Special issue

of the

Baltic Journal of Modern Computing

(Vol. 4(2016), No. 2)

Proceedings

of the

19th annual conference

of the

European Association
for Machine Translation
(EAMT)

Riga, Latvia, 2016

Foreword from the president of the European Association for Machine Translation

Mikel L. FORCADA

Departament de Llenguatges i Sistemes Informàtics,
Universitat d'Alacant, E-03690 Sant Vicent del Raspeig, Spain

mlf@ua.es

As newly-elected president of the European Association for Machine Translation (EAMT) it is a great pleasure for me to write the foreword to this special issue of the Baltic Journal of Modern Computing, which also serves as the proceedings of the first EAMT conference held under my presidency: the 19th annual conference of the EAMT in Riga, Latvia. It is also a great honour, but also a great responsibility, to be given the chance to continue the huge work done by Andy Way's during a fruitful six-year-long presidency. Taking advantage of the momentum impressed on the EAMT by Andy, I hope I can live up to the expectations of our membership and of the wider machine translation community. If I succeed, it will be thanks to the help of a hard-working Executive Committee, whose members are elected or ratified when our Association holds a General Assembly as part of each annual conference.

The EAMT started organizing annual workshops in 1996; then, these workshops have become annual conferences, and have been hosted all around Europe. In the last years, the venue has steadily moved from west to east: from Barcelona (2009) to Saint-Raphaël (2010) to Leuven (2011) to Trento (2012) and finally —after skipping one year to host the successful world-wide MT Summit 2013 in Nice— to Dubrovnik (2014), only to turn around a little bit after visiting Antalya in 2015 to come to Riga in 2016. Yes, you have guessed: EAMT 2017, our 20th annual conference, will surely be west from Riga. It will be announced at EAMT 2016 shortly after I am writing these lines. Those who miss our conference, will find out by visiting our Association's website, EAMT.org.

By the way, if you have not done so yet, please consider joining the EAMT. Our membership rates are low, particularly for students, and have not increased since the EAMT's inception. You will benefit from discounts when attending not only our conferences, but also the conferences held by our partner associations the Asia-Pacific Association for Machine Translation (AAMT) and the Association for Machine Translation in the Americas (AMTA). You will also have an exclusive chance to benefit from funding for your activities related to machine translation. And perhaps you can get even more involved and participate in serving the European machine translation community by becoming a member of the Executive Committee of the EAMT.

But let me go back to EAMT 2016. As in previous conferences, it is great to see the strong programme put together by our programme chairs: Antonio Toral, as research track chair, and Tony O'Dowd and Alexandru Ceausu, as user track co-chairs. There will

also be a projects and products session which showcases the advance of machine translation in Europe. And, last but not least, I also feel very fortunate to have Alex Fraser as our invited speaker to top it all off.

EAMT 2016 would have never been possible without the generous offer to host and the hard work done subsequently by the local organizing committee at the well-known machine translation and language technology company Tilde, headed by Andrejs Vasiļjevs, and without the hospitality of the University of Latvia represented by Juris Borzovs. I warmly thank them all!

Another round of heartfelt thank-yous go to the co-editors-in-chief of the Baltic Journal of Modern Computing, and a special one to managing co-editor Evalds Ikaunieks. This is the first time that our whole proceedings are given the unique chance of becoming a special issue in an indexed, open-access journal; I hope this multiplies the impact of the advances presented in our conference.

Finally, I would like to thank EAMT 2016 attendees for coming to Riga. I hope the conference leads to new friendships and fruitful collaboration.

Mikel L. Forcada
EAMT President
May 2016

Preface from the Programme Chairs

Antonio TORAL[1], Tony O'DOWD[2], Alexandru CEAUSU[3]

[1]ADAPT Centre, School of Computing, Dublin City University, Ireland
[2]KantanMT, , Invent Building, Dublin City University, Ireland
[3]AMPLEXOR International, Luxembourg, Luxembourg

atoral@computing.dcu.ie, tonyod@kantanmt.com,
alexandru.ceausu@amplexor.com

It is our pleasure to welcome you to the 19th annual conference of the European Association for Machine Translation (EAMT) to be held in Riga, Latvia. We have really enjoyed serving as programme chairs for this edition of the conference. The EAMT conference has become the most important event in Europe in the area of machine translation for researchers, users, professional translators, etc. As in previous editions, the conference is organised around three different tracks: research, user and projects/products. The research track concerns novel and significant research results in any aspect of machine translation and related areas while the user track reports users' experiences with machine translation, in industry, government, NGOs, etc. Finally, the project and product track offers projects and products the opportunity to be presented to the wide audience of the conference.

This year we have received 36 submissions to the research track, 8 submissions to the user track and 25 descriptions of projects and products. Overall, submissions come from 34 different countries. Each submission to the research and user tracks was peer-reviewed by three independent members of the Programme Committee. In the research track 19 papers out of 36 (53%) were accepted for publication while 5 papers out of 8 were accepted in the user track (63%).

Aside from regular papers from the three tracks, the programme includes an invited talk by Alexander Fraser, MT Group Leader at LMU München, on the hot topic of machine translation into morphologically rich languages. We will also have a presentation by the winner of the EAMT Best Thesis Award.

