

Georg Rehm / Hans Uszkoreit

Language Technology for multilingual Europe: Current status and future prospects

Abstract

Speaking one's mother tongue, be it Latvian, Hungarian, or Portuguese, must not become a social or economic disadvantage in the networked European information society of the 21st century. Language Technology has the potential to become the key solution to this crucial challenge if it is robust, cost-effective as well as available for all European languages and to all European citizens. However, in order to achieve these goals, the pace of research and development has to be accelerated by means of a major, dedicated push. A push with the magnitude needed can only be put into motion through a joint action of all stakeholder groups involved such as, among others, researchers, user and provider industries, technology integrators, language communities including the national institutions for language, politicians and society in general. To this end, META-NET – a European Network of Excellence that consists of 57 Language Technology research centres from 33 countries – is building META, the open and constantly growing Multilingual Europe Technology Alliance. This article introduces the META-NET White Paper Series in which we provide surveys on the state of language technology support for 30 European languages. Additionally we present three priority research themes as strategic and unifying umbrella topics for the next 10 to 15 years of future European language technology research.

1. Introduction

Many European languages run the risk of becoming victims of the digital age as they are under-represented and under-resourced online. Huge regional market opportunities remain untapped because of language barriers (Directorate-General for Translation of the European Commission 2009). If we do not take action now, speaking their mother tongue will become a severe social and economic disadvantage for many European citizens.

Innovative multilingual Language Technology (LT) is the ultimate intermediary that can help all European citizens to participate in an egalitarian, inclusive, and economically successful knowledge and information society. It can also help to establish and to further the single digital market. However, the degree to which LT is used and actually can be used in Europe varies enormously from language to language.

We are currently witnessing a revolution whose impact on language and society is comparable to that of Gutenberg's invention of the printing press. Digitisation and networked communication technology make an unlimited exchange of information and services – at any place, at any time possible. The downside is that certain groups (for example, people who live in rural areas or senior citizens) have difficulties participating in this new information-driven society. This problem is known as the digital divide.

Digital communication will have far-reaching and dramatic effects on Europe's languages, just like modern printing did five hundred years ago. Back then the new opportunities of large-scale communication triggered orthographic and grammatical standardisation for some languages and made the rapid dissemination of new scientific and intellectual ideas

possible. At the same time, small languages and regional dialects were rarely put to print. This turned out to be a considerable disadvantage as it limited their sphere of use to oral conversation and sometimes even contributed to their eventual extinction.

Today's multitude of official and unofficial languages as well as regional dialects is one of Europe's richest and most important cultural assets and it is also a vital part of its social success story. While big languages such as English or Chinese will certainly be well represented in the emerging digital society, many European languages are in real danger if we do not act now.

The key for protecting and furthering the highly heterogeneous group of more than 60 European languages is Language Technology. Research in this area has made considerable progress in the last few years. Machine Translation (MT) delivers a reasonable amount of accuracy, albeit only in specific domains, and experimental applications provide multilingual information and knowledge management as well as content production across languages. Relevant related areas are the development of intuitive language-based interfaces to technology ranging from household appliances, to heavy-duty machinery, vehicles and robots. The entertainment sector including games and mobile information services also holds many opportunities, as does the educational sector with computer assisted language learning and self-assessment software. While prototypes for several of these technologies exist, they are, however, by no means perfect and not ready for production use, yet. Nevertheless, it is safe to say that current progress opens a genuine window of opportunity.

Unfortunately, the current pace of technological progress is much too slow to arrive at substantial software products and services that are able to move communication in a multilingual environment significantly forward within the next 10 to 20 years. Those basic technologies that are already widely-used nowadays are usually monolingual and only available for a handful of languages. Well-known examples of the broad use of LT are the spelling and, recently, grammar correction features in modern text processing systems.

Applications for multilingual communication such as machine translation require a certain level of sophistication. Services that are available online, such as Google Translate or Bing Translator, are helpful when it comes to getting a rough idea of what a document in a foreign language is about. However, both products and also professional MT applications are fraught with multiple difficulties, especially if precise and also complete translations are needed.

