l'option de recherche indiquée.

Il est possible d'obtenir ces mêmes informations en utilisant les fonctions RANK et NO\_OF\_MATCHES. Voici un exemple :

```
INSERT INTO RESULT
  SELECT COMMENTHANDLE,
         RANK(COMMENTHANDLE, '"procédures mémorisées"'),
         NO_OF_MATCHES(COMMENTHANDLE, '"procédures mémorisées"')
  FROM DB2TX.SAMPLE
  WHERE CONTAINS(COMMENTHANDLE, '"procédures mémorisées") = 1
```

La fonction DesGetSearchResultTable ne peut être utilisée que sur des tables de base, mais elle peut s'avérer plus rapide que les fonctions DB2 Extension Texte si la requête ne porte que sur du texte. Elle va en effet directement au serveur DB2 Extension Texte pour obtenir le rang et le nombre des correspondances et ne boucle qu'autant de fois qu'il y a documents correspondants. Par contre, dans l'exemple de fonction UDF, la fonction CONTAINS est appelée une fois pour chaque ligne de la table puis, c'est au tour des fonctions RANK et NO\_OF\_MATCHES d'être appelées. Pour chaque document trouvé, trois recherches distinctes sont effectuées.

### Entrée

Les paramètres d'entrée sont :

- le descripteur de connexion à la base de données ;
- la table de recherche ;
- le nom de la colonne de descripteurs associée à la colonne de texte sur laquelle portera la recherche ;
- un argument de recherche ;
- les options de recherche ;
- une option d'affichage (pour renvoyer des informations d'affichage) ;
- le nom de la table dans laquelle le résultat doit être stocké.

### Sortie

Si une option d'affichage est précisée, cette fonction renvoie un pointeur sur les informations d'affichage.

---

## Affichage de texte

Ce groupe de fonctions (figure 16 à la page 121) détermine les termes à mettre en évidence. Il lance ensuite une session d'affichage, ouvre un document et extrait les correspondances sous forme d'un flux de données pouvant être analysé par un programme d'application qui appelle l'afficheur.

### Extraction d'informations d'affichage (DesGetBrowseInfo)

La fonction DesGetBrowseInfo reçoit un argument de recherche et un descripteur. Elle renvoie un pointeur sur les informations d'affichage nécessaires à la fonction DesStartBrowseSession. Les informations d'affichage comprennent une liste de tous les termes devant être mis en évidence.

Il est également possible d'extraire des informations d'affichage en indiquant l'option Browse dans la fonction DesGetSearchResultTable.

#### Entrée

Les paramètres d'entrée sont :

- le descripteur de connexion à la base de données ;
- un descripteur ;
- un argument de recherche.

#### Sortie

Cette fonction renvoie un pointeur sur les informations d'affichage.

### Démarrage d'une session d'affichage (DesStartBrowseSession)

La fonction DesStartBrowseSession démarre une session d'affichage en établissant l'environnement nécessaire à l'affichage d'un document texte et en mettant en évidence ses correspondances. Elle reçoit un pointeur sur les informations d'affichage, envoyé par DesGetBrowseInfo ou DesGetSearchResultTable, et renvoie un descripteur de session d'affichage qu'utiliseront les autres fonctions d'affichage.

#### Entrée

Le paramètre d'entrée est :

- un pointeur sur les informations d'affichage, envoyé par DesGetBrowseInfo ou DesGetSearchResultTable ;
- un ID utilisateur ;
- un mot de passe.

#### Sortie

Cette fonction renvoie un descripteur de session d'affichage.

### Ouverture d'un document (DesOpenDocument)

La fonction DesOpenDocument reçoit un pointeur de session d'affichage, un descripteur et une option DES\_FAST ou DES\_EXTENDED indiquant le type d'analyse linguistique à utiliser pour la mise en évidence des termes trouvés (reportez-vous à l'«Étape 2 : Correspondance étendue» à la page 266). L'option DES\_FAST utilise l'analyse de texte de base, sans recours au dictionnaire, pour déterminer les termes à mettre en évidence. L'option DES\_EXTENDED utilise la correspondance étendue.

## Affichage de texte

La fonction `DesOpenDocument` prépare le document texte correspondant au descripteur, afin d'extraire le texte et les informations de mise en évidence, et renvoie un descripteur de document utilisé pour l'appel itératif de `DesGetMatches`.

### Entrée

Les paramètres d'entrée sont :

- un descripteur de session d'affichage, envoyé par `DesStartBrowseSession` ;
- un descripteur de texte ;
- une option de correspondance : `DES_FAST` ou `DES_EXTENDED`.

### Sortie

Cette fonction renvoie un descripteur de document qui est utilisé par `DesGetMatches` et `DesCloseDocument`.

## Extraction de correspondances (`DesGetMatches`)

La fonction `DesGetMatches` renvoie un pointeur sur les informations de mise en évidence concernant le document texte décrit par un descripteur de document. Les informations de mise en évidence constituent un flux de données. Elles comprennent le contexte (au moins un paragraphe) et les informations de mise en évidence de texte dans ce contexte. Le flux de données est décrit à la section «Syntaxe du flux de données» à la page 238. Un programme d'application peut analyser le flux de données et le traiter via l'afficheur de l'utilisateur.

L'option `DesGetMatches` ne renvoie qu'une partie du flux de données, en indiquant la longueur de celle-ci dans la structure de sortie.

Une série d'appels de `DesGetMatches` permet d'extraire la totalité du contenu du document texte. Lorsque la fin du document est atteinte, un descripteur est renvoyé.

### Entrée

Les paramètres d'entrée sont :

- un descripteur de session d'affichage
- un descripteur de document envoyé par `DesOpenDocument`

### Sortie

Cette fonction renvoie un pointeur sur une structure contenant la partie de flux de données et sa longueur.

## Fermeture d'un document (`DesCloseDocument`)

La fonction `DesCloseDocument` ferme un document texte ouvert par `DesOpenDocument` et libère la mémoire allouée lors du renvoi du texte et des informations de mise en évidence.

### **Entrée**

Les paramètres d'entrée sont :

- un descripteur de session d'affichage
- un descripteur de document envoyé par DesOpenDocument

### **Sortie**

Aucune.

### **Arrêt d'une session d'affichage (DesEndBrowseSession)**

La fonction DesEndBrowseSession arrête une session d'affichage lancée par DesStartBrowseSession et libère la mémoire allouée à la session d'affichage.

### **Entrée**

Le paramètre d'entrée est :

- un descripteur de session d'affichage.

### **Sortie**

Aucune.

### **Libération des informations d'affichage (DesFreeBrowseInfo)**

La fonction DesFreeBrowseInfo libère la mémoire allouée aux informations d'affichage par DesGetBrowseInfo.

### **Entrée**

Le paramètre d'entrée est :

- un pointeur sur les informations d'affichage.

### **Sortie**

Aucune.

## Récapitulatif des modifications

---

## Partie 2. Référence



## Chapitre 9. Commandes de préparation de texte et d'administration du client

Commande	Objet	Page
CHANGE INDEX SETTINGS	Modifie les caractéristiques d'un index	131
CHANGE TEXT CONFIGURATION	Modifie les définitions de la configuration de texte	133
CONNECT	Connecte à une base de données	137
DELETE INDEX EVENTS	Supprime les événements d'index de la table de journalisation	139
DISABLE DATABASE	Désactive une base de données pour DB2 Extension Texte	140
DISABLE TEXT COLUMN	Désactive une colonne de texte pour DB2 Extension Texte et supprime les index qui y sont associés	141
DISABLE TEXT FILES	Désactive les fichiers de texte pour DB2 Extension Texte et supprime les index qui y sont associés	142
DISABLE TEXT TABLE	Désactive une table pour DB2 Extension Texte et supprime les index qui y sont associés	143
ENABLE DATABASE	Prépare une base de données à son utilisation par DB2 Extension Texte	144
ENABLE TEXT COLUMN	Prépare une colonne de texte à son utilisation par DB2 Extension Texte et crée un index de texte spécifique à cette colonne	145
ENABLE TEXT FILES	Prépare les fichiers de texte à leur utilisation par DB2 Extension Texte et crée un index de texte spécifique à ces fichiers	154
ENABLE TEXT TABLE	Crée un index de texte commun pour une table	157
GET ENVIRONMENT	Affiche les paramètres en cours des variables d'environnement	162
GET INDEX SETTINGS	Affiche les caractéristiques d'un index	163
GET INDEX STATUS	Affiche les informations relatives à l'état d'un index	165
GET STATUS	Affiche l'état activé de bases de données, de tables et de colonnes	166
GET TEXT CONFIGURATION	Affiche les définitions de la configuration de texte	167
GET TEXT INFO	Affiche les informations de texte pour une colonne de texte	168
QUIT	Quitte le mode interpréteur de commandes de DB2 Extension Texte	169

Commande	Objet	Page
REORGANIZE INDEX	Réorganise un index pour améliorer les recherches	170
RESET INDEX STATUS	Réinitialise l'état d'un index pour en permettre la réutilisation	171
UPDATE INDEX	Met à jour un index de texte	172

Le présent chapitre décrit la syntaxe des commandes de préparation et d'administration du client. Le «Chapitre 5. Préparation d'un texte pour la recherche» à la page 57 et le «Chapitre 7. Administration» à la page 101 décrivent l'utilisation de ces commandes.

Avant d'utiliser ces commandes, démarrez l'interpréteur de commandes DB2 Extension Texte en entrant la commande `db2tx` dans le shell UNIX d'OS/390. Vous passez alors en mode d'entrée interactif dans lequel toutes les commandes ultérieures sont interprétées comme des commandes de DB2 Extension Texte. L'interpréteur de commandes est normalement démarré en même temps que DB2.

Pour quitter ce mode, entrez `QUIT`.

Comme dans DB2, vous pouvez émettre des commandes directement à partir de l'invite du système d'exploitation en indiquant le préfixe `db2tx`, et inclure des requêtes de l'interpréteur de commandes dans un fichier de commandes d'un script shell.

#### Conseil

Une commande précédée du préfixe `db2tx` active la connexion à la base de données par défaut indiquée dans la variable d'environnement `DB2DBDFT`. Les commandes ci-dessous activent la base de données par défaut, et non `MYDATABASE`.

```
db2tx CONNECT TO MYDATABASE
db2tx ENABLE DATABASE
```

## CHANGE INDEX SETTINGS

Cette commande modifie les caractéristiques d'un index **après** que la base de données a été activée.

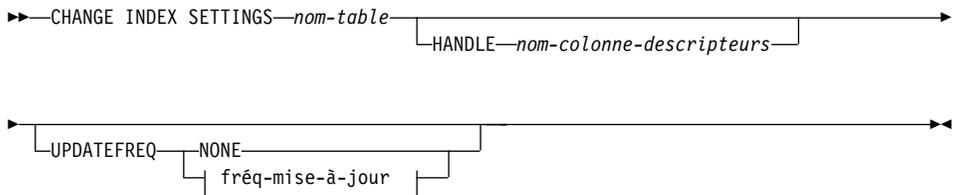
Pour modifier les définitions par défaut utilisées lorsqu'une base de données est activée pour la première fois, utilisez la commande «CHANGE TEXT CONFIGURATION» à la page 133.

### Droits

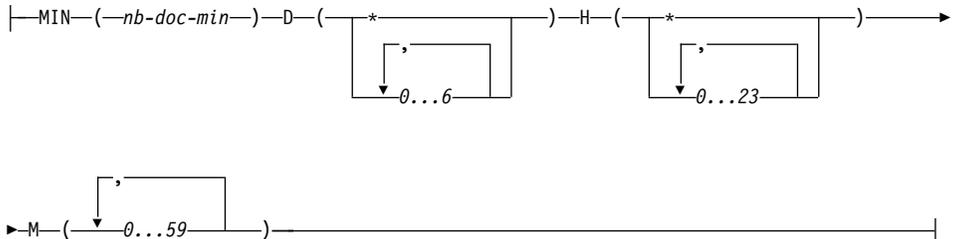
Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

### Syntaxe de la commande



### fréq-mise-à-jour:



### Paramètres de la commande

#### nom-table

Nom de la table de texte de la base de données connectée, qui contient la colonne de texte dont la fréquence de mise à jour d'index doit être modifiée. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### HANDLE nom-colonne-descripteurs

Nom de la colonne de descripteurs dont la fréquence de mise à jour d'index doit être modifiée. Ce mot clé est obligatoire si la colonne de

## Commande CHANGE INDEX SETTINGS

texte possède son propre index (c'est-à-dire si ce dernier a été créé par la commande ENABLE TEXT COLUMN),

mais il est ignoré si l'index a été créé par la commande ENABLE TEXT TABLE (la table possède donc un index de texte pour toutes les colonnes de texte).

### **UPDATEFREQ** fréq-mise-à-jour

Fréquence de mise à jour de l'index exprimée en fonction de la date et du nombre minimal de documents texte en attente dans la table de journalisation. Si cette dernière ne comporte pas suffisamment de documents texte au jour et à l'heure indiqués, l'index n'est pas mis à jour.

La syntaxe correspondante est décrite à la section «Mise à jour d'un index» à la page 46.

### **NONE**

Aucune autre mise à jour d'index n'est effectuée. Ceci concerne une colonne de texte qui ne fera plus l'objet de modifications.

Si vous omettez le mot clé UPDATEFREQ, les paramètres relatifs à la fréquence de mise à jour ne sont pas modifiés.

## CHANGE TEXT CONFIGURATION

Cette commande modifie les paramètres par défaut de la configuration de texte utilisée lorsqu'une base de données est activée. Il s'agit des définitions de la *configuration de texte*. Les paramètres de configuration de texte initiaux, définis lors de l'installation de DB2 Extension Texte, sont décrits à la section «Définitions de la configuration de texte» à la page 19.

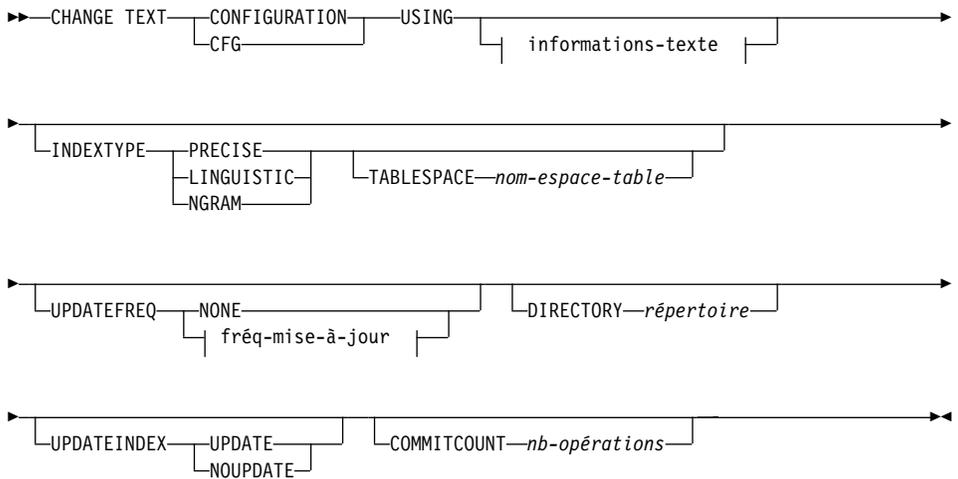
Pour modifier ces paramètres pour une base de données **après que** celle-ci a été activée, utilisez la commande «CHANGE INDEX SETTINGS» à la page 131.

### Droits

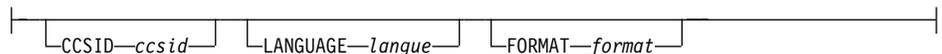
Vous devez disposer des droits suivants :

SYSADM  
DBADM

### Syntaxe de la commande

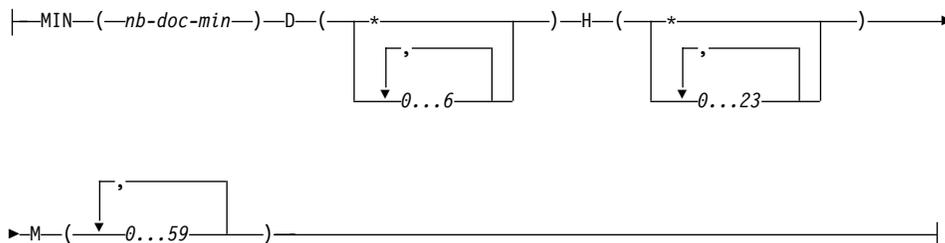


### informations-texte :



### fréq-mise-à-jour :

## Commande CHANGE TEXT CONFIGURATION



### Paramètres de la commande

#### INDEXTYPE

Pour modifier le type d'index par défaut, vous pouvez choisir l'une des options suivantes. Pour plus d'informations, reportez-vous à la section «Types de recherche» à la page 40.

#### PRECISE

Les termes sont indexés et recherchés sous la forme exacte dans laquelle ils apparaissent dans les documents texte.

#### LINGUISTIC

Les termes font l'objet d'un traitement linguistique avant d'être indexés, de même que les termes de recherche avant le début de la recherche.

#### NGRAM

L'indexation des termes se fait par analyse de séquences de caractères, au lieu de se fonder sur un dictionnaire. Il est obligatoire d'utiliser ce type de dictionnaire si les documents à indexer contiennent des caractères DBCS. Toutefois, un index Ngram peut être utilisé pour des documents SBCS.

#### TABLESPACE nom-espace-table

Indique le nom d'un espace table existant. L'espace table contient les tables propres index, créées par DB2 Extension Texte (à l'instar des tables de journalisation). Pour les tables volumineuses, utilisez un espace table séparé. Si vous n'en spécifiez pas, les tables seront créées dans l'espace table par défaut de DB2.

#### UPDATEFREQ fréq-mise-à-jour

Fréquence de mise à jour de l'index exprimée en fonction de la date et du nombre minimal de documents texte en attente dans la table de journalisation. Si cette dernière ne comporte pas suffisamment de documents texte au jour et à l'heure indiqués, l'index n'est pas mis à jour.

La syntaxe correspondante est décrite à la section «Mise à jour d'un index» à la page 46.

## Commande **CHANGE TEXT CONFIGURATION**

### **NONE**

Aucune autre mise à jour d'index n'est effectuée. Ceci concerne une colonne de texte qui ne fera plus l'objet de modifications.

### **DIRECTORY répertoire**

Répertoire dans lequel l'index de texte doit être stocké.

### **UPDATEINDEX**

Mot clé déterminant si les documents texte sont indexés immédiatement à l'issue de la commande utilisant cette option, sans attendre la prochaine indexation périodique définie par UPDATEFREQ. Les commandes concernées sont ENABLE TEXT COLUMN et ENABLE TEXT FILES.

### **UPDATE**

L'indexation des documents texte a lieu immédiatement après l'achèvement de la commande.

### **NOUPDATE**

L'indexation a lieu au moment défini par les paramètres de fréquence de mise à jour indiqués dans cette commande par UPDATEFREQ, ou dans les définitions de la configuration de texte.

### **COMMITCOUNT nb-opérations**

Valeur comprise entre 500 et 1 000 000, indiquant le nombre d'insertions ou de mises à jour effectuées avant qu'une instruction intermédiaire de validation DB2 UDB soit émise. L'utilisation de ce mot clé permet d'éviter que vous ne manquiez d'espace de journalisation lorsque vous activez des tables ou des colonnes volumineuses, ou un nombre élevé de fichiers externes.

### **CCSID ccsid**

Identificateur de jeu de caractères codés à utiliser lors de l'indexation des documents texte.

Pour plus d'informations sur les autres CCSID pouvant être pris en charge, reportez-vous à la section «CCSID» à la page 35.

### **LANGUAGE langue**

Langue dans laquelle le texte est rédigé. Ceci détermine le dictionnaire à utiliser lors de l'indexation de documents texte et des recherches dans ces derniers. Le «Chapitre 15. Traitement linguistique des index linguistiques et exacts» à la page 255, décrit le mode d'utilisation des dictionnaires.

Les langues prises en charge sont indiquées dans le tableau 5 à la page 54.

## Commande **CHANGE TEXT CONFIGURATION**

### **FORMAT format**

Type du document texte stocké, tel que WordPerfect, ou ASCII. Cette information est nécessaire à DB2 Extension Texte lors de l'indexation des documents. Vous trouverez la liste des formats de documents pris en charge à la section «Documents pris en charge» à la page 31.



## Commande CONNECT

Si vous utilisez un poste client OS/2 et qu'il s'agit de votre première connexion à une base de données depuis le démarrage du système, les fonctions Gestion des profils utilisateur (UPM) vous invitent à entrer un ID utilisateur et un mot de passe. Si vous êtes déjà connecté, les fonctions UPM tentent d'effectuer une connexion au serveur éloigné sous l'ID utilisateur et le mot de passe de la session en cours. Reportez-vous à la section «Démarrage de l'interpréteur de commandes DB2 Extension Texte» à la page 59.

## DELETE INDEX EVENTS

Cette commande supprime les événements d’indexation de la table de journalisation d’un index pour une colonne de descripteurs ou une table déterminée.

### Droits

Aucun.

### Syntaxe de la commande

```

▶▶—DELETE INDEX EVENTS—nom-table—┬──┬──▶
                                     └──HANDLE—nom-colonne-descripteurs—┘
    
```

### Paramètres de la commande

#### **nom-table**

Nom de la table de texte, dans la base de données connectée, dont les événements d’erreur doivent être supprimés de la table de journalisation. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs dont les messages doivent être supprimés de la table de journalisation.

### Commentaires

Si vous indiquez un nom de colonne de descripteurs, seuls les événements d’indexation la concernant seront supprimés.

Si vous nettoyez les tables de journalisation, vous devez supprimer les événements relatifs aux index après en avoir vérifié la cause et, éventuellement, supprimé la la source de l’erreur.

## Commande DISABLE DATABASE

---

### DISABLE DATABASE

Cette commande annule toute opération de préparation effectuée par DB2 Extension Texte pour une base de données et désactive toutes les tables de texte afin que DB2 Extension Texte ne les utilise pas.

#### Droits

Vous devez disposer des droits suivants :

SYSADM

DBADM

#### Syntaxe de la commande

►—DISABLE DATABASE—◄

#### Paramètres de la commande

Aucun.

#### Commentaires

Cette commande restaure la base de données connectée afin qu'elle ne fasse plus l'objet de recherches par DB2 Extension Texte, c'est-à-dire qu'elle désactive toutes les tables et les colonnes de texte de DB2 Extension Texte dans cette base de données. Toutes les modifications apportées à la base de données pour activer les tables et colonnes de texte de DB2 Extension Texte ainsi que les fichiers externes sont annulées : tous les index de texte correspondants, la vue de catalogue TEXTCOLUMNS de DB2 Extension Texte dans la base de données et tous les déclencheurs de DB2 Extension Texte sont supprimés.

### DISABLE TEXT COLUMN

Cette commande désactive une colonne de texte afin que DB2 Extension Texte ne l'utilise pas.

#### Droits

Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

#### Syntaxe de la commande

►►—DISABLE TEXT COLUMN—*nom-table*—HANDLE—*nom-colonne-descripteurs*—►►

#### Paramètres de la commande

##### **nom-table**

Nom de la table de texte de la base de données connectée qui contient la colonne à désactiver. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

##### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs à désactiver afin que DB2 Extension Texte cesse de l'utiliser.

#### Commentaires

L'index est supprimé.

La table de journalisation utilisée pour enregistrer les modifications apportées à la colonne de descripteurs (insertions, mises à jour et suppressions) est supprimée.

Les déclencheurs qui écrivent des entrées dans la table de journalisation sont supprimés.

La colonne de descripteurs n'est pas modifiée.

## Commande DISABLE TEXT FILES

---

### DISABLE TEXT FILES

Cette commande désactive un jeu de fichiers de texte externes pour que DB2 Extension Texte ne les utilise plus.

#### Droits

Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER

SELECT

UPDATE

#### Syntaxe de la commande

►—DISABLE TEXT FILES—*nom-table*—HANDLE—*nom-colonne-descripteurs*—◄

#### Paramètres de la commande

##### **nom-table**

Nom de la table de texte de la base de données connectée, qui contient la colonne de descripteurs concernant les fichiers de texte externes à désactiver. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

##### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs à désactiver afin que DB2 Extension Texte cesse de l'utiliser.

#### Commentaires

L'index est supprimé.

La table de journalisation utilisée pour enregistrer les modifications effectuées dans la colonne de descripteurs (insertions, mises à jour et suppressions) est supprimée. Les déclencheurs qui écrivent des entrées dans la table de journalisation sont également supprimés.

---

### DISABLE TEXT TABLE

Cette commande désactive toutes les colonnes de texte d'une table afin que DB2 Extension Texte ne les utilise pas.

#### Droits

Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

#### Syntaxe de la commande

►►—DISABLE TEXT TABLE—*nom-table*—◄◄

#### Paramètres de la commande

##### **nom-table**

Nom de la table de texte de la base de données connectée qui contient la colonne à désactiver. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### Commentaires

Cette commande rend toutes les colonnes de texte d'une table inutilisables par DB2 Extension Texte.

Si les colonnes de texte de cette table ont été activées individuellement par ENABLE TEXT COLUMN, tous les index de texte associés sont supprimés. (Pour désactiver des colonnes de texte et supprimer les index de texte qui leur sont associés séparément, utilisez la commande DISABLE TEXT COLUMN.) Si les colonnes de texte de cette table ont été activées collectivement par ENABLE TEXT TABLE, il existe un index commun à toutes ces colonnes. Cette commande supprime l'index commun.

Les tables de journalisation utilisées pour enregistrer les modifications effectuées dans les colonnes de texte (insertions, mises à jour et suppressions) sont supprimées. Les déclencheurs qui écrivent des entrées dans la table de journalisation sont supprimés.

## Commande ENABLE DATABASE

---

### ENABLE DATABASE

Cette commande active la base de données en cours afin qu'elle stocke des données de texte.

#### Droits

Vous devez disposer des droits suivants :

SYSADM

DBADM

#### Syntaxe de la commande

►—ENABLE DATABASE—◄

#### Paramètres de la commande

Aucun.

#### Commentaires

Avant d'exécuter cette commande, vous devez être implicitement ou explicitement connecté à une base de données (reportez-vous à la section consacrée à «CONNECT» à la page 137).

Cette commande prépare la base de données connectée en vue de son utilisation par DB2 Extension Texte. Cette étape est obligatoire avant d'activer une table ou une colonne de texte de DB2 Extension Texte dans la base de données .

ENABLE DATABASE crée une vue de catalogue de DB2 Extension Texte appelée DB2TX.TEXTINDEXES (décrite à la section «Utilisation de la vue de catalogue de DB2 Extension Texte» à la page 112) et une vue de catalogue appelée DB2TX.TEXTCOLUMNS, utilisée pour les requêtes relatives aux "performances".

Elle crée également les définitions de la configuration de texte décrites à la section «Définitions de la configuration de texte» à la page 19.

D'autres opérations sont également effectuées, comme la déclaration des types distincts DB2 Extension Texte et de fonctions DB2 Extension Texte.

## ENABLE TEXT COLUMN

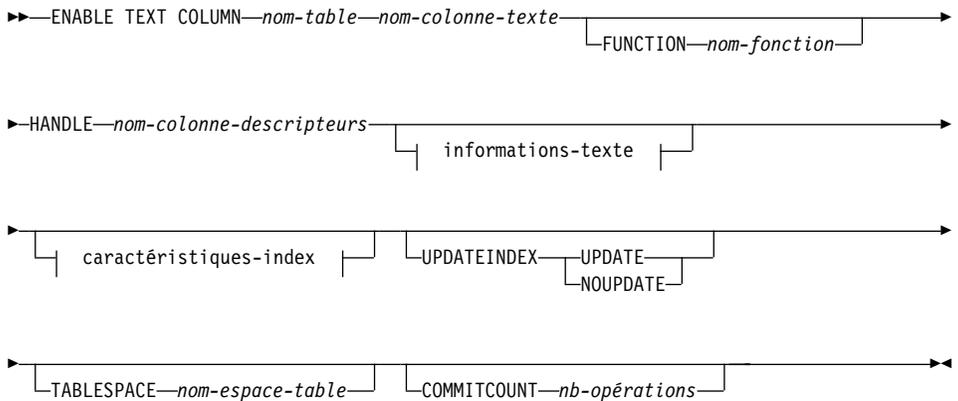
Cette commande active une colonne de texte en vue de son utilisation par DB2 Extension Texte.

### Droits

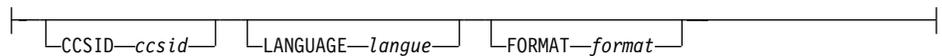
Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

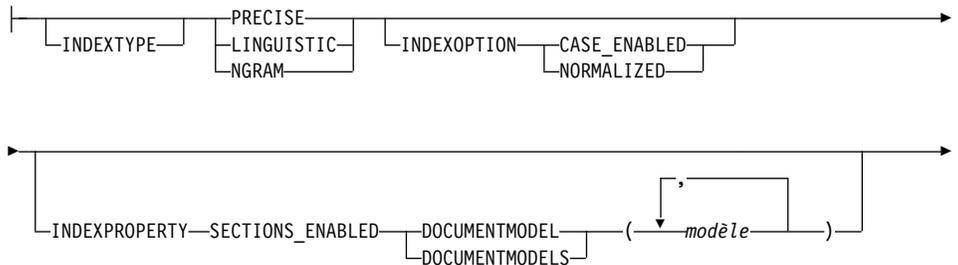
### Syntaxe de la commande



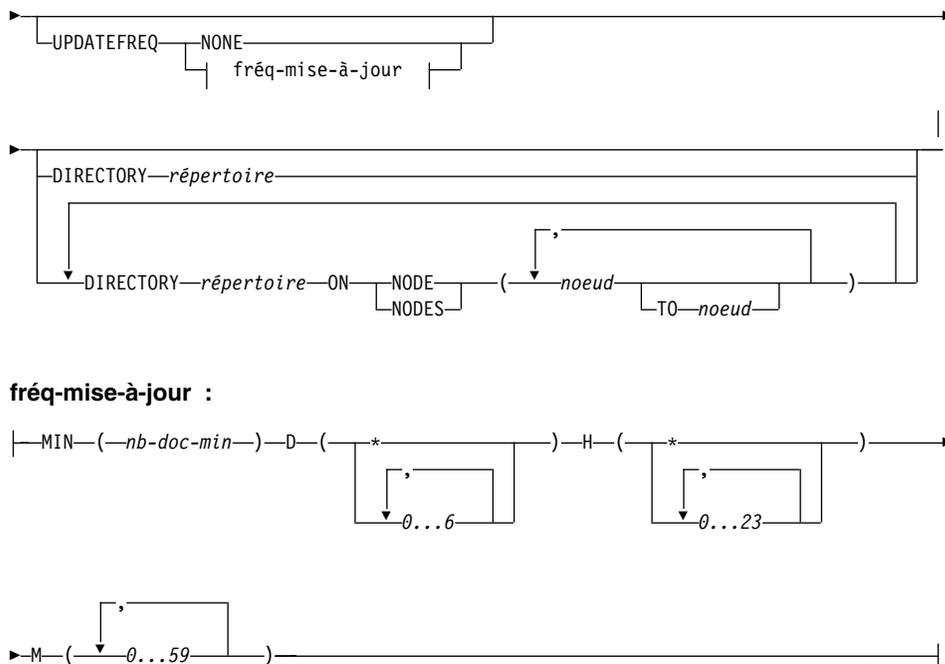
#### informations-texte :



#### caractéristiques-index :



## Commande ENABLE TEXT COLUMN



**freq-mise-à-jour :**

## Paramètres de la commande

### nom-table

Nom de la table de texte de la base de données connectée, qui contient la colonne à activer. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

### nom-colonne-texte

Nom de la colonne à activer en vue de son utilisation par DB2 Extension Texte. Cette colonne doit être de type `CHAR`, `VARCHAR`, `LONG VARCHAR`, `CLOB`, `DBCLOB`, `GRAPHIC`, `VARGRAPHIC` ou `LONG VARGRAPHIC`. Si le document n'est pas de l'un des types précédemment cités, utilisez le paramètre `FUNCTION` pour le convertir.

### FUNCTION nom-fonction

Nom de la fonction UDF qu'utilisera DB2 Extension Texte pour accéder aux documents texte qui se trouvent dans une colonne de type autre que `CHAR`, `VARCHAR`, `LONG VARCHAR`, `CLOB`, `DBCLOB`, `GRAPHIC`, `VARGRAPHIC` ou `LONG VARGRAPHIC`. Pour plus d'informations, reportez-vous à la section «Activation de colonnes de texte d'un type de données non pris en charge» à la page 71.

### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs à ajouter à la table en vue de son utilisation par les fonctions DB2 Extension Texte.

### **CCSID ccsid**

Identificateur de jeu de caractères codés à utiliser lors de l'indexation des documents texte.

Si vous indiquez un CCSID lors de l'activation d'une colonne de texte pour un index Ngram, il doit être identique à celui de la base de données , et le CCSID utilisé lors de la recherche (CCSID de la base de données ) doit correspondre à ce CCSID.

Pour déterminer le CCSID par défaut, utilisez la commande :

```
db2tx get text cfg
```

La valeur par défaut à l'installation est le CCSID de la base de données .

Si ce mot clé n'est pas indiqué, le CCSID spécifié dans les définitions de la configuration de texte est utilisé. Les modifications ultérieures apportées à ces définitions sont ignorées ; la valeur utilisée est celle qui existait au moment de l'activation de la colonne, et non celle indiquée lors de l'indexation des documents texte.

Pour plus d'informations sur les autres pages de code pouvant être prises en charge, reportez-vous à la section «CCSID» à la page 35.

### **LANGUAGE langue**

Langue dans laquelle le texte est rédigé. Ceci détermine le dictionnaire à utiliser lors de l'indexation de documents texte et des recherches dans ces derniers. Le «Chapitre 15. Traitement linguistique des index linguistiques et exacts» à la page 255, décrit le mode d'utilisation des dictionnaires.

La langue indiquée par ce mot clé s'applique à la colonne entière. Vous pouvez écraser cette valeur pour les documents texte insérés individuellement au moyen de la fonction INIT\_TEXT\_HANDLE dans une instruction INSERT.

Si ce mot clé n'est pas indiqué, la langue spécifiée dans les définitions de la configuration de texte est utilisée. Les modifications ultérieures apportées à ces définitions sont ignorées ; la valeur utilisée est celle qui existait au moment de l'activation de la colonne, et non celle indiquée lors de l'indexation des documents texte.

Les langues prises en charge sont indiquées dans le tableau 5 à la page 54.

## Commande ENABLE TEXT COLUMN

### FORMAT format

Type du document texte stocké, tel que WordPerfect ou ASCII. Cette information est nécessaire à DB2 Extension Texte lors de l'indexation des documents. Vous trouverez la liste des formats de documents pris en charge à la section «Documents pris en charge» à la page 31.

Les formats pris en charge pour les documents structurés sont les suivants :

- ASCII\_SECTIONS

Les documents dont le format est ASCII\_SECTIONS ne peuvent pas contenir de sections imbriquées. (Pour plus de détails sur ces dernières, reportez-vous à la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48.) Une marque de début de section s'achève par la marque de début suivante.

- HTML

Un fichier exemple de modèles de documents est fourni pour les documents HTML. Il contient un sous-ensemble des définitions HTML standard, que vous pouvez modifier si nécessaire. Les documents HTML ne peuvent contenir de sections imbriquées.

- XML

Le traitement des documents XML inclut l'évaluation des définitions de type de document (DTD - Document Type Definition). Le modèle affecté au document est vérifié par rapport à la DTD. Si les marques définies dans le fichier de modèles de documents ne sont pas définies dans la DTD, le document n'est pas indexé. Si aucun modèle n'a été défini pour une DTD reconnue, le document ne sera pas indexé. Les documents XHTML peuvent contenir des sections imbriquées.

Pour ces différents formats, vous devez spécifier les informations relatives à la structure dans un fichier de modèles de documents. Reportez-vous à la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48. Si le format TDS et INDEXPROPERTY SECTION\_ENABLED sont indiqués, le système considère que le format de document est ASCII\_SECTIONS.

Les marques non définies dans le fichier de modèles sont indexées de la manière habituelle, en fonction du type d'index.

Le format indiqué par ce mot clé s'applique à la colonne entière. Vous pouvez écraser cette valeur pour les documents texte insérés individuellement au moyen de la fonction INIT\_TEXT\_HANDLE dans une instruction INSERT.

Si ce mot clé n'est pas indiqué, le format spécifié dans les définitions de la configuration de texte est utilisé. Les modifications ultérieures apportées à ces définitions sont ignorées ; la valeur utilisée est celle qui existait au moment de l'activation de la colonne, et non celle indiquée lors de l'indexation des documents texte.

### **INDEXTYPE**

Type d'index à créer. Pour plus d'informations, reportez-vous à la section «Types de recherche» à la page 40.

### **PRECISE**

Les termes sont indexés et recherchés sous la forme exacte dans laquelle ils apparaissent dans les documents texte.

### **LINGUISTIC**

Les termes font l'objet d'un traitement linguistique avant d'être indexés, de même que les termes de recherche avant le début de la recherche.

### **NGRAM**

L'indexation des termes se fait par analyse de séquences de caractères, au lieu de se fonder sur un dictionnaire. Il est obligatoire d'utiliser ce type d'index si les documents à indexer contiennent des caractères DBCS. Toutefois, un index Ngram peut être utilisé pour des documents SBCS.

Si vous omettez le mot clé **INDEXTYPE**, la valeur spécifiée dans les définitions de la configuration de texte est utilisée.

Les documents au format XML ne sont pas pris en charge pour les index Ngram.

### **INDEXOPTION**

Options utilisables lors de la création d'un index.

### **CASE\_ENABLED**

Cette option est valable **uniquement pour les index Ngram**. Normalement, ces index n'acceptent pas la distinction entre majuscules et minuscules. En précisant **CASE\_ENABLED**, vos documents seront indexés de manière à permettre des recherches intégrant la distinction entre majuscules et minuscules. Pour plus d'informations sur cette option, reportez-vous à la section «Recherche ou recherche de correspondance partielle dans des documents DBCS» à la page 43.

## Commande **ENABLE TEXT COLUMN**

### **NORMALIZED**

Cette option est valable **uniquement pour les index exacts**.

Un index exact normalisé se distingue d'un index exact par les caractéristiques suivantes :

- Il ne tient pas compte de la distinction majuscules/minuscules ; tous les mots, à l'exception de ceux dont tous les caractères sont en majuscules, sont transformés en minuscules.
- Les mots dont tous les caractères sont en majuscules ne sont pas éliminés lors du filtrage des "mots vides" ; l'abréviation USA, par exemple, est indexée.
- La recherche de termes en français peut être étendue par lemmatisation grâce à un algorithme heuristique : si vous recherchez le mot maison, la forme maisons sera également recherchée.

### **INDEXPROPERTY SECTIONS\_ENABLED DOCUMENTMODEL(S)**

#### **nom-modèle**

Propriétés d'un type d'index sélectionné.

SECTIONS\_ENABLED précise que le type d'index sélectionné peut contenir des informations sur la structure du document.

DOCUMENTMODEL/DOCUMENTMODELS *nom-modèle* spécifie le ou les modèles devant être associés par défaut aux documents à indexer. Si la propriété d'index SECTIONS\_ENABLED est utilisée, un nom de modèle doit être indiqué. Si une liste de modèles est précisée, le modèle par défaut pour l'index est le premier de la liste. Le modèle par défaut est utilisé lors de l'indexation si le document ne comporte pas de référence à un modèle, ou qu'aucun modèle n'est indiqué lors de la recherche.

Les caractères admis pour un nom de modèle sont les suivants : a à z, A à Z et 0 à 9.

Le nom de modèle spécifié doit correspondre à une définition de modèle figurant dans le fichier `desmodel.ini`.

Pour modifier le ou les modèles associés à un index, procédez comme suit :

1. Utilisez **DISABLE TEXT COLUMN** pour désactiver l'index.
2. Utilisez **ENABLE TEXT COLUMN** pour ré-indexer les documents, en indiquant des noms de modèles de document différents.

### **UPDATEFREQ fréq-mise-à-jour**

Fréquence de mise à jour de l'index exprimée en fonction de la date et du nombre minimal de documents texte en attente dans la table de

journalisation. Si cette dernière ne comporte pas suffisamment de documents texte au jour et à l'heure indiqués, l'index n'est pas mis à jour.

La syntaxe correspondante est décrite à la section «Mise à jour d'un index» à la page 46.

Si vous omettez `UPDATEFREQ`, le système utilise la fréquence par défaut spécifiée dans les définitions de la configuration de texte.

### **Conseil**

Si vous disposez d'un grand nombre de tables, nous vous conseillons de ne pas utiliser les valeurs par défaut. En définissant une fréquence de mise à jour propre à chaque table, vous pouvez éviter l'indexation simultanée de toutes les tables et par conséquent l'utilisation intensive et intempestive des ressources du système.

### **NONE**

Aucune autre mise à jour d'index n'est effectuée. Ceci concerne une colonne de texte qui ne fera plus l'objet de modifications.

Ces paramètres relatifs à la fréquence de mise à jour sont ignorés s'ils ont déjà été définis pour la table entière par `ENABLE TEXT TABLE`.

### **DIRECTORY répertoire**

Chemin d'accès au répertoire dans lequel l'index de texte doit être stocké. La chaîne `""txins $nnn$ ""` est ajoutée à ce chemin, où  $nnn$  représente le numéro de noeud.

Il s'agit d'un répertoire déjà présent sur le système sur lequel le serveur DB2 Extension Texte s'exécute. Si le répertoire n'existe pas, il est créé. S'il existe, vérifiez, pour les plateformes UNIX, que ce répertoire dispose des droits en lecture et écriture pour le propriétaire de l'instance ; vérifiez également l'"ID utilisateur isolé", s'il en a été indiqué un différent de celui défini lors de la création de l'instance DB2.

Si vous omettez le mot clé `DIRECTORY`, la valeur spécifiée dans les définitions de la configuration de texte est utilisée.

Ce paramètre est ignoré s'il a déjà été défini pour la table entière par `ENABLE TEXT TABLE`.

## Commande **ENABLE TEXT COLUMN**

Si vous utilisez des bases de données partitionnées, vous pouvez spécifier un chemin d'accès aux index qui soit commun à tous les noeuds ou à une série de noeuds, ou des chemins d'accès différents pour chaque noeud.

### **ON NODE** *noeud* [**TO** *noeud*]

Numéro du noeud ou intervalle de noeuds auquel un chemin de répertoire est affecté.

### **UPDATEINDEX**

Mot clé déterminant si les documents texte associés à cette colonne de descripteurs sont indexés immédiatement après la fin de l'exécution de cette commande, sans attendre la prochaine indexation périodique définie par **UPDATEFREQ**.

### **UPDATE**

L'indexation des documents texte a lieu immédiatement après l'achèvement de la commande.

### **NOUPDATE**

L'indexation a lieu au moment défini par les paramètres de fréquence de mise à jour indiqués dans cette commande par **UPDATEFREQ**, ou dans les définitions de la configuration de texte.

Si vous omettez ce mot clé, le système utilise la valeur spécifiée dans les définitions de la configuration de texte.

### **TABLESPACE** *nom-espace-table*

Nom de l'espace table créé en interne sur la colonne de descripteurs. L'espace table doit avoir été créé préalablement.

### **COMMITCOUNT** *nb-opérations*

Valeur comprise entre 500 et 1 000 000, indiquant le nombre d'insertions ou de mises à jour effectuées avant qu'une instruction intermédiaire de validation DB2 UDB soit émise. L'utilisation de ce mot clé permet d'éviter que vous ne manquiez d'espace de journalisation lorsque vous activez des tables ou des colonnes volumineuses, ou un nombre élevé de fichiers externes.

## **Commentaires**

Cette commande ajoute une colonne de descripteurs à la table DB2 indiquée. Chaque colonne de descripteurs est associée à une colonne de texte et est utilisée par les fonctions de DB2 Extension Texte.

Si cette table n'a pas déjà été activée pour créer un index commun, un index est créé et associé à cette colonne de texte.

## Commande ENABLE TEXT COLUMN

Une table de journalisation est également créée dans la base de données. Elle est utilisée pour enregistrer les modifications effectuées dans une colonne de texte, à savoir les insertions, les mises à jour et les suppressions. Les déclencheurs d'insertion, de mise à jour et de suppression définis pour la colonne de texte permettent de la mettre à jour automatiquement.

Si la colonne de texte que vous activez se trouve dans une table appartenant à un groupe de noeuds, le répertoire d'index spécifié doit être disponible sur tous les noeuds physiques. Si vous utilisez le répertoire par défaut spécifié dans la configuration de texte, assurez-vous que le chemin d'accès est accessible à partir de tous les noeuds du groupe. Si ce n'est pas le cas, vous pouvez définir un chemin spécifique pour chaque noeud dans la commande ENABLE TEXT COLUMN.

Si vous modifiez la configuration d'un groupe de noeuds qui contient une table activée pour DB2 Extension Texte, vous devrez ré-indexer la table.

### Conseil

A ce stade, si vous manquez d'espace de journalisation, reportez-vous à la section «Activation d'une colonne de texte dans une table volumineuse» à la page 70, pour prendre connaissance des solutions possibles.

# Commande ENABLE TEXT FILES

## ENABLE TEXT FILES

Cette commande permet à DB2 Extension Texte d'effectuer des recherches dans des fichiers de texte extérieurs à une base de données DB2 UDB.

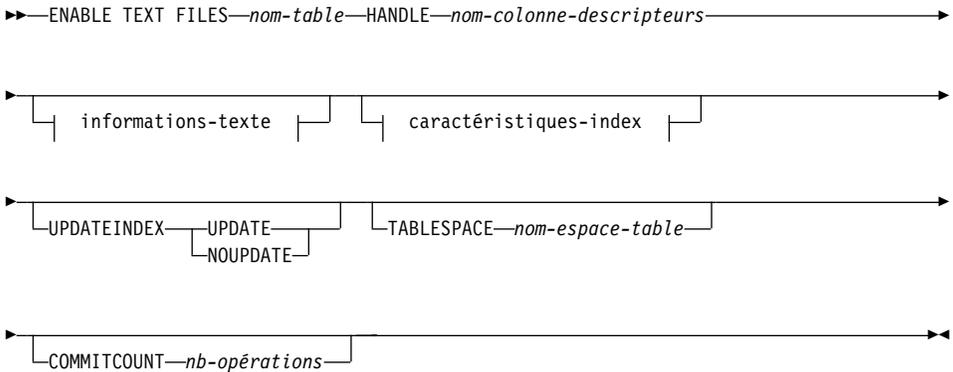
**Conseil**  
Cette commande ne peut pas être utilisée si les colonnes de texte de la table partagent un index commun. Pour plus de détails, reportez-vous à la section «Activation d'une table de texte (facultatif)» à la page 63.

