

TextSpy 1.5

User Guide

About TextSpy	2
TextSpy Indexer.....	2
TextSpy Query	2
Using TextSpy	
Creation of Indexes – TextSpy Indexer	3
Search in Indexes – TextSpy Query	5
Auto attaching of TextSpy indexes.....	9
Licensing options for TextSpy	10
Known Issues	10
Technical Support	10
Additional Information.....	11

About TextSpy

Amos Software TextSpy is a solution for fulltext search in Adobe Acrobat PDF files. The software can be used for indexing sets of PDF files to facilitate rapid search for words and phrases. Thanks to the fact that TextSpy is based on a universal unicode representation of characters, this solution is not only limited to texts written with the Western (Roman) or Eastern European character sets.

Amos Software TextSpy was developed by Techsoft Servis company exclusively for Amos Software. The program consists of a set of plug-ins for Adobe Acrobat which extends the possibilities for working with fulltext by adding languages that use different character sets than the original Western languages (Roman, Latin 1). The main difference from the fulltext tool supplied with Adobe Acrobat is a significantly extended language support that besides all Western European languages includes Central and Eastern European languages, such as Czech, Slovak, Polish, Slovenian, Croatian, Hungarian, Romanian, Latvian, Lithuanian, Estonian, Turkish, Greek, Bulgarian, Ukrainian, Russian and Serbian.

The TextSpy unicode fulltext solution is based on two modules – Indexer and Query.

TextSpy Indexer

TextSpy Indexer extracts text data from Adobe PDF files and creates indexes with information about the occurrence and position of individual words. Indexes are subsequently used with the TextSpy Query module. TextSpy Indexer is installed and works in Adobe Acrobat 5.0 CE only.

TextSpy Query

This module installed in Adobe Acrobat 5.0 CE and Acrobat Reader 5.0 CE is used for entering queries about the incidence of a given word or words. Using indexes, TextSpy Query finds all the words that match the entered criteria. The results are shown as a list in a dialog box which also displaying information about found files or opening selected pages in Adobe Acrobat or Acrobat Reader. When a page is opened, the found words are highlighted in a different color.

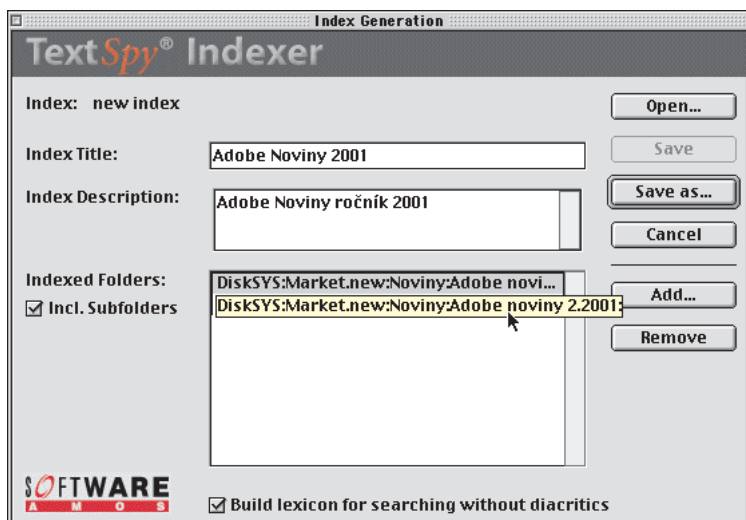
The following pages contain instructions for use of TextSpy Indexer and TextSpy Query.

Using TextSpy

Creation of Indexes – TextSpy Indexer

Indexes are created with TextSpy Indexer. TextSpy Indexer can be opened in Adobe Acrobat 5.0 or later by selecting *Edit > TextSpy > New Index* or by using the keyboard shortcut *Option + M* on Macintosh computers.

The Indexer dialog box opens.



The two upper fields serve for entering the title and a description of the index which is being created. The bottom field serves for adding or removing folders containing PDF files which are to be indexed.

