



# Chapter

# 21

## SQL Server全文檢索搜尋

### 21-1 | SQL Server 全文檢索搜尋

#### Memo

本章測試的 SQL 指令是使用【教務系統】資料庫，請重新啟動 SSMS 執行本書範例「Ch21\Ch21\_School.sql」的 SQL 指令碼檔案，來建立本章測試所需的資料庫、資料表和記錄資料。

SQL Server 整合全文檢索搜尋（Integrated Full-text Search，iFTS）是一種企業級搜尋功能，其高效能與強大的功能，可以替資料庫應用程式建立無與倫比的搜尋能力。

#### 21-1-1 全文檢索搜尋的基礎

全文檢索搜尋（Full-text Search）可以讓我們使用關鍵字尋找 SQL Server 資料庫中的文字內容資料，一種以特定語言規則為基礎（指分析過濾出文字資料中單字或片語的規則）的單字或片語搜尋。例如：Yahoo!或 Google 搜尋引擎的搜尋就屬於全文檢索搜尋，可以執行特定單字或單字組合的快速搜尋。

## 認識全文檢索搜尋

全文檢索搜尋特別適用在各種商務案例，可以用來查詢大量非結構化文字資料，例如：當有數百萬筆記錄資料時，使用 LIKE 運算子執行查詢可能需要花費好幾分鐘才能傳回結果；全文檢索搜尋通常只需幾秒鐘的時間，即可得到搜尋結果。

在 SQL Server 執行全文檢索搜尋前，我們需要建立全文檢索目錄和索引，其說明如下所示：

- 全文檢索目錄（Full-text Catalogs）：方便我們管理執行全文檢索搜尋所需的眾多全文檢索索引。在全文檢索目錄包含 0 至多個全文檢索索引，這些索引資料是儲存在 SQL Server 資料庫，全文檢索目錄本身只是一個虛擬物件，並不屬於任何檔案群組，只是在邏輯上參考一組全文檢索索引。
- 全文檢索索引（Full-text Indexes）：我們可以針對資料庫的資料表建立全文檢索索引，其內容是儲存單字或片語在指定記錄的位置資訊。簡單的說，SQL Server 就是使用這些索引資訊來執行全文檢索搜尋。

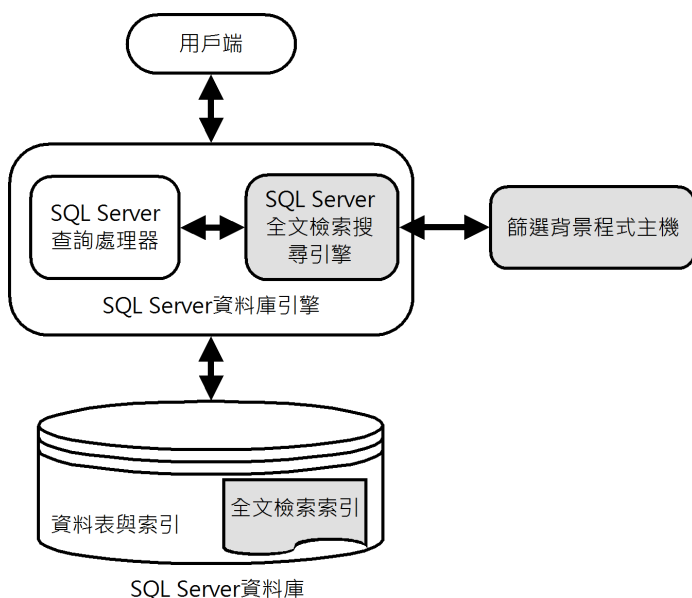
SQL Server 全文檢索搜尋除了可以替 char、varchar、nvarchar 和 xml 等類型的欄位建立全文檢索索引外，還能夠針對儲存在 varbinary(max) 二進位類型欄位的 Word 等文件內容建立全文檢索索引。

## SQL Server 整合全文檢索搜尋架構

SQL Server 資料庫引擎整合「全文檢索搜尋引擎」（Full-text Query Engine），可以大幅減少服務之間的資料交換，使用更有效率的最佳化方式來執行全文檢索搜尋，其架構如下圖所示。

圖例的 SQL Server 全文檢索搜尋引擎負責母體擴展與管理全文檢索目錄，支援全文檢索索引的建立和 SQL Server 全文檢索搜尋。

全文檢索引擎是透過「篩選背景程式主機」（Filter Daemon Host）來過濾資料表的記錄資料，分析出文字資料中的單字或片語來建立全文檢索索引。



## 21-1-2 下載與安裝 SQL Server 範例資料庫

SQL Server 安裝程式並不包含 AdventureWorks 範例資料庫，微軟已經將這些範例資料庫放在 GitHub，使用者如果需要，可以自行從 GitHub 下載和安裝範例資料庫。

本章內容是使用 SQL Server 的 AdventureWorks 範例資料庫為例，我們需要先下載和安裝 SQL Server 範例資料庫。請啟動瀏覽器進入 GitHub 來下載範例資料庫，在本書截稿前只支援至 2019 版，如下所示：

- <https://github.com/Microsoft/sql-server-samples/releases/tag/adventureworks>

### AdventureWorks (OLTP) full database backups

[AdventureWorks2019.bak](#)

[AdventureWorks2017.bak](#)

[AdventureWorks2016.bak](#)

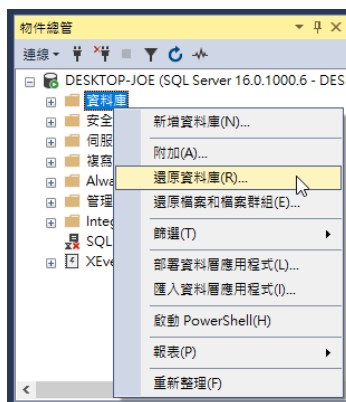
[AdventureWorks2016\\_EXT.bak](#)

Download size is 883 MB. This is an extended version of AdventureWorks, designed to showcase SQL Server 2016 features. To see the features in action, run the [SQL Server 2016 sample scripts](#) on this database.

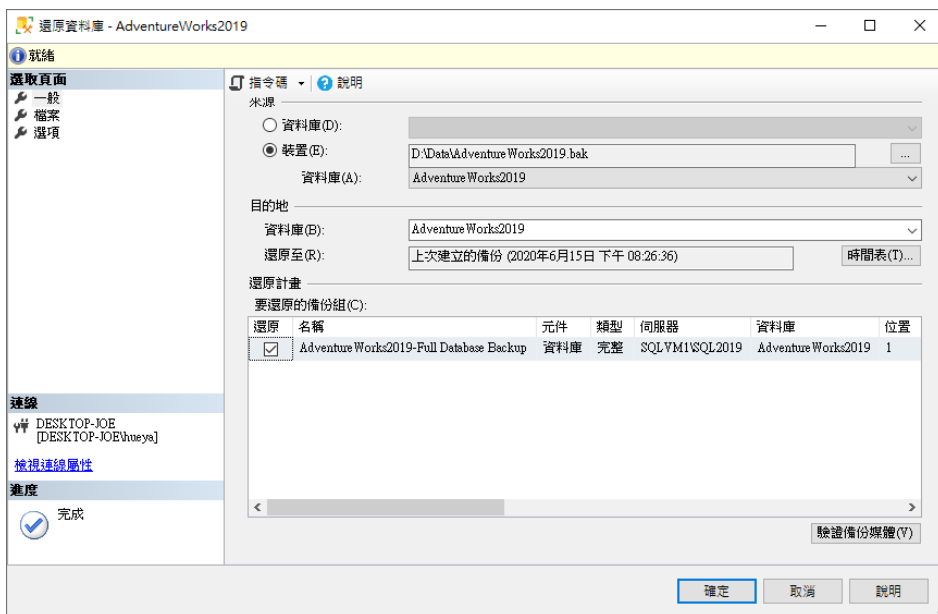
在上述範例資料庫下載網頁點選【AdventureWorks201?.bak】超連接，下載 SQL Server 版的範例資料庫 AdventureWorks，這是資料庫備份檔。

在成功下載副檔名.bak 的備份檔後，我們需要啟動 Management Studio 執行資料庫還原功能來安裝範例資料庫 AdventureWorks，其步驟如下所示：