### Droits

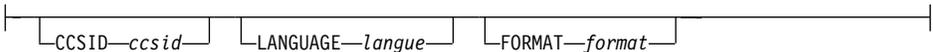
Vous devez disposer d'au moins un des droits suivants sur la table :

- ALTER
- SELECT
- UPDATE

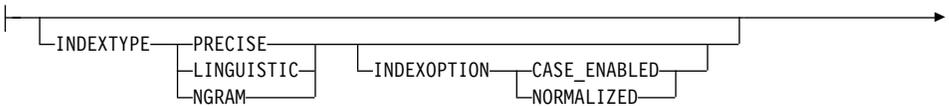
### Syntaxe de la commande



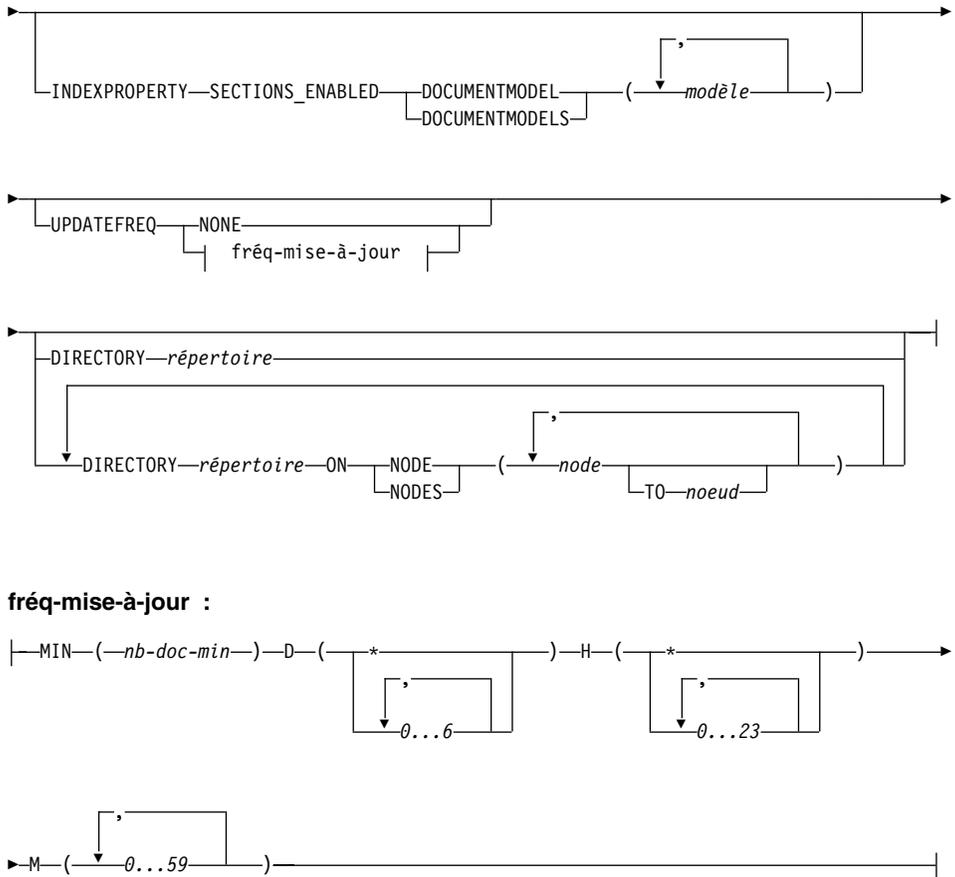
#### informations-texte :



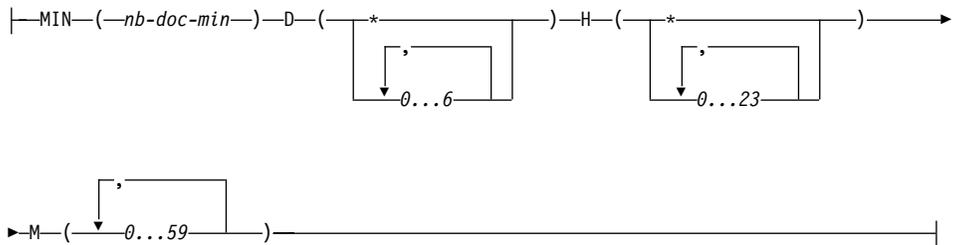
#### caractéristiques-index :



## Commande ENABLE TEXT FILES



### fréq-mise-à-jour :



## Paramètres de la commande

### nom-table

Nom de la table de texte de la base de données connectée, qui doit être associée aux fichiers de texte externes sur lesquels porte l'indexation. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

Pour obtenir une description des autres paramètres de la commande, reportez-vous à la section «ENABLE TEXT COLUMN» à la page 145.

## Commentaires

Cette commande ajoute une colonne de descripteurs à la table DB2 indiquée. Chaque colonne de descripteurs est associée à une colonne de texte et est utilisée par les fonctions de DB2 Extension Texte. Un index associé à ces fichiers est créé.

## Commande ENABLE TEXT FILES

Après avoir activé les fichiers de texte, initialisez les descripteurs dans la colonne correspondante, à l'aide de la fonction INIT\_TEXT\_HANDLE. Puis, actualisez l'index par la commande UPDATE INDEX.

Vous ne pouvez pas réutiliser un nom de colonne de descripteurs déjà utilisé dans une commande ENABLE TEXT FILES pour identifier une colonne de texte.

Une table de journalisation est également créée pour enregistrer les modifications apportées aux fichiers. Vous devez cependant activer manuellement les déclencheurs, étant donné que DB2 UDB ne supervise pas ces fichiers et ne gère donc pas de telles modifications. Pour plus de détails à ce sujet, reportez-vous à la section «Mise à jour d'un index pour y intégrer les références de fichiers externes» à la page 102.

A ce stade, si vous manquez d'espace de journalisation, reportez-vous à la section «Activation d'une colonne de texte dans une table volumineuse» à la page 70, pour prendre connaissance des solutions possibles.

## ENABLE TEXT TABLE

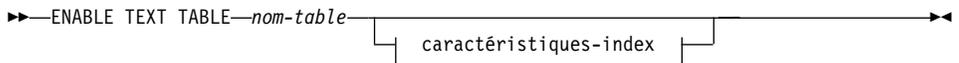
Cette commande crée un index commun à utiliser par les colonnes de texte de la table qui seront activées ultérieurement. Cette table devient alors une table à index commun. Une table dont les colonnes de texte activées ultérieurement créent leurs propres index est appelée table à index multiples.

### Droits

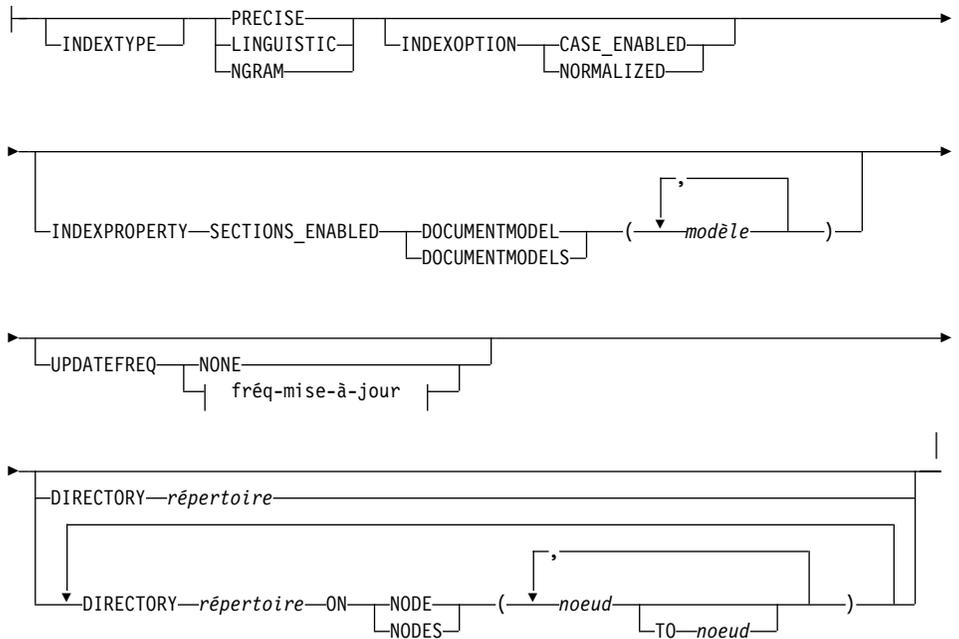
Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

### Syntaxe de la commande

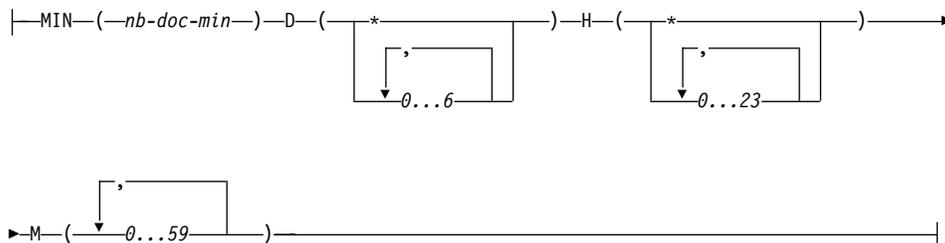


#### caractéristiques-index :



#### fréq-mise-à-jour :

## Commande ENABLE TEXT TABLE



### Paramètres de la commande

#### nom-table

Nom de la table de texte à activer dans la base de données connectée. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### INDEXTYPE

Type d'index à créer. Pour plus d'informations, reportez-vous à la section «Types de recherche» à la page 40.

#### PRECISE

Les termes sont indexés et recherchés sous la forme exacte dans laquelle ils apparaissent dans les documents texte.

#### LINGUISTIC

Les termes font l'objet d'un traitement linguistique avant d'être indexés, de même que les termes de recherche avant le début de la recherche.

#### NGRAM

L'indexation des termes se fait par analyse de séquences de caractères, au lieu de se fonder sur un dictionnaire. Il est obligatoire d'utiliser ce type de dictionnaire si les documents à indexer contiennent des caractères DBCS. Toutefois, un index Ngram peut être utilisé pour des documents SBCS.

Si vous omettez le mot clé `INDEXTYPE`, le système utilise la valeur spécifiée dans les définitions de la configuration de texte.

#### INDEXOPTION

Options utilisables lors de la création d'un index.

#### CASE\_ENABLED

Cette option est valable **uniquement pour les index Ngram**. Normalement, ces index n'acceptent pas la distinction entre majuscules et minuscules. En précisant `CASE_ENABLED`, vos documents seront indexés de manière à permettre des recherches intégrant la distinction entre majuscules et minuscules. Pour plus d'informations sur cette option,

reportez-vous à la section «Recherche ou recherche de correspondance partielle dans des documents DBCS» à la page 43.

### INDEXPROPERTY SECTIONS\_ENABLED DOCUMENTMODEL(S)

#### nom-modèle

Propriétés d'un type d'index sélectionné.

SECTIONS\_ENABLED précise que le type d'index sélectionné peut contenir des informations sur la structure du document.

DOCUMENTMODEL/DOCUMENTMODELS *nom-modèle* spécifie le ou les modèles devant être associés par défaut aux documents à indexer. Si la propriété d'index SECTIONS\_ENABLED est utilisée, un nom de modèle doit être indiqué. Si une liste de modèles est précisée, le modèle par défaut pour l'index est le premier de la liste. Le modèle par défaut est utilisé lors de l'indexation si le document ne comporte pas de référence à un modèle, ou qu'aucun modèle n'est indiqué lors de la recherche.

Les caractères admis pour un nom de modèle sont les suivants : a à z, A à Z et 0 à 9.

Le nom de modèle spécifié doit correspondre à une définition de modèle figurant dans le fichier `desmodel.ini`.

Pour modifier le ou les modèles associés à un index, procédez comme suit :

1. Utilisez DISABLE TEXT TABLE pour désactiver l'index.
2. Utilisez ENABLE TEXT TABLE pour ré-indexer les documents, en indiquant des noms de modèles de document différents.

### UPDATEFREQ fréq-mise-à-jour

*Fréquence* à laquelle l'index est mis à jour ainsi que le *nombre minimal de documents texte à mettre en file d'attente* dans la table de journalisation. Si cette dernière ne comporte pas suffisamment de documents texte au jour et à l'heure indiqués, l'index n'est pas mis à jour.

La syntaxe correspondante est décrite à la section «Mise à jour d'un index» à la page 46.

Si vous omettez UPDATEFREQ, le système utilise la fréquence par défaut spécifiée dans les définitions de la configuration de texte.

#### NONE

Aucune autre mise à jour d'index n'est effectuée. Ceci concerne une colonne de texte qui ne fera plus l'objet de modifications.

## Commande ENABLE TEXT TABLE

### Conseil

Si vous disposez d'un grand nombre de tables, nous vous conseillons de ne pas utiliser les valeurs par défaut. En définissant une fréquence de mise à jour propre à chaque table, vous pouvez éviter l'indexation simultanée de toutes les tables et par conséquent l'utilisation intensive et intempestive des ressources du système.

### DIRECTORY répertoire

Chemin d'accès au répertoire dans lequel l'index de texte doit être stocké. La chaîne `""txinsnnn"` est ajoutée à ce chemin, où *nnn* représente le numéro de noeud.

Il s'agit d'un répertoire déjà présent sur le système sur lequel le serveur DB2 Extension Texte s'exécute. Si le répertoire n'existe pas, il est créé. S'il existe, vérifiez, pour les plateformes UNIX, que ce répertoire dispose des droits en lecture et écriture pour le propriétaire de l'instance ; vérifiez également l'"ID utilisateur isolé", s'il en a été indiqué un différent de celui défini lors de la création de l'instance DB2.

Si vous omettez le mot clé DIRECTORY, la valeur spécifiée dans les définitions de la configuration de texte est utilisée.

Si vous utilisez des bases de données partitionnées, vous pouvez spécifier un chemin d'accès aux index commun à tous les noeuds ou à une série de noeuds, ou bien un chemin de répertoire d'index différent pour chaque noeud.

### ON NODE noeud [TO noeud]

Numéro du noeud ou intervalle de noeuds auquel un chemin de répertoire est affecté.

### Commentaires

Un nouvel index de texte est créé et associé à toutes les colonnes de texte de cette table. Ceci permet d'obtenir un index commun à toutes les colonnes de texte d'une table, plutôt qu'un index distinct pour chacune de ces colonnes.

Après avoir activé une table, vous devez exécuter la commande ENABLE TEXT COLUMN pour chacune des colonnes de texte dans lesquelles vous souhaitez effectuer une recherche.

Une table de journalisation est créée dans la base de données. Elle est utilisée pour enregistrer les modifications (insertions, mises à jour et suppressions) effectuées dans les colonnes de texte qui seront activées ultérieurement.

## Commande **ENABLE TEXT TABLE**

Lorsqu'une colonne de texte est activée, des déclencheurs sont créés afin de contrôler les modifications apportées au texte et de conserver automatiquement un enregistrement des documents qui doivent être indexés dans la table de journalisation.

DB2 Extension Texte indexe périodiquement les documents texte répertoriés dans cette table en fonction des valeurs du mot clé UPDATEFREQ.

## Commande GET ENVIRONMENT

---

### GET ENVIRONMENT

Cette commande affiche les paramètres des variables d'environnement.

#### Droits

Aucun.

#### Syntaxe de la commande

▶—GET ENVIRONMENT—▶

#### Paramètres de la commande

Aucun.

#### Commentaires

Les variables d'environnement affichées sont les suivantes :

##### **DB2INSTANCE**

Nom de l'instance DB2 UDB.

##### **DB2DBDFT**

Nom de la base de données par défaut.

##### **DB2TX\_INSTOWNER**

Nom de l'instance DB2 Extension Texte. Obligatoire uniquement en cas de connexion à un serveur UNIX.

##### **DB2TX\_INSTOWNERHOMEDIR**

Répertoire principal du propriétaire de l'instance Obligatoire uniquement en cas de connexion à un serveur UNIX.

## GET INDEX SETTINGS

Cette commande affiche les paramètres d'un index, indiquant les informations suivantes :

- Type d'index
- Option d'index (facultatif)
- Option de mise à jour de l'index
- Répertoire de l'index
- Fréquence de mise à jour
- Modèle par défaut

### Droits

Aucun.

### Syntaxe de la commande

```

▶▶ GET INDEX SETTINGS nom-table [HANDLE nom-colonne-descripteurs]
  
```

### Paramètres de la commande

#### **nom-table**

Nom de la table de texte de la base de données connectée, dont les paramètres d'index doivent être affichés. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs dont les paramètres d'index doivent être affichés.

### Commentaires

Si la table est activée en tant que table à index multiples, cette commande affiche les paramètres d'index de toutes les colonnes de texte activées dans la table. Si vous indiquez le paramètre *nom-colonne-descripteurs*, cette commande affiche les paramètres d'index de la colonne spécifiée.

S'il s'agit d'une table à index commun, les paramètres de l'index commun sont affichés. Si un *nom-colonne-descripteur* est fourni, il n'est pas pris en compte.

Si la table ou la colonne est activée avec la propriété d'index `SECTIONS_ENABLED`, la commande `GET INDEX SETTINGS` affiche le modèle par défaut pour l'index. Il s'agit du nom de modèle spécifié lors de l'activation ou du premier nom figurant dans une liste de noms de modèles.

## Commande GET INDEX SETTINGS

Voici un exemple :

Current index settings:

Index type	(INDEXTYPE) = LINGUISTIC
Default model	(DOCUMENTMODEL) = mymodel
Update index option	(UPDATEINDEX) = UPDATE
Update frequency	(UPDATEFREQ) = NONE
Node 1	
Index directory	(DIRECTORY) = /home/user1/db2tx/index

## GET INDEX STATUS

Cette commande affiche les informations suivantes sur l'état de l'index, pour une colonne de descripteurs ou une table donnée :

- Fonction de recherche disponible ou non
- Fonction de mise à jour de l'index disponible ou non
- Fonction de réorganisation disponible ou non
- Nombre de documents planifiés
- Nombre de documents indexés
- Nombre de documents indexés dans l'index primaire
- Nombre de documents indexés dans l'index secondaire
- Événements d'erreur

### Droits

Aucun.

### Syntaxe de la commande

```

▶▶—GET INDEX STATUS—nom-table—┬──┬──▶
                                └─HANDLE—nom-colonne-descripteurs—┘
    
```

### Paramètres de la commande

#### **nom-table**

Nom de la table de texte de la base de données connectée, qui contient les colonnes de texte dont l'état doit être affiché. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

#### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs dont l'état doit être affiché.

### Commentaires

Dans une table à index multiples, vous devez préciser le nom de la colonne de descripteurs.

## Commande GET STATUS

---

### GET STATUS

Cette commande affiche des informations relatives à l'état Activé (ou non) de bases de données, de tables ou de colonnes de texte.

#### Droits

Aucun.

#### Syntaxe de la commande

► GET STATUS ◀

#### Paramètres de la commande

Aucun.

#### Commentaires

Cette commande affiche l'état de la base de données (activée ou non), le nom des tables de texte activées dans la base de données, des colonnes de texte activées et des colonnes de descripteurs qui leur sont associées, ainsi que celui des colonnes de descripteurs associées aux fichiers externes.

---

### GET TEXT CONFIGURATION

Cette commande affiche les paramètres par défaut de la configuration de texte définie pour la base de données connectée.

Pour modifier ces paramètres par défaut, utilisez «CHANGE TEXT CONFIGURATION» à la page 133.

#### Droits

Vous devez disposer d'au moins un des droits suivants sur la table :

SELECT

#### Syntaxe de la commande

```
▶▶ GET TEXT — CONFIGURATION  
                |  
                — CFG
```

#### Paramètres de la commande

Aucun.

#### Commentaires

Vous trouverez un exemple des informations données en sortie par la commande GET TEXT CONFIGURATION dans la section «Affichage des définitions de la configuration de texte» à la page 107.

## Commande GET TEXT INFO

---

### GET TEXT INFO

Cette commande affiche les paramètres d'information de texte pour les colonnes de texte :

- CCSID
- Langue
- Format

#### Droits

Aucun.

#### Syntaxe de la commande

```
▶ GET TEXT INFO nom-table [HANDLE nom-colonne-descripteurs] ▶
```

#### Paramètres de la commande

##### **nom-table**

Nom de la table de texte de la base de données connectée, qui contient les colonnes de texte dont les informations de texte doivent être affichées. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

##### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs dont les paramètres d'information de texte doivent être affichés.

#### Commentaires

Si vous indiquez un nom de colonne de descripteurs, seules les informations de texte la concernant sont affichées.

Si vous n'en indiquez pas, le système affiche les informations de texte relatives à chacune des colonnes activées de cette table.

### QUIT

Cette commande interrompt l'interpréteur de commandes de DB2 Extension Texte et rend la main au système d'exploitation.

#### Droits

Aucun.

#### Syntaxe de la commande

»—QUIT—«

#### Paramètres de la commande

Aucun.

#### Commentaires

La connexion à la base de données est arrêtée.

---

### REORGANIZE INDEX

Si une colonne de texte est mise à jour fréquemment, une recherche dans l'index peut devenir inefficace. Pour rendre son efficacité à la recherche, l'index doit être *réorganisé*. DB2 Extension Texte est capable d'estimer le moment où un index a besoin d'être réorganisé et effectue cette opération automatiquement, en arrière-plan. Cependant, dans certains cas, il peut être nécessaire de réorganiser un index manuellement par la commande REORGANIZE INDEX. Vous pouvez utiliser la commande GET INDEX STATUS pour savoir si un index a besoin d'une réorganisation.

#### Droits

Aucun.

#### Syntaxe de la commande

►—REORGANIZE INDEX—*nom-table*—HANDLE—*nom-colonne-descripteurs*—◄

#### Paramètres de la commande

##### **nom-table**

Nom de la table de texte de la base de données connectée, dont l'index doit être réorganisé. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

##### **HANDLE nom-colonne-descripteurs**

Nom de la colonne de descripteurs dont l'index doit être réorganisé.

#### Commentaires

Dans une table à index multiples, vous devez préciser le nom de la colonne de descripteurs.

Des recherches peuvent être effectuées pendant l'exécution de REORGANIZE INDEX, mais non des mises à jour d'index.



## Commande UPDATE INDEX

### UPDATE INDEX

Cette commande démarre immédiatement l'indexation. Elle permet de mettre à jour l'index en fonction du contenu existant de la ou des colonnes de texte auxquelles il est associé.

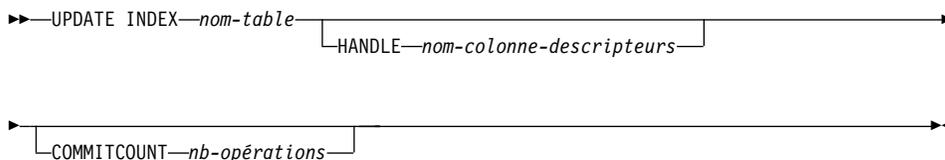
Pour mettre à jour des fichiers externes associés à l'index, vous devez forcer l'insertion d'une entrée de type "modification" dans la table de journalisation par une instruction UPDATE sur la colonne de descripteurs concernée. Vous trouverez un exemple de cette procédure à la section «Mise à jour d'un index pour y intégrer les références de fichiers externes» à la page 102.

#### Droits

Vous devez disposer d'au moins un des droits suivants sur la table :

ALTER  
SELECT  
UPDATE

#### Syntaxe de la commande



#### Paramètres de la commande

##### nom-table

Nom de la table de texte de la base de données connectée, qui contient la colonne de texte dont l'index doit être mis à jour. Il peut également s'agir du nom d'une table à index commun. Il doit comporter un nom de schéma explicite (qualificatif), sauf si ce dernier est identique à votre ID utilisateur.

##### HANDLE nom-colonne-descripteurs

S'il s'agit d'une table à index commun, *nom-colonne-descripteurs* n'est pas nécessaire et est ignoré. L'index à mettre à jour est associé à la table entière, et non à une colonne de texte spécifique.

S'il s'agit d'une table à index multiples, *nom-colonne-descripteurs* est le nom de la colonne de descripteurs dont l'index doit être mis à jour.

##### COMMITCOUNT nb-opérations

Valeur comprise entre 500 et 1 000 000, indiquant le nombre d'insertions ou de mises à jour effectuées avant qu'une instruction intermédiaire de validation DB2 UDB soit émise. L'utilisation de ce

## Commande UPDATE INDEX

mot clé permet d'éviter que vous ne manquiez d'espace de journalisation lorsque vous activez des tables ou des colonnes volumineuses, ou un nombre élevé de fichiers externes.

## Commande UPDATE INDEX

---

## Chapitre 10. Commandes d'administration du serveur

Le présent chapitre décrit la syntaxe des commandes d'administration du serveur. L'administratiop du serveur est constituée de tâches permettant de démarrer, d'arrêter ou de contrôler l'état du serveur DB2 Extension Texte, ainsi que de créer un modèle de bases de données et d'échantillonner une table. La section «Configuration et gestion d'un serveur DB2 Extension Texte» à la page 21, explique comment utiliser ces commandes.

Commande	Objet	Page
TXICRT	Crée une instance de DB2 Extension Texte	176
TXIDROP	Supprime une instance de DB2 Extension Texte	178
TXILIST (sous UNIX uniquement)	Dresse la liste des instances de DB2 Extension Texte (sous UNIX uniquement)	179
Commande TXIPCLEAN (sous UNIX uniquement)	Nettoie les ressources du système après une défaillance système (sous UNIX seulement)	180
TXNADD	Ajoute des serveurs de DB2 Extension Texte	181
TXNCHECK	Vérifie l'intégrité d'une base de données DB2 Extension Texte	182
TXNDROP	Supprime des serveurs de DB2 Extension Texte	183
TXSAMPLE	Crée et active un modèle de table	184
TXSTART	Démarre les fonctions de DB2 Extension Texte	185
TXSTATUS	Affiche l'état de la fonction de recherche	186
TXSTOP	Arrête les fonctions de DB2 Extension Texte	187
TXTHESC	Compile un fichier de définition de thésaurus	188
TXTHESN	Compile un fichier de définition de thésaurus Ngram	190
TXTRACE	Génère des informations de trace	192
TXVERIFY (commande)	Crée et active un modèle de base de données	197
TXWIZARD (commande)	Exemple de programme pour l'indexation et la recherche (AIX, Windows NT et Windows 2000 seulement)	198

# Commande TXICRT

## TXICRT

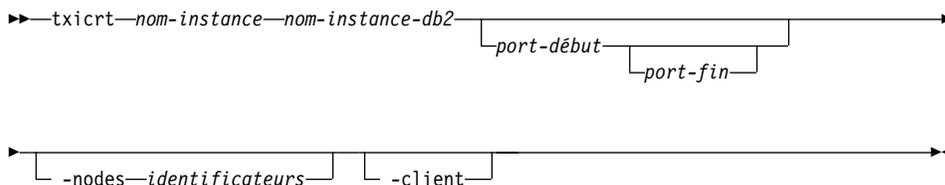
Cette commande permet de créer une instance de DB2 Extension Texte.

### Droits

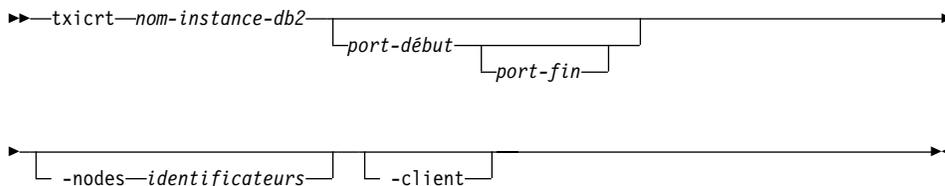
Root (sous UNIX uniquement)

### Syntaxe de la commande

Pour UNIX :



Pour Windows NT et Windows 2000 :



### Paramètres de la commande

*nom-instance*

(UNIX uniquement) ID utilisateur du propriétaire de l'instance.

*nom-instance-db2*

Nom de l'instance DB2 à associer à l'instance en cours de création.

*port-début/port-fin*

Intervalle de ports TCP/IP à utiliser si vous gérez plusieurs noeuds physiques. Non requis si vous n'utilisez que des noeuds logiques.

Cet intervalle doit être égal au nombre de noeuds à gérer. Dans le cas d'un noeud unique, spécifiez seulement le paramètre port-début. Pour prendre en charge un sous-ensemble de noeuds, utilisez le paramètre port-fin.

N'indiquez pas d'intervalle de ports si l'instance est créée sur une base de données monopartition.

Utilisez uniquement des numéros de port non répertoriés dans le fichier TCP/IP etc/services. Ne spécifiez pas de numéro de port pour les installations relatives aux seuls clients.

### **-nodes** *identificateurs*

Liste d'identificateurs de noeud. Exemple : -nodes 0 1 2.

Utilisez ce paramètre si votre table appartient à des noeuds spécifiques d'un groupe de noeuds.

Si ce paramètre est omis, tous les noeuds sont utilisés, c'est-à-dire tous ceux définis dans db2nodes.cfg.

Ne spécifiez pas d'identificateur de noeud pour les installations relatives aux seuls clients.

### **-client**

Spécifiez ce paramètre si vous disposez d'une installation relative aux seuls clients DB2 Extension Texte. S'il est omis, TXICRT configure une instance client-serveur de DB2 Extension Texte.

## **Commentaires**

L'activation de tables ou de colonnes de texte n'est possible que sur les noeuds où vous avez créé une instance de DB2 Extension Texte.

Une instance DB2 doit déjà être créée pour l'ID utilisateur.

Dans les installations UNIX, cette commande propre aux instances se trouve dans le répertoire /usr/lpp/db2tx\_07\_01/instance.

Sous Windows NT ou Windows 2000, une instance par défaut est créée lors de l'installation dans un environnement de bases de données non partitionnées.

## Commande TXIDROP

---

### TXIDROP

Cette commande permet de supprimer une instance de DB2 Extension Texte avec tous ses index.

#### Droits

Root (sous UNIX uniquement)

#### Syntaxe de la commande

Pour UNIX :

► `txidrop nom-instance nom-instance-db2` ◀

Pour Windows NT et Windows 2000 :

► `txidrop nom-instance` ◀

#### Paramètres de la commande

##### **nom-instance**

Nom de l'instance à supprimer.

##### **nom-instance-db2**

Nom de l'instance DB2 à associer à l'instance à supprimer.

#### Commentaires

Avant de supprimer une instance, désactivez toutes les bases de données activées dans le cadre de cette instance.

Dans les installations UNIX, cette commande propre aux instances se trouve dans le répertoire `/usr/lpp/db2tx_07_01/instance`.

### TXILIST (sous UNIX uniquement)

Cette commande dresse la liste des instances de la version en cours de DB2 Extension Texte.

#### Droits

SYSADM, DBADM

#### Syntaxe de la commande

▶▶—txilist—————▶▶

#### Paramètres de la commande

Aucun.

#### Commentaires

Si plusieurs versions de DB2 Extension Texte ont été installées, cette commande dresse la liste des instances de la version actuellement en cours d'utilisation, c'est-à-dire celle indiquée dans le répertoire en cours.

## Commande TXIPCLEAN (sous UNIX uniquement)

---

### Commande TXIPCLEAN (sous UNIX uniquement)

Cette commande nettoie les ressources système pour l'utilisateur connecté, autrement dit, il arrête le processus `desssdem`, et régénère la mémoire partagée, les files de messages relatives aux segments et les sémaphores.

#### Droits

Membre du groupe SMADMIN

#### Syntaxe de la commande

►—txipclean—◄

#### Paramètres de la commande

Aucun.

#### Commentaires

N'utilisez cette commande qu'avec la plus grande prudence car elle nettoie non seulement les ressources DB2 Extension Texte mais également toutes les ressources partagées de l'utilisateur connecté. Connectez-vous toujours en tant que propriétaire de l'instance afin de limiter le nettoyage aux ressources de l'instance.

Dans certains cas, vous ne pouvez pas démarrer ou arrêter correctement DB2 Extension Texte. Essayez alors d'arrêter d'abord DB2 et DB2 Extension Texte, puis d'exécuter TXIPCLEAN.

## TXNADD

Cette commande permet de créer un serveur DB2 Extension Texte supplémentaire sur le noeud spécifié. Utilisez ensuite TXSTART pour démarrer le serveur.

### Droits

(sous UNIX uniquement) SYSADM, DBADM

### Syntaxe de la commande

►► `txnadd` *numéro-noeud* `port` ◀◀

### Paramètres de la commande

#### numéro-noeud

Numéro du noeud à ajouter. Indiquez-le exactement tel qu'il apparaît dans le fichier `db2nodes.cfg`.

**port** Numéro du port à utiliser pour le nouveau noeud. Vous devez le spécifier si vous ajoutez un noeud physique, ou si votre configuration de noeuds contient des noeuds physiques.

### Commentaires

Utilisez TXNCHECK pour déterminer les index à recréer afin d'assurer la synchronisation avec votre base de données.

#### Conseil

Si vous ajoutez un noeud physique dans une configuration comportant des noeuds logiques, vous devez d'abord supprimer l'instance, puis la recréer en spécifiant des numéros de port pour tous les noeuds.

## Commande TXNCHECK

---

### TXNCHECK

Cette commande permet de vérifier l'intégrité de DB2 Extension Texte pour une base de données spécifique.

#### Droits

(sous UNIX uniquement) SYSADM, DBADM

#### Syntaxe de la commande

► txncheck nom-bdd ◄

#### Paramètres de la commande

##### nom-bdd

Nom de la base de données à vérifier. Si ce paramètre est omis, le nom de la base DB2 par défaut est utilisé.

#### Commentaires

Une table configurée pour DB2 Extension Texte doit être réindexée si le groupe de noeuds auquel elle appartient est modifié. Utilisez TXNCHECK pour afficher les tables de ce type qui nécessitent une réindexation en raison de la redistribution d'un groupe de noeuds.

---

**TXNDROP**

Cette commande permet de supprimer un serveur DB2 Extension Texte pour un noeud donné.

**Droits**

(sous UNIX uniquement) SYSADM, DBADM

**Syntaxe de la commande**

►►—txndrop—*numéro-noeud*—————►◄

**Paramètres de la commande****numéro-noeud**

Numéro du noeud à supprimer. Indiquez-le exactement tel qu'il apparaît dans le fichier `db2nodes.cfg`.

**Commentaires**

Utilisez TXNCHECK pour déterminer les index à recréer afin d'assurer la synchronisation avec votre base de données.

## Commande TXSAMPLE

---

### TXSAMPLE

Cette commande crée un modèle de table, charge des exemples de documents en anglais dans la colonne COMMENT et active la colonne de texte. Reportez-vous à la section «Préparation d'une base de données exemple pour vérifier l'installation» à la page 16.

#### Droits

(sous UNIX uniquement) SYSADM, DBADM

#### Syntaxe de la commande

► txsample *—nom-bdd—* *—id-utilisateur—mot-de-passe—* ◀

#### Paramètres de la commande

##### nom-bdd

Nom de la base de données , déjà activée par DB2 Extension Texte, dans laquelle le modèle de table sera créé.

##### id-utilisateur

Paramètre requis uniquement si vous travaillez à partir d'un poste client.

##### mot-de-passe

Paramètre requis uniquement si vous travaillez à partir d'un poste client.

#### Commentaires

Si le CCSID de votre base de données n'est pas 850, reportez-vous au modèle d'étape d'activation pour le type d'index Ngram et modifiez-y le CCSCSID afin qu'il corresponde à celui de votre base de données.

##### Conseil

Cette commande peut également être utilisée sur un poste de travail client.

## TXSTART

Cette commande démarre les fonctions de DB2 Extension Texte.

### Droits

(sous UNIX uniquement) SYSADM, DBADM

### Syntaxe de la commande

```

>> txstart [NODENUM numéro-noeud]
  
```

### Paramètres de la commande

#### NODENUM *numéro-noeud*

Numéro de noeud du serveur à démarrer. Si ce paramètre est omis, tous les serveurs sont démarrés.

### Commentaires

Exécutez cette commande :

- Lorsque vous êtes connecté sous un ID utilisateur appartenant au groupe SMADMIN.
- Chaque fois que vous arrêtez ou redémarrez le système serveur.

Si le démarrage des services DB2 Extension Texte suscite des difficultés, arrêtez DB2 et exécutez la commande TXIPCLEAN. Reportez-vous à la section «Commande TXIPCLEAN (sous UNIX uniquement)» à la page 180.

## Commande TXSTATUS

---

### TXSTATUS

Cette commande indique si DB2 Extension Texte est démarré.

#### Droits

(sous UNIX uniquement) SYSADM, DBADM

#### Syntaxe de la commande

```
► txstatus [NODENUM numéro-noeud] ◀
```

#### Paramètres de la commande

##### **NODENUM** *numéro-noeud*

Numéro de noeud du serveur dont l'état doit être vérifié. Si ce paramètre est omis, l'état de tous les serveurs est affiché.

## TXSTOP

Cette commande arrête les fonctions de DB2 Extension Texte.

### Droits

(sous UNIX uniquement) SYSADM, DBADM

### Syntaxe de la commande

```

▶▶ txstop [NODENUM—numéro-noeud—]

```

### Paramètres de la commande

#### NODENUM numéro-noeud

Numéro de noeud du serveur dont l'état doit être vérifié. Si ce paramètre est omis, tous les serveurs sont arrêtés.

### Commentaires

Cette commande n'arrête pas DB2.

## TXTHESC(commande)

### TXTHESC

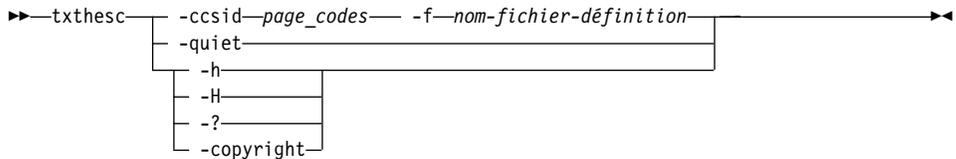
Cette commande compile un fichier de définition de thésaurus. Ce thésaurus ne peut être utilisé que pour des recherches sur des index exacts ou linguistiques.

Pour compiler un fichier de définition de thésaurus Ngram, reportez-vous à la commande «TXTHESN» à la page 190.

#### Droits

(sous UNIX uniquement) Aucun

#### Syntaxe de la commande



#### Paramètres de la commande

##### **-f** *nom-fichier-définition*

Nom du fichier SGML contenant la définition du thésaurus. Le nom du fichier doit contenir le chemin absolu ou relatif d'accès au fichier.

Le dictionnaire du thésaurus est créé dans le même répertoire que le fichier de définition. Il porte le même nom que ce dernier, mais avec les extensions th1 à th6.

##### **Conseil**

Les fichiers thésaurus étant écrasés s'ils portent le même nom, il est préférable d'utiliser un répertoire distinct pour chaque thésaurus.

##### **-ccsid** *page-de-codes*

Page de codes dans laquelle le fichier de définition du thésaurus est écrit. Actuellement, seule la page de codes 850 est prise en charge.

**-quiet** Les résultats ne sont pas affichés.

##### **-copyright**

Renvoie le numéro de compilation interne du produit. Utilisez ce numéro pour signaler les incidents.

##### **-h, -H, or -?**

Affiche l'aide.

**Commentaires**

Utilisez cette commande pour compiler un fichier de définition de thésaurus standard au format binaire. Le fichier de définition doit être au format SGML.

Pour utiliser un fichier thésaurus compilé, transférez-le dans le répertoire du dictionnaire de l'instance du serveur et indiquez l'emplacement des fichiers lors de la recherche.

Le répertoire du dictionnaire sur les systèmes OS/2 et Windows a la désignation suivante :

*unité*:\dmb\db2tx\dict

Sous AIX, HP-UX et SUN-Solaris, le répertoire du dictionnaire est :

*DB2TX\_INSTOWNER*/db2tx/dicts

## Commande TXTHESN

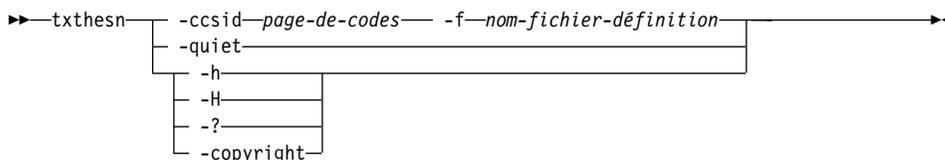
### TXTHESN

Cette commande compile un fichier de définition de thésaurus Ngram. Ce thésaurus ne peut être utilisé que pour des recherches sur des index Ngram.

#### Droits

(sous UNIX uniquement) Aucun

#### Syntaxe de la commande



#### Paramètres de la commande

##### **-f** *nom-fichier-définition*

Nom du fichier contenant la définition du thésaurus. Le nom du fichier doit contenir le chemin absolu ou relatif d'accès au fichier. Le nom de fichier est limité à 8+3 caractères. L'extension est facultative.

Le dictionnaire du thésaurus est créé dans le même répertoire que le fichier de définition. Il a le même nom que le fichier de définition mais avec les extensions : wdf, wdv, grf, grv, MEY, ROS, NEY, SOS, et 1kn, où *n* est un chiffre.

#### Conseil

Les fichiers thésaurus étant écrasés s'ils portent le même nom, il est préférable d'utiliser un répertoire distinct pour chaque thésaurus.

##### **-ccsid** *page-de-codes*

Page de codes dans laquelle le fichier de définition du thésaurus est écrit. Pour obtenir une liste des pages de codes prises en charge, reportez-vous à la section «CCSID» à la page 35.

**-quiet** Les résultats ne sont pas affichés.

##### **-copyright**

Renvoie le numéro de compilation interne du produit. Utilisez ce numéro pour signaler les incidents.

**-h, -H, or -?**

Affiche l'aide.

**Commentaires**

Utilisez cette commande pour compiler un fichier de définition de thésaurus au format binaire. Le fichier de définition doit être au format décrit à la section «Création d'un thésaurus Ngram» à la page 278.

Pour utiliser un fichier thésaurus compilé, transférez-le dans le répertoire du dictionnaire de l'instance du serveur et indiquez l'emplacement des fichiers lors de la recherche.

Le répertoire du dictionnaire sur les systèmes OS/2 et Windows a la désignation suivante :

*unité*: \dmb\db2tx\dict

Sous AIX, HP-UX et SUN-Solaris, le répertoire du dictionnaire est :

*DB2TX\_INSTOWNER*/db2tx/dicts

## TXTRACE(commande)

### TXTRACE

Cette commande génère l'écriture d'informations sur l'exécution en cours dans un tampon de trace situé en mémoire partagée. Ces informations peuvent être utilisées pour la trace des erreurs. Elles peuvent être transférées, en format binaire, du tampon de trace dans un fichier pour être formatées ultérieurement quand la trace sera désactivée, ou formatées et écrites directement dans un fichier alors que la trace est active.

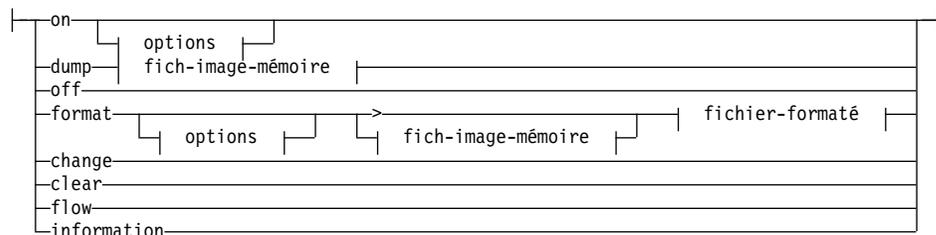
#### Droits

(sous UNIX uniquement) SYSADM, DBADM

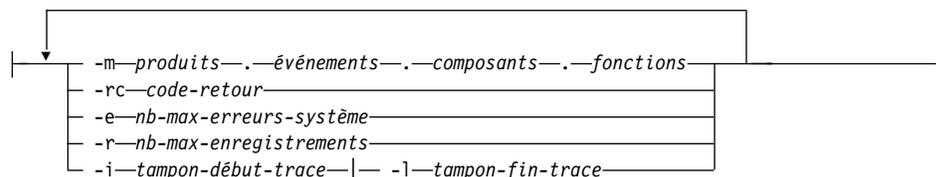
#### Syntaxe de la commande

►txtrace| paramètres |

#### paramètres :



#### options :



#### fich-image-mémoire :



fichier-formaté :



## Paramètres de la commande

**Remarque :**

Il existe également une option `-s` utilisable avec tous les paramètres de TXTRACE. Cette option permet d'afficher les informations concernant le paramètre en question.

**on** Démarre l'utilitaire de trace.

**dump | dmp**  
Ecrit les informations de la trace en binaire dans un fichier.

**off** Arrête l'utilitaire de trace.

**format | fmt**  
Formate les informations de trace de type binaire. Vous pouvez formater le fichier d'image mémoire lorsque la trace est désactivée. Pour cela, indiquez le nom du fichier d'image mémoire ainsi que celui du fichier cible où seront stockées les informations formatées. Pour formater les informations de trace, directement à partir du tampon de trace, alors que celle-ci est active, entrez : `destrc format`  
> `nom-fichier.tmp`.

**change | chg**  
Modifie le masque de la trace, le nombre maximum d'erreurs système ou le nombre maximum d'enregistrements.

**clear | clr**  
Efface la trace.

**flow | flw**  
Affiche le flux de contrôle de la trace.

**information | info | inf**  
Permet d'obtenir des informations sur la trace.

**options**  
Permettent de filtrer les informations de trace, soit au moment où la trace est activée (ce qui permet de réduire l'usage de la mémoire partagée), soit au moment de leur formatage. A moins que la trace soit très volumineuse, il est généralement plus aisé de générer la totalité des informations, puis de les filtrer lors du formatage.

## TXTRACE(commande)

**-m** Ajoute un "masque" permettant de déterminer les événements, composants et fonctions à inclure dans la trace. Par défaut, il n'y a aucun filtre. Un masque comporte quatre segments séparés par des points, par exemple : 2.2-6.1,3.\* Vous pouvez indiquer une plage ou une liste en utilisant respectivement "-" ou "," comme séparateur. Le segment 2-6 inclut uniquement les événements ayant un ID compris entre 2 et 6. Pour n'inclure que les événements 2 et 6, il faudrait indiquer 2,6.

### *produits*

ID Produit. Celui de DB2 Extension Texte est "2". L'ID produit de TextMiner est "3".

### *événements*

Types d'événement à inclure dans la trace :

0	system_error
1	system_error
2	system_error
3	non-fatal_error
4	non-fatal_error
5	api_errcode
7	fnc_errcode
8	trap error
10	api_entry
11	api_exit
13	api_retcode
15	api_data
30	fnc_entry
31	fnc_exit
33	fnc_retcode
35	fnc_data

### *composants*

Composants à inclure dans la trace.

Les ID composant pour DB2 Extension Texte sont :

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

Les ID composant pour TextMiner sont :

- 1 automachine
- 2 bgproc (traitement en arrière-plan)
- 3 cluster
- 4 common
- 5 commsrvc (fonctions communes)
- 6 communic (communication)
- 7 daemon
- 8 dsclient
- 9 environ (environnement)
- 10 glue
- 11 idxcomm (création d'index, tronc commun)
- 12 libsrv (fonctions de bibliothèque)
- 13 search
- 14 trace
- 15 guru
- 16 indexbld (création d'index, TM uniquement)
- 17 indexeng (moteur d'index, TM uniquement)
- 18 smsearch
- 19 searcheng (moteur de recherche, TM uniquement)
- 20 tmsearch
- 21 gtrcm (GTR, tronc commun)
- 22 gtrsrch (recherche, GTR uniquement)
- 23 gridx (création d'index, GTR uniquement)

### *fonctions*

Astérisque (\*). Ensemble des fonctions à inclure dans la trace. Utilisez un astérisque (\*) pour intégrer toutes les fonctions dans la trace, sauf si le support technique IBM vous demande de procéder autrement.

### *-rc code-retour*

Traite les *codes-retour* comme des erreurs système.

