OLANTO

myCAT

Self-Quote Detector User Manual

1. General Description

myCAT is a computer-assisted translation tool which includes several tools: a Text Aligner, a Quote Detector and a Self-Quote Detector. This document is the User Manual for the Self-Quote Detector.
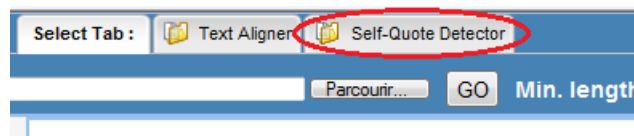
The Self-Quote Detector is a tool which detects whether some expressions are repeated within a document. If an expression is used several times across the text, it will find it and turn it into a hypertext so you can navigate from an occurrence of that expression to the next one.

This feature is useful if you want to see a specific expression in the various contexts in which it is used in the document to be translated, or if you have to share a translation between several translators and you want to harmonize the terminology.

The Self-Quote Detector also builds statistics tables at the end of the analyzed text in which expressions are classified according to their frequencies and which indicates the number of repetitions for each of them, along with other information.
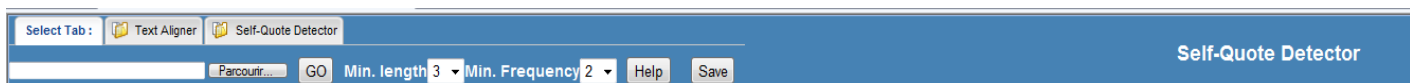
2. Accessing the Self-Quote Detector

The Self-Quote Detector is accessed from the initial interface by clicking on the "Self-Quote Detector" tab at the top of myCAT's home page:

Supported Browsers: myCAT supports Internet Explorer 8 or later and Mozilla Firefox 3.6 or later.
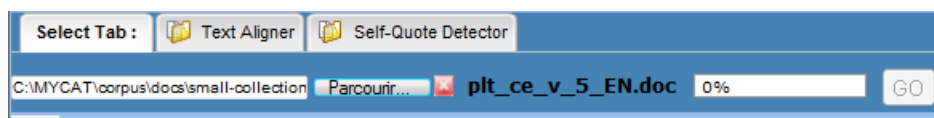
3. Using the Self-Quote Detector

Initially the Self-Quote Detector displays an empty screen with the following menu :

3.1. Uploading the Document to be Analyzed

To upload the document you want to analyze (that is, the document to be translated, for which you are looking for repeating expressions), click on the "Browse" button in the top left corner. A standard Browse box will open.

In that Browse box, go to the location of the document to be analyzed and click on the "Open" button. The name of the file being uploaded is displayed along with a progression bar:

while in the bottom bar the message "Uploading File..." is displayed.
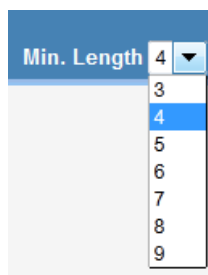
This operation is completed when the document appears in the central frame of the interface.

Notes:

- The supported formats are doc, docx, pdf, html and txt.

- There is a **limit to the size of the documents which can be uploaded**: by default, this limit is set to **3'072 Kb**, *i.e.* 3 Mb. This is to avoid submitting very large documents which would unnecessarily overload the Self-Quote Detector.
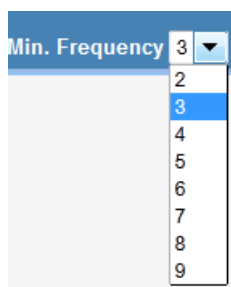
3.2. Setting Up the Minimum Length of Quotes

