

MTSummit 2025



MT SUMMIT

Geneva 2025

Machine Translation Summit XX

Volume 1

Edited by:

Pierrette Bouillon, Johanna Gerlach, Sabrina Girletti, Lise Volkart, Raphael Rubino, Rico Sennrich,
Ana C. Farinha, Marco Gaido, Joke Daems, Dorothy Kenny, Helena Moniz, Sara Szoc



**UNIVERSITÉ
DE GENÈVE**
FACULTÉ DE TRADUCTION
ET D'INTERPRÉTATION



EUROPEAN
ASSOCIATION
FOR **M**ACHINE
TRANSLATION

June 23-27, 2025

Geneva, Switzerland



The papers published in this proceedings are —unless indicated otherwise— covered by the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC-BY-NC ND 4.0). You may copy, distribute, and transmit the work, provided that you attribute it (authorship, proceedings, publisher) in the manner specified by the author(s) or licensor(s), and that you do not use it for commercial purposes. The full text of the licence may be found at:

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>

©2025 The authors

ISBN 978-2-9701897-0-1

Foreword from the General Chair

As president of the International Association for Machine Translation (IAMT) and General Chair of the 20th Machine Translation Summit, it is my utmost pleasure to write these opening words. Be most welcome to our MT Summit 2025!

The European Association for Machine Translation (EAMT) Executive Committee (EC) has been very busy. Mikel Forcada (treasurer) and Sara Szoc (secretary) have been tirelessly supporting all initiatives. Carolina Scarton and Sara Szoc took great care of our bursaries. Patrick Cadwell, André Martins, and Manuel Lardelli were our chairs for the Research Projects. Manuel Lardelli was also our policies chair, revisiting all our policies and contributing to inclusivity strategies. Our very own Mary Nurminen, chair of the bid proposals for our next events, has been busy selecting our next venue! EAMT 2026 venue will be disclosed in our closing ceremony in Geneva!

One of our core initiatives, the best thesis award – Rachel Badwen and Barry Haddow, chairs of the Best Thesis Award, had a very difficult time selecting a candidate, since the submissions were of very high quality. Our congratulations to Ricardo Rei’s thesis “Robust, Interpretable and Efficient MT Evaluation with Fine-tuned Metrics” (Unbabel, INESC-ID, Instituto Superior Técnico, Portugal), supervised by Maria Luísa Torres Ribeiro Marques da Silva Coheur and Alon Lavie. We would also like to congratulate for the highly commended thesis of Sara Papi (University of Trento & Fondazione Bruno Kessler), entitled “Direct Speech Translation in Constrained Contexts: the Simultaneous and Subtitling Scenarios” and supervised by Marco Turchi and Matteo Negri.

EAMT, as full sponsor of the MT Marathon, would also like to thank the Institute of Formal and Applied Linguistics (ÚFAL), Charles University for organizing the 17th MT Marathon. The event included MT lectures and labs, covering the basics and tutorials; keynote talks from experienced researchers and practitioners; presentations of research and open source tools related to MT; and hacking projects to advance tools or research in one week or start new collaborations. A special thank you to Jindřich Helcl his commitment and passion for this event!

MT Summit 2025 will be a moment to celebrate our IAMT Award of Honour!¹ We celebrate Professor Mikel Forcada, unanimously supported by all sister organizations (EAMT, AAMT, and AMTA), in recognition of his long-standing distinguished contribution to the EAMT and IAMT communities and for his impactful research on Machine Translation. Thank you for being an inspiration to us all!

Geneva, Switzerland, MT Summit 2025! Our conference will have a three-day, four-track programme put together by our chairs: Catarina Farinha and Marco Gaido (research: technical track chairs); Dorothy Kenny and Joke Adaems (research: translators & users track chairs); Samuel Lübbli and Martin Volk (implementations & case studies track chairs); Miguel Esplà and Vincent Vandeghinste (products & projects track chairs) and François Yvon and Sheila Castilho (workshop and tutorial chairs). Our filters of quality and alignment! We really appreciate your work. We will continue with our tradition and also have a two-day workshops and tutorials event.

Our gratitude to all our keynotes speakers. Sarah Ebling, Full Professor of Language, Technology and Accessibility at the University of Zurich. Joss Moorkens, Associate Professor at the School of Applied Language and Intercultural Studies in Dublin City University (DCU). Eva Vanmassenhove, Assistant professor in the Department of Cognitive Science and Artificial Intelligence at Tilburg University (TiU). Our outstanding keynote speakers will demonstrate their extensive and global impactful work in translation studies and translation technologies, in a multidisciplinary motto which is the core of our community.

¹<https://eamt.org/iamt-award-of-honour/>

MT Summit 2025 is the result of a very aligned, sharp, engaged, and hard working local organising team! What a diligent team! Our local co-chairs, Pierrette Bouillon, Johanna Gerlach, Sabrina Girletti and Lise Volkart (all from the University of Geneva, Switzerland) have put a lot of work in giving us a Geneva unforgettable event. To Sevita Caseres, Bastien David, Céline De Graaf, Julie Humbert-Droz, Rebeka Mali, Lucía Morado, Jonathan Mutal, Lucía Ormaechea, Aurélie Picton, Donatella Pulitano, Silvia Rodríguez, Raphael Rubino, Valentin Scourneau, Marianne Starlander, Irene Strasly, Nikolaos Tsourakis, Florine Voisard (all from the University of Geneva, Switzerland) and Rico Sennrich (University of Zurich, Switzerland), our deepest appreciation.

