LT 2020

Vision and Priority Themes for Language Technology Research in Europe until the Year 2020

Towards the META-NET Strategic Research Agenda

Do you have comments, ideas or suggestions with regard to the content of this document? Please send them to office@meta-net.eu or discuss them online: http://www.meta-net.eu/sra.

The development of this paper has been funded by the Seventh Framework Programme and the ICT Policy Support Programme of the European Commission under contracts T4ME (Grant Agreement 249119), CESAR (Grant Agreement 271022), METANET4U (Grant Agreement 270893) and META-NORD (Grant Agreement 270899).

Contents

1 Introduction to this Excerpt

For decades, it has been obvious that one of the last remaining frontiers of information technology is still separating our rapidly evolving technological world of mobile gadgets, PCs and Internet from the most precious and powerful asset of mankind, the human mind, the only system capable of thought, knowledge and emotion. Although we use the PC to write, the telephone to chat and the Web to search for knowledge, information technology as such has no access to the meaning, purpose and sentiment behind our trillions of written and spoken words. This is why it is unable to summarize a text, answer a question, respond to a letter and to translate reliably. In many cases it cannot even correctly pronounce a simple English sentence.

Visionaries such as Ray Kurzweil, Marvin Minsky and Bill Gates have long predicted that this border would eventually be overcome by artificial intelligence including language understanding whereas science fiction such as the Star Trek TV series suggested attractive ways in which the increased power of the technology would change our lives. Visualized fiction more than technical prophecies surely extended the imagination of an entire generation into the right direction by establishing the fantastic concept of an invisible computer that you have a conversation with and that is able to react to the most difficult queries by correct replies and also of technology that can reliably translate any human and non-human language.

Many enterprises had started much too early to invest in language technology research and development and then lost faith again after some long periods without any tangible progress. However, during the years of apparent technological standstill, research continued to conquer new ground. The results were: a deeper theoretical understanding of language, better machine-readable dictionaries, thesauri and grammars, specialized efficient language processing algorithms, hardware with greater computing power and storage capacities, large volumes of digitized text and speech data and, most importantly, powerful new methods of statistical language processing that could exploit the language data for learning hidden regularities governing our language use.

Lately Google's web search, Autonomy's text analytics, Nuance's speech technology, Google's online translation, IBM Watson's question answering and Apple Siri's personal assistance have given us but a glimpse of the massive potential behind the evolving language technologies. Leading-edge industry has already reacted, but this time much more decisively. IBM, SAP, Apple, Google, Amazon, Nokia, Nuance, Facebook and others have started acquiring language technology enterprises left and right, among them many small promising start-up companies.

However, all stakeholders know that today none of the industrial players possesses the needed know-how for unleashing the full potential of language technology for business and society since essential research results are still missing. Nevertheless, the speed of research keeps increasing and even small improvements can already be exploited for innovative products and services that are commercially viable.

Even if Google, Apple, Microsoft, IBM and Nuance seem to be far ahead of all competitors in language technology development, any breakthrough in one of the subareas of LT research can rapidly change this picture. New results are coming in from numerous places, among them many European universities and research organisations. Actually, because of a long history of basic-research funding and a lively LT industry of mainly sophisticated SMEs, the European LT scene is very well positioned in the race for the needed breakthroughs. To a large part, the strong European LT landscape is a direct consequence of the commercial and social interests stemming from the multilingual setting of the integrating Europe. As strange as it may sound, another advantage results from the well-known lack of pragmatism in European science. Where most of collaborative US research had abandoned computationally challenging knowledge-based or mathematically sophisticated approaches to linguistic and semantic processing, the typical European quest for truth and insight coupled with a sufficient level of funding for fundamental research became the basis for continued output of high-potential novel approaches and techniques. As a side effect, many of the pioneers in US industrial LT research and development as well as many industrial providers of technologically supported language services all over the world are of European origin.

Would this pole position in international research together with an appropriate increase in funding be sufficient for enabling Europe to profit substantially from the next revolution in IT? We have good reason to be sceptical; after all, Europe's track record in successful commercialization of advanced IT has not improved much since the loss of countless opportunities. Nevertheless, we strongly believe that LT offers a chance for a true success story because of several factors:

- A lively and diverse industrial landscape of sophisticated language technology and service providers,
- An alliance of research communities, language industries and other stakeholders behind a shared vision and long-term research programme,
- Large public users such as the translation and web-content services of the European Union and other national and European organizations,
- A deep cultural and political interest in the preservation and support of all European languages in the digital society,
- Access to immense volumes of multilingual data.

Neither politically nor economically can Europe afford to miss the boat on the foreseeable next revolution in IT because we must not leave the decisions on the degree of technological support of our languages and the cultural, social, scientific and economic exploitation of the aggregated digital contents of the world entirely to others. The danger and also threat of not taking this opportunity is real seen how our sometimes well founded and sometimes half-hearted hesitation has put us on the international back bench in future technologies such as genetics or solar energy.

In the following sections we summarize our shared vision of the role of language technology in the year 2020 in non-technical terms and outline three priority themes for large-scale research and innovation. The three thematic directions have been designed with the aim of turning the joint vision into reality and to letting Europe benefit from a technological revolution that will overcome barriers of understanding between people of different languages, between people and technology and between people and the accumulated knowledge of mankind.

The three priority themes described in this excerpt form the heart of the META-NET Strategic Research Agenda (SRA). In the SRA, they build the bridge between societal needs, Language Technology applications, and concrete roadmaps for the organization of research, development and scientific innovation. The three themes (see **Error! Reference source not found.**) as presenter here are self-contained and will be interlinked more tightly in future versions of the SRA.

The priority themes are contextualized in the advanced networked society and cover the main functions of language: storing, sharing and using of information and knowledge, as well as improving social interaction among humans and enabling social interaction between humans and technology. As multilingualism is at the core of European culture and becoming a global norm, one theme is devoted to overcoming language barriers.

Contributors

Jan Hajic, Philipp Koehn, Hans Uszkoreit, Josef van Genabith, Marcello Federico, Lucia Specia, Gabor Proszeky. (Theme 1)

Philipp Cimiano, Rob Gaizauskas, Marko Grobelnik, Bernardo Magnini, Stephan Oepen, Rudi Studer, Ido Dagan. (Theme 2)

Maja Pantic, Hervé Bourlard, Alex Waibel, Kristiina Jokinen, Giuseppe Riccardi, Steve Renals, Joseph Mariani. (Theme 3)

Editors

Aljoscha Burchardt, Georg Rehm, Hans Uszkoreit

2 Language Technology 2020 - A Plain Summary of the META-NET Technology Vision

People store and exchange information using the language they have known since early childhood. Computers were created for processing information but for a long time they were totally ignorant about the languages of their masters. It took a while until computers could handle scripts of languages different from English. It took even longer until computers could check the spelling of texts and read them aloud for the visually impaired.

On the web we can now get rough translations and we can search for texts containing a word, even if the word occurs in the text in a different form such as in plural number or in the genitive case. But when it comes to interpreting, actually making sense of certain input, and correctly responding, computers only "understand" simple artificial languages such as Java, PHP, Python, C++ and HTML.

After the next IT revolution, computers will master the languages of their users. Just as measures and formats for dates and times, the operating systems of tomorrow's IT, may they reside on the local CPU or in some cloud, will know human languages. They may not reach the linguistic performance of educated people and they will not yet know enough about the world to understand everything, but they will be much more useful than today and further enhance our work and life.

2.1 Communication among People

Since language is our most natural medium for interpersonal communication, computers cannot help much in regular conversations. However, when the communication partner speaks a *different* language this situation changes. With thousands of languages spoken on our planet, chances are high that we do not understand our partner. Rudimentary speech translation has been successfully demonstrated for limited numbers of languages and themes. By the year 2020, reliable and robust dialogue translation for face-to-face conversation and telecommunication can be achieved at least for hundreds of languages if research concentrates sufficient efforts on solving the problems of high-quality automatic translation and robust accurate speech recognition. Some technology experts predict this technology for 2013 or 2015, but these forecasts are too optimistic. This does not mean that there could not be useful new speech translation products for selected languages and purposes between now and 2020.

In written communication, we use the computer as a tool for producing texts as well as for reading – from short emails or instant messages to novels or complex technical documents. The machine checks spelling and grammar and its thesaurus suggests alternatives for words. LT products are already successfully employed in enterprises for checking conformance to corporate terminologies and style guidelines. In 2020 general authoring software will also check for appropriate style according to genre and purpose including comprehensibility. It will not only flag potential errors but also suggest appropriate corrections. On demand it will explain the errors to language learners. It will employ large authoring memories to proactively suggest completions of started sentences or whole paragraphs.

Today, Google Translate and other large online translation services provide affordable access to information and knowledge for hundreds of millions of users who do neither speak English nor any other of the languages that together make up most of the global web content. The technology is extremely important for personal use and for numerous professional applications, e.g., intelligence jobs at which analysts have to search through masses of texts for relevant information nuggets. However, automatic translations are still far away from the quality standards needed for a true impact on the translation and globalization markets. Although the European Commission uses similar technology provided by European research projects, the translations in their current quality can only be used internally, which is great progress but does not yet help with the skyrocketing costs for outbound translation. Many large and smaller translation services have started using machine translation but the economic breakthrough is still ahead of us. It will come in stages over the next ten years when the existing barriers for quality are overcome by new translation technologies that get closer to the structure and meaning behind human language.

