

# The Strategic Impact of META-NET on the Regional, National and International Level

Georg Rehm<sup>1</sup>, Hans Uszkoreit<sup>1</sup>, Sophia Ananiadou<sup>2</sup>, N ria Bel<sup>3</sup>, Audron  Bielevi ien <sup>4</sup>, Lars Borin<sup>5</sup>, Ant nio Branco<sup>6</sup>, Gerhard Budin<sup>7</sup>, Nicoletta Calzolari<sup>8</sup>, Walter Daelemans<sup>9</sup>, Radovan Garab k<sup>10</sup>, Marko Grobelnik<sup>11</sup>, Carmen Garc a-Mateo<sup>12</sup>, Josef van Genabith<sup>13</sup>, Jan Haji <sup>14</sup>, Inma Hern ez<sup>15</sup>, John Judge<sup>13</sup>, Svetla Koeva<sup>16</sup>, Simon Krek<sup>11</sup>, Cvetana Krstev<sup>17</sup>, Krister Lind n<sup>18</sup>, Bernardo Magnini<sup>19</sup>, Joseph Mariani<sup>20</sup>, John McNaught<sup>2</sup>, Maite Melero<sup>3</sup>, Monica Monachini<sup>8</sup>, Asunci n Moreno<sup>21</sup>, Jan Odijk<sup>22</sup>, Maciej Ogrodniczuk<sup>23</sup>, Piotr P zik<sup>24</sup>, Stelios Piperidis<sup>25</sup>, Adam Przepi rkowski<sup>23</sup>, Eir kur R gnvaldsson<sup>26</sup>, Mike Rosner<sup>27</sup>, Bolette Sandford Pedersen<sup>28</sup>, Inguna Skadi a<sup>29</sup>, Koenraad De Smedt<sup>30</sup>, Marko Tadi <sup>31</sup>, Paul Thompson<sup>2</sup>, Dan Tufi <sup>32</sup>, Tam s V radi<sup>33</sup>, Andrejs Vasiļjevs<sup>34</sup>, Kadri Vider<sup>35</sup>, Jolanta Zabarskaite<sup>4</sup>

DFKI GmbH  
Berlin, Germany<sup>1</sup>

University of Gothenburg  
Gothenburg, Sweden<sup>5</sup>

University of Antwerp  
Antwerp, Belgium<sup>9</sup>

Dublin City University, CNGI  
Dublin, Ireland<sup>13</sup>

University of Belgrade  
Belgrade, Serbia<sup>17</sup>

Universitat Polit cnica de Catalunya  
Barcelona, Spain<sup>21</sup>

R. C. "Athena", ILSP  
Athens, Greece<sup>25</sup>

University of Latvia  
Riga, Latvia<sup>29</sup>

Hungarian Academy of Sciences  
Budapest, Hungary<sup>33</sup>

University of Manchester  
Manchester, UK<sup>2</sup>

University of Lisbon  
Lisbon, Portugal<sup>6</sup>

E.  t r Institute of Linguistics  
Bratislava, Slovakia<sup>10</sup>

Charles University  
Prague, Czech Republic<sup>14</sup>

University of Helsinki  
Helsinki, Finland<sup>18</sup>

University of Utrecht  
Utrecht, The Netherlands<sup>22</sup>

University of Iceland  
Reykjavik, Iceland<sup>26</sup>

University of Bergen  
Bergen, Norway<sup>30</sup>

Tilde  
Riga, Latvia<sup>34</sup>

Universitat Pompeu Fabra  
Barcelona, Spain<sup>3</sup>

University of Vienna and  
Austrian Academy of Sciences<sup>7</sup>

Jo ef Stefan Institute  
Ljubljana, Slovenia<sup>11</sup>

University of the Basque Country  
Bilbao, Spain<sup>15</sup>

FBK  
Trento, Italy<sup>19</sup>

Polish Academy of Sciences  
Warsaw, Poland<sup>23</sup>

University of Malta  
Msida, Malta<sup>27</sup>

University of Zagreb  
Zagreb, Croatia<sup>31</sup>

University of Tartu  
Tartu, Estonia<sup>35</sup>

Institute of the Lithuanian Language  
Vilnius, Lithuania<sup>4</sup>

CNR  
Pisa, Italy<sup>8</sup>

University of Vigo  
Vigo, Spain<sup>12</sup>

Bulgarian Academy of Sciences  
Sofia, Bulgaria<sup>16</sup>

LIMSI-CNRS, IMMI  
Orsay, France<sup>20</sup>

University of Ł d   
Ł d , Poland<sup>24</sup>

University of Copenhagen  
Copenhagen, Denmark<sup>28</sup>

Romanian Academy  
Bucharest, Romania<sup>32</sup>

*Corresponding author:*  
georg.rehm@dfki.de<sup>1</sup>

## Abstract

This article provides an overview of the dissemination work carried out in META-NET from 2010 until early 2014; we describe its impact on the regional, national and international level, mainly with regard to politics and the situation of funding for LT topics. This paper documents the initiative's work throughout Europe in order to boost progress and innovation in our field.

**Keywords:** LR National/International Projects, Infrastructural/Policy Issues, Multilinguality, Machine Translation

## 1. Introduction and Overview

The multilingual setup of our European society imposes societal challenges on political, economic and social integration and inclusion, especially in the creation of the single digital market and unified information space targeted by the Digital Agenda (EC, 2010). Language technology is the missing piece of the puzzle, it is the key enabler and solution to boosting growth and strengthening Europe's competitiveness.