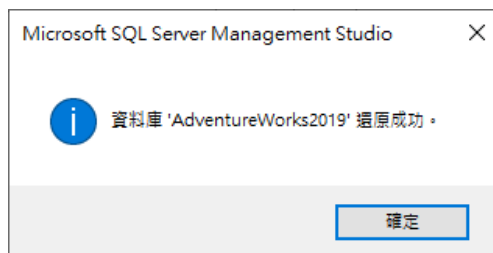
1 請啟動 Management Studio 建立連線後，在「物件總管」視窗的【資料庫】上，執行【右】鍵快顯功能表的【還原資料庫】命令。



2 在「還原資料庫」對話方塊選【裝置】後，按之後按鈕選擇加入下載檔案【AdventureWorks201?.bak】，可以在下方看到還原的資料庫名稱，按【確定】鈕開始還原。



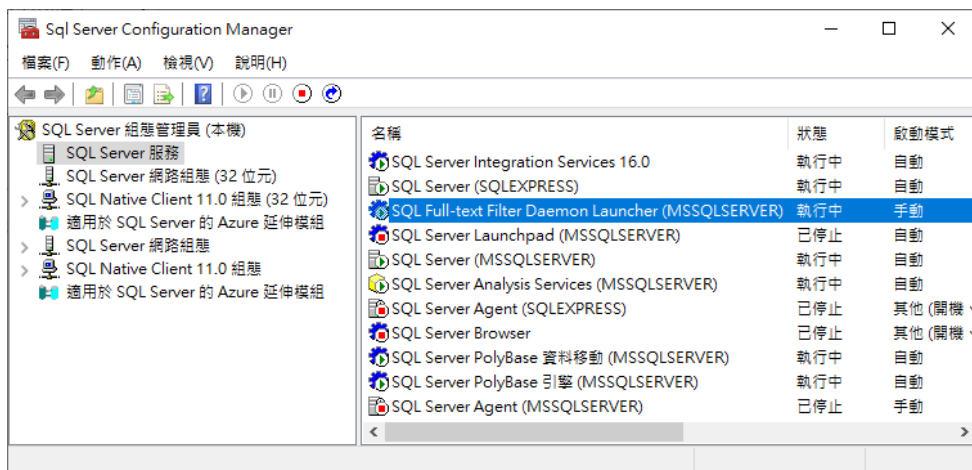
- 3 稍等一下，可以看到還原成功的訊息視窗，按【確定】鈕完成 AdventureWorks 範例資料庫的安裝。



- 4 為了方便說明本章 SQL 指令，請執行【右】鍵快顯功能表的【重新命名】命令，將範例資料庫更名為【AdventureWorks】。

### 21-1-3 啟用 SQL Server 的全文檢索搜尋

SQL Server 需要在 SQL Server 組態管理員執行「篩選背景程式啟動器服務」（SQL Full-text Filter Daemon Launcher），才能啟動篩選背景程式主機來啟用 SQL Server 的全文檢索搜尋，如下圖所示：



上述圖例的 SQL Full-text Filter Daemon Launcher 服務是正在執行中。

當啟用 SQL Server 全文檢索搜尋後，SQL Server 資料庫預設啟用全文檢索搜尋，如果發現資料庫沒有啟用全文檢索搜尋，例如：附加舊版資料庫時，我們可以使用系統預存程序啟用資料庫的全文檢索搜尋，其基本語法如下所示：

```
EXEC sp_fulltext_database {'enable' | 'disable'}
```

上述系統預存程序的參數為 **enable**，表示啟用全文檢索搜尋；**disable** 是取消全文檢索搜尋。



### SQL 指令碼檔：Ch21\_1\_3.sql

請使用系統預存程序啟用【教務系統】資料庫的全文檢索搜尋，如下所示：

```
USE 教務系統  
GO  
EXEC sp_fulltext_database 'enable'
```

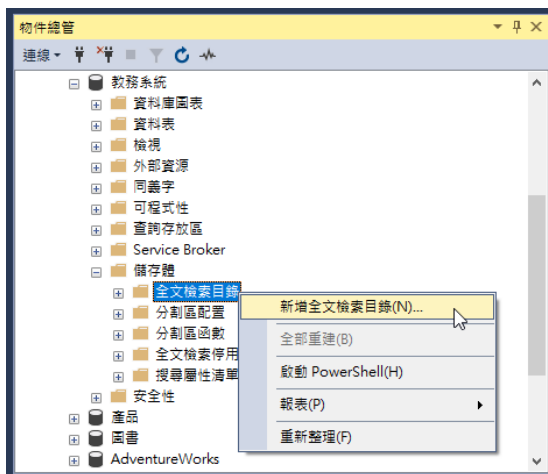
## 21-2 | 全文檢索目錄的建立

在 SQL Server 執行資料庫全文檢索搜尋的第一步，就是建立全文檢索目錄，我們可以使用 Management Studio 或 T-SQL 指令來建立全文檢索目錄。

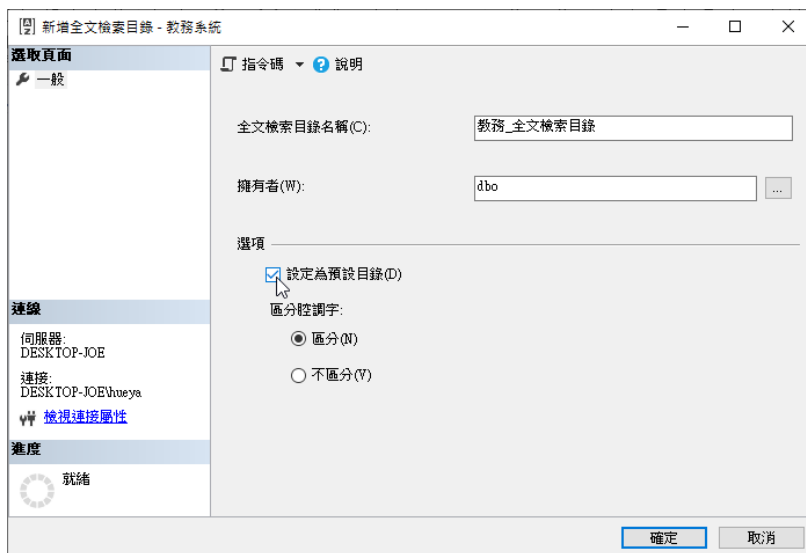
### 21-2-1 使用 Management Studio 建立全文檢索目錄

在 Management Studio 新增全文檢索目錄是一件十分簡單的工作。例如：在【教務系統】資料庫新增名為【教務\_全文檢索目錄】的全文檢索目錄，其步驟如下所示：

- 1 請啟動 Management Studio 建立連線後，在「物件總管」視窗展開【資料庫】下【教務系統】資料庫的【儲存體】項目，可以看到【全文檢索目錄】項目。
- 2 在【全文檢索目錄】上執行【右】鍵快顯功能表的【新增全文檢索目錄】命令。



- 3 在「新增全文檢索目錄」對話方塊輸入全文檢索目錄名稱【教務\_全文檢索目錄】後，選擇【擁有者】為【dbo】。

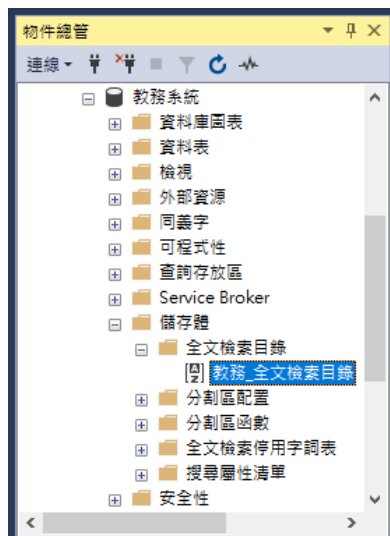


- 4 在下方選項勾選是否是預設目錄（當新增全文檢索索引時，如果沒有指定，預設就是新增至此目錄），和是否區分腔調字，按【確定】鈕建立全文檢索目錄。

## Memo

腔調字是指文字上如果有腔調符號時，是否將它和沒有腔調符號的文字視為相同的文字，例如：resumé和 resume 是否是相同字。

在 Management Studio 的「物件總管」視窗，展開【教務系統】資料庫下的【全文檢索目錄】項目，可以看到新建名為【教務\_全文檢索目錄】的全文檢索目錄，如右圖所示：