To enable entering queries without diacritical marks check the “Build lexicon for searching without diacritics” box, when creating index files. This can be useful when the index will be used on systems which does not enable to type texts with special characters with accents. In this case you can enter “Skoda” instead of “Škoda” when you want to find the word “Škoda”.

You can create new indexes or open, regenerate and modify existing indexes. Pressing Save and Save As starts the creation of the actual index. The command Save starts the updating of an existing index; Save As is used for creating and saving a new index or updating an existing index and saving it under a different name or in a different location.

In the left part of the dialog box you can check whether you want the program to index only files located in the selected folder or also files in all subfolders.

When saving an indexed file you have to pay attention whether the index file is located on the same disk (volume) as the indexed files. In this regard, there are two different ways of using indexes:

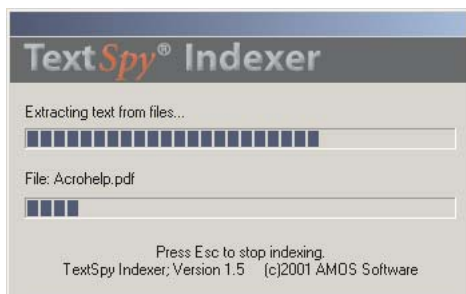
1. When TextSpy and its indexes are used for personal needs or within a workgroup in a fixed computer network, indexed files can be saved in any place (for example in a central location), regardless of where the PDF files to be searched are located. Just make sure you avoid changing the location of indexed PDF files in the future.
2. Concerning the creation of indexes which are to be distributed on various data carriers or CD-ROM, the index and the indexed PDF files have to be located within a single volume and when distributed or copied their relative position must be maintained. It means that entire folders containing both the index and the indexed PDF files have to be moved together.

The rules are basically the same as with the creation of HTML documents that contain links.

Especially in the case of CD-ROM publishing the search function should be tested before large quantities of CDs are produced.

Important: Public distribution of indexes created using this version of TextSpy Indexer is not allowed. You should use TextSpy Indexer CD-ROM Publishing Kit instead. See “Licensing options for TextSpy”.

The creation process is shown in the following box.



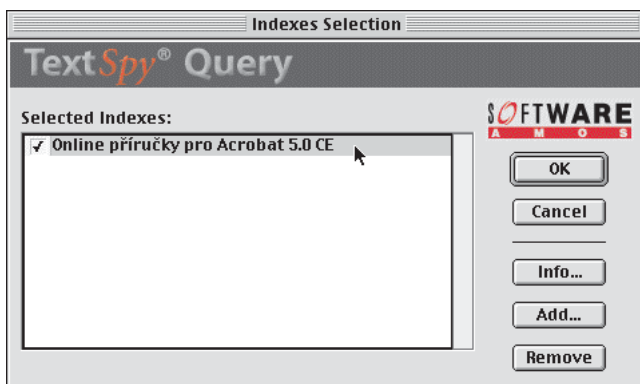
The upper progress bar shows the progress of creation of the entire index, the lower progress bar shows the processing of individual PDF files. Completed indexes are ready for use by TextSpy Query.

Search in Indexes –TextSpy Query



Searching is executed using the TextSpy Query module in Adobe Acrobat 5.0 CE or Acrobat Reader 5.0 CE or later.

First you have to select indexes which are to be searched in. To choose indexes select *Edit > TextSpy > Select Indexes* or use the keyboard shortcut *Option + X* on Macintosh computers.

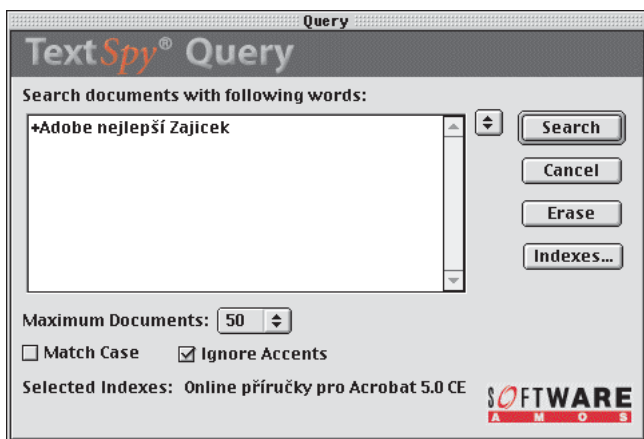
The following dialog box opens.



