EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION

THE 18TH ANNUAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION (EAMT 2015)

MAY 11-13, 2015 • WOW TOPKAPI PALACE • ANTALYA / TURKEY

www.eamt2015.org

Edited by

İlknur Durgar El-Kahlout • Mehmed Özkan • Felipe Sánchez-Martínez

Gema Ramírez-Sánchez • Fred Hollowood • Andy Way

Organized by

BOGAZICI UNIVERSITY 1863

EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION

TÜBİTAK BİLGEM

EAMT 2015

Proceedings of the 18th Annual Conference of the European Association for Machine Translation

Antalya, Turkey

May 11 - 13, 2015 at the WOW Topkapı Palace

Edited by

İlknur Durgar El-Kahlout

Mehmed Özkan

Felipe Sánchez-Martínez

Gema Ramírez-Sánchez

Fred Hollowood

Andy Way

Organized by

Contents

Foreword	**iii**
Preface by the Programme Chairs	**v**
Message from the Conference Chair	**vi**
EAMT 2015 Committees	**vii**
Sponsors	**x**
List of papers	**xi**
Invited talk by Olga Beregovaya	**1**
Research papers	**2**
User papers	**193**
Project/product descriptions	**210**
Author index	**230**

Foreword

It has been a huge honour for me to serve as president of the European Association for Machine Translation (EAMT) over the past six years. As I step down from office, I am delighted that the last EAMT annual conference under my presidency is being held in the beautiful location of Antalya, Turkey. This continues the policy I started in 2009 of bringing EAMT to new regions of Europe. This began with our first visit to the Iberian peninsula (Barcelona, 2009), our first conference in France in 2010 (St. Raphaël), followed by the Benelux region in 2011 (Leuven), continuing in 2012 with our first conference hosted in Italy (Trento), followed by the 2014 meeting in Croatia (Dubrovnik). Quite coincidentally, you may have noticed that we have journeyed step-by-step from West to East, almost 4,000km in fact! Just to give you a little hint, next year's conference starts to reverse this trend, but only just! Of course, you'll have to wait until the closing session of this year's meeting before you find out the actual 2016 conference location!

The EAMT organised its first Workshop/Conference back in 1996, and now we come to our Eighteenth Annual Conference in 2015. It is fair to say that for many, our EAMT conferences are pencilled in long in advance as not-to-be missed events. As well as running successful conferences over longer periods of time with much larger numbers of attendees than in the past, in 2012 we ran the very successful MT Summit in Nice on behalf of the whole MT community. All these things provide clear evidence to me that the EAMT as an organisation is continuing to grow and thrive. As I've often noted before, since its inception in 1997, the EAMT has not raised its membership rates, and we will continue to hold the cost of membership for 2015. Joining us is great value, especially in years like this when more than one IAMT-affiliated event takes place: as well as EAMT, we will converge later in the year on Miami, for MT Summit XV. The close cooperation with the other regional associations, which started with help from Alon Lavie and Hitoshi Isahara – including mutual conference discounts for all IAMT members – continues, despite both AMTA (George Foster) and AAMT (Hiromi Nakaiwa) having elected new presidents. As this is my last conference as EAMT president, I can only hope that this partnership continues and thrives in the future.

As ever, I would like to thank my colleagues on the EAMT Committee. They work tirelessly on behalf of all of us, and we are truly fortunate to have such a strong body of colleagues representing our Association. This year, there are significant changes to your Committee, which we will announce during the General Assembly. I urge all of you to consider contributing to this service to the community. I would like to thank all those who have helped me on the Committee over the past 6 years; it is testament to them all that our Association remains strong, and in good hands for the future.

As in the recent past, I am confident that the programme that has been assembled for this 18th Conference is a strong one, which will be of great interest to you all. I would like to thank the Programme Co-Chairs Felipe Sánchez-Martínez (Research track) and Gema Ramírez-Sánchez and Fred Hollowood (User track), for their assistance in helping assemble the programme you have before you, comprising Research and User tracks, poster sessions, and a terrific Invited Speaker in Olga Beregovaya. As in recent conferences, we continue to feature a special session featuring prominent FP7/H2020 projects, which has proven very popular in the past.

Last but not least, I would especially like to thank our local organizer, İlknur Durgar El-Kahlout, who very generously volunteered to hold the meeting in Antalya. We are very grateful to İlknur and her team for their excellent organization of this event.

Finally, thanks to all of you for coming. I hope 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.

