META NORD

**Baltic and Nordic Branch of the European
Open Linguistic Infrastructure**

Norræn upplýsingasetur í máltækni

- Norræn upplýsingasetur í máltækni
  - Nordic Documentation Centres (2001-2005)
- Norræna rannsóknaráætlunin í máltækni
  - The aim of these centres is to ensure that future and, as far as possible, existing research results, corpora, tools etc. in the field of language technology, are available and re-usable. Research results will be made as widely available as possible at both the national and international level.

NorDokNet

- Samstarfsnet setranna
  - [NorDokNet](#), 2001-(2005)
- Þátttakendur:
  - Háskóli Íslands (Eiríkur Rögnvaldsson)
  - Háskólinn í Kaupmannahöfn (Hanne Fersøe)
  - Háskólinn í Bergen (Koenraad de Smedt)
  - Háskólinn í Gautaborg (Robin Cooper)
  - Háskólinn í Helsinki (Kimmo Koskenniemi)

Íslensk upplýsingasetur

- Íslenskt upplýsingasetur (lexis.hi.is/isldok)
  - Rekið í samvinnu við Orðabók Háskólans, 2001-2004
- Upplýsingasetur um tungutækni (tungutaekni.is)
  - Rekið af verkefnisstjórn um tungutækni, 2000-2004
- Setrin sameinuð í tungutaekni.is
  - Rekið af Tungutæknisetri, 2005-2010
- Nýr vefur – maltaeknisetur.is, iclt.is
  - Rekinn af Máltæknisetri frá 2010

Samvinna við Eystrasaltslönd

- Nordic Graduate School of Language Technology
  - [NGSLT](), 2004-2009
- Norðurlönd og Eystrasaltslönd (og NV-Rússland)
  - Ísland (Eiríkur Rögnvaldsson, HÍ)
  - Danmörk (Peter Juel Henrichsen, CBS, Kaupmannahöfn)
  - Noregur (Torbjørn Svendsen, NTNU, Þrándheimi)
  - Svíþjóð (Robin Cooper, Háskólanum í Gautaborg)
  - Finnland (Kimmo Koskenniemi, Háskólanum í Helsinki)
  - Eistland (Mare Koit, Háskólanum í Tartu)
  - Lettland (Inguna Skadina, Háskólanum í Riga)
  - Litháen (Ruta Marcinkeviciene, Háskólanum í Kaunas)

Technologies for the Multilingual European Information Society – T4ME

- Linguistic diversity is an essential element of our multicultural European society. In the age of globalisation and the emerging information society this diversity is not yet sustainable. Languages that are not supported by advanced information technology will be in danger. The vision of an inclusive barrier-free European information society requires technologies for overcoming language boundaries at affordable costs.

- T4Me is a new Network of Excellence founded by thirteen European research and technology organizations. Its mission is to forge a strategic alliance for creating the technologies and applications needed for making the linguistic diversity of our multicultural society sustainable.

Þátttakendur í T4ME

Participant Acronym	Participant Organisation Name	Country
DFKI (Coordinator)	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH	Germany
BM	Barcelona Media - Centre d'Innovació	Spain
CNR	Consiglio Nazionale Ricerche - Instituto di Linguistica Computazionale "Antonio Zampolli"	Italy
ILSP	Institute for Language and Speech Processing, R.C. "Athena"	Greece
CUNI	Charles University in Prague	Czech Republic
CNRS	Centre National de la Recherche Scientifique - Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur	France
UU	Universiteit Utrecht	The Netherlands
Aalto	Aalto University	Finland
FBK	Fondazione Bruno Kessler	Italy
DCU	Dublin City University	Ireland
RWTH	Rheinisch Westfälische Technische Hochschule Aachen	Germany
JSI	Jožef Stefan Institute	Slovenia
ELDA	Evaluations and Language Resources Distribution Agency	France

META-NET

- META=NET is a Network of Excellence dedicated to fostering the technological foundations of a multilingual European information society. Language Technologies will:

- enable communication and cooperation across languages,

- secure users of any language equal access to information and knowledge,

- build upon and advance functionalities of networked information technology.

Hvað er META-NET?

- A concerted, substantial, continent-wide effort in language technology research and engineering is needed for realising applications that enable automatic translation, multilingual information and knowledge management and content production across all European languages. This effort will also enhance the development of intuitive language-based interfaces to technology ranging from household electronics, machinery and vehicles to computers and robots.

Markmið META-NET

- To this end META-NET is building the Multilingual European Technology Alliance (META). Bringing together researchers, commercial technology providers, private and corporate language technology users, language professionals and other information society stakeholders. META will prepare the necessary ambitious joint effort towards furthering language technologies as a means towards realising the vision of a Europe united as one single digital market and information space.