## 21-2-2 使用 T-SQL 指令建立全文檢索目錄

T-SQL 語言是使用 CREATE FULLTEXT CATALOG 指令建立全文檢索目錄，其基本語法如下所示：

```
CREATE FULLTEXT CATALOG 全文檢索目錄名稱  
[ AUTHORIZATION 擁有者名稱]  
[ AS DEFAULT ]  
[ WITH ACCENT_SENSITIVITY = { ON | OFF } ]
```

上述語法可以建立名為【全文檢索目錄名稱】的全文檢索目錄，AUTHORIZATION 關鍵字指定目錄的擁有者。AS DEFAULT 子句指定為預設的全文檢索目錄，WITH ACCENT\_SENSITIVITY 子句指定是否區分腔調字，ON 是區分；OFF 為不區分。



**SQL 指令碼檔：Ch21\_2\_2.sql**

請在【AdventureWorks】資料庫建立名為【產品文件\_全文檢索目錄】的全文檢索目錄，並且指定它為預設目錄，如下所示：

```
USE AdventureWorks
GO
CREATE FULLTEXT CATALOG 產品文件_全文檢索目錄
AS DEFAULT
```

## 21-3 | 建立全文檢索索引

在資料庫建立全文檢索目錄後，就可以建立指定資料表的全文檢索索引，我們可以将資料庫中的一至多個資料表建立全文檢索索引。在資料表建立全文檢索索引需要滿足一些條件，如下所示：

- 資料表需要擁有主索引鍵，或唯一值索引鍵，不過，索引鍵不可以是欄位組合的複合鍵。
- 每一個資料表只能建立一個全文檢索索引，但是可以包含多個資料表欄位。
- 建立全文檢索索引的欄位類型只能是：char、varchar、nchar、nvarchar、text、ntext、image、xml 和 varbinary。

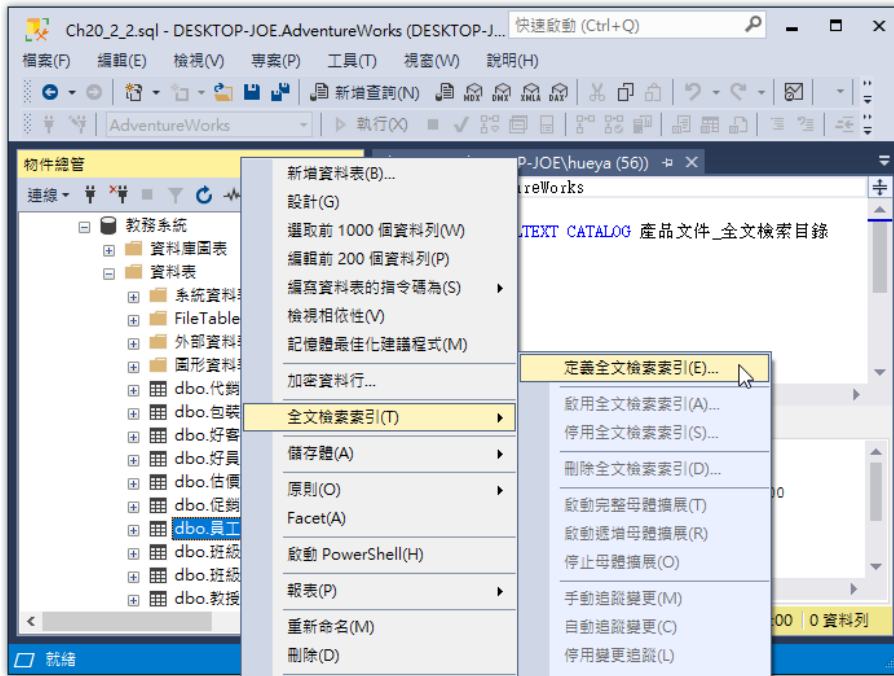
### 21-3-1 使用 Management Studio 建立全文檢索索引

在 Management Studio 提供全文檢索索引精靈來幫助我們建立指定資料表的全文檢索索引。例如：在【教務系統】資料庫的【員工】資料表建立全文檢索索引，其步驟如下所示：

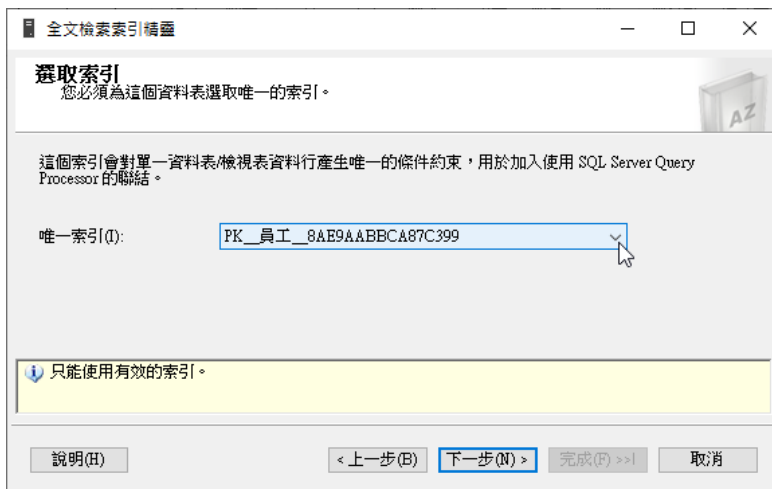
- 1 請啟動 Management Studio 建立連線後，在「物件總管」視窗展開【資料庫】下【教務系統】資料庫的【員工】資料表。



- 在【員工】資料表上執行【右】鍵快顯功能表的「全文檢索引>定義全文檢索引」指令，可以啟動全文檢索引精靈，看到歡迎畫面。



- 在歡迎畫面按【下一步】鈕，在資料表選取唯一索引。



- 4 在【唯一索引】欄選擇索引，通常就是主索引鍵，按【下一步】鈕勾選加入全文檢索索引的欄位。

**全文檢索索引精靈**

**選取資料表資料行**  
選取要適用於全文檢索查詢，以字元或以影像為主的資料行。

可用的資料行 ▲	斷詞工具的語言	類型資料行	統計語意
<input type="checkbox"/> 身份證字號			<input type="checkbox"/>
<input checked="" type="checkbox"/> 姓名			<input type="checkbox"/>
<input checked="" type="checkbox"/> 城市			<input type="checkbox"/>
<input checked="" type="checkbox"/> 街道			<input type="checkbox"/>
<input checked="" type="checkbox"/> 電話			<input type="checkbox"/>

說明(H)    < 上一步(B)    **下一步(N) >**    完成(F) >>I    取消

- 5 精靈會自動列出資料表可用的欄位清單，在勾選後，按【下一步】鈕選擇當資料表或檢視變更時，追蹤其變更的方式。

**全文檢索索引精靈**

**選取變更追蹤**  
選擇如何追蹤對資料表和檢視表的更新。

在這個資料表/檢視表發生變更時追蹤變更:

☒ 自動(A)  
☐ 手動(M)  
☐ 不要追蹤變更(D)

☒ 建立索引後啟動完整母體擴展(S)

