



T O V E K

---

# Tovek Tools: Query Editor

---

version 4.31 and above

---

## User Guide

---







# Copyright

---

Information contained in this document may change without notice. The software described herein is subject to and protected by copyright law and international agreements related to intellectual property. It can be used solely in strict accordance with the License Agreement. No part of this publication can be reproduced, transmitted, transcribed by any electronic or mechanical means including photocopying and storing in retrieval systems, for any purpose, without the prior explicit permission of Tovek.

© Copyright Tovek. All rights reserved.  
<http://www.tovek.com>

## Registered trademarks

TOVEK and InfoRating are registered trademarks of Tovek.

Analyst's Notebook is a registered trademark of i2 Limited.

## Other trademarks recognition

Microsoft, Excel, Windows and Windows NT are registered trademarks of Microsoft Corporation.

All other companies and product names are trademarks or registered trademarks of their respective owners.

**TOVEK**

<http://www.tovek.com>

# Contents

---

<b>Overview .....</b>	<b>7</b>
Manual Version .....	7
Publications Related to Tovek Tools .....	7
Help .....	7
Introduction.....	8
<b>Editing a Topic .....</b>	<b>10</b>
<b>Running Queries .....</b>	<b>12</b>
<b>Working with Topics .....</b>	<b>13</b>
Using of the Unicode and Multilingual Query.....	13
Edit/Cut, Edit/Copy .....	14
Edit/Paste .....	14
Edit/Change Node .....	15
Find and Replace .....	15
Arranging Child Nodes - Sorting.....	15
Arranging Child Nodes - Moving Up or Down.....	16
Topic Tree, Topic Outline .....	16
Collapsing and Expanding.....	17
Zooming .....	17
Insert File.....	18
Save Subtree.....	18
Save IR-TXT.....	18
Export Context for InfoRating .....	18
Import Text .....	19
Creating Categories.....	20
Query Report.....	23
Fonts .....	23
<b>Working with File .....</b>	<b>24</b>
Create a New Topic.....	24
Open an existing Topic.....	24
Save a Topic .....	24
Close a Topic .....	24
Properties .....	24
Link to another Topic.....	25
<b>Basic Operators and Modifiers .....</b>	<b>26</b>
Stemming, Word, Case .....	26
The Not Modifier .....	27
Searching the Fields.....	27
<b>Index.....</b>	<b>28</b>



## Overview

---

Tovek Query Editor is a sophisticated tool for advanced users. It helps you create complex queries with a hierarchical tree structure. These are called topics. Using topics you can very precisely define what you are looking for and so you will be able to find all the relevant documents easily.

## Manual Version

This manual corresponds to the Tovek Tools version stated on the first side of the manual. For further program versions Release Notes will be published, which will be available on [www.tovek.com](http://www.tovek.com).

## Publications Related to Tovek Tools

Tovek Tools Search Pack

Tovek Tools Analyst Pack

Tovek Tools: Query Editor

Tovek Tools: InfoRating

Tovek Tools: Harvester

Tovek Tools: Fulltext Plug-in for Analyst's Notebook

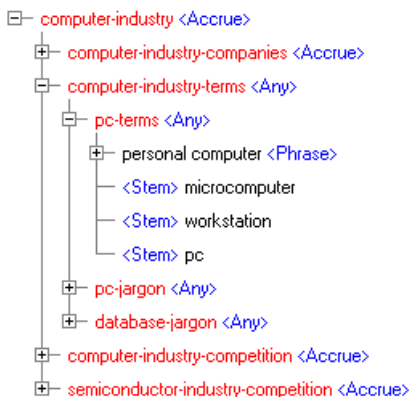
Verity Query Language

## Help

Tovek Tools contains help accessible from **Help** menu. After selecting the **Help Content F1** command you can see the overview help for all Tovek Tools programs and after selecting **About** command you will see information about the licence owner, expiration and Tovek Tools version. The version number is required for a support demand.

## Introduction

Example of a topic query might look like this:



Such topic representation is called a tree. The middle parts are called branches (or subtrees) while the end parts are called leaves. Leaves represent search items indicating a relationship of the actual topic to the subject.