In 2020 affordable instant high-quality translation for numerous domains and genres will be available among hundreds of languages if the necessary big push in research and innovation that we are proposing is implemented. We will be able to access such services online from any place for written as well as for spoken language.

At work, much of collective planning and controlling takes place in meetings. By 2020 many meetings will be tele-meetings utilizing very large displays and comfortable presentation technology. LT will be able to record and transcribe face-to-face and virtual meetings. It will produce drafts and also summaries of minutes. For both types of meetings, it will simultaneously translate (interpret) the contributions of participants into as many languages as needed. The incrementally drafted records and summaries will be used for displaying the state of the discussion including intermediate results and open issues. This software will be guided by partial understanding of the contents, i.e., by their semantic association with concepts in semantic models of domains and processes. Brainstorming will be facilitated by semantically driven automatic lookup and structured display of relevant data, proposals, charts, pictures and maps.

Finally, LT will also be used massively for helping with the ever-growing volume of correspondence. It will actively help to draft messages through automatic authoring techniques. Today many businesses and other organisations already employ e-mail response management software to filter, sort and route incoming email and to suggest replies to recognized types of requests based on earlier responses. By 2020, the business email communication will be embedded in semantically structured process models automating much of standardised communication. Already before 2020, email communication will be semantically analysed, checked for sentiment indicators and summarized in reports.

LT will also help to integrate the contents of all communication channels: telecommunication, reports, meetings, email and chat, among others. Semantic integration into the work processes, threading and response management will be applied across channels. Machine translation and analytics will be available for all communication channels.

The extremely popular and powerful Web 2.0 mechanisms – social networks and user-generated content – have confronted language technology with a new set of challenges. Every user can become a content producer and large numbers of people can participate in communications. This is interpersonal communication with large numbers of participants in highly dynamic networks. Some of these many-way mass communications have turned into effective instruments for support solicitation, idea creation, opinion formation and solution search. Communities can emerge in a matter of hours or days around admired works of art, shared preferences or social issues. Citizen action movements, international NGOs, patient self-help groups, expert circles and communities of concerned consumers can organize mutual support schemes, arrive at optimal and broadly supported social solutions and exert pressure on decision makers.

However, the emerging social web cannot yet unfold its true potential because the large volumes of user-generated content become intransparent and unmanageable in no time. Both for direct participants and for outside stakeholders or concerned decision makers, it requires considerable efforts to stay on top of new developments. Much of the often-cited wisdom of the crowds and quite a bit of the aggregated motivation is wasted because of information overflow. Language technology can and will eventually harness the inevitable information deluge resulting from active Web 2.0 communities and discussions. If dedicated research efforts are focussed and well supported as suggested in one of our three priority themes, technologies will be in place by 2020 that monitor, analyse, summarize, structure, document and visualize social media dynamics including many-way mass communication. Democracy and markets will be enriched by new powerful mechanisms for improved collective solution development and decision making based on massive involvement of stakeholder and expert communities.

Another important role of LT in interpersonal communication is the automatic conversion of language between different modes. Early examples are dictation systems and text-to-speech tools that convert between spoken and written language. Such conversions are also needed for providing persons with disabilities better access to communication. These technologies are already commercially successful but within the next few years they will reach full maturity opening up much larger markets. They will be complemented by reliable conversion from spoken or written language into sign language and vice versa. LT will also be utilized for improved methods of so-called supported communication and for conversion of everyday language into strongly simplified language for special types of disabilities. All these technology-supported means for compensating impairments are often subsumed under the term augmentative alternative communication. As this is a rather young field of research, much progress can be expected between now and the year 2020 so that improved language technology will enable many more people with disabilities to participate in social life.

2.1 Communication with technology and the rest of the world

Through language technology, human language will also become the paramount medium for communication between people and the rest of the world. Today's voice-control interfaces to smart-phones and the query fields of search engines are just the modest beginning of overcoming the communication barrier between humankind and the non-human part of the world.

This world consists of plants, animals and natural as well as man-made objects. The realm of artefacts ranges from small simple objects via all kinds of technical devices such as machines, appliances and vehicles all the way to highly complex units such as robots, airplanes, buildings, traffic systems and cities. But the artificially created world also consists of information and knowledge contained in books, films, sound recordings and digital storage. Virtually all information and knowledge will soon be available in digital form, even if other media continue to coexist. The volume of data and thus potential information created daily about all parts of our world keeps increasing at a fast rate. The result is a gigantic distributed digital model of our world, let's call it *second world*, which continuously gains in complexity and fidelity. Through massive networking of this information by meta-information services and the linking of open data, the second world is getting more useful as a resource for information, planning and knowledge creation.

Today we still have a rather clear distinction between intelligent beings, i.e., humans, artificial agents with some autonomous behaviour and all other kinds of objects. We can easily communicate with people and we would like to communicate with computers and robots. We do not feel a pressing need to speak with a cup or with a power drill. However, the situation keeps changing fast since more and more products come equipped with sensors, processors and some information services such as descriptions, specifications or manuals. Many of these objects are even connected to the internet (Internet of Things) or at least represented with passive or active information services on the web (Web of Things). Thus, eventually, we can and will communicate with such objects.

Depending on the function, complexity, relevance and relative autonomy of the artefacts, the nature of communication with artefacts can strongly vary. Some objects will come with interesting information, often represented in the second world, that we would like to query and explore (such as, for example, user and maintenance manuals, historical digests and consumer information). Other objects will provide information on their state and will also have their own individual memory that can be queried. Objects than can perform actions such as vehicles and appliances will accept and carry out voice commands.

One may well find the vision of communicating with dead matter to be resting on a rather questionable concept of communication, since real communication requires at least two agents with a certain degree of autonomy and different information states. However, if we communicate with an in-car entertainment system or with an autonomous vehicle, does it matter how the computer representing the artefact's information state is connected to the vehicle? If the computer were not part of the moving vehicle but connected to it via a wireless network, we would still call it communication.

The real question is whether we really want to talk to machines, buildings and other objects. By our evolution and culture we are used to communicating with people – would the microwave need at least a face with a mouth in order to be accepted as a conversation partner? This could easily be provided by a dedicated interface displayed either on the object or, more likely, on our mobile device. In a few years the anthropomorphic interface may well be a 3D projection appearing next to the object.

We must be careful not to mix up the interface with the objects we are interested in. Recently the old concept of a personal digital assistant has gained much popularity because of the successful launch of Siri on the iPhone 4S. In the near future, much more sophisticated virtual characters will follow equipped with expressive voices, faces and gestures. They will become an interface to any information that is provided on the web in the appropriate form. Thus this assistant could speak for or about machines, locations, the weather, the Empire State Building and the London Stock Exchange. The metaphor of a personal assistant is powerful and extremely useful, since such an assistant can be made sensitive to the user's preferences, habits, moods and goals. It can even be made aware of socio-emotional signals and processes and learn appropriate reactions from experience.

Realizing this ambitious vision will require a dedicated thoughtfully planned massive effort in research and innovation. By the year 2020 we could have a highly personalized, socially aware and socially interactive virtual assistant. Having been trained on the user's behaviour and educated from his digital information and communication space it will be proactive by offering valuable unrequested advice. Voice, gender, language and mentality of the virtual character could be adjusted to the user's preferences. The agent will be able to speak in the language and dialect of the user but also digest information in many other natural and artificial languages and formats. Because of these skills, the assistant can translate or interpret without the user even realizing it.

The existence of such a powerful personal assistant does not mean that all communication with the non-human world will have to go through one single digital helper. In fact, it is rather unlikely that all internet-based information, assistance and transaction services will interface to the assistant. Already today, many companies such as IKEA, Yello and HSBC employ specialized chatbots as access points to their information resources. Appearance, behaviour and style of these virtual agents have been carefully designed to meet the communication goals, intended product image and corporate identity of the enterprise.

In the future, many providers of information on products, services and touristic sites will try to present their information with a specific look and feel. The personality and functionality of the interface may also depend on the user-type; there may be special interfaces for children, foreigners and persons with disabilities. Command interfaces may need specialized functionalities anyway. Thus there will be space for many interfaces custom-tailored to the corporate identities of the providers or to the nature of the objects and services.

With the appropriate interfaces provided, people will not find it unnatural at all to communicate with their inanimate surroundings. For millennia our ancestors attributed souls and thought to objects. Smaller children still find it completely normal if objects speak to them and freely talk back. This is why stories about talking things have survived in fairy tales and children's books. Learning about the world will take on a completely new quality if the child can freely query objects and places.

By the year 2020 there will be a competitive landscape of intelligent interfaces to all kinds of objects and services employing language and other media such as manual and facial gestures for effective communication. Depending on the complexity of functionalities and provided information, the language coverage will range from simple commands to sophisticated dialogues. Many interface services will be offered as customizable cloud-based middleware others may be completely custom-tailored. The technologies needed for such dedicated interfaces to machines, objects and locations are all part of the socially aware virtual assistant so that our proposed priority theme also creates enabling technologies for other interface products.

Two large application domains are special in their demands and need for additional technologies: robotics and knowledge services.

