# 23rd Annual Conference of the European Association for Machine Translation (EAMT2022)

Ghent, Belgium
1-3 June 2022

**Editors:**

**Lieve Macken**
**Andrew Rufener**
**Joachim Van den Bogaert**
**Joke Daems**
**Arda Tezcan**
**Bram Vanroy**
**Margot Fonteyne**
**Loïc Barrault**
**Marta R. Costa-Jussà**
**Ellie Kemp**
**Spyridon Pilos**
**Christophe Declercq**
**Maarit Koponen**
**Mikel L. Forcada**
**Carolina Scarton**
**Helena Moniz**

ISBN: 978-1-7138-5346-6

**Printed from e-media with permission by:**

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2022) by the Association for Computational Linguistics
All rights reserved.

Copyright for individual papers remains with the authors and are licensed under a Creative Commons 3.0 Attribution-NonCommercial-NoDerivatives 3.0 International license, (CC-BY-ND 3.0) https://creativecommons.org/licenses/by-nc-nd/3.0/deed.en.

Printed with permission by Curran Associates, Inc. (2022)

For permission requests, please contact the Association for Computational Linguistics at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

**Additional copies of this publication are available from:**

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

# Contents

**Foreword from the General Chair** . . . . . **vii**

**Message from the Organising Committee Chairs** . . . . . **ix**

**Preface by the Programme Chairs** . . . . . **xi**

**EAMT 2022 Committees** . . . . . **xiii**

**Invited Speeches** . . . . . **1**

**EAMT 2020 and EAMT 2021 Best Thesis Award — Anthony C Clarke Award** **3**

Maha Elbayad. *Rethinking the Design of Sequence-to-Sequence Models for Efficient Machine Translation* . . . . . 5

Mattia Antonino Di Gangi. *Neural Speech Translation: From Neural Machine Translation to Direct Speech Translation* . . . . . 7

Danielle Saunders. *Domain Adaptation for Neural Machine Translation* . . . . . 9

**Research papers** . . . . . **11**

Minh-Quang Pham, Josep Crego and François Yvon. *Multi-Domain Adaptation in Neural Machine Translation with Dynamic Sampling Strategies* . . . . . 13

Rudy Loock, Sophie Léchauguette and Benjamin Holt. *The use of online translators by students not enrolled in a professional translation program: beyond copying and pasting for a professional use* . . . . . 23

Xabier Soto, Olatz Perez-De-Viñaspre, Gorka Labaka and Maite Oronoz. *Comparing and combining tagging with different decoding algorithms for back-translation in NMT: learnings from a low resource scenario* . . . . . 31

Dongqi Pu and Khalil Sima'an. *Passing Parser Uncertainty to the Transformer. Labeled Dependency Distributions for Neural Machine Translation.* . . . . . 41

Mark Pluymaekers. *How well do real-time machine translation apps perform in practice? Insights from a literature review* . . . . . 51

Ricardo Rei, Ana C Farinha, José G.C. de Souza, Pedro G. Ramos, André F.T. Martins, Luisa Coheur and Alon Lavie. *Searching for COMETINHO: The Little Metric That Could* . . . . . 61

Lise Volkart and Pierrette Bouillon. *Studying Post-Editese in a Professional Context: A Pilot Study* . . . . . 71

Minghan Wang, Jiaxin Guo, Yuxia Wang, Daimeng Wei, Hengchao Shang, Yinglu Li, Chang Su, Yimeng Chen, Min Zhang, Shimin Tao and Hao Yang. *Diformer: Directional Transformer for Neural Machine Translation* . . . . . 81

Taido Purason and Andre Tättar. *Multilingual Neural Machine Translation With the Right Amount of Sharing* . . . . . 91