This dialog box can be used to add, remove, activate or deactivate index files which are to be searched in. After selection of indexes TextSpy Query is ready to start searching. If an index is not available (for example if it was attached from a CD-ROM) the line with its name is dimmed and you cannot activate this index. This index will not be used for searching.

Search is activated in Adobe Acrobat 5.0 or later and in Acrobat Reader 5.0 or later by selecting *Edit > TextSpy > Search* or the keyboard shortcut *Ctrl + E* on Windows-based computers or *Edit > TextSpy > Query* or the keyboard shortcut *Option + F* on Macintosh computers. In addition, you can also click on the icon  (Windows), or  (Macintosh) on the Adobe Acrobat toolbar .

The following dialog box opens.

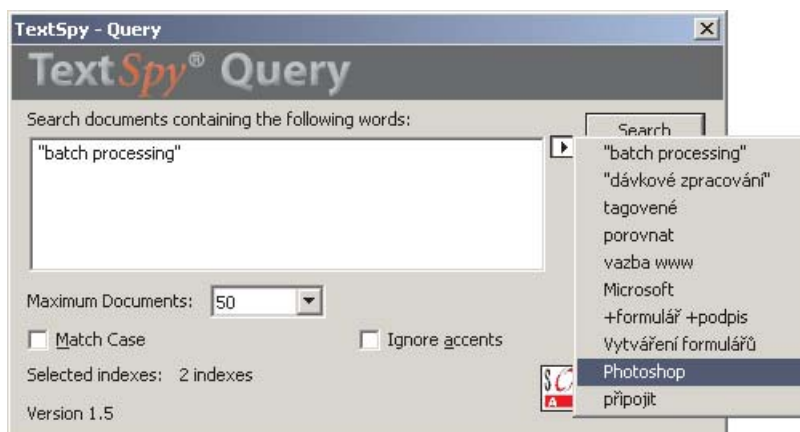


Searching is very intuitive. The syntax of queries is not based directly on Boolean operators, but uses a structure similar to common Internet search engines. Enter the searched words in the Query dialog box (you can enter more than one word) and operators as follows:

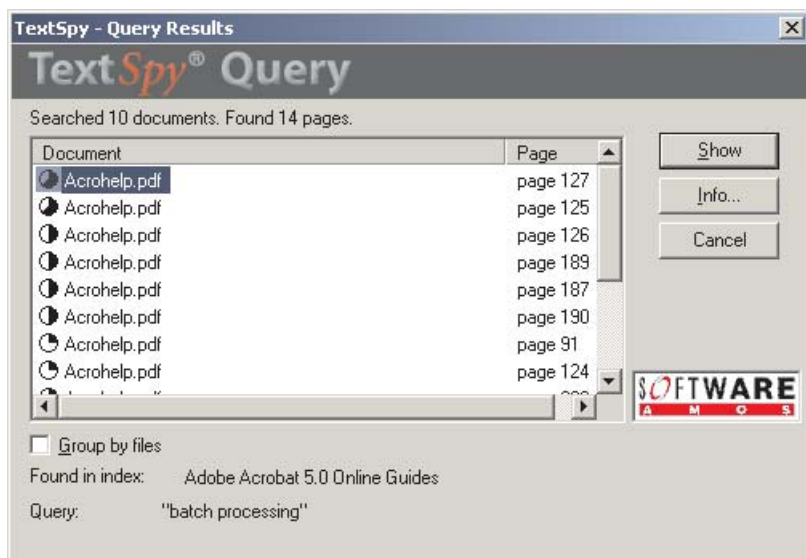
1. Inclusion or exclusion of words – if you want a word to be included in the results of a query place the (+) sign in front of the word in the Query dialog box. If you want a word excluded from the results, type the (-) sign in front of the word.
2. Using wildcards – by including an asterisk (*) or a question mark (?) in a searched word you can search for different forms of a word. An asterisk substitutes any number of characters; a question mark substitutes any single character.
3. Case sensitive – if you want to match the case, check the Match Case box.
4. Ignore accents – check this option if you want to search without respect to diacritical marks. This option has only effect for indexes, which have been generated with selected option Create index for searching without diacritics.
5. Search for exact phrases containing several words – place quotation marks before and after the phrase. The phrase will be searched for as a series of words regardless of the symbols which separate individual words.