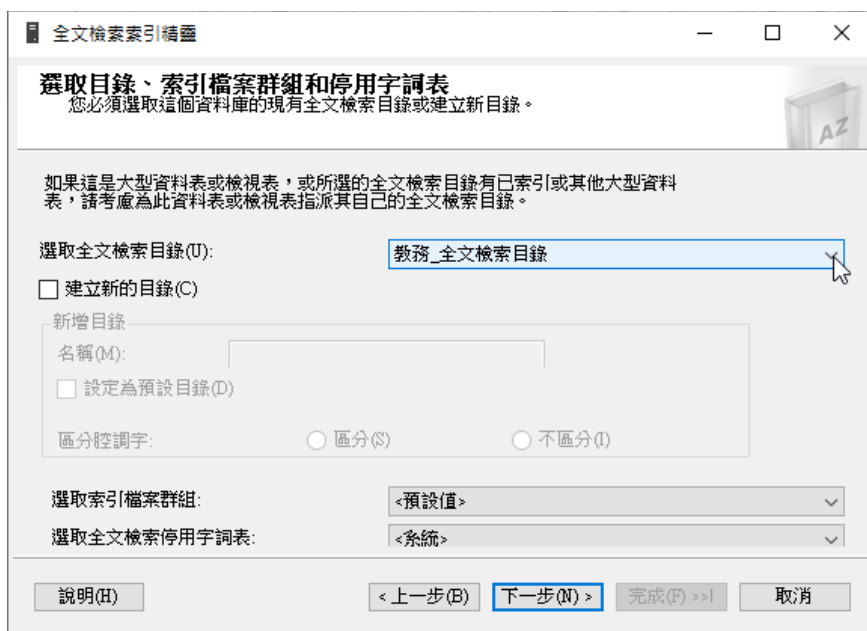
您定義自動或手動變更追蹤時，會進行索引的完整母體擴展。若要避免在這個精靈結束時的母體擴展，請選取 [不要追蹤變更] 選項，並取消選取 [建立索引後啟動完整母體擴展] 核取方塊。

說明(H)    < 上一步(B)    **下一步(N) >**    完成(F) >>I    取消

上述精靈步驟的三個變更追蹤選項說明，如下所示：

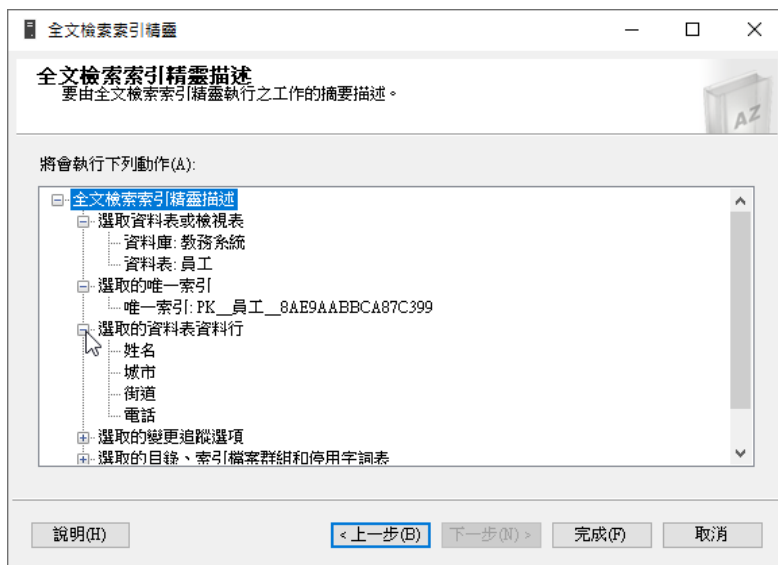
- 自動（Automatically）：當資料表的全文索引欄位值變更時，SQL Server 自動更新全文檢索引，此為預設值。
- 手動（Manually）：手動或使用 SQL Server Agent 以排程來定時更新全文檢索引，當自動變更追蹤會造成系統效能下降時，我們可以改為手動變更追蹤，以避免更新全文檢索引影響系統效能。
- 不要追蹤變更（Do not track changes）：SQL Server 不追蹤變更，此時只能使用 T-SQL 指令 ALTER FULLTEXT INDEX 指令，加上 START FULL 或 INCREMENTAL POPULATION 子句來母體擴展整個全文檢索引。

- 6 請使用預設值，不用更改，按【下一步】鈕選擇全文檢索引所屬的全文檢索引目錄。



- 7 在【選取全文檢索引目錄】欄選擇資料庫已經建立的目錄，如果尚未建立，請勾選下方【建立新的目錄】選項來新增全文檢索引目錄，按【下一步】鈕選擇母體擴展排程。

- 8 關於母體擴展排程的說明與設定是在第 21-5 節，目前請直接按【下一步】鈕，可以看到目前工作的摘要描述。



- 9 如果有任何問題，請按【上一步】鈕更改，沒有問題，按【完成】鈕開始建立全文檢索索引，可以看到目前的建立進度。



- 10 等到完成建立後，按【關閉】鈕完成全文檢索索引的建立。

## 21-3-2 使用 T-SQL 指令建立全文檢索引

T-SQL 語言是使用 CREATE FULLTEXT INDEX 指令來建立全文檢索引，其基本語法如下所示：

```
CREATE FULLTEXT INDEX ON 資料表名稱  
[ (欄位名稱 [ TYPE COLUMN 型態的欄位名稱 ]  
[ LANGUAGE 語言名稱 ] [, ...n ] ) ]  
KEY INDEX 索引名稱  
[ ON 全文檢索引目錄名稱 ]  
[ WITH { CHANGE_TRACKING  
{MANUAL | AUTO | OFF [, NO POPULATION]} }  
]
```

上述語法是針對【資料表名稱】的資料表來建立全文檢索引，在之後括號是欲建立索引的欄位名稱清單，KEY INDEX 子句指定唯一索引，ON 子句指定所屬的全文檢索引目錄。

在欄位清單的欄位如果是 varbinary(max)或 image 類型，我們需要使用 TYPE COLUMN 子句告訴 SQL Server 如何直譯欄位儲存的資料，即指定儲存副檔名的欄位名稱。LANGUAGE 關鍵字可以指定欄位儲存資料使用的語言。

### Memo

SQL Server 全文檢索引搜尋支援 varbinary(max)或 image 類型的欄位，其儲存內容可以是 Word (doc)、Excel (xls)、PowerPoint (ppt)、HTML 文件 (htm) 或純文字檔 (txt)。

不過，在資料表需要新增字串類型的欄位，用來指明 varbinary(max)或 image 類型欄位儲存是哪一種文件，其欄位值就是括號內的副檔名。

最後 WITH CHANGE\_TRACKING 子句指定當資料表或檢視表變更時，追蹤變更的方式：MANUAL 是手動處理、AUTO 是自動和 OFF 是不處理，NO POPULATION 是當全文檢索引建立後，並不執行母體擴展，進一步說明請參閱〈第 21-5 節：母體擴展的使用〉。

**Memo**

因為 SQL Server 的 AdventureWorks 範例資料庫已經建立名為【AW20??FullTextCatalog】的全文檢索目錄，內含 Production.Document 資料表的全文檢索索引，請先參考第 21-4-3 節的說明刪除 Production.Document 資料表的全文檢索索引，再執行本節 SQL 指令碼檔案。

**SQL 指令碼檔：Ch21\_3\_2.sql**

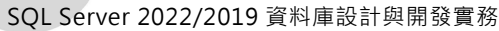
在【AdventureWorks】資料庫的【Product.Document】資料表儲存的是產品文件，我們準備在【產品文件\_全文檢索目錄】目錄建立 DocumentSummary 和 Document 欄位的全文檢索索引，如下所示：

```
USE AdventureWorks
GO
CREATE FULLTEXT INDEX ON Production.Document
(DocumentSummary, Document TYPE COLUMN FileExtension)
KEY INDEX PK_Document_DocumentNode
ON 產品文件_全文檢索目錄
WITH CHANGE_TRACKING AUTO
```

在上述 CREATE FULLTEXT INDEX 指令建立的全文檢索索引中，Document 欄位的類型是 varbinary(max)，TYPE COLUMN 子句指定的 FileExtension 欄位，這就是 Document 欄位內容的文件副檔名。

## 21-4 | 修改與刪除全文檢索目錄與索引

對於 SQL Server 資料庫已經建立的全文檢索目錄與索引，我們可以使用 Management Studio 或 T-SQL 指令來修改與刪除目錄與索引。



## 21-4-1 修改全文檢索目錄

SQL Server 可以使用 Management Studio 工具或 T-SQL 指令來修改全文檢索目錄。

## 在 Management Studio 修改全文檢索目錄

對於資料庫已經存在的全文檢索目錄，我們可以使用 **Management Studio** 修改其內容，也就是編輯屬性。請在「物件總管」視窗指定的全文檢索目錄上，執行【右】鍵快顯功能表的【屬性】指令，可以看到「全文檢索目錄屬性」對話方塊。