Andy Way
Deputy Director of ADAPT,
School of Computing,
Dublin City University.

President of the EAMT
away@computing.dcu.ie

Preface by the Programme Chairs

It is our pleasure to welcome you to the 18th Annual Conference of the European Association for Machine Translation (EAMT) to be held in Antalya, Turkey. We really enjoyed serving as a programme committee chairs for this year edition of which has become the most important event on machine translation in Europe for developers, researchers, users, professional translators and translation/localisation managers. As in previous editions, the conference is organised in three different tracks: a research track in which unpublished research results in machine translation and related areas are reported; a user track where users of machine translation companies, language service providers, government agencies and non-government organisations report their experience on using and adapting machine translation in their organisations; and a project and product description track where products and projects on machine translation have the opportunity to reach the broad audience of the conference.

We received 46 submissions to the research track, 7 submissions to the user track and 20 project/product descriptions. All these papers come from more than 20 different countries. Each of the research and user papers were peer-reviewed by three independent reviewers from the program committee. Following the reviewers' suggestion, one of the papers submitted to the research track was redirected to the user track; of the remaining 45 papers, 24 (53\%) were accepted for their publication in the conference proceedings: 10 of them were selected for oral presentation, whereas 14 will be presented as a poster. In the user track 4 papers were accepted (57\%), three of them were selected for oral presentation and one will be presented as a poster. In the project/product description track 18 papers were selected for presentation as a poster.

We will enjoy an invited talk by Olga Beregovaya, current Vice President of Language Tools at Welocalize, who has an extense experience of over 15 years in the localisation industry. We hope all attendees, researchers and users, will find her talk highly appealing. We will also have a presentation by the winner of the EAMT Best Thesis Award. Poster presenters will have a two-slides, two-minutes presentation of their papers in a poster boaster session.
We thank all the Programme Committee members and sub-reviewers, whose names are subsequently listed, for their detailed extensive reviews and useful recommendations which where vital in helping us to decide the papers to accept. We also thank all the authors, who tried their best to incorporate the reviewers' suggestions when preparing their camera-ready papers. For those papers that were not accepted we hope that the reviewers' comments will help them to improve their papers for their submission somewhere else. Special thanks goes to Mikel L. Forcada, who took care of the project/product description track.

Finally, a big thank you goes to the local organising committee and to all the authors who made this conference both possible and successful. We hope that the resulting selection of papers represents the best of machine translation research, development and real-world usage.

Felipe Sánchez-Martínez — Universitat d'Alacant
Gema Ramírez-Sánchez — Prompsit Language Enginnering
Fred Hollowood — Fred Hollowood Consulting

EAMT 2015 Programme Committee co-chairs

Message from the Conference Chair

It is my privilege and great pleasure to welcome you in WOW Topkapı Hotels for the 18th Conference of the European Association for Machine Translation. I am very proud that the EAMT conference is organized this time in Antalya, the tourism capital and most beautiful province of Turkey on the Mediterranean coast.

This year's conference format is the same as those of the last three years, two and half days of oral and poster sessions, followed by a social programme. I hope you will enjoy the high quality papers in three different tracks – research, user and product/project – that will give an overview of current developments and trends in Machine Translation.

The conference will be held at the WOW Topkapı Hotel Osmanlı Halls. Sessions and coffee breaks will be hosted in front of the conference room. Lunches will be served in the Hünkar restaurant in a separated area for EAMT'15 participants. I hope you will enjoy the social events that we have organized: the welcome reception in the Cariye Pool and the conference banquet on the beach of WOW Kremlin Palace. On the last day of the conference, a cultural tour will be organized to Kurşunlu and Perge.

I will have to express my gratitude to the EAMT Board for providing us with the opportunity to host the 18th EAMT conference; especially Andy Way, Mikel Forcada, Viggo Hansen, and Tony Clarke who helped me whenever needed. It was a great pleasure to work with you. I would like to thank the Program Chair, Felipe Sánchez-Martínez, and User Co-chairs Gema Ramírez-Sanchez and Fred Hollowood for taking care of the large number of submissions and for preparing the programme.

This conference would not be possible without the efforts of many people involved in its organization; our organizational partner DEKON particularly Kubilay Şahin and Kübra Şenkahveci who worked hard to prepare a successful setup for the conference, local organization co-chair Mehmed Özkan and our local organization committee members Kemal Oflazer, Coşkun Mermer, Yücel Bicil, Şeniz Demir and Alper Kanak who volunteered for the organization and helped during every stage of the preparations.

