

Preface to the Proceedings

The MT Summits have been established as the meeting place for people and companies with a professional interest in machine translation and other tools for translation, be it researchers, developers, providers, users, or watchers who want to be abreast of new developments.

This MT Summit VIII, which is the first MT Summit conference of the new century, contributes strongly to this tradition by bringing researchers, users and providers together for a multi-faceted week in Galicia. The programme features 68 papers, of which 6 will be presented in a poster session.

The term machine translation, MT, is interpreted in its widest sense at the MT Summits, and in the three sister organisations, EAMT, AAMT and AMTA, in charge of the MT Summits. This means that first of all MT refers not only to the original FAHQ (Fully Automatic High Quality) visions and their possible implementation, but also other tools for translation, be it translation memory, alignment, example-based machine translation, etc. But secondly, MT Summits cover language technology for other multilingual purposes than translation proper, in particular multilingual language technology, e.g. multilingual text generation, and multilingual information retrieval.

In this way, the MT Summits are responding to the development in society where communication is becoming more and more electronic, and more and more global, and hence across languages or multilingual. The Internet contributes of course greatly to this development by providing access to content in all languages, thereby creating a need for tools for translation or other multilingual treatment. The fact that the Internet is decreasingly English-language supports this demand. The Internet itself obviously also becomes a prime test case for machine translation, MT on the Web . All these developments are mirrored in the papers presented at this Summit.

One of the key aspects of this conference programme is the number of papers dealing with practical solutions. Experienced users are presenting their approaches to workflow management, to the use of auxiliary tools etc. Providers are presenting the newest developments, and a very important trend is the focus on creating tools for the production of MT. This concerns particularly methods and tools for the creation of large dictionaries or knowledge bases, be at the end-users' place or at the MT providers. It is with great pleasure that the programme committee has read and accepted these papers, as the only practical way forward to the creation of MT for all languages is the existence and use of tools for the provision of the basic resources. In this respect the work on standards for linguistic resources and of conversion tools is also important.

It is important to know how to measure progress during the development of an MT system. It is also important for users to be able to measure which system or which combination of systems best suits their needs. This whole question of evaluation of MT has always been a hot topic, and again this year we are happy to have a track on evaluation running over two days.

Apart from paper and poster presentations, there are three invited speakers, one from each region. Giovanni B. Varile, European Commission, Information Society DG, will share his ideas of the methods used by a funding agency to obtain good tools for MT/multilinguality. Susanne Andersson, Sun Microsystems, Multilingual Information Services, takes us to the private sector and the globalisation, and describes the visions for the translation pipeline in a company working with a

large number of languages. Finally, Hitoshi Iida, Sony Computer Science Laboratories will provide insights and visions about speech translation.

The fact that researchers and users with explicit needs are brought together at the summit, and the fact that there are a fair number of research presentations as well as a fair number of what we called practical presentations above, provides an excellent opportunity for mutual feedback. The programme committee and the organisers hope that we have been able to create an environment for interesting discussions and maybe even for the creation of new partnerships.

Before closing, I first want to thank all of the programme committee for their invaluable contribution, and for their excellent feedback to the authors during the review process. The committee members are listed at the end of this preface.

Secondly, I want to thank Trine Esdahl and Christian Meyer at Center for Sprogteknologi who worked hard to overcome the technical challenges associated with the production of these proceedings.

Bente Maegaard
Center for Sprogteknologi, Copenhagen, Denmark
Chairman of the MT Summit VIII programme committee

Programme Committee

Chairperson:

Bente Maegaard, Center for Sprogteknologi, Denmark

Members:

Annelise Bech, Lingtech, Denmark
Nuria Bel, University of Barcelona, Spain
Herman Caeyers, Lant, Belgium
David Carter, 20/20 Speech, United Kingdom
Key-Sun Choi, KORTERM, KAIST, Korea
Daniel Grasmick, SAP, Germany
Anna Sögvall Hein, University of Uppsala, Sweden
Ed Hovy, ISI, USA
John Hutchins, East Anglia, UK
Hitoshi Iida, Sony, Japan
Ian Johnson, Lexiquest, France
Maghi King, ETI, University of Geneva, Switzerland
Steven Krauwer, University of Utrecht, Netherlands
Elliott Macklovitch, University of Montreal, Canada
Jörg Schütz, IAI, Germany
Harold Somers, UMIST, United Kingdom
Keh-Yih Su, Taiwan
Gregor Thurmair, SAIL Labs, Germany
Jun-ichi Tsujii, University of Tokyo, Japan
Anne Tucker, European Parliament, Luxembourg
John White, Litton PRC, USA