Recognising Europe's exceptional demand and opportunities, 60 leading research centres in 34 European countries joined forces in META-NET, a Network of Excellence dedicated to the technological foundations of a multilingual, European information society. META-NET was partially supported through a total of four projects funded by the European Commission: T4ME (2010–2013, FP7), CESAR, METANET4U and META-NORD (2011–2013, ICT-PSP). META-NET is forging

the Multilingual Europe Technology Alliance (META) with more than 760 organisations and experts representing multiple stakeholders and signed collaboration agreements with more than 40 other projects and initiatives such as CLARIN and FLReNet.

META-NET's goal is monolingual, crosslingual and multilingual technology support for all languages spoken by a significant population in Europe (Rehm and Uszkoreit, 2013). We recommend focusing on three priority research themes connected to application scenarios that will provide European R&D with the ability to compete with other markets and achieve benefits for European society and citizens as well as opportunities for our economy and future growth. We are confident that the EU funding programmes Horizon 2020 (EC, 2012) and Connecting Europe Facility (CEF) (EC, 2014), combined with national/regional funding, can provide the necessary resources for accomplishing our joint vision.

The work carried out in META-NET is structured in three pillars. All aspects concerned with community building, formulating a shared vision and preparing a strategy belong to the META-VISION pillar. Additionally, META-NET carried out innovative research in META-RESEARCH and developed META-SHARE (Piperidis et al., 2014). This paper primarily discusses the impact of the work carried out in META-VISION.

## 2. Key Communication Instruments

Communication activities circled around three key instruments. The study *Europe's Languages in the Digital Age* describes, for 31 European languages, the level of support through Language Technology. The *Strategic Research Agenda* provides recommendations and a unified strategy for European LT to boost growth and innovation. META-NET also organised multiple conferences, roadshow events, and workshops.

### 2.1. Language White Paper Series

The META-NET Language White Paper Series “Europe’s Languages in the Digital Age” (Rehm and Uszko-reit, 2012) describes the current state of LT support for 30 European languages (including all 24 official EU languages). More than 200 experts participated as co-authors and contributors. In early 2014, the first additional volume of the series was published (Evas, 2014). The differences in support are dramatic and alarming. In all four areas we examined (MT, speech, text analytics, LRs), English is ahead of the other languages but even support for English is far from being perfect. While there are good quality software and resources available for a few larger languages and application areas, others, usually smaller or very small languages, have substantial gaps. Many languages lack basic technologies for text analytics and essential LRs. Currently no language has the technological support it deserves.

The whole series is available online.<sup>1</sup> On the website we also present the press release “At least 21 European Languages in Danger of Digital Extinction”, circulated on Sept. 26, 2012, the European Day of Languages 2012.

### 2.2. The META-NET Strategic Research Agenda

META-NET developed the Strategic Research Agenda for Multilingual Europe 2020 (SRA) (Rehm and Uszko-reit, 2013).<sup>2</sup> Our recommendations for multilingual Europe are based on a planning process involving more than one thousand experts. They envisage five lines of action for large-scale research and innovation. First, there are three priority research themes: *Translingual Cloud*, *Social Intelligence and e-Participation* and *Socially Aware Interactive Assistants*. Two additional

themes focus upon *Core technologies and resources for Europe's languages* and a *European service platform for language technologies*. Their objective is to turn our joint vision into reality and allow Europe to benefit from the LT revolution. This SRA is the first unified strategic plan for the entire European LT sector.

### 2.3. Conferences and Events

Since 2010 META-NET has been organising annual conferences: META-FORUM 2010 (Brussels), 2011 (Budapest), 2012 (Brussels), 2013 (Berlin). Additionally, we organised dozens of smaller events, primarily with a focus on the regional or national level, e. g., in Croatia, Germany, Hungary, Latvia, Lithuania and Poland.<sup>3</sup> The initiative also participated in many scientific, strategic, and political workshops and meetings.

## 3. Impact on the International Level

In this section we present a brief summary of the impact on the international level, while section 4. describes the impact on the national and regional levels.

### 3.1. Visibility of the Initiative and its Key Topics

Our work had significant effects with regard to the visibility not only of META-NET itself in our own research community but also of our key topics with regard to the public at large. First and foremost, the impact of the White Paper Series and the corresponding press release (“At least 21 European Languages in Danger of Digital Extinction”) published in 30 languages on Sept. 26, 2012, was beyond our imagination. It resulted in more than 600 pieces of media coverage internationally (online, print, radio, tv). News came in from 43 countries and in 35 different languages with all European countries covered. The press campaign resulted in 45+ broadcast interviews with META-NET representatives (ca. 30 radio interviews, ca. 25 tv reports). We had coverage in well known outlets such as, among others, Politiken (DK), Süddeutsche Zeitung (DE), Prosilipsis (GR), Wired (IT), Demokracija (SL), PTC1 (RS), El Mundo (ES), Huffington Post (UK) and Reddit (USA). Along with the increased coverage we noticed a huge increase of traffic on our website, 65% of which were new visitors. Of those new visitors, 11% of visits came from Brussels – the greatest for a single location. Similar trends could be observed for almost all capital cities across Europe. The echo generated by the press campaign shows that Europe is extremely passionate and concerned about its languages and that it is also very interested in the idea of establishing a solid technology base for overcoming language barriers. As an analysis of our website traffic shows, we also managed to raise a certain level of attention and awareness, both in the EC and the EP.