Lastly, I am grateful to our sponsors for their generous contributions; gold sponsor STAR, bronze sponsors KuveytTürk, Ebay, and Universal, Welcome Reception Sponsor SesTek and lastly lunch sponsor Welocalize.

I wish you a very successful 18th EAMT conference in Antalya. Enjoy both the academic program and Antalya during your stay!

İlknur Durgar El-Kahlout
TÜBİTAK-BİLGEM
ilknur.durgar@tubitak.gov.tr

EAMT 2015 Committees

Chairs

General Chair of the Conference: Andy Way - (Dublin City University, Ireland)
Research Programme Chair: Felipe Sánchez-Martínez - (Universitat d'Alacant, Spain)
User Programme Co-chair: Gema Ramírez-Sánchez - (Prompsit Language Engineering, Spain)
User Programme Co-chair: Fred Hollowood - (Fred Hollowood Consulting, Ireland)
Local Organization Co-chair: Mehmed Özkan -(Boğaziçi University)
Local Organization Co-chair: İlknur Durgar El-Kahlout -(TÜBİTAK-BİLGEM)

Local Organizing Committee

Kemal Oflazer (CMU- Qatar)
Coşkun Mermer (TÜBİTAK-BİLGEM)
Yücel Bicil (TÜBİTAK-BİLGEM)
Şeniz Demir (TÜBİTAK-BİLGEM)
Alper Kanak (TÜBİTAK-BİLGEM)

Research Committee

Jesús Andrés-Ferrer (Nuance Communications, USA)
Eleftherios Avramidis (German Research Center for Artificial Intelligence (DFKI), Germany)
Bogdan Babych (University of Leeds, UK)
Loïc Barrault (Université du Maine, France)
Núria Bel (Universitat Pompeu Fabra, Spain)
Nicola Bertoldi (Fondazione Bruno Kessler, Italy)
Laurent Besacier (Université J. Fourier, France)
Alexandra Birch (University of Edinburgh, UK)
Arianna Bisazza (University of Amsterdam, Netherlands)
Hervé Blanchon (Laboratoire d'Informatique de Grenoble, France)
Ondřej Bojar (Charles University in Prague, Czech Republic)
Nicola Cancedda (Applied Researcher at Microsoft, UK)
Michael Carl (Copenhagen Business School, Denmark)
Francisco Casacuberta (Universitat Politècnica de València, Spain)
Helena Caseli (Universidade Federal de São Carlos, Brazil)
Mauro Cettolo (Fondazione Bruno Kessler, Italy)
David Chiang (University of Notre Dame, USA)
Marta R. Costa-Jussà (Institute For Infocomm Research, Singapore)
Adrià De Gispert (University of Cambridge, UK)
Jinhua Du (Xi'an University of Technology, China)
Nadir Durrani (University of Edinburgh, UK)
Marc Dymetman (Xerox Research Centre Europe, France)
Andreas Eisele (Directorate-General for Translation (EC), Luxembourg)
Cristina España-Bonet (Universitat Politècnica de Catalunya, Spain)
Mireia Farrús (Universitat Pompeu Fabra, Spain)
Christian Federmann (Microsoft, USA)
Mark Fishel (University of Zurich, Switzerland)
Mikel L. Forcada (Universitat d'Alacant, Spain)