A brief description of the query syntax is available by selecting Adobe Acrobat from *Help > About Plug-Ins > TextSpy Query...*

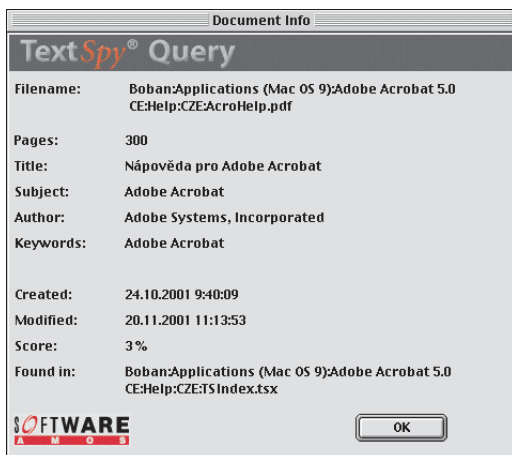
Click on the button located at the right of the query field to show up to ten previous searches:



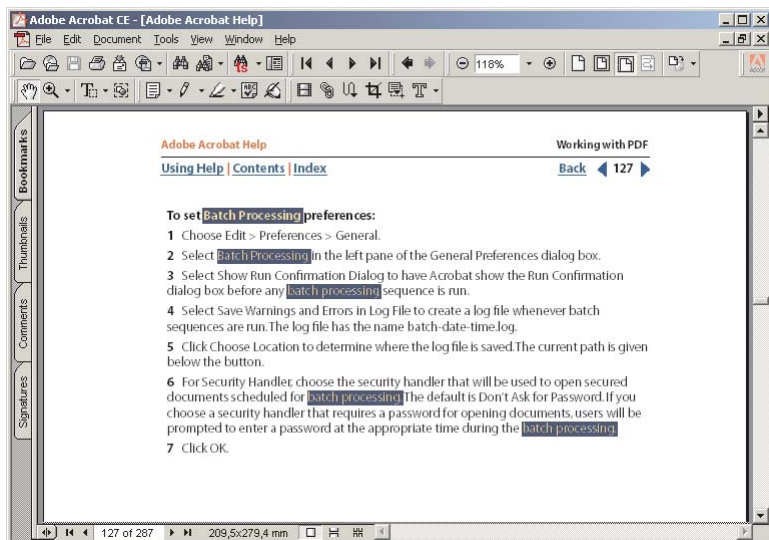
Click **Search** to search through the index files. If there are files matching entered query the following dialog box which displays a list of found pages opens:







The Query Results box shows the pages that match the search criteria in the order of incidence. This box shows the file name and page number. Additional information can be obtained by selecting a line from the query results list and clicking the *Info...* button.





To open selected page in Adobe Acrobat or Acrobat Reader with the highlighted searched items double-click on a line in the query results list or click *Show* button.



To go back to the query results list (found pages) select *Edit > TextSpy > Results* or use keyboard shortcut *Ctrl + ;* (Windows) or *Option + G* (Macintosh) or click on the icon  (Windows) or  (Macintosh) on the upper Adobe Acrobat toolbar.

To move forward in the query results list (found pages) select *Edit > TextSpy > Next Page* or use keyboard shortcut *Ctrl + >* (Windows) or *Apple + .* (Macintosh) or click the icon  (Windows) or  (Macintosh) on the Adobe Acrobat toolbar.

