# Baltic and Nordic Branch of the European Open Linguistic Infrastructure (META-NORD)

<u>Project ID card</u>
- Funded under: The Information and Communication Technologies Policy Support Programme
- Area: CIP-ICT-PSP-2010-4 Theme 6: Multilingual Web: Machine translation for the multilingual web
- Total cost: €4.5m
- EU contribution: €2.25m
- Project reference: 270899
- Execution: From 01/02/2011 to 31/01/2013
- Project status: Running
- Contract type: The Information and Communication Technologies Policy Support Programme PB Pilot Type B

<u>Challenge:</u>

In the last decade linguistic resources have grown rapidly for all EU languages, including lesser-resourced languages. However they are located in different places, have developed in different standards and in many cases are not well documented.

Currently users can conveniently benefit only from a very limited number of publicly available resources for commercial or non-commercial purposes. It is difficult or in many cases impossible to get access to resources that are scattered around different places, are not accessible online, reside within research institutions and companies. Many resources are in proprietary formats or have proprietary interfaces that further hinder their broader usage.

High fragmentation and a lack of unified access to language resources are among key factors that hinder European innovation potential in language technology development and research.

<u>Proposed solution</u>

The META-NORD project aims to establish an open linguistic infrastructure in the Baltic and Nordic countries to serve the needs of the industry and research communities. The project will focus on 8 European languages - Danish, Estonian, Finnish, Icelandic, Latvian, Lithuanian, Norwegian and Swedish - that each have less than 10 million speakers. The project will assemble, link across languages, and make widely available language resources of different types used by different categories of user communities in academia and industry to create products and applications that facilitate linguistic diversity in the EU.

<u>Project objectives:</u>

The **main objectives** of the META-NORD project as established under Objective 6.1 are:
- **to provide a description of the national (resp. language community) landscape** in terms of language use; language-savvy products and services, language technologies and resources; main actors (research, industry, government and society); public policies and programmes; prevailing standards and practices; current level of development, main drivers and roadblocks; and create this in a simple, clear, standardized format;
- to contribute to a pan-European digital resource exchange facility by **identifying**, **collecting resources in the Baltic and Nordic countries** and by documenting, processing, linking and upgrading them to agreed standards and guidelines;
- **to collaborate with** other partner projects, in particular **concurrent 6.1 pilot projects and the META-NET network of excellence**. Cooperation with other relevant multi-national forums or activities, e.g., FlaReNet, CLARIN, will ensure consistent approaches, practices and standards aimed at ensuring a wider accessibility of and easier access and reuse of quality language resources;

- **to help build and operate** broad, non-commercial, community-driven, inter-connected **repositories**, exchanges, and facilities that will be used by different categories of target user communities;
- **to mobilize national and regional actors**, public bodies and funding agencies by raising awareness, organizing meetings and other focused events.

Specific targets:

Besides general objectives META-NORD has set several **specific targets**:

- **to provide expertise** to other 6.1 pilots in fields where META-NORD partners have outstanding expertise: **treebanks/syntax databases, terminology resources, wordnets** and **finite-state techniques;**
- **to develop and document methodologies for** building language resources for the so-called under-resourced languages (i.e. languages with limited language resources) as efficiently as possible, with focus on **semi-automatic/machine assisted resource generation**;
- **to facilitate availability of BLARK resources** for META-NORD languages (Danish, Estonian, Finnish, Icelandic, Latvian, Lithuanian, Norwegian, and Swedish);
- to facilitate **knowledge transfer** between CLARIN and META-NORD, especially on standards and intellectual property rights (IPR) issues.

Target users:

The target users are developers and researchers both in industry and academia. This includes private and public institutions, companies and individuals involved in HLT research and development: industrial organizations and SMEs, academic institutions, research organizations, universities, individual researchers and students, national governments, EC institutions, and private investors.

Cooperation:

META-NORD will tightly cooperate and coordinate its activities with other ICT-PSP Objective 6.1. projects and META-NET network of excellence to ensure a coherent pan-European open linguistic infrastructure that will efficiently serve the needs of European industry and research communities for multilingual resources.

Contact information

Vasiļjevs Andrejs
VIENIBAS GATVE 75 A, 1004 RIGA
REPUBLIC OF LATVIA
Tel. +37167605001
Fax. +37167605750
E-mail: andrejs[at]tilde[dot]lv