EAMT has been supported by generous sponsors in its initiatives along the years.² This year is no exception in a summit year! In fact, it is a very exceptional year in terms of sponsoring activities. Our gratitude to our Platinum sponsors who will also be giving a research oral presentation, BIG Language Solution, STAR, WIPO. Our Gold sponsor Systran by ChapsVision. Our Silver sponsors: Translated, Reverso, and Unbabel. To our Bronze sponsors: AppTek, CrossLang, TransPerfect, and Zoo Digital. To all our Supporter sponsors: Apertium, iguanodon.ai, prompsit, Springer Nature (our Supporter sponsor for the Best Paper award) and Supertext. Finally, to our Media sponsors, MultiLingual and Slator. Your support is vital in our efforts to give back to our community through grants and other initiatives.

A note still to all our IAMT members and our participants! Without you no effort would make sense! Let us take this opportunity to create scientific collaboration and give constructive feedback. To fully enjoy the conference, please check our Code of Conduct.³ I'm looking forward to seeing you all and celebrating our community gathering!

Our sister organizations have been renewed with new board of Directors. The best wishes to AMTA's new board, represented by the President, Jay Marciano, and to the AAMT's Directors, Hisahiro Adachi, SunFlare Co., Ltd. (President of AAMT) and Masao Utiyama, National Institute of Information and Communications Technology, Japan (Vice President of AAMT). MT Summit 2027 will be held by AMTA! More soon!

It is our organisation's greatest wish to continue giving back to our community and to drive and be driven by our community's energy and enthusiasm. Reach out to us if you have new ideas or suggestions you would like to implement. We will try hard to accomplish it with you. Learn more about us.

Helena Moniz

President of the IAMT
General Chair of MT Summit 2025
University of Lisbon, Portugal

²<https://mtsummit2025.unige.ch/sponsors.html>

³<https://mtsummit2025.unige.ch/about.html#codeOfConduct>

Message from the Local Organising Committee

It is our great pleasure to welcome you to the Faculty of Translation and Interpreting (FTI) for this 20th edition of the MT Summit. We are particularly proud that for the first time in its history, the Summit is being hosted by a translation faculty, highlighting the importance of the human factor in today's technology. This is also a sign that technology has become an imperative in professional translation. Our faculty has long embraced this evolution, as illustrated by its translation technology department, first established back in the 1970s (first under the name of ISSCO, and then TIM). It was long spearheaded by Prof Maghi King, who, as some of you may recall, received the prestigious IAMT Award of Honour in 2005.

Our department has always been committed to building bridges between research in MT and professional translators. The conference taking place here today is further proof that this bridge is now well established and solid! The structure of the conference itself also reflects this dual focus, with two dedicated research tracks, one Technical, and the other for Translators and Users.

This year also brings an important new initiative: authors of papers involving computational experiments are encouraged to include sustainability reports. Most authors engaged with the initiative, reflecting the willingness of our community to embrace more transparent and thoughtful research practices.

We hope you will enjoy the rich and carefully curated program put together by our dedicated track chairs and made possible by the thorough work of our reviewers. We are also deeply grateful to our keynote speakers, as well as the organizers of the workshops and tutorials, whose contributions are crucial to the success of this conference.

We also want to thank our sponsors, more generous than ever before! Their presence is a strong indicator of the fruitful and trustworthy collaboration that exists between academia and industry in our field.

When we signed up to organise this conference, we had no idea of the summit that we would have to climb, nor how much determination, patience and endurance it would require of us. But thanks to our experience of the mountains, a dedicated team, and the valuable support of EAMT Executive Committee and previous organisers, we reached the (MT) Summit (almost) without problems. As in every climb, it is the strength of the team that gets you to the top!

We wish you an excellent MT Summit!

On behalf of the MT Summit 2025 Organising Committee:

Pierrette Bouillon

Johanna Gerlach

Sabrina Girletti

Lise Volkart

Department of Translation Technology (TIM)

Faculty of Translation and Interpreting

University of Geneva, Switzerland

Preface by the Programme Chairs

The **Research Technical track** received 57 submissions, out of which 28 were accepted, for an acceptance rate of 49%. 14 papers will be presented orally and the other 14 will be part of two poster sessions. The topics covered by the submitted papers include named entity aware translation, context-aware machine translation, domain-specific translation, multilingual and low-resource translation, and translation evaluation. We express our most heartfelt thanks to the 83 reviewers, who made this track possible, with a particular gratitude for the emergency reviewers who promptly accomplished their duties, enabling us to respect the timeline for author notification.