Although stationary industry robots have already taken over large parts of industrial production, the real era of robots is still ahead of us. But within this decade, specialized mobile robots will be deployed for personal services, rescue missions, household chores and tasks of guarding and surveillance. Natural language is by far the best communication medium for natural human-robot interaction. Since human language is very elaborate when it comes to speaking about perception, motion and action in space and time, the interaction in the physical world poses enormous challenges to language technology. Some of these challenges can be addressed within the priority theme of the digital assistant, but without additional language technology research in the robotics area, the communication skills of robots will lag behind their physical capabilities for a long time. By 2020 we will have robots around us that can communicate with us in human language, but their user-friendliness and acceptance will largely depend on progress in the coming years of language technology research.

The communication with knowledge services raises a different set of problems. Here it is the inherent complexity of the represented knowledge that requires considerable advances in technology. The complexity arises from both the intricate structures of the subject domains and the richness of linguistic expressivity, in particular the great variety of options to implicitly or explicitly express the same fact or question. Moreover, most information that we can learn from a text is not encoded explicitly but stands "between the lines." For the human reader it follows from the text but for language technology it needs to be derived by applying reasoning mechanisms and inference rules along with large amounts of explicitly encoded knowledge about the world.

Our current methods for fishing in the ocean of knowledge through a web search engine is already much better than listing all the books in a library in which the query words occur because it ranks the found sources and provides access to each of them by a single click. Looking the information up in the selected sources is extremely helpful but much less efficient and effective than soliciting the needed information directly from a knowledgeable person, which would be the method of choice if this person happened to be sitting next to you.

From watching the crew of spaceship Enterprise in the famous TV series Star Trek, we expect that, eventually, we will be able to just say "Computer" followed by any question. As long as an answer can be found or derived from the accumulated knowledge of mankind, it will come back in a matter of milliseconds. In a spectacular US broadcast of the TV knowledge quiz Jeopardy, the computer giant IBM Watson was recently able to find correct answers that none of its human competitors could provide. Erroneously one may think now that the problem of automatic question answering is solved. Undoubtedly Watson is a great achievement demonstrating the power of language technology. But some of the questions that were too hard for the human quiz champions were actually rather easy for a machine that has stored handbooks, decades of news, lexicons, dictionaries, bibles, databases and the entire Wikipedia. With clever lookup and selection mechanisms for the extraction of answers, Watson could actually find the right responses without a full analysis of the questions or clues.

Most questions that people might ask cannot be answered by today's technology, even if it has access to the entire web, because they require a certain degree of understanding of both the question and the passages containing potential answers. However, research on automatic question answering and textual inferencing progresses fast and by 2020 we will be able to use internet services that can answer huge numbers of non-trivial questions.

One prerequisite of this envisaged knowledge access through natural communication are novel technologies for offline processing of large knowledge repositories and massive volumes of other meaningful data which will be discussed in the following subsection.

2.2 Processing Knowledge and Information

Most knowledge on the web, by far, is formulated in human language. However, machines cannot yet automatically interpret the texts containing this knowledge. Machines can interpret knowledge represented in databases but databases are too simple in structure to express complex concepts and their relations. The logical formalisms of semanticists, on the other hand, that were designed to cope with the complexity of human thought, were too unwieldy for practical computation. Therefore computational logicians developed simpler logical representation languages as a compromise between desired expressivity and required computability. In these languages, knowledge engineers can formulate formal models of knowledge domains, so-called ontologies, describing the concepts of the domains by their properties and their relations to other concepts. Ontologies enable knowledge engineers to specify which things, people, places in the world belong to which concepts. Such a domain model can be queried like a database. Its contents can be automatically analysed and modified. The intellectual creation of domain models, however, turned out to be an extremely demanding and time-consuming task, requiring well-trained specialists.

Their encoding of knowledge seemed to be a promising alternative to the current web, so that the vision of the Semantic Web was born.

The main bottleneck of the Semantic Web vision is the problem of knowledge engineering. It is unrealistic to believe that the authors of web content will be able to encode the relevant knowledge in the semantic web languages based on description logics. Nor will there be any affordable services for the manual conversion of large volumes of content.

Since language technology did not have any means for automatically interpreting texts, language technologists had developed methods for at least extracting relevant pieces of information from such texts. This technique became a useful extension to information retrieval, which enables users to find entire documents such as in Google search. A relatively simple information extraction task is the reliable recognition of all person and company names, time and date expressions, locations and monetary expressions, so-called named-entity extraction. Much harder is the recognition of certain relevant relations such as the one between company and customer, company and employee or inventor and invention. Even more difficult are many-place relations such as the four-place relation of a wedding between groom and bride at a certain data and time. Events are typical cases of relations. However, events can have many more components such as the participants, costs, causes, victims and circumstances of accidents. Although research in this area is advancing fast, a reliable recognition of relations is not yet possible.

Even if information extraction was created for the recognition of individual relations as needed for answering certain questions, or for filling databases with events of certain types, it can also be used for populating ontologies. Texts and pieces of texts can be annotated by extracted data. These metadata can serve as a bridge between the "semantic" portions of the web and the traditional web containing unstructured data. Therefore, language technology is indispensable for the realization of the vision of a semantic web.

Besides this type of semantification of the web through powerful metadata, language technology can perform many other tasks in the processing of knowledge and information. It can sort, catalogue and filter content for certain applications and it can deliver the data for data mining in texts, which has been termed text data mining. Language technology can automatically connect web documents with meaningful hyperlinks and it can produce summaries of larger collections of texts. The LT techniques of opinion mining and sentiment analysis can find out what people think about products, personalities or problems and analyse their feelings about such topics.

Another class of techniques is needed for connecting between different media in the multimedia content of the web. Some of the tasks are annotating pictures, videos and sound recordings with metadata, interlinking them with texts, semantic linking and searching in films and video content and cross-media analytics including cross-media summarization.

In the next few years we will see considerable advances for all these techniques. For large parts of research and application development, language processing and knowledge processing will merge. The most dramatic innovations will draw from progress in multiple subfields. The predicted and planned use of language and knowledge technologies for social intelligence applications, one of our three priority areas, will involve text analytics, translation, summarization, opinion mining, sentiment analysis and several other technologies. If the planned massive endeavour in this direction can be realized, it will not only result in a new quality of collective decision-making in business and politics. In 2020, language technology will enable forms of knowledge evolution, knowledge transmission and knowledge exploitation that speed up scientific, social and cultural development. The effects for other knowledge-intensive application areas such as business intelligence, scientific knowledge discovery and multimedia production will be immense.

2.3 Learning Language

Soon every citizen on Earth will learn a second language, many will learn a third and a few will go beyond this by acquiring additional languages. If this learning takes place after the period of early childhood it will be hard. Learning a language is very different from acquiring scientific knowledge because it requires repetitious practicing by actual language use. The more natural the use, the more effective the practice is.

IT products that help to ease and speed up language learning have a huge market. Already today, the software market for computer-assisted language learning (CALL) grows at a fast rate. Current products are helpful complements to traditional language instruction, however, they are still limited in functionality because the software cannot reliably analyse and critique the language produced by the learner. This is true for written language and even more so for spoken utterances. Software producers are trying to circumvent the problem by strongly restricting the expected responses of the user. This helps for many exercises but it still rules out the ideal interactive CALL application, which is an automatic dialogue partner ready around the clock and capable to error-free conversation on many topics, a software that furthermore analyses and critiques the learner's errors and adapts its dialogue to the learner's problems and progress. Contemporary language technology cannot yet provide such functionality.

This is the reason why research on CALL applications has not yet come into full bloom. As research on language analysis and understanding and on dialogue systems progresses, we predict a boom in research and development in this promising and commercially attractive application area. Research toward the missing technologies is covered by our priority themes. We expect a strong increase in CALL applications research at some time between 2015 and 2020.

2.4 Learning Through Language

Since most K-12, academic and vocational instruction happens through language, spoken in classroom and read in textbooks, language technology can and will play a central role in learning. Currently LT is already applied at a few places in the preparation of multiple-choice tests and in the assessment of learners' essays.

As soon as human-language dialogue systems can robustly conduct nearly error-free dialogues based on provided knowledge, research can design ideal tutoring systems. But long before LT research will reach this point, we will be able to create systems that test for knowledge by asking questions and that provide knowledge to the learner by answering questions. Thus even adaptive loops of analytic knowledge diagnosis and customized knowledge transmission as they form the core of an effective learning system will become possible through language technology. Knowledge structuring and question answering is covered by our priority themes. The transfer to research and development toward educational applications should happen through close cooperation with the active and highly dynamic research scene in e-learning.

Although we do not expect to substitute human teachers by 2020, we predict that e-learning technology will have become much more effective and learner-friendly by that time through the integration of advanced language technology.

2.5 LT for Creative Contents and for Creative Work

One of the major cost factors in European TV and film production is the required subtitling and dubbing. Whereas some countries with multiple official languages or with strict legislation on subtitling or sign-language display for the hearing-impaired have a long tradition in providing these services, producers in many other countries still leave all subtitling and dubbing to importing distributors or media partners. With a single common digital market and the increase of productions for multiple language communities and with the strengthening of inclusion policies, the demand for fast and cost-effective subtitling and dubbing will grow. In some countries, the cheaper method of voice-over is widely used. A professional voice talent reads all translations, sometimes shortened, over the original sound track. This method is common in Eastern Europe and occasionally in some Asian countries.