The internet connects people, nationally and internationally, through a massively growing number of, in recent years especially mobile, devices. Multilingual LT enables instantaneous, cheap and effortless communication and interaction as well as business transactions across language borders. Once the necessary infrastructure is in place and research breakthroughs have been achieved, LT will allow people to collaborate, do business, share knowledge, and to participate in social and political opinion forming. LT offers tremendous opportunities for the European Union and its highly multilingual environment, both from the viewpoint of the economy and also from that of the citizen.

It also opens up transcontinental economic opportunities as experiences made developing, using and refining multilingual LT within the EU could be adapted to the specific needs of other multilingual communities, e.g., the citizens of India or Africa.

2. Three lines of action

META-NET is a European Network of Excellence forging the Multilingual Europe Technology Alliance (META) through a concerted effort to build a strong and powerful European community for and around LT (Rehm/Uszkoreit 2011). Its goal is to prepare the grounds for multilingual applications that enable automatic translation, information and knowledge management, including localisation, as well as content production and applications in related areas across all European languages. The objective is to advance LT so that communication and cooperation across languages becomes possible and to secure users of any language equal access to information and knowledge. META-NET, which started work on February 1, 2010, aims to advance research in LT as a means towards realising the vision of a Europe united in a single digital market and information space and is supporting these goals by pursuing three lines of action: META-VISION, META-SHARE and META-RESEARCH (see Figure 1).



Figure 1: META-NET's three lines of action

2.1 META-VISION: Fostering a dynamic and influential community around a shared vision and strategic research agenda

This line of action is concerned with a goal that is not only important but strategically indispensable for the overall success of the initiative: building up a coherent and homogeneous European LT community by bringing together representatives from the highly fragmented and heterogeneous stakeholder groups. These are comprised of, amongst others, researchers, user industries as well as provider industries, administrators, politicians, technology integrators and representatives of the language communities and national institutions for language. Significant steps towards realising this goal have been taken through various means such as, for example, by mobilising ca. 70 participants for three think tanks and focus groups called Vision Groups; these are made up of external experts from industry who provide seed ideas for innovative LT application scenarios for the future knowledge and information society. The Vision Groups focus on the areas of “Translation and Localisation”, “Media and Information Services” and “Interactive Systems”. Furthermore, META-NET has been engaged in intense dissemination activities with presentations at multiple national as well as international conferences.

In addition to the wide visibility provided by successful mobilisation activities, META-NET organised four conferences of its own: the METAnk 2010 was a brainstorming workshop at which about 100 researchers presented and discussed their long-term visions for the field of Human Language Technologies (June 4/5, Berlin). At Translingual Europe 2010 (June 7, Berlin) researchers discussed current problems and visions with representatives from the provider industries (such as Microsoft, Asia Online and ProMT) and Language Technology as well as Machine Translation users (European Patent Office, Symantec, EC DGT). At META-FORUM 2010 (November 17/18, Brussels, Belgium) the initial results of the vision building process were showcased to more than 250 participants. In a series of interactive sessions the participants provided their feedback and views on the visions presented by the project. At the follow-up conference, META-FORUM 2011 (June 27/28, Budapest, Hungary), more than 300 participants from research and industry came together for, among many other agenda items, an update on the vision building process, drafts of the META-NET Language White Paper Series (see section 3), a first outline of META-NET's Strategic Research Agenda (see section 4) and the META Prize and META Seal of Recognition award ceremonies; in the META Exhibition 40 exhibitors showcased LT products and recent research results. META-FORUM 2012 is to be held in Brussels again, on June 20/21, 2012, collocated with the Digital Agenda Assembly 2012.

2.2 META-SHARE: Creating an open resource exchange infrastructure

In its second line of action META-NET is building META-SHARE, a sustainable peer-to-peer network of repositories of language data, tools and web services that are documented with high-quality metadata and aggregated in inventories allowing for uniform search and access. Data and tools can be both open and with restricted access rights, free and for-a-fee. META-SHARE targets existing but also new and emerging language data, tools and systems required for building and evaluating new technologies as well as innovative products and services. In this respect, reuse, combination, repurposing and re-engineering of language data and tools play a crucial role. META-SHARE will eventually become an important component of an LT marketplace for researchers and developers, language professionals (translators, localisation experts, etc.), as well as for industrial players including SMEs and big enterprises. In this role, META-SHARE will cater for the full development cycle of LT, from research through to innovative products and services. In this regard, designing, building and successfully establishing META-SHARE as an important and valuable piece of infrastructure within the European and also global LT community is one of META-NET's decisive goals (Piperidis to appear 2012). Among the important relevant components of the META-SHARE infrastructure is a universal metadata scheme for the description of Language Resources and Language Technologies that was developed by a working group that consists of experts from within the initiative and several other European specialists (Gavrilidou et al. to appear 2012). We also explored thoroughly the landscape of language resources licensing and, with the help of legal experts, prepared a set of licensing templates. META-SHARE favours and aligns itself with the growing open data and open source movement, especially the Creative Commons Initiative. A first, fully functional prototype of META-SHARE was presented at META-FORUM 2010. Currently META-SHARE is in production use within the wider network

