

Deadline extended to June 10th - 3rd CfP - 4th Workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT 2019)

The 4th Workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT 2019)

Malaga, 27 September 2019

<http://www.lexytrad.es/europhras2019/mumttt-2019-2/>

Deadline extended to June 10th, 2019

In conjunction with EUROPHRAS 2019 – International Conference “Computational and Corpus-based Phraseology: Recent advances and interdisciplinary approaches”

Third Call for Papers

Following the success of the three previous editions of the workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT) we are announcing the fourth edition to be held in conjunction with the International Conference “Computational and Corpus-based Phraseology’ (Europhras’2019)” which is jointly organised by the European Association for Phraseology EUROPHRAS, the University of Malaga (Research Group in Lexicography and Translation), the University of Wolverhampton (Research Group in Computational Linguistics) and the Association for Computational Linguistics – Bulgaria.

The MUMTTT workshop will be held on the last day of the Europhras’2019 conference, namely on 27th September 2019. It will provide a forum for researchers and practitioners in the fields of (Computational) Linguistics, (Computational) Phraseology, Translation Studies and Translation Technology to discuss recent advances in the area of multi-word unit processing and to coordinate research efforts across disciplines in order to improve the integration of multi-word units in machine translation and translation technology tools.

Multi-word units are word combinations which range from compounds such as ‘credit card’ to idiomatic expressions such as “it is raining cats and dogs’ and are acknowledged as one of the major challenges in natural language processing (NLP), because of their lexical, syntactic, semantic, pragmatic and/or statistical idiosyncrasies.

In spite of the relative progress achieved in translation technology with the adoption of neural approaches and in the processing for particular types of units such as verb-particle constructions,

the identification, interpretation and translation of multi-word units in general still represent open challenges, both from a theoretical and a practical point of view. The idiosyncratic morpho-syntactic, semantic and translational properties of multi-word units pose many obstacles even to human translators, mainly because of intrinsic ambiguities, structural and lexical asymmetries between languages, and, finally, cultural differences.

In recent years, growing attention has been paid to integrating multi-word units (MWUs) in machine translation and translation technology tools, as it has been acknowledged that it is not possible to create large scale language solutions without properly handling MWUs of all types. As a matter of fact, researchers are now addressing the problems posed by MWU processing and translation using different formalisms and techniques, such as automatic recognition of MWUs in a monolingual or bilingual setting; alignment and paraphrasing techniques; development and use of (handcrafted) monolingual and bilingual language resources; creation of annotated monolingual and parallel corpora, development of strategies for handling syntactically flexible units in language analysis and translation modules, development of evaluation projects. The aim of the workshop is to bring together researchers and practitioners working on MWU processing from various perspectives, in order to enable cross-fertilisation and foster the creation of innovative solutions that can only arise from interdisciplinary collaborations. In particular, the workshop welcomes interactions between NLP researchers working on the computational treatment of multi-word units, experts in (computational) phraseology working on challenging topics of their discipline, as well as translation practitioners, to the benefit of applying their latest results to advance the state of the art in MWU processing.

Topics of Interest

The MUMTTT 2019 workshop invites the submission of papers reporting on original and unpublished research on topics related to MWU processing in machine translation and translation technology, including:

- Lexical, syntactic, semantic and translational aspects in MWU representation
- Theoretical approaches to MWUs (e.g., collostructional analysis of MWU, cognitive approaches to processing MWUs, etc.)
- Development of multilingual MWU resources
- Identification and acquisition of MWUs and variants
- Learning semantic information about MWUs from monolingual, parallel or comparable corpora
- Development and use of MWU resources in machine translation and translation technology
- Development of corpora for the extraction and translation of MWUs
- Compilation of resources for the extraction and translation of multiword units
- Creation of MWU-annotated corpora with a focus on translation aspects
- Paraphrasing of MWUs applied to the improvement of machine translation
- MWUs and word alignment techniques
- MWUs in machine translation
- MWUs in translation memory systems
- MWUs in term extraction
- MWU-centred machine translation evaluation
- Evaluation of MWU translation
- MWUs in CAT tools
- Multilingualism and MWU processing

- Psycholinguistic studies of MWU processing in a bilingual setting.

Submission and Publication

Submissions must consist of full-text papers and should not exceed 7 pages excluding references.

A selection of papers will be invited to submit an extended version for the second volume of Multiword Units in Machine Translation and Translation Technology to be published by John Benjamins by the end of 2019.

They must be formatted according to the ACL 2017 style guidelines available both for Word and LaTeX text processor. Each submission will be reviewed by at least three programme committee members. Accepted papers will be presented orally or as posters, as determined by the programme committee. There will be no distinction in the workshop proceedings between papers presented orally or as posters. The proceedings will be published both as e-proceedings with ISSN and as an electronic volume with ISBN and will be made available at the time of the conference.

All submissions must be made online using the Softconf START conference management system: <https://www.softconf.com/j/mumttt2019>.

Authors of accepted papers will receive further instructions on how to produce camera-ready versions of their papers for inclusion in the proceedings.

Important Dates

Deadline for paper submission: **10 June 2019**

Acceptance notification: **1 July 2019**

Final version: **1 September 2019**

MUMTTT 2019 workshop: **27 September 2019**

Workshop invited speakers

Jean Pierre Colson, Université Catholique de Louvain

Aline Villavicencio (Federal University of Rio Grande do Sul and University of Essex)

Workshop Chairs

Johanna Monti, Università degli Studi di Napoli "L'Orientale", Italy

Gloria Corpas Pastor, Universidad de Màlaga, Spain

Ruslan Mitkov, University of Wolverhampton, United Kingdom

Programme Committee

Giuseppe Attardi, University of Pisa
Verginica Barbu Mititelu, Romanian Academy Research Institute for Artificial Intelligence, Romania
Jean Pierre Colson, Université Catholique de Louvain
Matthieu Constant, Université de Lorraine
Maria Pia di Buono, TakeLab – Faculty of Electrical Engineering and Computing – University of Zagreb, Croatia
Antoine Doucet, University of La Rochelle
Thomas François, Université Catholique de Louvain
Marcello Federico, Amazon
Federico Gaspari, University for Foreigners “Dante Alighieri”
Tatjana Gornostaja, Tilde
Philipp Koehn, The Johns Hopkins University
Alessandro Lenci, University of Pisa
Rodolfo Maslias, Terminology Coordination Unit at the European Parliament
Michael Oakes, University of Wolverhampton
Carla Parra Escartín, ADAPT Centre, Dublin City University
Pavel Pecina, Charles University
Éric Poirier, Université du Québec à Trois-Rivières
Carlos Ramisch, Aix Marseille University
Michael Rosner, University of Malta
Federico Sangati, University of Naples “L’Orientale”
Agata Savary, Université François Rabelais Tours
Gerold Schneider, University of Zurich
Max Silberztein, Université de Franche-Comté
Kathrin Steyer, Institut für Deutsche Sprache, Mannheim
Amalia Todirascu, LILPA, Université de Strasbourg
Beata Trawinski, Institut für Deutsche Sprache, Mannheim
Marco Turchi, FBK
Agnes Tutin, Université Grenoble 3, LIDILEM
Aline Villavicencio, Federal University of Rio Grande do Sul
Veronika Vincze, Hungarian Academy of Sciences
Martin Volk, University of Zurich

Further information

For further information, please contact Ms. Annalisa Raffone from the workshop Organising Committee at mumttt2019@gmail.com