

Machine Translation Summit XVIII

Volume 2: MT Users & Providers Track

Online
16 – 20 August 2021

Editors:

**Ben Huyck
Stephen Larocca
Janice Campbell
Jay Marciano**

**Konstantin Savenkov
Alex Yanishevsky
Stephen Richardson**

ISBN: 978-1-7138-3920-0

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© (2021) by the Association for Computational Linguistics
All rights reserved.

Copyright for individual papers remains with the authors and are licensed under a Creative Commons 4.0 license, CC-BY-ND. (<https://creativecommons.org/licenses/by-nd/4.0/>)

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

- 1 **Roundtable: Building MT Capacity and Competence In-House**
Digital Marketing Globalization at NetApp: A Case Study of Digital Transformation utilizing Neural Machine Translation
Edith Bendermacher

- 7 **Roundtable: Building MT Capacity and Competence In-House**
Neural Machine Translation at Ford Motor Company
Nestor Rychtyckyj

- 17 **Roundtable: Building MT Capacity and Competence In-House**
Salesforce NMT System: A Year Later
Raffaella Buschiazzo

- 29 **Roundtable: Building MT Capacity and Competence In-House**
Autodesk: Neural Machine Translation – Localization and beyond
Emanuele Dias

- 38 **Government Track: Neural Translator Designed to Protect the Eastern Border of the European Union**
Artur Nowakowski and Krzysztof Jassem

- 44 **Government Track: Corpus Creation and Evaluation for Speech-to-Text and Speech Translation**
Corey Miller, Evelyne Tzoukermann, Jennifer Doyon and Elizabeth Mallard

- 54 From Research to Production: Fine-Grained Analysis of Terminology Integration
Toms Bergmanis, Mārcis Pinnis and Paula Reichenberg

- 78 Glossary functionality in commercial machine translation: does it help? A first step to identify best practices for a language service provider
Randy Scansani and Loïc Dugast

- 89 Selecting the best data filtering method for NMT training
Fred Bane and Anna Zaretskaya

- 98 A Review for Large Volumes of Post-edited Data
Silvio Picinini
- 131 Accelerated Human NMT Evaluation Approaches for NMT Workflow Integration
James Phillips
- 149 MT Human Evaluation – Insights & Approaches
Paula Manzur
- 166 A Rising Tide Lifts All Boats? Quality Correlation between Human Translation and Machine Assisted Translation
Evelyn Yang Garland and Rony Gao
- 175 Bad to the Bone: Predicting the Impact of Source on MT
Alex Yanishevsky
- 200 Machine Translation Post-Editing (MTPE) from the Perspective of Translation Trainees: Implications for Translation Pedagogy
Dominika Cholewska
- 211 Using Raw MT to make essential information available for a diverse range of potential customers
Sabine Peng
- 227 Field Experiments of Real Time Foreign News Distribution Powered by MT
Keiji Yasuda, Ichiro Yamada, Naoaki Okazaki, Hideki Tanaka, Hidehiro Asaka, Takeshi Anzai and Fumiaki Sugaya
- 233 A Common Machine Translation Post-Editing Training Protocol by GALA
Viveta Gene and Lucía Guerrero
- 246 Preserving high MT quality for content with inline tags
Konstantin Savenkov, Grigory Sapunov, Pavel Stepachev
- 277 Early-stage development of the SignON application and open framework – challenges and opportunities
Dimitar Shterionov, John J O’Flaherty, Edward Keane, Connor O’Reilly, Marcello Paolo Scipioni, Marco Giovanelli and Matteo Villa

- 291 Deploying MT Quality Estimation on a large scale: Lessons learned and open questions
Aleš Tamchyna
- 306 Validating Quality Estimation in a CAT Workflow: Speed, Cost and Quality Trade-off
Fernando Alva-Manchego, Lucia Specia, Sara Szcc, Tom Vanallemeersch and Heidi Depraetere
- 316 Neural Translation for European Union (NTEU)
Mercedes García-Martínez, Laurent Bié, Aleix Cerdà, Amando Estela, Manuel Herranz, Rihards Krišlauks, Maite Melero, Tony O'Dowd, Sinead O'Gorman, Marcis Pinnis, Artūrs Stafanovič, Riccardo Superbo and Artūrs Vasiļevskis
- 335 A Data-Centric Approach to Real-World Custom NMT for Arabic
Rebecca Jonsson, Ruba Jaikat, Abdallah Nasir, Nour Al-Khdour and Sara Alisis
- 353 Building MT systems in low resourced languages for Public Sector users in Croatia, Iceland, Ireland, and Norway
Róisín Moran, Carla Para Escartín, Akshai Ramesh, Páraic Sheridan, Jane Dunne, Federico Gaspari, Sheila Castilho, Natalia Resende and Andy Way
- 382 Using speech technology in the translation process workflow in international organizations: A quantitative and qualitative study
Pierrette Bouillon and Jeevanthi Liyanapathirana
- 396 Multi-Domain Adaptation in Neural Machine Translation Through Multidimensional Tagging
Emmanouil Stergiadis, Satendra Kumar, Fedor Kovalev and Pavel Levin
- 421 cushLEPOR uses LABSE distilled knowledge to improve correlation with human translations
Gleb Erofeev, Irina Sorokina, Lifeng Han and Serge Gladkoff
- 440 A Synthesis of Human and Machine: Correlating “New” Automatic Evaluation Metrics with Human Assessments
Mara Nunziatini and Andrea Alfieri
- 466 Lab vs. Production: Two Approaches to Productivity Evaluation for MTPE for LSP
Elena Murgolo