of excellence. An improved version will be rolled out for use both by META-NET and the public at large in the spring of 2012 (Federmann et al. to appear 2012).

2.3 META-RESEARCH: Building bridges to neighbouring technology fields

The third line of action consists of innovative research work with regard to leveraging advances in other fields to help LT. Specifically the work focuses on bringing more semantics into Machine Translation (MT), optimising the division of labour in hybrid MT, preparing an empirical base for MT and exploiting the context when computing an automatic translation. To this end, META-NET is carrying out research by building bridges to other fields and disciplines such as Machine Learning and the Semantic Web community. META-RESEARCH is concerned with collecting data, preparing data sets and language resources for evaluation purposes, compiling inventories of tools and methods, and organising workshops and advanced training events for its staff members. Among its current major outcomes are the clear identification of issues in Machine Translation in which semantics has shown potential to positively impact the state of the art, recommendations for approaching the problem of integrating semantic information in MT, and a list of tools and resources that could be employed for this purpose. A new language resource for MT, the Annotated Hybrid Sample MT Corpus, provides data for the language pairs English-German, English-Spanish and English-Czech. A third important outcome is software for the collection of multilingual hidden-web corpora. The tool clusters news articles in different languages discussing the same topic or event and clusters pages identified as being translations of each other. The research that is carried out in this line of action is meant to advance significantly the state of the art in MT.

2.4 Extension – impact – collaborations

META-NET has a founding consortium that consists of 13 partners in 10 countries. As the initiative operates on a European level we began to extend the network in November 2010. In the spring of 2012 the enlarged network consists of 57 partners in 33 countries (see Table 5). Most of the new members participate in three EU-funded projects that support the META-NET objectives by systematically collecting language resources and language technologies, curating and describing them with metadata records and making them available through META-SHARE, mobilising the communities in their respective countries and organising general awareness raising activities. These three projects – CESAR, METANET4U and META-NORD – commenced their work on February 1, 2011 – exactly one year after the start of META-NET.

In addition to the network of excellence and the open technology alliance META (Multilingual Europe Technology Alliance), META-NET drafted and signed individual collaboration agreements with more than 20 projects and initiatives funded by the European Union. Among projects these are machine translation projects such as ACCURAT, Let's MT, EuroMatrix Plus and iTranslate4.eu, Language Technology projects such as PANACEA and ATLAS, web- and W3C (World Wide Web Consortium)-related projects such as Multilingual Web and Multilingual Web-LT as well as large European initiatives and networks such as CLARIN and FLReNet.

We expect META-NET to have a significant, hitherto unprecedented, long-term impact on the European LT landscape. The three tightly integrated lines of action all aim at the same goal, albeit on different levels: to provide technological antidotes for the language barriers Europe has been facing for quite some time, in the form of robust, precise, high-performance, multilingual Language Technology; to stimulate the development of novel and innovative LT applications; to assemble and strengthen the European LT community; to raise awareness about the enormous potential Language Technology has for the European information society, for the single digital market and for society at large; to foster innovative LT research.

3. The META-NET Language White Paper Series: Language Technology support for Europe's languages

The META-NET Language White Paper Series describes the current state of language technology support for 30 different European languages and is a complement to the Strategic Research Agenda (see section 4). The individual volumes are meant to raise awareness for the topic of language technology. The target audience are mostly national but also international politicians, journalists, decision makers and the public at large.