Aðgerðaáætlun

- **META-NET is supporting these goals by pursuing three lines of actions:**

1. fostering a dynamic and influential community around a shared vision and strategic research agenda (META-VISION),

2. creating an open distributed facility for the sharing and exchange of resources (META-SHARE),

3. building bridges to relevant neighbouring technology fields.

Styrkir ESB

- ICT Policy Support Programme (ICT PSP)
  - The **ICT Policy Support Programme** (or ICT PSP) aims at stimulating innovation and competitiveness through the wider uptake and best use of ICT by citizens, governments and businesses.

- Styrkir eru í boði, m.a. undir yfirskriftinni
  - **Multilingual Web**
    - Open Linguistic Infrastructure

Multilingual Web

- This theme contributes to the creation of a pan-European digital *single market* by stimulating ICT-based cross-lingual communication, collaboration and participation. Successful deployment of cross-border platforms and services depends on how effectively language barriers are overcome.

- Existing systems tend to address a small number of languages and not all functionalities are available for all language communities. There is a need for a wider range of language-aware online services, and for lowering the associated development and operation costs.

Markmið

- The overall objective is to ease and speed up the provision of online services centred around computer-based translation and, cross-lingual information access and delivery. The focus is on assembling, linking across languages, and making widely available the basic language resources (models, tools and datasets) used by developers, professionals and researchers to build specific products and applications. Emphasis is placed on less-resourced languages, including those of the twelve member states that have joined the EU in recent years.

META-NORD

- Evrópusambandið vildi víkka META-NET út
  - þannig að það næði til allra tungumála EES
- Því voru stofnaðir þrír samstarfshópar
  - Í Norður-, Austur- og Suður-Evrópu
- Evrópusambandið fól Tilde í Riga
  - fyrirtæki með mikla reynslu í máltækniverkefnum
  - að mynda samstarfshóp Norður- og Eystrasaltslanda
- Úr því varð verkefnið META NORD

Consortium

Partner	Country
Tilde SIA	Latvia
Københavns Universitet	Denmark
Tartu Ulikool	Estonia
Universitetet i Bergen	Norway
Helsingin yliopisto	Finland
Háskóli Íslands (Máltæknisetur)	Iceland
Institute of Lithuanian Language	Lithuania
University of Gothenburg	Sweden

Project Goal

- The META-NORD project aims to **establish an open linguistic infrastructure** in the Baltic and Nordic countries (Denmark, Estonia, Finland, Iceland, Latvia, Lithuania, Norway, Sweden)

- The project will focus on European languages with less than 10 million speakers – EU official languages **Danish, Finnish, Swedish**, the languages of the recently accessed EU countries – **Estonian, Latvian and Lithuanian**, and the languages of the European Economic Area – **Icelandic and Norwegian**.

Main objectives, 1

- **to provide a description of the national landscape** in terms of language use; language-savy 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 **collecting resources in the Baltic and Nordic countries** and by documenting, linking and upgrading them to agreed standards and guidelines;

Main objectives, 2

- **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

- **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.

Horizontal Actions - Core Set

- **WordNets**: monolingual WordNets and cross-linked pilots are planned as outcomes of META-NORD project
- **Treebanks**: treebanks available in the consortium will be integrated on a uniform platform and linked across languages using parallel multilingual treebanking based
- **Terminology collections:** distributed terminology resources across languages and domains will be consolidated
- No resource type covered by all languages of META-NORD

Authorization and Access Interface

- Web service for granting digital rights to end users

- Provides a solution for granting access to
  - publicly available resources
  - resources with more limited availability

- Controls the access to the resources according to user category

Key results, 1

- **national scene charts** describing language community and role of language in the respective country, research community, language service and language technology industry, use of language technology by business and administration, legal provisions;

- **language resources charts** of actually or potentially available resources to the META-NORD consortium;

- **treebanks** for relevant languages accessible through a uniform web interface and state-of-the-art search tool and linked across languages using a parallel multilingual treebanking;

Key results, 2

- **wordnets** upgraded to agreed standards and used for creation of pilot multi-lingual lexicons for IR purposes using cross-language synset linking and validate these pilot cross-lingual resources;

- **monolingual and bilingual terminology collections** integrated into multilingual terminology bank with elaborated terminology data access and sharing mechanisms;

- **language resources batches** upgraded to agreed standards, extended and linked across different sources, aligned across languages and populated into digital exchange platform for pilot operation;

- an **Authorization and Access Interface (AAI)** web service for granting digital rights to end users.

