

Newsletter

Number 1, 2012, April-June

What is META-NORD?

To describe META-NORD in a few words is a difficult task, given the complexity and range of the project. The project is funded by the European Commission within the scope of the Information and Communication Technologies Policy Support Programme. The aim of the project is the establishment of an open linguistic infrastructure ensuring the development of language technologies and resources. Thus META-NORD serves the needs of the LT developing community.

The META-NORD project is composed of members from the eight countries of the Baltic and Nordic region of Europe: Denmark, Estonia, Finland, Iceland, Latvia, Lithuania, Norway and Sweden. All

of these countries have less than 10 million speakers and the situation of the language technology developed for their national languages is in some cases so critical that it can determine the decline of the languages in technologically based multilingual web. This is the challenge of META-NORD: to influence the use of these languages by developing technologies, assembling, linking, and making widely available language resources of different types used by different categories of user communities in academia and industry.

META-NORD is strongly related with META-NET, The Network of Excellence which is forging the Multilingual Europe

Technology Alliance (META). META-NET aims at consolidating the European research and industry communities engaged in development of multilingual LT. For this reason META-NET has created META-SHARE, an open resource and technology sharing and exchange facility. META-NORD and its related projects CESAR and METANET4YOU coordinate their activities with META-NET and contribute to the development of META-SHARE.

Dear reader,

We are proud to present you the first number of the META-NORD newsletter. From now on the newsletter will be issued almost every other month, depending on the number of news related to our project. We hope that this newsletter will help you find out about the initiatives of the META-NORD partners and their achievements as well as about the most important LT events in Europe.

On the first page you will find the description of the project. We hope that most of the people receiving this newsletter are already acquainted with META-NORD, but if not, this article will provide basic facts about the purpose, chal-

lenge and the scope of our linguistic initiative.

The newsletter aims at presenting META-NORD as an active and social project. This issue is almost entirely dedicated to events, workshops and conferences attended by project members and where META-NORD was presented. Of utmost significance was the META-FORUM 2012 conference organized in June by META-NET. At this event an important part of the LT research and industry community gathered for a conference that lasted two days. The related articles in this issue reveal the technical and linguistic importance of this conference.

After the articles on META-FORUM you will find descriptions of other major LT events like CHAT 2012 and LREC 2012.

You can also read here about the national workshops organized by the META-NORD partners in Scandinavia and the Baltic states.

The last page contains a list of organizations already involved in the META community in the Baltic and Nordic countries. We hope that you can find your organization in the list. If not, it is still not too late to register it at <http://www.meta-net.eu/join-forms/organisation-form>.

Enjoy your reading!

*Imre Bartis, University of Helsinki
Dovilė Bielevičiūtė, Institute of the
Lithuanian Language*

META FORUM 2012

An international conference on powerful technologies for the multilingual European information society. This year META-FORUM 2012 took place in Brussels, Belgium on June 20-21

The topic of the conference was “A Strategy for Multilingual Europe”. This idea was reflected in the sessions topics: “Industry and Innovation – Language Technology Made in Europe” and “Language technologies for Europe 2020”. The winners of the META Prize and the META Seal of Recognition chosen by the META Technology Council were announced during the final session of the first day.

The conference hosted the META Exhibition which included 40 exhibits, booths and stands from research and industry. The exhibition was a good opportunity for LT companies to network and showcase their work on a European stage where they were seen by decision makers, MEPs and other VIPs.

META-NORD was presented by the project coordinator Andrejs Vasiljevs (Tilde, Latvia) in the session “METASHARE: An Open Resource Exchange Infrastructure”. In his presentation Andrejs Vasiljevs emphasized the fact that the META-SHARE language resource exchange facility is devoted to the sustainable sharing and dissemination of LRs between its members and the community at large. META-SHARE thus increases the access to LRs at a global scale by supporting the harmonization of the laws governing copyright exceptions and limitations, fair use and fair dealing. How to ensure the sustainability of work done so far by harmonizing language resources after the EU funding stage is an issue of high importance, which must be clearly understood by all EU funded LT projects.

Andrejs Vasiljevs delivered also another presentation on translation clouds in the

session “Language Technologies for Europe 2020”. The META-NORD project partner Bolette Sandford Pedersen (University of Copenhagen, Denmark) presented “The META-NET Language White Paper Series: Overview and Key Results” in the session “Europe and its

Languages”. The conference offered an excellent opportunity for all the METANORD partners to change ideas with other related EU project members and to find out more about LT companies, language research and industry possibilities in other countries.