Some words are more relevant to the given subject than other. In the scope of concept search, a presence of search words is considered as an evidence of presence of the given concept. These relations are expressed by weights ranging from 0.01 to 1.00 (1 to 100 per cent). Weights can be assigned to every branch. If no value is specified, a default weight for given operator is used.

During the search the whole topic structure is used to determine score of every document. Score indicates relevance of the particular document to given subject. All documents containing at least one query item will be retrieved. When you sort the result list by score you will quickly see which documents are the most relevant to your query.

Topics have these important features:

- ▶ can be reused
- ▶ can be combined
- ▶ can be shared

Topics can be used repeatedly. This means that once a topic was created and used for searching through a document set, it can be saved and used again so that the search criteria remain unchanged between two searches.

Combining topics means the ability to create new topics which will use the already existing ones. After some time, this feature will make your work easier, since most of the (sub)topics will already exist and you will be able to create new topics very quickly.

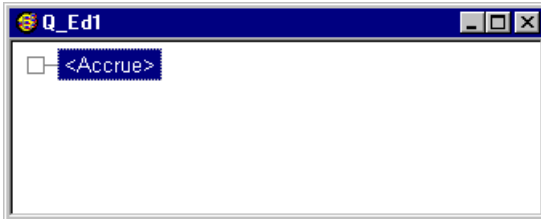
Topic sharing enables multiple users to use the same topics. This means that not only their authors, but also other people can make use of existing topics. Usually, topics are being constructed by experts, who then pass them to other users which are not supposed to have such deep knowledge of the given subject.

Thanks to this three features a large and valuable knowledge base can be created in any organisation actively using topics.

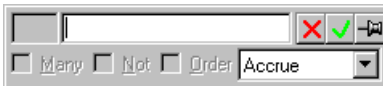
## Editing a Topic

---



Upon launching the Query Editor a window of a new topics will appear:



To define a topic item, double-click it or press **Enter**. This window will be shown:





Then proceed as follows:

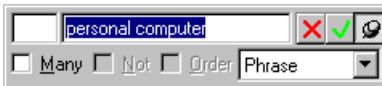
- ▶ Type in the topic name. Topic names should not contain any accented characters or spaces. Instead of spaces, words can be separated by hyphens, underscores, or by using capital letters.
- ▶ Select an operator from the pull-down menu.
- ▶ Finish the editing by pressing **Enter** key or clicking 
- ▶ To cancel the editing, press **Esc** key or click 

Adding new items:

- ▶ Decide whether the topic should be a sibling (an item on the same level) or a child (an item one level lower).
- ▶ To add a sibling, select the **Edit/Add Sibling** command, use the **Ctrl+D** keyboard shortcut, or right-click the item to which the sibling should be added and select **Add Sibling** command from the context menu.

- ▶ To add a child, select the **Edit/Add Child** command, use the **Ctrl+A** keyboard shortcut, or right-click the item to which the child should be added and select **Add Child** command from the context menu.
- ▶ Type in the appropriate search element - a word or a phrase - or a next branch label.
- ▶ Select an operator from the pull-down menu. For complete list of operators and modifiers, see the Verity Query Language description.
- ▶ Type weight into the leftmost field. If not specified, a default weight for given operator will be used.
- ▶ Depending on the operator, **Many**, **Not**, **Order** and **Case** modifiers might be selected.
- ▶ Finish the editing by pressing **Enter** key or clicking 
- ▶ To cancel the editing, press **Esc** key or click 

You can decide whether the editing window appears over the edited item or always in the same location by pressing the pin button (the one in the top right corner of the editing window).