In recent research very promising results could be obtained for automatic subtitling and subtitle translation. It turned out that the automatic translation of subtitles is easier than the translation of newspaper articles because of shorter and simpler sentences in spoken language. Some commercial services have already started using machine translation for subtitles and audio-description. If monolingual subtitling becomes the norm demanded by the law, the automation of subtitle translation could be deployed at large scale.

Open challenges are the automatic production of sign-language translations and dubbing. Especially automatic dubbing will be a hard task for speech technology since it requires the interpretation of the intonation in the source language and the generation of the adequate intonation in the target language. An easier method would be automatic voice-over for appropriate material and markets. In 2020 we will see wide use of automatic subtitling and first successful examples of automatic dubbing for a few languages.

Language can also be a medium for creative work, not only in literature. In traditional fine arts, creation mainly happens by a direct production of visual objects or images in two or three-dimensional space through drawing, sculpting, constructing, painting or photographing. In creative writing, the creation happens in language. But in many other areas of creative work, the creation happens through languages, ranging from musical notation to programming languages. Here the created work is specified in some suitable notation. Often natural language is used, so for instance in the formulation of storyboards and scripts for movies or in the design of processes, services or performances.

In computer science, the idea of writing programmes in natural language is almost as old as programming itself. This approach would require the translation of natural language into some programming language or assembler language. However, the inherent ambiguity and vagueness of natural language has remained a major problem. Another obstacle is the richness of language, i.e., there are often too many ways to express the same statement. Even if we could implement a system that would correctly translate a subset of a natural language into computer programs, how would one specify and memorize this subset?

Instead of restricting natural language through the definition of a suitable subset, computer scientists have created a number of easily learnable interpreted languages, often so-called scripting languages, whose syntax resembles simple sentence structures of English. The idea of natural language programming has recently received renewed attention because of the concept of ontology-assisted programming. The natural language statements are interpreted with respect to an ontology. We expect that the general concept of programming in natural language will bear fruit through progress in the semantic interpretation of natural language with respect to formal ontologies. Natural language may never become the programming language of choice for professional programmers who might always prefer the brevity and clarity of specifically designed artificial languages, but we will certainly witness means for specifying scripts and simple programs in natural language for the everyday computer user.

The ontology-based interpretation of natural language statements will also permit the specification of processes, services, objects which will then be automatically translated into formal descriptions and finally into actions, models, workflows or physical objects. By 2020 we can expect successful examples of natural language scripting and specification in a few suitable application areas.

2.6 Diagnosis and Therapy

Because of the central role of language in human life, psychological and medical conditions affecting language use belong to the most severe impairments people can suffer from. Deficiencies in language can also be strong indicators for other conditions that are harder to detect directly such as damage to brain, nerves or articulatory system. Language technology has been utilized for diagnosing the type und degree of brain damage after strokes. Since the administration of diagnosis and therapy are time-critical for a successful recovery of brain functions, permanently available software can support the immediate detection and treatment of stroke effects.

Language technology can also be applied to the diagnosis and therapy of aphasia resulting from other causes than strokes, e.g., from infections or physical injuries. Another application area is the diagnosis and therapy of innate or acquired speech impairments, especially in children.

Dyslexia is a widespread condition affecting skills in reading and orthography. Some effects of dyslexia can be greatly reduced by appropriate training methods. Recent advances in the development of software for the therapy of dyslexia give rise to the hope that specialized CALL systems for different age groups and types of dyslexia will eventually help to treat this condition early and effectively.

Technologies for augmentative alternative communication referred to in the subsection on interpersonal communication can also perform an important function in therapy since any improvement of communication for language-impaired patients opens new ways for the treatment of causal or collateral conditions. Expected progress in language technology together with advances in miniaturisation and endo- and exoprosthetics will certainly open new ways for helping people who cannot naturally enjoy the benefits of communication.

2.7 Language Technology as a Key-Enabling Technology

The wide range of novel or improved applications explicitly mentioned in our shared vision only represent a fragment of the countless opportunities for language technology to change our work and every life. Language-proficient technology will enable or enhance applications wherever language is present. It will change the production, management and use of patents, legal contracts, medical reports, recipes, technical descriptions, scientific texts and it will permit many new voice applications such as automatic services for the submission of complaints and suggestions, for accepting orders and for counselling in customer-care, e-government, education, community services

With so many applications and application areas, each of them confronted with different functionalities and types of language, we may be tempted to doubt that there is a common technology core. And indeed there has been a trend of excessive diversification in language technology software development. Many language analysis tools can only be used for one purpose. This is different from the way humans their language. Once we have learned our mother tongue we can easily obtain new skills, always employing the core knowledge acquired during childhood. We learn to read, to write, to skim texts, to summarize, to outline, to proof-read, to edit and to translate.

Currently we are witnessing a promising trend in language technology giving rise to hope for faster progress. Instead of relying only on highly specialized components, powerful core technologies are reused for many applications. We can now compose lists of processing components and tools that we need for every language since these will be adapted for and integrated into many applications.

In addition to these core processing technologies, we have also identified lists of core data, such as large text and speech corpora, and language descriptions, such as lexicons, thesauri and grammars, that are needed for a wide spectrum of purposes.

In information technology, we can differentiate between specialized application technologies, such as credit-card readers, and enabling technologies, such as microprocessors, that are needed for rather diverse types of applications. In hardware technology, certain key-enabling technologies have been identified, technology areas indispensable for projected essential progress. Among those are nanotechnology, microelectronics including semiconductors, advanced materials, biotechnology and photonics. Similar key-enabling technologies also exist on the software side of IT, such as database technology or network technology. Considering the broad range of language-technology-enabled applications and their potential impact on business and the entire society, language technology is certainly becoming a key-enabling technology for future generations of information technology. In contrast to some of the key-enabling technologies listed above, Europe has not lost yet a leadership role in this field. There is no reason to be discouraged or even paralyzed by the strong evidence of interest and expertise on the side of major commercial players in the US. In software markets the situation can change fast. Many of today's big players have

If Europe does not take a decisive stand for a substantial commitment to language technology research and innovation in the years to come, we may as well give up any ambition in future information technology altogether because there is no other software sector, in which European research can benefit from a similar combination of existing competitive competence, recognized economic potential, acknowledged societal needs and determined political obligation toward our unique wealth of languages.

3 Priority Theme 1: Translation Cloud

3.1 Solutions for the EU Society and for the Citizen

The goal 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. This will be a world where written and spoken communication is not hindered anymore by language boundaries and where even 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 simple 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 service and technology providers.

Special components and extensions of the permanent and ubiquitous service are:

- use and provision platform for providers of computer-supported creative top-quality human translation, multilingual text authoring and quality assurance by experts
- trusted service centers: certified service providers fulfilling highest standards for privacy, confidentiality and security of source data and translations
- quality upscale models: services permitting instant quality upgrades if the results of the requested service levels does not yet fulfil the quality requirements
- domain and task specialization models:
- translingual spaces: dedicated locations for ambient interpretation. Meeting rooms equipped with acoustic technology for accurate directed sound sensoring and emission

3.2 Novel Research Approaches and Targeted Breakthroughs

The core reason why HQMT has not been systematically addressed yet seems to be the Zipfian distribution of issues in MT: some improvements, the “low-hanging fruit” can be harvested with moderate effort in a limited amount of time. Yet, many more resources and a more fundamental, novel scientific approach -- that eventually runs across several projects and also calls -- are needed for significant and substantial improvements that cover the phenomena and problems that make up the Zipfian long tail. This is an obstacle in particular for individual research centres and SMEs given their limited resources and planning horizon.

Although recent progress in machine translation has already led to many new applications of this technology, radically different approaches are needed to accomplish the ambitious goal of this research including a true quality breakthrough. Among these new research approaches are:

- Systematic concentration on quality barriers, i.e. on obstacles for high quality
- Inclusion of translation professionals and enterprises in the entire research and innovation process
- A unified dynamic-depth weighted-multidimensional quality assessment model with task profiling
- Strongly improved automatic quality estimation
- Ergonomic work environments for computer-supported creative top-quality human translation and multilingual text authoring
- Semantic translation paradigm by extending statistical translation by semantic data such as linked open data, ontologies including semantic models of processes and textual inference models
- Exploitation of strong monolingual analysis and generation methods and resources
- Modular combinations of specialized analysis, generation and transfer models, permitting accommodation of registers and styles and also enabling translation within a language (e.g. between specialists and laypersons).

The expected breakthroughs will include:

- A modular analysis-transfer-generation translation technology that facilitates reuse and constant improvement of modules
- High-quality text translation and reliable speech translation
- Seemingly creative translation skills by analogy-driven transfer models
- Automatic subtitling of films
- Automatic dubbing of films
- Ambient translation

3.3 Solution and Technological Realisation

Special components and extensions of the permanent and ubiquitous service are:

- use and provision platform for providers of computer-supported creative top-quality human translation, multilingual text authoring and quality assurance by experts
- trusted service centers: certified service providers fulfilling highest standards for privacy, confidentiality and security of source data and translations
- quality upscale models: services permitting instant quality upgrades if the results of the requested service levels does not yet fulfil the quality requirements

- domain and task specialization models:
- translingual spaces: dedicated locations for ambient interpretation. Meeting rooms equipped with acoustic technology for accurate directed sound sensoring and emission

The envisaged technical solutions will benefit from new trends in IT such as software as a service, cloud computing, linked open data and semantic web, social networks, crowd-sourcing etc. For MT, a combination of translation brokering on a large scale and translation on demand is promising. The idea is to streamline the translation process such that it (a) becomes simpler to use and more transparent for the end user, and at the same time (b) respects important factors such as subject domain, language, style, genre, corporate requirements, user preferences etc. Technically, what is required is maximum interoperability of all components (corpora, processing tools, terminology, knowledge, maybe even pre-trained translation models) and a cloud or server/service farm of specialized language technology services for different needs (text and media types, domains, etc.) offered by SMEs or large companies.