We hope that the tremendous work of all those who made the conference happen and especially Prof. Hans Uszkoreit, Network Manager Georg Rehm, as well as the META-NET Communication Group led by John Judge will be rewarded. Their effort to ensure that META - A Network of Excellence forging the Multilingual Europe technology Alliance would attract more and more members and would become richer in innovative ideas on the future of language technology in Europe is remarkable. You can read more about the programme and the events of META FORUM 2012 at <http://www.meta-net.eu>.

*Audronė Bielevičienė,
Institute of the Lithuanian Language*



Impressions from META-FORUM 2012



As expected, META-FORUM 2012 was rich in interesting talks and inspirational presentations. The introductory talk by András Kornai (Hungarian Academy of Sciences, Budapest), “Language Death in the Digital Age” for example was a compelling and to-the-point presentation of a topic that unfortunately does not seem to be taken seriously enough by many governments, namely the importance of language technology in ensuring the future of languages.

Another presentation that impressed me was Lori Thicke’s (Lexcelera Localization, France) “Localization and Language Services”. Thicke spoke passionately about ways for an LT company to be socially responsible and demonstrated that LT services can actually save lives.

Finally, I would like to mention one more speaker whose convincing rhetoric did not seem to pass unnoticed to the audience. In his short presentation Karel Oliva, the head of Institute of the Czech Language of the Czech Academy of Sciences pointed out that for those organizations involved in T4ME, CESAR, METANET4U and META-NORD that develop language technology for “small languages” it is quite difficult to attract funding from LT companies. Thus measuring success by taking into consideration only the question of how much money these organizations manage to obtain from the business sector would be quite unfair.

*Imre Bartis,
University of Helsinki*

A fantastic event – LREC 2012!

Istanbul, May 21-26, main conference May 23-25



The eighth international conference on Language Resources and Evaluation (LREC) was organized by ELRA with the support of a wide range of international associations and organizations. The event run smoothly, judging from the participants and interested parts. During the event an overview of the new R&D directions and emerging trends was provided, exchanged information regarding LR's and their applications, evaluation methodologies and tools, industrial uses and needs, requirements coming from the e-society. Therefore conference like previous ones achieved a great success in the community of the researchers and creators of LR's.

In this year conference all the members of the Meta-Net projects: META-NORD, T4ME, CESAR, META4U were pre-

sented combined and unified like all joint family. Therefore all the projects took a fair space in project village, in the main conference and together arranged a workshop. Also projects were presented in various poster and oral sessions. META-NORD was presented by Andrejs Vasiljevs (TILDE) in a special session "EU projects, a broad overview" and separately in a poster session "Creation of an Open Shared Language Resource Repository in the Nordic and Baltic Countries".

During this year LREC conference META-NORD together with other META-NET members organized the META-RESEARCH workshop on *Advanced Treebanking*.

META-RESEARCH Workshop on Advanced Treebanking was initiated equally by META-NORD partner University of Bergen (UiB). Workshop was organized by Jan Hajič (T4ME), Koenraad De Smedt (META-NORD), Marko Tadic (CESAR), Antonio Branco (META4U). Gyri S. Losnegaard and Victoria Rosén (UiB) presented a paper on the workshop

with the title "What We Have Learned from Sofie: Extending Lexical and Grammatical Coverage in an LFG Parsebank", written by Gyri Losnegaard, Gunn Inger Lyse, Martha Thunes, Victoria Rosén, Koenraad De Smedt, Helge Dyvik and Paul Meurer. The paper presented some findings from the delivery from project INESS to project META-NORD's first batch of resources.

The workshop as a whole addressed EU and world-wide projects and research groups that are creating, standardizing, converting and/or using treebanks, thereby often tackling the same issues and reinventing methods and tools.



The role of terminology in a modern life and...

...the second workshop on the Creation, Harmonization and Application of Terminology resources (CHAT 2012) which was held in conjunction with the conference on Terminology and Knowledge Engineering (TKE) on June 22, 2012 in Madrid, Spain

Every day the volume of terminology is growing along with the increasing volume of information available on the web. Efficient terminology acquisition and management has become an essential component of intelligible translation, localization, technical writing and other professional language work. The current models for finding, sharing and using terminology data cannot keep up with a growing demand in multilingual Europe. The role of terminology however is today more important than ever to ensure that people communicate efficiently and precisely.

Consistent, harmonized and easily accessible terminology is an extremely important prerequisite for ensuring unambiguous multilingual communication in the European Union and throughout the world. The planned workshop aimed at bringing together academic and industrial players in the area of terminology and attracting

holders of terminology resources. The workshop also focused on fostering the cooperation between EU projects and research and development activities in the area of terminology along with sharing experience and discussing recent advances of the consolidation, harmonization, and distribution of terminology resources, as well as their practical application in various domains.