### *-e nb-max-erreurs-système*

Nombre entier. Une fois cette valeur atteinte, la fonction de trace est interrompue. La valeur par défaut est 1 ; dans ce cas, lorsque se produit la première erreur système, la trace de tous les événements suivants de moindre gravité est supprimée. Cette valeur n'est acceptable que si vous n'êtes intéressé que par le premier incident majeur. Il est toutefois recommandé d'indiquer un nombre plus élevé (tel que -e 50) si vous souhaitez consulter la trace complète après l'apparition de l'erreur système initiale. Les informations de trace se trouvent en mémoire partagée.

## TXTRACE(commande)

### **-r** *taille-max-enregistrements*

Nombre entier. Arrête la fonction de trace une fois l'espace indiqué occupé par les enregistrements. La valeur par défaut est 16 ko.

### **-i** *tampon-début-trace*

Nombre entier. Réserve cet espace pour les premiers enregistrements de la trace (indique que les enregistrements les plus anciens sont conservés). Si -i est spécifié, la valeur par défaut est 16 ko. Sur des systèmes UNIX, il est recommandé de porter la mémoire tampon à 2 Mo.

Si vous n'indiquez ni -i, ni -l, le système retiendra -l par défaut.

Si vous indiquez -i, aucune nouvelle entrée de la trace ne sera acceptée si le volume des enregistrements dépasse la valeur spécifiée pour *taille-max-enregistrement*, même si vous effacez toutes les entrées de la trace. Pour réactiver l'écriture de nouvelles entrées, augmentez la taille de la mémoire tampon, arrêtez la trace, puis redémarrez-la.

### **-l** *tampon-fin-trace*

Nombre entier. Réserve cet espace pour les derniers enregistrements de la trace (indique que les enregistrements les plus récents sont conservés). La valeur par défaut est 16 ko. Sur des systèmes UNIX, il est recommandé de porter la mémoire tampon à 2 Mo.

### **chemin**

Répertoire où sera stocké le fichier correspondant.

### **nom-fich-image-mémoire**

Nom du fichier contenant les informations de trace de type binaire.

### **nom-fichier-formaté**

Nom du fichier contenant les informations de trace formatées.

## Exemples

Reportez-vous à la section «Trace des incidents» à la page 23.



## Commande TXWIZARD

---

### TXWIZARD (commande)

Cette commande n'est disponible que sous AIX, Windows NT et Windows 2000. Elle fournit un exemple de création d'index sur des tables et des colonnes sélectionnées et de recherche sur ces index.

### Syntaxe de la commande

►—txwizard—◄

### Paramètres de la commande

Aucun.

### Commentaires

Cet assistant vous guide dans les différentes étapes à parcourir pour effectuer une recherche sur les colonnes sélectionnées dans une table à l'aide des fonctions de recherche DB2 Extension Texte. Vous pouvez utiliser le code source Java fourni avec le modèle d'assistant pour réaliser votre propre application de préparation et de recherche de texte ou simplement pour avoir une idée sur la manière d'utiliser et d'appeler les fonctions DB2 Extension Texte dans un environnement Java.

Sous AIX, le code source Java de l'assistant est stocké dans le répertoire  
`/usr/lpp/db2tx_06_01/samples/wizard`

Sous Windows NT et Windows 2000, il est stocké dans le répertoire  
`%DMBMPATH%\samples\wizard`

---

## Chapitre 11. Fonctions de recherche

DB2 Extension Texte fournit des fonctions SQL permettant d'effectuer des recherches dans les documents texte stockés dans DB2 UDB et d'utiliser les résultats de ces recherches. Certains paramètres de ces fonctions appartiennent à des types de données appelés *types distincts* fournis avec DB2 Extension Texte.

Le présent chapitre décrit les fonctions SQL et les types distincts de DB2 Extension Texte.

DB2 Extension Texte fournit un fichier d'entrée pour l'interpréteur de commandes DB2 UDB, appelé `txsample.udf`. Il contient des exemples de fonctions de DB2 Extension Texte utilisables avec les exemples de table présentés à la section «Préparation d'une base de données exemple pour vérifier l'installation» à la page 16. Utilisez ce fichier pour visualiser des exemples de la syntaxe de la préparation de texte et des fonctions de recherche.

---

### Les types distincts dans DB2 Extension Texte

Type distinct	Type de données source	Commentaires
DB2TEXTH	VARCHAR(60) FOR BIT DATA	<p><b>Descripteur de texte.</b> Chaîne de longueur variable contenant des informations nécessaires à l'indexation d'un document texte, stockée dans une colonne de type texte. Les informations contenues dans ce descripteur comprennent un ID document, le nom du serveur sur lequel le texte doit être indexé, le nom de l'index et les informations relatives au document texte.</p> <p>Les descripteurs sont stockés dans des colonnes créées par DB2 Extension Texte et associées à chaque colonne de type texte.</p> <p>Les fonctions <code>HANDLE</code> et <code>INIT_TEXT_HANDLE</code> renvoient ce type de données.</p>

## Types distincts dans DB2 Extension Texte

Type distinct	Type de données source	Commentaires
DB2TEXTFH	VARCHAR(210) FOR BIT DATA	<p><b>Descripteur de fichier.</b> Chaîne de longueur variable contenant des informations nécessaires à l'indexation d'un fichier de texte externe, c'est-à-dire un fichier non contrôlé par DB2 UDB. Les informations contenues dans ce descripteur comprennent un ID document, le nom du serveur sur lequel le texte doit être indexé, le nom de l'index, des informations relatives au fichier de texte et à son emplacement.</p> <p>Les descripteurs sont stockés dans des colonnes créées par DB2 Extension Texte et associées à chaque groupe de fichiers externes.</p> <p>Les fonctions FILE et INIT_TEXT_HANDLE renvoient ce type de données.</p>
DB2TEXTHLISTP	VARCHAR(16) FOR BIT DATA	<p><b>Pointeur de liste de descripteurs.</b> Pointeur sur une liste de descripteurs associé à des documents texte trouvés au cours d'une recherche.</p> <p>La fonction HANDLE_LIST renvoie ce type de données.</p>
DB2TEXTFHLISTP	VARCHAR(16) FOR BIT DATA	<p><b>Pointeur de liste de descripteurs.</b> Pointeur sur une liste de descripteurs associé à des fichiers externes trouvés au cours d'une recherche.</p> <p>La fonction HANDLE_LIST renvoie ce type de données.</p>

### Récapitulatif des fonctions DB2 Extension Texte

Fonction de recherche	Objet	Page
CCSID	Renvoi du CCSID d'un descripteur.	202
CONTAINS	Recherche de texte dans un document particulier.	203
FILE	Renvoi ou modification du chemin et du nom d'un fichier dans un descripteur existant.	204
FORMAT	Renvoi ou modification des paramètres de format de document dans un descripteur.	205
HANDLE <sup>1</sup>	Renvoi d'un descripteur à partir de la liste correspondante.	206
HANDLE_LIST <sup>1</sup>	Recherche et renvoi d'une liste de descripteurs.	207
INIT_TEXT_HANDLE	Renvoi d'un descripteur partiellement initialisé contenant des informations, telles que les paramètres de format et de langue.	208
LANGUAGE	Renvoi ou modification des paramètres de langue dans un descripteur.	209
NO_OF_DOCUMENTS <sup>1</sup>	Renvoi du nombre de documents figurant dans une liste de descripteurs.	210
NO_OF_MATCHES	Recherche et renvoi du nombre de correspondances trouvées.	211
RANK	Recherche et renvoi du rang d'un document texte trouvé.	212
REFINE	Renvoi d'un argument combiné à partir d'un argument de recherche simple et d'un argument de limitation de la recherche.	213
SEARCH_RESULT	Renvoi d'une table intermédiaire contenant le résultat de la recherche sur la chaîne spécifiée.	214

Vous trouverez des exemples de fonctions DB2 Extension Texte au «Chapitre 6. Mode de recherche» à la page 77.

1. Ces fonctions de recherche proviennent d'une précédente version de DB2 Extension Texte. Pour des raisons de compatibilité, elles continuent à être prises en charge mais la fonction SEARCH\_RESULT dispose de fonctionnalités supérieures.

## CCSID

---

### CCSID

La fonction CCSID renvoie le CCSID (type de données SMALLINT) à partir d'un descripteur. Il s'agit du paramètre CCSID utilisé pour l'indexation du document texte correspondant, qui est décrit à la section «CCSID» à la page 35. Ce CCSID est défini pour chaque colonne de type texte par la commande ENABLE TEXT COLUMN.

### Syntaxe de la fonction

►—CCSID—(—*descripteur*—)—————►◄

### Paramètres de la fonction

#### descripteur

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs à partir de laquelle le paramètre CCSID doit être renvoyé.

---

## CONTAINS

La fonction CONTAINS recherche du texte dans un document particulier. Elle renvoie la valeur de type INTEGER 1 si le document en contient. Dans le cas contraire, c'est la valeur 0 qui est renvoyée.

### Syntaxe de la fonction

►—CONTAINS—(—*descripteur*—,—*argument-recherche*—)—————►◄

### Paramètres de la fonction

#### **descripteur**

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs contenant les descripteurs des documents texte dans lesquels la recherche doit être effectuée.

#### **argument-recherche**

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

## FILE

---

## FILE

La fonction FILE exécute l'une des opérations suivantes :

- Renvoi du chemin et du nom de fichier dans un descripteur
- Modification du chemin et du nom de fichier dans un descripteur et renvoi des paramètres modifiés.

Le descripteur renvoyé est de type DB2TEXTFH.

### Syntaxe de la fonction

► FILE(*—descripteur—*)

► FILE(*—descripteur—, —nom-fichier—*)

### Paramètres de la fonction

#### descripteur

Expression dont le résultat est une valeur de type DB2TEXTFH. Il s'agit généralement du nom d'une colonne de descripteurs à partir de laquelle le nom de fichier doit être renvoyé.

#### nom-fichier

Chaîne de type VARCHAR(150) indiquant les nouvelles valeurs du chemin complet et du nom du fichier externe à associer au descripteur. Le chemin peut être, par exemple, une unité du réseau local ou une unité montée via NFS. Les droits d'accès à ce fichier doivent permettre au propriétaire de l'instance DB2 UDB d'accéder à ce fichier.

## FORMAT

La fonction FORMAT exécute l'une des opérations suivantes :

- Renvoi du format de document mentionné dans un descripteur
- Modification de la spécification de format dans un descripteur de document et renvoi du descripteur modifié.

Le format de document renvoyé est une chaîne de type VARCHAR(30). Le descripteur renvoyé est de type DB2TEXTFH ou DB2TEXTH.

Il s'agit du paramètre de format utilisé pour l'indexation du document texte correspondant. Vous trouverez la liste des formats de documents pris en charge à la section «Documents pris en charge» à la page 31.

### Syntaxe de la fonction

►►—FORMAT<sup>(1)</sup>—(—*descripteur*—)—————►►

#### Remarques :

- 1 Renvoie une valeur de format, de type VARCHAR(30).

►►—FORMAT<sup>(1)</sup>—(—*descripteur*—,—*format*—)—————►►

#### Remarques :

- 1 Renvoie un descripteur de type DB2TEXTFH ou DB2TEXTH.

### Paramètres de la fonction

#### descripteur

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs à partir de laquelle le paramètre de format est renvoyé ou défini.

#### format

Paramètre de format du nouveau document de type VARCHAR(30).

Lorsque *format* est mentionné, ce format de document est défini dans le descripteur. Dans ce cas, c'est le descripteur qui est renvoyé à la place du paramètre de format.

## HANDLE

---

## HANDLE

La fonction HANDLE renvoie un descripteur de type DB2TEXTH ou DB2TEXTFH, sélectionné en fonction du numéro d'ordre dans une liste de descripteurs.

La fonction HANDLE ne doit être utilisée que dans une instruction INSERT.

### Conseil

Cette fonction provient d'une précédente version de DB2 Extension Texte. Pour des raisons de compatibilité, elle continue à être prise en charge mais la fonction SEARCH\_RESULT dispose de fonctionnalités supérieures.

## Syntaxe de la fonction

►—HANDLE—(—*liste-descripteurs*—,—*entier*—)—————►

## Paramètres de la fonction

### **liste-descripteurs**

Expression dont le résultat est une valeur de type DB2TEXTHLISTP ou DB2TEXTFHLISTP, renvoyée par la fonction HANDLE\_LIST.

Il s'agit d'un pointeur sur une liste de descripteurs de documents trouvés lors d'une recherche.

**entier** Valeur de type INTEGER désignant le descripteur de la liste à renvoyer.

## HANDLE\_LIST

La fonction HANDLE\_LIST recherche des documents texte à l'aide d'un argument de recherche et renvoie une valeur de type DB2TEXTHLISTP ou DB2TEXTFHLISTP qui pointe sur une liste de descripteurs des documents trouvés.

**Remarque :** Le pointeur sur cette liste n'est accessible que dans la portée de l'instruction SQL utilisant cette fonction.

Si aucun document contenant le terme recherché n'est trouvé, la liste des descripteurs est vide. Utilisez la fonction NO\_OF\_DOCUMENTS pour déterminer si la liste est vide.

Si vous recherchez du texte stocké dans une colonne de table de base, n'utilisez cette fonction qu'une fois par instruction. Si vous recherchez du texte stocké dans une colonne de vue, et que la vue est constituée à partir de plusieurs tables de base (vue combinée), utilisez un descripteur pour chaque colonne comme argument d'entrée de *descripteur* pour la fonction HANDLE\_LIST.

### Conseil

Cette fonction provient d'une précédente version de DB2 Extension Texte. Pour des raisons de compatibilité, elle continue à être prise en charge mais la fonction SEARCH\_RESULT dispose de fonctionnalités supérieures.

## Syntaxe de la fonction

►►—HANDLE\_LIST—(—*descripteur*—,—*argument-recherche*—)—————►►

## Paramètres de la fonction

### descripteur

Expression dont le résultat est une valeur de type DB2TEXTH ou DB2TEXTFH. Il s'agit généralement de la colonne de type descripteur contenant les descripteurs des documents texte devant faire l'objet d'une recherche. Il peut également s'agir d'un descripteur type extrait de la vue de catalogue TEXTCOLUMNS.

### argument-recherche

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

## INIT\_TEXT\_HANDLE

---

### INIT\_TEXT\_HANDLE

La fonction INIT\_TEXT\_HANDLE renvoie un descripteur partiellement initialisé qui contient des valeurs prédéfinies pour la langue ou le format du texte. Ce descripteur peut être inséré dans une colonne de type descripteur. Cela est utile lorsque vous ajoutez une ligne contenant du texte dont la langue et le format sont différents de ceux indiqués dans les définitions de la configuration de texte.

Le descripteur renvoyé est de type DB2TEXTH.

Si vous voulez effectuer une recherche sur du texte stocké dans des *fichiers externes*, et non dans une table DB2 UDB, vous pouvez utiliser la fonction INIT\_TEXT\_HANDLE pour renvoyer un descripteur complètement initialisé, contenant des valeurs prédéfinies pour le CCSID, le format et la langue du texte, ainsi que pour l'emplacement du fichier.

Le descripteur renvoyé est de type DB2TEXTFH.

Utilisez la fonction INIT\_TEXT\_HANDLE pour insérer ou modifier les valeurs contenues dans un descripteur.

### Syntaxe de la fonction

►► INIT\_TEXT\_HANDLE(—*format*—, —*langue*—) ►►

►► INIT\_TEXT\_HANDLE(—*CCSID*—, —*format*—, —*langue*—, —*nom-fichier*—) ►►

### Paramètres de la fonction

#### **format**

Chaîne de type VARCHAR(30) indiquant le paramètre de format du nouveau document. Les formats pris en charge sont indiqués à la section «Documents pris en charge» à la page 31.

#### **langue**

Chaîne de type VARCHAR(30) indiquant le paramètre de langue du nouveau document. Les langues prises en charge sont indiquées à la section tableau 5 à la page 54.

#### **nom-fichier**

Chaîne de type VARCHAR(150) indiquant le chemin complet et le nom du fichier externe à associer au descripteur. Pour avoir accès au fichier sous UNIX, le propriétaire de l'instance DB2 UDB doit disposer des droits d'accès. Sous OS/2 et Windows, les droits d'accès au fichier doivent inclure l'ID utilisateur servant à la connexion.

---

**LANGUAGE**

La fonction LANGUAGE exécute l'une des opérations suivantes :

- Renvoi de la langue du document mentionnée dans un descripteur
- Modification de la spécification de la langue dans un descripteur de document et renvoi du descripteur modifié.

La langue est renvoyée sous forme de chaîne de type VARCHAR(30). Le descripteur renvoyé est de type DB2TEXTFH ou DB2TEXTH.

Il s'agit du paramètre de langue utilisé pour l'indexation du document texte correspondant. Les langues prises en charge sont indiquées au tableau 5 à la page 54.

### Syntaxe de la fonction

►►—LANGUAGE<sup>(1)</sup>—(—*descripteur*—)—————►►

#### Remarques :

- 1 Renvoie une valeur de langue, de type VARCHAR(30).

►►—LANGUAGE<sup>(1)</sup>—(—*descripteur*—,—*langue*—)—————►►

#### Remarques :

- 1 Renvoie un descripteur de type DB2TEXTFH ou DB2TEXTH.

### Paramètres de la fonction

#### descripteur

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs à partir de laquelle le paramètre de langue est renvoyé ou défini.

#### langue

Paramètre de langue du nouveau document de type VARCHAR(30).

Si le paramètre *langue* est mentionné, la langue du document est définie dans le descripteur et c'est ce dernier qui est renvoyé à la place du paramètre de langue.

## NO\_OF\_DOCUMENT

---

## NO\_OF\_DOCUMENTS

La fonction NO\_OF\_DOCUMENTS renvoie une valeur de type INTEGER indiquant le nombre d'éléments contenus dans une liste de documents texte trouvés au cours d'une recherche. La valeur renvoyée est le nombre d'entrées figurant dans une liste de descripteurs.

### Conseil

Cette fonction provient d'une précédente version de DB2 Extension Texte. Pour des raisons de compatibilité, elle continue à être prise en charge mais la fonction SEARCH\_RESULT dispose de fonctionnalités supérieures.

### Syntaxe de la fonction

►—NO\_OF\_DOCUMENTS—(*—liste-descripteurs—*)—◄

### Paramètres de la fonction

#### liste-descripteurs

Expression dont le résultat est une valeur de type DB2TEXTHLISTP ou DB2TEXTFHLISTP, renvoyée par la fonction HANDLE\_LIST.

Il s'agit d'un pointeur sur une liste de descripteurs de documents trouvés lors d'une recherche.

Les fonctions UDF HANDLE\_LIST et NO\_OF\_DOCUMENTS doivent se trouver dans la même instruction SQL car la liste n'existe que dans la portée de l'instruction.

---

## NO\_OF\_MATCHES

NO\_OF\_MATCHES peut effectuer une recherche dans des documents texte et renvoyer une valeur INTEGER indiquant le nombre de correspondances par document.

### Syntaxe de la fonction

►►—NO\_OF\_MATCHES—(—*descripteur*—,—*argument-recherche*—)—►►

### Paramètres de la fonction

#### **descripteur**

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs contenant les descripteurs des documents texte dans lesquels la recherche doit être effectuée.

#### **argument-recherche**

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

## RANK

---

### RANK

RANK peut effectuer une recherche dans des documents texte et renvoyer une valeur correspondant au rang de chaque document trouvé, en indiquant le degré de pertinence du document.

RANK renvoie une valeur de type DOUBLE comprise entre 0 et 1. La valeur de rang est une valeur absolue, indiquant le degré pertinence du document trouvé par rapport aux autres. Elle indique le nombre de correspondances trouvées dans le document par rapport à sa taille.

### Syntaxe de la fonction

►►RANK(—*descripteur*—,—*argument-recherche*—)◄◄

### Paramètres de la fonction

#### **descripteur**

Expression dont le résultat est une valeur de type DB2TEXTFH ou DB2TEXTH. Il s'agit généralement du nom d'une colonne de descripteurs contenant les descripteurs des documents texte dans lesquels la recherche doit être effectuée.

#### **argument-recherche**

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

---

**REFINE**

La fonction REFINE prend en compte deux arguments de recherche et renvoie un argument de recherche combiné de type LONG VARCHAR, se composant des deux arguments d'origine reliés par l'opérateur booléen AND.

**Syntaxe de la fonction**

►►REFINE(—*argument-recherche*—,—*argument-recherche*—)◄◄

**Paramètres de la fonction****argument-recherche**

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

L'argument de recherche ne doit pas comporter les paramètres de recherche IS ABOUT, THESAURUS ou EXPAND.

## SEARCH\_RESULT

---

### SEARCH\_RESULT

La fonction SEARCH\_RESULT renvoie le résultat d'une recherche dans une table intermédiaire. Cette fonction peut être utilisée dans une clause FROM d'une instruction SQL.

La table renvoyée présente la structure suivante :

Column Name	Data Type
HANDLE	DB2TX.DB2TEXTH, DB2TX.DB2TEXTFH
NUMBER_OF_MATCHES	INTEGER
RANK	DOUBLE

Les valeurs ne sont générées que pour les colonnes sélectionnées de la table intermédiaire. Select count(\*) ne génère que la colonne HANDLE. Le calcul des valeurs RANK consommant une importante part de ressources systèmes, il est préférable de ne pas sélectionner la valeur RANK d'une table intermédiaire si celle-ci n'est pas obligatoire.

Cette fonction est plus rapide que CONTAINS ou RANK, lorsqu'il s'agit de traiter des tables de grande taille.

### Syntaxe de la fonction

►►SEARCH\_RESULT(—schéma—,—table—,—descripteur—,—argument-recherche—)◄◄

### Paramètres de la fonction

#### schéma

Nom de schéma de la table dans laquelle vous effectuez des recherches.

**table** Nom de la table dans laquelle vous effectuez des recherches.

#### descripteur

Nom d'une colonne de descripteurs correspondant à la colonne contenant les documents dans lesquels s'effectue la recherche.

#### argument-recherche

Chaîne de type LONG VARCHAR contenant les termes à rechercher. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

### Exemples

Pour un exemple, reportez-vous à la section «Amélioration des performances de recherche» à la page 99, ou consultez les exemples de fonctions DB2 Extension Texte proposés dans le fichier décrit à la section «Localisation d'exemples de syntaxe pour les fonctions de recherche» à la page 77.

---

## Chapitre 12. Syntaxe des arguments de recherche

Un argument de recherche est la condition que vous indiquez lors de la recherche de termes dans des documents texte. Il se compose d'un ou de plusieurs termes et paramètres de recherche.

Vous trouverez des exemples d'arguments de recherche à la section «Définition des arguments de recherche» à la page 86, ainsi que dans le fichier `txsample.udf`. Il contient des exemples de fonctions de DB2 Extension Texte utilisables avec le modèle de table présenté à la section «Préparation d'une base de données exemple pour vérifier l'installation» à la page 16.

Les fonctions DB2 Extension Texte qui utilisent les arguments de recherche sont les suivantes :

- **CONTAINS**. Cette fonction utilise un argument de recherche pour rechercher du texte dans un document particulier. Elle renvoie la valeur de type `INTEGER` 1 si le document en contient. Dans le cas contraire, c'est la valeur 0 qui est renvoyée.
- **NO\_OF\_MATCHES**. Cette fonction utilise un argument de recherche pour effectuer une recherche dans des documents texte. Elle renvoie une valeur de type `INTEGER` indiquant le nombre de correspondances par document.
- **RANK**. Cette fonction utilise un argument de recherche pour effectuer une recherche dans des documents texte. Pour chaque document trouvé, elle renvoie une valeur indiquant le degré de pertinence du document trouvé.
- **REFINE**. Cette fonction prend en compte deux arguments de recherche et renvoie un argument de recherche combiné de type `LONG VARCHAR`, se composant des deux arguments d'origine reliés par l'opérateur booléen `AND`.
- **HANDLE\_LIST**. Cette fonction utilise un argument de recherche pour effectuer une recherche dans des documents texte. Elle renvoie une valeur de type `DB2TEXTHLISTP` ou `DB2TEXTFHLISTP` qui pointe sur une liste de descripteurs des documents trouvés.
- **SEARCH\_RESULT**. Cette fonction renvoie une table contenant les informations requises, à savoir le rang, le nombre de correspondances et le descripteur.

Les fonctions API qui utilisent les arguments de recherche sont les suivantes :

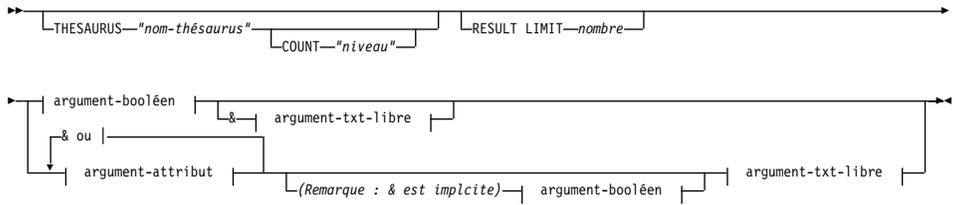
- **DesGetBrowseInfo**. Cette fonction utilise un argument de recherche pour effectuer une recherche dans du texte identifié par un descripteur. Elle renvoie un pointeur sur les informations d'affichage nécessaires à la fonction `DesStartBrowseSession` pour la mise en évidence des termes.

## Syntaxe des arguments de recherche

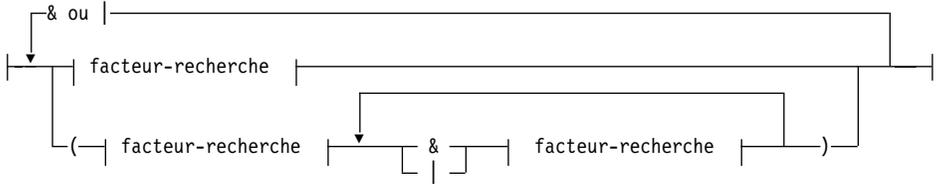
- DesGetSearchResultTable. Cette fonction utilise un argument de recherche pour effectuer une recherche dans des documents texte identifiés par une colonne de type texte. Le contenu du descripteur correspondant aux éléments de texte trouvés est inscrit dans une table de résultats. Les informations d'affichage concernant le rang et le nombre de correspondances peuvent aussi être inscrites dans la table de résultats.

Argument de recherche

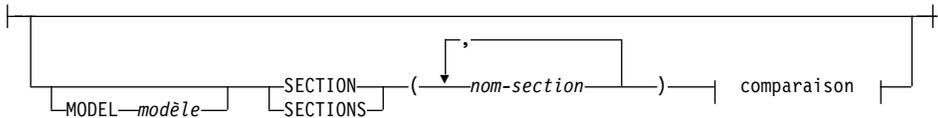
Syntaxe de l'argument de recherche



argument\_booléen :



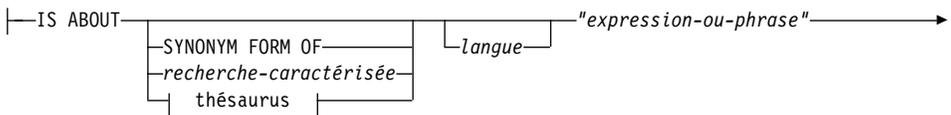
argument-attribut :



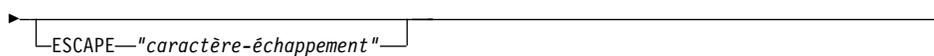
Comparaison :



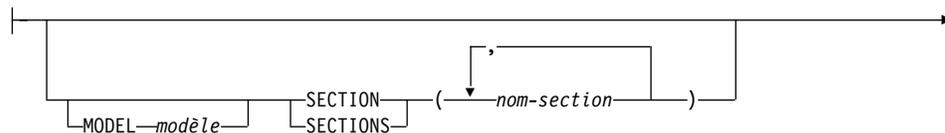
argument-txt-libre :



## Syntaxe des arguments de recherche

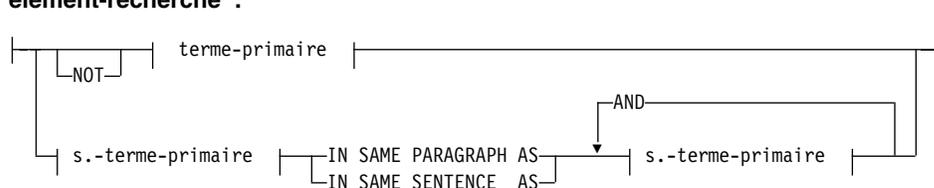


### facteur-recherche :

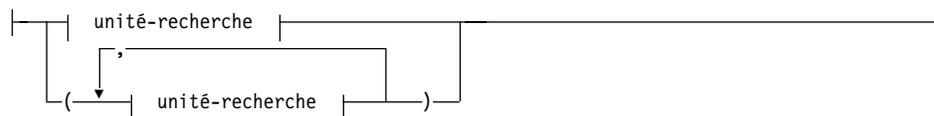


### élément-recherche

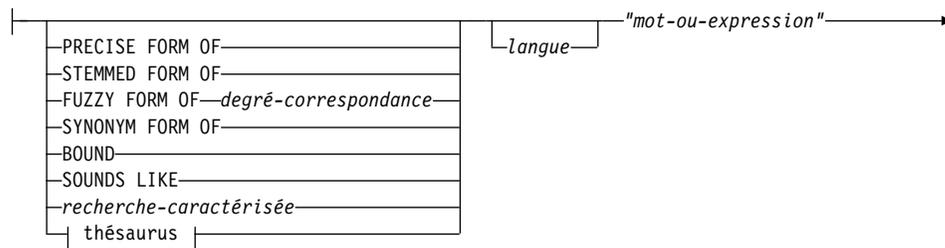
### élément-recherche :



### terme\_primaire :



### unité-recherche :



thésaurus (si THESAURUS est spécifié) :



### Exemples

Des exemples sont fournis à la section «Définition des arguments de recherche» à la page 86.

### Paramètres de recherche

#### IS ABOUT

Option permettant d'entrer un argument de recherche en texte libre, c'est-à-dire une expression ou une phrase en langage naturel décrivant le concept à trouver. Pour plus de détails, reportez-vous à la section «Recherche sur texte libre et recherche hybride» à la page 94.

#### MODEL *modèle*

Mot clé utilisé pour indiquer le nom du modèle de document à utiliser dans le cadre de la recherche. Le modèle de document décrit la structure des documents contenant des sections identifiables, afin de restreindre la recherche au contenu de ces sections.

Le nom du modèle doit être spécifié dans un fichier de modèles de documents décrit à la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48. Le nom du modèle peut être masqué en utilisant des caractères génériques.

Si vous ne spécifiez pas de modèle, le modèle par défaut spécifié lors de la création de l'index est utilisé.

#### SECTION(S) *nom-section*

Mot clé utilisé pour spécifier la ou les sections auxquelles la recherche doit être restreinte. Le nom de section doit être spécifié dans un modèle lui-même défini dans un fichier de modèles de documents, comme décrit à la section «Utilisation de documents structurés (prise en charge de sections)» à la page 48. Un nom de section peut être masqué au moyen des caractères génériques % et \_.

Certaines sections peuvent être imbriquées dans d'autres, comme par exemple :

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

**Restrictions** : Les recherches sur les sections imbriquées ne peuvent s'effectuer que dans les documents stockés dans des colonnes activées

## Syntaxe des arguments de recherche

avec le format XML. Pour les index Ngram, la recherche ne peut porter que sur un seul nom de section et le format XML n'est pas pris en charge.

*<valeur-attribut*

Valeur précédée d'un opérateur de comparaison pour les attributs énumérés dans la liste de sections qui précède. Une requête pour laquelle une valeur d'attribut doit être comprise dans une plage déterminée peut utiliser deux opérateurs de comparaison dans une même condition d'attribut.

Une combinaison d'opérateurs utilisant le même type de comparaison, par exemple  $\geq$  dans le premier et  $>$  dans le second, dans une même condition n'est pas autorisée. Il n'est également pas permis de définir deux comparaisons comportant l'opérateur  $=$ .

- $=$  Indique que l'attribut du document indexé doit être égal à la valeur d'attribut qui suit.
- $\geq$  Indique que l'attribut du document indexé doit être supérieur ou égal à la valeur d'attribut qui suit.
- $>$  Indique que l'attribut du document indexé doit être supérieur à la valeur d'attribut qui suit.
- $\leq$  Indique que l'attribut du document indexé doit être inférieur ou égal à la valeur d'attribut qui suit.
- $<$  Indique que l'attribut du document indexé doit être inférieur à la valeur d'attribut qui suit.

Certaines sections peuvent être imbriquées dans d'autres, comme par exemple :

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

**Restrictions** : Les recherches sur les sections imbriquées ne peuvent s'effectuer que dans les documents stockés dans des colonnes activées avec le format XML. Pour les index Ngram, la recherche ne peut porter que sur un seul nom de section et le format XML n'est pas pris en charge.

**THESAURUS** *nom-thésaurus*

Mot clé utilisé pour indiquer le nom du thésaurus à utiliser pour l'extension d'un terme de recherche. Le nom du thésaurus correspond au nom de fichier (sans l'extension) du thésaurus préalablement compilé à l'aide du compilateur de thésaurus TXTHESC ou TXTHESN. Les thésaurus par défaut, desthes et desnthes (thésaurus

Ngram), sont stockés dans le répertoire SAMPLE. Vous pouvez également spécifier le chemin du fichier. Le chemin par défaut est celui du dictionnaire.

### **COUNT** *niveau*

Mot clé utilisé pour indiquer le nombre de niveaux à prendre en compte dans la classification des termes du thésaurus, dans le cadre de l'extension d'une recherche pour une relation déterminée. Si vous n'indiquez pas ce mot clé, le système lui attribue par défaut la valeur 1.

### **RESULT LIMIT** *nombre*

Mot clé servant à indiquer le nombre maximal d'entrées à renvoyer dans la liste de résultats. *nombre* est une valeur comprise entre 1 et 32767. Si une recherche sur texte libre est spécifiée, les degrés de correspondance indiqués dans la liste de résultats se rapportent à la liste complète. Sinon, les degrés de correspondance indiqués dans le résultat de recherche limité ne se rapportent qu'aux entrées de la liste.

### **EXPAND** *relation*

Mot clé utilisé pour indiquer la relation, telle qu'une INSTANCE, existant entre le terme recherché (spécifié par TERM OF) et les termes du thésaurus à utiliser pour élargir le champ de la recherche. Le nom de la relation doit correspondre à une relation utilisée dans le thésaurus. Pour plus de détails, reportez-vous à la section «Thésaurus» à la page 268.

Pour un thésaurus Ngram, utilisez le nom membre-relation décrit à la section «Création d'un thésaurus Ngram» à la page 278. Pour les relations de membres définies par l'utilisateur, utilisez :RELATION *n* où *n* est le numéro de relation de membre indiqué dans :RELATED (*nombre*).

### **TERM OF** *"expression-ou-phrase"*

Terme ou expression à rechercher auquel seront ajoutés d'autres termes provenant du thésaurus dans le cadre de la recherche.

### **facteur-recherche**

Opérande pouvant être associé à d'autres opérandes pour constituer un argument de recherche. L'évaluation se fait de gauche à droite. L'opérateur logique AND (&) crée une liaison plus forte que l'opérateur logique OR (|). Exemple :

```
"passager" & "véhicule" | "transport" & "public"
```

est évalué comme suit :

```
("passager" & "véhicule") | ("transport" & "public")
```

## Syntaxe des arguments de recherche

Pour rechercher :

```
"passager" & ("véhicule" | "transport") & "public"
```

vous devez placer les parenthèses comme indiqué.

### **NOT terme-primaire**

Opérateur qui permet d'exclure de la recherche des documents texte contenant un terme particulier.

Lorsque NOT est utilisé dans un facteur de recherche, vous ne pouvez pas utiliser le mot clé SYNONYM FORM OF.

### **terme-primaire IN SAME PARAGRAPH AS terme-primaire**

Mot clé qui permet de rechercher une combinaison de termes figurant dans un même paragraphe.

L'argument de recherche suivant permet de rechercher des documents texte contenant le terme "trafic", uniquement s'il est associé au terme "aérien" dans le même paragraphe.

```
"trafic" IN SAME PARAGRAPH AS "aérien"
```

Vous ne pouvez pas utiliser le mot clé IN SAME PARAGRAPH AS lorsque NOT est utilisé dans un facteur de recherche.

### **terme-primaire IN SAME SENTENCE AS terme-primaire**

Mot clé qui permet de rechercher une combinaison de termes figurant dans une même expression. Similaire à IN SAME PARAGRAPH AS.

### **AND terme-primaire**

Mot clé permettant de combiner plusieurs termes primaires de recherche devant être recherchés dans la même expression ou le même paragraphe.

L'argument de recherche suivant permet de rechercher les mots "forêt", "pluie", "érosion" et "terre" situés dans la même phrase.

```
"forêt" IN SAME SENTENCE AS "pluie" AND "érosion" AND "terre"
```

### **unité-recherche**

Si vous reliez une série d'unités de recherche par des virgules, la recherche aboutit lorsque l'une des unités est trouvée. Chaque unité de recherche doit contenir au moins un mot ou une expression.

L'argument suivant est vrai si un ou plusieurs des arguments de recherche est (sont) trouvé(s).

```
CONTAINS (mondescripteur, '( "texte",  
"graphique",  
"audio",  
"vidéo" )') = 1
```

### PRECISE FORM OF, STEMMED FORM OF, FUZZY FORM OF, SYNONYM FORM OF, BOUND

Le tableau 8 présente les options qui correspondent aux divers types d'index. Par exemple, pour un index linguistique, toutes les options conviennent, excepté PRECISE FORM OF. Si vous indiquez l'option PRECISE FORM OF, elle n'est pas prise en compte. C'est la valeur par défaut qui est utilisée.

Le processus de recherche des termes est décrit plus en détail dans le tableau 9.

Tableau 8. Options linguistiques

Mot clé	Type d'index				
	Linguistique	Exact	Exact Normalisé	Ngram	Ngram dist. maj/min
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=paramètre par défaut O=fonction disponible

Tableau 9. Options de recherche de termes disponibles pour les index Ngram

Mot clé	Processus de recherche				
	Distinction majuscules/minuscules		Lemmisation	Correspondance	
	Active	Ignorée		Exacte	Partielle
PRECISE FORM OF	avec dist. maj/min	X		X	
STEMMED FORM OF		X	X		
FUZZY FORM OF		X			X

X=paramètre par défaut

## Syntaxe des arguments de recherche

Si, pour un type d'index déterminé, vous utilisez un mot clé non disponible, ce dernier sera ignoré et le système prendra en compte le mot clé par défaut ou renverra un message.

### PRECISE FORM OF

Mot clé qui permet de rechercher le mot (ou chaque mot de l'expression) suivant PRECISE FORM OF tel qu'il a été saisi, et non pas réduit à sa forme de base. Pour les index exacts, ce type de recherche tient compte de l'utilisation des majuscules et des minuscules. Par exemple, si vous recherchez cheval, vous ne trouverez pas "Cheval".

Il s'agit de l'option par défaut pour les index exacts. Pour les index exacts normalisés, le type de recherche par défaut ne tient pas compte des majuscules et des minuscules. Si vous indiquez ce mot clé pour un index linguistique, il est ignoré et c'est l'option STEMMED FORM OF qui est prise en compte.

### STEMMED FORM OF

Mot-clé qui permet de rechercher le mot (ou chaque mot de l'expression) suivant STEMMED FORM OF tel qu'il est après avoir été réduit à sa forme de base. Ce type de recherche ne tient pas compte des majuscules et des minuscules. Par exemple, si vous recherchez cheval, vous pouvez trouver "Cheval".

La façon dont les mots sont réduits à leur forme de base dépend de la langue utilisée.

Exemple : programming computer systems est remplacé par program compute system lorsque vous utilisez le dictionnaire d'anglais américain et par programme compute system lorsque vous utilisez le dictionnaire d'anglais britannique.

Cette expression de recherche permet de trouver "programmer computes system", "program computing systems", "programming computer system", etc.

Il s'agit de l'option par défaut pour les index linguistiques. Si vous indiquez ce mot clé pour un index exact, il est ignoré et c'est l'option PRECISE FORM OF qui est prise en compte.

### FUZZY FORM OF

Mot clé permettant d'effectuer une recherche de correspondance "partielle", sur des termes ayant une orthographe similaire au terme initial de la recherche. Cette option est particulièrement utile dans le cas de documents créés par un programme de reconnaissance optique des caractères (OCR). Ces documents contiennent souvent des mots mal orthographiés. Par exemple, le mot économie peut

être lu par un programme de reconnaissance optique sous la forme économique. *degré-correspondance* : Nombre entier de 1 à 5 indiquant le degré de similarité par ordre décroissant (5 indiquant un plus fort niveau de ressemblance que 1).

### SYNONYM FORM OF

Mot clé qui permet de rechercher le mot ou l'expression qui suit SYNONYM FORM OF, ainsi que ses synonymes. Les synonymes sont fournis par le dictionnaire indiqué par l'option *langue* ou par le dictionnaire par défaut.

Les synonymes d'une expression sont des expressions contenant toutes les combinaisons possibles de synonymes pouvant être obtenues en remplaçant chaque mot de l'expression d'origine par l'un de ses synonymes. L'ordre des mots reste identique à celui de l'expression d'origine.

Si vous indiquez ce mot clé pour un index exact, il est ignoré et c'est l'option PRECISE FORM OF qui est prise en compte.

Vous ne pouvez pas indiquer ce mot clé lorsque NOT est utilisé dans le facteur de recherche, ou lorsque le mot ou l'expression à rechercher contient des caractères génériques.

### BOUND

Mot clé permettant d'effectuer des recherches dans des documents utilisant le CCSID coréen. Si la variable *langue* est indiquée, elle est ignorée. Ce mot clé est, en effet, spécifique du coréen.

*langue* Variable qui détermine le dictionnaire utilisé pour l'analyse linguistique de documents lors de leur indexation et de leur extraction. Elle s'applique non seulement aux index linguistiques, mais aussi aux index exacts car ces derniers utilisent un dictionnaire pour analyser les mots vides.

L'analyse linguistique comprend le traitement des synonymes et des racines de mots.

Les langues prises en charge sont indiquées au tableau 5 à la page 54.

**Remarque :** Lorsque vous effectuez des recherches dans des documents rédigés dans d'autres langues que l'anglais américain, indiquez la langue correspondante dans l'argument de recherche *sans tenir compte de la langue par défaut*.

*"mot-ou-expression"*

Mot ou expression à rechercher. Les caractères pouvant être utilisés dans un mot dépendent de la langue utilisée. Il en est de même pour

## Syntaxe des arguments de recherche

la séparation des mots par des caractères de séparation. En anglais et dans la plupart des autres langues, les mots d'une expression doivent être séparés par un blanc.

Recherche exacte ou linguistique : DB2 Extension Texte peut effectuer une recherche en utilisant la forme exacte d'un mot ou d'une expression, ou une de ses variantes. Si vous ne précisez pas l'une des options du tableau 8 à la page 223, les options linguistiques par défaut sont utilisées en fonction du type d'index choisi.

Pour rechercher une chaîne de caractères contenant des guillemets, entrez deux fois des guillemets. Par exemple, pour rechercher le texte caractère "générique", entrez

```
"caractère "générique""
```

Caractères génériques : Un mot peut contenir les caractères génériques suivants :

**\_ (trait de soulignement)**

Représente tout caractère simple.

**% (pour-cent)**

Représente un nombre quelconque de caractères quelconques.

Si un mot se compose d'un unique signe %, il représente

n'importe quel mot d'une longueur quelconque.

Un mot ne peut pas être composé exclusivement de caractères génériques, sauf lorsque le signe % est utilisé pour représenter un mot quelconque.

Si vous utilisez un caractère générique, vous ne pouvez pas utiliser les options SYNONYM FORM OF, *recherche\_caractérisée*, ni THESAURUS.

**ESCAPE** *caractère-échappement*

Caractère qui identifie le caractère suivant comme devant être recherché et non comme un caractère générique.

Exemple : si *caractère-échappement* est \$, alors \$%, \$\_ et \$\$ représentent respectivement %, \_ et \$. Tout caractère % ou \_ non précédé de \$ représente un caractère générique.

**Résumé des règles et restriction:**

**Opérations booléennes**

NOT n'est pas autorisé après OR.

### **FUZZY FORM OF**

Les 3 premiers caractères doivent correspondre. Ne peut pas être utilisé si un mot de l'unité de recherche contient un caractère générique. Ne peut pas être utilisé avec NOT. S'utilise uniquement avec un index Ngram.

### **IN SAME PARAGRAPH AS**

Ne peut être utilisé si NOT est utilisé dans un facteur de recherche.

### **IN SAME SENTENCE AS**

Ne peut être utilisé si NOT est utilisé dans un facteur de recherche.

### **Index linguistique**

Interdit l'utilisation de PRECISE FORM OF. Prend en compte STEMMED FORM OF par défaut. Autorise l'utilisation de caractères génériques. Ne tient pas compte de la distinction majuscules/minuscules.

### **Caractères génériques**

Interdit l'utilisation de SYNONYM FORM OF et THESAURUS.

### **Index Ngram**

Autorise l'utilisation de caractères génériques, même non précédés de caractères non alphanumériques. Les recherches ne tiennent pas compte de la distinction majuscules/minuscules, sauf en cas d'utilisation des options CASE-ENABLED et PRECISE FORM OF.

**NOT** Interdit l'utilisation de SYNONYM FORM OF, IN SAME PARAGRAPH AS et IN SAME SENTENCE AS.

### **PRECISE FORM OF**

Option ignorée dans le cas d'un index linguistique.

### **Index exact**

Interdit l'utilisation de STEMMED FORM OF et SYNONYM FORM OF. Prend en compte PRECISE FORM OF par défaut. Autorise l'utilisation de caractères génériques. Tient compte de la distinction majuscules/minuscules.

### **STEMMED FORM OF**

Est ignoré pour un index exact, mais autorisé avec un index exact normalisé contenant des documents en anglais.

### **SYNONYM FORM OF**

Ne peut pas être utilisé si un mot de l'unité de recherche contient un caractère générique. Ne peut pas être utilisé avec NOT. Ne peut pas être utilisé dans un index exact.

## Syntaxe des arguments de recherche

---

## Chapitre 13. Fonctions API de recherche et d'affichage

DB2 Extension Texte fournit des fonctions en langage C permettant de rechercher et d'afficher des documents texte. Ces fonctions constituent l'interface de programmation d'applications (API) de DB2 Extension Texte. Le présent chapitre décrit les fonctions API dans l'ordre alphabétique.

Le «Chapitre 8. Utilisation des fonctions API pour la recherche et l'affichage» à la page 119, présente ces fonctions et décrit comment les utiliser conjointement.

