User Manual for Qscreener

Qscreener.exe is desktop application for quick file search by processing user queries and semantic tune up of result by search of the files similar to those selected by user. The program is open source c# project. The first step in processing is making of local database for top folder selected by user. Program enumerates all internal folders and process files with selected file types.

Qscreener - Query Processing & Semantic Search	
Indexing Querying Semantic tune-up	
Select folder with documents	
C:\APOLAR\PATENTCORPUS128	Browse
Project name: data	project name is name of local subfolder v
Select file types 📝 *.txt 📝 *.	
Processing	
Status: making words' index database	Start
	Stop

Making database requires some time but runs fast enough. For example, 5000 ASCII files representing technical description of granted patents needs between 1 and 2 minutes on regular laptop. The database name is entered as project name in edit control. In provided example it is call 'data'.

Name	Date modified	Туре	Size
鷆 data	6/17/2012 1:35 AM	File folder	
🕑 Qscreener.exe	6/17/2012 1:18 AM	Application	228 KB

When completed the database sits in the same folder with executable. It can be copied and transferred to any other location or storage media. The database contains information about processed data in form of indexes of processed words and dictionary. It is possible to restore approximately the original information having database. That means if original information was confidential the database also should be treated as confidential information. After first step is completed the program can be terminated and restarted. It is possible to created several different databases and execute queries later on each of them separately.

The next step is querying. It uses different tab in GUI.

🧿 Qscreener - (Query Processing & Semantic Search	- • ×
Indexing Que	erying Semantic tune-up	
Database:	C:\APOLAR\UTILITIES\Qscreener\Qscreener\bin\Release\data	Browse
Query:	SCANNING-PROBE TECHNIQUES OR APPARATUS; APPLICATIONS OF SCANNING-PROBE TECHNIQUES, E.G., SCANNING-PROBE MICROSCOPY [SPM]	Search
Return size	16 - Proximity 4 - KMAX 8 -	

Location of the database folder is entered into database edit control. The query goes to next multi-lines edit control. The search button executes the search. Return size parameter defines the number of returned files. The proximity parameter establish admissible range between words. KMAX defines preference of results with longer match relative to shorter matches. These two coefficients are explained in <u>http://semanticsearchart.com/researchinvertex.html</u>. Values 4 and 8 are determined by program developer after some experimenting. Every user can experiment and find the best combination for any particular case. After query is processed, program opens next tab. It presents list of found files accompanied by the check boxes, allowing user to examine the content of each file, select those that look right and refresh the search. The program returns list of similar files along with those that were checked, so user can continue, approaching the identification of all files representing specific interest.

Qscreener - Query Processing & Semantic Search		×
Indexing Querying Semantic tune-up Select group of files using check boxes and hit 'Refresh' butter The program will find files in the collection similar to selected. Tune-up: ✓ C:\APOLAR\PATENTCORPUS128\C850P7971266.txt ✓ C:\APOLAR\PATENTCORPUS128\C850P793356.txt ○ C:\APOLAR\PATENTCORPUS128\C850P7810166.txt ○ C:\APOLAR\PATENTCORPUS128\C850P7810166.txt ○ C:\APOLAR\PATENTCORPUS128\C850P7810166.txt ○ C:\APOLAR\PATENTCORPUS128\C850P7818816.txt	on. Refresh	

Andrew Polar, 6/17/2012