META-NORD was presented briefly at the CHAT 2012 by Tatiana Gornostay. Presentation entitled as "META-NORD: Baltic and Nordic Parts of the European Open Linguistic Infrastructure". META-NORD coordinator, Andrejs Vasiljevs, gave a presentation "EuroTermBank – Towards Dedicated Terminology Services for European Linguistic Infrastructure". Also participated in the discussion panel "Terminology Resources as part of Euro-

pean Open Linguistic Infrastructure" moderated by Tatiana Gornostay.

This workshop was organized by FP7 projects TaaS (Terminology as a Service) and TTC (Terminology Extraction, Translation Tools and Comparable Corpora), and ICT-PSP project META-NORD (Baltic and Nordic Branch of the European Open Linguistic Infrastructure) to continue a series of meetings that started as the first CHAT 2011 workshop in Riga in May 2011.

The other two sister projects CESAR and METANET4U have been invited to take part in the workshop to insure coherence between the three projects and seek potential synergies this time within the consolidation of multilingual terminology resources.

*Tatiana Gornostay,
Tilde SIA*

National workshops on the road!

One of the most important dissemination channels for every project are national workshops. Through them the information about newest achievements of the subject is spread and specialists are involved into tackling the common obstacles and sharing of the common experience.

META-NORD project partners also organize national workshops to increase visibility of the status and development of language technologies and resources. In every country level of consummation is different; therefore workshops aim at different target groups and different topics. A lot of countries are planning their national workshops in the autumn thought some of the countries like Iceland and Norway have already arranged them.

National workshop in Norway

The Norwegian partner University of Bergen (UiB) in META-NORD was represented

at the Workshop on the Interoperability of Metadata on June 5, 2012 in Oslo, Norway. The workshop was organized by the National Library of Norway, in cooperation with CLARINO and META-NORD. Koenraad De Smedt (UiB) played a leading role in organizing the workshop and chairing it. There were two presentations by META-NORD partners, one by Gunn Lyse (UiB) on the META-NORD experience with metadata in META-SHARE, and one by Jussi Pii-tulainen (UHEL) on making META-SHARE and CLARIN work together. Gyri S. Losnegaard (UiB) also participated at the workshop. The main objective of the workshop was to coordinate various initiatives aimed at metadata catalogues of language research data, such as CLARINO, META-SHARE, Språkbanken and the National Library's digital collections.

Koenraad De Smedt,
University of Bergen

Cooperation between UGOT and the Swedish Language Council

Lars Borin, director of Språkbanken (UGOT), and Rickard Domeij, of the Swedish Language Council have agreed to hold a national workshop in the fall of 2012. The workshop will promote the importance of Language Technology for the Swedish language not only to make sure it survives in the multilingual cultures of Europe, but also to ensure it flourishes. Furthermore, the workshop will promote META-NORD and the META-SHARE repository by highlighting their role w.r.t. Language Technology and Swedish in the European context. The workshop will take place in Stockholm on November 23rd 2012.

Martha D. Brandt,
Språkbanken, University of Gothenburg

National Workshop in Iceland grew into a Conference

Instead of organizing a workshop, the META-NORD partners in Iceland joined efforts with The Icelandic Centre for Language Technology and the Icelandic Language Council and held a conference, offering variety of talks on LT and language preservation

April 27, 2012, the LT conference *Máltækni fyrir alla* "Language Technology for Everyone" (<http://www.malfong.is/Malthing/>) took place at the University of Iceland. One of the primary objectives of META-NORD is to secure users of any language equal access to information and knowledge, regardless of what language they speak. By joining forces with the Icelandic Centre for Language Technology and the Icelandic Language Council it was possible to reach wider audience and send the message further. The inclusion of participants like the Minister of Education, Culture and Science gave the conference more weight.

The balanced program of the conference provided something for everyone. The first part focused on LT in general and particularly LT and language preservation, and included talks on LT and language policies, databases and the META-

program. The second half introduced several LT resources that are being prepared for the language, such as speech synthesizers, speech recognizers, spell-checkers and search engines. The speakers came from different parts of the community: a minister, two chairmen and one vice chairman of language councils, the Deputy President of EFNIL, the manager of META-NORD in Iceland, university researcher, chairman of the Icelandic society for the blind, professors, a student and a representative from the industry.

Approximately 70 attendees were at the conference from all aspects of society and they were extremely pleased with the program and quite impressed with the number of LT resources available, even though all agreed that more is needed.