We would like to thank the Programme Committee members and additional reviewers, whose names are listed below, for their high quality reviews and recommendations. These have been very useful for the Programme Chairs to make decisions. We would also like to thank all the authors for trying their best to incorporate the reviewers' suggestions when preparing the camera ready papers. For those papers that were not accepted, we hope that the reviewers' comments will be useful to improve them. Special thanks to Mikel L. Forcada, who took care of the projects and products track.

EAMT 2016 Committees

Chairs

General chair: Mikel Forcada (Universitat d'Alacant, Spain)
Research programme chair: Antonio Toral (Dublin City University, Ireland)
User programme co-chair: Tony O'Dowd (KantanMT, Ireland)
User programme co-chair: Alexandru Ceausu (Amplexor, Luxembourg)
Local organisation co-chair: Andrejs Vasiļjevs (Tilde, Latvia)
Local organisation co-chair: Rihards Kalniņš (Tilde, Latvia)
Local host: Juris Borzovs (University of Latvia)

Research Committee

Eleftherios Avramidis (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)
Arianna Bisazza (University of Amsterdam, Netherlands)
Hervé Blanchon (Laboratoire d'Informatique de Grenoble, France)
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)
Jinhua Du (Dublin City University, Ireland)
Nadir Durrani (University of Edinburgh, UK)
Andreas Eisele (Directorate-General for Translation (EC), Luxembourg)
Cristina España-Bonet (Universitat Politècnica de Catalunya, Spain)
Mireia Farrús (Universitat Pompeu Fabra, Spain)
Mark Fishel (University of Tartu, Estonia)
George Foster (National Research Council, Canada)
Federico Gaspari (Dublin City University, Ireland)

Ulrich Germann (University of Edinburgh, UK)

Barry Haddow (University of Edinburgh, UK)

Christian Hardmeier (University of Uppsala, Sweden)

Yifan He (Bosch RTC, USA)

Kenneth Heafield (Bloomberg Labs, USA)

Teresa Herrmann (Karlsruhe Institute of Technology, Germany)

Hieu Hoang (New York University, Abu Dhabi, UAE)

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)

Qun Liu (Dublin City University, Ireland; Chinese Academy of Sciences, China)

Pavel Pecina (Charles University in Prague, Czech Republic)

Juan Antonio Pérez-Ortiz (Universitat d'Alacant, Spain)

Maja Popovic (Humboldt University of Berlin, Germany)

Raphael Rubino (Universität des Saarlandes, Germany)

Víctor Sánchez-Cartagena (Prompsit Language Engineering, Spain)

Felipe Sánchez-Martínez (Universitat d'Alacant, Spain)

Kepa Sarasola (Euskal Herriko Unibertsitatea, Spain)

Lane Schwartz (University of Illinois, USA)

Rico Sennrich (University of Edinburgh, UK)

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 (DFKI, Germany)

Sara Stymne (Uppsala University, Sweden)

Jörg Tiedemann (Uppsala University, Sweden)

Dan Tufis (Academia Romana, Romania)

Marco Turchi (Fondazione Bruno Kessler, Italy)

Francis M. Tyers (UiT Norgga árktalaš universitehta, Norway)

Josef van Genabith (Universität des Saarlandes, 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

Nora Aranberri (Euskal Herriko Uni, Spain)
Diego Bartolomé (Tauyou, Spain)
Olga Beregovaya (Welocalize, US)
James Cogley (Microsoft, Ireland)
Stephen Doherty (University of New South Wales, Australia)
Kurt Eberle (Lingenio, Germany)
Andreas Eisele (European Commission, DGT, Luxembourg)
Tatjana Gornostaja (Tilde, Latvia)
Declan Groves (Microsoft, Ireland)
Manuel Herranz (Pangeanic, Spain)
Fred Hollowood (Fred Hollowood Consulting, Ireland)
Maxim Khalilov (bmmt, Germany)
Hilário Leal Fontes (European Commission, Luxembourg)
Pedro Luis Díez-Orzas (Linguaserve, Spain)
Paul Mangell (Alpha CRC, UK)
Jay Marciano (Lionbridge, US)
Daniel Marcu (Uni Southern Calif, US)
John Moran (Transpiral, Ireland)
Antoni Oliver (Universitat Oberta de Catalunya, Spain)
Niko Papula (Multilizer, Finland)
Mirko Plitt (Modulo Language Automation, Switzerland)
Bruno Pouliquen (World Intellectual Property Organization, Switzerland)
Gema Ramírez-Sánchez (Prompsit Language Engineering, Spain)
Spyridon Pilos (European Commission, Belgium)
Gregor Thurmair (Linguatec, Germany)
Ventsislav Zhechev (Yelp, US)

Additional Reviewers

Mikel Artetxe (Euskal Herriko Unibertsitatea, Spain)
Annette Rios Gonzales (University of Zurich, Switzerland)
Sanja Štajner (University of Mannheim, Germany)
Eva Vanmassenhove (Dublin City University, Ireland)