Machine Translation Summit XVII

Proceedings of Machine Translation Summit XVII
Volume 2: Translator, Project and User Tracks
https://www.mtsummit2019.com

19–23 August, 2019
Dublin, Ireland

© 2019 The authors. These articles are licensed under a Creative Commons 4.0 licence, no derivative works, attribution, CC-BY-ND.

Proceedings of Machine Translation Summit XVII
Volume 2: Translator, Project and User Tracks

https://www.mtsummit2019.com

19–23 August, 2019
Dublin, Ireland

© 2019 The authors. These articles are licensed under a Creative Commons 4.0 licence, no derivative works, attribution, CC-BY-ND.

Foreword by the President of the IAMT and the EAMT

Céad míle fáilte romhaibh!

It's a pleasure for me to warmly welcome you all to the 17th Machine Translation Summit.

Every two years, the International Association for Machine Translation (IAMT), an umbrella organization comprising the Asian Association for Machine Translation (AAMT), the Association for Machine Translation in the Americas (AMTA), and the European Association for Machine Translation (EAMT), jointly call everyone related in some way or another to machine translation and translation technologies to the most inclusive MT conference in the world, a real Summit. It brings together senior and junior researchers, developers, vendors and all kinds of users, coming from academia, industry, or even as freelancers, to share and become aware of any new developments in the field.

This is the sixth such summit held in Europe, after having visited Munich (1989), Luxembourg (1995), Santiago de Compostela (2001), Copenhagen (2007), and Nice (2013).

The organizers have assembled an excellent programme; after two days with a wide offer of tutorials and workshops, the main conference features four tracks: the research track, the users' track, a new translators' track, and the usual projects track, and includes three invited talks, poster sessions and oral sessions. Everyone will find something going on that interests them throughout the event.

Every six years, the EAMT organizes the MT Summit in Europe. The EAMT is a growing association, which organizes a yearly conference, sponsors research, development and community outreach initiatives, and annually grants a Best Thesis Award. Individuals, institutions and companies from Europe, Africa and the Middle East can join the EAMT for a modest fee and benefit from all these activities. In addition to that, EAMT members (as AMTA and AAMT members) enjoy attractive discounted fees when attending EAMT, AMTA and AAMT conferences. This is possible thanks to our members but especially to my colleagues in the EAMT Executive Committee —coming from both academia and industry— who work hard to make it all happen.

The XVII MT Summit would not be possible without the hard work of our local organizers, headed by my predecessor as EAMT president and current Executive Board member, Andy Way, who have, with the help of other MT actors from the Adapt Centre and the professional collaboration of Abbey Conference & Events, put together an excellent conference. I am very thankful for their hard work and for having put their local MT expertise at the disposal of the EAMT (and the IAMT).

Bainigí taitneamh as, that is, enjoy. Enjoy the programme, the company, and the city. Ten years ago, I lived here and worked here for a year and I'll miss it every day of my life. And I'll tell you something: it is especially the local people that makes Dublin —and all of Ireland— one of the best places in the world to hold a conference like our Summit. I'm sure you'll bring home sweet memories of it!

Baile Átha Cliath/Dublin, Lúnasa/August 2019

Mikel L. Forcada
President of the IAMT and the EAMT
Professor of Computer Languages and Systems Universitat d'Alacant
Alacant, Valencian Country, Spain.
Email: `mlf@ua.es`

Foreword by MT Summit 2019 Conference Chair

Back in 2017, on behalf of the International Association for Machine Translation (IAMT), the European Association for Machine Translation (EAMT) entrusted me with hosting this conference that you are currently attending.

While I was grateful for the trust shown in me, as a previous IAMT/EAMT president, I was acutely aware of the need to deliver; compared to our annual EAMT conferences, MT Summits provide us with the opportunity to show our Asian and American friends and colleagues that we can put on an event that all three regional associations and the IAMT can be duly proud of; if you mess up, Europe has to wait 6 years to try to put it right!

