

Are Europe's Languages in Danger?

META-FORUM 2011 sends a mix of alarming and optimistic messages

META-FORUM, to be held in Budapest, Hungary on June 27/28, is an international conference on powerful technologies for the multilingual European information society and an official event of the Hungarian EU Presidency

Cultural and linguistic diversity is a hallmark of European integration. While the European Union works in 23 official languages, a total of about 60 languages are spoken on our continent when regional languages are taken into account. Information provided to citizens, business partners, consumers and tourists keeps growing at a fast pace. Will businesses and public administrations be able to translate this vast quantity of text into 23 or 60 languages? For 23 languages, we already have 506 pairs of source and target languages, for 60 languages 3540 pairs are needed. Surely, we cannot afford to sacrifice our linguistic diversity. But, can we afford to maintain it?

European integration is removing the borders for people, goods and capital. The global Internet permits the free exchange of information, and human language is the only medium for storing and sharing mankind's knowledge, thus serving as the fabric of the Web. Yet, this Web consists of many languages, and non-English content is growing very fast. The last remaining borders that hinder the free flow of ideas and thought are our language boundaries. Huge regional market opportunities remain unused today because of language barriers. A recent UNESCO report on multilingualism states that languages are an essential medium for executing fundamental rights, such as political expression, education and participation in society.¹ Concerned citizens have started using the social media of the Internet for lively dialogues on pressing social issues such as securing sustainable energy, reforming the financial system and dealing with demographic changes; but, these discussions are still parcelled by language boundaries. Successful European e-democracy needs to extend across linguistic borders.

Language technology is anticipated to provide the means for overcoming language boundaries. Indeed, in the last few years, automatic translation has improved considerably. Nevertheless, research and development are still much too slow and fragmented to solve our language problems in time. For obvious economic reasons, most research and development is centred on English. The majority of European languages are severely under-resourced and some are almost completely neglected. In this sense, our languages are not yet future-proof.

META-FORUM 2011 will report on the findings of 30 language white papers each surveying the status of a European language in the digital age. The conference brings together representatives of top-notch European research centres; small and large technology corporations; translation services and other users of language technology; language communities; and policy makers responsible for supporting research and innovation.

The meeting is organized by META-NET, a Network of Excellence consisting of 47 research centres in 31 countries and funded by the European Commission. META-NET is forging the "Multilingual Europe Technology Alliance" uniting technology researchers, providers and users for a large European research and innovation effort. Representatives of more than 280 organisations from 40 countries have already joined the alliance.

In his opening speech, Zoran Stančič, Deputy Director-General for Information Society and Media in the European Commission, formulates clear expectations: "In the European Union we have lifted to a large extent the physical borders between countries, still there are many borders remaining, including linguistic ones. Access to information in all languages is a necessary condition to enhance the circulation of products and services, and to boost the advent of a seamless digital single market. I strongly believe that Europe can further develop its leadership in language technologies and deliver solutions that will benefit the European society and economy at large. The only way we can achieve this, however, is to combine efforts and build a strong partnership with all stakeholders concerned. The place of language technologies in the future European research and innovation landscape will depend heavily on the ability of the field to speak with one voice."

The participants of META-FORUM will debate the guiding visions and initial plans for the envisaged technology push. In three vision groups and in a public web dialogue, experts from more than 100 companies and research organisations have already assembled bold visions for future research and visions about powerful language technology applications that will change our work and everyday life. The visions will be presented and discussed at the Budapest conference. The shared vision will serve as the starting point for a strategic research agenda, whose first outline will also be discussed at META-FORUM. As Hans Uszkoreit, the META-NET coordinator, explained, "With the right vision, actors and agenda, we can secure the future of Europe's languages and the competitiveness of a European industry sector in a key area of technological growth. The public costs for such an effort might not be higher than those for 100 kilometres of highway in a new member state."

¹ UNESCO Director-General, *Intersectoral mid-term strategy on languages and multilingualism*, Paris, 2007 (<http://unesdoc.unesco.org/images/0015/001503/150335e.pdf>).