The white papers include a language-specific assessment of existing technologies and resources, shortcomings and gaps, as well as a cross-language comparison. The documents were written for the following European languages (including all 23 EU member state languages): Basque, Bulgarian, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Greek, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian (bokmål and nynorsk), Polish, Portuguese, Romanian, Serbian, Slovak, Slovene, Spanish, and Swedish. Each Language White Paper is written in the language it reports upon and includes a complete English translation.

Draft versions of the Language White Papers were distributed at META-FORUM 2011 in Budapest (June 27/28, 2011); the documents are also available online at www.meta-net.eu. In total, more than 160 authors from within META-NET and several additional experts contributed to preparing the Language White Papers.

The current state of LT support varies considerably from one language community to another. In order to compare the situation between languages, the Language White Papers introduce an evaluation based on two sample application areas (machine translation and speech processing) and one underlying technology (text analysis) as well as basic resources needed for building LT applications. LT support for the languages was categorised using a five-point scale (1. excellent support; 2. good support; 3. moderate support; 4. fragmentary support; 5. weak or no support) and measured according to the following key criteria:

- **Speech Processing:** quality of existing speech recognition and synthesis technologies, coverage of domains, number and size of existing speech corpora, amount and variety of available speech-based applications.

- **Machine Translation:** quality of existing MT technologies, number of language pairs covered, coverage of linguistic phenomena and domains, quality and size of existing parallel corpora, amount and variety of available MT applications.
- **Text Analysis:** quality and coverage of existing text analysis technologies (morphology, syntax, semantics), coverage of linguistic phenomena and domains, amount and variety of available applications, quality and size of existing (annotated) text corpora, quality and coverage of existing lexical resources (e.g., WordNet) and grammars.
- **Resources:** quality and size of existing text corpora, speech corpora and parallel corpora, quality and coverage of existing lexical resources and grammars.

The in total more than 160 contributing authors to the Language White Papers prepared an initial language-specific assessment of LT support using an approach in which ca. 25 different applications, tools and resources were assessed along seven different axes and criteria. Later on, the 30 individual and language-specific matrices were condensed in multiple iterations in order to arrive at a single score per language and area.

Tables 1 to 4 show that there are dramatic differences in language technology support between the various European languages and technology areas. For all LT areas, English is ahead of any other language but even support for English is far from being perfect. While there are good quality software and resources available for some languages and application areas, others, usually smaller languages, have substantial gaps. Many languages lack basic technologies for text analysis and essential resources. Others have basic tools and resources but the implementation of, for example, semantic methods is still far away. Therefore, a large-scale effort is needed to attain the ambitious goal of providing high-quality language technology support for all European languages, for example through high quality machine translation.

Excellent support	Good support	Moderate support	Fragmentary support	Weak/ no support
	English	Czech Dutch Finnish French German Italian Portuguese Spanish	Basque Bulgarian Catalan Danish Estonian Galician Greek Hungarian Irish Norwegian Polish Serbian Slovak Slovene Swedish	Croatian Icelandic Latvian Lithuanian Maltese Romanian

Table 1: Speech processing – state of LT support for 30 European languages

Excellent support	Good support	Moderate support	Fragmentary support	Weak/ no support
	English	French Spanish	Catalan Dutch German Hungarian Italian Polish Romanian	Basque Bulgarian Croatian Czech Danish Estonian Finnish Galician Greek Icelandic Irish Latvian Lithuanian Maltese Norwegian Portuguese Serbian Slovak Slovene Swedish

Table 2: Machine translation – state of LT support for 30 European languages

Excellent support	Good support	Moderate support	Fragmentary support	Weak/ no support
	English	Dutch French German Italian Spanish	Basque Bulgarian Catalan Czech Danish Finnish Galician Greek Hungarian Norwegian Polish Portuguese Romanian Slovak Slovene Swedish	Croatian Estonian Icelandic Irish Latvian Lithuanian Maltese Serbian

Table 3: Text analysis – state of LT support for 30 European languages

Excellent support	Good support	Moderate support	Fragmentary support	Weak/ no support
	English	Czech Dutch French German Hungarian Italian Polish Spanish Swedish	Basque Bulgarian Catalan Croatian Danish Estonian Finnish Galician Greek Norwegian Portuguese Romanian Serbian Slovak Slovene	Icelandic Irish Latvian Lithuanian Maltese

Table 4: Speech and text resources – state of LT support for 30 European languages

As the META-NET Language White Papers serve as important communication instruments, high-quality paper and also ebook copies are currently being produced with a scientific publishing house in order to ensure a wide distribution to non-specialists, journalists, politicians, administrators and other stakeholders (Rehm/Uszkoreit 2012).

