

Logana Content Analyzer

LOGANA INFORMATION RESEARCH TECHNOLOGY

István Nagy; Logana Information Research Center, analog@logana.com, www.logana.com

*Motto: Shadow, a gloom, or a mystery
in front of me can't stay,
because I drive Logana's fiery chariot
through lines, pages,
searching for secret signs, words.
(Ogana Yan)*

Logana Content Analyzer is a special application of Logana Information Research Technology. The essence of this technology is illustrated by the demo program at www.logana.com, one of the menu items of which is the Logana Product Strategy, which presents the operating principle and other potential applications.

Using the Logana Content Analyzer

!!_LoganaContentAnalysis-EN.bat [the name of the file to be examined]

where the file to be examined is a text file of type **doc**, **docx**, or **txt**, which should also be placed in this folder.

Example:

!!_LoganaContentAnalysis-EN.bat 00_SampleText.doc

Purpose of Logana Content Analyzer

Logana Content Analyzer automatically generates the list of specific Keywords for the User-selected Text to Process (at different Limit Similarity values), and the text environment of these keywords in order to provide the User with a quick overview about the content of the text to be processed.

Introducing the Logana Content Analyzer through the processing of a specific text

- 1.) The text to be processed: 00_SampleText.doc
- 2.) The text, which is converted to txt: 01_SampleText.txt
- 3.) The sets of Collected Keywords in case of different Limit Similarity (50%, 52%, and 55%):
 - 02_KeyWords-EN_50.txt,
 - 02_KeyWords-EN_52.txt,
 - 02_KeyWords-EN_55.txt

Comment:

The program collected the Keywords based on the \$ANALOG_EN31.DIC Reference Word Collection (containing the commonly used english words) in such a way that just those words from the original text were retained, which at least in the measure of Limit Similarity are different from the words of this reference word collection (i.e. they are at most similar in measure of Limit Similarity).

Of course, the words in this Reference Word Collection can be modified arbitrarily by the User.
- 4.) The Text Environment of the collected Keywords in the Converted Text of type txt for different Limit Similarities:
 - 03_KeyEnvironment-EN_50.txt,
 - 03_KeyEnvironment-EN_52.txt,
 - 03_KeyEnvironment-EN_55.txt