George Foster (National Research Council, Canada)
Federico Gaspari (Dublin City University, Ireland)
Ulrich Germann (University of Edinburgh, UK)
Jesús Giménez (Nuance Communications, USA)
Meritxell Gonzalez (Universitat Politècnica de Catalunya, Spain)
Barry Haddow (University of Edinburgh, UK)
Christian Hardmeier (University of Uppsala, Sweden)
Yifan He (New York University, USA)
Kenneth Heafield (Bloomberg Labs, USA)
Teresa Herrmann (Karlsruhe Institute of Technology, Germany)
Hieu Hoang (University of Edinburgh, UK)
Matthias Huck (University of Edinburgh, UK)
Gonzalo Iglesias (University of Cambridge, UK)
Jie Jiang (Applied Language Solutions, UK)
Marcin Junczys-Dowmunt (Adam Mickiewicz University, Poland)
Philipp Koehn (John Hopkins University, USA)
Roland Kuhn (National Research Council, Canada)
Alon Lavie (Carnegie Mellon University, USA)
Qun Liu (Dublin City University, Ireland; Chinese Academy of Sciences, China)
Yanjun Ma (Baidu, China)
Pavel Pecina (Charles University in Prague, Czech Republic)
Juan Antonio Pérez-Ortiz (Universitat d'Alacant, Spain)
Maja Popovic (German Research Center for Artificial Intelligence (DFKI), Germany)
Stefan Riezler (Heidelberg University, Germany)
Mike Rosner (University of Malta, Malta)
Raphael Rubino (Prompsit Language Engineering, Spain)
Kepa Sarasola (Euskal Herriko Unibertsitatea, Spain)
Lane Schwartz (University of Illinois, USA)
Holger Schwenk (University of Le Mans, France)
Kashif Shah (University of Sheffield, UK)
Khalil Simaan (University of Amsterdam, Netherlands)
Michel Simard (National Research Council, Canada)
Lucia Specia (University of Sheffield, UK)
Ankit Srivastava (Dublin City University, Ireland)
Sara Stymne (Uppsala University, Sweden)
Jörg Tiedemann (Uppsala University, Sweden)
Antonio Toral (Dublin City University, Ireland)
Dan Tufis (Academia Romana, Romania)
Marco Turchi (Fondazione Bruno Kessler, Italy)
Francis M. Tyers (UiT Norgga árktalaš universitehta, Norway)
Antal van Den Bosch (Radboud University Nijmegen, Netherlands)
Josef van Genabith (German Research Center for Artificial Intelligence (DFKI), Germany)
Vincent Vandeghinste (University of Leuven, Belgium)
David Vilar (Nuance Communications, USA)
Martin Volk (University of Zurich, Switzerland)
Andy Way (Dublin City University, Ireland)

Jürgen Wedekind (University of Copenhagen, Denmark)
François Yvon (Université Paris Sud, France)

User Committee

Jeff Allen (SaP, France)
Nora Aranberri (Euskal Herriko Uni, Spain)
Diego Bartolomé (Tauyou, Spain)
Olga Beregovata (Welocalise, US)
Eric Blassin (Lionbridge, France)
Arancha Caballero (Nuadda, Spain)
James Cogley (Microsoft, Ireland)
Arnaud Daix (Euroscript, Luxembourg)
Pedro Díez (Linguaserve, Spain)
Stephen Doherty (UNSW, Australia)
Kurt Eberle (Lingenio, Germany)
Wojceich Froelich (Argos, Poland)
Tatiana Gornostay (Tilde, Latvia)
Declan Groves (Microsoft, Ireland)
Manuel Herranz (Pangeanic, Spain)
Maxim Khalilov (bmmt, Germany)
Paul Mangell (alphacrc, UK)
Jay Marciano (Lionbridge, US)
Daniel Marcu (Uni Southern Calif, US)
John Moran (CNGL, Ireland)
Hideyuki Namiki (Sony, Japan)
Antoni Oliver (UOC, Spain)
Patricia Palidini (CA, Spain)
John Papaioannou (Lexcelera, France)
Niko Papula (Multilizer, Findland)
Mirko Plitt (Modulo Language Automation, Switzerland)
Phil Ritchie (Vistatec, Ireland)
Johann Roturier (Symantec, Ireland)
Javier Sastre (Ateknea Solutions, Spain)
Indra Semite (Tilde, Latvia)
Jean Senellart (Systran, France)
Pilos Spyridon (EC, Belgium)
Ventsislav Zhechev (Autodesk, Switzerland)

Sub-Reviewers

Sariya Karimova (Heidelberg University, Germany)
Arefeh Kazemi (University of Isfahan, Iran)
Gideon Maillette de Buij Wenninger (University of Amsterdam, Netherlands)
Carolina Scarton (University of Sheffield, UK)
Maarten van Gompel (Radboud University Nijmegen, Netherlands)
Longyue Wang (University of Macau, China)

Acknowledgments

The European Association for Machine Translation acknowledges with gratitude the support and sponsoring of the following institutions ahd companies:

Gold Sponsor

Bronze Sponsors

Welcome Reception Sponsor

Lunch Sponsor

List of papers

Invited Talk

What we want, what we need, what we absolutely can't do without – an enterprise user's perspective on machine translation technology and stuff around it
Olga Beregovaya 1

Research papers 2

Exploiting portability to build an RBMT prototype for a new source language
Nora Aranberri, Gorka Labaka, Arantza Díaz de Ilarraza and Kepa Sarasola 3

Building hybrid machine translation systems by using an EBMT preprocessor to create partialtranslations
Mikel Artetxe, Gorka Labaka, Kepa Sarasola 11