4. The Strategic Research Agenda for multilingual Europe

In addition to building up a coherent, dynamic and influential European LT community and to preparing the META-NET Language White Paper Series, another important goal of META-VISION is collaboratively – within and by the community – to prepare, establish and also promote a Strategic Research Agenda (SRA) for the European LT landscape. This SRA is intended to be a long-term instrument that will serve as a unifying umbrella for both industrial and academic research and development in the period leading up to 2020. The SRA contains high-level recommendations and suggestions for joint actions to be presented to the European Commission and national as well as regional bodies and funding agencies. The process of preparing the SRA is complex and includes representatives of META-NET, the abovementioned three Vision Groups (see section 2.1) and external experts.

After the three Vision Groups had collected literally hundreds of attractive and powerful technology visions in 2010 and early 2011 (Mariani/Magnini 2010; Koutsombogera/Piperidis 2010; Burchardt/Rehm 2010; Burchardt/Rehm/Sasaki 2011), the META Technology Council, a group that consists mostly of industry representatives and several researchers, took over in the complex process of preparing the SRA and discussed these visions in several meetings, reducing the number of potential visions to a shortlist of seven. The key criterion in these discussions was that the respective technology vision needed to be attractive and powerful enough to assemble behind it a very large propor-

tion of the European LT research and innovation landscape. At the same time the vision needed to present a convincing solution for the issue of technology-enabled multilingualism in Europe. Towards the end of 2011 we agreed upon three main research strands that we call priority themes (Burchardt/Rehm/Uszkoreit 2012):

1. **Translation Cloud** – The goal of this priority theme is a multilingual European society, in which all citizens can use any service, access all knowledge, enjoy all media and control any technology in their mother tongues. Written and spoken communication is not to be hindered anymore by language boundaries. Costs for large volume and specialized high-quality translation will be truly affordable. The citizen, the professional, the organization, or the software application in need of cross-lingual communication will use a single access point for channelling text or speech through a gateway that will instantly return the translations into the requested languages in the required quality and desired format. Behind this access point will be a network of generic and special-purpose services combining automatic translation or interpretation, language checking, post-editing, as well as human creativity and quality assurance where needed for achieving the demanded quality. The service will be free for small volume use and for high-volume base-line quality but it will offer extensive business opportunities for a wide range of language service and language technology providers.
2. **Social Intelligence and e-Participation** – The goal is to improve decision-making in business and society. The quality, speed and acceptance of decisions are the main factor for the success of social systems such as enterprises, communities, states and supranational organisations. Business intelligence and analytics programmes search online data for relevant information, decision support systems evaluate and order the information and apply decision rules. Social intelligence builds upon improved text analytics methodologies for the analysis of large volumes of social media, comments, blogs, forum postings etc. of citizens, customers, patients, consumers and other members of arbitrary stakeholder communities. Part of the analysis is directed to the status, opinions and acceptance associated with individual information units. As the formation of collective opinions and attitudes is highly dynamic, new developments need to be detected and trends to be analysed. Emotions and sentiment play an important part in actions such as voting, buying, supporting, donating and in collective opinion formation. Social intelligence does not just analyse but also support collective deliberation processes. Precise and robust multilingual technologies are needed to support discussion and deliberation processes on an international scale.
3. **Socially Aware Interactive Assistant** – Socially aware interactive assistants are conversational agents realized with or without a physical shell (from robots to different types of graphical or voice interfaces). Their behaviour leverages from the combination of analysis and synthesis of non-verbal, speech and semantic signals. It is the proper time to develop, implement and deploy socially aware and also multilingual assistants that can support and enhance the interaction of humans with their environment. This includes classical Human-Computer Interaction, Human-Artificial Agent (or robot) Interaction, and Computer-mediated Human-Human Interaction. Those assistants must be able to act in indoor environments (such as meeting

rooms, offices, apartments), outdoor environments (streets, cities, transportation, roads) and virtual environments, and also be able to communicate, exchange information and understand the other agents' intentions. They must be suitable and/or able to adapt to the user's needs and environment. They must have the capacity to learn incrementally from all interactions and other sources of information. The ideal socially aware multilingual assistant can interact naturally with humans, in any language and in any communication modality, it can adapt and be personalized to individual communication abilities, it can recognize and generate speech incrementally and fluently.