The importance of language technology is undisputed and within a small language community like Iceland, it is even

greater. A strong LT community is needed with researchers and developers creating and updating the language resources that are necessary to keep up with the fast changing world. This can only happen with increased support from government and society. As the conference attracted much attention from the media, in addition to the excellent attendance from the general public, it is safe to say that its goals to point out the issues and suggest solutions were met.

Kristín M. Jóhannsdóttir,
University of Iceland



More events with META-NORD

The conference “Crosslingual Language Technology in service of an integrated multilingual Europe - 20 years on”

The conference was held on 4-5 May, 2012 at the University of Hamburg, Germany. The conference was co-organized among others with the help of the META-NORD project. META-NORD was presented on May 5 during the section 5 "Language Resources and Tools in a Crosslingual environment", presentation "Recent advances in the development and sharing of language resources and tools for Latvian".

W3C workshop “The Multilingual Web – Linked Open Data and MultilingualWeb-LT Requirements”

On 11-13 June, 2012 in Dublin, Ireland W3C workshop took its place. At this workshop Tatiana Gornostay presented “Extending the Use of Web-Based Terminology Services”. The presentation focused on Tilde’s projects and advances (ACCURAT, LetsMT!, TTC, TaaS, META-NORD) in establishing cloud-based platforms for acquiring, sharing, and reusing language resources to improve automated natural language data processing (e.g., machine translation). Multilingual consolidated and harmonized terminology is already utilized as data in the process of human translation and now it is also being developed as a web-based service with machines (e.g. machine translation systems, indexing systems, and search engines) as users. This development has the potential to vastly enhance the degree of automation for linked open data technologies and reveal synergies between Linked Open Data and Multilingual Language Technologies.

During the workshop and “Terminology metadata” session, connections with the open data community were explored in order to increase the visibility of the published resources and tools outside the META-SHARE network

2nd Digital Agenda Assembly

The Digital Agenda for Europe is the European Union's roadmap for bringing the benefits of a digital society and economy to Europe's citizens. But it can only deliver if all stakeholders are involved in assessing problems and identifying solutions. The main objectives of the second Assembly on 21 June, 2012 were to assess progress to date on implementation towards the Digital Agenda's goals and actions and seek ways to improve delivery; to identify challenges ahead for the implementation of the Digital Agenda and for the information society in general; and to mobilize stakeholders' actions to make further progress and address challenges. Andrejs Vasiljevs participated in the “Workshop 5: Data” as a panelist during the session “Accessing and exploiting data across languages”.

LT-Innovate Summit 2012

On 19 June 2012, Brussels was the meeting point for Language Technology Industry to network, discuss needs, strategies, innovation opportunities and business trends. The LT-Innovate Summit brought together Language Technology vendors (speech, translation, intelligent content technologies) and other stakeholders (investors, buyers, integrators, researchers, policy makers). Andrejs Vasiljevs attended the Summit and participated in the two discussions as a panelist: “Innovation focus” session and “Showcase Presentation Session VI”.

Join META!

We invite institutions related with or interested in the development of LT to become a part of the Multilingual Europe Technology Alliance – META

www.meta-net.eu

Members

In the Nordic and Baltic countries

Country; Organization; Location; Representative

Denmark

Ankiro ApS; Copenhagen; Bo Vincents
Copenhagen Business School; Copenhagen;
Peter Juel Henriksen
Dansk Sprognævn; Copenhagen; Sabine
Kirchmeier-Andersen
GramTrams; Viby J; Eckhard Bick
Ordbogen; Odense; Peter Revsbech
RAP; København, Son; Keld Simonsen
Society for Danish Language and Literature;
Copenhagen; Jørg Asmussen
Technical University of Denmark; Lyngby; Jan
Larsen
The National Museum of Denmark; Copenha-
gen; Birgit Rønne
The Royal Library; Copenhagen; Anders Con-
rad
University of Copenhagen, Centre for Lan-
guage Technology; Copenhagen; Bente
Maegaard
University of Southern Denmark, Faculty of
Humanities; Kolding; Johannes Wagner
Wind Kommunikation; Copenhagen; Jørgen
Christian Wind Nielsen

Estonia

University of Tartu, Institute of Computer Sci-
ence; Tartu; Tiit Roosmaa

Finland

Aalto University, Computational Cognitive
Systems Research Group; Aalto; Timo Honkela
CSC, the Finnish IT Center for Science; Espoo;
Antti Pursula
Lingsoft Inc; Helsinki; Juhani Reiman
Sunda Systems Oy; Helsinki; Markku Kiiski
The Research Institute for the Language of
Finland; Helsinki; Toni Suutari
University of Helsinki; Helsinki; Martti Vainio