<sup>1</sup><http://www.meta-net.eu/whitepapers>

<sup>2</sup><http://www.meta-net.eu/sra>

<sup>3</sup><http://www.meta-net.eu/events>

### 3.2. The Multilingual Europe Technology Alliance

One of the goals of META-NET is to forge the Multilingual Europe Technology Alliance (META) as an open strategic technology alliance as a measure against the highly fragmented set of LT-related communities that we had in Europe before META-NET. By now (early March 2014), META has 764 members in 63 countries. Together with META-NET itself (60 research centres in 34 European countries) this group constitutes a very large and strong community that is finally able to speak with one voice and to present a shared vision and strategy as demonstrated by the unanimous support of the plans presented in the SRA. Its priority themes are the result of 80+ meetings and discussions between hundreds of experts from research and industry. META-NET is now a sustainable network and established brand with a lifetime beyond that of its first funded phase.

META-NET's work led to dozens of invited keynote speeches at multiple workshops and conferences and many invited papers, book chapters and contributions in industry journals. Our outreach programme led to the drafting and signing of more than 40 collaboration agreements with other projects and organisations. We built strong connections to other stakeholder communities, most importantly to organisations representing Language Service Providers (GALA, tekomp), bodies representing the language communities (EFNIL, NPLD, Council of Europe Committee of Experts on the Language Charter) (META-NET, 2013), many language professionals, the LOD community and standardisation organisations (W3C, ISO TC37/SC4). We established contacts and exchanged strategies with representatives of multilingual language communities beyond our own continent, most notably those of South Africa and India. The community building aspect also brought in LT industry which constitutes roughly one third of META's membership base. The industry's engagement is also demonstrated by our META-FORUM side-event META Exhibition which is mainly targeted at industry and which has always been booked to capacity.

META-SHARE had a huge impact on the community and is receiving and distributing an increasing number of language resources. It also consists of an increasing number of nodes as more and more organisations join the META-SHARE network in order to distribute their resources, such as, e. g., the Optima Activity of the Joint Research Centre (JRC) of the EC.

### 3.3. Impact on Politics and Funding Programmes

With regard to META-NET's impact on politics, administration and funding programmes, this section reports a selection of our successes in the past few years.

In April 2012, META-NET was invited to the EPP Working Group on Future Internet, "Internet Today and

Tomorrow", at the EP in Strasbourg. The WG unites MEPs from different committees to discuss Internet-related topics with consequences on the industry, culture, and legal affairs. This meeting led to sustainable contacts with several MEPs.

Our preferred option for the action suggested in the SRA is to set up a shared programme between the EC and the Member States as well as Associated Countries. First steps have been taken at META-FORUM 2012, when representatives of several funding agencies (BG, CZ, FR, HU, NL, SI) expressed the urgent need for such a shared programme (META-NET, 2012). This openness towards our topic we also see in discussions with politicians and administrators, for example, we managed to collect a multitude of quotes from high-ranking national politicians – several heads of states, several ministers – for the purpose of promoting the White Paper Series.

The EC is currently starting two funding programmes, Connecting Europe Facility (CEF) (EC, 2014) and Horizon 2020 (EC, 2012). As anticipated in the META-NET SRA, major aspects of our suggested European Service Platform for Language Technologies will be developed in the framework of CEF while the more concrete and applied research – the priority research themes (section 2.2.) – will be taken care of in H2020. The Digital Component of CEF (overall earmarked budget: 1B€) consists of five Digital Service Infrastructures (DSI) as essential building blocks that would serve and enable other DSIs. One of these core DSIs is, according to the amended proposal on the CEF Guidelines (EC, 2013), "automated translation". The inclusion of MT services as one of the five obligatory building blocks clearly is a milestone and important achievement of our field. It shows that the European Institutions recognise that they have the potential of overcoming language barriers, creating a single digital market.

In Dec. 2013 the first calls of H2020 were published. Of specific relevance for multilingual technologies is ICT-17 ("Cracking the language barrier") with a budget of 15M€. Even though the spirit of the call goes in the right direction, the budget is rather low. Nevertheless, the call text references key results of the White Paper Series, asking proposals to focus upon the 21 European digitally endangered languages; research projects are asked to focus upon the area of high-quality translation. By intensifying our dissemination work, especially around the SRA, we are confident that we can help to increase the budget available in future calls.

We established good connections with the cabinet of Commissioner Neelie Kroes, Vice-President of the EC, responsible for the Digital Agenda. A representative of her cabinet, Thibaut Kleiner, gave an opening speech at META-FORUM 2012. In a video message, Commissioner Kroes addressed the participants of LREC 2012. In another message, addressing the participants of the

LT Innovate Summit, she highlighted the opportunity of translation technologies to unlock the single digital market, also mentioning and acknowledging the work of META-NET (Kroes, 2013). A face-to-face meeting between Commissioner Kroes and representatives of the European LT community took place in October 2013. A meeting of the EU National Parliaments was held in the Lithuanian Parliament in Sept. 2013 (Vilnius-Meeting, 2013). The participants encouraged and welcomed initiatives which prioritise funding aimed at the reduction of the digital divide so that all European languages would come closer to the minimum EU-standard for LT development; apply PPP principles in designing and developing multilingual digital resources and tools; enable consistent integration of smaller languages into multilingual scientific infrastructures related to the design and development of digital resources and language technology, and ensure, at the European level, sustainability and international standards of digital language resources and technologies that build them.