Catarina Farinha (Unbabel)

Marco Gaido (Fondazione Bruno Kessler, Italy)

The **Translators and Users track** initially received 28 submissions, of which 21 could be considered for this track, the other 7 covered more technical aspects of machine translation and were therefore considered for the Technical track instead. Of these 21, 19 were accepted (an acceptance rate of 90%, showing the overall high quality of submission to the track). As track chairs, we noticed a few trends in these accepted papers, and we tried to group the submissions in sessions accordingly. The large language model trend, established in earlier EAMT conferences, clearly continues. Large language models are used for literary translation (post-editing) and emergency response text translation, and there is a clear interest in how these technologies are currently being used by students as well as perceived by professionals. From the text types that are being studied, it is obvious that 'literary translation' is most strongly represented in this track, with 5 submissions covering the topic. This is particularly striking, given that this MT Summit is also hosting a dedicated workshop on Creative-text Translation and Technology. The intersection of creativity, literature and automatic translation has clearly arrived as a field of inquiry. We thank all PC members for their time and dedication in delivering insightful feedback, ensuring the quality of the submissions to this track. Special thanks to the emergency reviewers who helped us avoid any delays. You all made this conference possible.

Joke Daems (Ghent University, Belgium)

Dorothy Kenny (Dublin City University, Ireland)

The **Implementation and Case Studies track** received 12 submissions out of which 9 were accepted for presentation at the MT summit (6 talks and 3 posters). The papers cover a broad range of topics, e.g. speech translation, LLM-based translation, low-resource settings, productivity evaluation and translator satisfaction. We would like to express our gratitude and appreciation to our reviewers from academia and industry for their time and effort in commenting and grading the submissions.

Samuel Läubli (Textshuttle/Supertext, Switzerland)

Martin Volk (University of Zurich, Switzerland)

The **Products and Projects track** received 22 submissions, of which 20 have been accepted for a short, two-page description and a poster presentation at the conference. Our selection highlights a diverse range of products and projects created by our community, covering research projects and cutting-edge services and innovations from distinguished industry and research leaders. Expect a lively session filled with poster boosters and engaging poster presentations. We wish to thank the 26 members of the program committee for this track for their timely and thorough reviews.

Miquel Esplà-Gomis (University of Alicante, Spain)
Vincent Vandeghinste (KU Leuven, Belgium)

The **Workshop and Tutorials** received seven workshop proposals, five of which were finally selected: four are reiterations of workshops that have already been collocated with MT conferences in the past: these are the “2nd Workshop on Creative-text Translation and Technology” (CTT 2025), the 3rd “International Workshop on Gender-Inclusive Translation Technologies” (GITT 2025), the 3rd “International Workshop on Automatic Translation for Signed and Spoken Languages” (AT4SSL), and the 11th “Workshop on Patent and Scientific Literature Translation” (PSLT 2025). We are also happy to see the start of a hopefully equally successful new series, with the 1st “Workshop on Artificial Intelligence and Easy and Plain Language in Institutional Contexts” (AI & EL/PL). With the exception of PSLT, they will all run for a full day, on the 23rd or on the 24th of June. Five half-day tutorials were also submitted, and three will be offered to the participants: “Understanding Large Language Model-Generated Translations”, “Leveraging Examples in Machine Translation”, and “Best practices for data quality in human annotation of translation datasets”. Our hope is that the choice between such diverse and exciting proposals will be a difficult one, and that these two pre-conference days will be as enjoyable and rewarding as possible, sparking new ideas, collaborations, and conversations in Geneva and beyond.

Sheila Castilho (Dublin City University, Ireland)
François Yvon (Sorbonne University, France)

EAMT 2024 Best Thesis Award (Anthony C. Clarke Award)

Six PhD theses defended in 2024 were received as candidates for the 2024 year edition of the EAMT Best Thesis Award, all of which were eligible. Eight external reviewers were recruited to examine and score the theses alongside five EAMT executive committee members. Each thesis was evaluated according to predefined criteria: how challenging the topic was, how relevant the results were to the MT field and the strength of its impact in terms of scientific publications. As in previous years, 2024 was another strong year for PhD theses in machine translation.

All PhD theses were of good quality, focused on interesting topics and were all highly appreciated by reviewers. A panel of two EAMT Executive Committee members (Barry Haddow and Rachel Bawden) was assembled to process the reviews and select a winner that was later ratified by the EAMT executive committee.

We are pleased to announce that the **winner of the 2024 edition of the EAMT Best Thesis Award is Ricardo Rei's thesis "Robust, Interpretable and Efficient MT Evaluation with Fine-tuned Metrics"** (Unbabel, INESC-ID, Instituto Superior Técnico, Portugal), supervised by Maria Luísa Torres Ribeiro Marques da Silva Coheur and Alon Lavie.

In addition, the committee judged that the thesis of **Sara Papi** (University of Trento & Fondazione Bruno Kessler) entitled "Direct Speech Translation in Constrained Contexts: the Simultaneous and Subtitling Scenarios" and supervised by Marco Turchi and Matteo Negri was **"highly commended"**.

The awardee will receive a prize of €500, together with an inscribed certificate. In addition, Dr. Rei will present a summary of their thesis at the 20th Machine Translation Summit in Geneva, Switzerland, receive complimentary membership to the EAMT in 2026 and will receive a travel bursary of €200.

