

Libellex : une plateforme multiservices pour la gestion des contenus multilingues

François Brown de Colstoun¹, Estelle Delpech^{1,2}, Étienne Monneret¹

- (1) LINGUA ET MACHINA, Laval Technopole, 6 rue Léonard de Vinci, 53001 Laval Cedex
et c/o Inria, Rocquencourt BP 105, 78 153 Le Chesnay Cedex
(2) LINA FRE CNRS 2729, 2 rue de la Houssinière BP 92208, 44322 Nantes Cedex 3
fbc(a)lingua-et-machina.com, ed(a)lingua-et-machina.com, em(a)lingua-et-machina.com

Cette démonstration industrielle présente *Libellex*, un prototype de plateforme multiservices pour la gestion des contenus multilingues. Cette plateforme vise à la fois des professionnels de la langue (traducteurs, terminologues) et un public plus général amené à communiquer en langue étrangère dans un contexte professionnel. *Libellex* assiste ces utilisateurs dans la production, la compréhension et la traduction de documents dans des langues ou des domaines qu'ils maîtrisent bien, imparfaitement ou pas du tout. Pour cela, *Libellex* propose une palette de services exploitant une base de connaissances linguistiques constituée automatiquement à partir de textes fournis par l'utilisateur. Ces services vont de la simple recherche d'expression à la pré-traduction et la constitution automatique de terminologies bilingues. Le processus d'extraction de connaissances linguistiques s'appuie sur des algorithmes issus des recherches en traduction automatique (Gale et Church 1993 ; Lardilleux, 2010), traduction assistée par ordinateur (Planas, 2000), terminologie computationnelle (Bourigault, 1994) et exploitation des corpus comparables (Fung, 1997 ; Morin et Daille, 2009).

Remerciements

Libellex a été en partie financé par l'ANR (subvention no. ANR-08-CORD-009) et par Oséo (subvention no. A1010034Z). Nous remercions également Guillaume Pelluau et Mikaël Morardo pour leur participation au développement de *Libellex*.

Références

- BOURIGAULT, D. (1994). LEXTER un Logiciel d'EXtraction de TERminologie. Application à l'extraction des connaissances à partir de textes. *Thèse en Mathématiques, Informatique Appliquée aux Sciences de l'Homme*. École des Hautes Études en Sciences Sociales.
- FUNG, P. (1997). Finding Terminology Translations from Non-parallel Corpora. Dans *5th Workshop on Very Large Corpora*, Hong Kong, p. 192-202.
- GALE, W. A., ET K. W. CHURCH. (1993). A program for aligning sentences in bilingual corpora. *Computational Linguistics* 19(1): 85-102.
- LARDILLEUX, A. (2010). Contribution des basses fréquences à l'alignement sous-phrastique multilingue : une approche différentielle. *Thèse en Informatique*. Université de Caen Basse-normandie.
- MORIN, E., ET B. DAILLE. (2009). Compositionality and lexical alignment of multi-word terms. Dans *Language Resources and Evaluation (LRE), Multiword expression: hard going or plain sailing*, P. Rayson, S. Piao, S. Sharoff, S. Evert, B. Villada Moirón, p. 79-95.
- PLANAS, E. (2000). Multi-level Similar Segment Matching Algorithm for Translation Memories and Example-Based Machine Translation. Dans *Proceedings of the 18th Conference on Computational Linguistics*, Saarbrücken, Germany, p. 622-627.