

Analyser des corpus de discussions multilingues avec CALICO - Ex. tirés de Galanet

*François-Marie BLONDEL, Christophe REFFAY, STEF, ENS-Cachan / IFÉ (ENS-Lyon),
Emmanuel GIGUET, GREYC, Université de Caen Basse-Normandie, CNRS,
Sandrine DEPREZ, Universitat Autònoma de Barcelona et Université Stendhal Grenoble3*

Qui a le plus contribué à une discussion ? En nombre de messages ou en volume ? Qui a répondu à qui ? Qui a parlé de quoi ? Dans quelle langue ? Avec quels mots ? ... Autant de questions qui peuvent se poser couramment dans l'analyse d'un corpus de communication médiée par ordinateur.

La plate-forme Calico a été développée pour offrir des moyens de partage et d'analyse de forums à un regroupement de chercheurs intéressés par l'emploi de ces modalités de communication dans la formation (Bruillard, 2008). L'intention qui en a guidé la conception était de fournir des instruments qui facilitent la confrontation et les comparaisons de méthodes et d'outils d'analyses, avec la visée de transformer et d'affiner ces instruments en vue de leur utilisation à terme par des formateurs et des tuteurs.

Calico est une des rares plates-formes de partage de corpus de discussions et d'outils d'analyse en accès libre. Les discussions analysées peuvent provenir de forums, de listes de diffusion, de blogs ou de listes de messages avec commentaires provenant d'un dispositif de communication écrite. Conçus à l'initiative des chercheurs ou à la demande des utilisateurs, les outils disponibles offrent des vues globales et synthétiques sur les discussions et permettent la recherche et la visualisation dans le texte des messages de marqueurs extraits du lexique des acteurs.

Depuis son ouverture à un large public en 2009, plusieurs centaines de discussions, dont une importante collection issue de l'expérimentation Simuligne, ont été mises en ligne et ont pu être analysées par des chercheurs extérieurs au projet initial (Blondel & Giguët, 2011).

Par leur indépendance vis-à-vis de la langue, les outils de Calico se prêtent particulièrement bien à l'analyse de discussions ayant lieu dans un contexte multilingue, comme c'est le cas des messages exprimés dans six langues romanes sur la plateforme Galanet.

L'objectif de cet atelier est de présenter la plate-forme et les outils de Calico et de discuter les analyses de données d'interaction qui peuvent être effectuées au moyen de ces outils.

Dans un premier temps, seront présentées les principales fonctions de la plate-forme et des outils : dépôt de corpus, visualisation du volume des messages, de la fréquence des messages, de la simultanéité des discussions, calcul du lexique des participants, coloration thématique des messages...

Dans un second temps, les participants pourront observer et commenter les relations entre des analyses effectuées et les outils de Calico, à partir de cas d'études effectuées sur des extraits de discussions issues de Galanet. Ces exemples permettront d'explorer quelques techniques utilisables pour rechercher des marques et des indices d'interaction plurilingue, de réactivité, d'étagage réciproque, ou pour quantifier la présence des différentes langues dans un forum, dans un sujet de discussion, voire pour estimer la fréquence des alternances codiques intra et inter-messages.

Références (et sitographie)

Blondel, F.-M. & Giguët, E. (2011). Analyses et partages de corpus de discussions avec Calico - Leçons tirées d'une expérience récente. In Dejean, C., Mangenot, F., Soubrié, T. (coord.). *Actes du colloque Epal 2011 (Échanger pour apprendre en ligne)*, Université Stendhal – Grenoble 3. http://w3.u-grenoble3.fr/epal/dossier/06_act/pdf/epal2011-blondel-giguët.pdf

Bruillard, É. (2008). Teacher development, discussion lists and forums: issues and results. In K. McFerrin, R. Weber, R. Carlsen & D.A. Willis (eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference, SITE 2008*. Chesapeake, USA: AACE. pp. 2950-2955.

Giguët, E., Lucas, N., Blondel, F.-M. & Bruillard, É. (2009). Share and explore discussion forum objects on the Calico website. In A. Dimitracopoulou, C. O'Malley, D. Suthers & P. Reimann (Eds.) *Computer Supported Collaborative Learning Practices: CSCL 09 Community Events Proceedings* (pp. 174-176). International Society of the Learning Sciences (ISLS). http://www.stef.ens-cachan.fr/annur/blondel/cscl09_calico_share.pdf

Calico : <http://woops.crashdump.net/calicorss/> ; groupes de chercheurs : <http://www.stef.ens-cachan.fr/calico/calico.htm> ; plate-forme Galanet : <http://www.galanet.eu>