From an infrastructure point of view a platform has to be designed and implemented for the resource and evaluation demands of large-scale collaborative MT research. An initial inventory of language tools and resources as well as and extensive experience in shared tasks and technology evaluation has been obtained in several EU-funded projects.

Together with LSPs, a common service layer supporting research workflows on HQMT must be established. As third-party (customer) data is needed for realistic development and evaluation, intellectual property rights and legal issues must be taken into account from the onset. The infrastructures to be built include:

- Service clouds with trusted service centres
- Interfaces for services (APIs)
- Workbenches for supporting creative translations
- Novel translation workflows (and improved links to content production and authoring)
- Showcases for services such as ambient and embedded translation

3.4 Impact

Mastering the 23 official languages of the European Union and some 60 other European languages is becoming an insurmountable obstacle for Europe's citizens, economy, political debate, and scientific progress. According to some estimates, the European market volume for translation, and interpretation, comprising software localisation and website globalisation was 5.7 billion EUR in 2008 and was expected to grow by 10% per annum.1 The EU's institutions spend about 1 billion EUR a year on maintaining the policy of multilingualism.

This existing capacity based mostly on human translation is by far not enough to satisfy future translation needs (for example, e-democracy, cross-border commerce, entertainment, games, social networks etc.). However, despite recent improvements, Machine Translation (MT) technology is also far from being ready yet to fill this technological gap. Following a study conducted by META-NET (http://www.meta-net.eu/whitepapers), 23 out of the 30 European languages that were examined currently suffer from very limited quality and performance of MT support, which is an alarming result.

HQMT in the Cloud will ensure and extend the value of the digital information space where everyone can contribute in her own language and be understood by members of other language communities. It will assure that diversity will no longer be a challenge, but a welcome enrichment for Europe both socially and economically. Based on the new technology, language-transparent Web and language-transparent media will help realise a truly multilingual mode of online and media interaction for every citizen regardless of age, education, profession, cultural background, language proficiency or technical skills. Showcase applications areas are:

- Multilingual content production (media, web, technical, legal documents)
- Cross-lingual communication, document translation
- Real-time subtitling and translating speech from live events
- Mobile interactive interpretation for business, social services, and security
- Translation workspaces for on-line services

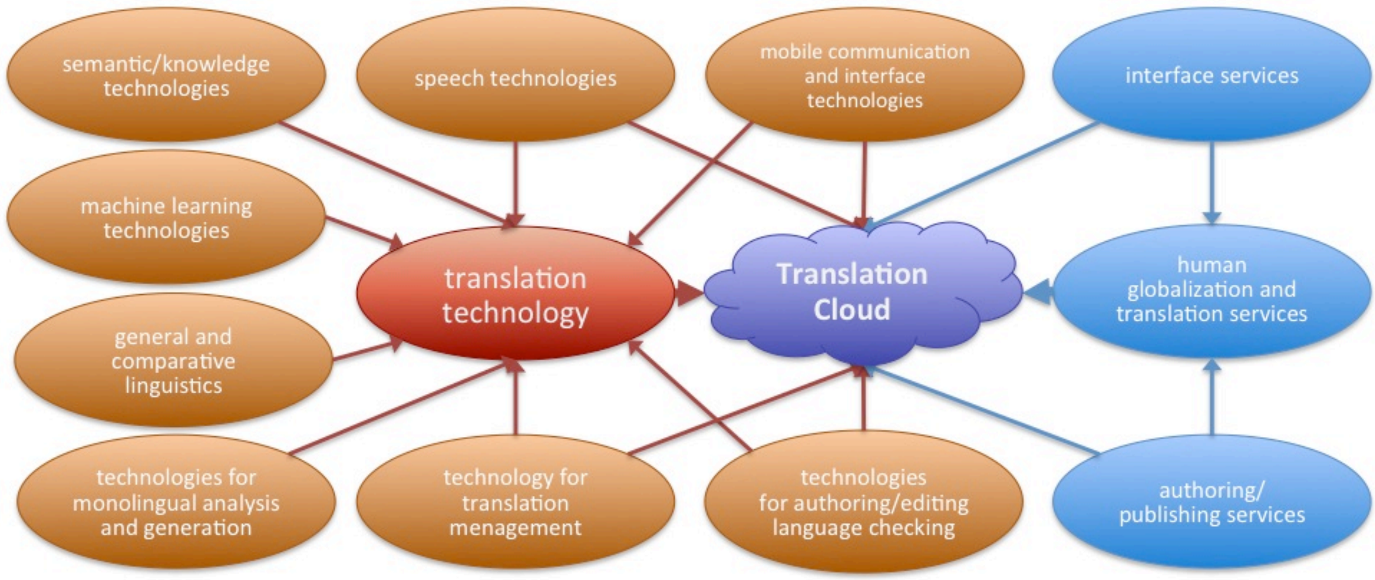
3.5 Organisation of Research

Several very large cooperating and competing lead projects will share an infrastructure for evaluation, resources (data and base technologies), and communication. Mechanisms for reducing or terminating partner involvements and for adding new partners or subcontracted contributors should provide the needed flexibility. A number of smaller projects including some national and regional projects will provide needed building blocks for particular languages, tasks, component technologies or resources. A special scheme will be designed for involving EC-funding, member states, industrial associations, and language communities.

Our stages approach foresees two major phases 2014-mid2017 and mid2017-2020. Certain services such as multilingual access to web-information across European languages should be transferred to implementation and testing at end of phase 2017. Internet-based real-time speech translation for a smaller set of languages will also get into service at this time as well as high-quality MT for some selected domains and tasks. A major mid-term revision with a thorough analytical evaluation will provide a possible breakpoint for replanning or termination.

A close cooperation of language technology and professional language services is planned. In order to overcome the quality boundaries we need to identify and understand the quality barriers. To this end, experienced professional translators and post-editors are required whose judgements and corrections will provide insights for the analytical approach and data for the bootstrapping methodology. The targeted cooperation scheme of research, commercial services and commercial translation technology is planned as a symbiosis since language service professionals working with and for the developing technology will at the same time be the first test users analytically monitored by the evaluation schemes. This symbiosis will lead to a better interplay of research and innovation.

Although the research strand will focus on advances in translation technology for innovation in the language/translation service sector, a number of other science, technology and service areas need to be integrated into the research from day one. Some technology areas such as speech technologies, language checking and authoring systems need to be represented by providers of state-of-the-art commercial products.



Supporting research and innovation in language technology should be accompanied by policy making in the area of multilingualism, but also in digital accessibility. Overcoming language barriers can greatly influence the future of the EU. Solutions for better communication and for access to content in the native languages of the users would reaffirm the role of the EC to serve the needs of the EU citizens. A substantial connection to the infrastructural program CEF could help to speed up the transfer the research results to badly needed services for the European economy and public.

At the same time, use cases should cover areas where the European social and societal needs massively overlap with business opportunities to achieve funding investment that pays back, ideally public-private partnerships.

Concerted activities sharing resources such as error corpora or test suites and challenges/shared tasks in carefully selected areas should be offered to accelerate innovation breakthrough and market-readiness for desperately needed technologies.

4 Priority Theme 2: Social Intelligence and e-Participation

4.1 Solutions for the EU Society and for the Citizen

The central goal behind this theme is to use networked information technology and the digital contents of the web for improving the output, effectiveness and efficiency of decision-making in business and society.

The quality, speed and acceptance of individual and collective decisions is the single main factor for the success of social systems such as enterprises, public services, communities, states and supranational organisations. The growing quantity and complexity of accessible relevant information poses a serious challenge to the efficiency and quality of decision processes. IT provides a wide range of instruments for intelligence applications. Business intelligence, military intelligence or security intelligence applications collect and pre-process decision-relevant information. Analytics programmes search the data for such information and decision support systems evaluate and order the information and apply problem-specific decision rules. Although much of the most relevant information is contained in texts, so-called text analytics programmes today only account for less than 1% of the more than 10 billion US$ business intelligence and analytics market. Because of their limited capabilities in interpreting texts, mainly business news, reports and press releases, their findings are still neither comprehensive nor reliable enough.

Social intelligence builds on improved text analytics methodologies but goes far beyond the analysis. One central goal is the analysis of large volumes of social media, comments, communications, blogs, forum postings etc. of citizens, customers, patients, employees, consumers and other members of stakeholder communities. Part of the analysis is directed to the status, opinions and acceptance associated with the 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 play an important part in individual actions such as voting, buying, supporting, donating and in collective opinion formation, the analysis of sentiment is a crucial component of social intelligence.