Chairs of the EAMT Best Thesis Award 2024
Rachel Bawden, Inria, Paris, France
Barry Haddow, University of Edinburgh, UK

Organising Committee

General Chair

Helena Moniz, Universidade de Lisboa / INESC-ID, Portugal

Local Organising Committee

Pierrette Bouillon, University of Geneva, Switzerland
Johanna Gerlach, University of Geneva, Switzerland
Sabrina Girletti, University of Geneva, Switzerland
Lise Volkart, University of Geneva, Switzerland

Local Support Team

Sevita Caseres, University of Geneva, Switzerland
Bastien David, University of Geneva, Switzerland
Céline De Graaf, University of Geneva, Switzerland
Julie Humbert-Droz, University of Geneva, Switzerland
Rebeka Mali, University of Geneva, Switzerland
Lucía Morado Vázquez, University of Geneva, Switzerland
Jonathan Mutal, University of Geneva, Switzerland
Lucía Ormaechea Grijalba, University of Geneva, Switzerland
Aurélien Picton, University of Geneva, Switzerland
Donatella Pulitano, University of Geneva, Switzerland
Silvia Rodríguez Vázquez, University of Geneva, Switzerland
Valentin Scourneau, Université de Mons
Marianne Starlander, University of Geneva, Switzerland
Irene Strasly, University of Geneva, Switzerland
Nikolaos Tsourakis, University of Geneva, Switzerland
Florine Voisard, University of Geneva, Switzerland

Publications Chair

Raphael Rubino, University of Geneva, Switzerland
Rico Sennrich, University of Zurich, Switzerland

Track Chair: Research Technical

Catarina Farinha, Unbabel, Portugal
Marco Gaido, Fondazione Bruno Kessler, Italy

Track Chair: Research Translators and Users

Joke Daems, Ghent University, Belgium
Dorothy Kenny, Dublin City University, Ireland

Track Chair: Implementations and Case Studies

Samuel Lübli, Textshuttle/Supertext, Switzerland
Martin Volk, University of Zurich, Switzerland

Track Chair: Products and Projects

Miquel Esplà-Gomis, University of Alicante, Spain
Vincent Vandeghinste, KU Leuven, Belgium

Workshops and Tutorials Chair

Sheila Castilho, Dublin City University, Ireland
François Yvon, Sorbonne University, France

Programme Committee

Track: Research Technical

Benyamin Ahmadnia	UC Davis
Dr Khetam Al Sharou	Imperial College London
Àlex R. Atrio	HEIG-VD / HES-SO & EPFL
Vicent Briva-Iglesias	SFI CRT D-REAL, Dublin City University
José G. C. de Souza	Unbabel
Vera Cabarrão	Unbabel Lda.; INESC-ID
Michael Carl	Kent State University
Luisa Coheur	INESC-ID
Mattia Antonino Di Gangi	AppTek
Siddharth Divi	SSN College of Engineering
Konstantin Dranch	Custom.MT
Kevin Duh	Johns Hopkins University
Hiroshi Echizenya	Hokkai-Gakuen University
Carlos Escolano	UPC - BSC
Miquel Esplà-Gomis	Universitat d'Alacant
Mikel Forcada	Universitat d'Alacant
Javier García Gilabert	Barcelona Super Computing Center (BSC)
Cyril Goutte	National Research Council Canada
Barry Haddow	University of Edinburgh
Rejwanul Haque	South East Technological University
Iikka Hauhio	University of Helsinki, Kielikone Oy
Javier Iranzo-Sánchez	Universitat Politècnica de Valencia
Josef Jon	Charles University
Swarang Joshi	IIIT Hyderabad
Alina Karakanta	Leiden University
Maria Kunilovskaya	University of Saarland
Natalie Kübler	University of Paris
Gorka Labaka	University of the Basque Country
Tsz Kin Lam	University of Edinburgh
Ekaterina Lapshinova-Koltunski	University of Hildesheim
Yves Lepage	Waseda University
Qun Liu	Huawei Noah's Ark Lab
John Mendonca	INESC-ID
Miguel Menezes	Lisboa, Inesc-ID, Unbabel
Thomas Moerman	Ghent University
Kenton Murray	Johns Hopkins University
Jonathan Mutal	UNIGE
Masaaki Nagata	NTT
Artur Nowakowski	Lanigo / Adam Mickiewicz University
Constantin Orasan	University of Surrey
David Orrego-Carmona	University of Warwick
Antonio Pareja-Lora	ATLAS (UNED) / FITISPos (UAH) / DMEG (UdG, México) / DSIC, ILSA (UCM)
Seong-Bae Park	Kyung Hee University
Patrícia Pereira	Instituto Superior Técnico
Andrea Piergentili	University of Trento
Esther Ploeger	Aalborg University

David Ponce	Vicomtech
Andrei Popescu-Belis	HEIG-VD / HES-SO
Maja Popovic	ADAPT Centre @ DCU
Bo Ren	Microsoft
Fatiha Sadat	UQAM
Beatrice Savoldi	Fondazione Bruno Kessler
Yves Scherrer	University of Oslo
Dimitar Shterionov	Tilburg University
Michel Simard	National Research Council Canada (NRC)
Patrick Simianer	Lilt, Inc.
Sokratis Sofianopoulos	ILSP / Athena R.C.
Rubén Solera-Ureña	INESC-ID Lisboa
Rui Sousa-Silva	University of Porto
Felix Stahlberg	Google Research
Katsuhito Sudoh	Nara Women's University
Marek Suppa	Comenius University in Bratislava
Felipe Sánchez-Martínez	Universitat d'Alacant
Marina Sánchez-Torrón	Smartling
Aleš Tamchyna	Phrase a.s.
Antonio Toral	Universitat d'Alacant
Marco Turchi	Zoom
Jannis Vamvas	University of Zurich
Vincent Vandeghinste	Instituut voor de Nederlandse Taal, Leiden // Centre for Computational Linguistics, KU Leuven
David Vilar	Google
Taro Watanabe	Nara Institute of Science and Technology
Guillaume Wisniewski	LLF - Université de Paris
Tong Xiao	Northeastern University (CN)
Jinan Xu	Beijing Jiaotong University
Rik van Noord	University of Groningen