In early Dec. 2013 the workshop “State of the Art of Machine Translation” took place in the EP, in which two representatives of META-NET as well as, among others, the EC and the EP presented their plans (STOA, 2013). In its resolution the participants agree on an urgent need for, among many other points, “establishing a Translingual Cloud services platform for all official EU languages and many additional languages”.

#### 4. Impact on the National Levels

**Austria:** META-NET has been serving as a catalyst for community building in pursuing common goals of improving the availability of and interoperability among LRs for research, industry, and eGovernment. Another strategic goal was achieved in closely linking national META-NET activities with the Austrian chapter of CLARIN (CLARIN-AT) in the context of Digital Humanities research as well as boosting projects for LR development and re-use for economic purposes, thus achieving synergies among inter-related activities.

**Belgium:** Flanders, as partner in the Dutch Language Union has in the past already invested considerably in LT R&D, among others by co-financing the STEVIN programme (Spyns and Odijk, 2013) which brought Dutch and Flemish LT R&D to a high level. The responsible civil servants at EWI (the Flemish government, Department for economy, science and innovation) use the META-NET documents (Rehm and Uszkoreit, 2012; Rehm and Uszkoreit, 2013; Odijk, 2012) in internal policy documents.

**Bulgaria:** The White Paper (Blagoeva et al., 2012) was used extensively to disseminate information about META-NET. The press release resulted in large amounts of media coverage in 2012. The META-NET roadshow in Sofia (May 2, 2012) was attended by 150 participants

and featured invited speakers from the EC and the Ministry of Education, Youth and Science.

**Croatia:** META-NET enabled the development of new LRs for Croatian as well as improving and standardising existing resources. The community was assembled around an international META-NET event held on Nov. 30, 2012 in Zagreb. This conference, organized under High Patronage of the President of the Republic of Croatia, gathered representatives from all Croatian research centres involved in computational linguistics and LT. The media attention helped promoting LTs for Croatian, particularly with respect to Croatian having become the 24th official language of the EU.

**Czech Republic:** The impact of META-NET in the Czech Republic is most visible in the area of LRs, where two infrastructural projects have been established: LINDAT and the Czech National Corpus project. Even though both the development of LRs and the Czech National Corpus existed before, now both activities have secured long-term funding from the Ministry of Education for the collection, development and preservation of LRs. More than 70 large, specialized resources have been made available through LINDAT; its metadata structure is compatible with both META-SHARE and CLARIN, it is being harvested by both networks.

**Denmark:** The White Paper (Pedersen et al., 2012) has resulted in comprehensive awareness in the society of the use of LT and the need for action. The idea of a Danish LT/LR collection has matured together with continued support to infrastructures in digital humanities (via the nationally funded DigHumLab and through CLARIN ERIC) and maintenance of META-SHARE. As a new initiative, an inter-institutional research project on semantic processing of Danish has been launched in Sept. 2013 with a grant from the Danish Research Council.

**Estonia:** The White Paper (Liin et al., 2012) found much response in the press, as the Estonian people tend to be vulnerable about the future of Estonian. This is also why the government had started, in 2006, to fund the National Programme for Estonian Language Technology (NPELT). The Center of Estonian Language Resources, the Estonian CLARIN partner, has a duty to deposit all NPELT results as resources and tools for preservation and long term access. The local META-SHARE node provides a suitable environment for this.

**Finland:** The impact of META-NET is visible in the strengthening of the Language Bank of Finland and its collection of LRs/LTs. These activities have secured long-term funding from the Ministry of Education for the collection, development and preservation of LRs/LTs. More than 190 specialized or large-scale resources have been made available via META-SHARE and the Language Bank of Finland, which is also part of CLARIN, and serves both humanities research and LT

needs. Awareness of a need for action has been raised thanks to the White Paper (Koskenniemi et al., 2012).

**France:** The White Paper (Mariani et al., 2012) was promoted widely and articles were published in widely distributed scientific journals. A long paper describing breakthroughs in MT, mentioning a number of European initiatives, was published in *Les Echos*, the leading French business and economics daily, on June 1, 2013. The agency in charge of the French language (DGLFLF) created a new position on LT and expressed great interest in the White Papers. The topic of LT for multilingualism was discussed at the “Multilingualism in the Overseas” summit, which took place in Guyana in Dec. 2011. DGLFLF is now proposing to initiate an interministerial national program with the aim of developing LTs and LRs for French and regional languages.

**Germany:** The press release and White Paper (Burcharth et al., 2012) generated a lot of interest and led to dozens of articles and multiple radio interviews. The German LT Day (Jan. 24, 2013) was attended by ca. 300 guests. The parallel industry exhibition was booked out almost immediately after the announcement. Representatives of almost all universities and research centres active in Germany in Computational Linguistics and related areas attended the event, so did several politicians and funding agencies. Another result of the dissemination work were many invited keynote presentations at events in Germany (and elsewhere) between 2010 and 2014. Additionally, through META-NET, we have been able to intensify discussions with two ministries and funding agencies around our topic. Two new funding programmes are currently on their way, initiated by the German Federal Ministry of Economics and Technology and the German Federal Ministry of Research. While Big Data is at the core of both programmes, LT is included as well, i. e., Text Analytics.