Social intelligence does not only analyse but also support collective deliberation processes. Today any collective discussion processes involving large numbers of participants are bound to become intransparent and incomprehensible rather fast. By recording, grouping, aggregating and counting opinion statements, pro and con arguments, supporting evidence, sentiments and new questions and issues, the discussion can be summarized and focussed. Decision processes can be structured, monitored, documented and visualized, so that joining, following and benefitting from them becomes much easier. The efficiency and impact of such processes can thus be greatly enhanced.

Since many collective discussions will involve participants in several countries, e.g., EU member states or enterprise locations, cross-lingual participation needs to be supported. Special support will also be provided for participants not mastering certain group-specific or expert jargons and for participants with disabilities affecting their comprehension.

4.2 Novel Research Approaches and Targeted Breakthroughs

A key enabler in the above vision will be language and semantic technologies that can map large, heterogeneous, and, to a large extent, unstructured volumes of on-line content to 'actionable' representations that support decision making and analytics tasks. Such mappings can range from the relatively 'shallow' to the relatively 'deep', encompassing for example coarse-grained topic classification at the document or paragraph level or the identification of named entities, as well as in-depth syntactic, semantic and rhetorical analysis at the level of individual sentences and beyond (paragraph level, text level) or the resolution of co-reference or modality cues within and across sentences.

Language technologies such as, for example, information extraction, data mining, automatic linking and summarization have to be made interoperable with modern knowledge representation approaches and semantic web methods such as ontological engineering. Drawing expertise from related areas such as knowledge management, information sciences, or social sciences is an important prerequisite to meet the challenge of modelling social intelligence. A new research approach should target the bottleneck of knowledge engineering by:

- "Semantification" of the web: bridging between the semantic parts and islands of the web and the traditional web containing unstructured data;
- Merging and integrating textual data with social network and social media data, especially along the dimension of time;
- Aligning and making comparable different genres of content like mainstream-news, social media (blogs, twitter, facebook etc.), academic texts, archives etc.;
- Extracting semantic representations from social media content, i.e., creating representations for reasoning and inferencing;
- Taking metadata and multimedia data into account.

The following list contains specific targeted breakthroughs to be sought in this scenario:

- Social intelligence by detecting and monitoring opinions, demands and needs;
- Detecting diversity of views, biases along different dimensions (e.g., demographic) etc. including temporal dimension (i.e., modelling evolution of opinions);
- Decision support for both decision makers and participants;
- Support of collective deliberation and collective knowledge accumulation;
- Vastly improved approaches to sentiment detection and sentiment scoring (going beyond the approach that relies on a list of positive and negative keywords).
- Introducing the approach of genre-driven text and language-processing (different genres need to be processed differently).
- Improved decision making and decision participation;

- Personalized recommendations of e-Participation topics to citizens;
- Proactive involvement in e-Participation activities ;
- Understanding influence diffusion across social media (identifying drivers of opinion spreading);
- More sophisticated methods for topic and event detection that are tightly integrated with the Semantic Web, Linked Open Data and machine-readable knowledge bases such as DBpedia.
- Modelling content and opinions flows across social networks;
- Evaluation of created methods by analytic/quantitative and sociological/qualitative means.

4.3 Solution and Technological Realisation

Individual solutions should be assembled from a repository of generic monolingual and cross-lingual language technologies, packaging state-of-the-art techniques in *robust*, *scalable*, *interoperable*, and *adaptable* components that are deployed across sub-tasks and sub-projects, as well as across languages where applicable (e.g., when the implementation of a standard data-driven technique can be trained for individual languages). These methods need to be combined with powerful analytical approaches that can aggregate all relevant data to support analytic decision making and develop new access metaphors and task-specific visualisations.

By *robust* and *scalable* we mean technologically mature, engineered and scalable solutions that can perform high-throughput analysis of web data at different levels of depth and granularity in line with the requirements of the respective applications. Technology should also be able to work with heterogeneous sources, ranging from completely unstructured (arbitrary text documents of any genre) to completely structured (ontologies, linked open data, databases).

To accomplish *interoperability* we suggest a strong semantic bias in the choice and design of interface representations: to the highest degree possible, the output (and at deeper levels of analysis also input) specifications of component technologies should be interpretable semantically, both in relation to natural language semantics (be it lexical, propositional, or referential) and extra-linguistic semantics (e.g., taxonomic world or domain knowledge). For example, grammatical analysis (which one may or may not decompose further into tagging, syntactic parsing, and semantic role labelling) should make available a sufficiently abstract, normalized, and detailed output, so that downstream processing can be accomplished without further recourse to knowledge about syntax. Likewise, event extraction or fine-grained, utterance-level opinion mining should operate in terms of formally interpretable representations that support notions of entailment and, ultimately, inference.

Finally, our *adaptability* requirement on component technologies addresses the inherent heterogeneity of information sources and communication channels to be processed in this scenario. Even in terms of monolingual analysis only, linguistic variation across genres (ranging from carefully edited, formal publications to spontaneous and informal social media channels) and domains (as in subject matters) often calls for technology adaptation, where even relatively mature basic technologies (e.g., part-of-speech taggers) may need to be customized or re-trained to deliver satisfactory performance. Further taking into account variation across downstream tasks, web-scale language processing typically calls for different parameterizations and trade-offs (e.g., in terms of computational cost vs. breadth and depth of analysis) than an interactive self-help dialogue scenario. For these reasons. relevant such trade-offs need to be documented empirically, and component technologies accompanied with methods and tools for adaptation and cost-efficient re-training, preferably in semi- and un-supervised settings.

The technical solutions needed include:

- Technologies and platforms for decision support, collective deliberation and e-participation.
    - A large public discussion platform for Europe-wide deliberation on pressing issues such as energy policies, financial system, migration, natural disasters, etc.;
    - Visualization of "social intelligence" related data and processes for decision support (for politicians, health providers, manufacturers, or citizens);
    - Documentation of social deliberation and decision making.
- High-throughput, web-scale content analysis techniques that can process multiple different sources, ranging from unstructured to completely structured, at different levels of granularity and depth by allowing to trade-off depth for efficiency as required.
    - Mining e-Participation content for recommendations, summarization and proactive engagement of less active parts of population;
    - Detection and prediction of events and trends from content and social media networks;
    - Extraction of knowledge and semantic integration of social content with sensory data and mobile devices (in near-real-time)
- Cross-lingual technology to increase the social reach and approach cross-culture understanding

4.4 Impact

The 21st century presents us with multiple challenges including efficient energy consumption, global warming and financial crises. It is obvious that no single individual can provide answers to challenging problems such as these, nor will top-down imposed measures find social acceptance as solutions. Language technology will enable a paradigm shift in transnational public deliberation.

The applications and technologies discussed in this section will change how business adapts and communicates with their customers. It will increase transparency in decision-making processes, e.g., in politics and at the same time give more power to the citizen. As a by-product, the citizens are encouraged to become better informed in order to make use of their right to participate in a reasonable way. Powerful analytical methods will help European companies to optimize marketing strategies or foresee certain developments in aspects by extrapolating on the basis of current trends. Leveraging social intelligence for informed decision making is recognized as crucial in a wide range of context and scenarios:

- Organisations will better understand the needs, opinions, experiences, communication patterns, etc. of their actual and potential customers so that they can react quickly to new trends and optimize their marketing and customer communication strategies.
- Companies will get the desparately needed instruments to exploit the knowledge and expertise of their huge and diverse workforces, the wisdom of their own crowds, which are the most highly motivated and most closely affected crowds.
- Political decision makers will be able to analyse public deliberation and opinion formation processes in order to react swiftly to ongoing debates or important, sometimes unforeseen events.
- Citizens and customers get the opportunity (and necessary information) to participate and influence political, economic and strategic decisions of governments and companies, ultimately leading to more transparency of decisions processes.

Thus, leveraging collective and social intelligence in developing new solutions to these 21st century challenges seems a promising approach in such domains where the complexity of the issues under discussion is beyond the purview of single individuals or groups.
The research and innovation will provide technological support for emerging new forms of issue-based, knowledge-enhanced and solution-centred participatory democracy involving large numbers of expert- and non-expert stakeholders distributed over large areas, using multiple languages.

4.5 Organisation of Research

Research in this area touches upon political as well as business interests and at the same time is scalable in reach from the regional to the European scale. Therefore, it is necessary to identify business opportunities and potential impact for society at different levels and to align EU level research with efforts on the national level.

5 Priority Theme 3: Socially Aware Interactive Assistants

5.1 Solutions for the EU Society and for the Citizen

Socially aware interactive assistants are conversational agents realized via explicit embodiment, robots or a different type of interface without explicit embodiment. Their socially-aware behaviour leverages from the combination of analysis and synthesis of non-verbal, speech and semantic signals.

It is the proper time to develop and make operational socially aware and also multilingual assistants that can support 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 various environments, instrumented indoor environments (such as meeting rooms, offices, apartments), instrumented outdoor environments (streets, cities, transportation, roads) and virtual environments (such as the World Wide Web, virtual worlds, and games), 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, wherever they are, in any environment. It can interact in any language and in any communication modality. It can adapt and be personalized to individual communication abilities, including special needs (for the visual, hearing, or motor impaired), affections, or language proficiencies. It can recognize and generate speech incrementally and fluently. It is able to self-assess its performance and recover from errors. It can learn, personalize itself and forget through natural interaction, act on objects in instrumented spaces (rooms, apartments, streets). It can assist in language training and in education in general, and provide synthetic multimedia information analytics. It recognizes people's identity, and their gender, accent, language, or style. If the agent is embodied in a robot, it can move, manipulate objects, and interact with people.