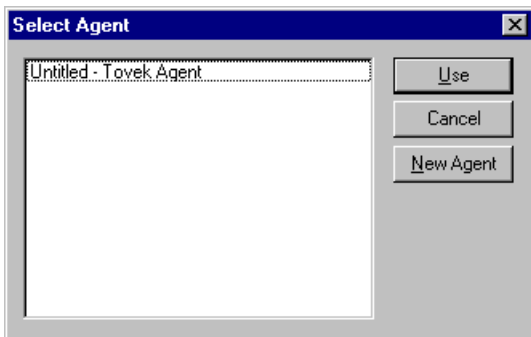


## Running Queries

---

When you are ready to test your topic, select **Tools/Run Query** command, click **Run Query** toolbar button, or press the **F5** key. This will launch a query containing only the selected branch of a tree. If you want to run query based on the complete topic tree, no matter which item is currently selected, hold down the **Ctrl** key while choosing the command.

When you run your query for the first time, you will need to select in which search window (an agent) it will be run:



This window contains a list of all currently open Tovek Agents. Highlight one entry and choose the **Use** button or double-click an entry in the list.

After you select an agent for your topic make sure that you check all the options such as where to search, required score etc. Next time you run the query, no dialog window will be shown and the search will be started immediately.

In selection window you can also create a new Tovek Tools Agent by choosing the **New Agent** button.

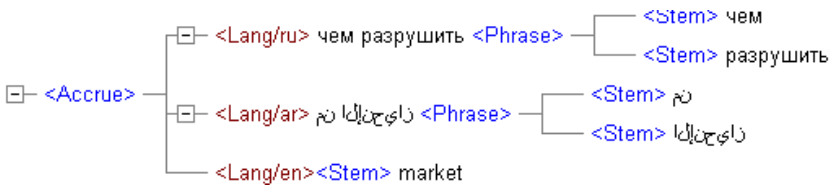
## Working with Topics

### Using of the Unicode and Multilingual Query

Several languages can be used at the same time in a topic in Unicode mode. Switching to work in Unicode for a particular topic can be done by **Tools / Switch to Utf-8 Mode**. This topic cannot then be switched back afterwards. After switching to Utf-8, the topic items can be entered in different languages.

You can enter a **<LANG/ID>** modifier for searching in collections created with multilingual locale (Unilocale) in a language other than the preferred one using the context menu of the created topic item. After right-clicking an already entered item with a **STEM** or **PHRASE** operator, choose from **<Lang/ID>** menu the modifier for the particular language (e.g. **<Lang/ru>** for an item in Russian).

An example of a topic items in various languages:



The **<LANG/ID>** modifier can be removed from a selected topic item using the **Remove Lang/ID** command from the **<Lang/ID>** menu from the item context menu.

The language list for **<LANG/ID>** modifier, as they appear in the **<Lang/ID>** menu in the item context menu, can be changed using the **Edit list** command from the same menu.

If a topic has Utf-8 mode turned on, this information is stated in the status bar.

A topic switched to Utf-8 mode is also saved in Utf-8 coding.

Switching to Utf-8 can be set as default for new topics using the **Tools / Utf-8 Mode is Default** command.

## Edit/Cut, Edit/Copy

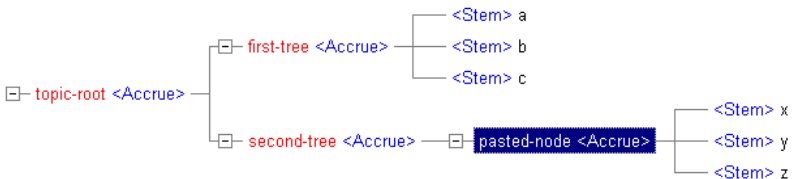
You can use all common Windows commands and techniques, such as Edit/Cut, Copy, Paste, tree expanding etc., but some of these commands are slightly enhanced for working with query trees. Commands and their differences are outlined below.

These commands work as usual, but they provide one more feature. When you choose **Edit/Cut** or **Edit/Copy**, the selected part of a tree is placed on the clipboard. It is stored there as an object of Query Editor type so it can be pasted back into the Query Editor or any OLE-compliant application. On the clipboard it is also stored as a text containing the query in Verity Query Language. This query can be pasted into any application that uses Verity's search technology, such as Tovek Agent, or into a web browser that is connected to Verity-enabled search server.