Fonction	Objet	Page
DesCloseDocument	Libère la mémoire allouée par DesOpenDocument.	231
DesEndBrowseSession	Ferme une session d'affichage et libère la mémoire allouée par DesStartBrowseSession.	232
DesFreeBrowseInfo	Libère la mémoire allouée par DesGetBrowseInfo.	233
DesGetBrowseInfo	Recherche du texte dans le document à l'aide d'un argument de recherche et crée des informations d'affichage.	234
DesGetMatches	Renvoie un pointeur sur les informations de mise en évidence pour le document texte décrit par un descripteur de document. Les informations de mise en évidence constituent un flux de données. Elles comprennent le contexte (au moins un paragraphe) et les informations de mise en évidence de texte dans ce contexte.	237
DesGetSearchResultTable	Utilise un argument de recherche afin de rechercher des documents de texte dans une colonne déterminée et stocke le résultat dans une table fournie par l'utilisateur. Peut également renvoyer des informations d'affichage.	243
DesOpenDocument	Reçoit un pointeur de session d'affichage, un descripteur et une option DES_FAST ou DES_EXTENDED indiquant le type d'analyse linguistique à utiliser pour la mise en évidence des termes trouvés. Prépare le document texte qui correspond au descripteur pour extraire le texte du document et les informations de mise en évidence, et renvoie un descripteur de document utilisé pour l'appel itératif de DesGetMatches.	248
DesStartBrowseSession	Ouvre une session d'affichage à l'aide des informations d'affichage provenant de DesGetBrowseInfo et renvoie un descripteur de session d'affichage qui sera utilisé par les autres fonctions d'affichage.	251

### Conseil

La plupart des fonctions API nécessitent un descripteur de connexion (hdbc). Vous devez le fournir via la fonction `SQLConnect`, mais cela ne vous interdit pas de lancer DB2 Extension Texte à partir de programmes en SQL intégré. Le manuel *DB2 Call Level Interface Guide and Reference* décrit comment associer des instructions CLI à des instructions en SQL intégré.

## DesCloseDocument

### Objectif

Ferme un document texte ouvert par DesOpenDocument et libère la mémoire allouée lors du renvoi du texte du document et des informations de mise en évidence.

### Syntaxe

```
DESRETURN
DesCloseDocument
( DESBROWSESESSION BrowseSession,
  DESHANDLE DocumentHandle);
```

### Arguments de la fonction

Tableau 10. Arguments de DesCloseDocument

Type de données	Argument	Utilisation	Description
DESBROWSESESSION	<i>BrowseSession</i>	Entrée	Descripteur de la session d'affichage.
DESHANDLE	<i>DocumentHandle</i>	Entrée	Descripteur renvoyé par DesOpenDocument identifiant un document texte ouvert.

### Codes retour

```
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
```

### Restrictions

Cette fonction ne peut être appelée qu'après l'ouverture d'un document texte par appel de DesOpenDocument.

## Fonction API DesEndBrowseSession

---

### DesEndBrowseSession

#### Objectif

Met fin à une session d'affichage démarrée par DesStartBrowseSession et libère la mémoire allouée pour la session d'affichage.

#### Syntaxe

```
DESRETURN  
DesEndBrowseSession  
(DESBROWSESESSION BrowseSession);
```

#### Arguments de la fonction

Tableau 11. Arguments de DesEndBrowseSession

Type de données	Argument	Utilisation	Description
DESBROWSESESSION	<i>BrowseSession</i>	Entrée	Descripteur de la session d'affichage.

#### Utilisation

Cette fonction ne libère pas la mémoire allouée par DesGetBrowseInfo pour la session d'affichage. Cette mémoire contient des informations d'affichage qui restent exploitables par d'autres sessions d'affichage. Pour libérer cette mémoire, appelez DesFreeBrowseInfo.

#### Codes retour

```
RC_SUCCESS  
  
RC_INVALID_SESSION  
RC_INVALID_PARAMETER  
RC_SE_UNEXPECTED_ERROR
```

#### Restrictions

Cette fonction ne peut être appelée qu'après le démarrage d'une session d'affichage par appel de DesStartBrowseSession.

## DesFreeBrowseInfo

### Objectif

Libère la mémoire allouée pour les informations d'affichage par DesGetBrowseInfo.

### Syntaxe

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

### Arguments de la fonction

Tableau 12. Arguments de DesFreeBrowseInfo

Type de données	Argument	Utilisation	Description
DESBROWSEINFO	<i>BrowseInfo</i>	Entrée	Informations d'affichage.

### Codes retour

RC\_SUCCESS

RC\_INVALID\_PARAMETER

### Restrictions

Cette fonction ne peut être appelée qu'après allocation de mémoire pour des informations d'affichage par appel de DesGetBrowseInfo.

## Fonction API DesGetBrowseInfo

### DesGetBrowseInfo

#### Objectif

Reçoit un argument de recherche pour effectuer une recherche dans du texte identifié par un descripteur. Elle renvoie un pointeur sur les informations d'affichage nécessaires à la fonction DesStartBrowseSession pour la mise en évidence des termes trouvés.

#### Syntaxe

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

#### Arguments de la fonction

Tableau 13. Arguments de DesGetBrowseInfo

Type de données	Argument	Utilisation	Description
SQLHDBC	<i>hdbc</i>	Entrée	Descripteur de connexion à une base de données.
SQLCHAR *	<i>pHandle</i>	Entrée	Pointeur sur un descripteur extrait de la base de données.
DESUSHORT	<i>HandleLength</i>	Entrée	Longueur de pHandle. DES_NTS ne peut pas être utilisé ici.
char *	<i>pSearchArgument</i>	Entrée	Pointeur sur l'argument de recherche de texte qui indique les informations à trouver.
DESSMALLINT	<i>ArgumentLength</i>	Entrée	Longueur de pSearchArgument (sans compter le caractère de fin d'octet NULL), ou DES_NTS.
DESBROWSEINFO *	<i>pBrowseInfo</i>	Sortie	Pointeur sur des informations d'affichage contenant les données nécessaires à l'affichage d'un document texte. Ce pointeur est transmis à DesStartBrowseSession.

Tableau 13. Arguments de DesGetBrowseInfo (suite)

Type de données	Argument	Utilisation	Description
DESMESSAGE *	<i>pErrorMessage</i>	Sortie	Texte de message défini au moment de la mise en oeuvre. Si une erreur se produit, DB2 Extension Texte renvoie un code et un message d'erreur. Le programme d'application alloue la mémoire tampon de taille DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> est le pointeur NULL, aucun message d'erreur n'est renvoyé.

### Utilisation

Votre programme d'application doit établir une connexion à la base de données avant d'appeler DesGetBrowseInfo.

Pour le pointeur sur l'argument de recherche, il est préférable d'utiliser *char\** plutôt que *SQLCHAR\**, car il est possible que la valeur de ce paramètre ne provienne pas de la base de données.

Pour le mappage entre les types de données SQL et C, vous devez utiliser le nom symbolique *SQL\_VARBINARY* pour le descripteur. Le type de variables SQL pointant sur la représentation en C des valeurs de *Handle* est *SQLCHAR\**.

DB2 Extension Texte alloue de la mémoire pour les informations d'affichage. Le programme d'application doit ensuite libérer la mémoire et les ressources associées en appelant DesFreeBrowseInfo.

Etant donné que les valeurs de *Handle* sont exprimées en binaire et contiennent plusieurs caractères '\0', vous devez indiquer la longueur de *pHandle*.

L'argument de recherche dans *pSearchArgument* est décrit dans le «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

### Codes retour

```
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
```

## Fonction API DesGetBrowseInfo

RC\_SE\_COMMUNICATION\_PROBLEM  
RC\_SE\_EMPTY\_INDEX  
RC\_SE\_EMPTY\_QUERY  
RC\_SE\_FUNCTION\_DISABLED  
RC\_SE\_FUNCTION\_IN\_ERROR  
RC\_SE\_INCORRECT\_HANDLE  
RC\_SE\_INDEX\_DELETED  
RC\_SE\_INDEX\_NOT\_ACCESSIBLE  
RC\_SE\_INDEX\_SUSPENDED  
RC\_SE\_INSTALLATION\_PROBLEM  
RC\_SE\_IO\_PROBLEM  
RC\_SE\_MAX\_NUMBER\_OF\_BUSY\_INDEXES  
RC\_SE\_MAX\_OUTPUT\_SIZE\_EXCEEDED  
RC\_SE\_NOT\_ENOUGH\_MEMORY  
RC\_SE\_PROCESSING\_LIMIT\_EXCEEDED  
RC\_SE\_QUERY\_TOO\_COMPLEX  
RC\_SE\_SERVER\_BUSY  
RC\_SE\_SERVER\_CONNECTION\_LOST  
RC\_SE\_SERVER\_NOT\_AVAILABLE  
RC\_SE\_UNEXPECTED\_ERROR  
RC\_SE\_UNKNOWN\_INDEX\_NAME  
RC\_SE\_UNKNOWN\_SERVER\_NAME  
RC\_SE\_WRITE\_TO\_DISK\_ERROR

Attention : Les codes retour suivants indiquent que la fonction a renvoyé un résultat, mais ce n'est pas forcément celui prévu.

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

### Restrictions

Cette fonction ne peut être appelée qu'après la connexion à une base de données et l'utilisation d'une fonction de DB2 Extension Texte pour extraire un descripteur de cette base.

## DesGetMatches

### Objectif

Renvoie un flux de données contenant des informations de mise en évidence pour le document texte décrit par un descripteur de document. Reportez-vous à la section «Syntaxe du flux de données» à la page 238. Les informations de mise en évidence sont composées du contexte du texte (au moins un paragraphe) et d'informations pour la mise en évidence du texte dans ce contexte.

L'option `DesGetMatches` ne renvoie qu'une partie du flux de données, en indiquant la longueur de celle-ci dans la structure de sortie.

Une série d'appels de `DesGetMatches` permet d'extraire la totalité du contenu du document texte. Lorsque la fin du document est atteinte, le code retour `RC_SE_END_OF_INFORMATION` est renvoyé.

### Syntaxe

```

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

```

### Arguments de la fonction

Tableau 14. Arguments de `DesGetMatches`

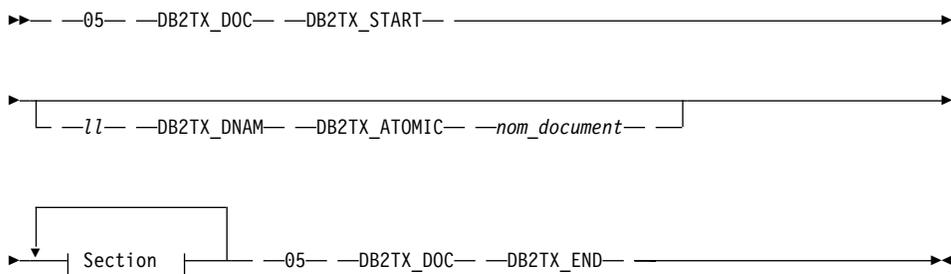
Type de données	Argument	Utilisation	Description
DESBROWSESESSION	<i>BrowseSession</i>	Entrée	Descripteur de la session d'affichage.
DESHANDLE	<i>DocumentHandle</i>	Entrée	Descripteur de document renvoyé par <code>DesOpenDocument</code> .
DESMATCHINFO *	<i>pMatchInfo</i>	Sortie	Pointeur sur une mémoire tampon contenant la partie de flux de données reçue. <code>DesGetMatches</code> alloue cette mémoire tampon.
DESULONG *	<i>pMatchInfoLength</i>	Sortie	Longueur de la partie de flux de données sur laquelle pointe <i>pMatchInfo</i> .

## Fonction API DesGetMatches

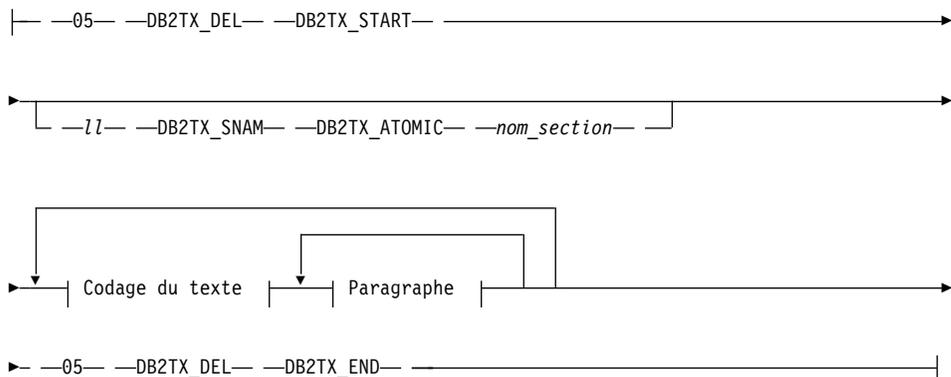
Tableau 14. Arguments de DesGetMatches (suite)

Type de données	Argument	Utilisation	Description
DESMESSAGE *	<i>pErrorMessage</i>	Sortie	Texte de message défini au moment de la mise en oeuvre. Si une erreur se produit, DB2 Extension Texte renvoie un code et un message d'erreur. Le programme d'application alloue la mémoire tampon de taille DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> est le pointeur NULL, aucun message d'erreur n'est renvoyé.

### Syntaxe du flux de données

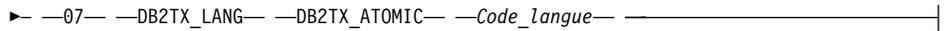


#### Section :

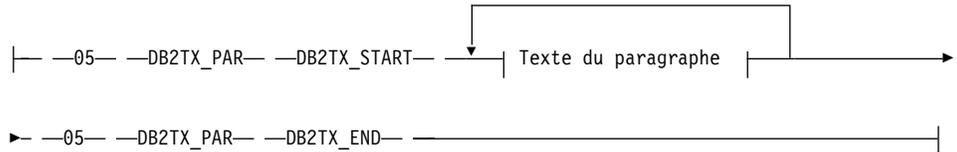


#### Codage du texte:

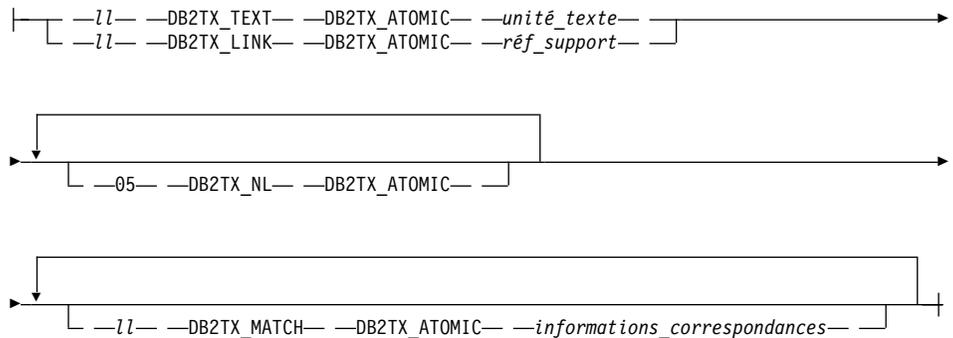




**Paragraphe :**



**Texte du paragraphe :**



Chaque segment du diagramme de syntaxe, tel que 05 DB2TX\_DOC DB2TX\_START, commence par une zone longueur de type INTEGER, qui, dans ce diagramme, est soit un nombre explicite tel que 05 ou une variable ll. La longueur du segment inclut la zone longueur sur 2 octets.

**Remarque :** La longueur est indiquée en format big-endian.

Chaque segment comprend l'un des identificateurs suivants sur 1 octet :

**DB2TX\_START**

Indique le début d'un segment, tel qu'un document ou un paragraphe.

**DB2TX\_END**

Indique la fin d'un segment.

**DB2TX\_ATOMIC**

Indique que l'élément qui suit est atomique, tel qu'un nom de document ou un identificateur de langue.

## Fonction API DesGetMatches

La longueur de chacun des éléments suivants du flux de données est de deux octets :

### DB2TX\_DOC

Indique le début ou la fin d'un document.

### DB2TX\_DNAM

Nom d'un document. Si aucun nom n'est indiqué, le système utilise l'identificateur du document.

### DB2TX\_DEL

Indique le début ou la fin d'un élément de document. Le seul type d'élément de document actuellement pris en charge est une *section de texte*.

### DB2TX\_SNAM

Indique le nom d'une section de texte. Actuellement, DB2 Extension Texte ne prend en charge qu'une section de texte et fournit automatiquement un nom par défaut. Si vous indiquez un nom de section, il n'est pas pris en compte.

### DB2TX\_PAR

Indique le début et la fin d'un paragraphe de texte se trouvant dans la section en cours.

### DB2TX\_TEXT

Indique une partie de texte se trouvant dans le paragraphe en cours. Habituellement, une *unité de texte* contient une ligne de texte et l'élément TEXT est suivi d'un élément DB2TX\_NL ; mais les lignes de texte peuvent également être divisées en plusieurs parties, chacune d'entre elles étant définie dans son propre élément DB2TX\_TEXT.

Le texte utilise le CCSID et la langue associés au paragraphe en cours.

### DB2TX\_LINK

Indique une référence hypermédia de DB2 Extension Texte. Il utilise le CCSID du paragraphe en cours.

### DB2TX\_NL

Indique le début d'une nouvelle ligne dans le paragraphe en cours.

### DB2TX\_MATCH

Contient des informations d'occurrence pour les correspondances dans la portion de texte en cours. Ces informations sont fournies sous la forme d'une séquence de paires de nombres binaires. Le premier nombre de chaque paire représente le décalage d'une correspondance dans la partie de texte en cours, le second représente la longueur, en caractères, de cette correspondance. La longueur donnée peut excéder la portion de texte donnée. Le décalage et la longueur sont des valeurs sur deux octets définies en format big-endian.

### DB2TX\_CCSID

CCSID pour le texte des paragraphes suivants jusqu'au prochain paragraphe précédé d'un nouvel élément DB2TX\_CCSID. Les CCSID suivants sont renvoyés :

#### DB2TX\_CCSID\_00500

pour du texte en Latin-1 EBCDIC page de codes 500.

#### DB2TX\_CCSID\_04946

pour du texte en Latin-1 ASCII page de codes 850.

#### DB2TX\_CCSID\_00819

pour du texte en ASCII page de codes 819.

Ces noms symboliques de CCSID sont définis dans le fichier DES\_EXT.H fourni avec DB2 Extension Texte. Les valeurs binaires sur deux octets sont indiquées en format big-endian.

### DB2TX\_LANG

Identificateur de langue pour le texte des paragraphes suivants jusqu'au prochain paragraphe précédé d'un nouvel élément DB2TX\_LANG. Le fichier DES\_EXT.H fourni avec DB2 Extension Texte définit les noms symboliques pour tous les identificateurs de langue pris en charge par DB2 Extension Texte. Les valeurs binaires sur deux octets sont indiquées en format big-endian.

## Utilisation

Lorsque la fin du document est atteinte, DesGetMatches renvoie le code retour RC\_SE\_END\_OF\_INFORMATION.

## Codes retour

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

Attention : Les codes retour suivants indiquent que la fonction a renvoyé un résultat, mais ce n'est pas forcément celui prévu.

RC\_SE\_DICTIONARY\_NOT\_FOUND

## Fonction API DesGetMatches

### Restrictions

Cette fonction ne peut être appelée qu'après l'ouverture d'un document texte par appel de DesOpenDocument.

## DesGetSearchResultTable

### Objectif

Utilise un argument de recherche pour effectuer une recherche dans des documents texte identifiés par une colonne de texte. Le contenu du descripteur correspondant aux éléments de texte trouvés est inscrit dans une table résultat. Les informations d'affichage concernant le rang et le nombre de correspondances peuvent aussi être inscrites dans la table résultat.

### Syntaxe

```

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);

```

### Arguments de la fonction

Tableau 15. Arguments de DesGetSearchResultTable

Type de données	Argument	Utilisation	Description
SQLHDBC	hdbc	Entrée	Descripteur de connexion à une base de données.
char *	pTableSchema	Entrée	Schéma de la table de base dans laquelle effectuer la recherche.
DESSMALLINT	TableSchemaLength	Entrée	Longueur de pTableSchema (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.
char *	pTableName	Entrée	Pointeur sur le nom de la table de base dans laquelle effectuer la recherche.
DESSMALLINT	TableNameLength	Entrée	Longueur de pTableName (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.

## Fonction API DesGetSearchResultTable

Tableau 15. Arguments de DesGetSearchResultTable (suite)

Type de données	Argument	Utilisation	Description
char *	pColumnName	Entrée	Pointeur sur le nom de la colonne concerné par la recherche de texte prévue. Cette colonne doit être de type DESTEXTH.
DESMALLINT	ColumnNameLength	Entrée	Longueur de pColumnName (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.
char *	pSearchArgument	Entrée	Pointeur sur l'argument de recherche de texte.
DESMALLINT	ArgumentLength	Entrée	Longueur de pSearchArgument (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.
char *	pResultSchema	Entrée	Pointeur sur le schéma contenant la table de résultats.
DESMALLINT	ResultSchemaLength	Entrée	Longueur de pSchemaName (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.
char *	pResultTableName	Entrée	Pointeur sur le nom de la table de résultats précédemment créée, dans laquelle sera stocké le résultat de la recherche. Pour connaître la structure de cette table, reportez-vous à la figure 17 à la page 245.
DESMALLINT	ResultTableNameLength	Entrée	Longueur de pResultTableName (sans compter le caractère de fin d'octet NULL), ou de DES_NTS.
DESSEARCHOPTION	SearchOption	Entrée	Option déterminant si vous demandez des informations relatives au rang, au nombre de correspondances ou uniquement aux descripteurs des documents texte correspondants. DES_RANK DES_MATCH DES_RANKANDMATCH DES_TEXTHANDLEONLY  Cette option détermine le contenu de la table résultat décrite à la section «Utilisation» à la page 245.
DESBROWSEOPTION	BrowseOption	Entrée	Réservé.

Tableau 15. Arguments de DesGetSearchResultTable (suite)

Type de données	Argument	Utilisation	Description
DESBROWSEINFO *	pBrowseInfo	Sortie	Pointeur sur des informations d'affichage ou sur une valeur NULL, en fonction de la valeur de BrowseOption.
DESMESSAGE *	pErrorMessage	Sortie	Texte de message défini au moment de la mise en oeuvre. Si une erreur se produit, DB2 Extension Texte renvoie un code et un message d'erreur. Le programme d'application alloue la mémoire tampon de taille DES_MAX_MESSAGE_LENGTH. Si pErrorMessage est le pointeur NULL, aucun message d'erreur n'est renvoyé.

### Utilisation

La connexion à la base de données doit être établie par le programme d'application qui appelle DesGetSearchResultTable.

Le nom *pResultTableName* fait référence à une table de résultats créée préalablement. L'utilitaire DESRESTB dans le répertoire SAMPLE crée une table de résultats pour les descripteurs de texte. Après l'appel de cette fonction, la table de résultats contient des informations identifiant des valeurs de texte correspondant à l'argument de recherche. La structure de la table de résultats est la suivante :

**TABLE DE RÉSULTATS**

TEXTHANDLE	RANK	MATCHES

Figure 17. Structure de la table de résultats

Le type de données de TEXTHANDLE est DB2TEXTH ou DB2TEXTFH. Le type de données de RANK est DOUBLE. Le type de données de MATCHES est INTEGER.

L'argument de recherche sur *pSearchArgument* est décrit dans le «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

Si la valeur de *BrowseOption* est *BROWSE*, DB2 Extension Texte renvoie des informations d'affichage provenant de son moteur de recherche implanté sur le serveur. *pBrowseInfo* pointe sur les informations d'affichage qui constituent l'entrée dans DesStartBrowseSession. Si la valeur de *BrowseOption* est *NO\_BROWSE*, *pBrowseInfo* pointe sur une valeur NULL.

## Fonction API DesGetSearchResultTable

### Codes retour

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

Attention : Les codes retour suivants indiquent que la fonction a renvoyé un résultat, mais ce n'est pas forcément celui prévu.

RC\_SE\_CONFLICT\_WITH\_INDEX\_TYPE  
RC\_SE\_DICTIONARY\_NOT\_FOUND

## Fonction API DesGetSearchResultTable

RC\_SE\_STOPWORD\_IGNORED  
RC\_SE\_UNKNOWN\_SECTION\_NAME  
RC\_SE\_DOCMOD\_READ\_PROBLEM

## Fonction API DesOpenDocument

### DesOpenDocument

#### Objectif

Reçoit un pointeur de session d'affichage, un descripteur et une option (DES\_EXTENDED ou DES\_FAST) indiquant si le document texte doit être analysé avec ou sans dictionnaire. Cette fonction prépare le document texte qui correspond au descripteur permettant d'extraire le texte du document et les informations de mise en évidence, et renvoie un descripteur de document utilisé pour l'appel itératif de DesGetMatches.

#### Syntaxe

```
DESRETURN  
DesOpenDocument  
( DESBROWSESESSION    BrowseSession,  
  SQLCHAR              *pHandle,  
  DESUSHORT            HandleLength,  
  DESMATCHMODE         MatchMode,  
  DESHANDLE            *pDocumentHandle,  
  DESMESSAGE           *pErrorMessage);
```

#### Arguments de la fonction

Tableau 16. Arguments de DesOpenDocument

Type de données	Argument	Utilisation	Description
DESBROWSESESSION	<i>BrowseSession</i>	Entrée	Descripteur de la session d'affichage.
SQLCHAR *	<i>pHandle</i>	Entrée	Pointeur sur un descripteur extrait de la base de données.
DESUSHORT	HandleLength	Entrée	Longueur de pHandle (DES_NTS ne peut pas être utilisé).
DESMATCHMODE	MatchMode	Entrée	Mode permettant de déterminer si les informations de mise en évidence sont trouvées à l'aide d'un dictionnaire.  <b>DES_FAST</b> Dictionnaire non utilisé  <b>DES_EXTENDED</b> Dictionnaire utilisé
DESHANDLE *	pDocumentHandle	Sortie	Descripteur de document pour l'appel itératif de DesGetMatches.

Tableau 16. Arguments de DesOpenDocument (suite)

Type de données	Argument	Utilisation	Description
DESMESSAGE *	<i>pErrorMessage</i>	Sortie	Texte de message défini au moment de la mise en oeuvre. Si une erreur se produit, DB2 Extension Texte renvoie un code et un message d'erreur. Le programme d'application alloue la mémoire tampon de taille DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> est le pointeur NULL, aucun message d'erreur n'est renvoyé.

### Utilisation

DES\_FAST et DES\_EXTENDED renvoient à l'utilisation d'un traitement linguistique pour déterminer les termes à mettre en évidence dans le texte affiché. Pour plus d'informations, reportez-vous à la section «Analyse linguistique pour l'affichage» à la page 266. Indiquez DES\_FAST pour utiliser une analyse de texte de base, et DES\_EXTENDED pour utiliser la procédure de mise en correspondance étendue.

Pour le mappage entre les types de données SQL et C, vous devez utiliser le nom symbolique SQL\_VARBINARY pour le descripteur. Le type de variables SQL pointant sur la représentation en C des valeurs de *TextHandle* est SQLCHAR\*.

DB2 Extension Texte alloue de la mémoire pour les informations d'affichage. Le programme d'application doit ensuite libérer cette mémoire et les ressources associées en appelant *DesFreeBrowseInfo*.

Etant donné que les valeurs de *TextHandle* sont exprimées en binaire et contiennent plusieurs caractères '\0', vous devez indiquer la longueur de *pHandle*.

Le demandeur doit avoir un accès en lecture à la table contenant le document texte auquel *pHandle* fait référence.

### Codes retour

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

## Fonction API DesOpenDocument

RC\_SE\_INCORRECT\_HANDLE  
RC\_SE\_IO\_PROBLEM  
RC\_SE\_LS\_FUNCTION\_FAILED  
RC\_SE\_LS\_NOT\_EXECUTABLE  
RC\_SE\_MAX\_NUMBER\_OF\_BUSY\_INDEXES  
RC\_SE\_NOT\_ENOUGH\_MEMORY  
RC\_SE\_REQUEST\_IN\_PROGRESS  
RC\_SE\_UNKNOWN\_INDEX\_NAME  
RC\_SE\_UNEXPECTED\_ERROR

### Restrictions

Cette fonction ne peut être appelée qu'après le démarrage d'une session d'affichage par appel de DesStartBrowseSession.

## DesStartBrowseSession

### Objectif

Démarre une session d’affichage en établissant l’environnement nécessaire à l’affichage d’un document texte et à la mise en évidence des correspondances trouvées. Elle reçoit un pointeur sur les informations d’affichage, envoyé par DesGetBrowseInfo ou DesGetSearchResultTable, et renvoie un descripteur de session d’affichage qu’utiliseront les autres fonctions d’affichage.