University of Helsinki, Department of General
Linguistics; Helsinki; Kimmo Koskeniemi
University of Joensuu; Joensuu; Jussi Niemi
University of Oulu; Oulu; Marketta Harju-Autti
University of Tampere; Tampere; Eero Sormu-
nen

Iceland

CLARA; Reykjavik; Jon Edvald Vignisson
HEILAHEILL; Reykjavik; Þórir Steingrímsson
Heyrnarhjálpi; Reykjavik; Sigurjón Einarsson
IceStat; Reykjavik; Snorri Guðmundsson
Málteknisetur (Icelandic Centre for Language
Technology); Reykjavik; Eiríkur Rögnvaldsson
Reykjavik University; Björn Þór Jónsson
Talþjálfun Reykjavíkur; Reykjavik; Þóra
Sæunn Úlfadóttir
The Árni Magnússon Institute for Icelandic
Studies; Reykjavik; Sigrún Helgadóttir
University of Iceland, School of Humanities;
Reykjavik; Eiríkur Rögnvaldsson

Latvia

ACCURAT; Riga; Aivars Berzins
Let's MT; Riga; Aivars Berzins
Tilde; Riga; Andrejs Vasiljevs
University of Latvia, Institute of Mathematics
and Computer Science; Riga; Inguna Skadiņa

Lithuania

Institute of the Lithuanian Language; Vilnius;
Jolanta Zabarskaitė
Vytautas Magnus University, Center of Com-
putational Linguistics; Kaunas; Rūta Marcinke-
vičienė

Sweden

ESTeam AB; Älgårås; Gudrun Magnusdóttir
Interverbium Technology; Linköping; Ioannis
Iakovidis
KTH; Stockholm; Joakim Gustafson
KTH Royal Institute of Technology; Stock-
holm; Rolf Carlson
Linköping University; Linköping; Lars Ahren-
berg
Lund University; Lund; Sven Strömqvist
MOLTO; Gothenburg; Emilia Rung
Swedish Institute of Computer Science; Kista;
Jussi Karlgrén
The Language Council of Sweden; Stockholm;
Rickard Domeij
Transmachina AB; Stockholm; Nicholas
Cottrell
Umeå University; Umeå; Patrik Svensson
University of Gothenburg, Department of
Swedish Language; Gothenburg; Lars Borin

University of Gothenburg, Department of
Swedish Language, CLT Dialogue Technology
Lab; Gothenburg; Staffan Larsson
Uppsala University; Uppsala; Joakim Nivre

Norway

Clapter Create; Stavanger; Tor Einar Enne
Clue Norge ASA; Oslo; Merete Kravik
CognIT a.s.; Oslo; Harald Falsen
Comperio; Oslo; Trond Renshusløyken
Computas; Lysaker; David Norheim
Dictatr; Bergen; Thomas Hagen
ESIS Norge AS; Oslo; Robert HP Engels
Include; Bergen; Sverre Andreas Hilditch Hol-
bye
Kaldera språkteknologi AS; Oslo; Lars Ny-
gaard
Lingit; Trondheim; Torbjørn Nordgård
MediaLT; Oslo; Magne Lunde
NAV The Norwegian Labour and Welfare Ad-
ministration; Oslo; Daniel Scheidegger
NHH Norwegian School of Economics; Ber-
gen; Gisle Andersen
Narvik Universit College; Narvik; Bernt Brem-
dal
National library of Norway; Oslo; Oddrun
Pauline Ohren
Norwegian School of Economics and Business
Administration; Bergen; Gisle Andersen
Norwegian University of Science and Technol-
ogy; Trondheim; Torbjørn Svendsen
Nynodata AS; Bø i Telemark; Bjørn Seljebotn
Opera Software; Oslo; Pål Eivind Nacobsen
Nes
SINTEF; Oslo; Diana Santos
Språkrådet; Oslo; Torbjørge Breivik
Standards Norway; Lysaker; Marit Sæter
Tansa Systems AS; Oslo; Henriette Edvarda
Berntsen
Tekstlaboratoriet, Universitetet i Oslo; Oslo;
Kristin Hagen
TextUrgy; Bergen; Trond Walker
The Language Council of Norway; Oslo;
Torbjoerg Breivik
The National Library of Norway; Oslo; Kristin
Bakken
Tobii Technology Norge AS; Bergen; Morten
Mjelde
Unifob AS; Bergen; Eli Hagen
Universitetet i Oslo; Oslo; Stephan Oepen
University of Tromsø, Det humanistiske fa-
kultet; Trond Trosterud
University of Bergen, Department of Linguis-
tic; Bergen; Koenraad De Smedt
Western Norway Research Institute; Sogndal;
Rajendra Akerkar