After two years of hard work, I can say with some confidence that we have achieved this. One of the first things I did was put together a very strong support team. I would really like to thank our seven co-chairs of the four tracks, namely:

- Research track co-chairs: Barry Haddow & Rico Sennrich (University of Edinburgh, UK)
- User track co-chairs: John Tinsley (Iconic Translation Machines, Ireland) & Dimitar Shterionov (ADAPT Centre, Dublin City University, Ireland)
- Translator track co-chairs: Celia Rico (Universidad Europea de Madrid, Spain) & Federico Gaspari (ADAPT Centre, Dublin City University, Ireland)
- Projects chair: Mikel L. Forcada (Universitat d'Alacant, Spain)

I am also very grateful to Laura Rossi (Lexis Nexis, The Netherlands) and Antonio Toral (University of Groningen, The Netherlands) for acting as excellent Tutorials and Workshops Chairs, respectively. I hope you all benefited from attending these pre-conference events!

For the most part, it is these 9 individuals who have put together the programme assembled before you. Each of them will comment on their Tracks later in these proceedings, but they all deserve our heartfelt thanks, as do the panels of reviewers they assembled which helped improve all our papers. From a personal point of view, I am delighted that we have – for the first but surely not the last time – included a Translator track; I have advocated for some time now that it is only through dialogue that MT developers and the translator community can advance our field. I have been very keen to take up a number of recent opportunities to speak at translator conferences, so I am especially pleased to welcome translators to this event; thank you for coming!

I am of course grateful to everyone who submitted a paper; whether your work was selected for presentation or not, if no-one had submitted, we wouldn't have had a conference. For those of you whose work was selected for presentation, many thanks for coming to Dublin, and to DCU, which have been my home and workplace now for 28 years, half my life. All of you would have interacted via EasyChair, and I am grateful to Carol Scarton, EAMT secretary, for her effort in setting up the various accounts which enabled the submission and notification processes to run so smoothly.

When you act as IAMT/EAMT president, or edit the Machine Translation journal, or act as track chair at major conferences, sometimes you have to be a bit of a pain, because you are often asking busy people to do things, mostly for free! Having been around the block a few times, I have lots of contacts in the industry, so I made myself responsible for bringing in sponsorship. I know they will say I was close to pestering them on many occasions, but I am truly grateful for the hugely generous support we obtained from our sponsors from the translation and CAT industry:

- Silver sponsors: Microsoft, and STAR
- Bronze sponsors: Pangeanic, text & form, CrossLang, Flitto, VistaTec

- Other sponsors: Welocalize, Iconic, XTM, Unbabel, DCU, ELRA, Tilde, Springer, Apertium

I am also extremely grateful to Fáilte Ireland for their generous support of this conference, which helped my ADAPT@DCU colleagues Joss Moorkens and Sharon O'Brien present our bid in Nagoya in 2017, as well as supporting our excellent invited speakers: Laura Casanellas, Helena Moniz, and Arianna Bisazza. With many women in our team, it's extremely important to have strong female role models, and we could not have asked for better from Laura, Arianna and Helena; many thanks to all of you for agreeing to share your expertise with us!

We took the decision a while back to try to be as green a conference as possible. You will already have noticed that, in order to reduce waste, there is no delegate bag. To reduce paper, we are not producing paper proceedings, and the normal programme booklet has been replaced by a smaller 'bradge' which doubles as a name badge. We are hoping to have a tree-planting ceremony during the conference in order to reduce the carbon footprint of the Summit. To reduce transport costs, we are using onsite accommodation at DCU, and will promote the use of public transport to the off-site events. Thanks to DCU Sustainability Manager Sam Fahy for her support in these efforts.

While we decided not to produce printed proceedings, they still needed to be put together in electronic form. I am grateful to Jenny Walsh for producing such an excellent conference logo, but huge thanks are due to Alberto Poncelas for putting together the proceedings, and for helping workshop chairs to produce theirs. Alberto has also liaised with Matt Post to ensure that your papers are indexed in perpetuity on the ACL Anthology!

I have two final people to thank. Firstly, I am very grateful to Grainne McQuaid and her team in Abbey Conference and Events for their professional support of the conference. You will have met them at registration, and they are available throughout the event to ensure your needs are met. We have been engaging with them for 2 years now, and this has been a true partnership that has made this journey an enjoyable one. Secondly, I am especially grateful to my colleague Jane Dunne, for managing the planning of the conference, and for managing me too. Jane has done this over and above her work on a European project, and I could not have chosen a better, nicer person to engage with over these past two years – thank you Jane on behalf of everyone; we are all deeply grateful for your huge effort in getting us to where we all are today!