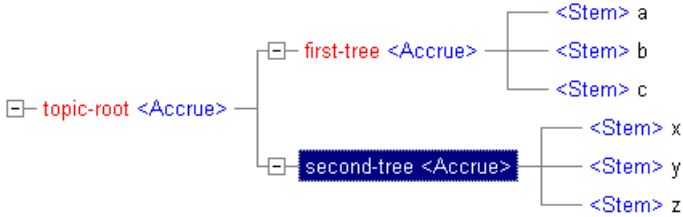
## Edit/Paste

There are two commands related to **Edit/Paste**. It is possible to paste plain text into the Query Editor, in which case what is actually happening is the **Tools/Import Text** command (see below) using the clipboard instead of a text file. It supports the same syntax and shortcuts and helps you to create bare topics more quickly.

The other command is **Ctrl+Del** shortcut and it might be helpful after the **Edit/Paste**. It deletes only the currently selected node and connects its child nodes to the parent node (if there is one). For example, there is a pasted subtree:



When **Ctrl+Del** is pressed, only the selected node will be deleted:



## Edit/Change Node

Opens selected node for editing. It can also be invoked by pressing **Enter** or **F2** key or by double-clicking a node.

- ▶ To finish editing, press **Enter** key or click
- ▶ To cancel editing, press **Esc** key or click

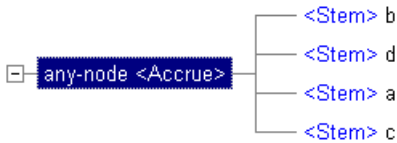
## Find and Replace

Find and replace commands from the Edit menu work slightly differently than in other applications. When searching within the topic tree the entered text is matched against all query parts of every node including the operator, ie you can search for nodes containing the operator accrue, for example. But during the replace operation the search is performed only in the score, evidence, condition expression, and possibly the proximity part of the Near/n operator. The other difference is the **Replace All** button which simply replaces found text of the current node and then searches for another node, but it does not perform automatic search and replace for the whole query tree.

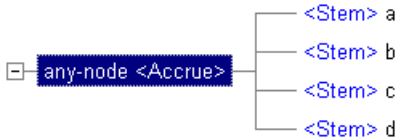
## Arranging Child Nodes - Sorting

Children of a selected node can be sorted by choosing the **Edit/Sort Child Nodes** command or by the respective shortcut key - see below.

Before sort:



After sort:



The default is alphabetical sort in ascending order. If you hold down **Shift** key when selecting **Edit/Sort Child Nodes** command, sort order will be reversed into alphabetical descending. If you hold down the **Ctrl** key, sorting by weights will be performed. The **Ctrl+Shift** key combination is also available and it reverses the weight order. When you highlight the **Edit/Sort Child Nodes** menu command, the status bar will display a help message with the possible options. This command can be used for all branch items except branches with the Phrase, Paragraph, or similar operators.

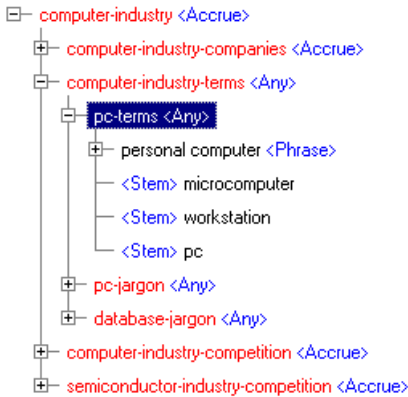
## Arranging Child Nodes - Moving Up or Down

A highlighted item can be moved up (within the same level) by selecting the **Edit/Move Up** command or by the **Ctrl+U** keyboard shortcut. If an item is already on the top, it will be moved to the last position.