This global solution includes several components:

- Interacting naturally with humans (in games, entertainment, education, communication, instrumented spaces, etc.) in an implicit (proactive manner, without explicit embodiment) or explicit (spoken dialogue and/or gesticulation) manner based on robust analysis of human user identity, age, gender, verbal and nonverbal behaviour, and the social context;
- Exhibiting robust performance everywhere (indoor and outdoor environments, mobile applications, augmented reality);
- Overcoming handicap obstacles by means of suitable technologies (sign language understanding, assistive applications, adapted communication to suit cognitively impaired, etc.);

- Interacting naturally with and in groups (in social networks and forums, with several humans or artificial agents/robots present);
- Exhibiting multilingual proficiency (speech-to-speech translation, interpretation in meetings and videoconferencing, cross-lingual information access);
- Referring to written support (speech transcription, close-captioning of audiovisuals, reading machines, multimedia books);
- Providing personalized training (computer-aided language learning, or computer-aided education and self-assessment in general).

5.2 Novel Research Approaches and Targeted Breakthroughs

Developing socially aware interactive assistants implies several research breakthroughs:

- **Better core speech and language technologies**
  - More basic research (incl. physiological, perception and cognitive processes)
  - Better speech recognition
    - Increased accuracy: open vocabulary, any speaker
    - More robust: noise, cross-talk, distant microphone
    - Lower maintenance: self-assessment, self-adaptation, personalization,
    - Error recovery, learning *and forgetting* of new/old information
  - Better speech synthesis
    - Improved naturalness and expression
    - Control parameters for linguistic/paralinguistic meaning, speaking style, voice conversion (mimicking someone's voice) and emotion
    - From "reading style" to incremental conversational speech, including filled pauses, hesitations, self-repairs
- **From recognition to understanding**
  - Speech is an aspect of human communication, not only STT/TTS
  - Human communication is multimodal (including speech, facial expressions, body gestures, postures, etc.), crossmodal and *flexi*modal: it is based on pragmatically best suited modalities.
  - Build semantic and pragmatic models of human communication
    - Parsing Accuracy: model temporal interdependencies within and between modalities as to maximise human-communication-prediction ability of the agent
    - Contextual awareness: model situational interdependencies between the observed context and

the best suited modalities for attaining robust analysis of human user communication
        - Self-assessment: What is plausible?
        - Multimodal content analytics: infer knowledge from multiple sensory modalities
    - Detect and recover interactively from mistakes
        - Learn continuously and incrementally from mistakes
        - Unsupervised, online and/or by interaction
    - Include paralinguistics (prosody analysis, visual cues): emotion, laughs
    - Production of adequate semantically annotated language and multimodal resources (huge effort)
    - Principled integration of recognition, synthesis, and understanding, including different levels of evaluation, and different levels of (automatic) annotation
    - Understanding of socio-emotional functions of communicative behaviour, including group dynamics, reputation management, relationship management
- **Towards Natural Dialogue**
    - Multimodal dialogue (speech, gesturing, body posture, facial expressions)
    - Socially aware systems for dialogue
    - “Transparent” systems
        - Open microphone, multiparty conversations (humans, artificial agents, robots)
        - Use of other sensor devices: GPS, RFID, cameras, etc.
    - More advanced dialogue models
        - Proactive (not only reactive)
        - Detect that a voice emission is in machine intention, interpret silence
        - Process direct/indirect speech acts, including lies, humour etc.
    - Define dialogue systems evaluation metrics/protocols
    - Produce language resources (acquisition/annotation) from the real world
        - Probably implies an incremental system design
        - Use of data available on the internet (conversations, talks shows)
    - Understanding and modelling the functions of interpersonal adaptation and alignment (in speech and other modalities) for the ease, robustness and efficiency of dialogue, towards integrated hybrid models that combine explicit grounding with implicit entrainment processes
- **Handling multilingualism**
    - Interactive systems should cover, or be easily portable at least to all EU languages (23 official languages and regional languages, Catalan, Basque, etc.)

- General language portability: from few to *many* languages (also: language support for European to/from non-European languages)
- Speech translation in human-human interaction (e.g., meetings, i.e., speech translation among multiple human users speaking different languages)
- Deal with languages, accents and dialects effectively
  - Should recognize language, gender and accents
  - Cross-cultural support, not only cross-lingual
- Provide cross-lingual access to information and knowledge
- Availability of language resources and language technology evaluation in all languages
- Methods to fully exploit limited language resources

5.3 Solution and Technological Realisation

The technological and scientific state-of-the-art is at a stage that allows developments towards the ideal socially aware multilingual assistant as specified above. Developments in machine learning, including adaptation, unsupervised learning from streams of data, continuous learning, and transfer learning make it possible to automatically learn the desired capabilities from data. In addition, existing language and multimodal resources make it possible to bootstrap systems. Interdisciplinary advances (e.g., social signal processing) provide a strong basis for the necessary advances.

Technological advances are continuously being achieved in the language technology and vision-based human behaviour analysis and synthesis fields. Ubiquitous technologies are now widely available (at lower costs and in reduced size). User-centric approaches have been largely studied and crowd-sourcing is being more and more widely used. Quantitative and objective language technology and human-behaviour understanding technology evaluations, allowing for assessing a technological readiness level (TRL), are carried out more widely, as best practice, and language resources and publicly-available annotated recordings of human spontaneous behaviour are now available.

However there are still some prohibitive factors. Language technology evaluation is still limited and is not conducted for all languages. There is limited availability of language resources (LR), and the necessary LR do not exist yet for all languages in sufficient quantity or with sufficient quality. Similarly, publicly-available recordings of spontaneous (rather than staged) human behaviour are still sparse, especially when it comes to continuous synchronised observations of multi-party interactions. Limited progress of the technology for automatic understanding of social behaviours like rapport, empathy, envy, conflict, etc., is mainly attributed to this lack of suitable resources. In addition, we still have very limited knowledge of human language and human behaviour perception processes and automated systems often face theoretical and technological complexity of modelling and handling these processes correctly.

5.4 Impact

The impact of this Priority Theme will be wide-ranging. It will impact the work environment and processes, creativity and innovation, leisure and entertainment, and the private life. Several societal and economical facts call for, but also allow for, improved and more natural interaction between humans and the real world through machines. The ageing society requests ambient intelligence. Globalization involves the capacity to interact in many languages, and offers a huge market for new products fully addressing that multilingual necessity.

The automation of society implies more efficiency and a 24/7 availability of services and information, while green technologies, such as advanced videoconferencing, need to be privileged. The continuously reduced costs and speed improvement of hardware allow for affordable and better technologies, that can now easily be made available online through app stores.

At the same time we still face prohibitive factors. The cultural, political and economical dimensions of language are well perceived, but not its technical dimension. There is still a psychological barrier for communicating with machines, although this gets more and more common through the use of smartphones and applications such as Skype or Facetime. There is an extra cost for developing personalized systems and the business models are difficult to define as humans are used to communicate by speech or in writing at no cost.

5.5 Organisation of Research

In order to increase research efficiency within a public-private partnership, the preferred infrastructure would be to handle the various applications in connection with the *cooperative* development of technologies, including regular objective evaluation of technology progress, and the production of the language and human naturalistic behaviour resources which are necessary to develop and test the technologies.

To maximise impact, it is necessary to have a substantial effort in the development of integrated systems based on open architectures, and a multilingual middleware to enable the developed functionalities to be incorporated in a wide range of software. This might best be achieved through a small number of coordinating projects, attached to a 'federation of STREPs' with complementary goals. These projects should be objective-driven, with clear research, technology and exploitation milestones, coordinated by an on-going road-mapping effort.

This includes the production of adequate language and human naturalistic behaviour corpora, semantically annotated including prosodic and non-verbal behavioural cues, which represents a huge effort. This also includes the production (acquisition and annotation) of dialogue corpora from the real world, which probably implies an incremental system design, and either the use of synchronised continuous observations of all involved parties, or the use of similar data available on Internet (conversation, talks shows).

Dialogue systems evaluation still needs research investigations on the choice of adequate metrics and protocols. The multilingual dimension that is targeted implies the availability of language resources and language technology evaluation for all languages. Handling them all together reduces however the overall effort, given the possibility to use the same best practices, tools and protocols.