Finally, I really hope that you all enjoy the conference, that you benefit from the excellent programme that has been assembled, and that you go away from here having made new friends. I am fortunate indeed that many of my very best friends are in the MT community, and I hope to meet as many of you as possible during the event.

Andy Way
Chair, MT Summit 2019
Deputy Director ADAPT Centre School of Computing Dublin City University
Dublin, Ireland.
Email: `Andy.Way@adaptcentre.ie`

Foreword by the Translator Track Program Chairs

For the first time ever, the call for papers of the 2019 edition of the MT Summit included a specific track aimed at translators, who are arguably the largest group of users of translation technologies. This exciting development built on the success of the 21st annual EAMT Conference, that was held in Alicante, Spain, roughly a year before, and featured a dedicated translator track which added an important new dimension to the multi-faceted contemporary debate on MT. The novel translator track in this year's MT Summit programme aimed at involving translators into the conversation on an equal footing with researchers, developers, vendors and users of translation technologies, as the co-chairs felt that it was high time for their voice to be heard loud and clear. With this objective in mind, specific topics of interest included in the call for papers for the translator track concerned issues that increasingly confront language and translation professionals on a daily basis, including productivity measurements and their impact on MT adoption, the role of MT in translators' work (pricing issues, post-editing tasks assignment and their acceptance among professionals), ethical and confidentiality considerations when using MT, psycho-social aspects of MT adoption, such as attitudes and (pre-)conceptions, etc. A total of 23 submissions were received for the translator track, and each of them was peer-reviewed by three independent members of the Programme Committee. After a thorough selection process, 15 papers were accepted (6 for the oral sessions and 9 for poster presentation). The Programme Committee included both academics and practitioners, also representing associations and bodies of professional translators, to reflect the key communities with a specific interest in the topics addressed by the call for papers. We were very pleased with the number and quality of the submissions to the track at its debut at the MT Summit, and were particularly delighted to receive the official support of the International Federation of Translators (FIT) for the conference, as we are convinced that this mutual recognition is essential to ensure collaboration and an open, honest debate on MT going forward. We are very grateful to the Programme Committee members for their high-quality reviews, which have been very useful to us as co-chairs to select the paper proposals that were accepted in the conference programme and to the relevant authors to improve their papers: Bogdan Babych, Valeria Barbero, Sarah Bawa-Mason, Katie Botkin, Oliver Czulo, Stephen Doherty, Ignacio Garcia, Luis González, Ana Guerberof Arenas, Adrià Martín-Mor, Joss Moorkens, Sharon O'Brien, Minako O'Hagan, Mary Phelan, Pilar Sánchez-Gijón, Mirko Silvestrini, Olga Torres-Hostench. We would also like to thank all the authors for trying their best to incorporate the reviewers' suggestions when preparing the final versions of their accepted papers that have been included in these proceedings. Finally, as regards the papers submitted to the translator track that were not accepted for the conference, we hope that the reviewers' constructive comments will be useful to improve them, so that the important and timely debate on the relationship between translators, MT and more broadly translation technologies can continue in subsequent editions of the MT Summit and at other similar conferences in the future.

Celia Rico and Federico Gaspari

Foreword by the User Track Program Chairs

Firstly, we want to acknowledge what an honour it has been to serve as co-chairs of the user track for the Machine Translation Summit. Between us, in various guises, we have attended, submitted to, and presented at this conference over its history. To be involved on the other side of the fence this time has been a great experience.

Continuing with a trend started at the last MT Summit in 2017, we solicited full papers for the user track, as opposed to abstracts, so that the work might leave a longer legacy and impact beyond what is presented at the conference itself. We also requested that all submissions have at least one industry partner to encourage true user scenarios, as opposed to simulations.

The result was 24 submissions for the user track, representing 22 companies and 12 academic institutions. Ten of the submissions came from collaboration between industry and academic partners, while the rest represented applied research, applications, and use cases directly from the commercial field. There were 80 distinct authors with on average 4 authors per submission.

Geographically, the submissions originated from 12 countries from Europe, and both North and South America: Ireland (8), Spain (6), Switzerland (5), Belgium (2), Latvia (2), United States (2), Argentina (1), Austria (1), Brazil (1), Italy (1), Portugal (1), United Kingdom (1).