WP1 - Management

- **Objectives**
  - Project coordination and management, administrative and financial management
  - Reporting and submission of deliverables, quality control and risk management
  - Internal and external communication, dissemination

WP2 - Analysis and selection of language resources
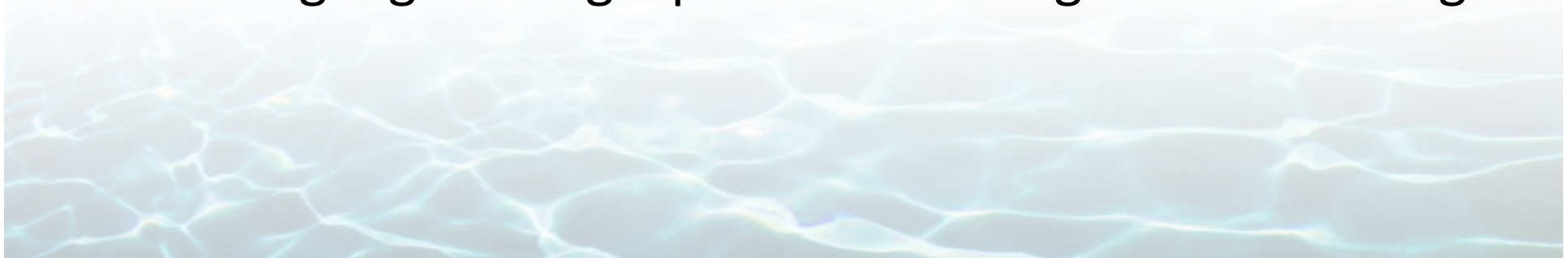
- **Objectives**
  - Charting of the national scene resp. the language community landscape.
  - Identify, select, acquire and/or obtain permission to rework and 'publish' own and third parties' data collections and software tools.

WP3 - Enhancing language resources

- **Objectives**
  - Deliver clean, documented and (re)usable resources through the open digital exchange provided by META-NET and where relevant other suitable channels and initiatives.

  - Make treebanks for relevant languages accessible through a uniform web interface and state-of-the-art search tool and to link treebanks across languages using a parallel multilingual treebanking.

WP3, frh.

- **Objectives**, frh.
  - Ensure that wordnets agree to modern standards regarding formats, derive/create pilot multi-lingual lexicons for IR purposes using cross-language synset linking and validate these pilot cross-lingual resources.
  - Link monolingual and bilingual terminology collections and integrate them into multilingual terminology bank with elaborated terminology data access and sharing mechanisms.

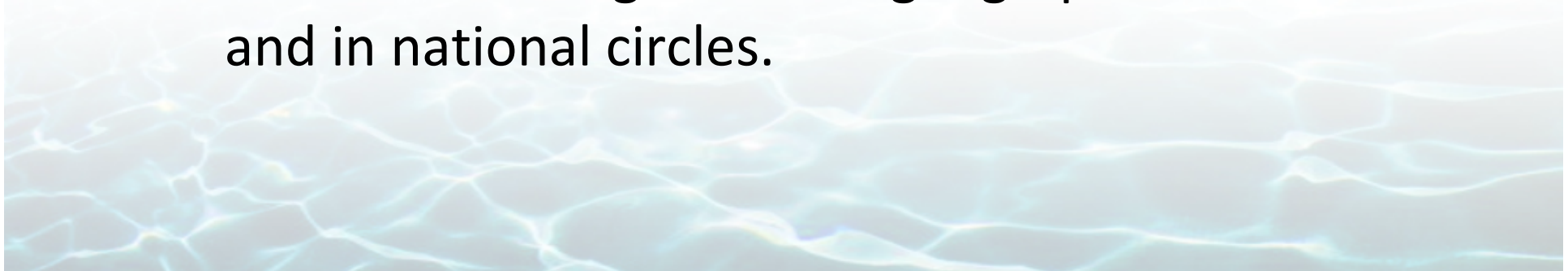# WP4 - Cross-national collaboration and Pilot service