Lieve Macken, Bram Vanroy, Luca Desmet and Arda Tezcan. *Literary translation as a three-stage process: machine translation, post-editing and revision* . . . . . 101
Àlex R. Atrio and Andrei Popescu-Belis. *On the Interaction of Regularization Factors in Low-resource Neural Machine Translation* . . . . . 111
Sebastian T. Vincent, Loïc Barrault and Carolina Scarton. *Controlling Extra-Textual Attributes about Dialogue Participants: A Case Study of English-to-Polish Neural Machine Translation* . . . . . 121
Nishant Kambhatla, Logan Born and Anoop Sarkar. *Auxiliary Subword Segmentations as Related Languages for Low Resource Multilingual Translation* . . . . . 131
Pedro Mota, Vera Cabarrão and Eduardo Farah. *Fast-Paced Improvements to Named Entity Handling for Neural Machine Translation* . . . . . 141
Alina Kramchaninova and Arne Defauw. *Synthetic Data Generation for Multilingual Domain-Adaptable Question Answering Systems* . . . . . 151
Tobias van der Werff, Rik van Noord and Antonio Toral. *Automatic Discrimination of Human and Neural Machine Translation: A Study with Multiple Pre-Trained Models and Longer Context* . . . . . 161
Khetam Al Sharou and Lucia Specia. *A Taxonomy and Study of Critical Errors in Machine Translation* . . . . . 171

**User papers** . . . . . **180**
Artur Nowakowski, Krzysztof Jassem, Maciej Lison, Rafał Jaworski, Tomasz Dwojak, Karolina Wiater and Olga Posesor. *nEYron: Implementation and Deployment of an MT System for a Large Audit & Consulting Corporation* . . . . . 181
Bianka Buschbeck, Jennifer Mell, Miriam Exel and Matthias Huck. *"Hi, how can I help you?" Improving Machine Translation of Conversational Content in a Business Context* . . . . . 189
Madalena Gonçalves, Marianna Buchicchio, Craig Stewart, Helena Moniz and Alon Lavie. *Agent and User-Generated Content and its Impact on Customer Support MT* . . . . . 199
Miguel Menezes, Vera Cabarrão, Pedro Mota, Helena Moniz, and Alon Lavie. *A Case Study on the Importance of Named Entities in a Machine Translation Pipeline for Customer Support Content* . . . . . 209
Maria Afara, Randy Scansani and Loïc Dugast. *Investigating automatic and manual filtering methods to produce MT-ready glossaries from existing ones* . . . . . 219
Celia Soler Uguet, Fred Bane, Anna Zaretskaya and Tània Blanch Miró. *Comparing Multilingual NMT Models and Pivoting* . . . . . 229
Kamal Kumar Gupta, Soumya Chennabasavraj, Nikesh Garera and Asif Ekbal. *Pre-training Synthetic Cross-lingual Decoder for Multilingual Samples Adaptation in E-Commerce Neural Machine Translation* . . . . . 239

**Translators' papers** . . . . . **247**
Justus Brockmann, Claudia Wiesinger and Dragoș Ciobanu. *Error Annotation in Post-Editing Machine Translation: Investigating the Impact of Text-to-Speech Technology* 249
Alina Karakanta, Luisa Bentivogli, Mauro Cettolo, Matteo Negri and Marco Turchi. *Post-editing in Automatic Subtitling: A Subtitlers' perspective* . . . . . 259
Sabrina Girletti. *Working with Pre-translated Texts: Preliminary Findings from a Survey on Post-editing and Revision Practices in Swiss Corporate In-house Language Services* . . . . . 269