**Greece:** META-NET has had catalyzing effects on reinforcing interest in LT in Greece and demonstrating the importance of language in the digital age. The White Paper (Gavrilidou et al., 2012) has generated lasting interest in the media, as well as in the Government, at ministerial level. This interest has led to a substantial improvement of the position of LT in the Greek research agenda and the national research infrastructures roadmap. The strategic planning activities coupled with META-SHARE (Piperidis, 2012) have greatly helped in building bridges with the Greek R&D communities active in linked open data, among others, leading to joint initiatives. The extension of the infrastructural initiative and the advocacy of language processing service platforms has drawn the attention of many SMEs and spin-offs, currently seeking LRs/LTs in various areas of text analytics and multilingual content processing.

**Hungary:** A series of roadshows was organised, involving decision makers, the media and local industry.

These one day events visited each of the six countries participating in CESAR at monthly regularity. They not only mobilised the whole LT community but received significant attention in the local media. The series culminated in the Budapest roadshow which found six state and government dignitaries (four of them of ministerial rank) sitting around a table, addressing the conference one by one. The conference had the title “The position of the Hungarian language in the Digital Age”, thus, proving that appealing to national pride in connection with the mother tongue can have a significant impact.

**Iceland:** Almost all LRs for Icelandic are now available through META-SHARE. This is a major achievement since many of these resources had either been unavailable up to now or only available through personal contacts. Public awareness of the importance, challenges and opportunities of LT for Icelandic increased greatly. The White Paper (Rögnvaldsson et al., 2012) received considerable attention in the media and was taken up for discussion in the Icelandic Parliament (Alþingi). According to a parliamentary proposal put forward in Jan. 2014, a specially appointed committee of experts on linguistics and computer science shall be given the mandate to devise an action plan for the development of Icelandic LT, with the aim of making Icelandic as widely usable as possible in the digital world. Since the proposal is supported by all political parties, it is expected that it will be approved in March.

**Ireland:** The White Paper (Judge et al., 2012) campaign generated awareness around LT in Ireland and received significant media coverage in many outlets including interviews, online and print articles and radio interviews. As recently as early Oct. 2013, some of our researchers were in meetings with government officials from the department overseeing the implementation on the “20 Year Strategy for the Irish Language”, presenting to them ways in which LT can help. The importance of LT for Ireland’s growth as a knowledge economy was underlined recently in the government and industry investing €19.8 million in CNGL intelligent content research. This investment supports 75 research jobs and indirectly supports many more and includes an investment of €6.3 million from 16 industry partners including Symantec, Microsoft, McAfee, Welocalize and Intel as well as €1.25 million in venture capital for CNGL spin out companies employing an additional 30 people.

**Italy:** The White Paper (Calzolari et al., 2012) generated significant interest in Italy. As a result, on two occasions the achievements of META-NET have been presented in meetings organized in the Italian Parliament. Almost all existing LRs for Italian are now available through META-SHARE which has become the natural repository for the dissemination of LRs produced within EU projects. One example is the FP7 project SAVAS, focusing on innovative products and services for mul-

tilingual subtitling, which has decided to distribute all datasets through META-SHARE. The project requires the collection of large datasets of training data from media partners; both raw data and their derivatives have significant commercial value for the speech community and require licensing agreements before being released. META-SHARE allows them to share resources with each other, while facilitating appropriate licenses for each resource when made available to external users.

**Latvia:** The White Paper and the conference “Language, Technologies and the Future of Europe” attracted significant interest from Latvian politicians, decision makers and funding agencies, researchers, developers and users (Skadiņa et al., 2012; Vasiljevs and Skadiņa, 2012). It was widely covered by Latvian and Russian media. The findings of the white paper were broadly discussed and initiated practical actions by language policy makers. For instance, the Latvian language agency formed a working group of researchers, policy makers and experts to create a strategy for LT development and support. LTs were among the topics discussed by the Prime Ministers of UK, Nordic and Baltic countries at the Northern Future Forum in Riga (Feb. 2013). Work on the most critical LT areas (speech, MT, semantic analysis) continues with support from Structural Funds. The importance of LT has been also recognized by the public sector: since the end of 2013, Tilde participates in a project to provide large-scale LRs for MT and to develop and maintain MT services for the integration into Latvia’s e-Government infrastructure.

**Lithuania:** META-NET allowed the improvement and standardization of LRs and made them available through META-SHARE. As reported in the White Paper (Vaišnien and Zabarskaitė, 2012) advanced technology requiring more thorough knowledge of linguistic processing and semantics is still in its infancy. The positive effects with respect to the Lithuanian government and research institutions lead to the obvious need to prepare a united strategy. The META-NET dissemination campaigns had a great influence in initiating a national LT strategy. Guidelines on the development of Lithuanian in ICT for the period 2014-2020, prepared by the State Commission of the Lithuanian Language, are still to be approved. Their main objective is to develop LT support for Lithuanian at the level that will enable the language to successfully function in the digital age.

