The 14th Conference of
The Association for Machine Translation
in the Americas

www.amtaweb.org

PROCEEDINGS

Vol. **2**: MT **User** Track

Editors:
Janice Campbell & Dmitriy Genzel (Commercial Users)
Ben Huyck & Patricia O'Neill-Brown (Government Users)

©2020 The Authors. These articles are licensed under a Creative Commons 3.0 license, no derivative works, attribution, CC-BY-ND.

Welcome to the 14th Biennial Conference of the Association for Machine Translation in the Americas

– AMTA 2020 Virtual!

AMTA conferences traditionally provide a unique opportunity for academic and commercial researchers to share their results with colleagues as well as to understand real-world user requirements. Business and government participants benefit from updates on leading-edge R&D in MT and have a chance to present and discuss their use cases. At the same time, students who attend gain a broad perspective and understanding of the fascinating field of MT.

This year's conference, however, is significant in at least two aspects. The first is that neural machine translation (NMT) has become a de facto standard in research and industry. At our last conference in March of 2018, generic NMT systems had just begun to be widely used during the preceding year, but it was later in 2018 that customizable NMT systems became widely available, enabling many companies, governments, and other organizations to benefit from an even higher level of MT quality for their specific applications. Since then, NMT customization and usage across the spectrum from individual translators to large corporations has continued to snowball.

The second aspect has been more of a difficulty than an advantage. The COVID-19 pandemic has resulted in transforming AMTA 2020 from an in-person event at a spectacular venue in Orlando, Florida to a completely online conference. While this transformation has presented many unique challenges, we now see some silver linings in this cloud. Without the need to travel and its associated costs, our attendance numbers have doubled from previous years, and participation has come from around the globe. We have been fortunate to receive tremendous support from our many sponsors, for which we are most grateful. Notably, Microsoft has provided their Teams platform to support the virtual conference sessions.

I wish to offer my sincerest thanks to our conference organizing committee, without whom this virtual conference would not have taken place. They have worked long hours to organize and prepare for this unique format, navigating uncharted waters and overcoming various roadblocks. I trust that all who attend will benefit from the results of their diligent efforts.

Steve Richardson
AMTA President

Introduction

Commercial Track

The Commercial MT Users and Translators Track at AMTA 2020 features twenty-two presentations from enterprises and individuals who supply or implement machine translation. These include technology and language service providers, as well as a host of commercial entities seeking to leverage the benefits of machine translation for better customer engagement.

A common theme that runs through many of the presentations this year is the use of metrics, such as MTPE and Quality Estimation, and setting acceptability thresholds therefor. The goals are to continually improve and interactively adapt models, and predict quality in order to increase language and content coverage, enhance linguists' productivity, ensure end-user trust, or even forego post editing, in some cases. Presentations explore novel applications of MT such as in building multilingual datasets; creating input for select NLP tasks; and translation memory alignment.

On the business side are presentations that explore the challenges an enterprise might face in adopting machine translation and in using technology and metrics to find the best engines or brands to meet their use cases. A unique approach to measuring the Return on Investment for adopting MT is detailed. Students put to the test various NMT claims to determine if valid or hype.

Whether a buyer or supplier, more organizations are building their own engines, thanks to a multitude of toolkits and available training data. Domain customization is the norm. Presentations discuss how to use metadata in source to fine tune customization and they detail strategies for handling tags and placeholders to achieve better output results.

On the practical side, there are presentations on scaling up MT specifically for software and continuous localization scenarios in order to reduce or delay human intervention and still achieve maximum customer impact.

Finally, what bodes for the very near future? Presenters offer that it is identifying and resolving societal biases encoded in machine learning systems, or simultaneously translating speech.

The Commercial Track Co-Chairs

Janice Campbell
Dmitriy Genzel

Government Track

The AMTA 2020 Government and Military MT Stakeholders Track brings together machine translation users, developers, and researchers in government, military and public service worldwide. The proceedings include eight presentations covering a broad range of topics. Two of these presentations include papers that provide in-depth detail and context to the presentations.

Several submissions describe how to effectively use MT in government, as well as how to augment human translation efforts, including the use of complementary NLP tools such as Speech-to-Text (STT) technologies. Others describe the practical application, insertion and measurement of MT into government space. One discusses video to text MT for sign language. Another presentation describes a custom MT engine trained using US Government data to assist with the COVID-19 crisis.

This track is made possible by the hard work and contributions of many individuals. We would like to thank Steve Richardson and all members of the conference committee for their organizational support, Jennifer Doyon and the rest of the organizing committee for guidance on the government track, and all of the AMTA 2020 authors and reviewers.

The Government Track Co-Chairs

Benjamin Huyck
Patricia O'Neill-Brown

Contents

Commercial Track

Government Track