全文檢索目錄屬性 - 教務\_全文檢索目錄

選取頁面

一般

資料表/檢視表

母體擴展排程

連線

伺服器:  
DESKTOP-JOE

連接:  
DESKTOP-JOE\\vhuaya

檢視連接屬性

進度

就緒

指令碼 說明

屬性(P)

名稱	教務_全文檢索目錄
預設目錄	True
上次母體擴展日期	2023/1/6 上午 09:53
母體擴展狀態	閒置
目錄大小	0 MB
項目計數	11
唯一索引鍵計數	64
區分腔調字	True
擁有者	dbo

目錄動作

☒ 無(N)

☐ 最佳化目錄(O)

☐ 重建目錄(B)

☐ 重新擴展目錄(R)

確定

取消

在上述屬性對話方塊分為三個頁面，其簡單說明如下所示：

- 一般：除了顯示目錄目前的狀態外，我們也可以更改一些目錄屬性，包含是否為預設目錄、擁有者、是否區分腔調字和下方的目錄動作。



- 資料表/檢視表：顯示目錄擁有的全文檢索引，也就是建立索引的資料表或檢視表清單，在選取後，可以在下方顯示索引使用的唯一索引和建立索引的欄位清單等資訊。
- 母體擴展排程：檢視目錄的母體擴展排程，在第 21-5 節有進一步的說明。

## 使用 T-SQL 指令修改全文檢索目錄

T-SQL 語言是使用 ALTER FULLTEXT CATALOG 指令來修改全文檢索目錄，其基本語法如下所示：

```
ALTER FULLTEXT CATALOG 全文檢索目錄名稱
{ REBUILD [WITH ACCENT_SENSITIVITY = {ON | OFF}]
  | REORGANIZE
  | AS DEFAULT
}
```

上述語法可以修改名為【全文檢索目錄名稱】的全文檢索目錄，REBUILD 可以重建全文檢索目錄，Management Studio 就是在目錄上，執行【右】鍵快顯功能表的【重建】指令來重建全文檢索目錄。

WITH ACCENT\_SENSITIVITY 子句指定是否區分腔調字，REORGANIZE 可以讓 SQL Server 執行最佳化全文檢索目錄內部的索引，AS DEFAULT 子句將它設為預設目錄。



### SQL 指令碼檔：Ch21\_4\_1.sql

請在【教務系統】資料庫重建名為【教務\_全文檢索目錄】的全文檢索目錄，如下所示：

```
USE 教務系統
GO
ALTER FULLTEXT CATALOG 教務_全文檢索目錄
REBUILD WITH ACCENT_SENSITIVITY = OFF
```

上述指令可以重建全文檢索目錄，並且關閉區分腔調字。

## 21-4-2 修改與停用全文檢索引

SQL Server 可以使用 Management Studio 或 T-SQL 指令來修改與停用全文檢索引。在 Management Studio 提供相關指令和介面來修改與停用全文檢索引。

### 在 Management Studio 修改全文檢索引

請啟動 Management Studio，在「物件總管」視窗指定全文檢索引目錄上，執行【右】鍵快顯功能表的【屬性】指令，可以看到「全文檢索引目錄屬性」對話方塊。選【資料表/檢視表】頁面，可以修改全文檢索引。

我們也可以在指定資料表上，執行【右】鍵快顯功能表的「全文檢索引>屬性」命令來修改全文檢索引。

### 在 Management Studio 停用全文檢索引

在 Management Studio 停用全文檢索引，請在「物件總管」視窗指定的資料表上，執行【右】鍵快顯功能表的「全文檢索引>停用全文檢索引」命令，再按【關閉】鈕即可停用全文檢索引。

對於已經停用的全文檢索引，可以在資料表上執行【右】鍵快顯功能表的「全文檢索引>啟用全文檢索引」命令來啟用全文檢索引。

### 使用 T-SQL 指令修改與停用全文檢索引

T-SQL 語言是使用 ALTER FULLTEXT INDEX 指令來修改與停用全文檢索引，其基本語法如下所示：

```
ALTER FULLTEXT INDEX ON 資料表名稱
{
  ENABLE | DISABLE
  | SET CHANGE_TRACKING {MANUAL | AUTO | OFF}
  | ADD (欄位名稱 [ TYPE COLUMN 型態的欄位名稱]
        [ LANGUAGE 語言名稱] [, ...n ])
        [ WITH NO POPULATION]
  | DROP (欄位名稱 [, ...n])
        [ WITH NO POPULATION]
}
```

上述語法修改【資料表名稱】資料表的全文檢索索引，ENABLE 是啟用；DISABLE 是停用，SET CHANGE\_TRACKING 子句指定當資料表或檢視表變更時，追蹤變更的方式：MANUAL 是手動處理、AUTO 是自動和 OFF 是不處理。

ADD 子句可以增加建立索引的欄位；DROP 子句刪除索引的欄位，NO POPULATION 是當全文檢索索引建立後，並不執行母體擴展。



SQL 指令碼檔：Ch21\_4\_2.sql

請修改【員工】資料表的全文檢索索引，新增索引欄位【身份證字號】，如下所示：

```
USE 教務系統
GO
ALTER FULLTEXT INDEX ON 員工
ADD (身份證字號)
```

## 21-4-3 刪除全文檢索索引

對於資料表不再需要的全文檢索索引，我們可以使用 Management Studio 或 T-SQL 指令來刪除全文檢索索引。

### 在 Management Studio 刪除全文檢索索引

在 Management Studio 的「物件總管」視窗展開資料庫下的資料表，在資料表上執行【右】鍵快顯功能表的「全文檢索索引>刪除全文檢索索引」命令，再按【確定】鈕，即可刪除全文檢索索引。

### 使用 T-SQL 指令刪除全文檢索索引

T-SQL 語言是使用 DROP FULLTEXT INDEX 指令來刪除全文檢索索引，其基本語法如下所示：

```
DROP FULLTEXT INDEX ON 資料表名稱
```

上述語法可以刪除 ON 子句資料表的全文檢索引。



#### SQL 指令碼檔：Ch21\_4\_3.sql

請在【教務系統】資料庫刪除【員工】資料表的全文檢索引，如下所示：

```
USE 教務系統
GO
DROP FULLTEXT INDEX ON 員工
```

## 21-4-4 刪除全文檢索引目錄

對於資料庫不再需要的全文檢索引目錄，我們可以使用 Management Studio 或 T-SQL 指令來刪除全文檢索引目錄。

### 在 Management Studio 刪除全文檢索引目錄

在 Management Studio 的「物件總管」視窗展開資料庫下的【儲存體】項目，再展開【全文檢索引目錄】後，即可在目錄上執行【右】鍵快顯功能表的【刪除】指令，再按【確定】鈕刪除全文檢索引目錄（請注意！必須是沒有全文檢索引的目錄才能刪除）。

### 使用 T-SQL 指令刪除全文檢索引目錄

T-SQL 語言是使用 DROP FULLTEXT CATALOG 指令來刪除全文檢索引目錄，其基本語法如下所示：

```
DROP FULLTEXT CATALOG 全文檢索引目錄名稱
```

上述語法可以刪除名為【全文檢索引目錄名稱】的全文檢索引目錄。

**SQL 指令碼檔：Ch21\_4\_4.sql**

請刪除【教務系統】資料庫名為【教務\_全文檢索目錄】的全文檢索目錄，如下所示：

```
USE 教務系統
GO
DROP FULLTEXT CATALOG 教務_全文檢索目錄
```