The planned large-scale effort will not only improve automatic translation but also create enabling technologies for many other applications. Language technology is generally acknowledged today as one of the key growth areas in information technology. Large international corporations such as Google, Microsoft, IBM and Nuance have invested substantially in this area. In Europe, hundreds of small and medium enterprises have specialized in certain language technology applications or services. Language technology allows people to collaborate, learn, do business and share knowledge across language borders and independently of their computer skills.

Already today, language technology supports us in everyday tasks, such as writing e-mails or buying tickets. We benefit from language technology when searching for and translating web pages; using a word processor's spell and grammar checking features; operating our car's entertainment system or our mobile phone with spoken commands; getting recommendations in an online book-store; or following the instructions spoken by a mobile navigation app. In the near future, we will be able to talk to computer programs as well as machines and appliances including those long awaited service robots that will soon enter our homes and work places. Wherever we are, when we need information, we will simply ask for it, and, when we need help, we will outspokenly demand it. Removing the communication barrier between people and technology will change our world.

At META-FORUM, keynote speakers Thomas Hofmann of Google Europe and Bran Boguraev of IBM USA will report on technological advances and plans in their large international corporations. Some of the foremost European language technology scientists will summarize the state of the art, disclose new breakthroughs and share success stories about European research. Representatives of large language technology users, such as the translation service of the European Commission, Daimler Corporation and Vodafone will speak about the benefits of language technology applications, and they will present their needs.

On Tuesday, 28 June 2011, the META-Prize, which honours outstanding research, technologies and services for the European multilingual information society, will be awarded, and several META-Seals of Recognition will be presented for innovative multilingual products and services. An industry exhibition that concurrently runs with the main conference will feature presentations and demonstrations from large and small businesses working across the field of language technologies and showcase recent R&D results by EU-funded projects.

Two countries with multilingual societies—India with its 19 “official” languages and South Africa with 11 national languages—are conducting systematic language technology programs. At META-FORUM, these programs will be represented together with EU and national research programs. In two panel discussions, problems and solution strategies of multilingual societies and lesser-resourced European languages will be compared and discussed.

A prerequisite for the creation of successful applications for a language are large volumes of collected and interpreted language data, such as texts and speech recordings. Another precondition is the existence of basic language analysis technology for each language. At the Budapest forum, META-NET will unveil a new service for sharing and maintaining such resources, called META-SHARE, which will greatly facilitate research and development. The existence and quality of such resources vary from language to language, depending on the commercial relevance of the language, the problems the language poses for automatic processing and the research already devoted to it. Before the META-NET white papers, no one had assessed the state of European languages in regards to technology support. Now, META-NET can show why most languages are facing serious problems, and pinpoint the most threatening gaps.

The main message of META-FORUM 2011: Although the EU and its member states have already supported numerous individual research projects, the technology gap between “big” and “small” languages still keeps widening. No effort of sufficient scale and coordination has yet been made in Europe to create the missing resources and technologies as well as transfer technology to the majority of languages. There are strong reasons for approaching this immense challenge in a community effort involving the EU, its member states and industry, including: the high per-capita financial burden for smaller language communities; the needed transfer of technologies between languages; the interoperability of resources, tools and services; and the fact that linguistic borders often do not coincide with political borders. Europe must take action to prepare its languages for the digital age. They are a precious component of our cultural heritage and, as such, they deserve future-proofing.

Contact:

Prof. Dr. Hans Uszkoreit
Coordinator of META-NET
META-NET Office c/o DFKI Berlin
Alt-Moabit 91c
10559 Berlin, Germany
Email: office@meta-net.eu

Background Information:

META-NET Website: <http://www.meta-net.eu>
META-FORUM Website: <http://www.meta-net.eu/events/meta-forum-2011/>
Press material incl. this release in other languages:
<http://www.meta-net.eu/media/meta-forum-2011/>