

6es Rencontres FORMIST, 15 juin 2006 : Exploitation et usages de l'information par les étudiants avancés

Outils de résumé et de traduction automatiques

Outils de résumé et de traduction automatiques. In *6es Rencontres FORMIST : Exploitation et usages de l'information par les étudiants avancés, l'ens sib à Villeurbanne, 15 juin 2006* [en ligne].
Format PDF.

Disponible sur : <<http://www.enssib.fr/bibliotheque-numerique/notice-1669>>

Ce document est « **tous droits réservés** ». Il est protégé par le droit d'auteur et le code de la propriété intellectuelle. Il est strictement interdit de le reproduire, dans sa forme ou son contenu, totalement ou partiellement, sans un accord écrit de son auteur.

L'ensemble des documents mis en ligne par l'ens sib sont accessibles à partir du site :

<http://www.enssib.fr/bibliotheque-numerique/>

Outils de résumé et de traduction automatiques

Pour finir cet exposé, il convient de présenter une dernière famille d'outils qui, s'ils ne sont pas encore parfaitement au point, peuvent rendre service. Il convient cependant d'alerter sur leur degré limité de pertinence et d'efficacité.

Ces outils permettent ainsi de réaliser des résumés et traductions automatiques, ce qui peut être très utile lorsque l'étudiant se trouve confronté à des documents de taille volumineuse pour en extraire rapidement l'information ou pour comprendre des documents écrits dans des langues qu'il ne maîtrise pas.

Le résumé automatique (*text mining*) consiste en une extraction automatique d'informations dans le corpus désigné, à travers des analyses linguistiques et mathématiques. Cette analyse aboutit à un résumé, mais peut aussi conduire à une catégorisation des données (génération automatique de classement), valable alors essentiellement pour des documents en grands nombres.

La traduction automatique se fait au mot à mot (comme un dictionnaire) ou porte sur l'intégralité d'un texte. Elle s'effectue soit à travers un logiciel de traduction installé sur l'ordinateur, soit directement en ligne, en fournissant l'URL de la page ou le fichier du document à traduire.

Copernic Summarizer est un outil de résumé automatique qui propose une version de démonstration <<http://www.copernic.com/en/products/summarizer/download.html>>.

Ces outils extraient essentiellement les premières lignes de certaines parties de document, définies comme significatives selon leurs algorithmes, ce qui donne lieu à un résumé qui peut sembler assez incohérent ! À noter qu'il est possible de déterminer le taux de résumé : plus celui-ci est élevé, plus le contenu extrait perd du sens.

À noter que des outils de traitement de texte très répandus, comme *Word*, proposent directement cette fonction de résumé automatique (Outils/Synthèse automatique).

La traduction automatique en ligne est souvent limitée en nombre de caractères. *Babelfish* <<http://babelfish.altavista.com/>>, disponible sous plusieurs versions, est un des outils les plus connus. *Reverso* <<http://www.reverso.net/>> est une autre possibilité. Il suffit de copier le texte à traduire, parfois d'en indiquer l'URL, de dire le sens de traduction, c'est-à-dire quelle est la langue

d'origine et la langue de destination, pour obtenir une traduction automatique, souvent « à la louche » et parfois très originale.

Panorama des outils de text mining. 2005 :

<<http://www.gillesbalmisse.com/blog/index.php?2005/02/28/39-panorama-des-outils-de-text-mining>>

Softissimo. Livre blanc Traduction

<<http://www.softissimo.com/products/infotrad.htm>>

ASSELIN C. Traduction gratuite sur Internet.

<<http://c.asselin.free.fr/french/traduction.htm>>

CETADIR. Perles de traductions.

<<http://www.cetadir.com/html/perles.php>>