## 21-5 | 母體擴展的使用

母體擴展（Population）是將資料加入全文檢索索引的程序，也就是維護全文檢索索引的程序，即在新增或更新記錄後，如何維護全文檢索索引資料的方式。

SQL Server 提供三種母體擴展，如下所示：

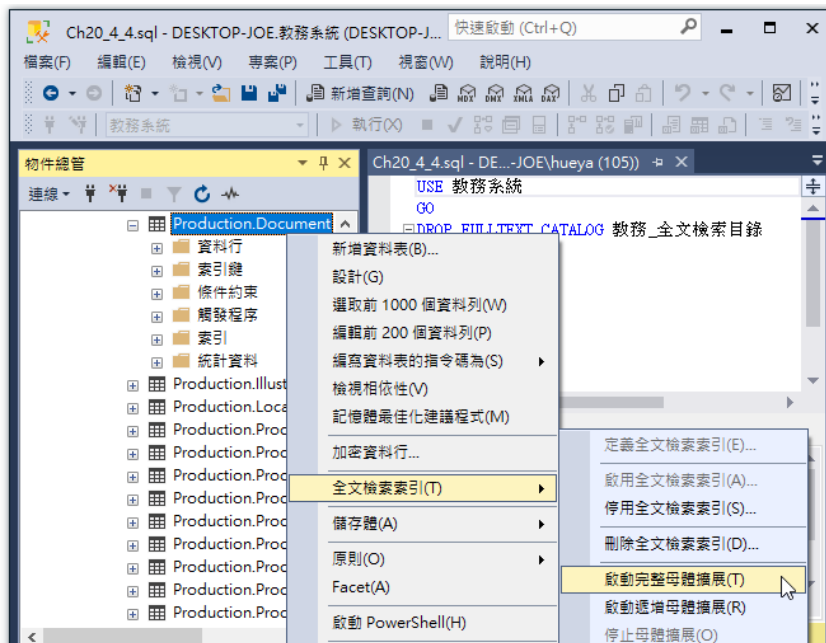
- 完整母體擴展（Full Population）：全部更新重建全文檢索索引。當建立過完整母體擴展後，就可以只執行遞增母體擴展或變更追蹤母體擴展來更新全文檢索索引。
- 遞增母體擴展（Incremental Population）：只更新上一次執行母體擴展後記錄有變動的部分，不過，當使用此類型的母體擴展，建立索引的資料表需要擁有 timestamp 類型的欄位，以判斷記錄是否更新。
- 變更追蹤母體擴展（Update Population）：使用全文檢索服務的自動追蹤變更功能來執行母體擴展，它會自動追蹤資料表的變更，並且將它記錄下來，我們只需執行指令套用變更記錄，即可執行全文檢索索引的變更追蹤母體擴展。

## 21-5-1 在 SQL Server 執行母體擴展

SQL Server 可以使用 Management Studio 或 T-SQL 指令來執行母體擴展。

### 在 Management Studio 執行母體擴展

在 Management Studio 啟動母體擴展，請在「物件總管」視窗的【Production.Document】資料表上，執行【右】鍵快顯功能表的【全文檢索引】命令，可以看到一個子功能表，如下圖所示：



在上述圖例執行【啟動完整母體擴展】命令可以執行完整母體擴展；【啟動遞增母體擴展】命令執行遞增母體擴展。

如果建立全文檢索引時是選【自動】變更追蹤，就只能執行【停用變更追蹤】命令停用變更追蹤後，才能再執行其他種類的母體擴展。執行【套用追蹤變更】命令是執行變更追蹤母體擴展。

## 使用 T-SQL 指令執行母體擴展

T-SQL 語言是使用 ALTER FULLTEXT INDEX 指令來執行母體擴展，其基本語法如下所示：

```
ALTER FULLTEXT INDEX ON 資料表名稱
{  START {FULL | INCREMENTAL | UPDATE} POPULATION
  |  STOP POPULATION
}
```

上述語法可以啟動或停止母體擴展資料表的全文檢索引，START 是啟動，FULL 是完整母體擴展、INCREMENTAL 是遞增母體擴展和 UPDATE 變更追蹤母體擴展，STOP 可以停用母體擴展。



### SQL 指令碼檔：Ch21\_5\_1.sql

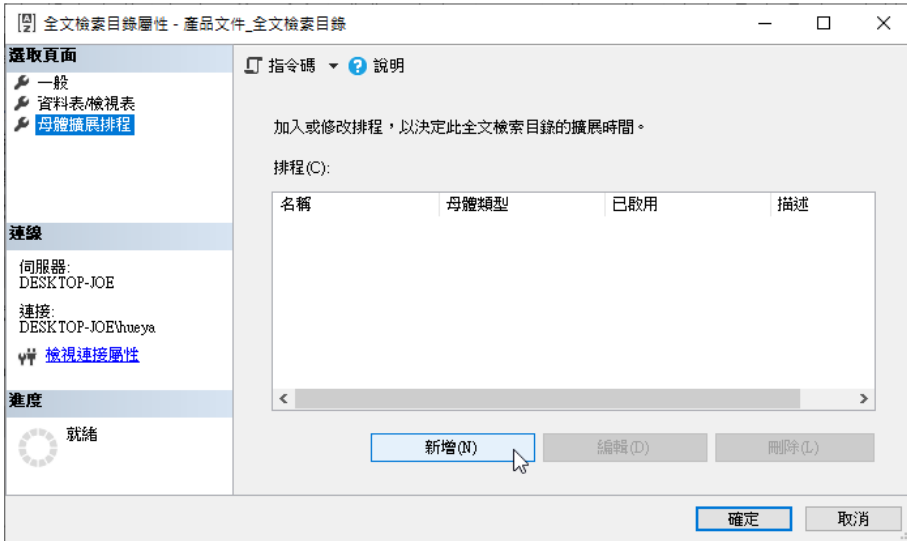
請在【AdventureWorks】資料庫的【Production.Document】資料表啟動完整母體擴展（如果已經啟用自動變更追蹤功能，需要先停用變更追蹤，才能成功執行完整母體擴展），如下所示：

```
USE AdventureWorks
GO
ALTER FULLTEXT INDEX ON Production.Document
START FULL POPULATION
```

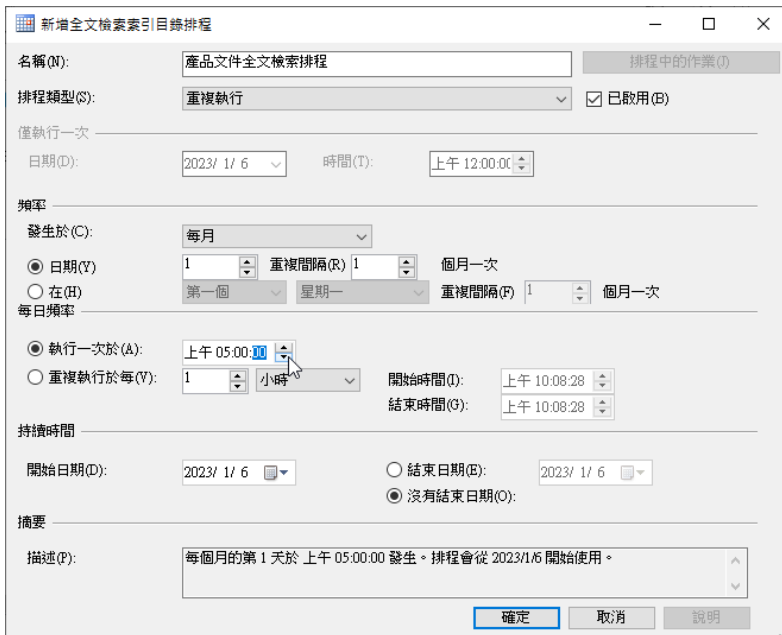
## 21-5-2 建立母體擴展排程

SQL Server 除了可以自行執行 T-SQL 指令來手動母體擴展全文檢索引外，也可以在全文檢索目錄建立母體擴展排程，使用排程在排程到達時，自動執行此目錄下所有全文檢索引的母體擴展，例如：建立排程在下班時間來母體擴展全文檢索引。

在 Management Studio 開啟【AdventureWorks】資料庫【產品文件\_全文檢索目錄】的「全文檢索目錄屬性」對話方塊，選【母體擴展排程】頁面，可以加入或修改排程，和決定何時執行母體擴展全文檢索引，如下圖所示：