**Malta:** META-NET has had two main positive effects upon the state of LT in Malta. First, a number of novel and useful LRs for Maltese have been created. Second, thanks to META-SHARE, these resources are now not only easy to discover but available for download under licenced conditions. These two factors alone have succeeded in raising the profile of LT, thus opening the way to better exploitation of its potential within the public and private sectors in the medium-term.

**Netherlands:** The White Paper (Odijk, 2012) and press release have generated a lot of attention, through newspapers, journals, blogs, and tv reports. The main findings were presented in combination with the launch of new versions of Nuance’s dictation software for Dutch. The awareness of the importance of LT for consolidating the position of the Dutch language in the digital age, which was already high through the STEVIN programme, has been further increased. The white paper has been brought to the attention of the Dutch Language Union, so that it can contribute to the policy for the Dutch language. Though the organisation of funding schemes is currently prohibitive of a national LT programme, LT R&D is further explored in CLARIN-NL, CLARIAH-SEED, and other projects.

**Norway:** The White Paper (Smedt et al., 2012a; Smedt et al., 2012b), together with other dissemination activities, was well received in Norway. The LR/LT collection for Norwegian (Språkbanken) was established by the government earlier, but was not very visible until a close cooperation with the META-NORD partner at the University of Bergen led to more resources to be included as well as complete metadata and licensing, and searchability through META-SHARE. The resources have also been useful input to CLARIN which has received national funding and is continuing best practice for making even more resources available.

**Poland:** The White Paper (Miłkowski, 2012) and META-NET’s goals were very well received in Poland. The community was mobilised through the “Human Language Technology Days 2012” (Sept. 27/28) under the Honorary Patronage of the Ministry of Science and Higher Education and the President of the Polish Academy of Sciences. The event gathered representatives from all Polish research centres involved in LT. Accompanying events included an LT Corner, an industrial Roadshow (Kilgray Translation Technologies, Samsung and SAS and others) and research projects presenting their work. The event attracted a lot of media attention and helped promote knowledge about LT and its potential on the national level. Many of the 76 new and upgraded LRs/LTs made available are currently further extended, going beyond the end of the initiative. A notable example is the national co-operation between research institutions in the new CLARIN.PL project.

**Portugal:** META-NET represents a major landmark in the area of LRs/LTs for the Portuguese language. A significant set of LRs/LTs developed at the top ranking centers working on Portuguese, both in Portugal and in Brazil, were made available in a concerted way. The White Paper (Branco et al., 2012) induced a wave of dissemination and awareness raising actions, in different sectors, that is still unfolding: the workshop where the white paper was launched, and its core message, had a widespread and significant impact in the media and in

the industry sector; it was a key factor in community building, of which the recent application for a national research infrastructure to support HLT is a notable outcome; the white paper has been serving as the key to meetings with top ranking decision makers, in the purpose of lobbying for having the Portuguese Language as one of the strategic scientific challenges in the national R&D plan for 2014-2020, which is in preparation.

**Romania:** As a result of META-NET, more than 70 LRs/LTs for Romanian have been made public and lots of projects were initiated. In 2014, the Romanian Academy approved an ambitious project on building a very large (300+ million words), reference corpus of contemporary Romanian. It is a project between the two IT institutes in the Romanian Academy (Research Institute for AI in Bucharest, Institute of Computer Science in Iasi) and it brings together several large publishing companies, news agencies, radio and TV broadcasting companies. The project has been officially launched on Feb. 3, 2014, at the Romanian Academy in the presence of a large audience of members of the Academy, researchers in humanities and language industry. The event has been very favorably described in the media. The corpus will be publicly launched in 2017.

**Serbia:** META-NET allowed the development of new LRs for Serbian as well as improvement and standardization of existing resources. The LT community was mobilized through the “Human Language Technology Day” held on Oct. 29, 2012 in Belgrade. The event gathered representatives from all Serbian research centers involved in language processing, as well as many representatives of academia and industry. It attracted a lot of attention of both traditional and digital media and helped promote LT for Serbian. A similar event was organized on Nov. 2, 2013 in Belgrade, “35 years of Computational Linguistics in Serbia”.

**Slovakia:** The survey, consolidation and update of Slovak LRs signified a major boost to LRs/LTs (the Slovak National Corpus went up from 770 million tokens in 2011 to 1200 million by the end of the project to 2500 million in 2013). Together with the White Paper (Šimková et al., 2012) and better availability and visibility of LRs/LTs it increased general awareness of modern language-related research among both academia and government organisations. The government realised the need for an increased support of language related technologies and created a special government programme “Building Slovak National Corpus and Digitalization of Language Research in Slovakia, 3rd period”. The project (a continuation of the Slovak National Corpus) has been signed in Jan. 2013, but was applied retroactively from Jan. 2012 to 2016.

**Slovenia:** The White Paper (Krek, 2012) had a major impact on language policy activities and ultimately led to the inclusion of LT topics in the Resolution on the

National Programme for Language Policy 2014–2018 which was passed in the Parliament in July 2013. Some of the findings from the White Paper were included in its LT part. Through various META-NET activities the need for a more systematic support of LTs for Slovene was recognized by the government which is now in the process of adopting a five-year action plan on language infrastructure development. META-SHARE enabled the results of a major LT project (Communication in Slovene) funded by the European Social Fund and the Ministry of Education, Science and Sports to be included in an international LRs/LTs repository.