The 24 submissions were reviewed by an elaborate programme committee of 38 experts in the fields of localisation, machine translation and natural language processing from industry and academia. Each submission received at least 3 reviews and in some cases even 4 and 5. In these latter cases, the additional reviews helped us make more definitive decisions on borderline cases.

In total, we accepted 14 submissions for publication and presentation at the conference, resulting in an acceptance rate of 58

The range of topics covered overall reflects the fact that, now more so than ever, machine translation is in wide commercial use, with a range of applications and stories from organisations of all shapes and sizes, including big tech, service providers, SMEs, government organisations and more. Neural approaches dominate in most of the workflows, whether it is for machine translation, quality estimation, post-editing or pre- and post-processing. Great amount of research also aims at optimised performance and quality by better exploiting of data, backtranslation and reuse of translation output.

Finally, we would like to thank the authors of all submissions, our diligent programme committee, and of course the conference organisers. We hope you enjoy the proceedings.

John Tinsley and Dimitar Shterionov

Organizers

President of the IAMT and the EAMT

Mikel Forcada	University of Alicante

MT Summit 2019 Conference Chair

Andy Way	ADAPT Centre

Programme Chairs

Research track co-chairs

Barry Haddow	The University of Edinburgh
Rico Sennrich	The University of Edinburgh

User track co-chairs

John Tinsley	Iconic Translation Machines
Dimitar Shterionov	Dublin City University

Translator track co-chairs

Celia Rico	Universidad Europea de Madrid
Federico Gaspari	Dublin City University

Projects track co-chair

Mikel L. Forcada	University of Alicante

Program Committee

Research Track

Alberto Poncelas	DCU
Aizhan Imankulova	Tokyo Metropolitan University
Aleš Tamchyna	Memsource a. s.
Alon Lavie	Carnegie Mellon University
Anabela Barreiro	INESC-ID
Andreas Maletti	Universität Leipzig
Andrei Popescu-Belis	HEIG-VD / HES-SO
Ankur Bapna	Google
Ann Clifton	Simon Fraser University
Annette Rios Gonzales	University of Zurich
Anoop Kunchukuttan	IIT Bombay
Antonio Toral	University of Groningen
Antonio Valerio Miceli Barone	University of Pisa
Arturo Oncevay	The University of Edinburgh
Arul Menezes	Microsoft
Arya McCarthy	Johns Hopkins University
Atsushi Fujita	National Institute of Information and Communications Technology

Barry Haddow	The University of Edinburgh
Boxing Chen	Alibaba Group
Carla Parra Escartín	Unbabel
Carolina Scarton	European Association for Machine Translation
Celia Rico	Universidad Europea de Madrid
Chenhui Chu	Osaka University
Chris Brockett	Microsoft
Christian Dugast	tech2biz
Christian Federmann	Microsoft
Christian Hardmeier	Uppsala University
Christoph Tillmann	IBM
Christos Baziotis	The University of Edinburgh
Colin Cherry	National Research Council Canada
Constantin Orasan	University of Wolverhampton
Cristina España-Bonet	UdS and DFKI
Dakun Zhang	Systran
Daniel Marcu	ISI/USC
Daniel Ortiz-Martínez	Unversitat Politecnica de Valencia
Dario Stojanovski	Ludwig Maximilian University of Munich
David Vilar	Amazon
Devendra Singh Sachan	CMU
Deyi Xiong	Tianjin University
Dimitar Shterionov	Dublin City University
Dušan Variš	Institute of Formal and Applied Linguistics; Charles University in Prague
Duygu Ataman	University of Trento; Fondazione Bruno Kessler; University of Edinburgh
Ekaterina Lapshinova-Koltunski	Saarland University
Ekaterina Vylomova	The University of Melbourne
Eleftherios Avramidis	German Research Center for Artificial Intelligence (DFKI)
Eva Hasler	The University of Edinburgh
Eva Vanmassenhove	DCU
Fatiha Sadat	UQAM
Federico Gaspari	Dublin City University
Felipe Sánchez-Martínez	Dep. de Llenguatges i Sistemes Informàtics. Universitat d'Alacant
Felix Hieber	Amazon
Felix Stahlberg	University of Cambridge
Ferhan Ture	Comcast Labs
Francis Tyers	Indiana University Bloomington
Francisco Casacuberta	Universitat Politècnica de València
Francisco Javier Guzman	Facebook
François Yvon	LIMSI/CNRS et Université Paris-Sud
Frederic Blain	The University of Sheffield
George Foster	NRC
Gholamreza Haffari	Simon Fraser University
Gonzalo Iglesias Iglesias	SDL
Graham Neubig	Carnegie Mellon University
Gregor Leusch	eBay inc
Hainan Xu	Johns Hopkins University