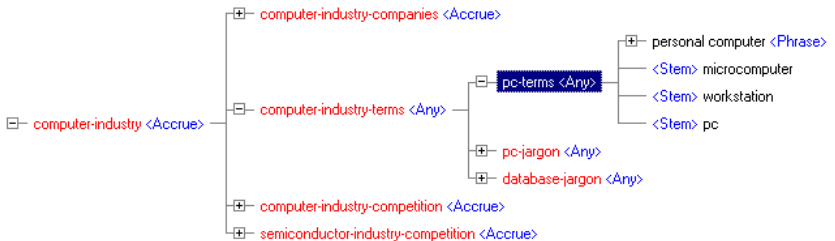
If you hold down the **Shift** key when selecting **Edit/Move Up** command (or press **Ctrl+Shift+U**), the direction will be reversed.

## Topic Tree, Topic Outline

You might choose between two view modes for a topic tree. Example of a topic displayed in an outline mode, after the **View/Topic Outline** command:



The same topic after the **View/Topic Tree** command:



## Collapsing and Expanding

Collapsing and expanding works as in other applications, such as Windows Explorer, and is controlled by common commands in the **View** menu. However, there is one extra command, **View/Expand One Level**. Its shortcut is **Ctrl+[+]**.

This command expands the selected node, but *only* to the next level. If the lower level had some subtrees expanded, they will be closed. It is different from the **View/Expand** command which opens all branches based on their previous state.

## Zooming

The whole topic tree can be zoomed in or out by the zoom commands on the bottom of the **View** menu, using the three toolbar buttons, or by

the shortcut keys. Zooming can also be specified even more precisely by the **View/Zoom/Zoom Custom** command, by allowing you to specify custom zoom settings with independent X- and Y-axis scales.

Another way to change a topic size is by using the zoom slider, which can be found in the status bar on the right. You can change the size by clicking the plus (+) or minus (-) button next to the slider or by moving the slider handle. You can return to the standard size (100%) by double-clicking the number next to the slider.



## Insert File

Into the selected node you can insert a subtree from a file by the **Tools/Insert File** command. Both OTL and plain text files can be inserted.

## Save Subtree

Selected branch of a topic tree can be saved into a separate file by the **Tools/Save Subtree** command.

## Save IR-TXT

The **File/Save IR-TXT** command saves subtrees of the selected node as a set of queries into a plain text file. This file can then be used for the html export with the InfoRating analysis in the Tovek Agent, or the queries can be loaded into the InfoRating application itself. This command cannot be used with Utf-8 mode switched on, see Using Unicode and Multilanguage Query.

## Export Context for InfoRating

Exporting the InfoRating context is similar to saving IR-TXT file but this command sends queries directly to the InfoRating application, starting it if necessary. This command cannot be used with Utf-8 mode switched on, see Using Unicode and Multilanguage Query.

## Import Text

You can import a plain text file into a topic tree using the **Tools/Import Text** command.

Basic rule is that every word or phrase should be on a separate line. However, you can take advantage of these possibilities:

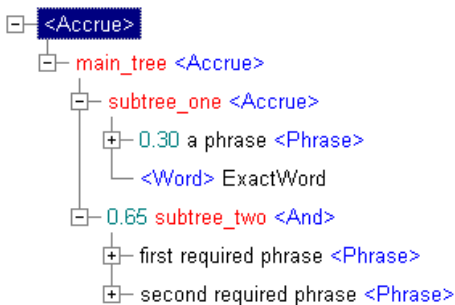
- ▶ Use spaces, tab characters, or asterisks to indent topic levels.
- ▶ Specify weights in front of the text.
- ▶ Specify operators at the end of line.

For example, this file:

```

sample.txt - Notepad
File Edit Search Help
main tree
    subtree one
        0.30 a phrase
            "ExactWord"
* 0.65 subtree two <and>
** first required phrase
** second required phrase
  
```

will be imported this way:

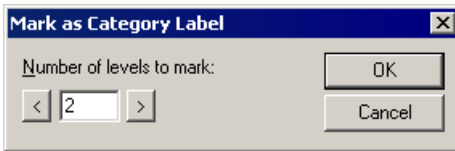


The same rules and options can be used when pasting text from clipboard into a topic tree.