**Spain:** META-NET raised awareness about the importance of LTs/LRs at the Spanish national and regional levels. The warm reaction that companies have given to the availability of resources, especially those related to regional languages, has been noticeable. Despite the economic situation, which has considerably reduced the governmental budget for R&D, support of the language infrastructures, including META-SHARE, is still guaranteed by several ongoing projects like the deployment of a UPF-Competence Centre for LTs in connection with CLARIN, co-funded by EU Structural Funds and the Generalitat de Catalunya, or SpeechTech4All, funded by the government, that focus on LTs for all official languages of Spain.

**Sweden:** META-NET generated increased awareness of the central role of LT for the future of the Swedish language. The White Paper (Borin et al., 2012) played a central role in the preparation of a report commissioned by the Department of Culture, on present and future societal needs for LRs and LT (ISOF, 2012). The publication of the white paper and arrangement of a national workshop served as a powerful reinforcement of the report’s message. A large number of LRs are now available through META-SHARE. The activities have done a great deal towards reducing fragmentation in the Swedish LT community, instilling a sense of common cause on the national level, as evidenced by a number of joint proposals and awarded grants. The Swedish Research Council has approved a proposal for Swedish membership in the CLARIN ERIC, negotiations have started about setting up SWE-CLARIN.

**United Kingdom:** META-NET has been successful in promoting awareness about the importance of LT in the UK Government. This has been demonstrated by endorsements and interest from both the House of Commons (David Willetts MP, Minister of State for Universities and Science) and the House of Lords (Baroness Coussins, Chair, All-Party Parliamentary Modern Languages Group) following dissemination of the White Paper (Ananiadou et al., 2012). The UK’s Engineering and Physical Sciences Research Council (EPSRC) conducted a “Shaping Capability” review of which subject areas to retain for funding, which to increase and which

to reduce. A decision was made to maintain funding for NLP research. Amongst the evidence cited by EPSRC for this decision was the META-NET SRA.

## 5. Conclusions and Next Steps

The intense communication and dissemination work carried out by META-NET had a significant impact on the European LT scene, helped to provide focus and to shape several national language policy and development strategies. We have also been able to provide input to CEF and Horizon 2020. While the first funded phase of META-NET and its four projects is over, the initiative itself is continuing its work and established brand under the umbrella of new projects. Of specific importance are increasing the membership base of META (especially industry) and the goal of setting up a shared programme between the EU/EC and Member States as well as Associated Countries.

META-NET was co-funded by FP7 and ICT-PSP of the European Commission through the contracts T4ME (grant agreement no.: 249 119), CESAR (no.: 271 022), METANET4U (no.: 270 893) and META-NORD (no.: 270 899).