**Project/product descriptions** . . . . . **279**

Arda Tezcan. *Dynamic Adaptation of Neural Machine-Translation Systems Through Translation Exemplars* . . . 281
Diego Bartolome and Chris Jacob. *Language I/O Solution for Multilingual Customer Support* . . . 283
Fernando Alva-Manchego and Matthew Shardlow. *Towards Readability-Controlled Machine Translation of COVID-19 Texts* . . . 285
Joke Daems and Janiça Hackenbuchner. *DeBiasByUs: Raising Awareness and Creating a Database of MT Bias* . . . 287
Mikel L. Forcada, Pilar Sánchez-Gijón, Dorothy Kenny, Felipe Sánchez-Martínez, Juan Antonio Pérez Ortiz, Riccardo Superbo, Gema Ramírez Sánchez, Olga Torres-Hostench and Caroline Rossi. *MultitraiNMT Erasmus+ project: Machine Translation Training for multilingual citizens (multitrainmt.eu)* . . . 289
Senja Pollak and Andraž Pelicon. *EMBEDDIA project: Cross-Lingual Embeddings for Less- Represented Languages in European News Media* . . . 291
Masaru Yamada, Takanori Mizowaki, Longhui Zou and Michael Carl. *Trados-to-Translog-II: Adding Gaze and Qualitivity data to the CRITT TPR-DB* . . . 293
Margot Fonteyne, Maribel Montero Perez, Joke Daems and Lieve Macken. *Writing in a second Language with Machine translation (WiLMa)* . . . 295
Eirini Kaldeli, Mercedes García-Martínez, Antoine Isaac, Paolo Sebastiano Scalia, Arne Stabenau, Iván Lena Almor, Carmen Grau Lacal, Martín Barroso Ordóñez, Amando Estela and Manuel Herranz. *Europeana Translate: Providing multilingual access to digital cultural heritage* . . . 297
Pierrette Bouillon, Johanna Gerlach, Jonathan Mutal and Marianne Starlander. *The PASSAGE project : Standard German Subtitling of Swiss German TV content* . . 299
Marta Bañón, Miquel Esplà-Gomis, Mikel L. Forcada, Cristian García-Romero, Taja Kuzman, Nikola Ljubešić, Rik van Noord, Leopoldo Pla Sempere, Gema Ramírez-Sánchez, Peter Rupnik, Vít Suchomel, Antonio Toral, Tobias van der Werff and Jaume Zaragoza. *MaCoCu: Massive collection and curation of monolingual and bilingual data: focus on under-resourced languages* . . . 301
Sheila Castilho and Natália Resende. *MT-Pese: Machine Translation and Post-Editese* 303
Elena Murgolo, Javad Pourmostafa Roshan Sharami and Dimitar Shterionov. *A Quality Estimation and Quality Evaluation Tool for the Translation Industry* . . . 305
Toms Bergmanis, Marcis Pinnis, Roberts Rozis, Jānis Šlapiņš, Valters Šics, Berta Bernāne, Guntars Pužulis, Endijs Titomers, Andre Tättar, Taido Purason, Hele-Andra Kuulmets, Agnes Luhtaru, Liisa Rätsep, Maali Tars, Annika Laumets-Tättar and Mark Fishel. *MTee: Open Machine Translation Platform for Estonian Government* . . . 307
Raúl Vázquez, Michele Boggia, Alessandro Raganato, Niki A. Loppi, Stig-Arne Grönroos and Jörg Tiedemann. *Latest Development in the FoTran Project – Scaling Up Language Coverage in Neural Machine Translation Using Distributed Training with Language-Specific Components* . . . 309
Gabriele Sarti and Arianna Bisazza. *InDeep × NMT: Empowering Human Translators via Interpretable Neural Machine Translation* . . . 311
José G.C. de Souza, Ricardo Rei, Ana C. Farinha, Helena Moniz and André F. T. Martins. *QUARTZ: Quality-Aware Machine Translation* . . . 313
Artur Nowakowski, Krzysztof Jassem, Maciej Lison, Kamil Guttmann and Mikołaj Pokrywka. *POLENG MT: An Adaptive MT Platform* . . . 315
Carlos Amaral and Peggy van der Kreeft. *plain X - AI Supported Multilingual Video Workflow Platform* . . . 317