Using on-line available sources of bilingual information for word-level machine translation quality estimation
Miquel Esplá-Gomis, Felipe Sánchez-Martínez, Mikel L. Forcada 19

A general framework for minimizing translation effort: towards a principled combination of translation technologies in computer-aided translation
Mikel L. Forcada and Felipe Sánchez-Martínez 27

Can Translation Memories afford not to use paraphrasing?
Rohit Gupta, Constantin Orasan, Marcos Zampieri, Mihaela Vela, Josef van Genabith 35

Dependency-based Reordering Model for Constituent Pairs in Hierarchical SMT
Arefeh Kazemiy, Antonio Toral, Andy Way, Amirhassan Monadjemiy, Mohammadali Nematbakhshy 43

The role of artificially generated negative data for quality estimation of machine translation
Varvara Logacheva and Lucia Specia 51

Document-Level Machine Translation with Word Vector Models
Eva Martinez Garcia, Cristina Espana-Bonet, Lluis Marquez 59

The potential and limits of lay post-editing in an online community
Linda Mitchell 67

Post-Editing Evaluations: Trade-offs between Novice and Professional Participants
Joss Moorkens and Sharon O'Brien 75

Benchmarking SMT Performance for Farsi Using the TEP++ Corpus
Peyman Passban, Andy Way, Qun Liu 82

Dynamic Terminology Integration Methods in Statistical Machine Translation
Marcis Pinnis 89

Identifying main obstacles for statistical machine translation of morphologically rich South Slavic languages
Maja Popovic and Mihael Arcan 97

Poor man's lemmatisation for automatic error classification
Maja Popovic, Mihael Arcan, Eleftherios Avramidis, Aljoscha Burchardt, Arle Lommel 105

Truly Exploring Multiple References for Machine Translation Evaluation
Ying Qin and Lucia Specia 113

Searching for Context: a Study on Document-Level Labels for Translation Quality Estimation
Carolina Scarton, Marcos Zampieri, Mihaela Vela, Josef van Genabith, Lucia Specia 121

Stripping Adjectives: Integration Techniques for Selective Stemming in SMT Systems
Isabel Slawik, Jan Niehues, Alex Waibel 129

Evaluating machine translation for assimilation via a gap-filling task
Ekaterina Ageeva, Francis M. Tyers, Mikel L. Forcada, Juan Antonio Pérez-Ortiz 137

Unsupervised training of maximum-entropy models for lexical selection in rule-based machine translation
Francis M. Tyers, Felipe Sánchez-Martinez, Mikel L. Forcada 145

Assessing linguistically aware fuzzy matching in translation memories
Tom Vanallemeersch and Vincent Vandeghinste 153

Re-assessingthe WMT2013 Human Evaluation with Professional Translators Trainees
Mihaela Vela and Josef van Genabith 161

Integrating a Large, Monolingual Corpus as Translation Memory into Statistical Machine translation
Katharina Wäschle and Stefan Riezler 169

TargetSide Generation of Prepositions for SMT
Marion Weller, Alexander Fraser, Sabine Schulte im Walde 177

Word Alignment Based Parallel Corpora Evaluation and Cleaning Using Machine Learning Techniques
Ieva Zariņa, Pēteris Ņikiforovs, Raivis Skadiņš 185

User papers 193

Content Translation: Computer-assisted translation tool for Wikipedia articles
Niklas Laxström, Pau Giner, Santhosh Thottingal 194

Pre-reordering for Statistical Machine Translation of Non-fictional Subtitles.
Magdalena Plamada, Gion Linder, Phillip Ströbel, Martin Volk 198

SMT at the International Maritime Organization: experiences with combining in-house corpus with more general corpus
Bruno Pouliquen, Marcin Junczys-Dowmunt, Blanca Pinero, Michał Ziemski 202

Evaluation of the domain adaptation of MT systems in ACCURAT
Gregor Thurmair 206

Project/product descriptions 210

MixedEmotions: Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets
Mihael Arcan and Paul Buitelaar 211

The ACCEPT Academic Portal: Bringing Together Pre-editing, MT and Post-editing into a Learning Environment
Pierrette Bouillon, Johanna Gerlach, Asheesh Gulati, Victoria Porro, Violeta Seretan 212

Russian-Chinese Sentence-level Aligned News Corpus
Wenjun Du, Wuying Liu, Junting Yu, Mianzhu Yi 213

