TILBURG ◆ ◆ UNIVERSITY

Tilburg University

TraMOOC (Translation for Massive Open Online Courses): Providing Reliable MT for MOOCs

Kordoni, Valia; Birch, Lexi; Buliga, Ioana; Cholakov, Kostadin; Egg, Markus; Gaspari, Federico; Georgakopoulou, Yota; Gialama, Maria; Hendrickx, I.H.E.; Jermol, Mitja; Kermanidis, Katia; Moorkens, Joss; Orlic, Davor; Papadopoulos, Michael; Popovic, Maja; Sennrich, Rico; Sosoni, Vilelmini; Tsoumakos, Dimitrios; van den Bosch, Antal; van Zaanen, Menno; Way, Andy

TraMOOC (Translation for Massive Open Online Courses): Providing Reliable MT for MOOCs

Valia KORDONI, Lexi BIRCH, Ioana BULIGA, Kostadin CHOLAKOV,
Markus EGG, Federico GASPARI, Yota GEORGAKOPOULOU,
Maria GIALAMA, Iris HENDRICKX, Mitja JERMOL,
Katia KERMANIDIS, Joss MOORKENS, Davor ORLIC,
Michael PAPADOPOULOS, Maja POPOVIĆ, Rico SENNRICH,
Vilelmini SOSONI, Dimitrios TSOUMAKOS,
Antal van den BOSCH, Menno van ZAANEN, Andy WAY

Humboldt Universität zu Berlin (Coordinator, Germany), Dublin City University (Ireland),
University of Edinburgh (UK), Ionian University (Greece), Stichting Katholieke Universiteit
(Radboud University, Netherlands), EASN-Technology Innovation Services BVBA (Belgium),
Deluxe Media Europe Ltd (UK), Stichting Katholieke Universiteit Brabant (Tilburg University,
Netherlands), Iversity GMBH (Germany), Knowledge 4 All Foundation LBG (UK)

`info@tramooc.eu`

Abstract. TraMOOC is a 3-year EU-funded Horizon 2020 collaborative project that started in February 2015 (ICT-17-2014: Cracking the language barrier; project reference: 644333) which aims at providing reliable MT for MOOCs. The main outcome of the project will be a high-quality semi-automated MT service for all types of educational textual data available on a MOOC platform. The service will support nine European languages, in addition to Russian, and simplified Chinese. Progress may be monitored via the official project website at TraMOOC.eu, which is updated regularly with publications, news, and public deliverables.

TraMOOC: Reliable MT for Massive Open Online Courses

Language barriers constitute major impediments to sharing massive open online courses (MOOCs) with all peoples and in the attempt to educate all citizens. The EU-funded TraMOOC (Translation for Massive Open Online Courses) project aims at tackling this major issue by developing high-quality MT for all types of text included in MOOCs (e.g. assignments, tests, presentations, lecture subtitles, blog text) from English into eleven languages (DE, IT, PT, EL, DU, CS, BG, HR, PL, RU, ZH). The core of the TraMOOC platform will be open-source with some premium add-on services, enabling the integration of any MT solution in the educational domain, for any language. The project is conducting a major crowdsourcing drive, creating more training and evaluation data, while also performing comparative evaluations of phrase- and syntax-based SMT systems and state-of-the-art neural network MT engines, using standard metrics of fluency and adequacy, along with more detailed error taxonomies.