Track: Research Translators and Users

Sergi Alvarez-Vidal	UPF
Fabio Alves	UFMG
Nora Aranberri	University of the Basque Country
Lynne Bowker	Université Laval
Vicent Briva-Iglesias	SFI CRT D-REAL, Dublin City University
Patrick Cadwell	Dublin City University
Dragos Ciobanu	University of Vienna
Helle Dam Jensen	Aarhus University
Christophe Declercq	Utrecht University
Silvana Deilen	University of Hildesheim
Félix Do Carmo	CTS - University of Surrey
Aletta G. Dorst	Leiden University
Maria Fernandez-Parra	Swansea University
Federico Gaspari	ADAPT Centre, Dublin City University
Ana Guerberof Arenas	University of Groningen
Sari Hokkanen	Tampere University
Maarit Koponen	University of Eastern Finland
Ekaterina Lapshinova-Koltunski	University of Hildesheim
Manuel Lardelli	University of Graz

Rudy Loock	Université de Lille, France, & CNRS “Savoirs, Textes, Langage” re- search unit
Lieve Macken	Ghent University
Joss Moorkens	Dublin City University
Lucas N Vieira	University of Bristol
Masaaki Nagata	NTT
Mary Nurminen	University of Eastern Finland and Tampere University
Antoni Oliver	Universitat Oberta de Catalunya
Constantin Orasan	University of Surrey
David Orrego-Carmona	University of Warwick
John Ortega	Columbia and New York Universities
Jun Pan	Hong Kong Baptist University
Celia Rico	Universidad Complutense de Madrid
Akiko Sakamoto	Kansai University
Vilelmini Sosoni	Ionian University
Sanjun Sun	Beijing Foreign Studies University
María Del Mar Sánchez Ramos	Universidad de Alcala
Susana Valdez	Leiden University Centre for Linguistics
Kirti Vashee	Translated Srl
Mihaela Vela	Universität des Saarlandes
Callum Walker	University of Leeds

Track: Implementations and Case Studies

Chantal Amrhein	Supertext
Thomas Brovelli	Google
Oliver Czulo	Universität Leipzig
Marcello Federico	AWS AI Labs
Mark Fishel	University of Tartu
Tim Graf	Supertext
Ana Guerberof Arenas	University of Groningen
Silvia Hansen-Schirra	Johannes Gutenberg-Universität Mainz
Martin Kappus	Zürcher Hochschule für Angewandte Wissenschaften
Judith Klein	STAR Group
Maarit Koponen	University of Eastern Finland
Alon Lavie	Phrase
Christian Lieske	SAP
Helena Moniz	University of Lisbon
Mary Nurminen	University of Eastern Finland and Tampere University
Carla Parra Escartín	RWS Language Weaver
Matiss Rikters	Tilde
Florian Schottnann	Supertext
Sara Szoc	CrossLang
Carlos Teixeira	Universitat Rovira i Virgili
Jannis Vamvas	University of Zurich
Masaru Yamada	Rikkyo University
Maike Züfle	Karlsruher Institut für Technologie

Track: Products and Projects

Sergi Alvarez-Vidal	UPF
Eleftherios Avramidis	German Research Center for Artificial Intelligence (DFKI)
Romane Bodart	Université catholique de Louvain

Pedro Luis Díez-Orzas	Linguaserve I.S. S.A.
Judith Klein	STAR Group
Rebecca Knowles	National Research Council Canada
Ekaterina Lapshinova-Koltunski	University of Hildesheim
Manuel Lardelli	University of Graz
Marie-Aude Lefer	Université catholique de Louvain
Lieve Macken	Ghent University
Maite Melero	UPF
Yasmin Moslem	ADAPT Centre, Dublin City University
Vlad Niculae	Instituto de Telecomunicacoes, Lisboa
Mary Nurminen	University of Eastern Finland and Tampere University
Antoni Oliver	Universitat Oberta de Catalunya
Juan Antonio Pérez-Ortiz	Universitat d'Alacant, Departament de Llenguatges i Sistemes Informàtics
Shenbin Qian	University of Surrey
Felipe Sánchez-Martínez	Universitat d'Alacant
Arda Tezcan	Ghent University
Antonio Toral	Universitat d'Alacant
Daniel Torregrosa	WIPO
Tom Vanallemeersch	CrossLang NV
Bram Vanroy	Instituut voor de Nederlandse Taal
Rik van Noord	University of Groningen

Keynote Talk

Sign Language Machine Translation

Sarah Ebling
University of Zurich (UZH)

Abstract: In this talk, I will highlight the challenges of automatic translation between spoken languages and sign languages, touching on the topics of representation, data, and ethics. Additionally, I will introduce preprocessing tasks and discuss their state of the art. I will present research conducted in our group in the different areas.