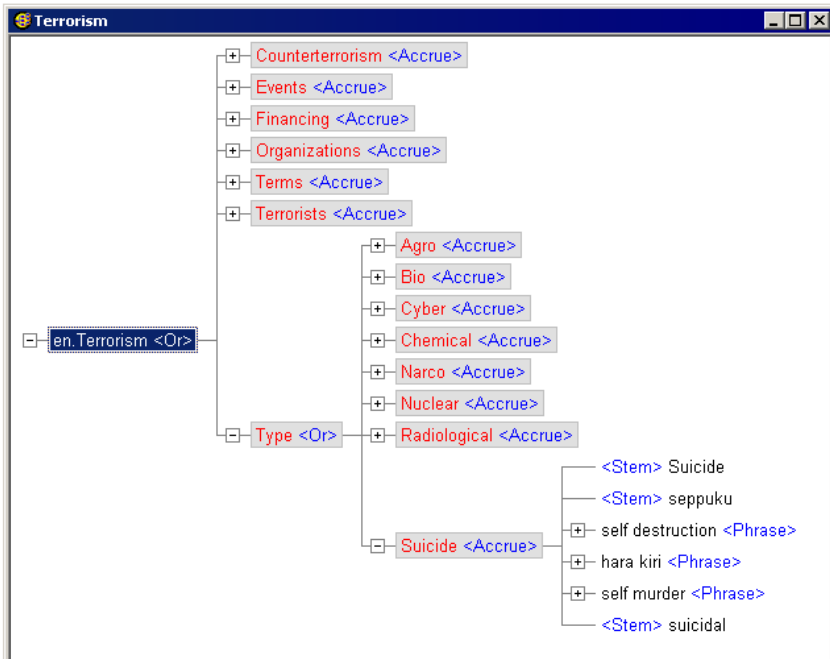
## Creating Categories

You can create categories for Tovek Server categorisation from a tree or its part in the Query Editor. First you need to mark the required branches using the **Edit / Mark as Category Label** command or the **F9** key. The currently selected branch and its ancestor names on higher levels of the hierarchical structure will be marked as a category. These names will have a grey background.

Several tree levels can be marked as categories at once using the **Edit / Mark Multiple Levels** command. In the Mark as Category Label dialog select number of levels which are to be marked as categories, together with currently selected branch:



If you select the top level branch and choose two levels to mark, for example, the tree with marked categories can look as follows:



You can unmark branches as categories using the **Edit / Unmark Category Label** command or **Shift+F9** shortcut.

**Notes:** Nodes marked as category names can contain spaces and diacriticals.

It is recommended to work in utf-8 mode (see Using of the Unicode and Multilingual Query), so no further conversion to utf-8 is needed for the OTL file (the topic) or the XML file.

From the branches marked as categories an XML file can be created, which can be loaded to the server and used for categorisation (see Tovek Server: Categorisation manual), using the **Tools / Export TS Categories XML** command.

After entering the file name and path, the category names will be created from the marked branch names, while the branch content will form the queries for the categories. The tree structure from the Query Editor will remain in the category structure.

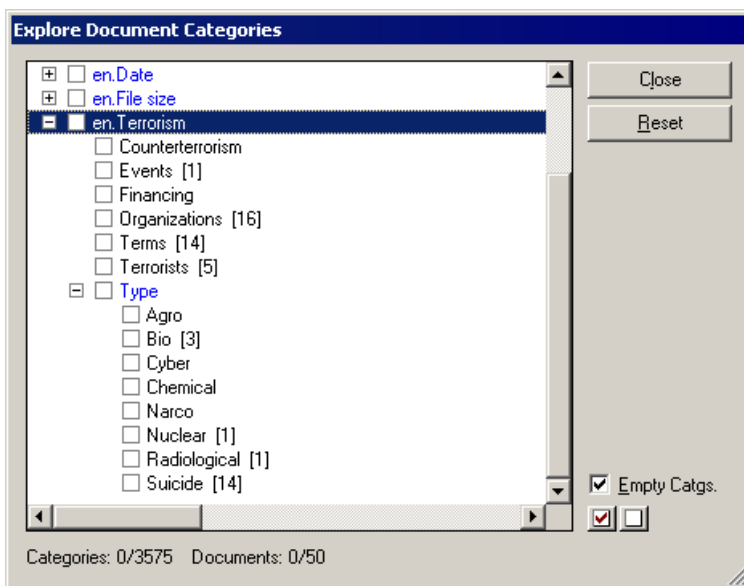
If a node marked as a category name has some unmarked nodes on a lower level, only marked branches will be exported to the XML file and the unmarked branches will be ignored.

