

DEUTSCH IN EUROPA

Sprachpolitisch, grammatisch, methodisch

METHODENMESSE

Mittwoch, 11. März 2020, 16:00 Uhr

MULTILINGWIS: EINE WEBBASIERTE SUCHMASCHINE ZUR ERFORSCHUNG VON PARALLELEN UND MULTIPARALLELEN KORPORA

Anastassia Shaitarova

Das Institut für Computerlinguistik der Universität Zürich hat eine bedeutende Sammlung von offen verfügbaren parallelen Korpora (d.h. großen Sammlungen übersetzter Texte) zusammengestellt. Multilingwis ist eine webbasierte Anwendung, die es der Öffentlichkeit ermöglicht, die Sammlung bequem zu durchsuchen und von ihrem Sprachwissen zu profitieren. Im Wesentlichen ist Multilingwis ein Konkordanzsystem für parallele Korpora, das Beispielsätze in mehreren Sprachen zeigt. Für Übersetzer und Sprachlerner ist Multilingwis von großem Nutzen, da es nicht nur verschiedene Anwendungsbeispiele und Übersetzungsvarianten bietet, sondern auch die Bestimmung von Synonymen durch sprachübergreifende Spiegelung ermöglicht. Darüber hinaus bietet Multilingwis die Möglichkeit, die Übersetzung von Mehrwortausdrücken zu bewerten, Übersetzungsergebnisse in verschiedenen Kontexten und Domänen zu vergleichen und die Häufigkeitsstatistiken von Übersetzungsvarianten zu beobachten. Das Tool kann hier eingesehen werden: <<http://pub.cl.uzh.ch/purl/multilingwis>>.

Auf der Methodenmesse werden wir die verfügbaren Korpora präsentieren und sie mit Multilingwis durchsuchen. Wir zeigen Ihnen die Vorteile und Herausforderungen, die mit diesem Tool verbunden sind.

Literatur:

- Clematide, S., J. Graen, and M. Volk (2016). "Multilingwis – A Multilingual Search Tool for Multi-Word Units in Multiparallel Corpora". In: *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives/Fraseología computacional y basada en corpus: perspectivas monolingües y multilingües*. Ed. by G. C. Pastor. Geneva: Tradulex.
- Graen, J., T. Kew, A. Shaitarova, and M. Volk (July 2019). "Modelling Large Parallel Corpora: The Zurich Parallel Corpus Collection". In: *Challenges in the Management of Large Corpora (CMLC-7)*. Ed. by P. Banski, A. Barbaresi, H. Biber, E. Breiteneder, S. Clematide, M. Kupietz, H. Lungen, and C. Iliadi. Leibniz-Institut für Deutsche Sprache.
- Graen, J., D. Sandoz, and M. Volk (May 2017). "Multilingwis2 – Explore Your Parallel Corpus". In: *Proceedings of the 21st Nordic Conference on Computational Linguistics, NoDaLiDa*. Linköping Electronic Conference Proceedings 131. Linköping University Electronic Press, pp. 247–250.