### Syntaxe

```

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

### Arguments de la fonction

Tableau 17. Arguments de DesStartBrowseSession

Type de données	Argument	Utilisation	Description
DESBROWSEINFO	<i>BrowseInfo</i>	Entrée	Pointeur sur des informations à l’affichage et à la mise en évidence des correspondances dans un document texte. Ce pointeur est renvoyé par DesGetSearchResultTable ou DesGetBrowseInfo.
char *	<i>pUserId</i>	Entrée	ID utilisateur pour la base de données.
DESSMALLINT	<i>UserIdLength</i>	Entrée	Longueur de l’ID utilisateur pour la base de données.
char *	<i>pPassword</i>	Entrée	Mot de passe pour la base de données.
DESSMALLINT	<i>PasswordLength</i>	Entrée	Longueur du mot de passe pour la base de données.
DESBROWSESESSION *	<i>pBrowseSession</i>	Sortie	Descripteur de session d’affichage qu’utiliseront les autres fonctions d’affichage.

## Fonction API DesStartBrowseSession

Tableau 17. Arguments de DesStartBrowseSession (suite)

Type de données	Argument	Utilisation	Description
DESMESSAGE *	<i>pErrorMessage</i>	Sortie	Texte de message défini au moment de la mise en oeuvre. Si une erreur se produit, DB2 Extension Texte renvoie un code et un message d'erreur. Le programme d'application alloue la mémoire tampon de taille DES_MAX_MESSAGE_LENGTH. Si <i>pErrorMessage</i> est le pointeur NULL, aucun message d'erreur n'est renvoyé.

### Utilisation

Cette fonction ouvre une session d'affichage de documents texte. Le système vous invite à entrer votre ID utilisateur et votre mot de passe pour vérifier votre droit d'accès à la base de données.

Pour fermer la session d'affichage, appelez DesEndBrowseSession.

*BrowseInfo* dépend de l'argument de recherche et de la colonne de texte de base utilisés pour constituer les informations d'affichage.

### Codes retour

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

### Restrictions

Vous devez appeler DesGetBrowseInfo ou DesGetSearchResultTable avec l'option d'affichage appropriée avant d'appeler cette fonction.

---

## Chapitre 14. Modèle de programme API

DB2 Extension Texte fournit un modèle de programme, DESSAMP1.C, situé dans le répertoire SAMPLES.

DESSAMP1.C est un exemple de programme qui utilise la fonction DesGetSearchResultTable ainsi que votre propre programme d'affichage. Il suit la séquence d'appels de fonctions API présentée dans la figure 16 à la page 121.

Vous devez accéder à une base de données et à une colonne de texte actives. Pour exécuter ce programme, procédez comme suit :

1. Facultatif. Copiez le fichier source DESSAMP1.C dans un répertoire local de votre poste client.
2. Utilisez le fichier Makefile fourni (dessamp.mak pour OS/2, Windows NT et Windows 2000, dessamp pour les systèmes UNIX) pour compiler et lier les fichiers exemples.
3. Lancez l'utilitaire DESRESTB pour créer une table de résultats dans la base de données que vous voulez utiliser avec le modèle de programme :

DESRESTB *nom\_db*

Cette table sert au stockage d'informations, telles que les résultats de la recherche. Elle présente la structure suivante :

Colonne	Type de données
HANDLE	DB2TX.DB2TEXTH ou DB2TX.DB2TEXTFH
RANK	DOUBLE
MATCHES	INTEGER

## Modèle de programme API

---

## Chapitre 15. Traitement linguistique des index linguistiques et exacts

DB2 Extension Texte permet une analyse linguistique dans les domaines suivants :

- **Indexation.** Lorsque DB2 Extension Texte analyse des documents pour en extraire des termes devant être stockés dans l'index de texte, le texte subit une analyse linguistique visant à extraire les termes appropriés pour l'index. Cela permet d'obtenir une extraction aussi simple et rapide que possible.
- **Extraction.** Lorsque DB2 Extension Texte effectue des recherches dans un index afin d'y trouver les noms des documents contenant des occurrences des termes recherchés, ces derniers subissent également une analyse linguistique afin de correspondre aux termes indexés.
- **Affichage.** Lorsque vous affichez un document trouvé à la suite d'une recherche, l'analyse linguistique est utilisée pour mettre en évidence les termes trouvés dans le document.

---

### Analyse linguistique lors de l'indexation

Lorsque DB2 Extension Texte indexe et extrait des documents, elle effectue une analyse linguistique du texte. Le tableau suivant permet de voir que l'étendue de l'analyse linguistique dépend du type d'index. Pour les index Ngram, aucune analyse linguistique n'est appliquée.

L'analyse linguistique utilisée pour l'indexation de documents se compose des opérations suivantes :

- Analyse de texte de base
  - Reconnaissance de termes (marquage sémantique)
  - Normalisation des termes au format standard
  - Reconnaissance de phrases
- Réduction des termes à leur forme de base
- Filtrage des mots vides
- Décomposition (éclatement des termes composés).

Le tableau 18 à la page 256, illustre la façon dont les termes sont indexés dans un index de type **linguistique** et lorsqu'aucune autre caractéristique d'index n'a été demandée.

## Analyse linguistique lors de l'indexation

Tableau 18. Extraction des termes pour un index linguistique

Texte du document	Terme dans l'index	Analyse linguistique
Cheval Käfer	cheval kaefer	Analyse de texte de base (normalisation)
chevaux acheté	cheval acheter	Réduction à la forme de base
orienté-objets  Wetterbericht	orienté-objets, orientation wetterbericht, wetter bericht	Décomposition
un rapport sur les animaux	rapport animal	Filtrage des mots vides. Les mots vides sont un et sur.

Par comparaison, le tableau 19 illustre la façon dont les termes sont indexés dans un index de type **exact**.

Tableau 19. Extraction des termes pour un index exact

Texte du document	Terme dans l'index	Analyse linguistique
Cheval Käfer	Cheval Käfer	Pas de normalisation
chevaux acheté	chevaux acheté	Pas de réduction à la forme de base
un rapport sur les animaux	rapport animaux	Filtrage des mots vides. Les mots vides sont un et sur.
orienté-objets Wetterbericht	orienté-objets Wetterbericht	Pas de décomposition

### Analyse de texte de base

DB2 Extension Texte analyse le texte de base sans utiliser le dictionnaire électronique.

### Reconnaissance des termes contenant des caractères non alphanumériques

Lorsque des documents sont indexés, les termes sont reconnus même s'ils contiennent des caractères non alphanumériques, par exemple : "14225,23 ", "belle-mère" et "22/10/95".

Les éléments suivants sont considérés comme des parties de termes :

- Accents

- Signes monétaires

- Caractères de séparation de chiffres (comme "/" ou ".")

- Le caractère "@" des adresses électroniques (anglais seulement)

Le signe “+”.

Les règles spécifiques à chaque langue sont également utilisées pour reconnaître les termes contenant :

- Des préfixes avec apostrophe dans les langues latines, tels que l' *aventure* en Français.
- Des formats nationaux pour les dates, les heures et les nombres.
- Les alternatives, telles que *mission/responsabilité*, indiquées en français par le caractère “/”.
- Les apostrophes de fin en Italien comme dans *securita'*. Il est fréquent, en italien écrit, lorsque le jeu de caractères ne comprend pas les caractères accentués, de taper les accents *après* le caractère. Par exemple, “à” est entré sous la forme “a”.

### Normalisation des termes au format standard

La normalisation réduit les termes contenant des majuscules et des minuscules, ou ceux contenant des caractères accentués ou spéciaux, à un format standard. Elle est effectuée par défaut si l'index est de type linguistique. (Dans un index exact, le format des lettres est laissé tel quel car la recherche tient compte des majuscules et des minuscules).

Par exemple, le terme *Ordinateur* est indexé sous sa forme *ordinateur*. La lettre majuscule est convertie en minuscule. La recherche du terme *ordinateur* aboutit aux occurrences *ordinateur* et *Ordinateur*. La normalisation durant l'indexation a pour effet d'indexer les termes de façon identique, qu'ils contiennent des majuscules ou pas.

La normalisation est appliquée non seulement lors de l'indexation, mais aussi lors de l'extraction. Les majuscules d'un terme recherché sont transformées en minuscules avant l'exécution de la recherche. Lorsque le terme recherché est, par exemple, *Ordinateur*, le terme utilisé pour la recherche est *ordinateur*.

Les termes accentués et spéciaux sont normalisés de la même manière. Toute variante de *école*, telle que *École*, aboutit aux termes *école*, *Ecole*, etc. En allemand, le terme *Bürger* aboutit à *buerger*, et le terme *Maße* conduit au terme *masse*.

Si le terme recherché comprend des caractères génériques, la normalisation est faite avant le traitement de ces caractères. Exemple: *Bür\_er* devient *buer\_er*.

### Reconnaissance de phrases

Vous pouvez rechercher des termes figurant dans une même phrase. Pour que cela soit possible, chaque document est analysé durant l'indexation pour rechercher la fin de chaque phrase.

## Analyse linguistique lors de l'indexation

DB2 Extension Texte permet d'identifier la fin de chaque phrase par deux méthodes :

- Le marqueur sémantique Universal Unicode Tokenizer pour les langues autres l'arabe et l'hébreu.

Il s'agit de la méthode la plus simple et la plus rapide. Le marqueur sémantique recherche un point ou un point d'exclamation ou d'interrogation, précédé d'un caractère, un lettre par exemple, et suivi d'un espace, d'une tabulation ou d'un caractère nouvelle ligne. Pour vérifier qu'il s'agit bien de la fin d'une phrase, et non d'une abréviation se terminant par un point, une comparaison est établie avec une liste d'abréviations propres à chaque langue.

- Le marqueur sémantique basé sur POE pour l'arabe et l'hébreu

Ce marqueur sémantique est plus évolué du point de vue linguistique mais exige plus de puissance de traitement. Le marqueur sémantique détermine la fin des phrases essentiellement par concordance de ponctuation, mais également par réception de signaux provenant de types d'entrées particuliers et en fonction du nombre de mots.

### Marqueur sémantique basé sur POE pour l'identification de phrases

Le marqueur sémantique basé sur POE détermine les limites d'une phrase (ou d'un fragment de phrase) par les règles de ponctuation et un traitement propre à la langue notamment pour les abréviations, bien que le niveau de fonction varie considérablement selon la langue considérée. A la plupart des langues qui utilisent des pages de codes à un seul octet est associé un dictionnaire complémentaire d'abréviations (Abbreviation Addenda Dictionary) fourni avec POE. Du fait que les langues double octet n'utilisent généralement pas des abréviations avec points, les dictionnaires complémentaires d'abréviations n'existent pas pour ces langues.

La fin d'une phrase est donc essentiellement déterminée par concordance de la ponctuation. Le tableau suivant indique les caractères de ponctuation de fin de phrase ou de segment de phrase ainsi que les identificateurs GCGID correspondants.

GCGID de caractères SBCS	GCGID de caractères DBCS	Description
SP110000	SP110080	Point
SP020000	SP020080	Point d'exclamation
SP150000	SP150080	Point d'interrogation
SP140000	N/A	Point-virgule (point d'interrogation grec)
N/A	JQ730080	Point double octet

## Analyse linguistique lors de l'indexation

Un caractère de ponctuation de fin, un point, un point d'exclamation ou d'interrogation, par exemple, est censé signaler la fin d'une phrase, sauf dans l'un des cas suivants :

- Le caractère de fin est suivi d'un des caractères de ponctuation fermants indiqués dans le tableau ci-après, une fermeture de parenthèse, par exemple.

GCGID de caractères SBCS	GCGID de caractères DBCS	Description
SP070000	SP070080	Parenthèse fermante
SP040000	SP040080	Guillemets
SP050000	SP050080	Apostrophe
SP180000	SP070083	Guillemets angulaires
N/A	SM140080	Closing brace (Non utilisé en français)
N/A	SM080080	Closing bracket (Non utilisé en français)
N/A	JQ720080	Single square quote (Non utilisé en français)
N/A	JQ720081	Double square quote (Non utilisé en français)
N/A	SP200080	Closing single hook quote (Non utilisé en français)
N/A	SP220080	Closing double hook quote (Non utilisé en français)
N/A	SP070081	Closing carapace bracket (Non utilisé en français)
N/A	SP070082	Closing single angle quote (Non utilisé en français)
N/A	SP070084	Closing cornered parenthesis (Non utilisé en français)
N/A	SP370080	Vertical closing single square quote (Non utilisé en français)
N/A	SP370081	Vertical closing double square quote (Non utilisé en français)
N/A	SP250084	Vertical closing squared parenthesis (Non utilisé en français)

## Analyse linguistique lors de l'indexation

GCGID de caractères SBCS	GCGID de caractères DBCS	Description
N/A	SP250080	Vertical closing parenthesis (Non utilisé en français)
N/A	SP350080	Vertical closing brace (Non utilisé en français)
N/A	SP250081	Vertical closing carapace (Non utilisé en français)
N/A	SP250083	Vertical closing double angled quote (Non utilisé en français)
N/A	SP250082	Vertical closing single angled quote (Non utilisé en français)

**Remarque :** POE ne considère pas les éléments repérés par N/A comme des caractères de ponctuation fermants et les caractères de ponctuation fermants verticaux ne sont pris en charge que pour le chinois.

Exemple :

...cette phrase se termine par deux parenthèses.)

Dans l'exemple, la seconde parenthèse est détectée en tant que fin de la phrase. En allemand toutefois, une apostrophe fermante ne dénote pas la fin d'une phrase lorsqu'elle est suivie d'une virgule.

- Le caractère de ponctuation de fin est suivi d'un autre. Exemple :

Ceci est une forte exclamation !!!

Le dernier point d'exclamation dénote la fin de la phrase.

- Le caractère de ponctuation fermant est précédé d'un chiffre, soit d'un caractère de ponctuation et suivi d'un caractère numérique. Ceci interdit à des chaînes de caractères telles que '1.25' et '.314' de mettre fin à une phrase.
- Le caractère de ponctuation fermant est un point et fait partie d'une abréviation non autorisée à la fin d'une phrase. Un traitement limité des abréviations est effectué pour chaque langue.
- Le caractère de ponctuation de fin est un point et n'est pas suivi d'un blanc, par exemple un espace ou un caractère nouvelle ligne. Ceci permet d'éviter que des en-têtes, tels que 'III.IV' soit détectés comme fin d'une phrase.

Le marqueur sémantique basé sur POE traite également les abréviations, afin de déterminer si un point fait partie d'une abréviation ou s'il dénote la fin d'une phrase. Vous pouvez ajouter des abréviations à un dictionnaire

supplémentaire des abréviations. Si aucun dictionnaire n'est transmis au marqueur sémantique basé sur POE, toutes les lettres isolées suivies d'un point sont repérées en tant qu'abréviations ; aucun autre traitement d'abréviation n'a lieu.

Le fait qu'un élément de texte soit ou non une abréviation est souvent ambigu il arrive qu'une abréviation soit interprétée comme un mot normal suivi d'un point. Prenons, par exemple, les caractères "no." dans les phrases suivantes en anglais

Enter the no. of exemptions you are claiming.  
Answer each question with yes or no.

Même quand un élément de texte est identifié en tant qu'abréviation, la question se pose toujours de savoir s'il met fin à la phrase. Certaines abréviations ne se placent jamais en fin de phrase, alors que, pour d'autres, cela est possible. Prenons, par exemple, l'abréviation "Hwy." (Highway, à savoir Autoroute en anglais) dans les phrases suivantes

The drive along Hwy. 1 to Santa Cruz was beautiful.  
Many people speak highly of the Pacific Coast Hwy.

Du fait que les abréviations puissent être ambiguës et que certaines ne puissent pas se situer en fin de phrase, POE tente de classer les abréviations détectées. Si un point fait partie d'une abréviation qui peut parfois se situer en fin de phrase, le traitement se poursuit. Si POE détermine que l'abréviation ne se trouve pas à la fin de la phrase, la marque représentant le point est jointe à celle correspondant au texte de l'abréviation. Dans tout autre cas, la marque représentant le point reste distincte.

Le traitement des abréviations basé sur POE met en oeuvre trois jeux de critères pour déterminer si un point fait partie d'une abréviation :

- Toute lettre isolée suivie d'un point est considérée comme étant une abréviation. Les abréviations en une seule lettre sont classées comme pouvant constituer une fin de phrase.
- Les mots du dictionnaire supplémentaire d'abréviations fourni en entrée sont toujours considérés comme des abréviations. L'information associée à une abréviation dans le supplément détermine si cette abréviation peut se situer en fin de phrase. in the addenda. Par exemple, "Mr." est repéré dans le dictionnaire supplémentaire des abréviations comme ne pouvant pas se situer en fin de phrase, alors que "etc." est une abréviation que l'on peut parfois retrouver en fin de phrase.
- Tout mot en deux à six caractères suivi d'un point qui n'a été localisé dans aucun dictionnaire d'entrée, ni dans aucun dictionnaire supplémentaire d'abréviations, est également considéré comme abréviation. Ceci permet de prendre en compte des mots tels que "Jrnl. Comp. Ling.". Les abréviations déterminées par recherche dans un dictionnaire sont toujours considérées comme pouvant se situer en fin de phrase.

## Analyse linguistique lors de l'indexation

Si une abréviation est considérée comme pouvant se situer en fin de phrase, POE examine le texte qui la suit pour vérifier si c'est le cas ; à cet effet, il contrôle si le mot suivant commence par une majuscule.

Si une abréviation est suivie de plusieurs éléments de données dénotant une nouvelle ligne, une nouvelle phrase ou un nouveau paragraphe, POE suppose une fin de phrase. De plus, si un point d'interrogation ou d'exclamation inversé constitue le texte suivant, une marque de fin de phrase est insérée dans la sortie.

Si POE détermine que le point fait partie d'une abréviation qui n'est pas une fin de phrase, il poursuit la recherche d'un délimiteur de phrase. Sinon, il vérifie d'autres exceptions de caractère de ponctuation de fin (autre ponctuation de fin ou fermante) avant d'indiquer une fin de phrase.

### Réduction des termes à leur forme de base (lemmisation)

Dans un index linguistique, vous pouvez rechercher le terme `cheval`, par exemple, et trouver `chevaux`. Les termes sont réduits à leur forme de base pour l'indexation ; le terme `chevaux` est indexé sous `cheval`. Par la suite, lorsque vous rechercherez le terme `cheval`, vous trouvez le document. Vous le trouverez également si vous utilisez le terme `chevaux`.

Vous pouvez donc trouver les documents contenant des informations sur les chevaux, quelle que soit la variante du terme `cheval` utilisée dans le document ou comme terme de recherche.

De même, les verbes conjugués sont réduits à leur forme infinitive. Ainsi, `acheté` devient `acheter`.

### Filtrage des mots vides

Les mots vides sont des mots tels que les prépositions ou les pronoms qui reviennent très fréquemment dans les documents et ne sont donc pas exploitables pour une recherche. Ces mots figurent dans une liste associée à chaque dictionnaire et sont exclus du processus d'indexation.

Le filtrage des mots vides ne fait pas de différence entre les majuscules et les minuscules. Ainsi, le mot vide `lorsque` exclut également le mot `Lorsque` placé en début de phrase. Les listes de mots vides sont fournies en plusieurs langues et elles peuvent être modifiées.

Un index Ngram ne comporte pas de liste de mots vides.

### Décomposition (éclatement des termes composés)

Les langues germaniques, telles que l'allemand ou le néerlandais, sont riches en termes composés comme `Versandetiketten`, qui signifie étiquettes (Etiketten) courrier (Versand). De tels termes composés peuvent être éclatés.

## Analyse linguistique lors de l'indexation

Dans le cas d'un index exact, les termes composés sont indexés en un seul mot. Par contre, dans un index linguistique, les termes composés sont décomposés lors de l'indexation. Lorsque vous effectuez une recherche, les termes composés sont décomposés si vous utilisez un index linguistique.

Les composants du mot composé sont retrouvés s'ils figurent dans une même phrase. Par exemple, lorsque vous recherchez le mot allemand *Wetterbericht* (bulletin météo), un document contenant l'expression *Bericht über das Wetter* (bulletin concernant la météo) sera également sélectionné comme résultat de recherche.

Un terme est décomposé :

- si la langue dans laquelle il est exprimé utilise des termes composés ;
- s'il possède une longueur minimale ;
- si le terme n'est pas lui-même une entrée du dictionnaire électronique—les mots composés couramment utilisés comme le mot allemand *Geschäftsbericht* (rapport commercial) *figurent* dans le dictionnaire allemand.

Lorsqu'une décomposition est possible, les composants du terme sont réduits à leur forme de base. Voici quelques exemples en danois, allemand et néerlandais :

Terme composé	Composants
børsmæglerselskab	børsmæglerselskab børs mægler selskab
Kindersprachen	kindersprache kind sprache
probleemkinderen	probleemkinderen probleemkind kind probleem

---

### Analyse linguistique pour l'extraction

L'analyse linguistique a pour objectif d'affaiblir des termes de recherche, pour augmenter le taux de rappel des recherches, c'est-à-dire le nombre de documents correspondants trouvés. Pour atteindre cet objectif, les termes de la recherche subissent l'une des deux opérations suivantes, ou les deux : extensions et réductions. En outre, certaines opérations de recherche de termes impliquent à la fois expansion et réduction.

## Analyse linguistique pour l'extraction

- Les extensions prennent, dans un terme de recherche, un mot ou un terme composé de plusieurs mots et l'associent à un ensemble de termes de recherche de remplacement, chacun pouvant être lui-même un terme de plusieurs mots. L'expression source et l'ensemble d'expressions cible constituent une expression booléenne de type OR dans le langage de requête de DB2 Extension Texte. Comme les extensions laissent le terme source inchangé, elles sont, jusqu'à un certain point, indépendantes du type d'index. Les extensions peuvent être de plusieurs sortes :

Extension de synonymes

Extension de thésaurus

- Les réductions ramènent le terme de la recherche à une forme plus générale que celle indiquée par l'utilisateur. Comme elles modifient le terme de recherche, les réductions dépendent du type d'index, afin qu'il existe une correspondance avec le terme modifié. Par conséquent, DB2 Extension Texte dérive les informations qui concernent la réduction du ou des types d'index qui serviront pour la recherche. Voici quelques exemples de réductions :

Réduction à une forme de base (reportez-vous à la section «Réduction des termes à leur forme de base (lemmisation)» à la page 262)

Normalisation (reportez-vous à la section «Normalisation des termes au format standard» à la page 257).

Mots vides (reportez-vous à la section «Filtrage des mots vides» à la page 262).

- Certaines opérations modifient le terme de recherche et créent une extension à partir d'un ensemble de termes de remplacement. La réduction inhérente fait que ces termes de remplacement dépendent eux aussi des informations de l'index. Voici les opérations qui entrent dans cette catégorie :

Masquage de caractères et de mots

Extension phonétique

### Synonymes

Les synonymes sont des mots proches par leur sémantique. Généralement, ils appartiennent à la même catégorie de mots (comme la catégorie des noms, des verbes, etc.) que le terme source. Les synonymes se trouvent dans un fichier différent selon la langue concernée. Ils sont toujours renvoyés sous leur forme de base et, à quelques exceptions près, ne sont pas composés de plusieurs mots. Lors d'une recherche de synonymes, les mots des termes de recherche sont toujours réduits à leur forme de base. Voici quelques exemples de synonymes d'un mot en trois langues :

- Anglais

*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

- Français

*Mot :*

expression parole terme vocable lettre billet missive épître  
plaisanterie

- Allemand

*Wort:*

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

### Extension de thésaurus

Un terme de recherche peut être étendu à l'aide des termes d'un thésaurus, accessible via une relation spécifique. Il peut s'agir d'une relation hiérarchique (comme une relation de type "Terme plus restreint"), associative (comme une relation de type "Terme associé") ou synonymique. Un terme du thésaurus est souvent constitué de plusieurs mots.

Pour plus de détails, reportez-vous à la section «Thésaurus» à la page 268.

Le terme de recherche (terme initial) n'est pas normalisé lors de la recherche dans le thésaurus. Les mots obtenus suite à la recherche dans le thésaurus sont ramenés à leur forme de base en fonction du type d'index.

### Extension phonétique

L'extension phonétique permet de constituer à partir d'un même mot un ensemble de mots phonétiquement ressemblants. Elle est particulièrement utile si vous ne connaissez pas l'orthographe exacte du terme à rechercher.

### Masquage de caractères et de mots

Le masquage (utilisation de caractères génériques) est une technique d'extension non linguistique, selon laquelle une expression normale est décomposée et étendue à tous les mots indexés correspondants. Ni une expression masquée ni ses extensions ne font l'objet d'une réduction à leur forme de base, d'une extraction de mots vides ou d'une quelconque autre technique d'extension. Le résultat est que, dans le cas d'un mot comme le verbe conjugué vient, quand la recherche porte sur vi\*, ce terme est comparé à un index exact et non à un index linguistique dans lequel le verbe a fait l'objet d'une réduction à sa forme de base pour devenir venir.

Lorsque vous utilisez un masquage de mots, les opérations peuvent être lentes, spécialement dans le cas de recherche sur des index importants.

### Analyse linguistique pour l'affichage

Une analyse linguistique est également effectuée lorsque vous affichez des documents trouvés lors d'une recherche. Elle s'effectue en deux étapes :

1. Analyse de texte de base : normalisation et extension des termes
2. Correspondance étendue

#### Etape 1 : Normalisation et extension des termes

La première étape s'effectue sans utilisation du dictionnaire électronique.

##### Normalisation

Le processus de normalisation est décrit à la section «Analyse de texte de base» à la page 256.

##### Extension des termes

L'extension de terme est l'opération inverse de la réduction d'un terme à sa forme de base. Lorsque l'index est de type linguistique, les termes de recherche sont alors réduits à leur forme de base avant le début de la recherche.

De même, si vous utilisez un index linguistique, les termes d'un document sont réduits à leur forme de base avant d'être ajoutés à l'index. Les documents sont donc recherchés en fonction de la forme de base des termes.

Lorsque vous affichez un document trouvé à la suite d'une recherche, vous souhaitez généralement la mise en évidence de toutes les variantes du terme de base. Pour cela, il faut que ce dernier soit étendu.

Toutes les variantes d'un terme figurant dans les dictionnaires peuvent être générées. Voici quelques variantes du mot aller :

vais va allez irai ira irez irais allés allées  
vas allons vont iras irons iront allé aller allât

#### Etape 2 : Correspondance étendue

La seconde étape est l'opération de correspondance étendue, qui peut être utilisée dans les rares cas où l'analyse de texte de base et la normalisation ne permettent pas de mettre en évidence un terme trouvé. Cette fonction permet d'extraire les correspondances les moins évidentes.

Pour activer la fonction de correspondance étendue, indiquez le paramètre `DES_EXTENDED` dans la fonction API `DesOpenDocument`.

La correspondance étendue fait intervenir la même analyse linguistique que celle utilisée lors de l'indexation linguistique.

Voici les cas dans lesquels la fonction de correspondance étendue peut permettre de trouver des correspondances supplémentaires :

- Le terme recherché comprend des caractères de masquage et est une variante.

Les caractères de masquage sont analysés et une réduction à la forme de base est faite sur le terme recherché. Les documents correspondants sont trouvés. Sans utilisation de la fonction de correspondance étendue, certains textes correspondant au critère de recherche précisé ne seraient pas mis en évidence.

Exemple : Un document contient la variante vient.

- Lors de l'indexation, le terme est réduit à la forme venir.
  - Si le terme de recherche est veni%, le document est trouvé car la réduction à la forme de base est venir.
  - Sans utilisation de la fonction de correspondance étendue, seuls les mots correspondant au terme veni% sont mis en évidence. Si cette fonction est activée, la variante vient est également mise en évidence.
- Des mots composés ont été indexés.

Lorsqu'un document en langue allemande contient un mot composé et est indexé à l'aide d'un index linguistique, l'index du document prend en compte les parties composant le mot et le mot composé lui-même. Lorsque vous recherchez une partie de mot composé, les documents contenant le mot composé sont obtenus, mais sans la fonction de correspondance étendue, le mot n'est pas mis en évidence.

Exemple : Un document en allemand contient le mot Apfelbaum (pommier).

- Lors de l'indexation linguistique, le terme est réduit sous les formes apfel et baum.
  - Lorsque le terme baum est recherché dans l'index, le terme Baum et le document dans lequel il apparaît sont obtenus comme résultat de recherche.
  - Sans la fonction de correspondance étendue, aucun terme n'est mis en évidence car le document contient Apfelbaum, mais pas Baum. Avec la fonction de correspondance étendue, le mot composé Apfelbaum est décomposé et la partie Baum est mise en évidence.
- Les mots sont coupés par un tiret en fin de ligne.

Si le tiret est inséré automatiquement par un traitement de texte, le mot le contenant est trouvé et mis en évidence. Par contre, si le trait d'union a été entré par l'utilisateur, les documents contenant le mot sont trouvés, mais, sans la fonction de correspondance étendue, le mot n'est pas mis en évidence.

## Analyse linguistique pour l’affichage

Exemple : Un document contient le mot conteneur coupé en fin de ligne de la manière suivante :

Les produits se trouvaient dans un con-  
teneur.

- Durant l’indexation, le mot est normalisé en conteneur.
- Lorsque le terme conteneur est recherché dans l’index, le terme et le document dans lequel il apparaît sont obtenus comme résultat de recherche.
- Une tentative est faite pour mettre en évidence les mots correspondant à conteneur. Sans la fonction de correspondance étendue, une correspondance n’est retenue que si le tiret de con-teneur a été inséré par le traitement de texte et non par l’utilisateur.

---

## Thésaurus

Un thésaurus regroupe des termes voisins par leur sémantique et traitant d’un sujet spécifique. Il peut être visualisé comme un réseau sémantique dans lequel chaque terme est représenté par un noeud. Si deux termes sont en relation, les noeuds correspondants sont reliés par une liaison étiquetée du nom de la relation. On peut accéder à tous les termes en relation directe avec un terme donné en suivant toutes les connexions qui partent du noeud. On peut accéder par étapes aux termes reliés indirectement en suivant toutes les connexions qui partent des noeuds atteints à une étape antérieure. La figure 18 à la page 269 présente la structure d’un petit thésaurus.

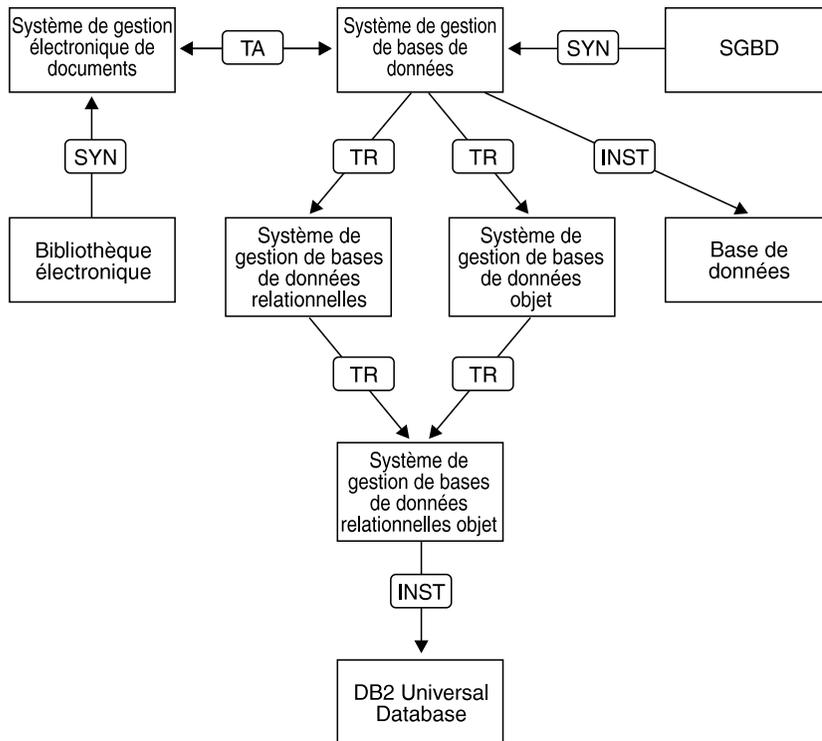


Figure 18. Thésaurus représenté sous la forme d'un réseau

Avec DB2 Extension Texte, il est possible de compléter une extension par des termes supplémentaires provenant d'un thésaurus déjà créé. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215 pour savoir comment utiliser l'extension de thésaurus dans une requête.

Pour créer un thésaurus utilisable dans une application de recherche, il est nécessaire de disposer d'un fichier de définition de thésaurus compilé en format interne, le dictionnaire du thésaurus.

Le format de dictionnaire utilisé par un index linguistique et un index exact diffère de celui utilisé par un index Ngram. Deux compilateurs de thésaurus différents sont donc fournis avec le produit. Leurs concepts de base diffèrent légèrement mais, de plus, les formats sources utilisés sont différents. Il vous faut donc décider quel type d'index vous utiliserez avant de commencer à définir les thésaurus de votre application de recherche.

Les composants de base d'un thésaurus sont les "termes" et les "relations".

# Thésaurus

## Termes

Un terme est un mot ou une expression désignant un concept appartenant au sujet auquel le thésaurus est consacré. Par exemple, les termes suivants pourraient faire partie d'un ou plusieurs thésaurus :

Traitement des données  
Hélicoptère  
Produit national brut

Dans un thésaurus DB2 Extension Texte, il y a deux catégories de termes : les descripteurs et les non descripteurs. Un *descripteur* est, dans un groupe de synonymes, le terme qui est choisi pour l'indexation et la recherche. Les autres termes du groupe sont appelés *non descripteurs*. Par exemple, contour et forme sont synonymes, et forme peut être choisi comme descripteur et contour comme non descripteur.

Un thésaurus Ngram ne fait pas la différence entre les descripteurs et les non descripteurs.

## Relations

Une relation est l'expression d'une association entre deux termes. Les relations ont les propriétés suivantes :

- La *profondeur* d'une relation est le nombre de niveaux sur lesquels la relation s'étend. Dans la syntaxe de la recherche, elle est précisée par le mot-clé THESDEPTH. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215 pour savoir comment utiliser l'extension de thésaurus dans une requête.
- La *direction* d'une relation précise si cette dernière est également vraie d'un terme à l'autre (bidirectionnelle), ou si elle n'est vraie que dans une direction (unidirectionnelle).

L'extension de thésaurus peut utiliser toutes les relations définies dans le thésaurus. Il est possible de préciser la profondeur de l'extension. Il s'agit du nombre maximal de transitions d'un terme source vers un terme cible. Cependant, l'ensemble de termes peut subir une croissance exponentielle avec l'augmentation de la profondeur.

L'exemple suivant montre la liaison entre l'ajout de termes et l'augmentation de la profondeur.

santé

sécurité sociale, paramédical, médicament, maladie

allergologie, virologie, médecine vétérinaire, toxicologie,  
chirurgie, stomatologie, rhumatologie, radiothérapie, psychiatrie, médecine préventive,  
pathologie, odontologie, diététique, médecine nucléaire,  
neurologie, néphrologie, bilan médical, médecine du travail,  
hématologie, médecine générale, épidémiologie, examen clinique,  
cardiologie, cancérologie

### Relations dans un thésaurus DB2 Extension Texte

Les types de relations fournis par un thésaurus DB2 Extension Texte sont les suivants

- Associative
- Synonyme
- Hiérarchique
- Autre

Dans un thésaurus DB2 Extension Texte, aucune relation n'est prédéfinie. Il est possible d'attribuer un nom à chaque relation, comme TERME ELARGI, qui peut également être une abréviation mnémotecnique, comme TE. Voici les relations couramment utilisées dans la conception d'un thésaurus :

- TE ou TERME ELARGI
- TR ou TERME RESTREINT
- TA ou TERME ASSOCIE
- SYN ou SYNONYME
- UTIL ou UTILISATION
- CTX ou CONTEXTE

**Associative** : Une relation associative est une relation bidirectionnelle établie entre des descripteurs s'étendant à toute profondeur. Elle relie deux termes non équivalents et sans relation hiérarchique entre eux, mais qui sont associés par leur sémantique de telle sorte que le lien qui est établi entre eux peut suggérer l'utilisation d'autres termes dans une indexation ou dans une extraction.

Les relations associatives sont couramment désignées par la mention Terme Associé (TA). En voici quelques exemples :

Chien TA sécurité  
Animal de compagnie TA vétérinaire

**Synonyme** : Quand une distinction est établie entre descripteurs et non descripteurs, ce qui est le cas dans un thésaurus DB2 Extension Texte, une relation synonyme est une relation unidirectionnelle établie entre deux termes de signification identique ou similaire. Dans un groupe de synonymes, l'un des termes est désigné comme descripteur. Les autres termes sont alors appelés non descripteurs. Reportez-vous à la section «Relations dans un thésaurus Ngram» à la page 272 qui définit la relation synonyme en l'absence de distinction entre descripteurs et non descripteurs.

La désignation courante UTILISATION amène d'un non descripteur donné vers son descripteur. La désignation courante CONTEXTE amène du descripteur vers chacun des non descripteurs. Par exemple :

félin UTILISATION chat

juriste CONTEXTE avocat

**Hiérarchique** : Une relation hiérarchique est une relation unidirectionnelle établie entre des descripteurs. Elle déclare qu'un des termes est plus spécifique, ou moins général, que l'autre. Cette différence amène à une représentation des termes sous forme de hiérarchie, où un terme représente une catégorie, dont ses subordonnés sont les membres. Par exemple, le terme "cheval" appartient à la catégorie "équidés".

TERME ELARGI et TERME RESTREINT désignent des relations hiérarchiques. Par exemple :

Voiture TR limousine  
Equidé TE cheval

**Autre** : Les relations de type *autre* sont les plus générales. Elles représentent une association qui ne tombe pas facilement dans une des autres catégories. Une relation de type *autre* peut être bidirectionnelle ou unidirectionnelle, sans restriction de niveau, et peut exister entre des descripteurs et des non descripteurs.

Cette relation est souvent associée aux nouveaux termes d'un thésaurus jusqu'à ce que leur relation correcte avec les autres termes soit déterminée.

Il vous est bien sûr possible de définir votre propre relation synonyme bidirectionnelle en utilisant le type de relation *associative* pour une relation synonyme entre descripteurs ou le type de relation *autre* pour une relation synonyme entre des termes arbitraires.

### Relations dans un thésaurus Ngram

Un thésaurus Ngram prend en charge les types de relations suivants :

- Associative
- Synonyme

Deux types de relations sont prédéfinies, sur la base de ces deux types. Vous pouvez définir vos propres relations en utilisant la formule *associative*. Pour plus de détails, reportez-vous à la section «Création d'un thésaurus Ngram» à la page 278.

#### **Associative**

Une relation associative est une relation bidirectionnelle entre deux termes qui n'expriment pas le même concept mais ont un rapport entre eux. La relation prédéveloppée *RELATED\_TO* et toutes les relations définies par l'utilisateur sont basées sur ce type de relation.

En voici quelques exemples :

tennis RELATED\_TO raquette

allemand RELATED\_TO saucisse

### Synonyme

Une relation synonyme est une relation bidirectionnelle entre deux termes de signification identique ou similaire pouvant être utilisés l'un pour l'autre. Cette relation peut, par exemple, s'appliquer à un terme et son abréviation. La relation prédéveloppée SYNONYM\_OF est la seule relation basée sur ce type.

En voici quelques exemples :

sali SYNONYM\_OF taché  
USA SYNONYM\_OF Etats-Unis

### Création d'un thésaurus

Reportez-vous également à la section «Création d'un thésaurus Ngram» à la page 278.

Le répertoire samples dans le chemin d'installation contient un exemple de compilateur de thésaurus anglais, dans le fichier d'entrée desthes.sgm . Le répertoire dictionnaire sur les systèmes OS/2 et Windows a la désignation suivante :

*unité*:\dmb\db2tx\samples

Sur les systèmes AIX, HP-UX et SUN-Solaris, le répertoire porte le nom :

*DB2TX\_INSTOWNER* /db2tx/samples

Une version compilée de ce thésaurus et de son fichier d'entrée SGML est mémorisée dans le répertoire du dictionnaire.

*unité*:\dmb\db2tx\dict  
ou  
*DB2TX\_INSTOWNER*/db2tx/dicts

Vous trouverez les fichiers appartenant à ce thésaurus sous les noms desthes.th1, desthes.th2, ... et desthes.th6..

Pour créer un thésaurus, commencez par définir son contenu dans un fichier. Il est recommandé d'utiliser un répertoire ordinaire pour chaque thésaurus que vous définissez. Le fichier peut avoir tout type d'extension sauf les fichiers th1 à th6, utilisés pour le dictionnaire du thésaurus. Si vous utilisez le même répertoire pour un thésaurus Ngram, reportez-vous à la section «Création d'un thésaurus Ngram» à la page 278 pour connaître les autres extensions exclues.

Compilez ensuite le fichier en exécutant la commande :

```
txthesc
-f nom-fichier -c ccsid
```

## Thésaurus

où *nom-fichier* ne peut comporter que les caractères a-z, A-Z et 0-9.

Actuellement, seul le paramètre CCSID 850 est pris en charge.

La commande `txthesc` crée des fichiers de thésaurus portant le nom que vous indiquez (*nom-fichier*) avec l'extension `th1` à `th6`, situés dans le répertoire où se trouve le fichier de définition. Si un thésaurus porte déjà ce nom, le nouveau se substitue à l'ancien sans avertissement.

Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215 pour savoir comment utiliser un thésaurus dans une requête.

Pour préciser le contenu d'un thésaurus, utilisez le langage SGML (Standard Generalized Markup Language). Le diagramme suivant indique la syntaxe à respecter lors de la création d'un thésaurus.

►<thesaurus><header><thname>*nom-thésaurus*</thname>—————►

►<rldf>—————| définition-relation |—————</rldf></header>—————►

►—————| entrée-thésaurus |—————</thesaurus>—————►

### définition-relation :

|<r|s><r|name>*nom-relation*</r|name><r|type> ASSOCIATIVE  
SYNONYMOUS  
HIERARCHICAL  
OTHER </r|type>—————►

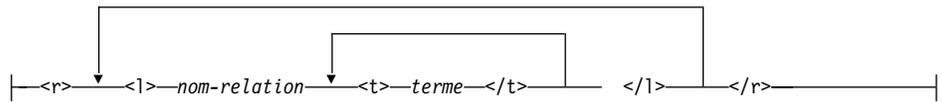
►</r|s>—————|

### entrée-thésaurus :

|<en>*numéro-unique*,  $\left[ \begin{array}{l} 1 \\ 2 \end{array} \right]$  <t>*terme*</t>—————| termes-associés |—————►

►</en>—————|

termes-associés :



*nom-relation* ne peut comporter que les caractères a-z, A-Z et 0-9.

La figure 19 à la page 276, contient la définition du thésaurus de la figure 18 à la page 269.

## Thésaurus

```
<thesaurus>
<header>
<thnom>Exemple de thésaurus thesc </thnom>
<rldef>

<rls>
<rlname>Terme associé</rlname>
<rltype>associative</rltype>
</rls>

<rls>
<rlname>Terme restreint</rlname>
<rltype>hierarchical</rltype>
</rls>

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

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

<en> 2, 1
<t>database management system</t>
<r>
  <l>Terme restreint
  <t>oo database management system</t>
  <t>relational database management system</t>
  </l>

  <l>Synonyme
  <t>DBMS</t>
  </l>

  <l>Terme associé
  <t>document management system</t>
  </l>

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

Figure 19. Définition d'un thésaurus simple (Numéro 1 de 2)

```

<en> 5, 1
<t> relational database management system </t>
<r>
  <l>Terme restreint
<t>object relational database management system</t>
  </l>
</r>
</en>

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

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

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

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

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

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

```

Figure 19. Définition d'un thésaurus simple (Numéro 2 de 2)

## Création d'un thésaurus Ngram

Le répertoire du dictionnaire dans le chemin d'installation contient un exemple de compilateur de thésaurus anglais Ngram, `desnthes.def`. Le répertoire du dictionnaire sur les systèmes OS/2 et Windows a la désignation suivante :

`unité:\dmb\db2tx\dict`

Sous AIX, HP-UX et SUN-Solaris, le répertoire du dictionnaire est :

`DB2TX_INSTOWNER /db2tx/dicts`

Une version compilée de cet exemple de thésaurus y est également mémorisée. Les fichiers appartenant à ce thésaurus sont nommés `desnthes.<extension>` avec l'extension suivante, n étant un chiffre :

- Pour les fichiers dictionnaires : `wdf, wdv, grf, grv, MEY, ROS, NEY, SOS, lkn`
- Pour les fichiers temporaires : `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 et $21`

Pour créer un thésaurus Ngram, commencez par définir son contenu dans un fichier. Vous pouvez avoir plusieurs thésaurus dans le même répertoire, mais il est recommandé d'avoir un répertoire séparé pour chaque thésaurus. La longueur du nom de fichier sans extension ne doit pas dépasser 8 caractères. L'extension est facultative mais limitée à 3 caractères et doit être différente des extensions répertoriées ci-dessus.

Si vous utilisez le même répertoire pour d'autres thésaurus DB2 Extension Texte, n'utilisez pas les extensions répertoriées dans la section «Création d'un thésaurus» à la page 273.

Compilez ensuite le fichier en exécutant la commande :

`txthesn -f nom-fichier-définition-ccsid page-codes`

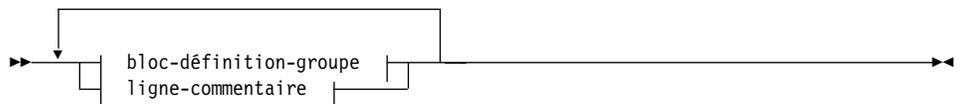
Voici la liste des pages de codes prises en charge par un thésaurus Ngram :

932	AIX, OS/2	Japonais
942	OS/2	Japonais
943	OS/2, Windows	Japonais
949	OS/2	Coréen
950	AIX, HP-UX, OS/2, SUN-Solaris, Windows	Chinois traditionnel
970	AIX, HP-UX, SUN-Solaris	Coréen
1381	OS/2, Windows	Chinois simplifié
1363	Windows	Coréen

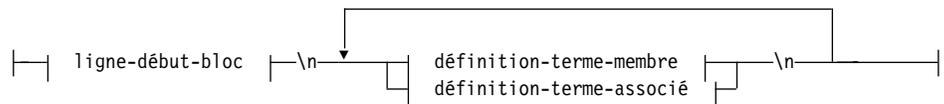
1383	AIX, HP-UX, SUN-Solaris	Chinois simplifié
850	AIX, OS/2	Latin-1
1252	Windows	Latin-1

La commande `txthesn` crée des fichiers de thésaurus portant le nom indiqué *nom-fichier-définition* et les extensions mentionnées ci-dessus. Les fichiers sont créés dans le même répertoire que le fichier de définition. Si un thésaurus de ce répertoire porte déjà ce nom, le nouveau se substitue à l'ancien sans avertissement.

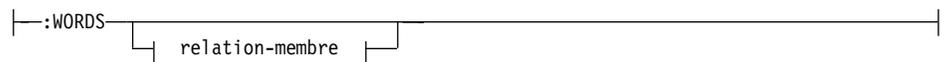
Spécifiez le contenu du thésaurus en utilisant le diagramme de syntaxe suivant :



**bloc-définition-groupe :**



**ligne-début-bloc :**



**relation-membre :**

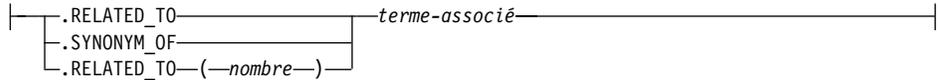


**définition-terme-membre :**



# Thésaurus

## définition-terme-associé :



## ligne-commentaire :



Chaque terme membre doit occuper une seule ligne. Chaque terme associé doit être précédé par le nom de la relation. Si les termes membres sont liés entre eux, indiquez leur relation.

La longueur des termes membres et celle des termes associés est limitée à 164 caractères. Les caractères simple octet et double octet pour une même lettre sont considérés comme identiques. Il n'y a pas de distinction entre majuscules et minuscules. Un terme peut contenir un caractère vide mais on peut utiliser le point "." ou les deux points ":" en caractère simple octet.

Les relations définies par l'utilisateur sont toutes basées sur le type de relation *associative*. Elles sont identifiées par des numéros uniques entre 1 et 128.

Si une application demande à utiliser des noms symboliques pour des relations thésaurus au lieu d'un nom et d'un numéro de relation, elle doit administrer la correspondance elle-même. Par exemple, si la relation `OPPOSITE_OF` a été affectée à `RELATED_TO(1)`, l'application doit mapper ce nom sur le nom de la relation interne `RELATED_TO(1)`. Reportez-vous au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215 pour savoir comment utiliser l'extension de thésaurus dans une requête.

---

## Chapitre 16. Fichiers de configuration

Le présent chapitre décrit les fichiers de configuration. Ceux-ci sont automatiquement générés lors de la création d'une instance DB2 Extension Texte par la commande `txicrt` ou de la création d'un profil client par la commande `descrc1`. Ces fichiers sont générés en page de codes 819.

Vous pouvez modifier ces fichiers afin de régler votre système, mais veuillez cependant à utiliser la page de codes correcte lors de cette opération. Les noms de sections et d'options peuvent être indiqués indifféremment en majuscules et/ou minuscules. Le point virgule est utilisé comme délimiteur de commentaire.

Si l'option est une valeur booléenne, les valeurs TRUE, YES, ON et 1 sont considérées comme vraies, qu'elles soient en majuscules ou en minuscules ; toutes les autres valeurs sont considérées comme fausses.

---

### Fichier de configuration du client

**Nom de fichier**

DESCL.INI

**Emplacement** Sous UNIX : DB2TX\_INSTOWNERHOMEDIR/db2tx

Sous Windows NT et Windows 2000 :

%DMBMMPATH%\instance\%DB2INSTANCE%\db2tx

Les options modifiées s'appliquent à partir de l'exécution suivante de la fonction `StartSession`.

Section	Option	Valeur par défaut	Description
[INSTANCE]	DESWORKCL		Pointe sur un répertoire de travail utilisé pour les fichiers temporaires.
	DESNLPCL		Pointe sur le répertoire des ressources.
	DESDTDPATHSRV		Pointe sur le répertoire dans lequel sont stockées les DTD pour les documents XML.

## Fichier de configuration du client

Section	Option	Valeur par défaut	Description
	DESDTDPATHCL		Pointe sur le chemin dans lequel doivent être stockés les fichiers DTD. Utilisé pour les fichiers XML.
[BUFFER]	BUFFERSEGMENTSIZ	32 000	Taille, en octets, des segments de blocs utilisés pour la mise en mémoire tampon.
	BUFFERSEGMENTCOUNT	3	Nombre maximum de segments utilisés avant que les tampons soient basculés dans des fichiers temporaires. Un segment tampon est défini par BUFFERSEGMENTSIZ.
[DOCUMENTFORMAT]	USEREXIT		Nom de l'exit utilisateur mis en oeuvre pour travailler avec des formats de documents non pris en charge. Indiquez soit un nom de fichier, si l'exit utilisateur est stocké dans un répertoire faisant partie de l'instruction PATH, soit un nom de fichier qualifié complet.  Pour plus de détails sur l'utilisation de l'exit utilisateur pour la conversion de format, reportez-vous à la section «Utilisation des formats de documents non pris en charge» à la page 33.
	FORMATRECOGNITION	TRUE	Déclenche la reconnaissance des formats de documents.  TRUE : reconnaissance de format activée  FALSE : reconnaissance de format désactivée

## Fichier de configuration du client

Section	Option	Valeur par défaut	Description
	UseExitForAllFormats	FALSE	<p>Détermine l'appel de l'exit utilisateur pour travailler sur des formats de documents ne figurant pas dans le fichier deslsdef.h. Vous devez définir une valeur pour USEREXIT.</p> <p>TRUE : l'exit utilisateur est toujours appelé. Si cette valeur est définie, FORMATRECOGNITION n'est pas prise en compte.</p> <p>FALSE : l'exit utilisateur est appelé pour tous les formats de documents dont la valeur se situe au-dessus de celle définie pour EHW_USER_FORMATS.</p>

## Fichier de configuration du serveur

### Nom de fichier

DESSRV.INI

### Emplacement

Sous UNIX :

DB2TX\_INSTOWNERHOMEDIR/db2tx/TXINSn

Sous Windows NT et Windows 2000 :

%DMBMMPPATH%\instance\%DB2INSTANCE%\db2tx\txinsnn

Ce fichier n'existe qu'en un seul exemplaire par instance de serveur.

Les options modifiées s'appliquent lors du démarrage suivant de l'instance du serveur.

Tableau 20. Options du fichier de configuration serveur

Section	Option	Valeur par défaut	Description
[INSTANCE]	DESWORKSRV		Pointe sur un répertoire de travail utilisé pour les fichiers temporaires.
	DESNLPSSRV		Pointe sur le répertoire des ressources.

## Fichier de configuration du serveur

Tableau 20. Options du fichier de configuration serveur (suite)

Section	Option	Valeur par défaut	Description
[DAEMON]	MaxMtEntries	30 sous AIX, Windows, OS/390  10 sous Solaris	Nombre maximum d'index utilisables en parallèle à un moment quelconque. Réduisez ce nombre si vous disposez de ressources insuffisantes, par exemple de sémaphores ou de mémoires partagées. La disponibilité en ressources dépend de la plateforme et il en est de même des valeurs par défaut.
	MaxIndexEntries	1000	Nombre maximum d'index utilisés. Réduisez ce nombre si vous disposez de ressources insuffisantes, par exemple de mémoires partagées.
[BUFFER]	BUFFERSEGMENTSZIE	32 000	Taille, en octets, des segments de blocs utilisés pour la mise en mémoire tampon. Cette valeur est utilisée par EhWUpdate.
	BUFFERSEGMENTCOUNT	3	Nombre maximum de segments utilisés lors du processus de mise à jour des index, avant que les tampons soient basculés dans des fichiers temporaires. Augmentez cette valeur si vos collections de documents comportent des documents de grande taille.
	BUFFERSORTSIZE	20 000 000	Taille, en octets, du tampon utilisé pour trier les fichiers de travail temporaires.

Tableau 20. Options du fichier de configuration serveur (suite)

Section	Option	Valeur par défaut	Description
[DOCUMENTFORMAT]	USEREXIT		<p>Nom de l'exit utilisateur mis en oeuvre pour travailler avec des formats de documents non pris en charge. Indiquez soit un nom de fichier, si l'exit utilisateur est stocké dans un répertoire faisant partie de l'instruction PATH, soit un nom de fichier qualifié complet.</p> <p>Pour plus de détails sur l'utilisation de l'exit utilisateur pour la conversion de format, reportez-vous à la section «Utilisation des formats de documents non pris en charge» à la page 33.</p>
	FORMATRECOGNITION	TRUE	<p>Déclenche la reconnaissance des formats de documents énumérés dans le fichier deslsdef.h.</p> <p>TRUE : reconnaissance de format activée</p> <p>FALSE : reconnaissance de format désactivée</p>

## Fichier de configuration du serveur

Tableau 20. Options du fichier de configuration serveur (suite)

Section	Option	Valeur par défaut	Description
	UseExitForAllFormats	FALSE	<p>Détermine l'appel de l'exit utilisateur pour travailler sur des formats de documents ne figurant pas dans le fichier deslsdef.h. Vous devez définir une valeur pour USEREXIT.</p> <p>TRUE : l'exit utilisateur est toujours appelé. Si cette valeur est définie, FORMATRECOGNITION n'est pas prise en compte.</p> <p>FALSE : l'exit utilisateur est appelé pour tous les formats de documents dont la valeur se situe au-dessus de celle définie pour EHW_USER_FORMATS.</p>
[LINGPREC] Pour tous les index dont le type de base est linguistique ou exact.	UPDATETHRESHOLD	4 000 000	Un processus de mise à jour d'index est éclaté en interne en plusieurs passes de mise à jour et de réorganisation. Cette valeur précise le nombre de mots à collecter au cours d'une étape de mise à jour.
	UPDATESLICE	1	Nombre de passes de mise à jour avant que ne démarre un processus de réorganisation interne. Une passe de mise à jour est définie par l'option UPDATETHRESHOLD.
[NGRAM] Pour tous les index prenant en charge DBCS.	UPDATETHRESHOLD	10 000 000	Taille totale, en octets, des documents ajoutés à un index au cours d'une passe de mise à jour. S'il y a dépassement du seuil défini, un processus de réorganisation démarre automatiquement.

Tableau 20. Options du fichier de configuration serveur (suite)

Section	Option	Valeur par défaut	Description
	UPDATESLICE	10 000	Nombre maximum de documents dans un index secondaire. Ce nombre est vérifié après chaque passe de mise à jour. Si le nombre de documents est supérieur à cette valeur, un processus de réorganisation démarre automatiquement.

---

### Fichier de configuration d'analyse de texte

**Nom de fichier**

DESTAF.TCF

**Emplacement** Sous UNIX :

DB2TX\_INSTOWNERHOMEDIR/db2tx/TXINSnnn

Sous Windows NT et Windows 2000 :

%DMBMPATH%\instance%\DB2INSTANCE%\db2tx\txinsnnn

Ce fichier n'existe qu'en un seul exemplaire par instance de serveur.

Le fichier de configuration d'analyse de texte comporte des sections pour le traitement des mises à jour et des requêtes pour tous les types d'index, à l'exception des index Ngram. Vous ne pouvez apporter de modifications qu'aux options décrites ici. Toute autre modification risque d'entraîner des erreurs lors du traitement des mises à jour ou des requêtes.

#### Options communes

Les options suivantes s'appliquent :

- [TAF\_Application:UPDATE:PREC]) pour les index exacts (PRECISE)
- [TAF\_Application:UPDATE:LING]) pour les index linguistiques (LINGUISTIC)
- [TAF\_Application:UPDATE:NORM]) pour les index exacts normalisés (PRECISE NORMALIZED)
- [TAF\_Application:QUERY:PREC])
- [TAF\_Application:QUERY:LING])
- [TAF\_Application:QUERY:NORM])

## Fichier de configuration d'analyse de texte

Option	Valeur par défaut	Description
FilterSet	Aucune, obligatoire	Nom de la section, dans le fichier de configuration, contenant les règles pour le filtrage des mots vides, par exemple, StwUseAllUpCs, StwIgnAllUpCs.
Normalize	Non	Indique si la normalisation est appliquée à l'index.  Si l'option est définie par Oui, la normalisation convertit tous les caractères en minuscules et remplace les caractères avec marque diacritique par des caractères de base.  Si l'option est définie par Non, les entrées pour NormalizeUmlauts et ConvertAllUpperToLower ne sont pas prises en compte.
NormalizeUmlauts	Oui	Indique si les trémas doivent être remplacés, par exemple, ä par ae. Ceci ne s'applique qu'aux documents en allemand, danois, suédois ou norvégien.
ConvertAllUpperToLower	Non	Indique si les mots intégralement indiqués en majuscules doivent être convertis en minuscules.

### [TAF\_Filter:STWUSEALLUPCS] [TAF\_Filter:STWIGNALLUPCS]

Les options de ces sections décrivent les ensembles de filtrage des mots vides. La section [TAF\_Filter:STWUSEALLUPCS] traite tous les mots en majuscules comme mots en majuscules/minuscules et la section [TAF\_Filter:STWIGNALLUPCS] définit un traitement spécial pour tous les mots en majuscules.

Option	Valeur par défaut	Description
FilterType	TAF_FILTER_STW	Section du fichier de configuration contenant les règles pour le filtrage des mots vides.
MinTokenLength	3	Nombre minimum de caractères nécessaires pour une marque. Si la longueur de la marque est inférieure à cette valeur, elle ne sera pas indexée. Entier $\geq 1$ .
MinUpperCaseTokenLength	2	Nombre minimum de caractères nécessaires pour une marque toute en majuscules. Si la longueur de la marque est inférieure à cette valeur, elle ne sera pas indexée. Entier $\geq 1$ .

## Fichier de configuration d'analyse de texte

Option	Valeur par défaut	Description
UseAlternateTerritories	Non	Indique si un autre dictionnaire doit être utilisé pour le traitement des mots vides lorsque le dictionnaire propre à la langue considérée ne peut être localisé.  Si l'option est définie par Oui, le premier dictionnaire de mots vides correspondant à la langue est utilisé. Par exemple, si la langue en cours est en-gb mais qu'il n'existe pas de fichier de mots vides en-gb.tsw, le fichier en-us.tsw sera utilisé.  Si l'option est définie par Non, aucun filtrage de mots vides n'est effectué.

### [TAF\_AnnotationType:TAF\_ThesaurusEntry ]

La présente section définit les valeurs par défaut pour les interrogations impliquant une extension du thésaurus.

Mot clé	Valeur par défaut	Description
ThesBaseName	desthes	Définit le nom par défaut du thésaurus utilisé pour des interrogations impliquant une extension du thésaurus. Le nom est une chaîne de caractères.
Profondeur	1	Définit le chemin par défaut traversant le thésaurus utilisé pour des interrogations impliquant une extension du thésaurus. Entier ≥ 1.

### [TAF\_Plugin:TAF\_POE ]

Vous pouvez définir une langue dans l'option PreloadLanguages afin de modifier l'environnement de langue par défaut préchargé au démarrage de Text Search Engine.

Option	Valeur par défaut	Description
PreloadLanguages	EN-US	Définit un ensemble de ressources propres à une langue, qui doivent être préchargées au démarrage. Par exemple, si une application s'exécute sur des documents en italien, les ressources correspondant à l'italien sont chargées. La valeur de la zone s'exprime sous la forme XX-YY, où XX est la langue et YY le pays.

## Fichier de configuration d'analyse de texte

---

## Chapitre 17. Codes retour

Le présent chapitre répertorie les codes renvoyés par l'API de DB2 Extension Texte suite à un appel de fonction. Ils sont répertoriés par ordre alphabétique.

Tous les appels API de DB2 Extension Texte renvoient un code retour numérique en tant que valeur de la fonction C. Les codes retour sont définis dans le fichier d'inclusion DES\_EXT.H fourni avec DB2 Extension Texte.

L'API de DB2 Extension Texte intercepte les erreurs et les signale, accompagnées d'un code retour.

Les applications qui appellent les fonctions API de DB2 Extension Texte API doivent toujours vérifier le code retour avant de tenter de traiter d'autres paramètres de sortie. Les codes retour correspondant à chaque appel sont répertoriés avec leurs paramètres dans le «Chapitre 13. Fonctions API de recherche et d'affichage» à la page 229.

Dans certains cas, une entrée incorrecte telle qu'un pointeur de session obsolète peut entraîner un arrêt anormal des fonctions API qui ne pourra pas être détecté par DB2 Extension Texte.

---

### RC\_ALLOCATION\_ERROR

**Explication :** Impossible d'allouer de la mémoire à usage interne.

**Action :** Assurez-vous qu'il existe suffisamment de mémoire disponible.

---

### RC\_FILE\_IO\_PROBLEM

**Explication :** DB2 Extension Texte n'a pas pu lire ou écrire sur un fichier.

**Action :** Vérifiez qu'il existe suffisamment d'espace disque et de mémoire disponibles sur le serveur. Vérifiez que les variables d'environnement et les définitions de la configuration de texte sont correctement définies.

---

### RC\_INVALID\_BROWSE\_INFO

**Explication :** Les informations d'affichage renvoyées par DesGetSearchResultTable ou par DesGetBrowseInfo et utilisées en entrée par DesStartBrowseSession ne sont pas correctes.

**Action :** Vérifiez si une erreur de programmation n'écrase pas les informations d'affichage.

---

### RC\_INVALID\_BROWSE\_OPTION

**Explication :** L'option d'affichage dans DesGetSearchResultTable n'est pas correcte.

**Action :** Vérifiez que l'option est bien définie par BROWSE ou NO\_BROWSE.

## Codes retour

---

### RC\_INVALID\_MATCH\_OPTION

**Explication** : L'option de correspondance utilisée dans DesOpenDocument n'est pas correcte.

**Action** : Vérifiez que l'option est bien définie par FAST ou EXTENDED.

---

### RC\_INVALID\_PARAMETER

**Explication** : L'un des paramètres d'entrée est incorrect.

**Action** : Lisez le message d'erreur renvoyé par DB2 Extension Texte afin d'en déterminer la cause.

---

### RC\_INVALID\_SEARCH\_OPTION

**Explication** : L'option de recherche dans DesGetSearchResultTable n'est pas correcte.

**Action** : Vérifiez que l'option est bien définie par DES\_TEXTHANDLEONLY, DES\_RANK, DES\_MATCH ou DES\_RANKANDMATCH.

---

### RC\_INVALID\_SESSION

**Explication** : Le pointeur de session indiqué dans l'appel de fonction en cours est incorrect ou obsolète.

**Action** : Sauvegardez les informations pouvant vous aider à trouver la cause de cette erreur, puis mettez fin à l'application.

---

### RC\_NO\_BROWSE\_INFO

**Explication** : Aucune information d'affichage n'est renvoyée par DB2 Extension Texte car l'argument de recherche a abouti à un résultat de recherche vide. Ceci n'est pas une erreur.

**Action** : Aucune action n'est nécessaire.

---

---

### RC\_PARSER\_INVALID\_ESCAPE\_CHARACTER

**Explication** : Le critère de recherche contient un caractère d'échappement incorrect. Cette erreur se produit si un blanc est utilisé en tant que caractère d'échappement ou si, pour un mot ou une expression, plusieurs caractères d'échappement sont indiqués dans le critère de recherche. Exemple : ESCAPE " " ou ESCAPE "#\$".

**Action** : Vérifiez la syntaxe de l'argument de recherche, puis relancez la recherche.

---

### RC\_PARSER\_INVALID\_USE\_OF\_ESCAPE\_CHAR

**Explication** : La syntaxe du caractère d'échappement dans le critère de recherche ne peut pas être interprétée.

**Action** : Vérifiez la syntaxe du caractère d'échappement. Par exemple, si \$ est le caractère d'échappement indiqué, le mot ou l'expression ne peut contenir que \$\$, \$\_ ou \$%, où \_ et % sont les deux symboles génériques autorisés.

---

### RC\_PARSER\_SYNTAX\_ERROR

**Explication** : La syntaxe du critère de recherche ne peut pas être interprétée.

**Action** : Vérifiez la syntaxe de l'argument de recherche, en vous reportant au «Chapitre 12. Syntaxe des arguments de recherche» à la page 215.

---

### RC\_RESULT\_TABLE\_NOT\_EXIST

**Explication** : La table dans laquelle vous tentez de stocker le résultat d'une recherche n'existe pas.

**Action** : Créez une table de résultats, comme indiqué dans la figure 17 à la page 245.

---

### RC\_SE\_BROWSER\_TIME\_OUT

**Explication** : Le processus d'affichage a été lancé mais n'a pas réagi dans les délais prévus.

DB2 Extension Texte a donc annulé le processus en attente.