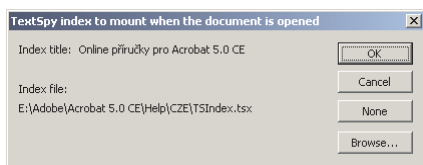
To go back select *Edit > TextSpy > Previous Page* or the keyboard shortcut *Ctrl + <* (Windows) or *Apple + ,* (Macintosh) or click the icon  (Windows) or  (Macintosh) on the upper Adobe Acrobat toolbar.

You can also move between entire documents and not only on individual pages. To move forward select *Edit > TextSpy > Next Document* or use keyboard shortcut *Ctrl + Shift + >* (Windows) or *Shift + Apple + .* (Macintosh). To move backward select *Edit > TextSpy > Previous Document* or use keyboard shortcut *Ctrl + Shift + <* (Windows) or *Shift + Apple + ,* (Macintosh).

Auto attaching of TextSpy indexes

You can add to PDF document the information about the index which will be automatically inserted into the list of searched indexes when the document is opened.

To attach an index to a PDF document or to change the attached index, choose *Edit > TextSpy > Attached Index* in Adobe Acrobat. Dialog box opens with information about attached index.



Click on *Browse* button to select index. Then click on *OK* and save the PDF document.

Licensing options for TextSpy

Index files created with TextSpy Indexer are for your individual use only or for sharing within your workgroup. These indexes are not for public distribution on CD-ROMs and other media. If you want to distribute the collection of PDF documents with generated index for TextSpy, you have to obtain license for the distribution and create the index using TextSpy Indexer CD-ROM Publishing Kit. For more information see <http://www.textspy.com> or <http://www.amsoft.cz>.

Known Issues

TextSpy cannot create an index correctly and function reliably with PDF files that use incorrectly localized fonts. These fonts do not work in new versions of Adobe products (InDesign, Photoshop 5.0 and later, Illustrator 8.0 and later) and in new versions of Microsoft Office. If you encounter this problem, request corrected fonts from the manufacturer of the fonts and generate new PDF files. To some extent you can verify whether the fonts are correct by selecting text with diacritical marks in a PDF file using the text tool and copying it to the clipboard. Open the Search dialog box (*Edit > Find*) and insert the text from the clipboard. The text in the dialog box should contain the same characters as in the page. On the contrary, even if the text is displayed incorrectly after this test, indexation may proceed accurately because TextSpy uses Eastern European coding better than Adobe Acrobat. This problem is not due to an error of the program, and the only solution of potential difficulties is to use correct fonts.

The creation of indexes can be memory intensive. On Mac OS 9 we recommend to increase memory partition for Adobe Acrobat 5 CE in Finder.

Technical Support

To get Amos Software TextSpy technical support please contact Amos Software technical support:

Amos Software, spol. s r. o.

Patočková 61

169 00 Praha 6 - Břevnov

Czech Republic

Tel.: +420/2/84011222

Fax: +420/2/84011212

E-mail: textspy@amsoft.cz

Additional Information

All commercial and technical information about Amos Software TextSpy, its use, licensing options and prices can be obtained from Amos Software:

Amos Software, spol. s r. o.

Patočková 61

169 00 Praha 6 - Břevnov

Tel.: +420/2/84011211

Fax: +420/2/84011212

E-mail: sales@amsoft.cz

<http://www.textspy.cz>

<http://www.textspy.com>

<http://www.amsoft.cz>

© 2000-2001 Amos Software, spol. s r. o. All Rights reserved.

This manual, as well as the software described in it, are furnished under license and may only be used or copied in accordance with the terms of such license. The information in this manual is furnished for informational use only, is subject to change without notice, and should not be construed as a commitment by Amos Software, spol. s r.o. Amos Software, spol. s r.o. assumes no responsibility or liability for any errors or inaccuracies that may appear in this book. Except as permitted by such license, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, recording, or otherwise, without the prior written permission of Amos Software, spol. s r.o.

TextSpy® is a registered trademark of AMOS Software, spol. s r.o. registered in the Czech republic and other countries.