Helena Caseli	Federal University of São Carlos (UFSCar)
Helmut Schmid	Ludwig Maximilian University of Munich
Hiroshi Echizenya	Hokkai-Gakuen University
Houda Bouamor	Carnegie Mellon University
Huda Khayrallah	Johns Hopkins University
Iacer Calixto	University of Amsterdam
Isao Goto	NHK
Jan Niehues	Maastricht University
Jesús González-Rubio	WebInterpret
Jiajun Zhang	Institute of Automation Chinese Academy of Sciences
Joachim Daiber	Apple Inc.
John Henderson	The MITRE Corporation
John Tinsley	Iconic Translation
Joke Daems	Ghent University
Jonathan Mallinson	The University of Edinburgh
Joost Bastings	University of Amsterdam
Jörg Tiedemann	University of Helsinki
José G. C. de Souza	ebay
Josep Crego	SYSTRAN
Joss Moorkens	ADAPT Centre
Juan Antonio Pérez-Ortiz	Universitat d'Alacant; Departament de Llenguatges i Sistemes Informàtics
Julia Ive	King's College London
Julia Kreutzer	Heidelberg University
Julian Schamper	DeepL GmbH
Katsuhito Sudoh	Nara Institute of Science and Technology
Ke Hu	ADAPT Centre; Dublin City University
Kenji Imamura	National Institute of Information and Communications Technology
Kenton Murray	Carnegie Mellon University School of Computer Science
Laura Jehl	Institut für Computerlinguistik; Universität Heidelberg
Laurent Besacier	Laboratoire d'Informatique de Grenoble
Lemao Liu	NICT
Lexi Birch	The University of Edinburgh
Lieve Macken	Ghent University
Linfeng Song	University of Rochester
Luisa Bentivogli	FBK
Maja Popovic	ADAPT Centre; DCU
Manny Rayner	University of Geneva
Mara Chinea Rios	Universitat Politècnica de València
Marc Dymetman	Xerox Research Centre Europe
Marcello Federico	Amazon AI
Marcin Junczys-Dowmunt	Microsoft
Marco Turchi	Fondazione Bruno Kessler
Marianna Apidianaki	CNRS
Marija Brkic	Department of Informatics; University of Rijeka
Marion Weller-Di Marco	University of Amsterdam
Mark Fishel	University of Tartu

Markus Freitag	Google AI
Marta R. Costa-Jussà	Institute For Infocomm Research
Martin Popel	UFAL; Faculty of Mathematics and Physics; Charles University
Martin Volk	University of Zurich
Masaaki Nagata	NTT
Masao Utiyama	NICT
Mathias Müller	University of Zurich
Matīss Rikters	Tilde
Matt Post	Johns Hopkins University
Matteo Negri	Fondazione Bruno Kessler (FBK-irst)
Matthias Huck	Ludwig Maximilian University of Munich
Mattia Antonino Di Gangi	Fondazione Bruno Kessler; University of Trento
Maximiliana Behnke	The University of Edinburgh
Mercedes García-Martínez	Pangeanic SL
Meriem Beloucif	The Hong Kong University of Science and Technology
Michael Carl	Kent State University
Michel Simard	National Research Council Canada (NRC)
Miguel Domingo	Universitat Politècnica de València
Mihael Arcan	Insight Centre for Data Analytics; National University of Ireland Galway
Mihaela Vela	Universität des Saarlandes
Miloš Stanojević	The University of Edinburgh
Miquel Esplà	Universitat d'Alacant
Mireia Farrús	Universitat Pompeu Fabra
Mirjam S. Maučec	FERI; University of Maribor
Myle Ott	Facebook
Nicola Ueffing	eBay
Nikolay Bogoychev	The University of Edinburgh
Niyu Ge	
Nizar Habash	Columbia University
Núria Bel	Universitat Pompeu Fabra
Orhan Firat	Google
Ozan Çağlayan	Le Mans University
Parnia Bahar	RWTH Aachen University
Patrick Simianer	Lilt.
Paul Michel	Carnegie Mellon University - LTI
Pavel Pecina	Charles University In Prague
Petra Barančíková	Charles University in Prague
Philip Williams	The University of Edinburgh
Philipp Koehn	Johns Hopkins University
Praveen Dakwale	Informatics Institute; University of Amsterdam
Preethi Raghavan	IBM TJ Watson Research
Qun Liu	Huawei Noah's Ark Lab
Rabih Zbib	Raytheon
Rachel Bawden	The University of Edinburgh
Raj Dabre	IIT Bombay
Rajen Chatterjee	Apple Inc.
Rebecca Knowles	Johns Hopkins University
Rebecca Marvin	Johns Hopkins University