Cette erreur peut se produire lorsque votre système manque d'espace mémoire ou est surchargé.

**Action :** Mettez fin à la session d'affichage en appelant `DesEndBrowseSession`, libérez de la mémoire allouée en appelant `DesFreeBrowseInfo` puis relancez le processus.

---

#### RC\_SE\_CAPACITY\_LIMIT\_EXCEEDED

**Explication :** La fonction demandée ne peut être exécutée. La mémoire ou l'espace disque sont insuffisants.

**Action :** Mettez fin au programme et vérifiez les ressources de votre système.

---

#### RC\_SE\_COMMUNICATION\_PROBLEM

**Explication :** La communication avec le serveur DB2 Extension Texte a échoué. Cette erreur peut être due à un manque d'espace mémoire ou à l'installation incorrecte de DB2 Extension Texte.

**Action :** Sauvegardez les informations pouvant vous aider à trouver la cause de cette erreur, puis mettez fin à l'application.

---

#### RC\_SE\_CONFLICT\_WITH\_INDEX\_TYPE

**Explication :** La spécification linguistique du terme de recherche de la requête ne correspond pas au type d'index. Par exemple, `PRECISE FORM OF` ne peut pas être utilisé avec un index linguistique. La spécification linguistique par défaut est utilisée comme indiqué dans le tableau 8 à la page 223.

**Action :** Adaptez votre application pour éviter que ne soient définies des options de requête provoquant des conflits avec le type d'index.

---

#### RC\_SE\_DICTIONARY\_NOT\_FOUND

**Explication :** Les fonctions linguistiques de DB2 Extension Texte ne trouvent pas les dictionnaires. La requête est traitée sans prise en charge linguistique. Les fichiers dictionnaires correspondant au code de langue indiqué ne se trouvent pas dans le chemin prévu.

**Action :** Vous pouvez continuer à lancer des appels API. Pour UNIX, vérifiez que le dictionnaire voulu figure dans le chemin d'accès `{DB2TX_INSTOWNERHOMEDIR}/db2tx/dicts`. Pour OS/2, vérifiez que le dictionnaire voulu figure dans le chemin défini dans le paramètre de configuration de texte `DB2TX_DATA`. Si nécessaire, installez le dictionnaire voulu.

---

#### RC\_SE\_DOCMOD\_READ\_PROBLEM

**Explication :** Lorsqu'une instance de DB2 Extension Texte est créée, un fichier de modèles de documents appelé `desmodel.ini` est rangé dans le répertoire de cette instance. Lorsque vous créez un index, un fichier `desmodel.ini` est également rangé dans le répertoire `index Ixnxxxxx`. Ce fichier de modèles de documents n'a pas pu être lu.

**Action :** Vérifiez qu'un fichier de modèles de documents a été créé et qu'il se trouve dans le répertoire adéquat.

---

#### RC\_SE\_DOCUMENT\_NOT\_ACCESSIBLE

**Explication :** Le document texte demandé a été trouvé, mais il n'est pas accessible pour l'instant.

**Action :** Vérifiez si le document est accédé en mode exclusif par une autre tâche ou un autre utilisateur.

---

#### RC\_SE\_DOCUMENT\_NOT\_FOUND

**Explication :** Le document texte demandé n'a pas été trouvé. Ce document a très probablement été supprimé de la mémoire, mais pas de l'index

## Codes retour

de DB2 Extension Texte. Cette erreur peut également se produire si vous tentez d'afficher un document identifié par un descripteur altéré.

**Action :** Dans la plupart des cas, vous pouvez ignorer ce code retour. Il ne s'affichera plus après la prochaine mise à jour de l'index.

S'il continue à s'afficher, vérifiez que votre programme d'application transfère le descripteur trouvé correctement en vue de l'affichage.

---

### RC\_SE\_EMPTY\_INDEX

**Explication :** L'index de DB2 Extension Texte correspondant à la colonne de descripteurs concernée par la demande de recherche est vide. Aucun document texte n'a été ajouté à cet index ou tous les documents texte en ont été supprimés.

Ceci peut se produire lorsqu'une colonne de texte a été activée, mais que les documents s'y trouvant n'ont pas encore été indexés. Autrement dit, vous avez indiqué dans la commande `ENABLE TEXT COLUMN` que l'index doit être créé ultérieurement, au moment déterminé par les paramètres d'indexation périodique.

Cette erreur peut également se produire lorsqu'une table de texte a été activée pour créer un index commun vide pour toutes les colonnes de texte, mais qu'aucune de ces dernières n'a été activée.

**Action :** Si `ENABLE TEXT TABLE` a été utilisée pour créer un index commun vide pour toutes les colonnes de texte, exécutez `ENABLE TEXT COLUMN` pour au moins une des colonnes de texte qui contient du texte dans lequel effectuer la recherche. Dans cette commande, vous pouvez déterminer si l'index est créé immédiatement ou à un moment déterminé par les paramètres d'indexation périodique.

Lancez la commande `GET INDEX STATUS` pour vérifier que l'index a bien été créé.

---

### RC\_SE\_EMPTY\_QUERY

**Explication :** Le critère de recherche indiqué a été analysé et traité au niveau linguistique par DB2 Extension Texte. Il peut s'agir d'une erreur de programmation qui a entraîné la création d'une requête ne contenant pas de termes de recherche ou du fait que tous les termes de recherche étaient des mots vides (mots non indexés par DB2 Extension Texte) qui sont supprimés d'une requête. Le résultat est l'absence de terme de recherche.

**Action :** Reformulez la requête. Si cette erreur persiste, recherchez une éventuelle erreur de programmation.

---

### RC\_SE\_END\_OF\_INFORMATION

**Explication :** Ceci n'est pas une erreur. Vous avez atteint la fin du document. `DesGetMatches` ne peut disposer d'aucune autre information.

**Action :** Utilisez ce code retour pour mettre fin au traitement itératif du document par `DesGetMatches`.

---

### RC\_SE\_FUNCTION\_DISABLED

**Explication :** La fonction demandée a appelé une fonction de DB2 Extension Texte qui a été neutralisée par l'administrateur.

**Action :** Demandez l'aide de votre administrateur. Il peut s'avérer nécessaire d'arrêter et de redémarrer DB2 Extension Texte (`txstop/txstart`).

---

### RC\_SE\_FUNCTION\_IN\_ERROR

**Explication :** La fonction demandée a été verrouillée en raison d'une erreur qui s'est produite sur le serveur de DB2 Extension Texte. L'appel d'API ne peut être traité.

**Action :** Vérifiez l'état de l'index ainsi que l'espace disponible dans le répertoire d'index. Réinitialisez l'état puis relancez la commande.

---

**RC\_SE\_INCORRECT\_HANDLE**

**Explication :** Un descripteur spécifié dans un paramètre d'entrée tel que *descripteur\_session\_affichage* n'est pas correct. Il doit s'agir d'un descripteur renvoyé par un appel précédent et qui ne soit pas obsolète.

**Action :** Sauvegardez les informations pouvant vous aider à trouver la cause de cette erreur, puis mettez fin à la session en appelant `DesEndBrowseSession`.

Vérifiez si une erreur de programmation n'a pas généré de descripteur incorrect.

---

**RC\_SE\_INDEX\_DELETED**

**Explication :** L'index de DB2 Extension Texte en cours d'accès est supprimé.

**Action :** Contactez l'administrateur de DB2 Extension Texte pour recréer l'index.

---

**RC\_SE\_INDEX\_NOT\_ACCESSIBLE**

**Explication :** L'accès à l'index de DB2 Extension Texte est impossible et l'appel en cours ne peut être traité.

**Action :** Demandez à l'administrateur de DB2 Extension Texte de vérifier l'accessibilité de l'index.

---

**RC\_SE\_INDEX\_SUSPENDED**

**Explication :** DB2 Extension Texte a reçu une demande relative à l'un de ses index suspendu à partir d'une autre session ou de la session en cours.

**Action :** Demandez à l'administrateur de DB2 Extension Texte de vérifier l'état de l'index.

---

**RC\_SE\_INSTALLATION\_PROBLEM**

**Explication :** Un incident lié à l'installation s'est produit au niveau de DB2 Extension Texte.

**Action :** Vérifiez la définition des variables d'environnement `DB2INSTANCE`, `DB2TX_INSTOWNER` et `DB2TXINSTOWNERHOMEDIR`. Utilisez `descfgcl -d` et `descfgsv -d -i txinsnn`

---

**RC\_SE\_IO\_PROBLEM**

**Explication :** Une erreur s'est produite lorsque le serveur a tenté d'ouvrir ou de lire l'un de ses fichiers index, et ce pour l'une des raisons suivantes :

- Action involontaire de l'administrateur telle que la suppression d'un fichier index de DB2 Extension Texte
- Définition incorrecte de la configuration de texte

**Action :** Mettez fin à l'application. Vérifiez auprès de l'administrateur que :

- Tous les fichiers de l'index de DB2 Extension Texte en cours existent.
- Les définitions de la configuration de texte sont correctes.

---

**RC\_SE\_LS\_FUNCTION\_FAILED**

**Explication :** Une fonction qui a accédé à la base de données pour extraire des documents texte pour affichage a échoué. La base de données n'est plus accessible à l'utilisateur ou ce dernier ne dispose pas des droits nécessaires pour la table de texte.

**Action :** Vérifiez que l'entrée permettant d'accéder à la fonction, l'ID utilisateur par exemple, est correcte. Vérifiez que la base de données est accessible et que l'utilisateur dispose des droits correspondant à la tâche.

## Codes retour

---

### RC\_SE\_LS\_NOT\_EXECUTABLE

**Explication :** Une fonction qui tente d'accéder à la base de données pour extraire des documents texte pour affichage ne peut être exécutée.

**Action :** Vérifiez que DB2 Extension Texte est correctement installée. Si cette erreur persiste, contactez votre agent commercial IBM.

---

### RC\_SE\_MAX\_OUTPUT\_SIZE\_EXCEEDED

**Explication :** Un nombre anormalement élevé de correspondances a été trouvé. Le volume des informations d'affichage a dépassé le maximum pouvant être géré. La demande ne peut pas être traitée.

**Action :** Affinez votre requête ou faites en sorte de disposer d'une plus grande quantité de mémoire système.

---

### RC\_SE\_MAX\_NUMBER\_OF\_BUSY\_INDEXES

**Explication :** La fonction de recherche a interdit l'exécution de la fonction demandée car le nombre maximal d'index actifs a été atteint.

**Action :** Attendez quelques instants et appelez à nouveau cette fonction. En règle générale, cet incident est temporaire.

---

### RC\_SE\_NO\_DATA

**Explication :** Ceci n'est pas une erreur. Aucun document texte ne correspond au critère de recherche. Si vous avez demandé des informations d'affichage, aucune n'est renvoyée. Aucune mémoire n'est allouée pour les informations d'affichage.

**Action :** Aucune action n'est nécessaire.

---

### RC\_SE\_NOT\_ENOUGH\_MEMORY

**Explication :** L'espace mémoire sur le poste client ou sur le système serveur est insuffisant.

La demande en cours ne peut être traitée.

**Action :** Libérez de l'espace mémoire et mettez fin à l'application.

---

### RC\_SE\_PROCESSING\_LIMIT\_EXCEEDED

**Explication :** La demande de recherche en cours a dépassé la taille de résultat maximale ou le temps de traitement maximal indiqué pour votre environnement client/serveur. La demande a été annulée.

**Action :** Affinez votre requête. Le cas échéant, augmentez le temps de traitement maximal.

---

### RC\_SE\_QUERY\_TOO\_COMPLEX

**Explication :** La requête indiquée est trop complexe.

**Action :** Adaptez votre application pour éviter toute utilisation excessive de caractères génériques et de synonymes.

Une utilisation excessive de symboles génériques ou de l'option SYNONYM peut porter une requête à une taille telle qu'elle ne soit pas gérable par DB2 Extension Texte.

---

### RC\_SE\_REQUEST\_IN\_PROGRESS

**Explication :** Une fonction API d'affichage de DB2 Extension Texte a été appelée alors qu'une autre demande API d'affichage était active pour la même session.

**Action :** Mettez fin à la session en appelant DesEndBrowseSession et libérez de la mémoire en appelant DesFreeBrowseInfo.

L'API d'affichage de DB2 Extension Texte ne prend pas en charge des accès simultanés à la même session d'affichage.

Toutes les applications s'exécutant simultanément dans le même processus doivent gérer leurs propres sessions d'affichage.

---

### RC\_SE\_SERVER\_BUSY

**Explication :** Le client DB2 Extension Texte ne parvient pas actuellement à établir une connexion avec le serveur DB2 Extension Texte demandé ou la liaison de données de ce dernier a été interrompue et ne peut être rétablie.

Le serveur DB2 Extension Texte a été démarré correctement, mais le nombre maximal de processus de serveur en parallèle a été atteint.

**Action :** Si cette erreur persiste, modifiez la configuration de communication sur le serveur.

---

### RC\_SE\_SERVER\_CONNECTION\_LOST

**Explication :** La communication entre le poste client et le serveur a été interrompue et ne peut pas être rétablie.

Il se peut que la tâche du serveur DB2 Extension Texte ait été arrêtée par un administrateur ou que le poste de travail serveur ait été éteint.

**Action :** Vérifiez si l'une de ces situations s'est produite et agissez en conséquence.

---

### RC\_SE\_SERVER\_NOT\_AVAILABLE

**Explication :** Les fonctions API de DB2 Extension Texte ne sont pas parvenues à établir une session avec le serveur DB2 Extension Texte demandé.

Il se peut que ce dernier n'ait pas été démarré.

**Action :** Vérifiez que le serveur DB2 Extension Texte a été démarré correctement. Si l'erreur persiste, elle est due à l'installation.

---

### RC\_SE\_STOPWORD\_IGNORED

**Explication :** Ce code informatif est renvoyé lorsque la requête indiquée contenait au moins un terme de recherche uniquement composé de mots vides. Le terme de recherche a été ignoré lors du traitement de la requête.

**Action :** Vous pouvez continuer à lancer des appels API. Évitez d'utiliser des mots vides dans les requêtes DB2 Extension Texte.

---

### RC\_SE\_UNEXPECTED\_ERROR

**Explication :** L'erreur qui s'est produite résulte peut-être d'une installation incorrecte de DB2 Extension Texte.

**Action :** Mettez fin à l'application, en sauvegardant les informations susceptibles de vous aider à déterminer la cause de l'erreur.

---

### RC\_SE\_UNKNOWN\_INDEX\_NAME

**Explication :** Le nom de l'index de texte associé à une colonne de texte fait partie du descripteur.

**Action :** Faites en sorte que le descripteur utilisé en entrée dans DesGetBrowseInfo soit correct.

---

### RC\_SE\_UNKNOWN\_SECTION\_NAME

**Explication :** Un nom de section indiqué ne provient ni d'un modèle spécifié dans un fichier de modèles de documents, ni du modèle utilisé par défaut.

**Action :** Indiquez un nom de section issu du modèle spécifié ou du modèle par défaut.

---

### RC\_SE\_UNKNOWN\_SERVER\_NAME

**Explication :** Le nom du serveur DB2 Extension Texte fait partie du descripteur.

**Action :** Faites en sorte que le descripteur utilisé en entrée dans DesGetBrowseInfo soit correct.

---

### RC\_SE\_WRITE\_TO\_DISK\_ERROR

**Explication :** Une erreur d'écriture s'est produite. Elle peut être due au fait que le disque sur le poste serveur DB2 Extension Texte est

## Codes retour

plein ou que l'installation de DB2 Extension Texte est incorrecte.

**Action** : Mettez fin à l'application, en sauvegardant les informations susceptibles de vous aider à déterminer la cause de l'erreur. Vérifiez que l'espace disque disponible sur le serveur est suffisant.

---

### RC\_SQL\_ERROR\_WITH\_INFO

**Explication** : Une erreur SQL s'est produite. Un message d'erreur est renvoyé.

**Action** : Prenez connaissance du message d'erreur renvoyé par DB2 Extension Texte pour obtenir de plus amples informations, telles que le message d'erreur SQL, le SQLState et le code d'erreur SQL natif.

---

### RC\_SQL\_ERROR\_NO\_INFO

**Explication** : Une erreur SQL s'est produite. Aucun message d'erreur n'est renvoyé.

---

### RC\_TEXT\_COLUMN\_NOT\_ENABLED

**Explication** : La colonne de descripteurs indiquée ne se trouve pas dans la table spécifiée.

**Action** : Vérifiez si le nom de la colonne de descripteurs indiqué est correct. Assurez-vous que la colonne de texte de cette table a été activée.

---

## Chapitre 18. Messages

Le présent chapitre décrit :

- **Les SQLSTATE renvoyés par les fonctions DB2 Extension Texte** : ces messages peuvent s'afficher lorsque vous utilisez des fonctions DB2 Extension Texte.
- **Les messages provenant de l'interpréteur de commandes DB2TX** : ces messages peuvent s'afficher lorsque vous entrez des commandes à l'aide de l'interpréteur de commandes DB2TX. Chaque numéro de message comporte le préfixe DES.

---

### SQLSTATE renvoyés par les fonctions DB2 Extension Texte

Les fonctions SQL fournies par DB2 Extension Texte peuvent renvoyer des messages d'erreur. Exemple :

```
SQL0443N La fonction utilisateur
"DB2TX.CONTAINS" (nom spécifique "DES5A")
a renvoyé un SQLSTATE d'erreur
accompagné du texte de diagnostic "Ouverture du
fichier de messages impossible".
SQLSTATE=38702
```

Les messages de cette section sont classés par numéro SQLSTATE.

---

#### 01H10 Impossible d'ouvrir le fichier *nom\_fichier.*

**Action** : Assurez-vous que ce fichier existe, et que le nom d'instance DB2 dispose des droits requis pour l'ouvrir.

---

#### 01H11 Un descripteur de texte est incomplet

**Explication** : Une tentative a été effectuée pour utiliser un descripteur qui a été initialisé, mais est incomplet. Un descripteur partiel a été créé via INIT\_TEXT\_HANDLE et contient les valeurs prédéfinies pour la langue et le format du document. Cependant, le descripteur n'a pas été activé par un déclencheur.

**Action** : N'utilisez que des descripteurs de texte complets. Si le descripteur concerné est stocké dans une colonne de descripteurs, activez

ou réactivez sa colonne de texte.

---

#### 01H12 Arguments de recherche trop longs. Le second argument de recherche a été ignoré.

**Explication** : La fonction REFINE a été utilisée pour combiner deux arguments de recherche, mais leur longueur totale est supérieure au maximum autorisé pour un texte de type LONG VARCHAR. REFINE renvoie le premier argument de recherche et non l'argument combiné.

**Action** : Réduisez la longueur de l'un ou des deux arguments, puis relancez la requête.

## SQLSTATE renvoyés par les fonctions DB2 Extension Texte

---

### 01H13 L'argument de recherche contient des mots vides.

**Explication :** La requête indiquée contient au moins un terme de recherche constitué uniquement de mots vides. Le terme de recherche a été ignoré lors du traitement de la requête.

**Action :** Evitez d'utiliser des mots vides dans les requêtes de DB2 Extension Texte.

---

### 01H14 Un dictionnaire manque pour le traitement linguistique.

**Explication :** Les fonctions linguistiques de DB2 Extension Texte ne trouvent pas les dictionnaires. La requête est traitée sans prise en charge linguistique. Les dictionnaires correspondant au(x) code(s) langue indiqué(s) ne figurent pas dans le chemin d'accès prévu.

**Action :** Pour UNIX, vérifiez que le dictionnaire voulu figure dans le chemin d'accès {DB2TX\_INSTOWNERHOMEDIR}/db2tx/dicts. Pour OS/2, vérifiez que le dictionnaire voulu figure dans le chemin d'accès indiqué dans les définitions de la configuration de texte. Si nécessaire, installez le dictionnaire voulu.

---

### 01H15 La spécification d'un terme de recherche linguistique ne correspond pas au type d'index.

**Explication :** La spécification linguistique du terme de recherche de la requête ne correspond pas au type d'index. Par exemple, PRECISE FORM OF ne peut pas être utilisé avec un index linguistique. La spécification linguistique par défaut est utilisée comme indiqué dans le tableau 8 à la page 223.

**Action :** Adaptez votre application pour éviter que ne soient définies des options de requête provoquant des conflits avec le type d'index.

---

### 38700 La bibliothèque de l'Extension Texte n'est pas active.

**Explication :** Une tentative a été effectuée pour utiliser un descripteur pouvant être interprété

uniquement par une version ultérieure de DB2 Extension Texte.

**Action :** Assurez-vous que le chemin d'accès à la version de la bibliothèque en cours est défini correctement et que vous disposez des droits requis pour y accéder.

Pour savoir quelle fonction a généré cette erreur, consultez la vue de catalogue DB2 SYSCAT.FUNCTIONS dans la colonne IMPLEMENTATION.

---

### 38701 Impossible d'ouvrir le fichier de trace fichiertrace.

**Explication :** Une tentative a été effectuée pour utiliser la fonction de trace écrivant dans le fichier DB2TX\_TRACEFILE du répertoire DB2TX\_TRACEDIR. Le fichier n'existe pas ou n'a pas été trouvé, ou les droits requis pour y accéder ne sont pas disponibles.

---

### 38702 Impossible d'ouvrir le fichier de messages fichier-messages.

**Explication :** Suite à une erreur, DB2 Extension Texte a tenté de renvoyer un message. Le fichier contenant les messages n'existe pas ou n'a pas été trouvé, ou vous ne disposez pas des droits requis pour y accéder.

**Action :** Assurez-vous que ce fichier existe, que le chemin d'accès est défini correctement et que vous disposez des droits requis pour ouvrir le fichier.

---

### 38704 Le format du descripteur de texte est incorrect.

**Explication :** Un descripteur de format incorrect a été utilisé en tant qu'argument pour une fonction DB2 Extension Texte.

**Action :** Assurez-vous que le descripteur n'a pas été généré par INIT\_TEXT\_HANDLE.

---

### 38705 *nomudf* Déclaration UDF incorrecte.

**Explication :** Le nom spécifique d'une fonction DB2 Extension Texte a été modifié dans le script

## SQLSTATE renvoyés par les fonctions DB2 Extension Texte

où les fonctions sont déclarées. Les noms de fonctions DB2 Extension Texte peuvent être modifiés, mais pas leurs noms spécifiques.

**Action :** Vérifiez le script DESCVDF.DDL qui contient les déclarations de fonction DB2 Extension Texte, afin de vous assurer que les noms corrects sont toujours utilisés. Comparez-les à ceux du support de distribution d'origine.

---

**38706**      *attribut* **Attribut non reconnu.**

**Explication :** Une tentative a été faite pour définir un CCSID, un format ou une langue par une valeur inconnue.

**Action :** Pour connaître les valeurs correctes, reportez-vous au «Chapitre 4. Planification de vos besoins en matière de recherche» à la page 29.

---

**38707**      **La fonction requise n'est pas encore prise en charge.**

**Explication :** La fonction requise n'est pas encore prise en charge.

**Action :** Vérifiez la fonction indiquée.

---

**38708**      *code retour* **Echec de la recherche.**

**Explication :** Une erreur s'est produite lors du traitement de la recherche.

**Action :** Pour obtenir une description du code retour concerné, reportez-vous au «Chapitre 17. Codes retour» à la page 291.

---

**38709**      **Mémoire disponible insuffisante.**

**Explication :** La mémoire disponible est insuffisante pour exécuter la fonction DB2 Extension Texte.

**Action :** Fermez les applications dont vous n'avez pas besoin pour libérer de la mémoire, puis réexécutez la fonction.

---

**38710**      *numéro-erreur* **Accès aux résultats de recherche impossible.**

**Explication :** Une erreur s'est produite lors d'une tentative de lecture de la liste des documents trouvés (liste de résultats) renvoyée par la fonction de recherche.

**Action :** Tentez de relancer la recherche. Si elle échoue, redémarrez la fonction de recherche. Si l'erreur persiste, consultez votre partenaire commercial IBM, en lui indiquant le numéro affiché.

---

**38711**      **Erreur interne grave.**

**Explication :** Une erreur grave s'est produite.

**Action :** Faites part de cette erreur à votre partenaire commercial IBM, en indiquant les circonstances dans lesquelles elle s'est produite.

---

**38712**      *nom-index* **Descripteur de texte incorrect dans cet index.**

**Explication :** Un descripteur a été altéré.

**Action :** Utilisez UPDATE INDEX pour reconstituer l'index.

---

**38714**      **Variable d'environnement abrégée DB2TX\_INSTOWNERHOMEDIR.**

**Explication :** Le nom du répertoire personnel du propriétaire de l'instance ne doit pas dépasser 256 caractères.

**Action :** Utilisez des liens pour réduire la longueur du nom du répertoire.

---

**38717**      **Le thésaurus indiqué est introuvable.**

**Explication :** Le thésaurus indiqué est introuvable.

**Action :** Vérifiez le nom de thésaurus indiqué.

## SQLSTATE renvoyés par les fonctions DB2 Extension Texte

---

**38718**            **Le nom de relation indiqué est introuvable dans le thésaurus.**

**Explication :** La relation indiquée n'existe pas dans le thésaurus spécifié.

**Action :** Vérifiez que la relation indiquée existe.

---

**38719**            **Une erreur s'est produite lors de la recherche. code anomalie : rc.**

**Explication :** La recherche n'a pas pu être effectuée pour la raison indiquée.

**Action :** Essayez de résoudre l'incident signalé par le code anomalie. Si ce code anomalie ne vous aide pas à résoudre l'incident qu'aucune information supplémentaire ne figure dans le fichier `desdiag.log`, créez une trace et transmettez les informations correspondantes à votre partenaire commercial IBM.

---

**38720**            **Une erreur s'est produite lors de la connexion à la mémoire partagée.**

**Explication :** Le système ne peut pas accéder à la mémoire partagée.

**Action :** Vérifiez votre configuration système et augmentez les ressources partagées, ou vérifiez les ressources partagées actuellement utilisées (ipcs) et procédez au nettoyage des ressources qui ne sont plus requises.

---

**38721**            **Une erreur s'est produite lors de la création ou de l'accès au sémaphore.**

**Explication :** Le système ne peut créer un sémaphore ou accéder à un sémaphore.

**Action :** Vérifiez votre configuration système et augmentez les ressources partagées, ou vérifiez les ressources partagées actuellement utilisées (ipcs) et procédez au nettoyage des ressources qui ne sont plus requises.

---

---

**38722**            **Un processus de recherche n'a pas abouti.**

**Explication :** Une erreur s'est produite lors du traitement de la recherche.

**Action :** Contrôlez `descfgcl` dans votre configuration système et vérifiez que tous les noeuds sont installés et opérationnels.

---

**38723**            **Les CCSID d'index et de base de données ne correspondent pas.**

**Explication :** Le CCSID de base de données utilisé pour la chaîne de requête n'est le même que le CCSID d'index de texte.

**Action :** Désactivez l'index de texte et recréez-le avec le CCSID de la base de données.

---

**38724**            **Le nom de modèle ou de section est incorrect.**

**Explication :** Le nom de modèle ou de section indiqué dans la requête est incorrect.

**Action :** Vérifiez le nom de modèle ou de section.

---

**38726**            **Une erreur de lecture du fichier de modèle s'est produite.**

**Explication :** Le fichier de définition de modèle est introuvable ou ne peut être ouvert.

**Action :** Vérifiez que le fichier de définition de modèle existe dans le répertoire d'index.

---

---

**Messages de DB2 Extension Texte**

L'identificateur de chaque message est constitué d'un préfixe (DES), d'un numéro et d'un suffixe (lettre). Le suffixe indique la gravité de l'occurrence qui a généré le message :

- I Message informatif
- W Message d'avertissement
- N Message d'erreur (ou "négatif")
- C Message d'erreur grave

---

**DES0001N Nombre d'arguments incorrect pour la commande db2txinstance.**

**Explication :** La commande db2txinstance exige deux arguments.

**Action :** Entrez-la à nouveau avec les arguments suivants :

db2txinstance nomInstance nomInstancedb2

où *nomInstance* est le nom de connexion d'un utilisateur UNIX existant désigné comme propriétaire de cette instance, et *nomInstancedb2* celui du propriétaire de l'instance DB2 correspondante.

---

**DES0002N NomInstance incorrect.**

**Explication :** Le nom d'instance indiqué doit être le nom de connexion d'un utilisateur UNIX existant.

**Action :** Indiquez un nom d'instance correct, sélectionnez un utilisateur UNIX ou créez-en un qui soit propriétaire de l'instance.

Entrez à nouveau la commande db2txinstance sous la forme suivante :

db2txinstance nomInstance

où *nomInstance* est le nom de connexion de l'utilisateur UNIX sélectionné.

---

**DES0004N L'instance indiquée existe déjà. La commande ne peut pas être traitée.**

**Explication :** *NomInstance* indique le nom de connexion d'un utilisateur UNIX propriétaire de l'instance. Ce dernier dispose déjà d'un répertoire db2tx dans son répertoire personnel.

**Action :** Pour créer l'instance, supprimer celle existante, puis relancez la commande.

---

**DES0005N Le catalogue des messages d'installation est introuvable.**

**Explication :** Le catalogue des messages nécessaires aux scripts d'installation ne se trouve pas sur le système ; il se peut qu'il ait été supprimé ou que les produits de la base de données aient été chargés de manière incorrecte.

**Action :** Vérifiez que l'option de module client db2tx\_01\_01\_0000. est installée correctement. Si vous trouvez des erreurs, réinstallez cette option.

---

**DES0015W La spécification d'un terme de recherche linguistique ne correspond pas au type d'index.**

**Explication :** La spécification linguistique du terme de recherche de la requête ne correspond pas au type d'index. Par exemple, PRECISE FORM OF ne peut pas être utilisé avec un index linguistique. La spécification linguistique par défaut est utilisée comme indiqué dans le tableau 8 à la page 223.

**Action :** Adaptez votre application pour éviter que ne soient définies des options de requête

## Messages de DB2 Extension Texte

provoquant des conflits avec le type d'index.

---

**DES0016W Une langue indiquée pour le type d'index n'est pas prise en charge.**

**Explication :** La langue indiquée n'est pas prise en charge par le type d'index spécifié.

**Action :** Consultez la documentation pour connaître les langues prises en charge pour ce type d'index.

---

**DES0017W La fonction d'extraction n'est pas activée.**

**Explication :** Vous avez utilisé un argument de recherche caractérisée dans votre requête, mais l'index a été créé sans l'option FEATURE\_EXTRACTION.

**Action :** Modifiez l'option d'index de manière à activer FEATURE\_EXTRACTION.

---

**DES0018W option non prise en charge par le type d'index en cours.**

**Explication :** Vous avez introduit une option de recherche qui n'est prise en charge ni pour le type d'index, ni pour l'option d'index.

**Action :** Vérifiez quel type ou option d'index supporte cette option de recherche. Pour plus de détails, reportez-vous au tableau 8 à la page 223.

---

**DES0121N La mémoire n'a pas pu être allouée.**

**Explication :** Aucun espace de mémoire n'a pu être réservé pour l'application.

**Action :** Augmentez l'espace de pagination.

---

**DES0333N Un incident d'entrée-sortie s'est produit au niveau du fichier d'index.**

**Explication :** Le client DB2 Extension Texte ne peut pas établir une session avec le serveur demandé.

**Action :** Vérifiez que le serveur de DB2

Extension Texte a été démarré. Sinon, lancez TXSTART.

---

**DES0709W Le dictionnaire *nom.dictionnaire* n'est pas installé.**

**Explication :** DB2 Extension Texte ne trouve pas les dictionnaires.

**Action :** Installez ou réinstallez le dictionnaire correspondant à la langue indiquée.

---

**DES0377N Un incident d'entrée-sortie s'est produit au niveau du fichier d'index.**

**Explication :** DB2 Extension Texte ne peut pas accéder à l'index de texte. Cet incident peut se produire si la valeur attribuée à DIRECTORY dans la configuration de texte désigne un répertoire non valide.

**Action :** Vérifiez les définitions de la configuration de texte.

---

**DES0700N Le numéro de noeud '*noeud*' ne se trouve pas dans la définition du groupe de noeuds.**

**Explication :** Le numéro de noeud spécifié est incorrect.

**Action :** Vérifiez le numéro de noeud DB2.

---

**DES0701N Le numéro de noeud '*noeud*' est en dehors des limites autorisées.**

**Explication :** Le numéro de noeud spécifié est incorrect.

**Action :** Vérifiez le numéro de noeud DB2.

---

**DES0704N Le format '*format*' requiert la spécification de propriétés d'index.**

**Explication :** Le format du document est incompatible avec les données relatives au type d'index.

**Action :** Spécifiez une propriété d'index compatible avec le format du document.

---

**DES0705N** Le nom de modèle de document indiqué, '*modèle*', est introuvable dans le fichier de définition de modèle.

**Explication :** Le nom du modèle de document est introuvable dans le fichier de définition de modèle. A noter : dans un nom de modèle, les majuscules sont différenciées des minuscules.

**Action :** Utilisez un nom de modèle spécifié dans le fichier de définition de modèles.

---

**DES0706N** L'accès au fichier de définition de modèle est impossible sur le serveur DB2 Extension Texte.

**Explication :** Le fichier de définition de modèle est introuvable ou ne peut être ouvert.

**Action :** Vérifiez l'existence du fichier de définition de modèles.

---

**DES0707N** Le format '*format*' ne prend pas en charge la propriété d'index indiquée.

**Explication :** Le format du document ne prend pas en charge la propriété d'index indiquée.

**Action :** Spécifiez une propriété d'index compatible avec le format du document.

---

**DES0710N** Un pointeur nul n'est pas autorisé pour le paramètre '*paramètre*'.

**Explication :** Aucune valeur n'est spécifiée pour le paramètre *paramètre*.

**Action :** Spécifiez une valeur pour le paramètre.

---

**DES0711N** Une erreur interne de l'Extension Texte s'est produite. Information de diagnostic : *message*.

**Explication :** Une erreur de traitement interne s'est produite.

**Action :** Vérifiez le message de diagnostic pour résoudre le problème. Si l'erreur interne n'est pas due à un problème d'installation, vous trouverez peut-être des informations supplémentaires dans

le fichier *desdiag.log* ou dans un fichier de trace. Si ces données ne vous sont d'aucune aide, rassemblez les informations disponibles et contactez votre partenaire commercial IBM.

---

**DES0712N** Le paramètre '*paramètre*' est trop long.

**Explication :** Le paramètre spécifié dépasse la longueur autorisée.

**Action :** Spécifiez le paramètre en utilisant un nombre de caractères autorisé.

---

**DES0713N** La longueur du paramètre '*paramètre*' est incorrecte : %d1.

**Explication :** Le paramètre spécifié dépasse la longueur autorisée.

**Action :** Spécifiez le paramètre en utilisant un nombre de caractères autorisé.

---

**DES0714N** Indiquez le paramètre *paramètre* directement ou au moyen d'une variable d'environnement.

**Explication :** Le CCSID, le format ou la langue n'ont pas été indiqués, et il n'existe aucune définition de configuration de texte pour cette valeur.

**Action :** Indiquez le paramètre manquant directement dans la commande `ENABLE TEXT COLUMN` ou attribuez-lui une valeur dans les définitions de la configuration de texte.

---

**DES0715N** Le type de données *schéma.type* n'est pas pris en charge pour les données de type texte.

**Explication :** *schéma.type* est le nom de schéma et de type de la colonne de texte ou du résultat d'une fonction d'accès. Le type de données pour une colonne de texte n'est pas pris en charge par DB2 Extension Texte. Il doit s'agir du type `CHAR`, `GRAPHIC`, `VARGRAPHIC`, `LONG VARGRAPHIC`, `DBCLOB`, `VARCHAR`, `LONG VARCHAR` ou `CLOB`. Si tel n'est pas le cas, vous devez fournir une fonction d'accès dont l'entrée est le type de données de la colonne de texte et

## Messages de DB2 Extension Texte

dont la sortie est de type VARCHAR, LONG VARCHAR ou LOB.

**Action :** Si *schéma.type* désigne une colonne de texte, vous devez enregistrer une fonction d'accès avec un résultat de type VARCHAR, LONG VARCHAR ou LOB. Si *schéma.type* est le résultat d'une fonction d'accès, il ne peut être utilisé. Fournissez une fonction d'accès avec un résultat du type requis.

---

**DES0716N Le format *format* n'est pas pris en charge.**

**Explication :** *format* est un format qui n'est pas pris en charge par DB2 Extension Texte.

**Action :** Consultez la liste des formats pris en charge à la section «Documents pris en charge» à la page 31.

---

**DES0717N La langue *langue* n'est pas prise en charge.**

**Explication :** *langue* est une langue qui n'est pas prise en charge par DB2 Extension Texte ou la langue sélectionnée n'est pas prise en charge par le type d'index spécifié.

**Action :** Consultez la liste des langues prises en charge au tableau 5 à la page 54.

---

**DES0718N Le CCSID *valeur\_ccsid* n'est pas pris en charge.**

**Explication :** La valeur indiquée pour le CCSID est incorrecte.

**Action :** Consultez la documentation pour connaître la liste des CCSID pris en charge.

---

**DES0719N Un appel du programme DB2 Extension Texte *programme* a échoué avec un code retour *rc*.**

**Explication :** Une erreur s'est peut-être produite pendant l'installation. Les codes retour sont répertoriés dans le fichier DES\_EXT.H.

**Action :** Vérifiez que l'installation s'est effectuée correctement. Vérifiez que les variables d'environnement DB2TX\_INSTOWNER et

DB2TX\_INSTOWNERHOMEDIR sont définies correctement.

---

**DES0720N La fonction d'accès *schéma.fonction* n'est pas enregistrée dans la base de données.**

**Explication :** Le nom de cette fonction est incorrect ou n'a pas été enregistré dans la base de données.

**Action :** Vérifiez le nom de la fonction d'accès. S'il est correct, vérifiez que cette fonction est connue du système de base de données. Utilisez CREATE FUNCTION pour enregistrer la fonction d'accès dans la base de données.

---

**DES0721N La base de données est incohérente. Une vue de catalogue de DB2 Extension Texte manque.**

**Explication :** L'une des vues de catalogue de DB2 Extension Texte ne se trouve pas dans la base de données.

**Action :** Utilisez la commande DISABLE DATABASE pour supprimer les vues de catalogue restantes, puis entrez de nouveau ENABLE DATABASE. Les données de l'index sont perdues. Réindexez les documents texte.

---

**DES0722N La table *schéma.table* n'est pas une table de base dans la base de données.**

**Explication :** Cette table n'existe pas dans la base de données ou est une table de résultats ou une vue. Une colonne de texte doit se trouver dans une table de base avant de pouvoir être activée pour DB2 Extension Texte.

**Action :** Assurez-vous que le nom de la table est correct et qu'il s'agit d'une table de base.

---

**DES0723N La création d'un index pour la colonne de descripteurs *nom-colonne-descripteurs* dans la table *schéma.table* a échoué.**

**Explication :** Aucun index de texte n'a pu être créé pour la colonne de descripteurs.

**Action :** Utilisez `txstatus` pour vérifier l'état du serveur. Si les fonctions sur le serveur s'exécutent correctement, utilisez `DISABLE TEXT COLUMN` ou `DISABLE TEXT TABLE` pour obtenir à nouveau un état cohérent. Puis réactivez la colonne à l'aide de `ENABLE TEXT COLUMN` ou `ENABLE TEXT TABLE`.

---

**DES0724N** Une entrée manque dans la vue de catalogue TextIndices pour la colonne de descripteurs *colonne-descripteurs* dans la table *schéma.table*.

**Explication :** La vue de catalogue TextIndices est altérée.

**Action :** Utilisez `DISABLE TEXT COLUMN` ou `DISABLE TEXT TABLE` pour obtenir à nouveau un état cohérent. Puis réactivez la colonne de texte à l'aide de `ENABLE TEXT COLUMN` ou `ENABLE TEXT TABLE`.

---

**DES0727N** La colonne *colonne* dans la table *schéma.table* est déjà active.

**Explication :** Ce message s'affiche si la table a été supprimée puis recrée à l'aide de la même colonne de texte, sans que cette dernière soit d'abord désactivée.

**Action :** Désactivez cette colonne, puis recommencez l'opération.

---

**DES0728N** La colonne *colonne* n'existe pas dans la table *schéma.table*.

**Explication :** La colonne de texte que vous tentez d'activer n'existe pas.

**Action :** Indiquez un autre nom de table ou de colonne, puis recommencez cette opération.

---

**DES0729N** La colonne de descripteurs *colonne\_descripteurs* n'existe pas dans la table *schéma.table*.

**Explication :** La colonne de descripteurs que vous tentez d'utiliser n'existe pas.

**Action :** Utilisez la commande `GET STATUS` pour vérifier si cette colonne de descripteurs

existe et que son nom a été indiqué correctement.

---

**DES0730N** La table *schéma.table* est déjà active en tant que table à index commun.

**Explication :** La table que vous tentez d'activer l'a déjà été en tant que table à index commun.

**Action :** Continuez sans activer cette table ou exécutez la commande `DISABLE TEXT TABLE` pour désactiver la table avant de la réactiver.

---

**DES0731N** La table *schéma.table* n'est pas active pour DB2 Extension Texte. Elle ne peut donc pas être désactivée.

**Explication :** La table que vous tentez de désactiver n'a pas été activée.

**Action :** Vérifiez le nom de la table.

---

**DES0732N** La fréquence de mise à jour est incorrecte en position *emplacement* dans la commande. La fréquence attendue était *paramètre*.

**Explication :** La spécification de *paramètre* pour la fréquence de mise à jour n'était pas correcte.

**Action :** Vérifiez le paramètre de la fréquence de mise à jour et entrez à nouveau la commande.

---

**DES0733N** La table *schéma.table* contient une colonne activée. Elle ne peut donc pas être activée comme table à index commun.

**Explication :** Cette table contient une colonne de texte qui possède déjà son propre index. Il est impossible de créer un index commun pour toutes les colonnes de texte tant que cet index existe.

**Action :** Utilisez `DISABLE TEXT COLUMN` pour désactiver les colonnes activées, puis entrez à nouveau la commande `ENABLE TEXT TABLE`.

## Messages de DB2 Extension Texte

---

**DES0734N** La colonne de descripteurs *colonne-descripteurs* appartient à la table à index commun *schéma.table*. Elle ne peut pas être désactivée séparément.

**Explication :** Il est impossible de désactiver une colonne de texte distincte dans une table qui a été activée en tant que table de type texte partiel.

**Action :** Désactivez la table de type texte partiel.

---

**DES0736N** *colonne-descripteurs* est déjà utilisée comme colonne de descripteurs dans la table *schéma.table*.

**Explication :** Le nom de la colonne de descripteurs que vous tentez d'utiliser existe déjà.

**Action :** Entrez la commande à nouveau, en indiquant un autre nom de colonne de descripteurs.

---

**DES0737N** La table *schéma.table* est active en tant que table à index commun avec l'option `STORAGE option`.

**Explication :** Il est impossible d'activer une table à index commun pour des fichiers externes.

**Action :** Si vous voulez activer une table pour des fichiers externes, utilisez une table à index multiples.

---

**DES0738N** La fonction d'accès *schéma.fonction* contient des paramètres incorrects.

**Explication :** Les paramètres d'entrée ou de sortie de *schéma.fonction* sont incorrects.

- Il ne peut exister qu'un seul paramètre d'entrée qui doit être du même type de données que celui de la colonne de texte à activer.
  - Le paramètre de sortie doit être de type CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR ou CLOB.
- 

**DES0739W** Le programme de mise à jour d'index pour la table *schéma.table*, colonne de descripteurs *colonne-descripteurs*, ne peut pas être lancé.

**Explication :** Le programme qui met les index à jour ne peut pas être lancé. Une erreur s'est peut-être produite pendant l'installation.

**Action :** Vérifiez que l'installation s'est effectuée correctement. Vérifiez que les variables d'environnement `DB2TX_INSTOWNER` et `DB2TX_INSTOWNERHOMEDIR` sont définies correctement.

---

**DES0740N** La colonne de descripteurs '*colonne*' ne peut être réutilisée que pour une colonne texte de type *type-colonne*.

**Explication :** La colonne de descripteurs a déjà été utilisée pour un autre type de colonne de descripteurs.

**Action :** Spécifiez un nouveau nom de colonne de descripteurs.

---

**DES0741N** Le programme ou le fichier *paramètre* est introuvable ou ne peut pas être lancé.

**Explication :** La commande `ENABLE DATABASE` ou `DISABLE DATABASE` ne peut pas ouvrir le fichier *paramètre*. Une erreur s'est peut-être produite pendant l'installation.

**Action :** Vérifiez que l'installation s'est effectuée correctement.

---

**DES0742N** Cette table utilise des index de colonnes. Définissez une colonne de descripteurs dans la commande.

**Explication :** La table spécifiée est activée en tant que table à index multiples. Pour travailler avec une colonne particulière, spécifiez la colonne de descripteurs associée.

**Action :** Spécifiez une colonne de descripteurs.

---

**DES0745N** Le propriétaire de l'instance DB2TX *propriétaire-instance* n'a pas un ID utilisateur correct.

**Explication :** La variable d'environnement DB2TX\_INSTOWNER ne contient pas d'ID utilisateur correct.

**Action :** Corrigez la variable d'environnement.

---

**DES0747N** Le CCSID en cours n'est pas pris en charge par le type d'index *type\_index*.

**Explication :** Vous avez indiqué un CCSID non pris en charge par le type d'index demandé.

**Action :** Consultez la documentation pour connaître la liste des CCSID pris en charge.

---

**DES0748N** La valeur du nombre de COMMIT '*nombre-validations*' n'est pas prise en charge par DB2 Extension Texte.

**Explication :** Le nombre de validations (commit) spécifié n'est pas pris en charge.

**Action :** Spécifiez un nombre de validations (commit) correct.

---

**DES0749N** La valeur de mise à jour d'index '*valeur-index*' est inconnue de DB2 Extension Texte.

**Explication :** La valeur de 'de mise à jour d'index' est incorrecte.

**Action :** Indiquez une valeur de mise à jour d'index correcte.

---

**DES0750N** La valeur de type d'index '*type-index*' est inconnue de DB2 Extension texte.

**Explication :** La valeur de type d'index indiquée est incorrecte.

**Action :** Spécifiez une valeur de type d'index correcte.

---

**DES0751N** Vous ne disposez pas de droits pour exécuter l'opération indiquée.

**Explication :** Vous ne disposez du droit d'administrateur de bases de données requis pour effectuer cette opération.

**Action :** Demandez à un administrateur de bases de données d'effectuer cette opération.

---

**DES0756N** La base de données n'est pas active pour DB2 Extension Texte.

**Explication :** La base de données doit être activée avant de pouvoir exécuter cette commande.

**Action :** Exécutez ENABLE DATABASE, puis soumettez à nouveau cette commande.

---

**DES0763N** La fréquence de mise à jour est incorrecte.

**Explication :** La valeur de 'fréquence de mise à jour' indiquée est incorrecte.

**Action :** Indiquez une 'fréquence de mise à jour' correcte.

---

**DES0765N** La base de données est déjà active pour DB2 Extension Texte.

**Explication :** La base de données que vous tentez d'activer est déjà active.

**Action :** Continuez sans activer la base de données ou utilisez DISABLE DATABASE pour la désactiver avant de l'activer à nouveau.