**Notes:** When creating the XML file, the category set name from the top level tree node which is currently selected when starting the export will be used.

If in the Category property (see Properties) the topic has the following: `ID:NN`, where NN is a number between 1 and 255, the number will be used as category set ID and will also be recorded into the XML file. If the Category property is not set this way, no ID will be recorded.

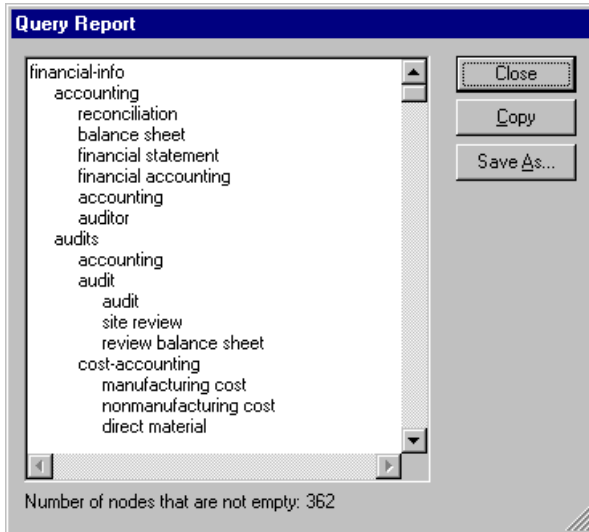
Information about categorisation on the Tovek Server can be found in the Tovek Server: Categorisation manual.

While exploring categories (see Tovek Tools, Categorisation manual), the tree from the above example can look as follows, for example:



## Query Report

Summary showing tree elements of a selected branch can be shown by the **Tools/Query Report** command. Text can be copied using the **Copy** button or saved into a text file by the **Save As** button.



## Fonts

**Tools/Change Font** command changes font for one topic tree. Font of the current topic can be applied to all new topics as a default by the **Tools/Set Font As Default** command.

## Working with File

---

### Create a New Topic



Pull up a new topic window by selecting **New** command from the **File** menu, clicking the **New** button on the toolbar or pressing **Ctrl+ N** shortcut.

### Open an Existing Topic



To open an existing topic select **Open** command from the **File** menu, click **Open** button on the toolbar, press **Ctrl+O** shortcut or choose one from the list of recent topics in the **File** menu.

### Save a Topic



To save a topic select **Save** (or **Save As**) command from the **File** menu, click the **Save** button on the toolbar or press **Ctrl+S** shortcut.

### Close a Topic



You can close an active topic by choosing **Close** command from the **File** menu or clicking the cross button in the top right hand corner of the topic window .

### Properties

**File/Properties** command brings up the dialog box for entering and editing additional information about the topic tree, such as who was the author and what is a general description of the topic. This information can then be, for example, retrieved from OTL files and used to categorize a set of topics.

## Link to another Topic

If a <Topic> operator is used in an evidence topic (end item), you can associate with it an OTL file with a topic definition using **Open Topic OTL** command from **Tools** menu or **Ctrl+T** shortcut. A dialog for finding OTL file opens, where the topic is stored. This topic opens in a new Query Editor window and you can view and update it.

When you use the **Open Topic OTL** command with the same name (in any OTL file) next time, the associated OTL file opens with the topic definition.

If you need to change the link to OTL file for an evidence topic, hold down the **Shift** key while running the **Open Topic OTL** command. The dialog for finding and linking another OTL file opens.

The topics with <Topic> operator can be run only after the administrator enters them into the system (e.g. indexing in the Tovek Server).