## 6. References

- Ananiadou, S., McNaught, J., and Thompson, P. (2012). *The English Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Blagoeva, D., Koeva, S., and Murdarov, V. (2012). *Българският език в дигиталната епоха – The Bulgarian Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Borin, L., Brandt, M.D., Edlund, J., Lindh, J., and Parkvall, M. (2012). *Svenska språket i den digitala tidsåldern – The Swedish Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Branco, A., Mendes, A., Pereira, S., Henriques, P., Pellegrini, T., Meinedo, H., Trancoso, I., Quaresma, P., de Lima, V.L. Strube, and Bacelar, F. (2012). *A lingua portuguesa na era digital – The Portuguese Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Burchardt, A., Egg, M., Eichler, K., Krenn, B., Kreutel, J., Leßmöllmann, A., Rehm, G., Stede, M., Uszkoreit, H., and Volk, M. (2012). *Die Deutsche Sprache im digitalen Zeitalter – German in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Calzolari, N., Magnini, B., Soria, C., and Speranza, M. (2012). *La Lingua Italiana nell’Era Digitale – The Italian Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- EC. (2010). A Digital Agenda for Europe. European Commission. [http://ec.europa.eu/information\\_society/digital-agenda/publications/](http://ec.europa.eu/information_society/digital-agenda/publications/).
- EC. (2012). Horizon 2020: The Framework Programme for Research and Innovation. European Commission. <http://ec.europa.eu/research/horizon2020/>.
- EC. (2013). Amended proposal for a Regulation of the European Parliament and of the Council on guidelines for trans-European telecommunications networks and repealing Decision No 1336/97/EC. European Commission. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2013:0329:FIN:EN:PDF>.
- EC. (2014). Connecting Europe Facility. European Commission. <https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/cef.pdf>.
- Evas, J. (2014). *Y Gymraeg yn yr Oes Ddigidol – The Welsh Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Gavrilidou, M., Koutsombogera, M., Patrikakos, A., and Piperidis, S. (2012). *H Ελληνική Γλώσσα στην Ψηφιακή Εποχή – The Greek Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- ISOF. (2012). Infrastruktur för språken i Sverige. Förslag till nationell språkinfrastruktur för det digitala samhället. Infrastructure for Sweden’s languages. Report commissioned from the Swedish Language Council by the Swedish Department of Culture.
- Judge, J., Chasaide, A.N., Dhubhda, R.N., Scannell, K.P., and Dhonnchadha, E.U. (2012). *An Ghaeilge sa Ré Dhigiteach – The Irish Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Koskeniemi, K., Lindén, K., Carlson, L., Väinö, M., Arppe, A., Lennes, M., Westerlund, H., Hyvärinen, M., Bartsis, I., Nuolijärvi, P., and Piehl, A. (2012). *Suomen kieli digitaalisella aikakaudella – The Finnish Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Krek, S. (2012). *Slovenski jezik v digitalni dobi – The Slovene Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Kroes, N. (2013). Speech of Neelie Kroes at LT-Innovate Summit 2013, July. <http://www.youtube.com/watch?v=XH-iRBdv3ho>.
- Liin, K., Muischnek, K., Müürisepp, K., and Vider, K. (2012). *Eesti keel digiajastul – The Estonian Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Mariani, J., Paroubek, P., Francopoulo, G., Max, A., Yvon, F., and Zweigenbaum, P. (2012). *La langue française à l’Ère du numérique – The French Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- META-NET. (2012). META-FORUM 2012: A Strategy for Multilingual Europe. Panel discussion “Plans for LT Research and Innovation in Member States and Regions”. Videos available at <http://www.meta-net.eu/events/meta-forum-2012/programme>.
- META-NET. (2013). META-FORUM 2013: Connecting Europe for New Horizons. Panel discussion “Connecting Towards New Horizons”, September. Videos available at <http://www.meta-net.eu/events/meta-forum-2013/programme>.
- Milkowski, M. (2012). *Język polski w erze cyfrowej – The Polish Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Odiijk, J. (2012). *Het Nederlands in het Digitale Tijdperk – The Dutch Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Pedersen, B. Sandford, Wedekind, J., Böhm-Andersen, S., Henrichsen, P.J., Hoffensetz-Andersen, S., Kirchmeier-Andersen, S., Kjørnum, J.O., Larsen, L.B., Maegaard, B., Nimb, S., Rasmussen, J.-E., Revsbech, P., and Thomsen, H. Erdman. (2012). *Det danske sprog i den digitale tidsalder – The Danish Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Piperidis, S., Papageorgiou, H., Spurk, C., Rehm, G., Choukri, K., Hamon, O., Calzolari, N., del Gratta, R., Magnini, B., and Girardi, C. (2014). META-SHARE: One year after. In *Proc. of the 9th Language Resources and Evaluation Conf. (LREC 2014)*, Reykjavik, Iceland.
- Piperidis, S. (2012). The META-SHARE Language Resources Sharing Infrastructure: Principles, Challenges, Solutions. In *Proc. of the 8th Language Resources and Evaluation Conf. (LREC 2012)*, Istanbul, Turkey.
- Rehm, G. and Uszkoreit, H., editors. (2012). *META-NET White Paper Series: Europe’s Languages in the Digital Age*. Springer, Heidelberg etc. 32 volumes on 31 European languages.
- Rehm, G. and Uszkoreit, H., editors. (2013). *The META-NET Strategic Research Agenda for Multilingual Europe 2020*. Springer, Heidelberg etc.
- Rögnvaldsson, E., Jóhannsdóttir, K.M., Helgadóttir, S., and Steingrímsson, S. (2012). *Íslensk tunga á stafrænni öld – The Icelandic Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Skadiņa, I., Veisbergs, A., Vasiljevs, A., Gornostaja, T., Keiša, I., and Rudzīte, A. (2012). *Latviešu valoda digitālajā laikmetā – The Latvian Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Smedt, K. De, Lyse, G.I., Gjesdal, A. Müller, and Losnegaard, G.S. (2012a). *Norsk i den digitale tidsalderen (bokmålsversjon) – The Norwegian Language in the Digital Age (Bokmål Version)*. White Paper Series. Springer, Heidelberg etc.
- Smedt, K. De, Lyse, G.I., Gjesdal, A. Müller, and Losnegaard, G.S. (2012b). *Norsk i den digitale tidsalderen (nynorskversjon) – The Norwegian Language in the Digital Age (Nynorsk Version)*. White Paper Series. Springer, Heidelberg etc.
- Spyns, P. and Odiijk, J., editors. (2013). *Essential Speech and Language Technology for Dutch. Results by the STEVIN-programme*. Theory and Applications of Natural Language Processing. Springer.
- STOA. (2013). European Parliament, Science and Technology Options Assessment: Workshop “State of the Art of Machine Translation: Current Challenges and Future Opportunities”. <http://www.europarl.europa.eu/stoa/cms/cache/offonce/home/events/workshops/translation>.
- Vaišnien, D. and Zabarskaitė, J. (2012). *Lietuvių kalba skaitmeniniame amžiuje – The Lithuanian Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.
- Vasiljevs, A. and Skadiņa, I., editors. (2012). *Latvian Language Resources and Tools: Assessment, Description and Sharing. Human Language Technologies – The Baltic Perspective – Proceedings of the Fifth International Conference Baltic HLT 2012*.
- Vilnius-Meeting. (2013). Resolution of the conference “united in diversity: Political and social development aspects of eu languages and cultures”. Vilnius, Lithuania. <http://renginiai.lrs.lt/renginiai/EventDocument/98af77a5-8014-4a6e-afe5-55c3f26fe0bf/Vilnius%20Conclusions%20-%20Education%20and%20Information%20Society%20EN.pdf>.
- Šimková, M., Garabík, R., Gajdošová, K., Laclavik, M., Ondrejovič, S., Juhár, J., Genčí, J., Furdík, K., Ivoriková, H., and Ivanecký, J. (2012). *Slovenský jazyk v digitálnom veku – The Slovak Language in the Digital Age*. White Paper Series. Springer, Heidelberg etc.