---

**DES0766N** Une opération a entraîné un dépassement de la taille maximale de ligne autorisée ou d'une table temporaire.

**Explication :** La commande ENABLE TEXT COLUMN ajoute une colonne de descripteurs à la table. Si celle-ci est déjà de grande taille, cela risque d'entraîner le dépassement de la valeur maximale de ligne autorisée, 4005.

La commande ENABLE TEXT COLUMN crée

## Messages de DB2 Extension Texte

également une table temporaire dont la taille est proportionnelle au nombre de colonnes de texte déjà activées. Si de nombreuses colonnes de texte sont déjà activées, la taille de la table temporaire risque de dépasser la valeur maximale.

**Action :** Utilisez la commande ENABLE TEXT COLUMN uniquement sur les tables qui n'entraînent pas le dépassement de cette limite.

---

**DES0769W Avertissement : Les caractéristiques de l'index ont été indiquées mais seront ignorées. La table '*nom-table*' est une table à index commun.**

**Explication :** La table spécifiée est une table à index commun, par conséquent, vous ne pouvez pas indiquer de caractéristiques.

**Action :** Aucune action n'est requise.

---

**DES0770N La variable d'environnement *variable-env* n'est pas définie.**

**Explication :** Un paramètre de commande n'a pas été indiqué et le système a tenté de lire la valeur par défaut à partir de la variable d'environnement *variable-env*, mais cette dernière n'est pas définie.

**Action :** Définissez la variable d'environnement concernée.

---

**DES0774N La longueur de valeur *longueur* pour la variable '*variable*' dépasse la longueur autorisée.**

**Explication :** La longueur de valeur du paramètre dépasse la longueur autorisée.

**Action :** Spécifiez le paramètre en utilisant un nombre de caractères autorisé.

---

**DES0775N La valeur de répertoire d'index '*répertoire*' est incorrecte.**

**Explication :** La valeur de répertoire d'index est incorrecte ; c'est peut-être la longueur du répertoire qui ne convient pas.

**Action :** Spécifiez une valeur de répertoire d'index correcte.

---

**DES0776N Le nom d'espace table '*espace-table*' est incorrect.**

**Explication :** Le nom d'espace table spécifié est incorrect ; c'est peut-être la longueur qui ne convient pas.

**Action :** Spécifiez une valeur d'espace table correcte.

---

**DES0777N Le nom d'espace table *espace-table* est inconnu du système de gestion de bases de données.**

**Explication :** L'espace table spécifié n'est pas reconnu par le système de bases de données.

**Action :** Vérifiez que l'espace table spécifié existe dans la base de données.

---

**DES0778N '*espace-table*' n'est pas un espace table de type REGULAR. Il a été créé avec le mot clé '*mot-clé*'.**

**Explication :** Le type de données n'est pas pris en charge pour l'espace table spécifié.

**Action :** Spécifiez un espace table normal.

---

**DES0779I L'indexation a bien démarré. Pour vérifier l'avancement de cette opération, utilisez la commande 'GET INDEX STATUS'.**

**Explication :** Le programme d'indexation a démarré. Vous pouvez utiliser la commande 'GET INDEX STATUS' pour suivre son déroulement.

**Action :** Vérifiez le résultat de l'exécution de la commande GET INDEX STATUS.

---

**DES0780I** La réorganisation d'index a débuté correctement. Pour vérifier l'avancement de cette opération, utilisez la commande 'GET INDEX STATUS'.

**Explication :** Le programme de réorganisation a démarré. Vous pouvez utiliser la commande 'GET INDEX STATUS' pour suivre son déroulement.

**Action :** Vérifiez le résultat de l'exécution de la commande GET INDEX STATUS.

---

**DES0789W** Avertissement : La mappe de partitionnement du groupe de noeuds en cours a été mise à jour. Veuillez appeler l'utilitaire TXNCHECK.

**Explication :** La mappe de partitionnement du groupe de noeuds en cours a été mise à jour.

**Action :** Utilisez la commande TXNCHECK.

---

**DES0800I** La commande *commande* a abouti.

**Explication :** La commande spécifiée a abouti.

**Action :** Aucune action n'est requise.

---

**DES0810N** Les guillemets de fermeture manquent.

**Explication :** Les guillemets d'ouverture existent, mais les guillemets de fermeture manquent.

**Action :** Vérifiez la syntaxe de la commande et relancez-la.

---

**DES0811N** "*marque*" inattendue. Le système attend un mot clé valable.

**Explication :** Les caractéristiques de l'index ou les informations relatives au texte sont incorrectes.

**Action :** Vérifiez la syntaxe et relancez la commande.

---

---

**DES0812N** La table *schéma.table* n'existe pas ou n'est pas active pour DB2 Extension Texte.

**Explication :** Au cours de l'exécution de la commande GET, le nom d'une table de base de données est incorrect ou la table n'existe pas, ou elle n'est pas encore activée.

**Action :** Si le nom de la table est correct, utilisez GET STATUS pour vérifier si elle est déjà active. Activez la table et retentez l'opération.

---

**DES0813N** La table *schéma.table* n'existe pas, n'est pas active pour DB2 Extension Texte ou ne contient pas de colonne de descripteurs *colonne*.

**Explication :** Au cours de l'exécution de la commande GET, aucune entrée pour la colonne de descripteurs n'est trouvée dans la table. Si la table existe, elle n'est pas active ou ne contient pas de colonne de descripteurs.

**Action :** Si le nom de la table est correct, utilisez GET STATUS pour vérifier si elle est déjà active. Activez la table et retentez l'opération.

---

**DES0814N** La table *nom-table* n'existe pas, n'est pas active pour DB2 Extension Texte ou aucune colonne de texte n'est active dans cette table.

**Explication :** La table spécifiée n'est pas active pour DB2 Extension Texte.

**Action :** Activez la table pour DB2 Extension Texte.

---

**DES0815N** Des guillemets vides (") ont été détectés. Un nom doit être indiqué entre ces guillemets.

**Explication :** Deux paires de guillemets consécutives sans texte ont été détectées.

**Action :** Vérifiez la syntaxe et relancez la commande.

---

## Messages de DB2 Extension Texte

---

**DES0816N** Le mot "*marque*" n'est pas le mot attendu. Utilisez le mot clé *mot-clé* ou *mot-clé*.

**Explication** : Une marque inattendue a été détectée.

**Action** : Dans la commande, utilisez l'un des mots clés indiqué dans le message.

---

**DES0817N** "*marque*" inattendue. Utilisez le mot clé *mot-clé*.

**Explication** : Une marque inattendue a été détectée.

**Action** : Dans la commande, utilisez le mot clé indiqué dans le message.

---

**DES0818N** Fin de commande inattendue. Le mot clé *mot-clé* est requis.

**Explication** : Un mot clé manque.

**Action** : Dans la commande, utilisez le mot clé indiqué dans le message.

---

**DES0819N** Fin de commande inattendue. L'un des mots clés suivants doit être indiqué : *mot-clé* ou *mot-clé*.

**Explication** : Un mot clé manque.

**Action** : Dans la commande, utilisez l'un des mots clés indiqué dans le message.

---

**DES0820N** L'option d'index *option* n'est pas prise en charge pour le type d'index *type\_index*.

**Explication** : Vous avez indiqué une option d'index non prise en charge pour le type d'index demandé.

**Action** : Reportez-vous à la documentation pour connaître les option d'index prises en charge avec un type d'index donné.

---

**DES0821N** Le nom "*marque*" est trop long. *nn* caractères seulement sont autorisés dans les noms *variable*.

**Explication** : Un nom est trop long.

**Action** : Indiquez un nom ayant une longueur autorisée.

---

**DES0822N** La commande contient une marque "*marque*" non reconnue. La fin de la commande est attendue.

**Explication** : La fin de la commande est détectée, mais un mot clé est attendu.

**Action** : Vérifiez la syntaxe de la commande et relancez-la.

---

**DES0823N** Un nom de table doit être indiqué après "*schéma*".

**Explication** : Il manque un nom de table ou de fonction après le ".".

**Action** : Vérifiez la syntaxe de la commande et relancez-la.

---

**DES0824N** Fin de la commande inattendue. Une valeur *mot-clé* doit être indiquée.

**Explication** : Le mot clé indiqué dans le message manque dans la syntaxe.

**Action** : Vérifiez la syntaxe de la commande et relancez-la.

---

**DES0825N** '*alias*' n'est pas un nom d'alias de base de données connu. *nn* caractères seulement sont autorisés.

**Explication** : L'alias de base de données spécifié n'est pas reconnu par le système de bases de données.

**Action** : Vérifiez que l'alias de base de données spécifié est correct.

---

**DES0826N** L'alias de base de données *alias* ne doit pas être indiqué entre guillemets.

**Explication :** Le nom indiqué dans le message a été interprété comme un alias de base de données. Il ne doit pas être indiqué entre guillemets.

**Action :** Vérifiez la syntaxe de la commande et relancez-la.

---

**DES0827N** Le CCSID "*ccsid*" n'est pas pris en charge.

**Explication :** Le CCSID ne fait pas partie de ceux pris en charge par DB2 Extension Texte.

**Action :** Pour obtenir une liste des CCSID pris en charge, reportez-vous à la documentation.

---

**DES0829N** L'ID utilisateur *IDutil* ne doit pas être indiqué entre guillemets.

**Explication :** Vous avez entré l'ID utilisateur entre guillemets.

**Action :** Supprimez les guillemets.

---

**DES0830N** Le paramètre "*paramètre*" dans la commande *enable/disable DATABASE* n'est pas reconnu. La commande va être interrompue.

**Explication :** Les commandes ENABLE DATABASE et DISABLE DATABASE n'utilisent pas de paramètres.

**Action :** Entrez la commande à nouveau sans indiquer de paramètres.

---

**DES0831N** Fin de commande inattendue. Il manque le nom de la table.

**Explication :** Cette commande requiert un nom de table.

**Action :** Entrez le nom de table approprié.

---

**DES0832N** Fin de commande inattendue. Il manque le nom de la base de données.

**Explication :** Cette commande requiert un nom de base de données.

**Action :** Entrez le nom de base de données approprié.

---

**DES0833N** Fin de commande inattendue. Il manque le nom de la colonne.

**Explication :** Cette commande requiert un nom de colonne.

**Action :** Entrez le nom de colonne approprié.

---

**DES0899N** Commande DB2TX inconnue : *commande*.

**Explication :** La commande spécifiée n'est pas prise en charge par DB2 Extension Texte

**Action :** Tapez **db2tx ?** pour obtenir une liste des commandes.

---

**DES0971N** Le répertoire d'index peut être indiqué une fois seul ou plusieurs fois accompagné du noeud.

**Explication :** La spécification du ou des répertoires d'index est incorrecte.

**Action :** Vérifiez la spécification du répertoire d'index. Vous pouvez spécifier un répertoire d'index sans indiquer de noeuds ou spécifier de multiples répertoires d'index en indiquant un noeud.

---

**DES0972N** Le noeud indiqué est incorrect. Une valeur numérique sans signe est attendue.

**Explication :** Une valeur non numérique est spécifiée pour le numéro de noeud.

**Action :** Spécifiez une valeur non signée pour le numéro de noeud.

## Messages de DB2 Extension Texte

---

**DES0973N** Le noeud indiqué est incorrect. Un ou plusieurs caractères inattendus ont été trouvés avant la parenthèse droite.

**Explication** : La spécification de noeud est incorrecte d'un point de vue syntaxique.

**Action** : Vérifiez la syntaxe et essayez de nouveau.

---

**DES0974N** La séquence de numéros de noeuds dans une clause TO est incorrecte (second noeud inférieur au premier).

**Explication** : La spécification de noeud est incorrecte d'un point de vue syntaxique.

**Action** : Vérifiez la syntaxe et essayez de nouveau.

---

**DES0975N** La syntaxe de la spécification des informations relatives au noeud est incorrecte.

**Explication** : La spécification de noeud est incorrecte d'un point de vue syntaxique.

**Action** : Vérifiez la syntaxe et essayez de nouveau.

---

**DES0976N** Les informations relatives au noeud sont incomplètes.

**Explication** : Les données relatives au noeud sont incomplètes ; il manque des informations.

**Action** : Vérifiez la syntaxe et essayez de nouveau.

---

**DES0977N** Les informations relatives au noeud sont incomplètes. Il manque la parenthèse gauche.

**Explication** : Les données relatives au noeud sont incomplètes ; il manque la parenthèse ouvrante.

**Action** : Corrigez la spécification de noeud et essayez de nouveau.

---

**DES0998N** La longueur du nom du modèle de document est incorrecte.

**Explication** : La longueur du 'nom du modèle de document' est incorrecte.

**Action** : Vérifiez la valeur du nom du modèle et essayez de nouveau.

---

**DES0999N** La syntaxe de spécification du/des modèle(s) de document est incorrecte.

**Explication** : La spécification du ou des noms de modèle est incorrecte d'un point de vue syntaxique.

**Action** : Vérifiez la syntaxe et essayez de nouveau.

---

**DES9994N** Une erreur du moteur de recherche de texte s'est produite.  
**Code anomalie** : code-anomalie

**Explication** : Le moteur de recherche de texte utilisé par DB2 Extension Texte a signalé une erreur.

**Action** : Vérifiez le code anomalie du moteur de recherche (reportez-vous au «Chapitre 19. Codes anomalie du moteur de recherche» à la page 317). Si le code anomalie indiqué ne vous aide pas à résoudre le problème, vous trouverez peut-être des informations complémentaires dans le fichier desdiag.log ou dans un fichier de trace. Si ces données ne vous sont d'aucune aide, rassemblez les informations disponibles et contactez votre partenaire commercial IBM.

---

**DES9995N** Une erreur s'est produite dans DB2 Extension Texte. Texte du message : texte de message

**Explication** : Une erreur DB2 Extension Texte s'est produite.

**Action** : Utilisez le message fourni par DB2 Extension Texte pour résoudre le problème. Si le message indiqué ne vous aide pas à résoudre le problème, vous trouverez peut-être des informations complémentaires dans le fichier desdiag.log ou dans un fichier trace. Si ces

données ne vous sont d'aucune aide, rassemblez les informations disponibles et contactez votre partenaire commercial IBM.

---

**DES9996N Une erreur interne DB2 Extension Texte s'est produite. Code anomalie : *code\_anomalie***

**Explication :** Une erreur de traitement interne s'est produite.

**Action :** Vérifiez que l'installation de DB2 Extension Texte a effectivement abouti dans son intégralité. Si oui, notez le code anomalie et contactez votre partenaire commercial IBM.

---

**DES9997N Une erreur SQL s'est produite. SqlState : *état* ; code erreur SQL : *rc* ; message d'erreur SQL : *message***

**Explication :** Une erreur SQL s'est produite.

**Action :** Pour corriger cette erreur, consultez le message d'erreur SQL affiché avec le message.

---

**DES9998N Une erreur SQL s'est produite. Aucune information supplémentaire n'est disponible.**

---

**DES9999N Pas de message d'erreur correspondant.**

**Explication :** Une erreur de traitement interne s'est produite.

**Action :** Vérifiez le message de diagnostic pour résoudre le problème. Si l'erreur interne n'est pas due à un problème d'installation, vous trouverez peut-être des informations supplémentaires dans le fichier *desdiag.log* ou dans un fichier de trace. Si ces données ne vous sont d'aucune aide, rassemblez les informations disponibles et contactez votre partenaire commercial IBM.



---

## Chapitre 19. Codes anomalie du moteur de recherche

Ce chapitre répertorie les codes anomalie renvoyés par le moteur de recherche de DB2 Extension Texte.

Tableau 21. Codes anomalie du moteur de recherche

Codes anomalie	Valeurs
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

## Codes anomalie du moteur de recherche

Tableau 21. Codes anomalie du moteur de recherche (suite)

Codes anomalie	Valeurs
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

## Codes anomalie du moteur de recherche

Tableau 21. Codes anomalie du moteur de recherche (suite)

Codes anomalie	Valeurs
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

## Codes anomalie du moteur de recherche

---

## Chapitre 20. Codes anomalie dus à des événements d'erreur

Ce chapitre répertorie les événements d'erreur susceptibles de survenir lorsque DB2 Extension Texte indexe des documents. Ces événements peuvent se produire, par exemple, lorsque :

- des documents n'ont pu être indexés ;
- des documents sont indexés, mais qu'une erreur survient ;
- un dictionnaire n'a pu être trouvé pour une langue donnée.

### Conseil

Si un code anomalie n'est pas documenté :

1. Vérifiez que l'espace disque disponible est suffisant.
2. Collectez toutes les informations d'erreur disponibles :
  - Fichier desdiag.log
  - Message d'événement
3. Contactez votre technicien de maintenance IBM.

1 Mémoire saturée sur le serveur : réduisez la charge de travail.

**116**

Erreur de syntaxe de flux de données.

**280**

Le document n'a pas été indexé. Impossible d'ouvrir l'un des fichiers d'index.

**281**

Le document n'a pas été indexé. Impossible de lire l'un des fichiers d'index.

**441**

Le document n'a pas été indexé. Ce message est généré pour les index de type Ngram uniquement. La page de codes du document est différente de celle utilisée pour la création de l'index. Cela peut se produire pour des documents HTML et XML si l'index n'a pas été créé en page de codes UTF8.

**500**

Le document n'a pas été indexé. Impossible de charger les Fonctions de bibliothèque. Assurez-vous que la DLL est disponible et que le chemin d'accès de la ressource est correct.

## Codes anomalie dus à des événements d'erreur

**501**

Le document n'a pas été indexé. Echec de Lib\_Init dans les Fonctions de bibliothèque. Sur les systèmes de fichiers plats : fichier DIT introuvable ou situé dans un répertoire inapproprié, ou contenu de DIT incorrect.

**502**

Le document n'a pas été indexé. Une erreur s'est produite lors de la lecture du contenu du document dans la fonction de bibliothèque LIB\_read\_doc\_content.

**503**

Le document n'a pas été indexé. Une erreur s'est produite dans la fonction de bibliothèque LIB\_access\_doc.

**504**

Le document n'a pas été indexé. La fonction de bibliothèque LIB\_doc\_index\_status a renvoyé une erreur.

**505**

Echec de la fermeture du document. La fonction de bibliothèque LIB\_close\_doc a renvoyé une erreur.

**506**

Echec lors de l'arrêt des Fonctions de bibliothèque. La fonction LIB\_end a renvoyé une erreur.

**507**

L'appel de fonction de bibliothèque LIB\_read\_doc\_content a échoué avec un code retour inattendu.

**508**

L'appel de fonction de bibliothèque LIB\_close\_doc a renvoyé l'erreur RC\_TERMINATION.

**545**

Le document n'a pas été indexé. Impossible d'ouvrir l'un des fichiers d'index temporaires.

**546**

Le document n'a pas été indexé. Impossible de fermer l'un des fichiers d'index temporaires.

**548**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**549, 550**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail.

**551-564**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer

## Codes anomalie dus à des événements d'erreur

l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**565**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail.

**566-587**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index.

**588-590**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**591-604**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index.

**605**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail.

**606-623**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**624**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail.

**625-631**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**632**

Impossible d'ouvrir l'un des fichiers temporaires créés lors de l'indexation avec des droits d'accès en écriture. Vérifiez les droits d'accès.

**633**

Impossible de fermer l'un des fichiers temporaires créés lors de l'indexation.

**634**

Impossible d'écrire dans l'un des fichiers temporaires créés lors de l'indexation. Vérifiez que l'espace disponible dans le répertoire de travail de l'index est suffisant.

**635**

Impossible de lire l'un des fichiers temporaires créés lors de l'indexation.

## Codes anomalie dus à des événements d'erreur

- 636**  
Impossible d'ouvrir l'un des fichiers temporaires créés lors de l'indexation avec des droits d'accès en lecture. Vérifiez les droits d'accès.
- 659**  
Impossible d'ouvrir l'un des fichiers temporaires créés lors de l'indexation.
- 660**  
Impossible d'écrire dans l'un des fichiers temporaires créés lors de l'indexation.
- 661**  
Impossible de fermer l'un des fichiers temporaires créés lors de l'indexation.
- 662**  
Impossible d'ouvrir l'un des fichiers temporaires créés lors de l'indexation.
- 663**  
Impossible d'écrire dans l'un des fichiers temporaires créés lors de l'indexation.
- 664**  
Impossible de fermer l'un des fichiers temporaires créés lors de l'indexation.
- 665**  
Impossible d'ouvrir l'un des fichiers temporaires créés lors de l'indexation.
- 667**  
Impossible d'écrire dans l'un des fichiers temporaires créés lors de l'indexation.
- 668-669**  
Le document n'a pas été indexé. Un problème de correspondance est survenu entre les marques de sections de documents et celles définies dans le fichier de modèles de documents.
- 670-672**  
Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.
- 673**  
Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail.

**674**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**675-687**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**688, 690**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail. Indiquez des valeurs inférieures dans le fichier de configuration.

**689**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**691-695**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**696-707**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**708**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail. Indiquez des valeurs inférieures dans le fichier de configuration.

**709-718**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**719-721**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**722, 729**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail. Indiquez des valeurs inférieures dans le fichier de configuration.

**730, 732, 733, 735-738**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

## Codes anomalie dus à des événements d'erreur

**731, 739-742, 744-746, 749, 755-758, 760-761, 767**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**743, 748, 750-754, 759, 765-766, 768-770**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**747, 763, 764**

Mémoire insuffisante (échec d'allocation) sur le serveur : réduisez la charge de travail. Indiquez des valeurs inférieures dans le fichier de configuration.

**815**

Deux causes sont possibles :

- L'un des fichiers de ressources requis pour le support de langue utilisé pour le document entraînant l'erreur est introuvable.
- La langue requise par ce document n'est pas prise en charge par DB2 Extension Texte.

**831**

Le document n'a pas été indexé. Aucun texte n'a pu être trouvé. La taille du document est de 0 octet.

**860**

Erreur d'ouverture de fichier. Des dictionnaires ou thésaurus sont introuvables. Vérifiez le chemin d'accès des dictionnaires. Si vous avez indiqué l'emplacement des thésaurus lors de la recherche, vérifiez ces informations ainsi que le nom des fichiers.

**954-956**

Erreur interne. Transmettez les informations consignées dans le journal de diagnostic à votre partenaire commercial IBM.

**957-967**

Le document n'a pas été indexé. Impossible d'ouvrir, lire, éditer ou fermer l'un des fichiers d'index. Vérifiez que l'espace disque disponible est suffisant pour l'index et que les droits d'accès sont corrects.

**1000**

Une erreur est survenue à l'ouverture du fichier. Vérifiez les droits d'accès.

**1001**

Une erreur est survenue lors de l'ajout de données dans le fichier. Vérifiez les droits d'accès.

**1002**

Une erreur est survenue à la lecture du fichier. Il est peut-être altéré.

## Codes anomalie dus à des événements d'erreur

### 1003

Une erreur est survenue lors de l'écriture dans le fichier. Vérifiez l'espace disque disponible et les droits d'accès.

### 1005

Une erreur est survenue à la lecture du fichier (positionnement dans le fichier). Il est peut-être altéré.

### 1006

Une erreur est survenue lors du changement de nom d'un fichier temporaire. Vérifiez les droits d'accès.

### 1007

Une erreur est survenue à la création du fichier. Vérifiez les droits d'accès.

### 1008

Une erreur est survenue à la compression du fichier. Vérifiez les droits d'accès.

### 1009

Une erreur est survenue à la fermeture du fichier. Il est peut-être altéré.

### 1010

Le nom d'index spécifié est déjà utilisé. Choisissez-en un autre.

### 1011

Le chemin d'accès spécifié est déjà utilisé. Choisissez un autre emplacement.

### 1012

Le même chemin d'accès est utilisé pour les répertoires de données et de travail. Choisissez un autre emplacement.

### 1013

Le nom d'index spécifié est incorrect. Il doit comporter des lettres majuscules ou des chiffres (8 caractères au maximum).

### 1014

Une erreur est survenue à la copie du fichier. Vérifiez les droits d'accès et l'espace disque.

### 1017

Nom d'index inconnu. Vérifiez son orthographe.

### 1019

Une erreur est survenue à la suppression du fichier. Vérifiez les droits d'accès. Ce message d'erreur peut être généré en tant que "message d'erreur secondaire". Consultez le fichier de diagnostic pour voir si une erreur précédemment consignée contient des informations supplémentaires.

## Codes anomalie dus à des événements d'erreur

**1020**

Erreur générale de fichier. Vérifiez les droits d'accès.

**1070-1074**

Le document n'a pas été indexé. La page de codes spécifiée est incorrecte de manière générale, ou incorrecte pour l'index utilisé.

**1085**

Le document n'a pas été indexé. Une erreur est survenue à la lecture de la file d'attente d'index.

**1086**

Le document n'a pas été indexé. La file d'attente d'index est vide.

**1116-1117**

Le document n'a pas été indexé. Les informations issues du fichier d'initialisation de l'instance serveur n'ont pas pu être traitées. Assurez-vous que les entrées contenues dans ce fichier sont correctes et que ce dernier est accessible par l'application.

**1129**

Aucun document n'a été indexé. Echec du démarrage du traitement en arrière-plan.

**1158**

Une erreur s'est produite lors du changement de nom d'un fichier. Vérifiez les droits d'accès et l'espace disque.

**1162**

Les fichiers d'indexation d'un index de type Ngram sont peut-être altérés.

**1163, 1164**

Le document n'a pas été indexé en raison d'une erreur inattendue.

**1165**

Le document n'a pas été indexé en raison d'une condition de fin de fichier inattendue.

**1176**

Aucun document supplémentaire ne peut être indexé pour un index de type Ngram. Une condition de dépassement s'est produite pour les numéros de document (dépassement de type long). Si de nombreuses suppressions ou des mises à jour successives d'un même document ont eu lieu, essayez d'appeler EhwrReorg pour résoudre l'incident. Sinon, vous pouvez envisager d'utiliser un second index pour les documents nouveaux.

**1177**

Le document n'a pas été indexé. Il a été considéré comme étant trop long par le programme d'indexation Ngram.

### 1189

Le document n'a pas été indexé. Un incident lié à une séquence de limites s'est produit (concerne uniquement le Coréen).

### 1198-1200

Le document n'a pas été indexé. Un incident lié à l'accès aux index s'est produit. L'index est peut-être altéré.

### 1201

Le document n'a pas été indexé. La page de codes du document n'a pas pu être convertie dans celle de l'index. Cette erreur ne concerne que les index de type Ngram dont la page de codes est UTF8.

### 1202

Le document n'a pas été indexé. La page de codes du document n'a pas pu être convertie dans celle de l'index en raison de données incorrectes contenues dans le document. Cette erreur ne concerne que les index de type Ngram dont la page de codes est UTF8.

### 1500-1505

Des incidents se sont produits au niveau du composant d'analyse des documents. Il n'a pas pu être initialisé (vérifiez la variable LIBPATH et le contenu du fichier de configuration) ou son utilisation a échoué en raison d'incidents internes. Pour plus de détails, consultez le fichier de diagnostic.

### 1904

Le document n'a pas été indexé. Un incident s'est produit lors de l'accès au modèle de document pour un index activé par section. Vérifiez les droits d'accès et l'existence du fichier.

### 2000

Le document n'a pas été indexé. Ce type de document n'est pas pris en charge. La fonction de bibliothèque Lib\_access\_doc a renvoyé un type de document incorrect.

### 2001

Le document n'a pas été indexé. Une séquence de zones incorrecte a été détectée dans le flux de données du document.

### 2002

Le document n'a pas été indexé. car une zone incorrectement structurée a été détectée dans son flux de données.

### 2003

Le document n'a pas été indexé. Il ne doit contenir qu'une seule section de texte dans le format texte de DB2 Extension Texte.

### 2005

Le document n'a pas été indexé. car une langue indiquée dans son flux de données n'est pas prise en charge.

## Codes anomalie dus à des événements d'erreur

**2006**

Le document n'a pas été indexé. car un CCSID indiqué dans son flux de données n'est pas pris en charge.

**2007**

Le format de document attendu, spécifié par la bibliothèque ou la règle par défaut, est incorrect. L'en-tête de document est inapproprié pour ce format. Vérifiez si la règle par défaut est un document doté d'un en-tête spécial ; effectuez les modifications nécessaires si cette règle est incorrecte.

**2008**

Le document n'a pas été indexé car il est inaccessible.

**2009**

Le document n'a pas été indexé car il était en cours d'utilisation, donc inaccessible.

**2010**

Le document n'a pas été indexé. Le CCSID indiqué n'est pas correct.

**2011**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. La marque de fin de page doit être le dernier élément de contrôle du corps du document.

**2012**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. Une zone structurée contient un paramètre de longueur incorrect.

**2013**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. Un élément de contrôle incorrect a été détecté dans le document.

**2014**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. Un élément de contrôle multi-octet ou une zone incorrectement structurée a été détecté dans le document.

**2015**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. Des paramètres en double ont été détectés.

**2016**

Le document n'a pas été indexé car il n'est pas dans un format IBM RFT ou FFT valide. Le programme a détecté une unité de texte vide.

**2018**

Le format du document n'est pas pris en charge ou la base DIT contient

## Codes anomalie dus à des événements d'erreur

une entrée EXCLUDE relative à l'extension du document. Vérifiez que le document porte une extension autorisant son indexation.

### 2020

Le document n'a pas été indexé. car il ne s'agit pas d'un document ou fichier WordPerfect.

### 2021

Le document n'a pas été indexé. Il s'agit d'un fichier de type WordPerfect, mais non d'un document WordPerfect.

### 2022

Le document n'a pas été indexé. Cette version de WordPerfect n'est pas prise en charge.

### 2023

Le document n'a pas été indexé. Il s'agit d'un fichier WordPerfect chiffré. Sauvegardez le document sans chiffrement.

### 2026

END\_TXT trouvé dans une note de bas de page ou de fin de document. Vérifiez le fichier WordPerfect ; il est peut-être altéré.

### 2028

L'analyseur syntaxique a renvoyé du texte n'appartenant pas au document. Vérifiez le contenu du fichier et principalement les mots spécifiques du format. Vérifiez si le format du document est pris en charge. Si l'opération de reconnaissance automatique du format échoue, assurez-vous que l'analyseur syntaxique approprié est appelé.

### 2030

Le document n'a pas été indexé. Il ne s'agit pas d'un fichier Microsoft Word, ou cette version de Word n'est pas prise en charge.

### 2031

Le document n'a pas été indexé. Une marque de fin de fichier inattendue a été détectée dans un document Microsoft Word.

### 2032

Le document n'a pas été indexé. Un élément de contrôle incorrect a été détecté dans un document Microsoft Word.

### 2033

Le document n'a pas été indexé. Il a été sauvegardé au format *complexe* avec l'option *fastsave*. Enregistrez-le de nouveau après avoir désactivé cette option.

### 2034

Le document n'a pas été indexé. Il manque une marque de fin de zone obligatoire dans un document Microsoft Word.

## Codes anomalie dus à des événements d'erreur

**2035**

Le document est chiffré. Enregistrez-le au format Microsoft Word sans chiffrement.

**2036**

Il s'agit d'un document Word pour Macintosh ; traitement impossible. Enregistrez-le au format Word pour Windows.

**2037**

Ce document Word contient des objets OLE inclus.

**2040**

Le document n'a pas été indexé car il n'est pas dans un format ECTF valide.

**2041**

Le document n'a pas été indexé. Il contient un élément de contrôle .SO LEN non suivi d'un nombre.

**2042**

Le document n'a pas été indexé. Il contient un élément de contrôle .SO LEN suivi d'un nombre incorrect. Ce nombre doit être compris entre 1 et 79.

**2043**

Le document n'a pas été indexé. Un seul élément de contrôle .SO DOC est autorisé. Sauvegardez chaque document ECTF dans un fichier séparé.

**2044**

Le document n'a pas été indexé. Un élément de contrôle .SO HDE doit être suivi de marques de début et de fin.

**2046**

Le document n'a pas été indexé. L'élément de contrôle .SO DOC du document est précédé de texte.

**2047**

Le document n'a pas été indexé. L'élément de contrôle .SO PID du document est précédé de texte.

**2048**

Le document n'a pas été indexé. Il manque une marque de fin après une marque de début.

**2050**

Le document n'a pas été indexé. Des marques incorrectes ont été détectées après l'élément de contrôle .SO HDE.

**2051**

Le document n'a pas été indexé. Une marque de fin de ligne a été détectée après un élément de contrôle .SO.

### 2052

Le document n'a pas été indexé. Une marque de fin de texte inattendue a été détectée.

### 2060

Le document n'a pas été indexé. Il ne s'agit pas d'un document AmiPro, ou cette version d'AmiPro n'est pas prise en charge.

### 2061

Le document n'a pas été indexé. Un élément de contrôle trop long a été détecté dans un document AmiPro.

### 2062

Le document n'a pas été indexé. Cette version d'AmiPro n'est pas prise en charge. Seule l'architecture AmiPro Version 4 est prise en charge.

### 2063

Des feuilles de style AmiPro n'ont pas été indexées.

### 2064

Le document n'a pas été indexé. Un jeu de caractères incorrect a été détecté. Seul le jeu de caractères Lotus "Character Set 82" (Windows ANSI) est pris en charge.

### 2065

Le document n'a pas été indexé. Une marque de fin de fichier inattendue a été détectée dans un document AmiPro.

### 2072

Impossible de lire le document car il est chiffré.

### 2073

Le format du document est incohérent.

### 2074

Le bit indicateur "fichier défectueux" (bad file) du document est activé.

### 2080

Le document n'a pas été indexé. Il ne s'agit pas d'un document RTF, ou cette version du format RTF n'est pas prise en charge.

### 2081

Le document n'a pas été indexé. Un libellé d'élément de contrôle RTF trop long a été détecté.

### 2083

Le document n'a pas été indexé. La page de codes Macintosh n'est pas prise en charge.

### 2084

Le document n'a pas été indexé. Ce format RTF n'est pas pris en charge. Seul RTF version 1 est accepté.

## Codes anomalie dus à des événements d'erreur

**2090**

Le document n'a pas été indexé. Il s'agit d'un document HTML qui contient une marque considérée comme étant trop longue par l'analyseur syntaxique.

**2093**

Le document n'a pas été indexé. Il s'agit d'un document XML qui a été rejeté par l'analyseur syntaxique.

**2100**

Le document est altéré ou illisible pour une raison indéterminée. Un nouvel analyseur syntaxique courant pourrait résoudre l'incident.

**2101**

Impossible d'indexer le document car il est vide ou ne contient pas de texte. Vérifiez si le document contient uniquement des graphiques.

**2102**

Impossible d'indexer le document car il est chiffré ou protégé par un mot de passe.

**2105**

Type de document reconnu, mais filtre non disponible.

**2106**

Impossible d'indexer le document car il est vide.

**2107**

Impossible d'indexer le document car le système ne parvient pas à l'ouvrir. Vérifiez les droits d'accès au document.

**2112**

Impossible d'indexer le document car il s'agit d'un fichier exécutable.

**2113**

Impossible d'indexer le document car il est compressé.

**2114**

Impossible d'indexer le document car il s'agit d'un graphique. Si ce format de document graphique renvoie une portion de texte acceptable, demandez qu'il soit inclus dans le processus d'indexation.

**2120**

Le fichier de sortie de l'exit utilisateur n'existe pas ou est inaccessible. Un nouvel analyseur syntaxique courant pourrait résoudre l'incident.

**2121**

Le fichier de sortie est vide ou ne peut pas être ouvert en lecture. Un nouvel analyseur syntaxique courant pourrait résoudre l'incident.

### 2122

Tentative d'utilisation d'un fichier de sortie d'exit utilisateur, mais aucun nom de fichier n'a été spécifié ou défini dans l'objet.

### 2130

Impossible d'exécuter le programme d'exit utilisateur. Vérifiez si l'exécutable figure dans l'un des chemins d'accès définis par la variable d'environnement PATH. Créez une trace et un cliché pour obtenir des informations supplémentaires sur les codes retour (errno) d'environnement.

### 2131

Echec du programme d'exit utilisateur ; code retour erroné. Créez une trace et un cliché pour obtenir des informations supplémentaires sur les codes retour (errno) d'environnement.

## Récapitulatif des modifications

---

## Partie 3. Annexes



---

## Remarques

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Pour plus de détails, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial IBM. Toute référence à un produit, logiciel ou service IBM n'implique pas que seul ce produit, logiciel ou service puisse être utilisé. Tout autre élément fonctionnellement équivalent peut être utilisé, s'il n'enfreint aucun droit d'IBM. Il est de la responsabilité de l'utilisateur d'évaluer et de vérifier lui-même les installations et applications réalisées avec des produits, logiciels ou services non expressément référencés par IBM.

IBM peut détenir des brevets ou des demandes de brevet couvrant les produits mentionnés dans le présent document. La remise de ce document ne vous donne aucun droit de licence sur ces brevets ou demandes de brevet. Si vous désirez recevoir des informations concernant l'acquisition de licences, veuillez en faire la demande par écrit à l'adresse suivante :

IBM EMEA Director of Licensing  
IBM Europe Middle-East Africa  
Tour Descartes  
La Défense 5  
2, avenue Gambetta  
92066 - Paris-La Défense CEDEX  
France

Pour le Canada, veuillez adresser votre courrier à :

IBM Director of Commercial Relations  
IBM Canada Ltd  
3600 Steeles Avenue East  
Markham, Ontario  
L3R 9Z7  
Canada

Les informations sur les licences concernant les produits utilisant un jeu de caractères double octet peuvent être obtenues par écrit à l'adresse suivante :

IBM World Trade Asia Corporation  
Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106, Japon

**Le paragraphe suivant ne s'applique ni au Royaume-Uni, ni dans aucun pays dans lequel il serait contraire aux lois locales :** LE PRESENT DOCUMENT EST LIVRE « EN L'ETAT». IBM DECLINE TOUTE RESPONSABILITE, EXPRESSE OU IMPLICITE, RELATIVE AUX INFORMATIONS QUI Y SONT CONTENUES, Y COMPRIS EN CE QUI CONCERNE LES GARANTIES DE QUALITE MARCHANDE OU D'ADAPTATION A VOS BESOINS. Certaines juridictions n'autorisent pas l'exclusion des garanties implicites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Le présent document peut contenir des inexactitudes ou des coquilles. Il est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. IBM peut modifier sans préavis les produits et logiciels décrits dans ce document.

Les références à des sites Web non IBM sont fournies à titre d'information uniquement et n'impliquent en aucun cas une adhésion aux données qu'ils contiennent. Les éléments figurant sur ces sites Web ne font pas partie des éléments du présent produit IBM et l'utilisation de ces sites relève de votre seule responsabilité.

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie des informations qui lui seront fournies.

Les licenciés souhaitant obtenir des informations permettant : (i) l'échange des données entre des logiciels créés de façon indépendante et d'autres logiciels (dont celui-ci), et (ii) l'utilisation mutuelle des données ainsi échangées, doivent adresser leur demande à :

IBM Canada Limited  
Department 071  
1150 Eglinton Ave. East  
North York, Ontario  
M3C 1H7  
CANADA

Ces informations peuvent être soumises à des conditions particulières prévoyant notamment le paiement d'une redevance.

Le logiciel sous licence décrit dans ce document et tous les éléments sous licence disponibles s'y rapportant sont fournis par IBM conformément aux termes du Contrat sur les produits et services IBM, des Conditions internationales d'utilisation des logiciels IBM ou de tout autre accord équivalent.

Les données de performance indiquées dans ce document ont été déterminées dans un environnement contrôlé. Par conséquent, les résultats peuvent varier de manière significative selon l'environnement d'exploitation utilisé. Certaines mesures évaluées sur des systèmes en cours de développement ne sont pas garanties sur tous les systèmes disponibles. En outre, elles peuvent résulter d'extrapolations. Les résultats peuvent donc varier. Il incombe aux utilisateurs de ce document de vérifier si ces données sont applicables à leur environnement d'exploitation.

Les informations concernant des produits non IBM ont été obtenues auprès des fournisseurs de ces produits, par l'intermédiaire d'annonces publiques ou via d'autres sources disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances ni leur compatibilité. Elle ne peut recevoir aucune réclamation concernant des produits non IBM. Toute question concernant les performances de produits non IBM doit être adressée aux fournisseurs de ces produits.

Toute instruction relative aux intentions d'IBM pour ses opérations à venir est susceptible d'être modifiée ou annulée sans préavis, et doit être considérée uniquement comme un objectif.

Ce document peut contenir des exemples de données et des rapports utilisés couramment dans l'environnement professionnel. Ces exemples mentionnent des noms fictifs de personnes, de sociétés, de marques ou de produits à des fins illustratives ou explicatives uniquement. Toute ressemblance avec des noms de personnes, de sociétés ou des données réelles serait purement fortuite.

#### LICENCE DE COPYRIGHT :

Le présent logiciel peut contenir des exemples de programmes d'application en langage source destinés à illustrer les techniques de programmation sur différentes plateformes d'exploitation. Vous avez le droit de copier, de modifier et de distribuer ces exemples de programmes sous quelque forme que ce soit et sans paiement d'aucune redevance à IBM, à des fins de développement, d'utilisation, de vente ou de distribution de programmes d'application conformes aux interfaces de programmation des plateformes pour lesquelles ils ont été écrits ou aux interfaces de programmation IBM. Ces exemples de programmes n'ont pas été rigoureusement testés dans toutes les conditions. Par conséquent, IBM ne peut garantir expressément ou implicitement la fiabilité, la maintenabilité ou le fonctionnement de ces programmes.

Toute copie totale ou partielle de ces programmes exemples et des oeuvres qui en sont dérivées doit comprendre une notice de copyright, libellée comme suit :

© (nom de votre société) (année). Certaines parties de ce code sont dérivées des programmes exemples d'IBM Corp. © Copyright IBM Corp. \_indiquez l'année ou les années\_. All rights reserved.

---

## Marques

Les termes qui suivent, accompagnés d'un astérisque (\*) dans le document, sont des marques d'International Business Machines Corporation dans certains pays.

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

Les termes qui suivent sont des marques d'autres sociétés :

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation dans certains pays.

Java, ou toutes les marques et logos incluant Java, et Solaris sont des marques de Sun Microsystems, Inc.

UNIX est une marque enregistrée d'Open Group aux Etats-Unis et/ou dans certains pays.

D'autres sociétés sont propriétaires des autres marques, noms de produits ou logos qui pourraient apparaître dans ce document.



---

## Glossaire

Ce glossaire définit les principaux termes, abréviations et acronymes utilisés dans le présent document. Si vous ne trouvez pas le terme que vous recherchez, reportez-vous à l'index ou à la *Terminologie du traitement de l'information*, GCF2-0076, publiée par IBM France, ou encore à un dictionnaire d'informatique générale.

### A

**activer** : Préparer une base de données , une table ou une colonne de texte en vue de son utilisation par DB2 Extension Texte.

**afficher** : Visualiser un texte sur un écran.

**afficheur** : Fonction de DB2 Extension Texte qui permet d'afficher du texte sur un écran.

**affiner** : Ajouter un critère de recherche provenant d'une recherche précédente à d'autres critères de façon à réduire le nombre de *correspondances*.

**API** : Interface de programmation d'applications.

**argument de recherche** : Conditions mentionnées lors de l'exécution d'une recherche, se composant d'un ou plusieurs termes et de paramètres de recherche.

### B

**base de données partitionnée** : Base de données constituée de plusieurs parties, chacune étant gérée par un serveur de partitions de bases de données distinct.

**BOUND** : Option de recherche spécifique de la langue coréenne.

### C

**caractère d'échappement** : Caractère indiquant que le caractère suivant ne doit pas être interprété comme un *caractère générique*.

**caractère générique** : Caractère utilisé pour représenter des caractères facultatifs au début, au milieu ou à la fin d'un terme de recherche. Les caractères génériques sont normalement utilisés pour rechercher les variantes d'un terme dans un index exact.

**caractéristiques d'index** : Propriétés d'un *index de texte* déterminant :

le répertoire dans lequel il est stocké ;

son type ;

sa fréquence de mise à jour ;

le moment de sa première mise à jour.

**CCSID** : Identificateur de jeu de caractères codés.

**colonne de texte** : Colonne contenant des *documents texte*.

**configuration de texte** : Valeurs par défaut définissant l'index, le texte et le traitement.

**correspondance** : Occurrence d'un terme recherché dans un document de type texte.

**correspondance étendue** : Processus impliquant l'utilisation d'un *dictionnaire*, qui met en évidence les termes qui n'ont pas de correspondance immédiate avec le terme recherché.

**COUNT** : Mot clé utilisé pour indiquer le nombre de niveaux à prendre en compte dans la classification des termes du thésaurus, dans le cadre de l'extension d'une recherche pour une relation déterminée.

## D

**DBCS** : Jeu de caractères codés sur deux octets.

**DB2 Extension** : Groupe de programmes permettant de stocker et d'extraire des types de données autres que les traditionnelles données numériques et caractères : données image, audio et vidéo, et documents complexes.

**déclencheur** : Mécanisme qui ajoute automatiquement des informations relatives à des documents à indexer, dans une *table de journalisation*, lors de l'ajout, de la modification ou de la suppression d'un document dans une colonne de texte.

**désactiver** : Ramener une base de données, une table ou une colonne de type texte à l'état qui était le sien avant qu'elle ne soit activée pour DB2 Extension Texte en supprimant les éléments créés durant le processus d'activation.

**descripteur** : Valeur binaire qui identifie un document de type texte. Cette valeur comprend :

- un ID document
- le nom et l'emplacement de l'index associé
- les *informations de texte* relatives au document
- le chemin et le nom du fichier, si le document est situé dans un fichier externe ne dépendant pas de DB2.

Un descripteur est créé pour chaque document texte dans une colonne de texte lors de l'*activation* de cette dernière en vue de son utilisation par DB2 Extension Texte.

**descripteur de document** : Voir *descripteur*.

**descripteur de fichier** : Voir *descripteur*.

**descripteur initialisé** : *Descripteur* préparé à l'avance, contenant uniquement le format du texte et/ou la langue du texte.

**dictionnaire** : Ensemble d'informations linguistiques utilisées par DB2 Extension Texte durant l'analyse de texte, l'indexation, l'extraction et la mise en évidence de documents dans une langue spécifique.

**document** : Voir *document texte*.

**document texte** : Texte de type CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR ou CLOB, stocké dans une table DB2.

## E

**extension** : Voir *DB2 Extension*.

**extension par thésaurus** : Opération consistant à ajouter à un terme objet d'une recherche des termes supplémentaires issus d'un thésaurus.

**extraire** : Trouver un document de type texte, à l'aide d'un argument de recherche, dans l'une des fonctions de recherche de DB2 Extension Texte.

## F

**fichier externe** : Document texte se présentant sous forme de fichier stocké dans le système de fichiers du système d'exploitation, au lieu d'être sous forme de cellule dans une table contrôlée par DB2.

**flux de données** : Informations renvoyées par une fonction API, se composant de texte (au moins un paragraphe) contenant le terme recherché, et d'informations pour la mise en évidence du terme trouvé.

**fonction** : Voir *fonction d'accès*.

**fonction d'accès** : Fonction fournie à l'utilisateur, qui convertit le type des données du texte d'une colonne en un type pouvant être traité par DB2 Extension Texte.

**fonction UDF** : Fonction SQL créée par un utilisateur DB2, par opposition à une fonction SQL fournie par DB2. DB2 Extension Texte fournit des fonctions d'administration et de recherche, telles que CONTAINS, sous forme de fonctions UDF.