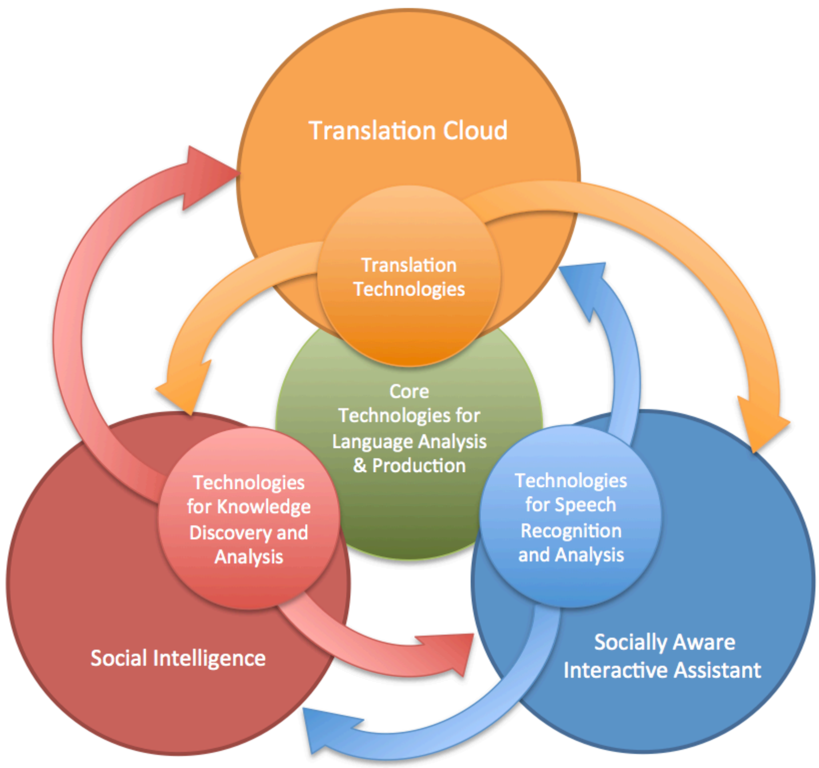
6 Organization of Research and Innovation

6.1 Structure and principles of research organisation

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

All three priority themes 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 themes will concentrate on one central area of language technology: 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 Socially Aware Interactive Assistant will take on interface technologies such as speech and multimodal interfaces.

A highly simplified illustration of the topics for scientific cooperation among the three priority themes is shown in the following figure.



Except for a few large national projects and programmes such as TechnoLanguage and Quaero in France, Verbmobil and Theseus in Germany and DARPA Communicator and GALE in the US the field of language technology does not have experience with research efforts of the scope required for the targeted advances and planned in this SRA. Nevertheless, our technology area has to follow developments in other key engineering disciplines and speed up technology evolution by massive collaboration based on competitive division of labour and sharing of resources and results. In our reflection on optimal schemes for organizing the programme we tried to draw lessons from our own field's recent history and to capitalize on experience from other fields by adopting approaches that proved successful and evading encountered pitfalls.

The final model for the organisation of collaboration will have to be guided by a thoughtful combination of the following basic approaches.

Flexible collaborative approach: For each priority theme, one or several very large cooperating and competing lead projects will share an infrastructure for evaluation, resources (data and base technologies), and communication. Mechanisms for reducing or terminating partner involvements and for adding new partners or subcontracted contributors should provide the needed flexibility. A number of smaller projects including some national and regional projects will provide needed building blocks for particular languages, tasks, component technologies or resources. A special cooperation scheme will be designed for effectively involving EC-funding, contributions from member states, industrial associations, and language communities. The choice of suggested funding instruments will be determined in due time.
Staged approach: Two or three major phases are foreseen. For better concertation the major phases should be synchronized among priority themes and individual projects, both on the international and also on the national level.
Evolutionary approach: Instead of banking on one selected paradigm, a number of competing approaches will be followed in parallel with shared and efficient schemes for evaluation, merging, adopting and also discontinuing research threads so that the two elements of successful evolutionary research approaches, selection and cross-fertilization, are exploited to the maximum extent possible.
Analytical approach: Instead of the currently predominant search for an ideal one-size-fits-all approach, the research will focus on observed quality barriers and not shun computationally expensive dedicated solutions for overcoming particular obstacles.
Bootstrapping approach: Better systems can be derived from more and better data and through new insights. Improved systems can also be used to gain better data and new insights. Thus the combination of the analytical evolutionary approach with powerful machine learning techniques will be the basis for a technology bootstrapping, which has been the by far most fruitful scheme for the development of highly complex technologies.
Close cooperation of language technology and relevant areas of service and technology industries: In order to increase chances of successful commercialisation and to obtain convincing and suffi-

ciently tested demonstrations of novel applications and technology showcases, the relevant industrial sectors of industry must be strongly integrated into the entire research cycle.

Tighter research-innovation cycle: Through the collaboration between research, commercial services and commercial technology industries, especially through the shared evaluation metrics and continuous testing, the usual push-model of technology transfer will hopefully be substituted by a pull-model, in which the commercial technology users can ask for specific solutions. In the envisaged research scheme incentives will be created for competing teams each composed of researchers, commercial users and commercial developers by the participating enterprises for initiating successful innovations

Interdisciplinary approach: Several science, technology and service areas need to be integrated into the research programme from day one. Some technology areas such as speech technologies, language checking and authoring systems need to be represented by providers of state-of-the-art commercial products. In addition, a tighter integration of neighbouring fields and disciplines into the programme such as the Semantic Web and the Machine Learning communities is foreseen to arrive at the needed level of technology, information and knowledge integration.

Supporting research and innovation in language technology should be accompanied by policy making in the area of multilingualism, but also in digital accessibility. Overcoming language barriers can greatly influence the future of the EU. Solutions for better communication and for access to content in the native languages of the users would reaffirm the role of the EC to serve the needs of the EU citizens. A substantial connection to the infrastructural programme CEF could help to speed up the transfer of research results to badly needed services for the European economy and public.

At the same time, use cases should cover areas where the European societal needs massively overlap with business opportunities to achieve funding investment that pays back, ideally public-private partnerships.

The coordination among the three priority research strands poses additional administrative challenges. Because of the described interdependencies and also because of the need to maintain and improve the obtained level of cohesion and community spirit in the European LT community, a coordinating body is needed. Whether such an entity is jointly carried by the three areas or by a separate support project, needs to be determined in the upcoming discussion on the appropriate support instruments for the identified research priorities.

6.2 Sharing of Resources and Results

By now the European academic and industrial technology community is fully aware of the need for sharing resources such as language data (e.g., text and speech corpora), language descriptions (e.g., lexicons, thesauri, grammars), tools (e.g., taggers, stemmers, tokenizers) and core technology components (e.g., morphological, syntactic, semantic processing). Initiatives like FLaReNet and CLARIN have prepared the ground for this sharing, META-SHARE has provided the technological platform as well as legal and organisational schemes.

6.3 Powerful mechanisms for showcasing and innovation

As a key-enabling technology, language technology usually enters the markets in combination with other technologies as an essential component of novel products and services. Existing widely used platforms that would benefit immediately from improved language technology include hardware platforms such as the iOS devices from Apple, but also service platforms such as Google, Amazon, Facebook and finally software platforms such as the Microsoft Office or Adobe application suites.

In the future many more platforms will appear whose services are enhanced by language technology. Actually, language technology could become the essential feature for the success of new services.

LT research projects, however, do not have the resources, expertise and time to create the necessary platforms and to integrate their results into real-life services.

For each of the three priority research strands, chances for successful showcasing and successful commercial innovation will increase tremendously if usable services could be offered on a platform of required strength and reliability. Given the nature of the priority themes, these platforms will most likely be web-based service platforms offering single and combined services to users. They will furthermore provide facilities for new services to join and combine with existing ones.

By the envisaged dimensions of the three research strands and by the foreseen massive involvement of European language technology industry such platforms can actually be built and put to use. They will serve as a major enabling factor for successful commercialization in Europe.

7 An Overview of META-NET

META-NET is a Network of Excellence partially funded by the European Commission. The network currently consists of 54 research centres in 33 European countries. META-NET forges the Multilingual Europe Technology Alliance (META), a growing community of language technology professionals and organisations in Europe. META-NET fosters the technological foundations for a truly multilingual European information society that:

- makes communication and cooperation possible across languages;
- grants all Europeans equal access to information and knowledge regardless of their language;
- builds upon and advances functionalities of networked information technology.

The network supports a Europe that unites as a single digital market and information space. It stimulates and promotes multilingual technologies for all European languages. These technologies support automatic translation, content production, information processing and knowledge management for a wide variety of subject domains and applications. They also enable intuitive language-based interfaces to technology ranging from household electronics, machinery and vehicles to computers and robots.

Launched on 1 February 2010 META-NET has already conducted various activities in its three lines of action META-VISION, META-SHARE and META-RESEARCH.

META-VISION fosters a dynamic and influential stakeholder community that unites around a shared vision and a common strategic research agenda (SRA). The main focus of this activity is to build a coherent and cohesive LT community in Europe by bringing together representatives from highly fragmented and diverse groups of stakeholders. White Papers were produced for 30 languages, each one describing the status of one language with respect to its state in the digital era and existing technological support. The shared technology vision was developed in three sectorial Vision Groups. The META Technology Council was established that is currently drafting the SRA based on the vision in close interaction with the entire LT community.

META-SHARE creates an open, distributed facility for exchanging and sharing resources. The peer-to-peer network of repositories will contain language data, tools and web services that are documented with high-quality metadata and organised in standardised categories. The resources can be readily accessed and uniformly searched. The available resources include free, open-source materials as well as restricted, commercially available, fee-based items.

META-RESEARCH builds bridges to related technology fields. This activity seeks to leverage advances in other fields and to capitalise on innovative research that can benefit language technology. In particular, the action line focuses on conducting leading-edge research in machine translation, collecting data, preparing data sets and organising language resources for evaluation purposes; compiling inventories of tools and methods; and organising workshops and training events for members of the community.