Rico Sennrich	The University of Edinburgh
Roland Kuhn	National Research Council of Canada
Roman Grundkiewicz	The University of Edinburgh; School of Informatics
Rudolf Rosa	Charles University
Saab Mansour	RWTH Aachen University
Sadao Kurohashi	Kyoto University
Sameen Maruf	Monash University
Sameer Bansal	The University of Edinburgh
Samuel Läubli	University of Zurich
Sarah Ebling	University of Zurich
Sergio Penkale	Lingo24
Shahram Khadivi	eBay
Shankar Kumar	Google
Sharon O'Brien	Dublin City University
Sheila Castilho	Dublin City University/ADAPT Centre
Stephan Peitz	Apple
Surafel Melaku Lakew	University of Trento
Tamer Alkhouli	RWTH Aachen University
Taro Watanabe	NICT
Teresa Herrmann	Fujitsu
Tim Anderson	Wright-Patterson Air Force Research Laboratory
Tomáš Musil	Charles University in Prague
Toshiaki Nakazawa	The University of Tokyo
Tsuyoshi Okita	Kyushu Institute of Technology
Ulrich Germann	The University of Edinburgh
Víctor M. Sánchez-Cartagena	Transducens Research Group; Departament de Llenguatges i Sistemes Informàtics; Universitat d'Alacant
Viktor Hangya	Ludwig Maximilian University of Munich
Vincent Vandeghinste	Instituut voor de Nederlandse Taal, Centre for Computational Linguistics, KU Leuven
Vishal Chowdhary	Microsoft
Vu Hoang	The University of Melbourne
Wei Wang	Google
Xing Niu	University of Maryland
Yinfei Yang	Redfin Inc.
Yuki Arase	Osaka University
Yunsu Kim	RWTH Aachen University
Yvette Graham	Dublin City University

Translator Track

Adrià Martín-Mor	UAB
Alberto Poncelas	DCU
Ana Guerberof Arenas	DCU/ADAPT Centre
Bogdan Babych	University of Leeds
Carolina Scarton	European Association for Machine Translation
Celia Rico Perez	Universidad Europea de Madrid
Federico Gaspari	Dublin City University
Ignacio Garcia	University of Western Sydney
Joss Moorkens	Dublin City University

Katie Botkin	MultiLingual magazine
Mary Phelan	Dublin City University
Minako O'Hagan	The University of Auckland
Mirko Silvestrini	UNILINGUE
Olga Torres-Hostench	Universitat Autònoma de Barcelona
Oliver Czulo	Universität Leipzig
Pilar Sanchez-Gijón	Autonomous University of Barcelona
Sarah Bawa-Mason	University of Portsmouth/Institute of Translation and Interpreting
Sharon O'Brien	Dublin City University
Stephen Doherty	The University of New South Wales
Valeria Barbero	MT Summit