From the short descriptions of the three priority themes one can easily see that the proposed research strands overlap in technologies and challenges. This intended overlap reflects the coherence and maturation of the field. At the same time, the division of labour and sharing of resources and results is a precondition for the realization of this highly ambitious research programme.

All three areas need to benefit from progress in core technologies of human language analysis and production such as morphological, syntactic and semantic parsing and generation. But each of the three areas will concentrate on one central area of LT: the Translation Cloud will focus on cross-lingual technologies such as translation and interpretation, the Social Intelligence strand will take care of knowledge discovery, text analytics and related technologies, and the research dedicated to the Interactive Assistant will take on interface technologies such as speech and multimodal interfaces (see Figure 2).

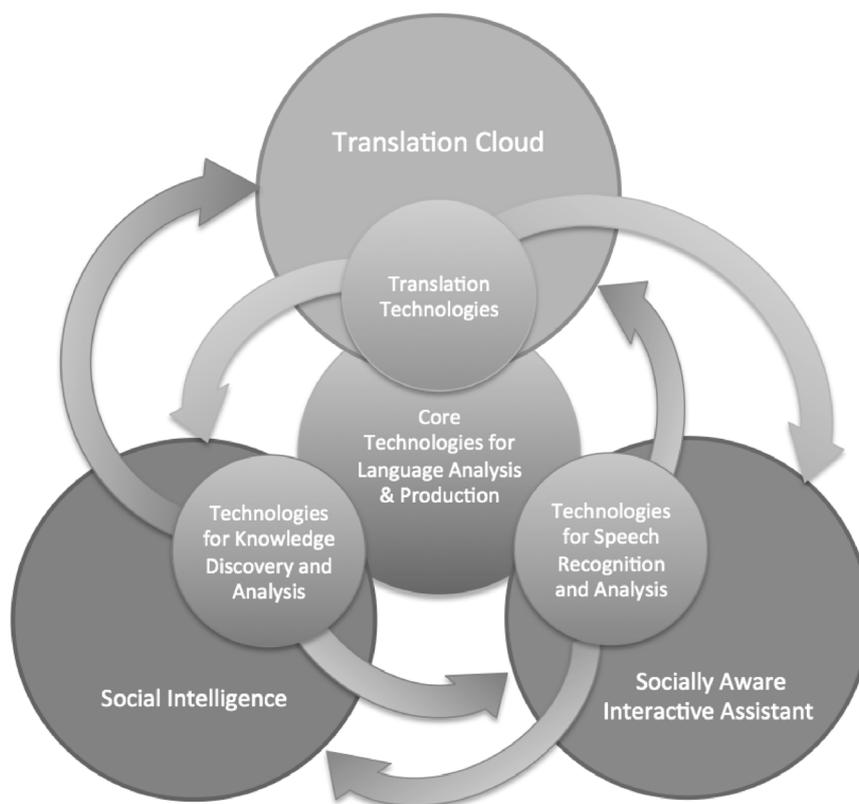


Figure 2: The three priority themes proposed by the META-NET Strategic Research Agenda and topics for scientific cooperation among the themes

In addition to these three priority themes the Strategic Research Agenda contains lists of technology and application visions, plans, suggestions for the organisation of research and roadmaps for the path to a truly multilingual Europe, realised through high-performance, robust and precise Language Technology. The SRA is currently being finalised and will be presented at the META-FORUM 2012 conference in Brussels in June.

5. Concluding remarks: Get involved and participate

With a large and diverse community behind our goals, META-NET and META can achieve the critical mass needed to really make a difference as to how Language Technology can enable and secure multilingualism in Europe's future (European Commission 2008; Directorate-General of the UNESCO 2007). To researchers, technologists, professionals and administrators developing, providing or using language technologies and also to the European language communities the Multilingual Europe Technology Alliance (META) offers a unique opportunity to stay informed, contribute ideas or advocate on behalf of specific languages, while participating in expert discussions, working groups and planning activities that will shape Europe's linguistic future. META is an open and growing technology alliance that currently has more than 300 members including multiple research centres and universities, companies that develop and provide as well as companies that make use of language technologies and also many organisations that represent Europe's language communities such as, for example, the Institute of the Lithuanian Language, Dansk Sprognævn and also EFNIL, the European Federation of National Institutions for Language.