**format** : Type d'un document (par exemple, ASCII ou WordPerfect).

**fréquence de mise à jour :** Fréquence à laquelle un index de texte est mis à jour, exprimée en jours, heures et minutes, et nombre minimal de noms de document devant être listés dans la *table de journalisation* pour indexation, avant l'exécution de cette opération.

## G

**groupe de noeuds :** Sous-ensemble nommé d'un ou de plusieurs serveurs de partition de base de données. *Noeud* affecté à une machine physique distincte. Voir aussi *noeud logique*.

## I

**indexation périodique :** Indexation à des intervalles prédéterminés, exprimés en jours, heures et minutes, et nombre minimal de noms de document devant être répertoriés dans la *table de journalisation* pour indexation, avant l'exécution de cette dernière.

**index de texte :** Ensemble de termes significatifs extraits de documents texte. Chaque terme est associé au document dont il est extrait. Il est possible d'améliorer considérablement les temps de recherche en effectuant la recherche sur l'index au lieu de l'effectuer sur le document lui-même. Voir aussi *index exact* et *index linguistique*.

**indexer :** Extraire des termes significatifs du texte et les stocker dans un *index de texte*.

**index exact :** *Index de texte* contenant des termes sous la forme exacte dans laquelle ils apparaissent dans le document dont ils sont extraits. Voir aussi *index exact* et *index Ngram*.

**index linguistique :** *Index de texte* contenant des termes réduits à leur forme de base par une procédure linguistique. Ainsi, "chevaux" sera indexé sous "cheval". Voir aussi *index exact* et *index Ngram*.

**Index Ngram :** *Index de texte* prenant en charge les documents DBCS et la recherche de

correspondance partielle dans des documents SBCS. Voir aussi *index linguistique*, *index exact* et *index double effet*.

**informations de mise en évidence :** Voir *flux de données*.

**informations de texte :** Caractéristiques d'un *document texte* décrivant :

- le *CCSID*,
- le *format*,
- la *langue*.

**instance :** Environnement logique de DB2 Extension Texte. Il peut y avoir plusieurs instances de DB2 Extension Texte sur un même poste de travail, mais une seule par instance DB2. Vous pouvez utiliser ces instances pour :

- séparer l'environnement de développement de l'environnement de production ;
- restreindre l'accès aux informations sensibles à un groupe de personnes particulier.

**interface de programmation d'applications (API) :** Interface générale entre des programmes d'applications et les fonctions d'extraction d'informations de DB2 Extension Texte.

**interpréteur de commandes :** Programme appelé DB2TX qui :

- permet d'entrer des commandes de DB2 Extension Texte ;
- traite les commandes ;
- affiche le résultat correspondant.

## L

**langue :** Nom d'un *dictionnaire* utilisé lors de l'*indexation*, la recherche et l'*affichage*.

## M

**masquage de caractères et de mots :** Voir *caractère générique*.

**modèle de document :** Définition de la structure d'un document en fonction des sections qui le compose. Un modèle de document permet à DB2

Extension Texte de reconnaître les sections d'un document lors de l'indexation. Un modèle de document énumère les balises qui identifient les sections. Pour chaque balise, vous pouvez spécifier un nom descriptif de section, qui sera utilisé lors de recherches dans ce type de section. Vous pouvez spécifier un ou plusieurs modèles de document dans un fichier de modèles de document.

**mot vide :** Mot courant, tel que "avant", dans un *document texte*, qui doit être exclu de l'*index de texte* et ignoré lorsqu'il figure dans un *argument de recherche*.

## N

**noeud :** Représentation d'un serveur dans un environnement de *base de données partitionnée*. Voir aussi *noeud logique*, *noeud physique* et *groupe de noeuds*.

**noeud logique :** *Noeud* affecté à la même machine physique que d'autres noeuds. Voir aussi *noeud physique*.

**noeud physique :** *Noeud* affecté à une machine physique distincte. Voir aussi *noeud logique*.

## O

**occurrence :** Synonyme de *correspondance*.

## P

**page de codes :** Affectation de caractères graphiques et de fonctions correspondantes à tous les points de code. Par exemple, affectation de caractères et de fonctions aux 256 points d'un code à 8 bits.

**profil :** Voir *profil d'environnement*.

**profil d'environnement :** Script fourni avec DB2 Extension Texte contenant les définitions des *variables d'environnement*.

**profil de shell :** Voir *profil d'environnement*.

## R

**rang :** Valeur absolue de type DOUBLE comprise entre 0 et 1 qui indique dans quelle mesure un document répond aux critères de recherche par rapport à d'autres documents trouvés. La valeur indique le nombre de correspondances trouvées dans le document par rapport à sa taille.

**recherche booléenne :** Recherche sur un ou plusieurs termes faisant intervenir des opérateurs booléens.

**recherche caractérisée :** Recherche sur des noms propres (tels que noms de personnes, lieux ou organisations) effectuée dans un index linguistique créé avec l'option `FEATURE_EXTRACTION`.

**recherche de correspondance partielle :** Recherche intégrant les mots dont l'orthographe est proche de celle du terme recherché.

**recherche hybride :** Recherche combinant les caractéristiques d'une *recherche booléenne* et d'une *recherche sur texte libre*.

**recherche sur texte libre :** Recherche où le terme recherché est exprimé sous forme libre, c'est-à-dire sous forme d'expression ou de phrase décrivant en langage naturel l'objet de la recherche.

## S

**SBCS :** Jeu de caractères codés sur un seul octet.

## T

**table à index commun :** Table DB2 dont les colonnes de texte partagent un index de texte commun. Voir aussi *table à index multiples*.

**table à index multiples :** Table DB2 dont les colonnes de texte disposent d'*index de texte* individuels. Voir aussi *table à index commun*.

**table de journalisation :** Table créée par DB2 Extension Texte, contenant des informations sur

les documents de type texte à indexer. Des *déclencheurs* sont utilisés pour stocker ces informations dans une table de journalisation chaque fois qu'un document est ajouté, modifié ou supprimé dans une colonne de texte activée.

**table de texte :** Table DB2 contenant des *colonnes de texte*.

**trace :** Opération consistant à stocker dans un fichier des informations qui pourront servir ultérieurement à identifier l'origine d'une erreur.

**type d'index :** Caractéristique d'un *index de texte* déterminant s'il contient des formes exactes et/ou des formes linguistiques de termes du document. Voir aussi *index exact*, *index linguistique* et *index Ngram*.

**type distinct :** Voir *type utilisateur*.

**type UDT :** Type de données créé par un utilisateur DB2, par opposition à un type de données fourni par DB2 tel que LONG VARCHAR.

## U

**UDF :** Fonction UDF (User-defined function).

**UDT :** Type UDT (User-defined type).

## V

**variable d'environnement :** Variable utilisée pour fournir des valeurs par défaut pour l'environnement DB2 Extension Texte.

**variable d'instance :** Variable utilisée pour fournir une valeur par défaut pour le nom du propriétaire de l'*instance* ou pour le nom du répertoire principal de ce dernier.

**vue de catalogue :** Vue d'une table système créée par DB2 Extension Texte à des fins d'administration. Une vue de catalogue contient des informations sur les tables et les colonnes activées pour être utilisées par DB2 Extension Texte.



---

# Index

## A

### abréviations

- listes 54
- modification d'un fichier d'abréviations 55

### administration

- affichage de l'état d'une base de données, d'une table et d'une colonne 106
- affichage de l'état de l'index 108
- affichage de l'état du serveur 22
- affichage de paramètres d'information de texte 111
- affichage des paramètres d'un index 110
- arrêt du serveur DB2 Extension Texte 187
- commande CHANGE INDEX SETTINGS 131
- commande CHANGE TEXT CONFIGURATION 133
- commande CONNECT 137
- commande DISABLE DATABASE 140
- commande DISABLE TEXT COLUMN 141
- commande DISABLE TEXT FILES 142
- commande DISABLE TEXT TABLE 143
- commande GET ENVIRONMENT 162
- commande GET INDEX STATUS 165
- commande GET TEXT CONFIGURATION 167
- commande REORGANIZE INDEX 170
- commande RESET INDEX STATUS 171
- commande TXICRT 176
- commande TXIDROP 178
- commande TXIPCLEAN 180
- commande TXNADD 181
- commande TXNCHECK 182
- commande TXNDROP 183
- commande TXSAMPLE 184
- commande TXSTART 185
- commande TXSTATUS 186

### administration (suite)

- commande TXSTOP 187
- commande TXTHESC 188
- commande TXTHESN 190
- commande TXTRACE 192
- commande TXVERIFY 197
- commande TXWIZARD 198
- commandes du client (résumé) 129
- compilation d'un fichier de définition de thésaurus 188
- compilation d'un fichier de définition de thésaurus Ngram 190
- création d'un exemple de base de données 16
- création d'un modèle de base de données 197
- création d'une instance DB2 Extension Texte 21
- DB2TX (commande) 130
- DELETE INDEX EVENTS 139
- DELETE INDEX EVENTS (commande) 139
- démarrage du serveur DB2 Extension Texte 185
- désactivation d'une base de données 117
- désactivation d'une colonne de texte 115
- désactivation d'une table de texte 116
- désactivation de fichiers de texte 116
- fichier d'abréviations, modification 55
- fichier de mots vides, modification 55
- gestion des index de texte 101
- GET INDEX SETTINGS (commande) 163
- GET STATUS (commande) 166
- GET TEXT INFO (commande) 168
- informations d'état, extraction 105
- interpréteur de commandes 130
- mise à jour d'un index pour des fichiers externes 102

### administration (suite)

- mise à jour immédiate d'un index 101
  - modification de la configuration du texte 20
  - modification des fichiers de mots vides et d'abréviations 55
  - modification des paramètres d'un index 102
  - QUIT (commande) 169
  - réinitialisation de l'état d'un index 103
  - réorganisation d'un index 105
  - résumé des commandes du client 129
  - résumé des commandes du serveur 175
  - sauvegarde et restauration 22
  - suppression d'une instance DB2 Extension Texte 21
  - suppression des événements d'un index 104
  - trace des incidents 23
  - TXILIST (commande) 179
  - UPDATE INDEX (commande) 172
  - variables d'environnement 17
- ### affichage
- analyse linguistique 266
  - exemple de programme 253
  - utilisation de votre propre afficheur 122
- ### affinement d'une recherche 95
- aide sur les commandes 60
  - AmiPro, format de document 31
  - analyse de texte
    - affichage 266
    - pour l'indexation 256
  - analyse de texte de base
    - mise en évidence 266
    - normalisation 257
    - pour l'indexation des termes 256
    - reconnaissance de phrases 257
    - termes contenant des caractères non alphanumériques 256
- ### AND
- mot clé dans argument de recherche 222

- AND (*suite*)
  - opérateurs booléens 86
- argument de recherche
  - AND (mot clé) 222
  - BOUND (mot clé) 223
  - COUNT (mot clé) 221
  - description 215
  - EXPAND (mot clé) 221
  - FUZZY FORM OF (mot clé) 223
  - IN SAME PARAGRAPH AS 222
  - IN SAME SENTENCE AS 222
  - MODEL (mot clé) 219
  - NOT (mot clé) 222
  - opérateur & (AND) 221
  - opérateur | (OR) 221
  - PRECISE FORM OF (mot clé) 223
  - recherche à l'aide de & et | 86
  - recherche à partir de fragments de termes 87
  - recherche avec NOT 91
  - recherche d'homophones 93
  - recherche de correspondance partielle 92, 223
  - recherche de plusieurs termes 86
  - recherche de synonymes 90
  - recherche de termes conjoints ou disjoints 89
  - recherche de termes d'un même paragraphe 89
  - recherche de termes d'une même phrase 89
  - recherche de termes dans des sections de document 89
  - recherche de termes dans plusieurs langues 90
  - recherche des variantes d'un terme 87
  - Recherche hybride 94
  - recherche par thésaurus 93
  - recherche segmentée 93
  - recherche sur texte libre 94
  - RESULT LIMIT (mot clé) 221
  - résumé des règles et restriction 226
  - SECTION (mot clé) 219
  - spécification 86
  - STEMMED FORM OF (mot clé) 223
  - SYNONYM FORM OF (mot clé) 223
  - syntaxe 217
  - TERM OF (mot clé) 221
  - THESAURUS (mot clé) 220
- argument de recherche (*suite*)
  - utilisation de caractères génériques 87
  - valeur d'attribut 220
- Argument de recherche booléen 221
- arrêt du serveur DB2 Extension Texte 187
- ASCII, format de document 31
- ASCII à plat, format de document 31
- assistant 27
  - commande TXWIZARD 198
  - description 27
- B**
- base de données
  - activation 61
  - commande CONNECT 137
  - Commande ENABLE DATABASE 144
  - connexion 60
  - désactivation 117
  - DISABLE DATABASE (commande) 140
  - GET STATUS (commande) 166
  - informations d'état, affichage 106
  - sauvegarde et restauration 22
- base de données exemple création (TXVERIFY) 16
- BOUND (mot clé) 223
- C**
- caractère d'échappement
  - syntaxe 226
  - utilisation 88
- caractères génériques
  - dans une séquence de recherche 87
  - masquage de mots 265
- caractéristiques d'index
  - affichage 110
  - dans ENABLE TEXT FILES 154
  - ENABLE TEXT COLUMN (commande) 145
  - ENABLE TEXT TABLE (commande) 157
  - paramètres par défaut de la configuration de texte 19
- caractéristiques de texte
  - CCSID 35
  - dans ENABLE TEXT FILES 154
  - ENABLE TEXT COLUMN (commande) 145
  - format 31
- caractéristiques de texte (*suite*)
  - langue 35
  - paramètres par défaut de la configuration de texte 19
- caractéristiques de traitement
  - paramètres par défaut de la configuration de texte 20
- CASE\_ENABLED (mot clé)
  - ENABLE TEXT COLUMN (commande) 149
  - ENABLE TEXT TABLE (commande) 158
- CCSID
  - dans CHANGE TEXT CONFIGURATION 135
  - description 35
  - ENABLE TEXT COLUMN (commande) 147
  - éviter les incidents de pages de codes 37
  - extraction des informations d'un descripteur 98
  - fonction 202
  - GET TEXT INFO (commande) 168
  - initialisation de descripteurs 208
  - liste 35
  - paramétrage par défaut 19
- Codes anomalie 321
- Codes anomalie du moteur de recherche 317
- codes retour 291
- colonne
  - activation 66
  - activation d'une table volumineuse 70
  - activation pour des types d'index différents 69
  - commande DISABLE TEXT COLUMN 141
  - commande ENABLE TEXT COLUMN 145
  - désactivation 115
  - commande CHANGE INDEX SETTINGS
    - procédure 102
    - syntaxe 131
  - commande CHANGE TEXT CONFIGURATION
    - procédure 20
    - syntaxe 133
  - commande DISABLE DATABASE
    - syntaxe 140
    - utilisation 117

- commande DISABLE TEXT COLUMN
    - syntaxe 141
    - utilisation 115
  - commande DISABLE TEXT FILES
    - syntaxe 142
    - utilisation 116
  - commande DISABLE TEXT TABLE
    - syntaxe 143
    - utilisation 116
  - commande ENABLE DATABASE
    - syntaxe 144
    - utilisation 61
  - commande ENABLE TEXT COLUMN
    - syntaxe 145
    - utilisation 66
  - commande ENABLE TEXT FILES
    - syntaxe 154
    - utilisation 73
  - commande ENABLE TEXT TABLE
    - syntaxe 157
    - utilisation 63
  - commandes
    - CHANGE INDEX SETTINGS 131
    - CHANGE TEXT CONFIGURATION 133
    - commande TXWIZARD 198
    - CONNECT 137
    - DB2TX 130
    - DELETE INDEX EVENTS 139
    - DISABLE DATABASE 140
    - DISABLE TEXT COLUMN 141
    - DISABLE TEXT FILES 142
    - DISABLE TEXT TABLE 143
    - ENABLE DATABASE 144
    - ENABLE TEXT COLUMN 145
    - ENABLE TEXT FILES 154
    - ENABLE TEXT TABLE 157
    - GET ENVIRONMENT 162
    - GET INDEX SETTINGS 163
    - GET INDEX STATUS 165
    - GET STATUS 166
    - GET TEXT CONFIGURATION 167
    - GET TEXT INFO 168
    - QUIT 169
    - REORGANIZE INDEX 170
    - RESET INDEX STATUS 171
    - résumé des commandes du client 129
    - résumé des commandes du serveur 175
    - TXICRT 176
  - commandes (*suite*)
    - TXIDROP 178
    - TXILIST 179
    - TXIPCLEAN 180
    - TXNADD 181
    - TXNCHECK 182
    - TXNDROP 183
    - TXSAMPLE 184
    - TXSTART 185
    - TXSTATUS 186
    - TXSTOP 187
    - TXTHESC 188
    - TXTHESN 190
    - TXTRACE 192
    - TXVERIFY (commande) 197
    - UPDATE INDEX 172
  - COMMITCOUNT (paramètre de configuration)
    - dans CHANGE TEXT CONFIGURATION 135
    - description 70
    - ENABLE TEXT COLUMN (commande) 152
    - ENABLE TEXT TABLE (commande) 172
    - paramétrage par défaut 20
    - protection de l'espace de journalisation 71
  - compilation d'un fichier de définition de thésaurus 188
  - compilation d'un fichier de définition de thésaurus Ngram 190
  - concepts du parallélisme 7
  - Configuration 17
  - CONNECT (commande)
    - syntaxe 137
    - utilisation 60
  - connexion à une base de données
    - commande CONNECT 137
    - procédure 60
  - CONTAINS (fonction)
    - exemple 84
    - syntaxe 203
  - correspondance
    - à partir de
      - DesGetSearchResultTable 121
      - dans un résultat de recherche 85
      - DesGetMatches (fonction) 237
      - NUMBER\_OF\_MATCHES (fonction) 211
    - correspondance étendue 266
    - Correspondance étendue 266
    - COUNT (mot clé) 221
  - création d'un modèle de base de données 197
  - création d'un modèle de table
    - commande TXSAMPLE 184
  - création d'une instance DB2 Extension Texte
    - commande TXICRT 176
- ## D
- DB2 Extensions
    - exemple d'utilisation 4
  - DB2DBDFT (variable d'environnement) 17
  - DB2INSTANCE (variable d'environnement) 17
  - DB2TEXTFH (type distinct) 199
  - DB2TEXTFHLISTP (type distinct) 200
  - DB2TEXTH (type distinct) 199
  - DB2TEXTHLISTP (type distinct) 200
  - DB2TX\_ (variables d'environnement)
    - affichage 107
    - description 17
  - DB2TX (interpréteur de commandes)
    - syntaxe 130
    - utilisation 59
  - DB2TX\_INSTOWNER (variable d'environnement) 17
  - DB2TX\_INSTOWNERHOMEDIR (variable d'environnement) 18
  - DB2TX.SAMPLE (table)
    - description 78
    - suppression 117
    - utilitaire de création 16
  - déclencheurs
    - création 68
    - description 30
  - décomposition des termes composés 262
  - définitions de la configuration de texte
    - affichage 107
    - modification 20
    - paramètres d'installation par défaut 19
  - DELETE INDEX EVENTS (commande)
    - exemple 104
    - syntaxe 139
  - démarrage de session d'affichage (fonction API)
    - description 251
    - procédure 123

- démarrage du serveur DB2
    - Extension Texte 185
  - DES\_BROWSE, option dans
    - DesGetSearchResultTable 244
  - DES\_EXT.H (fichier d'en-tête) 119
  - DES\_EXTENDED, option dans
    - DesOpenDocument 248
  - DES\_FAST, option dans
    - DesOpenDocument 248
  - DES\_MATCH, option dans
    - DesGetSearchResultTable 244
  - DES\_NOBROWSE, option dans
    - DesGetSearchResultTable 244
  - DES\_RANK, option dans
    - DesGetSearchResultTable 244
  - DES\_RANKANDMATCH, option dans
    - DesGetSearchResultTable 244
  - DES\_TEXTHANDLEONLY, option dans
    - DesGetSearchResultTable 244
  - DESL.INI 281
  - DesCloseDocument (fonction)
    - description 231
    - procédure 124
  - descripteur
    - CCSID (fonction) 202
    - définition et extraction d'informations 97
    - description 81
    - extraction d'une liste de descripteurs 206
    - extraction du CCSID, du format et de la langue 98
    - FORMAT (fonction) 205
    - initialisation 97
    - LANGUAGE (fonction) 209
    - modification du format et de la langue 98
    - pour fichiers externes 82
    - type distinct DB2TEXTFH 199
    - type distinct DB2TEXTH 199
    - utilisation de listes pour améliorer les performances 99
  - DesEndBrowseSession (fonction)
    - description 232
    - procédure 125
  - DesFreeBrowseInfo (fonction)
    - description 233
    - procédure 125
  - DesGetBrowseInfo (fonction)
    - description 234
    - procédure 123
  - DesGetMatches (fonction)
    - description 237
  - DesGetMatches (fonction) (*suite*)
    - procédure 124
  - DesGetSearchResultTable (fonction)
    - description 243
    - procédure 121
  - DESMODEL.INI 48
  - DesOpenDocument (fonction)
    - description 248
    - procédure 123
  - DESRESTB, création d'une table
    - résultat 245
  - DESSRV.INI 283
  - DesStartBrowseSession (fonction)
    - description 251
    - procédure 123
  - détection des incidents 23
  - DIRECTORY (mot clé)
    - affichage du paramètre en cours 110
    - dans CHANGE TEXT CONFIGURATION 135
    - ENABLE TEXT COLUMN (commande) 151
    - ENABLE TEXT TABLE (commande) 160
    - paramétrage par défaut 19
  - document
    - affichage des paramètres 111
    - CCSID 35
    - conversion de format 33
    - conversion de types de données 71
    - format (description) 31
    - format dans CHANGE TEXT CONFIGURATION 136
    - format dans ENABLE TEXT COLUMN 148
    - formats pris en charge 31
    - GET TEXT INFO (commande) 168
    - indexation 29
    - informations 111
    - langue 35
    - structure 48
    - types de données pris en charge 146
  - document (fermeture par fonction API)
    - description 231
    - procédure 124
  - documents DBCS, recherche dans 43
  - documents de type fichier à plat, prise en charge de section 51
  - documents HTML, prise en charge de section 51
  - documents structurés HTML 48
  - documents structurés XML 48
  - documents XML, prise en charge de section 52
- ## E
- environnement client-serveur 4
  - environnement multi-noeud 47
  - espace de journalisation insuffisant 70
  - Espace disque pour les index 45
  - Espace disque requis pour les index 45
  - espace libre pour informations d'affichage (fonction API)
    - description 233
    - procédure 125
  - Espace table 68
  - espaces table et groupes de noeuds 68
  - état d'un index
    - affichage 108
    - affichage de l'état en cours 165
    - réinitialisation 103
    - réinitialisation après erreur 103
  - état de l'index, affichage
    - exemple et données de sortie 108
    - syntaxe 165
  - état de l'index, réinitialisation
    - exemple 103
    - syntaxe 171
  - état de la mise à jour, affichage
    - exemple et données de sortie 108
    - syntaxe 165
  - état de la mise à jour, réinitialisation
    - exemple 103
    - syntaxe 171
  - état de la recherche, affichage
    - exemple et données de sortie 108
    - syntaxe 165
  - état de la recherche, réinitialisation
    - exemple 103
    - syntaxe 171
  - état du serveur DB2 Extension Texte 186
  - événements d'erreur
    - affichage 110
    - codes anomalie 321
    - commande GET INDEX STATUS 165
    - DELETE INDEX EVENTS 139

- événements d'erreur (*suite*)
    - données collectées 68
    - suppression 104
  - événements d'indexation
    - codes anomalie 321
  - événements d'indexation, suppression
    - exemple 104
    - syntaxe 139
  - exit utilisateur, conversion de format de document 33
  - EXPAND (mot clé) 221
  - extension de sons
    - description 265
    - exemple 93
  - Extension des termes pour la mise en évidence 266
  - Extensions
    - exemple d'utilisation 4
  - Extraction (analyse linguistique) 263
  - extraction d'informations d'affichage (fonction API)
    - description 234
    - procédure 123
  - extraction de correspondances (fonction API)
    - description 237
    - procédure 124
  - extraction de la table de résultats de recherche (fonction API)
    - description 243
    - procédure 121
- F**
- FFT, format de document 31
  - Fichier d'en-tête des\_ext.h 119
  - fichier de modèles de documents, contenu 51
  - fichiers de configuration 281
  - fichiers externes
    - activation 73
    - commande DISABLE TEXT FILES 142
    - commande ENABLE TEXT FILES 154
    - désactivation 116
    - descripteurs 82
    - extraction du chemin/nom de fichier dans un descripteur 98
    - FILE (fonction) 204
    - initialisation de descripteurs 208
    - mise à jour de l'index 102
    - modification du chemin/nom de fichier dans un descripteur 98
  - fichiers externes (*suite*)
    - renvoi ou modification d'un nom de fichier dans un descripteur 204
  - FILE (fonction)
    - exemple 98
    - syntaxe 204
  - fin de session d'affichage (fonction API)
    - description 232
    - procédure 125
  - fonction
    - conversion de types de données 71
    - définition du chemin pour les fonctions DB2 Extension Texte 82
    - fonctions API 119
    - fonctions de recherche 77
    - SET CURRENT FUNCTION PATH (instruction) 82
  - fonction d'accès
    - description 71
    - ENABLE TEXT COLUMN (commande) 146
  - fonctions
    - affinement d'une recherche 95
    - amélioration des performances de recherche 99
    - CCSID 202
    - chemin de fonction 82
    - CONTAINS 203
    - définition et extraction d'informations dans les descripteurs 97
    - description 77
    - FILE 204
    - FORMAT 205
    - HANDLE 206
    - HANDLE\_LIST 207
    - informations de référence 199
    - INIT\_TEXT\_HANDLE 208
    - LANGUAGE 209
    - NO\_OF\_DOCUMENTS 210
    - NUMBER\_OF\_MATCHES 211
    - présentation 201
    - RANK 212
    - recherche de texte 83
    - REFINE 213
    - SEARCH\_RESULT 214
    - spécification des arguments de recherche 86
    - SQLSTATE renvoyés 299
  - fonctions d'affichage 119
  - fonctions de recherche 119
  - FORMAT (fonction)
    - exemple 98
    - syntaxe 205
  - format de documents texte
    - conversion non prise en charge 33
    - dans CHANGE TEXT CONFIGURATION 136
    - description 31
    - ENABLE TEXT COLUMN (commande) 148
    - extraction des informations d'un descripteur 98
    - FORMAT (fonction) 205
    - FORMAT (mot clé) 136, 148
    - GET TEXT INFO (commande) 168
    - initialisation de descripteurs 208
    - liste 31
    - modification dans un descripteur 98
    - paramétrage par défaut 19
  - Format de documents texte 31
  - formats d'heure, prise en charge de section 50
  - formats de date, prise en charge de section 49
  - formats de nombres entiers, prise en charge de section 51
  - formats en virgule flottante pris en charge 50
  - forme de base (réduction des termes) 262
  - fréquence de mise à jour
    - description 46
    - GET INDEX SETTINGS (commande) 163
    - modification 103
    - paramétrage par défaut 20
    - syntaxe 46
    - UPDATEFREQ dans CHANGE INDEX SETTINGS 132
    - UPDATEFREQ dans CHANGE TEXT CONFIGURATION 134
  - FUNCTION (mot clé)
    - description 71
    - ENABLE TEXT COLUMN (commande) 146
  - FUZZY FORM OF (mot clé) 223
- G**
- GET ENVIRONMENT (commande)
    - exemple et données de sortie 107
    - syntaxe 162

- GET INDEX SETTINGS (commande)
  - exemple et données de sortie 110
  - syntaxe 163
- GET INDEX STATUS (commande)
  - exemple et données de sortie 108
  - syntaxe 165
- GET STATUS (commande)
  - exemple et données de sortie 106
  - syntaxe 166
- GET TEXT CONFIGURATION (commande)
  - exemple et données de sortie 107
  - syntaxe 167
- GET TEXT INFO (commande)
  - exemple et données de sortie 111
  - syntaxe 168
- groupe de noeuds multipartition 7
- H**
- HANDLE (fonction)
  - procédure 99
  - syntaxe 206
- HANDLE\_LIST (fonction)
  - procédure 99
  - syntaxe 207
- HTML, format de document 31
- I**
- IN SAME PARAGRAPH AS (mot clé) 222
- IN SAME SENTENCE AS (mot clé) 222
- inclusion du fichier des\_ext.h 119
- index
  - affichage des paramètres en cours 110
  - calcul de taille 45
  - CASE\_ENABLED (option) 44
  - CHANGE TEXT CONFIGURATION (commande) 133
  - commande CHANGE INDEX SETTINGS 131
  - commande GET INDEX STATUS 165
  - commande GET TEXT CONFIGURATION 167
  - création de différents types pour une colonne de texte 69
  - exact 43
  - fréquence de mise à jour 46
  - index (suite)
    - généralités 29
    - gestion 101
    - GET INDEX SETTINGS (commande) 163
    - INDEXOPTION dans ENABLE TEXT COLUMN 149
    - INDEXOPTION dans ENABLE TEXT TABLE 158
    - INDEXPROPERTY dans CHANGE TEXT CONFIGURATION 134
    - INDEXTYPE dans ENABLE TEXT COLUMN 149
    - INDEXTYPE dans ENABLE TEXT TABLE 158
    - linguistique 41
    - mise à jour immédiate 101
    - mise à jour périodique des index 46
    - mise à jour pour des fichiers externes 102
    - modification de la configuration du texte 20
    - modification de la fréquence de mise à jour 103
    - modification des paramètres actuels 102
    - modification du type d'index 44
    - multiple, utilisation 44
    - Ngram 43
    - paramétrage par défaut 20
    - planification 29
    - réorganisation 105
    - sauvegarde et restauration 22
    - table à index commun 44
    - TABLESPACE dans CHANGE TEXT CONFIGURATION 134
    - types 40
    - UPDATE INDEX (commande) 172
  - index exact
    - description 43
    - valeurs par défaut des options de recherche 223
  - index linguistique
    - description 41
    - valeurs par défaut des options de recherche 223
  - index multiple, utilisation 44
  - index Ngram
    - CASE\_ENABLED (option) 44
    - description 43
    - valeurs par défaut des options de recherche 223
- Indexation (analyse linguistique) 255
- INDEXOPTION (mot clé)
  - dans CHANGE TEXT CONFIGURATION 134
  - ENABLE TEXT COLUMN (commande) 149
  - ENABLE TEXT TABLE (commande) 158
- INDEXPROPERTY (mot clé)
  - ENABLE TEXT COLUMN (commande) 150
  - ENABLE TEXT TABLE (commande) 159
- INDEXTYPE (mot clé)
  - dans CHANGE TEXT CONFIGURATION 134
  - ENABLE TEXT COLUMN (commande) 149
  - ENABLE TEXT TABLE (commande) 158
- informations de mise en évidence
  - flux de données 124
  - syntaxe du flux de données 237
- informations sur les documents de texte
  - affichage du paramètre en cours 111
  - CCSID 35
  - format 31
  - GET TEXT INFO (commande) 168
  - langue 35
  - types 31
- INIT\_TEXT\_HANDLE (fonction)
  - exemple 97
  - syntaxe 208
- initialisation d'un descripteur
  - INIT\_TEXT\_HANDLE (fonction) 208
  - procédure 97
- instances
  - commande pour obtenir la liste des, TXILIST 179
  - création 21
  - suppression 21
  - variables d'environnement 17
- interface de programmation d'applications (API)
  - arrêt d'une session d'affichage 125
  - codes retour 291
  - démarrage d'une session d'affichage 123

- interface de programmation d'applications (API) *(suite)*
  - DesCloseDocument (fonction) 231
  - DesEndBrowseSession (fonction) 232
  - DesFreeBrowseInfo (fonction) 233
  - DesGetBrowseInfo (fonction) 234
  - DesGetMatches (fonction) 237
  - DesGetSearchResultTable (fonction) 243
  - DesOpenDocument (fonction) 248
  - DesStartBrowseSession (fonction) 251
- exemple de programme 253
- extraction d'informations d'affichage 123
- extraction d'un pointeur sur des informations de mise en évidence 237
- extraction de correspondances 124
- fermeture d'un document 124
- fonctions d'affichage 119
- fonctions de recherche 119
- informations de mise en évidence 237
- informations de référence 229
- libération de la mémoire des informations d'affichage 125
- messages 299
- obtention d'une table de résultats de recherche 121
- ouverture d'un document pour affichage 123
- présentation 120
- récapitulatif 229
- recherche de texte 121
- interpréteur de commandes
  - aide 60
  - DB2TX (commande) 130
  - démarrage 59
  - QUIT (commande) 169
- L**
- LANGUAGE (fonction)
  - exemple 98
  - syntaxe 209
- LANGUAGE (mot clé) 135, 147
- langue des documents de texte
  - dans un argument de recherche 90
- langue des documents de texte *(suite)*
  - description 35
  - extraction des informations d'un descripteur 98
  - GET TEXT INFO (commande) 168
  - initialisation de descripteurs 208
  - LANGUAGE (fonction) 209
  - liste 35
  - modification dans un descripteur 98
  - paramétrage par défaut 19
- M**
- masquage
  - analyse linguistique 265
  - dans une séquence de recherche 87
- masquage de caractères 265
- messages 299
- Microsoft, format de document 31
- mise à jour d'un index de texte
  - modification de la fréquence 103
  - périodique 46
  - UPDATEFREQ dans CHANGE INDEX SETTINGS 132
  - UPDATEFREQ dans CHANGE TEXT CONFIGURATION 134
- mise en route 25
- modèle de document
  - description 48
  - MODEL (mot clé pour la syntaxe de recherche) 219
  - modification du fichier des modèles de documents 48
  - SECTION (mot clé pour la syntaxe de recherche) 219
  - valeur d'attribut dans la syntaxe de recherche 220
- Modèle de programme API 253
- Modèle de programme DESSAMP1 253
- modèle de table
  - commande TXSAMPLE 184
  - description 78
  - suppression 117
- modèles de fonctions DB2 Extension
  - Texte
    - exécution 77
- mots vides
  - dans une analyse de texte de base 262
  - description 29
- mots vides *(suite)*
  - listes 54
  - modification d'un fichier de mots vides 55
- N**
- NO\_OF\_DOCUMENTS (fonction)
  - syntaxe 210
- NODE (mot clé)
  - ENABLE TEXT COLUMN (commande) 152
  - ENABLE TEXT TABLE (commande) 160
- noeud
  - affectation 7
  - environnement multi-noeud 47
  - groupes de noeuds et espaces table 68
  - TXNADD, commande d'ajout de noeud 181
  - TXNCHECK, commande de vérification de noeuds 182
  - TXNDROP, commande de suppression de noeud 183
- noeud logique, affectation 7
- noeud physique, affectation 7
- Noms de fichiers dictionnaires 54
- Normalisation des termes 257
- NORMALIZED (mot clé)
  - ENABLE TEXT COLUMN (commande) 150
- NOT
  - mot clé dans argument de recherche 222
  - opérateur booléen 91
- NUMBER\_OF\_MATCHES (fonction), syntaxe 211
- O**
- occurrences d'un terme recherché 211
- ON NODE (mot clé)
  - ENABLE TEXT COLUMN (commande) 152
  - ENABLE TEXT TABLE (commande) 160
- opérateur & (AND) dans un argument de recherche
  - syntaxe des arguments de recherche 221
  - utilisation 86
- opérateur | (OR) dans un argument de recherche
  - syntaxe des arguments de recherche 221
  - utilisation 86

opérateurs booléens  
& (AND) et | (OR) 86  
NOT 91

OR (opérateur booléen) 86  
ouverture de document (fonction  
API)  
description 248  
procédure 123

## P

paramètres de langue, liste 54  
paramètres LOGPRIMARY,  
LOGSECOND et LOGFILSIZ dans  
DB2 UDB 70  
performances, amélioration 99  
pointeur sur la liste des descripteurs  
(type distinct  
DB2TEXTFHLISTP) 200  
Pointeur sur la liste des descripteurs  
(type distinct  
DB2TEXTHLISTP) 200  
PRECISE FORM OF (mot clé) 223  
présentation de DB2 Extension  
Texte 3  
prise en charge des bases de  
données partitionnées 7  
Profils de shell 17

## Q

QUIT (commande)  
syntaxe 169  
utilisation 75

## R

rang  
à partir de  
DesGetSearchResultTable 121  
dans un résultat de recherche 85  
RANK (fonction)  
exemple 85  
syntaxe 212  
recherche de correspondance  
partielle 41  
recherche de correspondance  
partielle, exemple 92  
recherche de texte  
affinement d'une recherche 95  
amélioration des  
performances 99  
création d'une liste des  
documents trouvés 207  
création d'une requête 84  
exemple de programme 253  
extraction du nombre de  
correspondances trouvées 85

recherche de texte (*suite*)  
extraction du rang d'un  
document trouvé 85  
fonction SEARCH\_RESULT 214  
généralités 83  
REFINE (fonction) 213  
syntaxe 217  
utilisation de l'API 121  
recherche exacte 43  
recherche hybride, exemple 94  
recherche par thésaurus  
268  
commande TXTHESN 190  
compilation d'un fichier de  
définition de thésaurus 188  
compilation d'un fichier de  
définition de thésaurus  
Ngram 190  
création d'un thésaurus 273  
création d'un thésaurus  
Ngram 278  
exemple 93  
syntaxe 220  
THESAURUS (mot clé) 220  
TXTHESC (commande) 188  
recherche segmentée, exemple 93  
recherche sur texte libre  
exemple 94  
reconnaissance de phrases 257  
réduction des termes à leur forme de  
base 262  
REFINE (fonction)  
exemple 95  
syntaxe 213  
Règles applicables aux arguments de  
recherche 226  
REORGANIZE INDEX (commande)  
exemple 105  
syntaxe 170  
Répertoire d'index  
GET INDEX SETTINGS  
(commande) 163  
RESET INDEX STATUS (commande)  
exemple 103  
syntaxe 171  
Restrictions applicables aux  
arguments de recherche 226  
RESULT LIMIT (mot clé) 221

## S

sauvegarde et restauration 22  
SEARCH\_RESULT (fonction)  
exemple 99  
syntaxe 214

sections de document  
activation de la prise en charge  
de section 48  
DESMODEL.INI 48  
documents de type fichier à  
plat 51  
documents HTML 51  
Documents XML 52  
exemple de recherche 89  
fichier de modèles de documents,  
contenu 51  
formats d'heure pris en  
charge 50  
formats de date pris en  
charge 49  
formats de nombres entiers pris  
en charge 51  
formats en virgule flottante pris  
en charge 50  
MODEL (mot clé pour la syntaxe  
de recherche) 219  
SECTION (mot clé pour la  
syntaxe de recherche) 219  
sections avec attribut 49  
valeur d'attribut dans la syntaxe  
de recherche 220  
Séparation des mots 43  
Séparation des phrases 43  
serveur  
affichage de l'état 186  
arrêt 187  
commande TXICRT 176  
commande TXIDROP 178  
commande TXIPCLEAN 180  
commande TXNADD 181  
commande TXNCHECK 182  
commande TXNDROP 183  
commande TXSAMPLE 184  
commande TXSTART 185  
commande TXSTATUS 186  
commande TXSTOP 187  
commande TXVERIFY 197  
commande TXWIZARD 198  
configuration et gestion 21  
démarrage 185  
trace des incidents 23  
TXILIST (commande) 179  
TXTRACE(commande) 192  
SET CURRENT FUNCTION PATH  
(instruction) 82  
Spécification du niveau des termes  
dans un thésaurus 221  
SQLSTATE renvoyés par les  
fonctions DB2 Extension  
Texte 299

- STEMMED FORM OF (mot clé) 223
- structure de document
  - activation de la prise en charge de section 48
  - exemple de recherche 89
  - MODEL (mot clé pour la syntaxe de recherche) 219
  - SECTION (mot clé pour la syntaxe de recherche) 219
  - valeur d'attribut dans la syntaxe de recherche 220
- support de langue nationale 37
- suppression d'une instance
  - commande TXIDROP 178
  - procédure 21
- synonymes
  - dans un argument de recherche 90
  - description 264
  - SYNONYM FORM OF (mot clé) 223
- synaxe du flux de données 238
- T**
  - table à index commun
    - commande ENABLE TEXT TABLE 157
    - création 63
    - description 44, 45
  - table de configuration
    - affichage 107
    - commande CHANGE TEXT CONFIGURATION 133
    - commande GET TEXT CONFIGURATION 167
    - création 62
  - table de journalisation
    - affectation à un espace table 68
    - création 68
    - description 30
    - extraction des événements d'erreur 110
  - table de texte
    - activation d'une colonne dans une table volumineuse 70
    - commande ENABLE TEXT TABLE 157
    - DISABLE TEXT TABLE (commande) 143
    - sauvegarde et restauration 22
  - Table résultat 245
  - tables volumineuses, activation 70
  - TERM OF (mot clé) 221
  - termes composés (éclatement) 262
  - TEXTINDEXES (vue de catalogue) contenu 112
  - TEXTINDEXES (vue de catalogue) (*suite*)
    - création 62
    - suppression 117
  - trace des incidents
    - commande TXTRACE 192
    - configuration 23
  - traitement linguistique
    - affichage 266
    - analyse de texte de base 256
    - correspondance étendue 266
    - description 255
    - éclatement des termes composés 262
    - extension des termes 266
    - extension phonétique 265
    - filtrage des mots vides 262
    - lors de l'indexation 255
    - masquage 265
    - masquage de caractères 265
    - masquage de mots 265
    - pour l'extraction 263
    - réduction des termes à leur forme de base 262
    - synonymes 264
  - TXICRT (commande)
    - création d'une instance DB2 Extension Texte 21
    - syntaxe 176
  - TXIDROP (commande)
    - syntaxe 178
  - TXILIST (commande)
    - syntaxe 179
  - TXIPCLEAN (commande)
    - syntaxe 180
  - TXNADD (commande)
    - syntaxe 181
  - TXNCHECK (commande)
    - syntaxe 182
  - TXNDROP (commande)
    - syntaxe 183
  - TXSAMPLE (commande)
    - procédure 16
    - syntaxe 184
  - TXSAMPLE.UDF
    - exécution 77
  - TXSTART (commande)
    - procédure 22
    - syntaxe 185
  - TXSTATUS (commande)
    - procédure 22
    - syntaxe 186
  - TXSTOP (commande)
    - procédure 22
    - syntaxe 187
  - TXTHESC (commande)
    - syntaxe 188
  - TXTHESN (commande)
    - syntaxe 190
  - TXTRACE (commande)
    - procédure 23
    - syntaxe 192
  - TXVERIFY (commande)
    - création d'une base de données exemple 16
    - syntaxe 197
  - TXWIZARD (commande)
    - syntaxe 198
  - type d'index, modification
    - création de différents types pour une colonne de texte 69
    - modification 44
  - type de données DATALINK 72
  - types d'index de texte
    - CASE\_ENABLED (option) 44
    - commande CHANGE TEXT CONFIGURATION 134
    - exact 43
    - GET INDEX SETTINGS (commande) 163
    - INDEXTYPE dans CHANGE TEXT CONFIGURATION 134
    - INDEXTYPE dans ENABLE TEXT COLUMN 149
    - INDEXTYPE dans ENABLE TEXT TABLE 158
    - linguistique 41
    - Ngram 43
    - paramétrage par défaut 20
    - valeurs par défaut des options de recherche 223
  - types de données de documents
    - texte
      - activation de colonnes contenant le type de données DATALINK 72
      - fonction de conversion 71, 72
      - pris en charge 146
  - types distincts 199
- U**
  - UPDATE INDEX (commande)
    - exemple 101
    - syntaxe 172
  - UPDATEFREQ (mot clé)
    - CHANGE INDEX SETTINGS 132
    - dans CHANGE TEXT CONFIGURATION 134
    - ENABLE TEXT COLUMN (commande) 150

UPDATEFREQ (mot clé) *(suite)*  
  ENABLE TEXT TABLE  
    (commande) 159

UPDATEINDEX (mot clé)  
  affichage du paramètre en  
    cours 110  
  dans CHANGE TEXT  
    CONFIGURATION 135  
  ENABLE TEXT COLUMN  
    (commande) 152  
  GET INDEX SETTINGS  
    (commande) 163  
  paramétrage par défaut 20

## V

variables

  affichage des variables  
    d'environnement 107  
  commande GET  
    ENVIRONMENT 162  
  description des variables  
    d'environnement 17

variables d'environnement 17

  affichage 107  
  commande GET  
    ENVIRONMENT 162  
  description 17

vérification de l'installation 16

vue de catalogue

  contenu 112  
  création 62  
  suppression 117

## W

WordPerfect, format de  
  document 31

## X

XML, format de document 31

---

## Comment prendre contact avec IBM

En cas d'incident technique, consultez et effectuez les opérations suggérées par le manuel *Troubleshooting Guide* avant de vous adresser au Service client DB2. Ce manuel indique les informations susceptibles d'aider le Service clients à mieux répondre à vos besoins.

Pour obtenir des informations ou commander des produits DB2, prenez contact avec votre partenaire commercial IBM.

Si vous vivez aux Etats-Unis, vous disposez des numéros suivants :

- 1-800-237-5511 pour obtenir le Service clients,
- 1-888-426-4343 pour connaître les options de service disponibles.

---

### Infos produit

If you live in the U.S.A., then you can call one of the following numbers:

- Pour commander des produits ou obtenir des informations générales, composez le 1-800-IBM-CALL (1-800-426-2255) ou 1-800-3IBM-OS2 (1-800-342-6672).
- Pour commander des manuels, composez le 1-800-879-2755.

**<http://www.ibm.com/software/data/>**

Les pages DB2 World Wide Web fournissent des informations sur DB2, des descriptions de produit, les programmes de formation et d'autres informations.

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

DB2 Product and Service Technical Library permet d'accéder à des forums Q&A (questions/réponses), d'obtenir des correctifs et les dernières informations techniques sur DB2.

**Remarque :** Il est possible que ces informations ne soient disponibles qu'en anglais.

**<http://www.elink.ibm.com/pbl/pbl/>**

Le site Web de commande internationale de manuels fournit les informations correspondantes.

**<http://www.ibm.com/education/certify/>**

Le programme Professional Certification Program du site Web IBM fournit des informations sur les tests de certification concernant différents produits IBM, dont DB2.

**ftp.software.ibm.com**

Etablissez une connexion anonyme. Des démonstrations, des correctifs, des informations et des outils associés à DB2 ou à des produits connexes sont disponibles dans le répertoire /ps/products/db2.

**comp.databases.ibm-db2, bit.listserv.db2-l**

Ces groupes de discussion sont accessibles à tous ceux qui souhaitent partager leurs expériences sur les produits DB2.

**Sur CompuServe : GO IBMDB2**

Exécutez cette commande pour accéder aux forums IBM DB2. Tous les produits DB2 sont pris en charge sur ces forums.

En dehors des Etats-Unis, pour savoir comment prendre contact avec IBM, consultez l'annexe A du manuel IBM Software Support Handbook. Pour accéder à ce document, allez sur le site Web :

<http://www.ibm.com/support/>, puis effectuez une recherche sur le mot clé «handbook».

**Remarque :** Dans certains pays, les distributeurs agréés peuvent contacter leur centre d'assistance au lieu de prendre contact avec le centre de support IBM.





SC11-1683-00



Spine information:



DB2 Universal Database

DB2 Extension Texte Administration et  
programmation

Version 7