Sheila Castilho. *DELA Project: Document-level Machine Translation Evaluation* . . . 319
Giorgio Bernardinello and Judith Klein. *Background Search for Terminology in STAR MT Translate* . . . 321
Dimitar Shterionov, Mirella De Sisto, Vincent Vandeghinste, Aoife Brady, Mathieu De Coster, Lorraine Leeson, Josep Blat, Frankie Picron, Marcello Paolo Scipioni, Aditya Parikh, Louis ten Bosh, John O'Flaherty, Joni Dambre and Jorn Rijckaert. *Sign Language Translation: Ongoing Development, Challenges and Innovations in the SignON Project* . . . 323
André F. T. Martins, Ben Peters, Chrysoula Zerva, Chunchuan Lyu, Gonçalo Correia, Marcos Treviso, Pedro Martins and Tsvetomila Mihaylova. *DeepSPIN: Deep Structured Prediction for Natural Language Processing* . . . 325
Dimitra Anastasiou, Anders Ruge, Radu Ion, Svetlana Segărceanu, George Suciu, Olivier Pedretti, Patrick Gratz and Hoorieh Afkari. *A Machine Translation-Powered Chatbot for Public Administration* . . . 327
Natalia Resende. *MTrill: Machine Translation Impact on Language Learning* . . . 329
Jourik Ciesielski and Heidi Van Hiel. *Connecting client infrastructure with Yamagata Europe machine translation using JSON-based data exchange* . . . 331
Alina Karakanta, Luisa Bentivogli, Mauro Cettolo, Matteo Negri and Marco Turchi. *Towards a methodology for evaluating automatic subtitling* . . . 333
Ekaterina Lapshinova-Koltunski, Maja Popović and Maarit Koponen. *DiHuTra: a Parallel Corpus to Analyse Differences between Human Translations* . . . 335
Peggy van der Kreeft, Alexandra Birch, Sevi Sariisik, Felipe Sánchez-Martínez and Wilker Aziz. *GoURMET – Machine Translation for Low-Resourced Languages* . . . 337
Tamás Váradi, Marko Tadić, Svetla Koeva, Maciej Ogrodniczuk, Dan Tufiş, Radovan Garabík, Simon Krek and Andraž Repar. *Curated Multilingual Language Resources for CEF AT (CURLICAT): overall view* . . . 339
Joachim Van den Bogaert, Laurens Meeus, Alina Kramchaninova, Arne Defauw, Sara Szoc, Frederic Everaert, Koen Van Winckel, Anna Bardadym and Tom Vanallemeersch. *Automatically extracting the semantic network out of public services to support cities becoming Smart Cities* . . . 341
Artūrs Vasiļevskis, Jānis Ziediņš, Marko Tadić, Željka Motika, Mark Fishel, Hrafn Loftsson, Jón Guðnason, Claudia Borg, Keith Cortis, Judie Attard and Donatienne Spiteri. *National Language Technology Platform (NLTP): overall view* . . . 343
Anabela Barreiro, José GC de Souza, Albert Gatt, Mehul Bhatt, Elena Lloret, Aykut Erdem, Dimitra Gkatzia, Helena Moniz, Irene Russo, Fabio Kepler, Iacer Calixto, Marcin Paprzycki, François Portet, Isabelle Augenstein and Mirela Alhasani. *Multi3Generation: Multitask, Multilingual, Multimodal Language Generation* . . . 345
Petra Bago, Sheila Castilho, Jane Dunne, Federico Gaspari, Andre Kåsen, Gauti Kristmannsson, Jon Arild Olsen, Natalia Resende, Níels Rúnar Gíslason, Dana D. Sheridan, Páraic Sheridan, John Tinsley and Andy Way. *Achievements of the PRINCIPLE Project: Promoting MT for Croatian, Icelandic, Irish and Norwegian* 347
Mattia Di Gangi, Nick Rossenbach, Alejandro Pérez, Parnia Bahar, Eugen Beck, Patrick Wilken and Evgeny Matusov. *Automatic Video Dubbing at AppTek* . . . 349
Itziar Aldabe, Jane Dunne, Aritz Farwell, Owen Gallagher, Federico Gaspari, Maria Giagkou, Jan Hajic, Jens Peter Kückens, Teresa Lynn, Georg Rehm, German Rigau, Katrin Marheinecke, Stelios Piperidis, Natalia Resende, Tea Vojtěchová and Andy Way. *Overview of the ELE Project* . . . 351
Maarit Koponen, Kais Allkivi-Metsoja, Antonio Pareja-Lora, Dave Sayers and Márta Seresi. *LITHME: Language in the Human-Machine Era* . . . 353

Ana Guerberof Arenas and Antonio Toral. *CREAMT: Creativity and narrative engagement of literary texts translated by translators and NMT* . . . . . 355
Pintu Lohar, Guodong Xie and Andy Way. *Developing Machine Translation Engines for Multilingual Participatory Spaces* . . . . . 357
Luisa Bentivogli, Mauro Cettolo, Marco Gaido, Alina Karakanta, Matteo Negri and Marco Turchi. *Extending the MuST-C Corpus for a Comparative Evaluation of Speech Translation Technology* . . . . . 359
Carlos Amaral and Sebastião Miranda. *Monitio - Large Scale MT for Multilingual Media Monitoring* . . . . . 361

**Sponsors** . . . . . **363**