



HOW TO ENABLE THE THESAURUS SEARCH FOR IP DIRECTOR 6

06/11/2012



TABLE OF CONTENTS

Table of contents	1
<i>Introduction</i>	1
<i>User Setting</i>	2
<i>SQL Fulltext Search configuration</i>	3
IP Director or [DB] Server.....	3
tsenu.xml.....	3

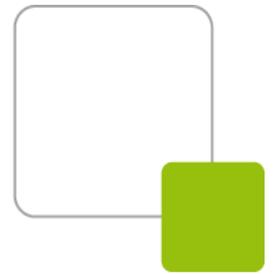
INTRODUCTION

IP Director 6 introduces the ability to perform searches against the SQL Fulltext Search with the use of a thesaurus.

This thesaurus is a list of words grouped together because they have the same meaning (synonyms).

It can also be used to have an auto-translated search if words are grouped with their translations in different languages.

This document describes how to enable the thesaurus search used by IP Director 6.



USER SETTING

A new user setting has been added to IP Director 6.

This setting defines

- if the user searches for any occurrence of a word (Classic fulltext)
- or if he searches for occurrences of the word and its synonyms (With thesaurus)

The “Classic fulltext” is the default setting

The value of this setting can be changed on a user basis in the settings window or globally in the User Manager.

The screenshot shows the 'IPDirector Settings' window with a sidebar on the left containing categories like 'Autoname', 'Clips', 'Playlist', 'Control Panel', and 'Ingest Scheduler'. The main area displays various settings. The 'Freetext search behaviour' setting is highlighted with a red box. It is located under the 'General' category and is currently set to 'With thesaurus'. Other visible settings include 'Display XT Structure', 'Datagrid settings', and 'Date Format Recognized in Text Searches'.

In addition to the setting value, the station hosting the IP Director database must be configured with a thesaurus file. This step is described in the next chapter.

In case of multilingual synonyms, search results are always displayed in their encoded language. The thesaurus search is not a translation service.



SQL FULLTEXT SEARCH CONFIGURATION

IP DIRECTOR OR [DB] SERVER

EVS IP Director hardware and EVS [DB] servers are configured with an instance of Microsoft SQL 2005 (Express edition for IP Directors and Standard edition for DB servers). Each of these platforms supports the SQL fulltext search and the SQL thesaurus.

The thesaurus is described in an XML file located in the SQL data folder.

These xml files are stored in **C:\Program Files\Microsoft SQL Server\MSSQL.1\MSSQL\FTData** on an IP Director or in **S:\MSSQL.1\MSSQL\FTData** on a DB server

TSENU.XML

The FTData folder contains a set of tsXXX.xml files. Each file describes the thesaurus linked to a specific language used in the fulltext indexation.

The language used in the IP Director fulltext indexes is "English (ENU)".

The thesaurus file to update is then **tsENU.xml**

By default, the tsenu.xml contains this structure:

```
<XML ID="Microsoft Search Thesaurus">
- <!-- Commented out

  <thesaurus xmlns="x-schema:tsSchema.xml">
<diacritics_sensitive>0</diacritics_sensitive>
  <expansion>
    <sub>Internet Explorer</sub>
    <sub>IE</sub>
    <sub>IE5</sub>
  </expansion>
  <replacement>
    <pat>NT5</pat>
    <pat>W2K</pat>
    <sub>Windows 2000</sub>
  </replacement>
  <expansion>
    <sub>run</sub>
    <sub>jog</sub>
  </expansion>
</thesaurus>
-->
</XML>
```

You will have to remove the XML comments to enable the thesaurus. The file will then look like that :

```
<XML ID="Microsoft Search Thesaurus">
= <thesaurus>
  <diacritics_sensitive>0</diacritics_sensitive>
= <expansion>
```





```
<sub>Internet Explorer</sub>
<sub>IE</sub>
<sub>IE5</sub>
</expansion>
= <replacement>
  <pat>NT5</pat>
  <pat>W2K</pat>
  <sub>Windows 2000</sub>
</replacement>
= <expansion>
  <sub>run</sub>
  <sub>jog</sub>
</expansion>
</thesaurus>
</XML>
```

Each <expansion> tag describes a group of words considered as synonyms.

You can add an <expansion> tag for each word you want to add to the Thesaurus and a <sub> tag for the different synonyms of this word

```
<expansion>
  <sub>tir</sub>
  <sub>shot</sub>
</expansion>
```

When the update of the thesaurus is completed, you will have to save the file in Unicode format (not UTF-8).

The SQL Server Fulltext Search service will have to be restarted to take the changes into account.

This operation can be done through:

- Computer management -> Services and Applications -> Services -> SQL Server Fulltext Search (MSSQLSERVER) -> Right-click -> restart
- Microsoft SQL Server Management Studio -> Management -> Fulltext Search -> Right-Click -> restart (only on DB servers)

The operation must be done on both servers in a mirroring configuration

For more information about the SQL Thesaurus file : Configuring Thesaurus Files :

[http://msdn.microsoft.com/en-US/library/ms345186\(v=sql.90\).aspx](http://msdn.microsoft.com/en-US/library/ms345186(v=sql.90).aspx)

A word can be used in only one expansion. The same word in two different expansions will make the thesaurus import fail.

The thesaurus file must be saved in Unicode