User Track

Alberto Poncelas	DCU
Aljoscha Burchardt	DFKI
Alon Lavie	Carnegie Mellon University
Bram Bulté	Katholieke Universiteit Leuven
Bruno Pouliquen	World Intellectual Property Organization
Carlos Collantes Fraile	TransPerfect / Universidad Complutense de Madrid
Carmen Heger	SZ
Carolina Scarton	European Association for Machine Translation
Chao-Hong Liu	ADAPT Centre, Dublin City University
Charlotte Tesselaar	LexisNexis Univentio
Chris Wendt	Microsoft
Christian Federmann	Microsoft
Christian Eisold	berns language consulting GmbH
Dag Schmidtke	Microsoft
Daniel Stein	eBay Inc.
David Vilar	Amazon
Declan Groves	Microsoft
Dimitar Shterionov	Dublin City University
Eva Martínez Garcia	Vicomtech / Universitat Politècnica de Catalunya
Eva Vanmassenhove	DCU
Evgeny Matusov	eBay
Félix Do Carmo	ADAPT Centre
Fred Blain	The University of Sheffield
Gema Ramírez-Sánchez	Prompsit Language Engineering, S.L.
Guodong Xie	ADAPT Centre, Dublin City University
Heidi Depraetere	Cross Language NV
Jean Senellart	SYSTRAN
John Tinsley	Iconic Translation Machines
José G. C. de Souza	eBay Inc.
Keith J. Miller	The MITRE Corporation
Kim Harris	text&form GmbH
Laurent Chevrette	Mondzo
Maxim Khalilov	Unbabel
Mercedes García-Martínez	Pangeanic SL
Nathalie DeSutter	Untranslate
Nicola Ueffing	eBay

Olga Beregovaya	Welocalize
Patrik Lambert	Pompeu Fabra University
Phil Ritchie	Vistattec
Raj Pate	CDAC
Rohit Gupta	Iconic Translations Machines Ltd.
Saša Hasan	Apple
Silvio Picinini	eBay
Steve Richardson	The Church of Jesus Christ of Latter-day Saints
Thierry Etchegoyhen	Vicomtech-IK4
Tony O'Dowd	Xcelerator Machine Translations Ltd.
Yury Sharshov	LexisNexis Univentio

Project Track

Alberto Poncelas	DCU
Carolina Scarton	The University of Sheffield
Mikel Forcada	University of Alicante

Contents

Competitiveness Analysis of the European Machine Translation Market **1**
Andrejs Vasiļjevs, Inguna Skadiņa, Indra Sāmīte, Kaspars Kauliņš, Ēriks Ajausks, Jūlija Meļņika and Aivars Bērziņš

Improving CAT Tools in the Translation Workflow: New Approaches and Evaluation **8**
Mihaela Vela, Santanu Pal, Marcos Zampieri, Sudip Naskar and Josef van Genabith

Hungarian translators' perceptions of Neural Machine Translation in the European Commission **16**
Ágnes Lesznyák

Applying Machine Translation to Psychology: Automatic Translation of Personality Adjectives **23**
Ritsuko Iwai, Daisuke Kawahara, Takatsune Kumada and Sadao Kurohashi

Evaluating machine translation in a low-resource language combination: Spanish-Galician. **30**
María Do Campo Bayón and Pilar Sánchez-Gijón

MTPE in Patents: A Successful Business Story **36**
Valeria Premoli, Elena Murgolo and Diego Cresceri

User expectations towards machine translation: A case study **42**
Barbara Heinisch and Vesna Lušicky

Does NMT make a difference when post-editing closely related languages? The case of Spanish-Catalan **49**
Sergi Alvarez, Antoni Oliver and Toni Badia

Machine Translation in the Financial Services Industry: A Case Study **57**
Mara Nunziatini

Pre-editing Plus Neural Machine Translation for Subtitling: Effective Pre-editing Rules for Subtitling of TED Talks **64**
Yusuke Hiraoka and Masaru Yamada

Do translator trainees trust machine translation? An experiment on post-editing and revision **73**
Randy Scansani, Silvia Bernardini, Adriano Ferraresi and Luisa Bentivogli

On reducing translation shifts in translations intended for MT evaluation **80**
Maja Popovic

Comparative Analysis of Errors in MT Output and Computer-assisted Translation: Effect of the Human Factor **88**
Irina Ovchinnikova and Daria Morozova

A Comparative Study of English-Chinese Translations of Court Texts by Machine and Human Translators and the Word2Vec Based Similarity Measure's Ability To Gauge Human Evaluation Biases **95**
Ming Qian, Jessie Liu, Chaofeng Li and Liming Pals

Translating Terminologies: A Comparative Examination of NMT and PBSMT Systems **101**
Long-Huei Chen and Kyo Kageura

NEC TM DATA PROJECT **109**
Alexandre Helle and Manuel Herranz

APE-QUEST **110**
Joachim Van den Bogaert, Heidi Depraetere, Sara Szoc, Tom Vanallemeersch, Koen Van Winckel, Frederic Everaert, Lucia Specia, Julia Ive, Maxim Khalilov, Christine Maroti, Eduardo Farah, Artur Ventura

