The Second Workshop on
Human-Informed Translation
and Interpreting Technology

**Second Workshop on Human-Informed
Translation and Interpreting Technology (HiT-IT 2019)**

Proceedings of the Workshop

September 5 - 6,
2019 Varna, Bulgaria

Preface

The Second Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT 2019) took place in Varna, Bulgaria and spanned over two days (5-6 September 2019), as a post-RANLP 2019 conference event. This is a continuation of the one-day workshop HiT-IT 2017 (`http://rgcl.wlv.ac.uk/hit-it/`) organised in conjunction with RANLP 2017. Both workshops proved very popular attracting the highest number of participants across all the RANLP workshops organised in their respective years.

In addition to academic submissions, this year's workshop also welcomed submissions from industry (translation agencies or companies developing translation and interpreting technologies) and practitioners (translators and interpreters). This was to reflect the fact that the HiT-IT events seek to act as a meeting points for researchers working in translation and interpreting technologies, practicing technology-minded translators and interpreters, companies and freelancers providing services in translation and interpreting as well as companies developing tools for translators and interpreters, including researchers and developers of machine translation engines.

Translation Technology (TT) has the core objectives of speeding up and easing the translation process, and of assisting human translators. The emerging field of Interpreting Technology also seeks to support the work of interpreters. TT relies heavily on methods developed in the field of Natural Language Processing (NLP) and Computational Linguistics (CL). Typical examples are Computer-Assisted Translation (CAT) tools, electronic dictionaries, concordancers, spell-checkers, terminological databases, terminology extraction tools, translation memories, partial machine translation of template documents, speech recognition systems for automatic subtitling, to name just a few. However, quite often these tools do not address the actual needs and quality required by translation and interpreting professionals.

Natural Language Processing and Machine Translation (MT) make use of the knowledge and expertise of professional translators and interpreters in order to build and improve models for automatic translation or for developing more advanced CAT tools. This includes using parallel aligned human translations and speech interpretation corpora for machine learning, human evaluation of machine translation outputs and human annotations, or by trying to learn from humans, in order to fine-tune their algorithms. However, NLP and MT researchers rarely meet in person with professional translators and interpreters to learn from their expertise and points of view.

By organising HiT-IT, we aimed to provide a discussion forum for professionals working on or with translation and interpreting technologies (including machine translation), in order to find ways to make machine translation output closer to human quality. These discussions also focused on ensuring that CAT tools are being developed in a way that will ease and speed-up translation and interpreting practitioners' job.

This year HiT-IT 2019 featured a large Programme Committee, consisting of well-known experts coming from industry, universities, and active practitioners. The nationalities and affiliations of PC members covered almost all continents and included: Europe (Belgium, Bulgaria, Czech Republic, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Portugal, Spain, Switzerland, UK), Australia, South America (Argentina, Brazil), Asia (Hong Kong, Japan, Jordan, Qatar), Africa (South Africa).

HiT-IT 2019 attracted a variety of participants with various backgrounds, including: machine translation development specialists, practicing translators and interpreters, owners and representatives of translation agencies from Bulgaria and from overseas, developers of tools and resources for translators and interpreters (including of translation memories), researchers working on the evaluation of machine translation, and computer-assisted translation tools, including automatic translation of subtitles,

researchers, applying Natural Language Processing techniques on human translation and interpreting, as well as TT students. This year HiT-IT invited submissions on theoretical ideas, practical applications, and position papers promoting new ideas, challenging the current status of the fields and proposing how to take them forward.