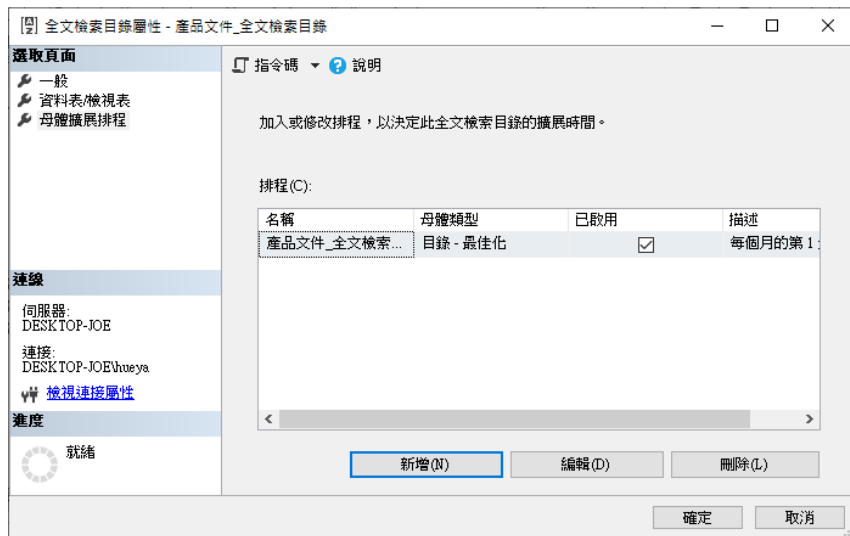


在【母體擴展排程】頁面按下方【新增】鈕，可以新增全文檢索目錄的排程，如下圖所示：





請輸入排程名稱、選擇排程種類和勾選【已啟用】選項表示啟用此排程，然後在下方指定排程頻率或時間後，按【確定】鈕新增母體擴展排程，如下圖所示：

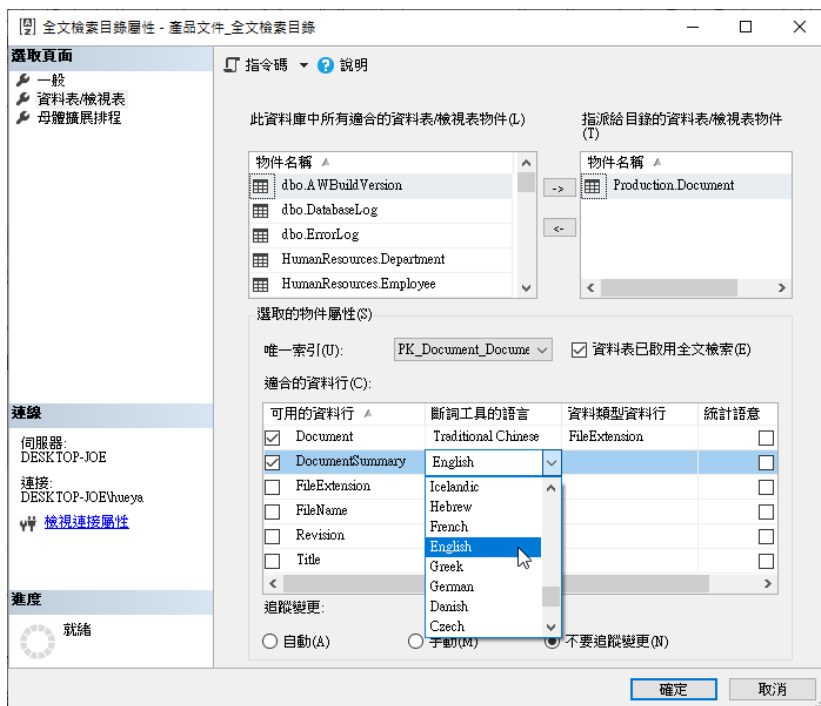


## 21-6 | 使用全文檢索搜尋

當建立全文檢索目錄與索引後，我們就可以使用 T-SQL 語言的 SELECT 指令來執行全文檢索搜尋，主要是在 HAVING 或 WHERE 子句建立全文檢索的搜尋條件，或在 FROM 子句取得合併查詢所需的記錄。

在本節範例的全文檢索搜尋是搜尋英文字，因為英文字搜尋包含衍生字，例如：單複數和不同時式等，中文字並沒有衍生字，因為之前建立全文檢索索引時沒有指定斷詞工具的語言，所以預設是繁體中文，我們需要將它改為 English，才能執行衍生字搜尋。

請在 Management Studio 開啟【產品文件\_全文檢索目錄】的屬性對話方塊，如下圖所示：



選【資料表/檢視表】頁面，在下方【適合的資料行】的【DocumentSummary】欄位，選【斷詞工具的語言】為【English】後，按【確定】鈕更改全文檢索引的屬性設定。

## 21-6-1 使用 FREETEXT 執行搜尋

FREETEXT 指令敘述可以在指定搜尋條件的句子中，分開執行句子中每一個單字的搜尋，包含完全符合和衍生字，其基本語法如下所示：

```
FREETEXT ( {欄位名稱 | 欄位名稱清單 | *}, '搜尋字串')
```

上述語法的第 1 個參數是欲搜尋的欄位名稱、欄位名稱清單，「\*」代表全文檢索引的所有欄位。第 2 個參數是使用單引號括起的搜尋字串，可以是完整句子或空白分隔的多個單字。



### SQL 指令碼檔：Ch21\_6\_1.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含'replace pedal'各別單字和衍生字的記錄，如下所示：

```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE FREETEXT (DocumentSummary, 'replace pedal')
```

上述 SELECT 指令可以執行全文檢索搜尋，WHERE 子句的條件是使用 FREETEXT 指令，執行結果找到包含 replace 和 pedal 單字的記錄，包含衍生字共有 3 筆記錄，如下圖所示：

	DocumentNode	DocumentSummary
1	0x7B40	Reflectors are vital safety components of your bicycle. Always ensure your front and back refle...
2	0x7BC0	Detailed instructions for replacing pedals with Adventure Works Cycles replacement pedals. Ins...
3	0x7C20	Worn or damaged seats can be easily replaced following these simple instructions. Instructions a...

## 21-6-2 使用 CONTAINS 執行單字搜尋

CONTAINS 指令敘述可以執行全文檢索的單字搜尋，不過，它並不包含衍生字搜尋，其基本語法如下所示：

```
CONTAINS ( {欄位名稱 | 欄位名稱清單 | *}, '包含的搜尋條件')
```

上述語法的第 1 個參數是欲搜尋的欄位名稱、欄位名稱清單，「\*」代表全文檢索引引的所有欄位。第 2 個參數是使用單引號括起的搜尋條件。



### SQL 指令碼檔：Ch21\_6\_2.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含單字 bicycle 的記錄，如下所示：



```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE CONTAINS (DocumentSummary, 'bicycle')
```

上述 SELECT 指令可以執行全文檢索搜尋，WHERE 子句的條件是使用 CONTAINS 指令，執行結果可以找到包含 bicycle 單字的記錄，共有 4 筆記錄，如下圖所示：

	DocumentNode	DocumentSummary
1	0x5B40	It is important that you maintain your bicycle and keep it in good repair. Detailed repair and serv...
2	0x6B40	Guidelines and recommendations for lubricating the required components of your Adventure W...
3	0x7B40	Reflectors are vital safety components of your bicycle. Always ensure your front and back reflec...
4	0x7BC0	Detailed instructions for replacing pedals with Adventure Works Cycles replacement pedals. Inst...

## 使用邏輯運算子

如果 CONTAINS 指令敘述搜尋的單字不只一個，我們可以使用 AND 或 OR 邏輯運算子來連接，請注意！其連接的每一個單字都需要使用雙引號括起。



### SQL 指令碼檔：Ch21\_6\_2a.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含單字 bicycle 和 pedals 的記錄，如下所示：

```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE CONTAINS (DocumentSummary, '"bicycle" AND "pedals"')
```

上述 CONTAINS 指令敘述的條件有兩個單字，所以使用 AND 連接，各單字分別使用雙引號括起，可以找到 1 筆記錄，如下圖所示：

	DocumentNode	DocumentSummary
1	0x7BC0	Detailed instructions for replacing pedals with Adventure Works Cycles replacement pedals. Inst...