PRINCIPLE: Providing Resources in Irish, Norwegian, Croatian and Icelandic for the Purposes of Language Engineering **112**
Andy Way and Federico Gaspari

iADAATPA Project: Pangeanic use cases **114**
Mercedes García-Martínez, Amando Estela, Laurent Bié, Alexandre Helle and Manuel Herranz

MICE **116**
Joachim Van den Bogaert, Heidi Depraetere, Tom Vanallemeersch, Frederic Everaert, Koen Van Winckel, Katri Tammsaar, Ingmar Vali, Tambet Artma, Piret Saartee, Laura Katariina Teder, Artūrs Vasiļevskis, Valters Sics, Johan Haelterman and David Bienfait

ParaCrawl: Web-scale parallel corpora for the languages of the EU **118**
Miquel Esplà, Mikel Forcada, Gema Ramírez-Sánchez and Hieu Hoang

Pivot Machine Translation in INTERACT Project **120**
Chao-Hong Liu, Andy Way, Catarina Silva and André Martins

Global Under-Resourced Media Translation (GoURMET) **122**
Alexandra Birch, Barry Haddow, Ivan Tito, Antonio Valerio Miceli Barone, Rachel Bawden, Felipe Sánchez-Martínez, Mikel L. Forcada, Miquel Esplà-Gomis, Víctor Sánchez-Cartagena, Juan Antonio Pérez-Ortiz, Wilker Aziz, Andrew Secker, Peggy van der Kreeft

Neural machine translation system for the Kazakh language **123**
Ualsher Tukeyev and Zhandos Zhumanov

Leveraging Rule-Based Machine Translation Knowledge for Under-Resourced Neural Machine Translation Models **125**
Daniel Torregrosa, Nivranshu Pasricha, Maraim Masoud, Bharathi Raja Chakravarthi, Juan Alonso, Noe Casas and Mihael Arcan

Bootstrapping a Natural Language Interface to a Cyber Security Event Collection System using a Hybrid Translation Approach **134**
Johann Roturier, Brian Schlatter and David Silva Schlatter

Improving Robustness in Real-World Neural Machine Translation Engines **142**
Rohit Gupta, Patrik Lambert, Raj Patel and John Tinsley

Surveying the potential of using speech technologies for post-editing purposes in the context of international organizations: What do professional translators think? **149**
Jeevanthi Liyanapathirana, Pierrette Bouillon and Bartolomé Mesa-Lao

Automatic Translation for Software with Safe Velocity **159**
Dag Schmidtke and Declan Groves

Application of Post-Edited Machine Translation in Fashion eCommerce **167**
Kasia Kosmaczewska and Matt Train

Morphological Neural Pre- and Post-Processing for Slavic Languages **174**
Giorgio Bernardinello

Large-scale Machine Translation Evaluation of the iADAATPA Project **179**
Sheila Castilho, Natália Resende, Federico Gaspari, Andy Way, Tony O'Dowd, Marek Mazur, Manuel Herranz, Alex Helle, Gema Ramírez-Sánchez, Víctor Sánchez-Cartagena, Mārcis Pinnis and Valters Šics

Collecting domain specific data for MT: an evaluation of the ParaCrawl pipeline **186**
Arne Defauw, Tom Vanallemeersch, Sara Szoc, Frederic Everaert, Koen Van Winckel, Kim Scholte, Joris Brabers and Joachim Van den Bogaert

Monolingual backtranslation in a medical speech translation system for diagnostic interviews - a NMT approach **196**
Jonathan Mutal, Pierrette Bouillon, Johanna Gerlach, Paula Estrella and Hervé Spechbach

Improving Domain Adaptation for Machine Translation with Translation Pieces **204**
Catarina Silva

Raising the TM Threshold in Neural MT Post-Editing: a Case Study on Two Datasets **213**
Anna Zaretskaya

Incremental Adaptation of NMT for Professional Post-editors: A User Study **219**
Miguel Domingo, Mercedes García-Martínez, Álvaro Peris, Alexandre Helle, Amando Estela, Laurent Bié, Francisco Casacuberta and Manuel Herranz

When less is more in Neural Quality Estimation of Machine Translation. An industry case study **228**
Dimitar Shterionov, Félix Do Carmo, Joss Moorkens, Eric Paquin, Dag Schmidtke, Declan Groves and Andy Way