Research and technology development projects are invited to join META-SHARE to access a pool of language resources and technologies while helping to validate and further to shape its services. Commercial enterprises are welcome to contribute their visions for products and services, to participate in our planning process and to use META to grow profitable partnerships. Schools and educators, journalists and the media, politicians, public institutions and organisations are encouraged to participate in open discussions on the vision of and way towards a truly multilingual information society.

Your voice is important, just like the language in which you express yourself. In joining our open technology alliance META and spreading the word, you'll be helping to shape the future of the European linguistic and also language technology landscape. Interested companies, research centres, institutions, organisations and individuals can join META without any financial obligations through a simple registration form: www.meta-net.eu/join.

6. Acknowledgements

This article is based on joint work with our colleagues Aljoscha Burchardt, Kathrin Eichler, Tina Klüwer and Felix Sasaki (all DFKI), the 57 member organisations of the META-NET network of excellence, the ca. 70 members of the Vision Groups, the ca. 30 members of the META Technology Council and the more than 160 authors of and contributors to the META-NET Language White Papers.

The META-NET Network of Excellence is co-funded by the 7th Framework Programme of the European Commission through the following grant agreements: T4ME Net (no. 249119), CESAR (no. 271022), METANET4U (no. 270893) and META-NORD (no. 270899).

More information on META-NET and META is available at www.meta-net.eu and via office@meta-net.eu.

7. References

- Burchardt, A./Rehm, G. (eds.) (2010): *Vision Group Translation and Localisation – Results of first two meetings*. META-NET. Internet: www.meta-net.eu/vision/reports/VisionGroup-TranslationLocalisation-draft.pdf.
- Burchardt, A./Rehm, G./Sasaki, F. (eds.) (2011): *The Future European Multilingual Information Society – Vision paper for a strategic research agenda*. META-NET. Internet: www.meta-net.eu/vision/reports/meta-net-vision-paper.pdf.
- Burchardt, A./Rehm, G./Uszkoreit, H. (eds.) (2012): *LT 2020 – vision and priority themes for language technology research in Europe until the year 2020. Towards a draft of the META-NET Strategic Research Agenda*. META-NET.
- Directorate-General of the UNESCO (2007): *Intersectoral mid-term strategy on languages and multilingualism*. Paris: UNESCO. Internet: <http://unesdoc.unesco.org/images/0015/001503/150335e.pdf>.
- Directorate-General for Translation of the European Commission (2009): *Size of the Language Industry in the EU*. Internet: <http://ec.europa.eu/dgs/translation/publications/studies>.
- European Commission (2008): *Multilingualism: An asset for Europe and a shared commitment*. Internet: http://ec.europa.eu/languages/pdf/comm2008_en.pdf.
- Federmann, C./Georgantopoulos, B./Girardi, C./Hamon, O./Mavroeidis, D./Minutoli, S./Schröder, M. (to appear 2012): META-SHARE Version 1: An Open Network of Repositories for Language Resources including Data and Tools. To appear in: *Proceedings of the Eighth International Conference on Language Resources and Evaluation – LREC 2012, 23-25 May, Istanbul, Turkey*.
- Gavriliidou, M./Labropoulou, P./Desipri, E./Piperidis, S./Monachini, M./Frontini, F./Declerck, T./Francopoulo, G./Arranz, V./Mapelli, V. (to appear 2012): The META-SHARE Metadata Schema for the Description of Language Resources. To appear in: *Proceedings of the Eighth International Conference on Language Resources and Evaluation – LREC 2012, 23-25 May, Istanbul, Turkey*.
- Koutsombogera, M./Piperidis, S. (eds.) (2010): *Vision Group Media and Information Services – Results of first two Meetings*. META-NET. Internet: www.meta-net.eu/vision/reports/VisionGroup-MediaInformationServices-draft.pdf.
- Mariani, J./Magnini, B. (eds.) (2010): *Vision Group Interactive Systems – Results of first two Meetings*. META-NET. Internet: www.meta-net.eu/vision/reports/VisionGroup-InteractiveSystems-draft.pdf.
- Piperidis, S. (to appear 2012): The META-SHARE Language Resources Sharing Infrastructure: Principles, Challenges, Solutions. The META-SHARE Metadata Schema for the Description of Language Resources. To appear in: *Proceedings of the Eighth International Conference on Language Resources and Evaluation – LREC 2012, 23-25 May, Istanbul, Turkey*.