## 使用萬用字元執行搜尋

在 CONTAINS 指令敘述的搜尋條件也可以在英文字首之後，使用萬用字元「\*」代表任何以此字首開始的單字。例如：pa\*代表所有以 pa 為字首的單字。



### SQL 指令碼檔：Ch21\_6\_2b.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含 pa 字首單字的記錄，如下所示：

```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE CONTAINS (DocumentSummary, '"pa*")
```

上述 CONTAINS 指令敘述是使用萬用字元執行搜尋，此時的單字也需要使用雙引號括起，可以找到 3 筆記錄，如下圖所示：

	DocumentNode	DocumentSummary
1	0x6B40	Guidelines and recommendations for lubricating the required components of your Adventure W...
2	0x7BC0	Detailed instructions for replacing pedals with Adventure Works Cycles replacement pedals. Inst...
3	0x7C20	Worn or damaged seats can be easily replaced following these simple instructions. Instructions a...

## 執行衍生字搜尋

對於英文單字來說，單字擁有現在式、過去式、未來式和單複數等多種衍生字，CONTAINS 指令敘述如果需要同時搜尋所有衍生字，我們需在搜尋條件使用 FORMSOF 來執行衍生字搜尋，如下所示：

```
CONTAINS (DocumentSummary,
          'FORMSOF (INFLECTIONAL, replace)')
```



### SQL 指令碼檔：Ch21\_6\_2c.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含 replace 衍生字的記錄，如下所示：



```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE CONTAINS (DocumentSummary,
                'FORMSOF (INFLECTIONAL, replace)')
```

上述 CONTAINS 指令敘述是執行衍生字搜尋，可以找到 3 筆記錄，如下圖：

	DocumentNode	DocumentSummary
1	0x7B40	Reflectors are vital safety components of your bicycle. Always ensure your front and back reflec...
2	0x7BC0	Detailed instructions for replacing pedals with Adventure Works Cycles replacement pedals. Inst...
3	0x7C20	Worn or damaged seats can be easily replaced following these simple instructions. Instructions a...

## 位置接近的字串搜尋

CONTAINS 指令敘述的搜尋條件可以搜尋位置接近的兩個單字，例如：在 oil 單字附近有 grease 單字，此時是使用 NEAR 運算子。



SQL 指令碼檔：Ch21\_6\_2d.sql

請在【AdventureWorks】資料庫的【Production.Document】資料表使用全文檢索搜尋來搜尋包含 guidelines 單字附近有 recommendations 單字的記錄，如下所示：

```
USE AdventureWorks
GO
SELECT DocumentNode, DocumentSummary
FROM Production.Document
WHERE CONTAINS (DocumentSummary,
                'guidelines NEAR recommendations')
```

上述 CONTAINS 指令敘述是使用 NEAR 運算子執行位置接近的字串搜尋，可以找到 1 筆記錄，如下圖所示：

	DocumentNode	DocumentSummary
1	0x6B40	Guidelines and recommendations for lubricating the required components of your Adventure W...

## 21-6-3 全文檢索的符合程度搜尋

符合程度搜尋（Ranked Searching）的搜尋結果會傳回系統自動產生的符合程度等級。在 SQL Server 可以使用 FREETEXTTABLE()和 CONTAINSTABLE()函數建立全文檢索的符合程度搜尋。

不同於本節前的全文檢索搜尋是在 WHERE 子句建立搜尋條件，FREETEXTTABLE()和 CONTAINSTABLE()函數是使用在 FROM 子句，可以取回滿足搜尋條件的記錄資料。

在執行本節的 T-SQL 範例前，請先執行 Ch21\_6\_3.sql 的 SQL 指令碼檔，可以建立 Production.ProductModel 資料表的全文檢索引，欄位包含 Name、CatalogDescription 和 Instructions，後 2 個欄位的資料類型是 xml。

### FREETEXTTABLE()函數的符合程度搜尋

FREETEXTTABLE()函數除了可以傳回滿足搜尋條件的記錄資料外，還會額外傳回 KEY 和 RANK 欄位，KEY 欄位是全文檢索引的唯一索引鍵值；RANK 欄位是系統自動產生的符合程度等級（值是 0~1000），其基本語法如下所示：

```
FREETEXTTABLE (資料表名稱, {欄位名稱 | 欄位名稱清單 | *},
               '搜尋字串', 語言)
```

上述語法和 FREETEXT 類似，只是多了第 1 個參數的資料表名稱，這就是全文檢索引所在的資料表，最後一個參數是斷詞工具的語言，以英文來說，就是 LANGUAGE 1033。



SQL 指令碼檔：Ch21\_6\_3a.sql

請在【AdventureWorks】資料庫【Production.ProductModel】資料表使用合併查詢來進行 FREETEXTTABLE()函數的符合程度搜尋，以便搜尋包含'washer weld polish'各別單字的記錄，查詢結果是以 RANK 欄位進行排序，如下所示：

```
USE AdventureWorks
GO
SELECT R_Table.RANK, L_Table.ProductModelID, L_Table.Name
FROM Production.ProductModel L_Table
INNER JOIN FREETEXTTABLE(Production.ProductModel,
    Instructions, 'washer weld polish', LANGUAGE 1033) R_Table
    ON L_Table.ProductModelID = R_Table.[KEY]
ORDER BY R_Table.RANK DESC
```

上述 INNER JOIN 合併查詢除了顯示 Production.ProductModel 資料表的欄位外，也顯示 RANK 欄位（由 FREETEXTTABLE() 函數傳回的欄位）。在 FROM 子句是使用 Production.ProductModel 資料表和 FREETEXTTABLE() 函數取回的記錄來建立合併查詢。

FREETEXTTABLE() 函數可以傳回內含 washer、weld 和 polish 單字的記錄資料，包含衍生字，合併條件是 ProductModelID 和 KEY 欄位，因為 KEY 本身也是關鍵字，所以使用方括號括起，最後使用 ORDER BY 子句以 RANK 欄位進行排序，可以找到 4 筆記錄，如下圖所示：

	RANK	ProductModelID	Name
1	157	7	HL Touring Frame
2	157	10	LL Touring Frame
3	149	47	LL Touring Handlebars
4	149	48	HL Touring Handlebars

## CONTAINSTABLE() 函數的符合程度搜尋

CONTAINSTABLE() 函數的功能和 FREETEXTTABLE() 函數相當，可以執行符合程度搜尋，其語法和 CONTAINS 類似，筆者就不重複說明。

不過，CONTAINSTABLE() 函數和 CONTAINS 指令敘述都可以替搜尋條件加上權值（Weight）來執行符合程度搜尋，其指定的權值會影響符合程度的等級，換句話說，CONTAINSTABLE() 函數符合程度的搜尋結果是可以透過權值來進行調整。



在 CONTAINSTABLE()函數和 CONTAINS 指令敘述替個別單字加上權值是使用 ISABOUT 指令，其語法如下所示：

```
ISABOUT {搜尋單字} [WEIGHT (權值)] [, ...n]
```

上述語法可以指定各別搜尋單字的權值，權值範圍是在 0.0 至 1.0 之間，如果單字不只一個，請使用「,」逗號分隔。



### SQL 指令碼檔：Ch21\_6\_3b.sql

請在【AdventureWorks】資料庫【Production.ProductModel】資料表使用合併查詢進行 CONTAINSTABLE()函數的符合程度搜尋，並且使用 ISABOUT 來替 washer、weld 和 polish 單字加上權值，如下所示：

```
USE AdventureWorks
GO
SELECT R_Table.RANK, L_Table.ProductModelID, L_Table.Name
FROM Production.ProductModel L_Table
INNER JOIN CONTAINSTABLE(Production.ProductModel,
    Instructions, 'ISABOUT (washer WEIGHT (1.0),
        weld WEIGHT (0.5), polish WEIGHT (0.1))') R_Table
ON L_Table.ProductModelID = R_Table.[KEY]
ORDER BY R_Table.RANK DESC
```

上述合併查詢的 FROM 子句是使用 CONTAINSTABLE()函數取回的記錄，並且使用 ISABOUT 指定各別單字的權值，共可以找到 4 筆記錄，如下圖所示：

	RANK	ProductModelID	Name
1	23	48	HL Touring Handlebars
2	20	47	LL Touring Handlebars
3	10	7	HL Touring Frame
4	10	10	LL Touring Frame