- **Objectives**
  - Enhance the availability and suitability of language resources.
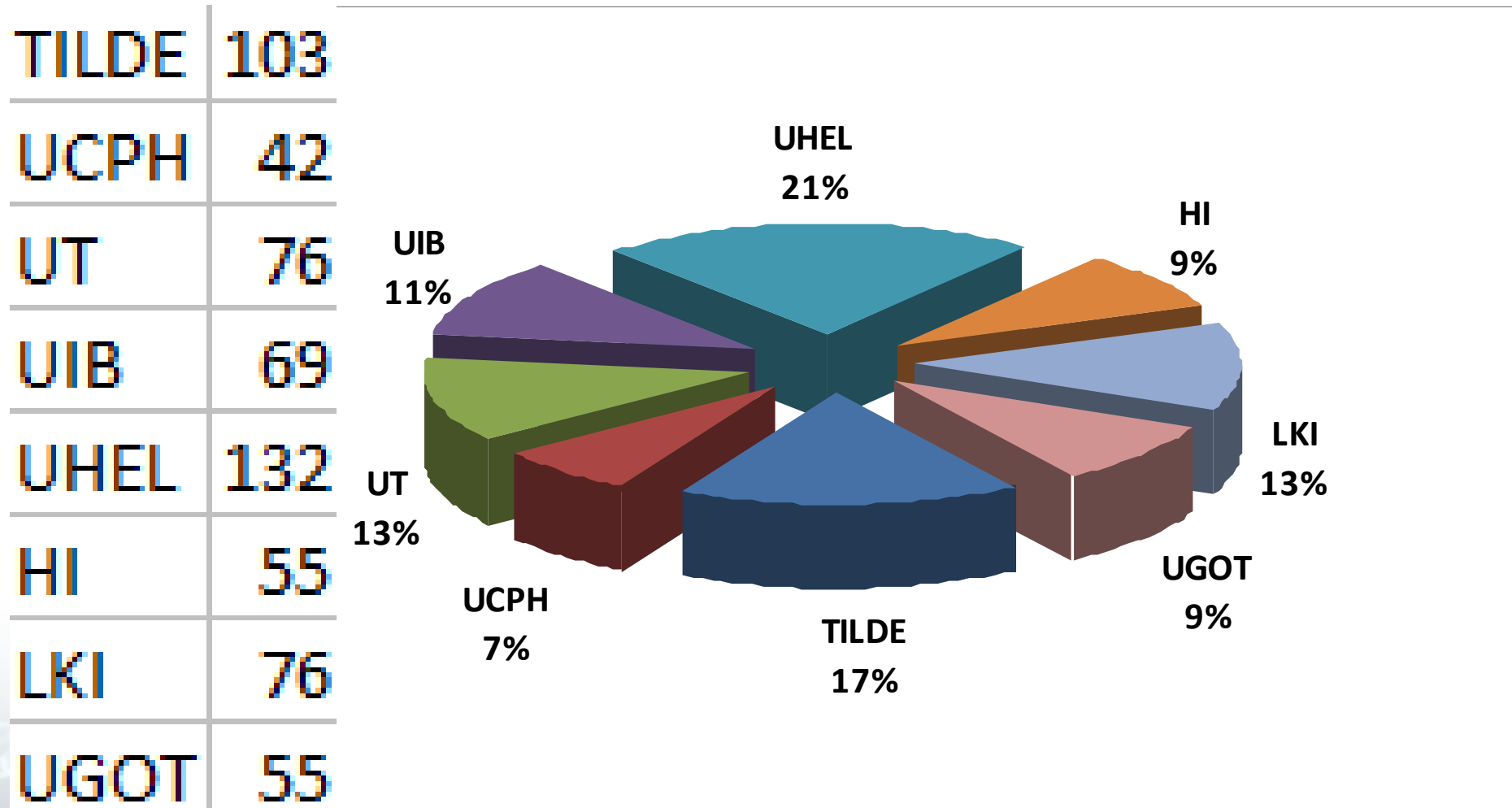 - Take an active part in the launch of the digital resource exchange.

WP5 - Outreach, awareness and sustainability

- **Objectives**
  - Make the project sustainable beyond the end of the EU-funded phase. Ensure the continuation and coordination of national efforts after the project's end.
  - To this end, promote language research, technology, resources and applications within META-NORD linguistic and geographic boundaries and in national circles.

Mannmánuðir á þátttakanda

TILDE	103
UCPH	42
UT	76
UIB	69
UHEL	132
HI	55
LKI	76
UGOT	55

Fjárhagsáætlun

Participant no.	Participant name (EUR)	Personnel costs (EUR)	Sub-contracting	Other specific direct costs	Indirect costs (EUR)	TOTAL COSTS (EUR)	Requested reimbursement	Requested EC contribution (EUR)
1	TILDE	494.400	7.000	26.070	148.320	675.790	0,5	337.895
2	UCPH	288.120	0	22.000	86.436	396.556	0,5	198.278
3	UT	266.000	0	22.000	79.800	367.800	0,5	183.900
4	UIB	488.382	2.000	28.000	146.515	664.897	0,5	332.448
5	UHEL	726.000	1.000	22.000	217.800	966.800	0,5	483.400
6	HI	293.810	0	22.000	88.143	403.953	0,5	201.977
7	LKI	380.000	0	22.000	114.000	516.000	0,5	258.000
8	UGOT	374.000	0	22.000	112.200	508.200	0,5	254.100
	TOTAL	3.310.712	10.000	186.070	993.214	4.499.996		2.249.998

Að lokum

- Umsóknin er að upphæð 350 milljónir króna
  - þar af 31,5 milljónir til Íslands
- Styrkveiting er ekki endanlega frágengin
  - en verið er að vinna að lokafrágangi samnings
  - sem verður væntanlega undirritaður á næstu vikum
- Verkefnið á svo að hefjast 1. febrúar
  - og standa í tvö ár
- Við þurfum á liðsinni ykkar að halda!