You can choose the minimum number of words you would like to see in the detected quotes by using the "Minimum Length" scroll-down menu. This function avoids detecting very small quotes, which could be numerous but not very useful. The lowest length is 3 consecutive words; the highest is 9. Please note that this highest possible value only applies to the <u>minimum</u> number of words to appear in a quote; there is no maximum length for the quote. In fact the system always looks for the longest possible quote.
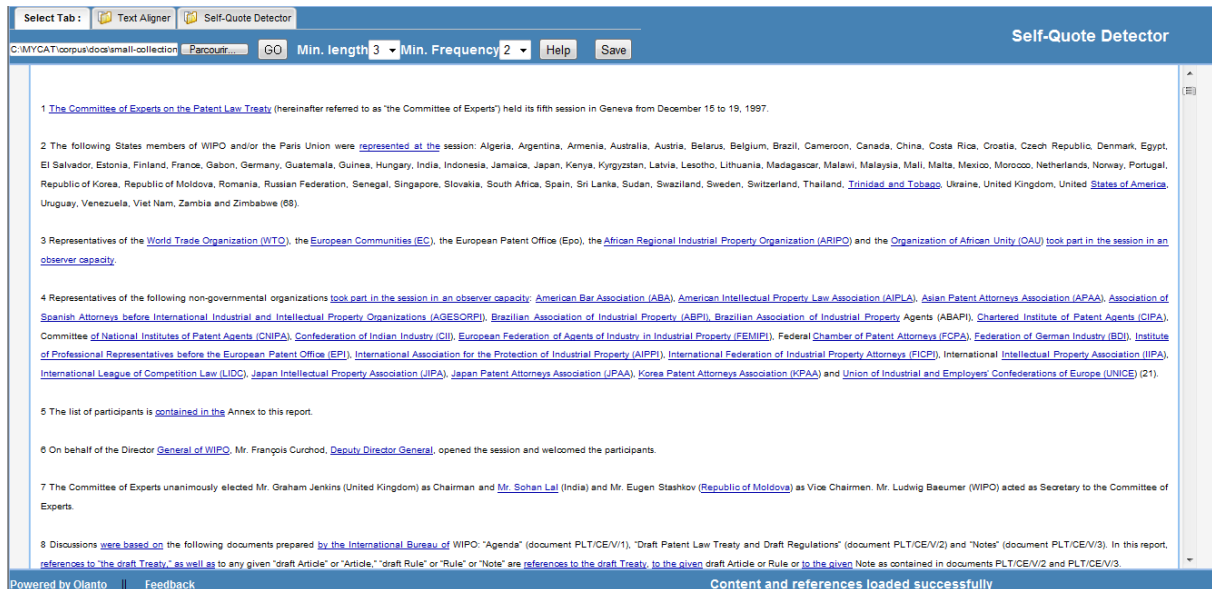

3.3. Setting Up the Minimum Number of Occurrences

You can also define the minimum number of times the expression should appear in the text in order to be considered by the Self-Quote Detector. If you want all of the repeated expressions (*i.e.* the expressions which appear at least twice in the text) select 2 in the "Minimum Frequency" menu:

3.4. Launching the Analysis

When all the parameters are set up you can click on the "GO" button to launch the analysis. The task is completed when the document to be analyzed is displayed in the top frame with one or several expressions appearing in hypertext :

Clicking on any reference (hypertexted words) will position the text at the following occurrence of the clicked terms. The following occurrence is normally displayed at the top of the screen. If you are positioned on the last occurrence of the expression, the Self-Quote Detector takes you back to the first occurrence.

3.5. Using the "Save" Button

After a document was analyzed it is possible to save the resulting file with all its internal references. This document can later be open in any browser to view its self-references and navigate among them.

To do so, after the analysis is completed, simply click on the "Save" button. The browser will propose to either open or save the file; choose "Save as" and specify the place where you want to store the HTML file.

3.6. Understanding the Statistics Created with the "Save" Button

The Self-Quote Detector builds, at the end of the analyzed text, a statistics section which includes the following elements:

- **General statistical information**:

  parsing of document: C:\MYCAT\corpus\docs\small-collection\WIPO\PLT\plt_ce_v_5_EN.doc

  Number of expressions found: 1019

  Minimum number of consecutive words: 3

  Minimum number of occurrences: 2

- **Expressions sorted by number of occurrences and then by length**

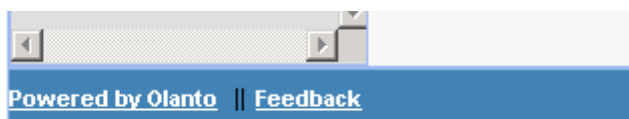| #	Expression (Source Language)	No. of Occurences	Length
1	of the application	7	3
2	of intellectual property	6	3
3	paragraph 2 b	6	3
4	of the text	5	3
5	paragraph 2 a	5	3
6	would not be	5	3
7	and that the	5	3
8	in lin 2	5	3
9	the delegation of the united states of america stated that	4	10
10	this rule was approved as proposed draft article	4	8
11	the international bureau explained that this	4	6
12	the provisions of the	4	4
13	should be replaced by	4	4
14	the representative said that	4	4
15	the delegation of japan	4	4

etc...

3.7. Using the "Help" Button

This User Manual can be accessed at anytime from the Self-Quote Detector interface by clicking on the Help button at the top right of the screen.

3.8. Coming Back to the Text Aligner

It is possible at anytime to come back to the Text Aligner environment by clicking on the "Text Aligner" tab. Alternatively you can choose to open each tool in a different window so you can keep both of them open at the same time.

3.9. Links in the Bottom Bar

Two links are provided at the bottom left corner of the screen, as illustrated :

- **Powered by Olanto :** This link points to the web site of the Olanto Foundation, which is the publisher of myCAT (see http://olanto.org ; all the software published by Olanto is distributed under the AGPL open source license).

- **Feedback :** This links allows you to send an email in order to provide your feedback about myCAT. The default email address is info@olanto.org; it should be changed for the email of your Translation Support Unit.