HimL (Health in my Language)
Barry Haddow 214

MT-enhanced fuzzy matching with Transit NXT and STAR Moses
Nadira Hofmann 215

HandyCAT - An Open-Source Platform for CAT Tool Research
Chris Hokamp and Qun Liu 216

TraMOOC: Translation for Massive Open Online Courses
Valia Kordoni, Kostadin Cholakov, Markus Egg, Andy Way, Lexi Birch, Katia Kermanidis, Vilelmini Sosoni, Dimitrios Tsoumakos, Antal van den Bosch, Iris Hendrickx, Michael Papadopoulos, Panayota Georgakopoulou, Maria Gialama, Menno van Zaanen, Ioana Buliga, Mitja Jermol and Davor Orlic 217

Streamlining Translation Workflows with Style Scorer
David Landan and Olga Beregovaya 218

Estonian-English Reversible Smart Phone Dictionary of Military Terms and Relevant Vocabulary
Epp Leete 219

FALCON: Federated Active Linguistic data CuratiON
David Lewis 220

Tapadóir
Eimear Maguire, John Judge and Teresa Lynn 221

Okapi+QuEst: Translation Quality Estimation within Okapi
Gustavo Henrique Paetzold, Lucia Specia, Yves Savourel 222

CRACKER: Cracking the Language Barrier
Georg Rehm 223

LTC KnowHow: Empowering the Social Enterprise in the Language Industry
Adriane Rinsche and Sabine Rinsches 224

Multi-Dialect Machine Translation (MuDMat)
Fatiha Sadat 226

Abu-MaTran: Automatic building of Machine Translation
Antonio Toral, Tommi A Pirinen, Andy Way, Gema Ramírez-Sánchez, Sergio Ortiz Rojas, Raphael Rubino, Miquel Esplà, Mikel Forcada, Vassilis Papavassiliou, Prokopis Prokopidis and Nikola Ljubešić 227

MNH-TT: A Platform to Support Collaborative Translator Training
Masao Utiyama, Kyo Kageura, Martin Thomas, Anthony Hartley 228

Smart Computer Aided Translation Environment
Vincent Vandeghinste, Tom Vanallemeersch, Frank Van Eynde, Geert Heyman, Sien Moens, Joris Pelemans, Patrick Wambacq, Iulianna Van der Lek - Ciudin, Arda Tezcan, Lieve Macken, Véronique Hoste, Eva Geurts and Mieke Haesen 229

Invited Talk
Olga Beregovaya

What we want, what we need, what we absolutely can't do without – an enterprise user's perspective on machine translation technology and stuff around it

As the gains made by current phrase-based MT systems approach a horizontal asymptote, those of us in industry who benefit from cutting-edge research are giving serious thought to what's next. The goal of this talk is to rekindle the dialog between research and industry circles on the immediate and future needs of the user community and finding mutually beneficial and interesting avenues for collaboration.
Of course, there are too many topics to cover in an hour, but I'll try to touch on some of the most relevant…

These are the areas of utmost interest and importance for the global business community as seen by a user from a major Language Service Provider:

Translation Quality: It's clear that we've nearly hit the wall with what more we can squeeze out of phrase-based MT systems, so where do we go from here? Are factored models the way to go, or should we in industry start giving serious consideration to systems using deep learning techniques and/or deep syntactic/semantic structures?

New real-life applications: User-generated content (UGC) is rapidly becoming a huge opportunity for translation & localization. Whether we're talking about blogs, reviews, or live chat, there are challenges for the MT status quo: normalization, ungrammatical/uncommon syntax, extreme sensitivity to maintaining proper negation or sentiment.

Domain adaptation: How much can we get out of minimal amounts of data? A little more data? Lots of out-of-domain data? How can we seamlessly integrate client/user dictionaries into standard SMT workflows?

Using and interpreting metadata: We're seeing a trend where we often have as much or more metadata than actual text to translate. How can this be leveraged to improve results, and make translators'/post-editors' lives easier?

Quality evaluation, utility prediction: What do we even mean by "utility"? Let's work together on establishing a standard. Can we get robust, reliable QE from limited (or no) bilingual data? (We've got some evidence that the answer is yes…) What can we say about quality *vis à vis* functionality?

Collaboration: We want to help! We're very interested to see research that's directly applicable, and we want to find ways to facilitate academic/industry partnerships. We can work with clients to try to make large amounts of domain-specific (non-parliamentary!) data available. Please reach out to us so we know what you're working on. Better yet, let's work on finding mutually beneficial projects.