The topics of submissions included:

User needs:

- interpreting and translation tools: user needs and user requirements

- incorporating human knowledge into translation and interpreting technology

- what existing tools for translators (including subtitlers) and interpreters do not offer

- user requirements for electronic resources for translators and interpreters

- translation and interpreting workflows in larger organisations and the tools employed for translation and interpreting

Existing methods and resources:

- latest developments in translation and interpreting technology

- electronic resources for translators and interpreters

- annotation of corpora for translation and interpreting technology

- crowdsourcing techniques for creating resources for translation and interpreting

- latest advances in pre-editing and post-editing of machine translation

- human-informed (semi-)automatic generation of interlingual subtitles

- technology for subtitling

Evaluation:

- (human) evaluation of translation and interpreting technology

- crowdsourcing techniques for evaluating translation and interpreting

- evaluation of discourse and other linguistic phenomena in (machine) translation and interpreting

- evaluation of existing resources for translators and interpreters

- human evaluation of neural machine translation

Other:

- position papers discussing how machine translation should be improved to incorporate the expertise of translators and interpreters

- translation and interpreting technologies' impact on the market

- comparison between human and machine translation

- changes in the translators and interpreters' professions in the new technology era especially as a result of the latest developments in Neural Machine Translation

HiT-IT 2019 featured four invited talks and one special presentation. Yves Champollion from WordFast gave two presentations: the "Rosetta Stone Decipherment" and "MT Acceptance Among Translators: Are We Nearing the Tipping Point". Dr Vilelmini Sosoni from the Ionian University, Greece, gave a talk on "Translators and Technology: Dancing a Tango Nuevo". Dr Carla Parra Escartín, the Director of Linguistic Services at Unbabel Portugal gave a talk on "The story of how academic and industrial research meet to put humans in the loop of Artificial Intelligence".

Professor Dr Ruslan Mitkov, director of the Research Institute in Information and Language Processing and Head of the Research Group in Computational Linguistics at the University of Wolverhampton, UK gave a special presentation on "The world's first Erasmus Mundus Masters programme in technology for translation and interpreting and the new generation of translators and interpreters".

The authors came from a large number of countries: Algeria, Austria, Belgium, Bulgaria, Brazil, Egypt, France, Greece, Hong Kong, Portugal, Qatar, Russia, Spain and Switzerland. The HIT-IT 2019 presentations attracted a lot of interest and generated active discussions between listeners and presenters. The discussed topics included: human and machine evaluation of machine translation; using NLP, automatic tools and techniques to study and extract meaningful patterns from interpreting, human translation, and manual post-editing of machine translation; new tools and resources for translators and interpreters; learning from practitioners in order to make speech machine translation closer to human quality; human evaluation of CAT tools, including subtitling technologies; translators and translation agencies experiences in adopting and improving the usage of CAT tools; analysis of translation technologies performance in translating wordplay.

The HiT-IT organisers would like to thank the authors for submitting their articles to the HiT-IT Workshop, the expert members of the large Programme Committee for providing exhaustive reviews, and the RANLP 2019 conference, which hosted HiT-IT 2019. We would also like to thank the other members of the organising committee: Souhila Djabri (University of Alicante, Spain), Rocío Caro (University of Wolverhampton, UK) and Encarnación Núñez (University of Malaga, Spain) for their great assistance with HIT-IT organisation.

Irina Temnikova, Constantin Orăsan, Gloria Corpas Pastor, and Ruslan Mitkov (the HIT-IT 2019 workshop chairs)

Workshop Chairs

Irina Temnikova (Bulgarian Academy of Sciences, Bulgaria)
Constantin Orasan (University of Wolverhampton, UK)
Gloria Corpas Pastor (University of Malaga, Spain)
Ruslan Mitkov (University of Wolverhampton, UK)

Members of the Organising Committee

Souhila Djabri (University of Alicante, Spain)
Rocío Caro (University of Wolverhampton, UK)
Encarnación Núñez (University of Malaga, Spain)

Program Committee:

Anja Rütten (Conference Interpreter, Member of AIIC, Germany)
Anna Zaretskaya (TransPerfect, Spain)
Bart Defrancq (Ghent University, Belgium)
Carla Parra (Unbabel, Portugal)
Claudia Angelelli (Heriot-Watt University, UK)
Claudio Bendazzoli (Università Degli Studi di Torino, Italy)
Claudio Fantinuoli (Johannes Gutenberg-Universität Mainz/Germersheim, Germany)
David Orrego-Carmona (Aston University, UK)
Dicken Minta (Televic, UK)
Dragos Ciobanu (University of Leeds, UK)
Eleanor Cornelius (University of Johannesburg, South Africa)
Eleni Zisi (EL-Translations, Greece)
Eva Dolezalova (MemSource, Czech Republic)
Federico Gaspari (ADAPT Centre, Ireland)
Filip Šanca (Memsource, Czech Republic)
Gabriela Gonzales (E-Trad, Argentina)
Haris Ghinos (ELIT Language Services, Greece)
Hendrik J. Kockaert (Hamad bin Khalifa University, Qatar)
Janice Jun Pan (Hong Kong Baptist University, Hong Kong)
Johanna Monti ("L'Orientale" University of Naples, Italy)
Joss Goldsmith (AIIC Interpreter, Geneve)
Joss Moorkens (Dublin City University, Ireland)
Juanjo Averallilo (HERMES Traducciones, Spain)
Kim Ludvigsen (Interprefy, Switzerland)
Lieve Macken (University of Ghent, Belgium)
Maja Popovic (ADAPT, DCU, Ireland)
Manuel Herrandez (Pangeanic, Spain)
Marcello Federico (Amazon, USA)
Maria Kunilovskaya (University of Wolverhampton, UK)
María Mercedes Enríquez Aranda (University of Malaga, Spain)
Maria Stambolieva (New Bulgarian University, Sofia)
Maria Stasimioti (Ionian University, Greece)
Mark Shuttleworth (Hong Kong Baptist University, Hong Kong)
Masaru Yamada (Kansai University, Japan)
Mercedes Garcia Martinez (Pangeanic, Spain)

Michael Carl (Copenhagen Business School, Denmark)
Michael Ustaszewski (Universität Innsbruck, Austria)
Mina Ilieva (Mitra Translations, Bulgaria)
Nannan Liu (The University of Hong Kong, Hong Kong)
Nieves Jiménez Carra (University of Malaga, Spain)
Omar Atari (University of Petra, Jordan)
Peter Reynolds (MemoQ, Hungary)
Pieter Demuytere (Televic, Belgium)
Pierrette Bouillon (University of Geneva, Switzerland)
Preslav Nakov (Qatar Computing Research Institute, HBKU, Qatar)
Raisa McNab (Sandberg Translation Partners Ltd, UK)
Rozane Rebechi (University Rio Grande do Sur, Brazil)
Sabrina Baldo (University of Evry Val d'Essonne, France)
Santanu Pal (Saarland University, Germany)
Sara Moze (University of Wolverhampton, UK)
Sharon O'Brien (Dublin City University, Ireland)
Sheila Castilho (ADAPT Centre, Ireland)
Silvia Bernadini (University of Bologna, Italy)
Sin-Wai Chan (Chinese University of Hong Kong, Hong Kong)
Stephen Doherty (The University of New South Wales, Australia)
Veronique Hoste (Ghent University, Belgium)
Verónica Pérez Guarnieri (Colegio de Traductores, Argentina)
Vilelmini Sosoni (Ionian University, Greece)
Yota Georgakopoulou (Athena Consultancy, Greece)
Yves Champollion (Wordfast, US/France)

Invited Speakers:

- Yves Champollion (WordFast, US/France): Talk 1: *MT Acceptance Among Translators: Are We Nearing the Tipping Point*, Talk 2: *Rosetta Stone Decipherment*

- Vilelmini Sosoni (Ionian University, Greece): *Translators and Technology: Dancing a Tango Nuevo*

- Carla Parra Escartín (Unbabel, Portugal): *The story of how academic and industrial research meet to put humans in the loop of Artificial Intelligence*

- Special presentation: Ruslan Mitkov (University of Wolverhampton, United Kingdom): *The world's first Erasmus Mundus Master's programme in technology for translation and interpreting*

Table of Contents

Research papers

Users papers