## Basic Operators and Modifiers

---

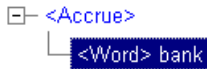
Only basic operators and modifiers are listed here. Extensive description of operators is available from [tovek.com](http://tovek.com) web pages in a separate guide.

Operator	Description
<b>Accrue</b> ,	<p>Retrieves documents containing at least one of the search elements. It can also be transcribed as a comma (,).</p> <p><b>Accrue</b> operator ranks documents based on presence of search elements by following rule "the more the better". In other words, the more search elements the document contains, the higher is its score.</p>
<b>And</b>	Retrieves documents containing all specified search elements.
<b>Or</b>	Retrieves documents containing at least one search element. Contrary to the <b>Accrue</b> operator, document score is somehow less accurate, which manifests itself when searching through a large number of documents. Also some search elements cannot be combined with the <b>Accrue</b> operator so <b>Or</b> operator has to be used instead.

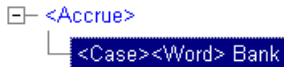
## Stemming, Word, Case

If you specify the search word *bank*, the query will find not only documents containing the very word bank, but also *banking*, *banks*, etc. This is possible due to so-called stemming.

If you have to search for an exact form of a word, use the **Word** operator:

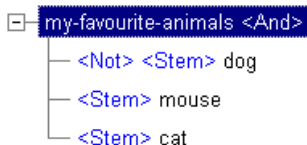


By default, searches are not case sensitive. If you want to search for a word in exact case, use the **Case** modifier:



## The Not Modifier

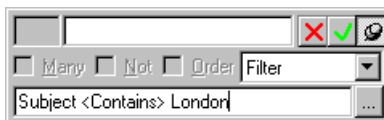
The **Not** modifier is used to exclude documents containing a certain word or a phrase. For example, to select documents containing the words *cat* and *mouse* but not those containing the word *dog*:



The **Not** modifier can be used only with the **And** and **Or** operators.

## Searching the Fields

For searching in the fields you can use **Filter** of **Field** operator, which also slightly modifies the query node editing window:



# Index

---

## A

Accrue, 26  
Add Child, 11  
Add Sibling, 10  
And, 26  
Arranging Child Nodes, 16

## B

Branch, 8

## C

Case, 11, 26  
Categories, 20  
Close, 24  
Collapse, 17  
Copy, 14  
Cut, 14

## E

Editing a Topic, 10  
Expand, 17  
Expand One Level., 17  
Export Context for InfoRating,  
18  
Export TS Categories XML, 21

## F

Field, 27  
Filter, 27  
Find, 15  
Font, 23

## H

Help, 7

## Ch

Change Font, 23  
Change Node, 15  
Child, 11  
Child Node, 15

## I

Import Text, 14, 19  
InfoRating, 18  
Insert File, 18  
IR-TXT, 18

## L

LANG, 13

## M

Manual Version, 7  
Many, 11  
Mark as Category Label, 20  
Mark Multiple Levels, 20  
Move Up, 16  
Multilingual Query, 13

## N

New, 24  
New Agent, 12  
Node, 15  
Not, 11, 27

## O

Open, 24  
Open Topic OTL, 25  
Or, 26  
Order, 11  
Other Publications, 7  
Outline, 16

**P**

Paste, 14  
Properties, 24  
Publications, 7

**Q**

Query, 12  
Query Report, 23

**R**

Remove Lang/ID, 13  
Replace, 15  
Report, 23  
Run Query, 12

**S**

Save, 24  
Save IR-TXT, 18  
Save Subtree, 18  
Set Font As Default, 23  
Sibling, 10  
Sort Child Nodes, 15  
Sorting, 15  
Subtree, 8, 18  
Switch to Utf-8 Mode, 13

**T**

Text, 19  
Topic, 8  
Topic Outline, 16  
Topic Tree, 16  
Tree, 8, 16

**U**

Unicode, 13, 21  
Unilocale, 13  
Unmark Category Label, 21  
Utf-8, 13, 18  
Utf-8 Mode is Default, 13

**V**

Version, 7

**W**

Word, 26

**Z**

Zoom, 17