Bio: Sarah Ebling is Full Professor of Language, Technology and Accessibility at the University of Zurich. Based in the field of computational linguistics, her research focuses on language-based assistive technologies in the context of persons with disabilities. Specifically, Sarah Ebling's research takes place in the context of deafness and hearing impairment, blindness and visual impairment, cognitive impairment, and language disorders. She is conducting research on sign language technologies, automatic text simplification, technologies for the audio description process, and computer-aided language sample analysis. Sarah Ebling is involved in international and national projects and is the PI of a large-scale Swiss innovation project entitled Inclusive Information and Communication Technologies"(2022-2026; <https://www.iict.uzh.ch/>).

Keynote Talk

Losing Our Tail – Again: Unnatural Selection and Translation Technologies

Eva Vanmassenhove
Tilburg University (TiU)

Abstract: Language is humanity’s primary tool to preserve and transmit knowledge, evolving alongside and with cultural technologies. Today, multilingual large language models (LLMs) represent the latest leap. Emerging evidence, however, suggests that LLMs might subtly (or not so subtly) distort language over time, amplifying frequent patterns while eroding linguistic richness, a phenomenon linked to *model collapse* which had already been observed in Neural Machine Translation (NMT) systems even before it was formally named. Unlike the visible artefacts that have already been observed in the AI-generated images created by computer vision models, linguistic shifts, such as the loss of the long tails of language, risk going unnoticed. Yet, they may have profound implications for language, translation, diversity, and the integrity of communication across different languages. This keynote will explore these ideas and connect them to specific translation issues, asking: What is (or will be) at stake when our world of words becomes increasingly shaped by multilingual LLMs.

Bio: Eva Vanmassenhove is a researcher specializing in Machine Translation and Language Technology, with a strong focus on tackling gender and algorithmic biases in translation systems. She earned her PhD from Dublin City University and now serves as an assistant professor in the Department of Cognitive Science and Artificial Intelligence at Tilburg University (TiU). At TiU, she contributes to the Computation and Psycholinguistics Research unit and the Inclusive and Sustainable Machine Translation Research Line. Her work aims to enhance machine translation by addressing biases, especially in gender representation, while preserving linguistic richness.

Keynote Talk

Ethics and MT Evaluation: An Exploded View

Joss Moorkens
Dublin City University (DCU)

Abstract: This talk reflects on ethical issues with MT using LLMs, looking particularly at a recent evaluation study in the medical domain. This study, and the potential for its findings to be used as a basis for action, bring abstract ethical issues into focus. More broadly, the heightened attention and potential for impact of MT and LLM research brings an added sense of responsibility for researchers, although this might be balanced with opportunities to contribute to the common good.

Bio: Joss Moorkens is an Associate Professor at the School of Applied Language and Intercultural Studies in Dublin City University (DCU), Science Lead at the ADAPT Centre, and member of DCU's Institute of Ethics and Centre for Translation and Textual Studies. He has published over 60 articles and papers on the topics of translation technology interaction and evaluation, translator precarity, and translation ethics. He is General Co-Editor of the journal *Translation Spaces* with Prof. Dorothy Kenny, co-editor of a number of books and journal special issues, and co-author of the textbooks *Translation Tools and Technologies* (Routledge 2023) and *Automating Translation* (Routledge 2024). He sits on the board of the European Masters in Translation Network.

Tutorial

Understanding Large Language Model-Generated Translations: How Can They Adapt to Different Translation Specifications and Pass the Translation Turing Test?

Longhui Zou¹, Michael Carl², Alan Melby³, Brandon Torruella⁴, Masaru Yamada⁵

¹University of Montana, ²Kent State University - CRITT, ³International Federation of Translators,

⁴Brigham Young University, ⁵Rikkyo University

Abstract: This tutorial explores the practical application of the Translation Turing Test (TTT) within today’s evolving generative AI landscape, addressing the growing need for human-centered approaches to translation project management and machine translation evaluation. While substantial research has examined large language models (LLMs)’ translation quality, little attention has been paid to their potential in managing the complex human interactions that characterize real-world translation project negotiations.

The TTT is a translation-specific adaptation of the classic Turing Test, evaluating whether a machine-managed translation project can successfully imitate a professional human project manager. In the TTT, a requester interacts with both human and computer systems to negotiate translation specifications and conduct a complete translation project. The machine passes if the requester cannot distinguish between the two managers more than 30% of the time.

This half-day tutorial guides participants through current language industry practices and the three major TTT components: specification negotiation, target text quality assessment, and complaint negotiation. By comparing three translation project cycles (managed by a human professional, a trained amateur, and a generative AI agent), we evaluate whether LLM-powered agents can handle complex coordination tasks characteristic of language service providers.

The program includes four sessions: introduction to the TTT, demonstration of requester-provider negotiations, translation quality evaluation including MQM customization and syntactic complexity analysis, and complaint negotiations. Participants gain both theoretical understanding and practical experience assessing the feasibility of integrating LLMs into real-world translation projects that support or enhance human project managers’ roles.