Rehm, G./Uszkoreit, H. (2011): Multilingual Europe: A challenge for language tech. In: *Multilingual* 22(3) (April/May), 51-52.

Rehm, G./Uszkoreit, H. (eds.) (2012): META-NET White Paper Series. Heidelberg: Springer.

8. Addendum: Current composition of META-NET Network of Excellence

Country	Member (Affiliation)	Contacts
Austria	Universität Wien	Gerhard Budin
Belgium	University of Antwerp	Walter Daelemans
	University of Leuven	Dirk van Compernelle
Bulgaria	Bulgarian Academy of Sciences	Svetla Koeva
Croatia	Zagreb University	Marko Tadic
Cyprus	University of Cyprus	Jack Burston
Czech Rep.	Charles University in Prague	Jan Hajic
Denmark	University of Copenhagen	Bolette Sandford Pedersen, Bente Maegaard
Estonia	University of Tartu	Tiit Roosmaa
Finland	Aalto University	Timo Honkela
	University of Helsinki	Kimmo Koskenniemi, Krister Linden
France	CNRS, LIMSI	Joseph Mariani
	ELDA	Khalid Choukri
Germany	DFKI	Hans Uszkoreit, Georg Rehm
	RWTH Aachen	Hermann Ney
	Saarland University	Manfred Pinkal
	University of Stuttgart	Jonas Kuhn, Hinrich Schütze
	Karlsruhe Institute of Technology	Alex Waibel
Greece	ILSP, R.C. "Athena"	Stelios Piperidis
Hungary	Hungarian Academy of Sciences	Tamás Váradí
	Budapest Technical University	Géza Németh, Gábor Olaszy
Iceland	University of Iceland	Eiríkur Rögnvaldsson
Ireland	Dublin City University	Josef van Genabith
Israel	Bar-Ilan University	Ido Dagan
Italy	Consiglio Nazionale Ricerche	Nicoletta Calzolari
	Fondazione Bruno Kessler	Bernardo Magnini
Latvia	Tilde	Andrejs Vasiljevs
	University of Latvia	Inguna Skadina
Lithuania	Institute of the Lithuanian Language	Jolanta Zabarskaitė
Luxembourg	Arax Ltd.	Vartkes Goetcherian
Malta	University of Malta	Mike Rosner
Netherlands	Universiteit Utrecht	Jan Odijk
	University of Groningen	Gertjan van Noord

Country	Member (Affiliation)	Contacts
Norway	University of Bergen	Koenraad De Smedt
	University of Oslo	Stephan Oepen
Poland	Polish Academy of Sciences	Adam Przepiórkowski
	University of Łódź	Barbara L.-Tomaszczyk
	Adam Mickiewicz University	Zygmunt Vetulani
Portugal	University of Lisbon	Antonio Branco
	Institute for Systems Engineering and Computers	Isabel Trancoso
Romania	Romanian Academy of Sciences	Dan Tufis
	University Alexandru Ioan Cuza	Dan Cristea
Serbia	Belgrade University	Dusko Vitas, Cvetana Krstev
	Pupin Institute	Sanja Vranes
Slovakia	Slovak Academy of Sciences	Radovan Garabik
Slovenia	Jozef Stefan Institute	Marko Grobelnik
Spain	Barcelona Media	Toni Badia
	Technical University of Catalonia	Asunción Moreno
	University Pompeu Fabra	Núria Bel
	University of the Basque Country	Inma Hernaez Rioja
	University of Vigo	Carmen García Mateo
Sweden	University of Gothenburg	Lars Borin
Switzerland	Idiap Research Institute	Hervé Bourlard
UK	University of Manchester	Sophia Ananiandou
	University of Edinburgh	Steve Renals
	University of Wolverhampton	Ruslan Mitkov
	University of Sheffield	Rob Gaizauskas

Table 5: Current composition of the META-NET Network of Excellence (Spring 2012)