Tutorial

Leveraging Examples in Machine Translation: A Guide to Retrieval and Integration Strategies

Maxime Bouthors¹, Josep Maria Crego²

¹ISIR - Sorbonne Université - CNRS, ²SYSTRAN by ChapsVision

Abstract: Retrieval-Augmented Generation (RAG) systems are growing popular in the era of Large Language Models (LLM). Nonetheless, retrieval augmentation has a long time story tied to Machine Translation (MT). This tutorial aims to put in perspective the various techniques used to (1) retrieve relevant examples for databases; (2) integrate them into MT models. We will uncover how the selection of examples can be performed (fuzzy matching, cross-lingual retrieval), some of the model architectures (edit-based models, augmented encoder-decoder generation models, LLMs), as well as how the augmentation affects the output. The target audience are academics and industry professionals wishing to incorporate examples to improve their translation quality.

Tutorial

Best Practices for Data Quality in Human Annotation of Translation Datasets

Marina Sánchez Torrón¹, Jennifer Wong¹

¹Smartling

Abstract: High-quality human annotations are essential for developing and evaluating machine learning (ML) models. However, annotation is a complex task, and creating reliable annotation datasets requires addressing multiple challenges. This tutorial provides comprehensive guidance on best practices for managing data quality in human annotation of translation datasets using the Multidimensional Quality Metrics (MQM) framework. Drawing from both academic research and industry experience, we cover the complete annotation lifecycle: from initial setup and annotator management to quality evaluation and improvement strategies. Through theoretical foundations and a practical demonstration, participants will learn concrete guidelines they can apply to create more reliable and consistent annotation datasets.

Table of Contents

Best Thesis Award

<i>Robust, interpretable and efficient MT evaluation with fine-tuned metrics</i>	
Ricardo Rei	1
<i>Direct Speech Translation in Constrained Contexts: the Simultaneous and Subtitling Scenarios</i>	
Sara Papi	2

Research – Technical

<i>Investigating Length Issues in Document-level Machine Translation</i>	
Ziqian Peng, Rachel Bawden and François Yvon	4
<i>Investigating the translation capabilities of Large Language Models trained on parallel data only</i>	
Javier García Gilabert, Carlos Escolano, Aleix Sant, Francesca De Luca Fornaciari, Audrey Mash, Xixian Liao and Maite Melero	24
<i>Improve Fluency Of Neural Machine Translation Using Large Language Models</i>	
Jianfei He, Wenbo Pan, Jijia Yang, Sen Peng and Xiaohua Jia	54
<i>Optimizing the Training Schedule of Multilingual NMT using Reinforcement Learning</i>	
Alexis Allemann, Àlex R. Atrio and Andrei Popescu-Belis	65
<i>Languages Transferred Within the Encoder: On Representation Transfer in Zero-Shot Multilingual Translation</i>	
Zhi Qu, Chenchen Ding and Taro Watanabe	81
<i>Decoding Machine Translationese in English-Chinese News: LLMs vs. NMTs</i>	
Delu Kong and Lieve Macken	99
<i>OJ4OCRMT: A Large Multilingual Dataset for OCR-MT Evaluation</i>	
Paul McNamee, Kevin Duh, Cameron Carpenter, Ron Colaianni, Nolan King and Kenton Murray	113
<i>Context-Aware or Context-Insensitive? Assessing LLMs’ Performance in Document-Level Translation</i>	
Wafaa Mohammed and Vlad Niculae	126
<i>Context-Aware Monolingual Evaluation of Machine Translation</i>	
Silvio Picinini and Sheila Castilho	138
<i>Culture-aware machine translation: the case study of low-resource language pair Catalan-Chinese</i>	
Xixian Liao, Carlos Escolano, Audrey Mash, Francesca De Luca Fornaciari, Javier García Gilabert, Miguel Claramunt Argote, Ella Bohman and Maite Melero	150
<i>Instruction-tuned Large Language Models for Machine Translation in the Medical Domain</i>	
Miguel Rios	162
<i>Lingonberry Giraffe: Lexically-Sound Beam Search for Explainable Translation of Compound Words</i>	
Théo Salmenkivi-Friberg and Iikka Hauho	173
<i>Testing LLMs’ Capabilities in Annotating Translations Based on an Error Typology Designed for LSP Translation: First Experiments with ChatGPT</i>	
Joachim Minder, Guillaume Wisniewski and Natalie Kübler	190

<i>Name Consistency in LLM-based Machine Translation of Historical Texts</i>	
Dominic P. Fischer and Martin Volk	204
<i>Non-autoregressive Modeling for Sign-gloss to Texts Translation</i>	
Fan Zhou and Tim Van de Cruys	220
<i>Exploring the Feasibility of Multilingual Grammatical Error Correction with a Single LLM up to 9B parameters: A Comparative Study of 17 Models</i>	
Dawid Wiśniewski, Antoni Solarski and Artur Nowakowski	231
<i>Do Not Change Me: On Transferring Entities Without Modification in Neural Machine Translation - a Multilingual Perspective</i>	
Dawid Wiśniewski, Mikołaj Pokrywka and Zofia Rostek	248
<i>Intrinsic vs. Extrinsic Evaluation of Czech Sentence Embeddings: Semantic Relevance Doesn't Help with MT Evaluation</i>	
Petra Barančíková and Ondřej Bojar	265
<i>Metaphors in Literary Machine Translation: Close but no cigar?</i>	
Alina Karakanta, Mayra Nas and Aletta G. Dorst	276
<i>Synthetic Fluency: Hallucinations, Confabulations, and the Creation of IrishWords in LLM-Generated Translations</i>	
Sheila Castilho, Zoe Fitzsimmons, Claire Holton and Aoife Mc Donagh	287
<i>Patent Claim Translation via Continual Pre-training of Large Language Models with Parallel Data</i>	
Haruto Azami, Minato Kondo, Takehito Utsuro and Masaaki Nagata	300
<i>The Devil is in the Details: Assessing the Effects of Machine-Translation on LLM Performance in Domain-Specific Texts</i>	
Javier Osorio, Afraa Alshammari, Naif Alatrush, Dagmar Heintze, Amber Converse, Sultan Al-sarra, Latifur Khan, Patrick T. Brandt and Vito D'Orazio	315
<i>Improving Japanese-English Patent Claim Translation with Clause Segmentation Models based on Word Alignment</i>	
Masato Nishimura, Kosei Buma, Takehito Utsuro and Masaaki Nagata	333
<i>Progressive Perturbation with KTO for Enhanced Machine Translation of Indian Languages</i>	
Yash Bhaskar, Ketaki Shetye, Vandan Mujadia, Dipti Misra Sharma and Parameswari Krishnamurthy	344
<i>Leveraging Visual Scene Graph to Enhance Translation Quality in Multimodal Machine Translation</i>	
Ali Hatami, Mihael Arcan and Paul Buitelaar	353
<i>Are AI agents the new machine translation frontier? Challenges and opportunities of single- and multi-agent systems for multilingual digital communication</i>	
Vicent Briva-Iglesias	365
<i>bytF: How Good Are Byte Level N-Gram F-Scores for Automatic Machine Translation Evaluation?</i>	
Raj Dabre, Kaing Hour and Haiyue Song	378
<i>Quality Estimation and Post-Editing Using LLMs For Indic Languages: How Good Is It?</i>	
Anushka Singh, Aarya Pakhale, Mitesh M. Khapra and Raj Dabre	388
Research – Translators and Users	
<i>Revisiting Post-Editing for English-Chinese Machine Translation</i>	
Hari Venkatesan	399

<i>Is it AI or PE that worry translation professionals: results from a Human-Centered AI survey</i>	
Miguel A. Jiménez-Crespo and Stephanie A. Rodríguez	407
<i>Prompt engineering in translation: How do student translators leverage GenAI tools for translation tasks</i>	
Jia Zhang, Xiaoyu Zhao and Stephen Doherty	420
<i>Can postgraduate translation students identify machine-generated text?</i>	
Michael Farrell	432
<i>MT or not MT? Do translation specialists know a machine-translated text when they see one?</i>	
Rudy Loock, Nathalie Moulard and Quentin Pacinella	442
<i>The Challenge of Translating Culture-Specific Items: Evaluating MT and LLMs Compared to Human Translators</i>	
Bojana Budimir	455
<i>Investigating the Integration of LLMs into Trainee Translators' Practice and Learning: A Questionnaire-based Study on Translator-AI Interaction</i>	
Xindi Hao and Shuyin Zhang	468
<i>Introducing Quality Estimation to Machine Translation Post-editing Workflow: An Empirical Study on Its Usefulness</i>	
Siqi Liu, Guangrong Dai and Dechao Li	485
<i>Human- or machine-translated subtitles: Who can tell them apart?</i>	
Ekaterina Lapshinova-Koltunski, Sylvia Jaki, Maren Bolz and Merle Sauter	496
<i>Extending CREAMT: Leveraging Large Language Models for Literary Translation Post-Editing</i>	
Antonio Castaldo, Sheila Castilho, Joss Moorkens and Johanna Monti	506
<i>To MT or not to MT: An eye-tracking study on the reception by Dutch readers of different translation and creativity levels</i>	
Kyo Gerrits and Ana Guerberof Arenas	516
<i>Translation Analytics for Freelancers: I. Introduction, Data Preparation, Baseline Evaluations</i>	
Yuri Balashov, Alex Balashov and Shiho Fukuda Koski	538
<i>ITALERT: Assessing the Quality of LLMs and NMT in Translating Italian Emergency Response Text</i>	
Maria Carmen Staiano, Lifeng Han, Johanna Monti and Francesca Chiusaroli	566
<i>Optimising ChatGPT for creativity in literary translation: A case study from English into Dutch, Chinese, Catalan and Spanish</i>	
Shuxiang Du, Ana Guerberof Arenas, Antonio Toral, Kyo Gerrits and Josep Marco Borillo .	578
<i>Improving MT-enabled Triage Performance with Multiple MT Outputs</i>	
Marianna J. Martindale and Marine Carpuat	592
<i>The GAMETRAPP project: Spanish scholars' perspectives and attitudes towards neural machine translation and post-editing</i>	
Cristina Toledo-Báez and Luis Carlos Marín-Navarro	608
<i>Using Translation Techniques to Characterize MT Outputs</i>	
Sergi Alvarez-Vidal, Maria Do Campo, Christian Olalla-Soler and Pilar Sánchez-Gijón	619