EASTIN-CL

**Crosslingual and multimodal search
in a Portal for Support of Assisted Living
Project no. 250432**

EASTIN-CL Annual Report 2010

Deliverable D8.11

Version: December 2010

Date: 2010-12-01

EASTIN-CL
1ˢᵗ Annual report 2010

Project Ref. no.:   250432

Acronym: **EASTIN-CL**

Reporting period: **March 2010 to November 2010**

Project coordinator:

Name:	Gregor Thurmair
Organisation:	Linguatec GmbH
Phone	+49 89 896664 0
Fax:	+49 89 88 91 99 33
Email:	g.thurmair@linguatec.de

Content

1. Introduction and Project Objective

Relevance

Participation and full citizenship, also called **e-Inclusion**, is one of the prominent goals of the European community. This includes the full participation of people with disabilities to social life.

In a 2003 study, the European Commission states that „assistive technologies play an increasingly central role in equalising opportunities for people with disabilities in all aspects of life". This statement must be seen in the context that the number of people with disabilities, in an ageing society, will increase from now 11% to 17% in 2020, so the problem will aggravate in the future.

However, access to information about assistive products and technology is still restricted, with language being a significant barrier in the information flow.

Being able to understand foreign language documents, and being able to freely select the communication channel to access such information, as offered by EASTIN-CL, will foster the principles of e-Inclusion.

Objective

The project will support the social participation of disabled and elderly people, by providing crosslingual and multimodal support for accessing information bases on assistive tools and technology.

Recent efforts have linked national assistive technology information bases into a European portal called EASTIN (www.eastin.eu). This portal will be enhanced and made more accessible using language technology:

- **Multilingual** technology will allow users to search the data in their native language;
- **Multimodal** technology will allow them to access the portal not just in written but also in spoken communication.

Target users currently are mainly professionals, like doctors, physiotherapists, insurances, or manufacturers. EASTIN-CL will help to open the information bases on assistive technologies also for end users (e.g. ageing people).

Main tasks

EASTIN-CL will support access to information bases on assistive domain along three action lines:

- providing easy and **intuitive access** also for end users, by supporting query formulation;
- offering **language-transparent** access, by allowing users to access foreign-language documents in their native language, and have the search results retranslated into their native language;

- adding a **speech communication** channel, by accepting spoken queries, and by converting search results into spoken language, making communication more barrier-free.
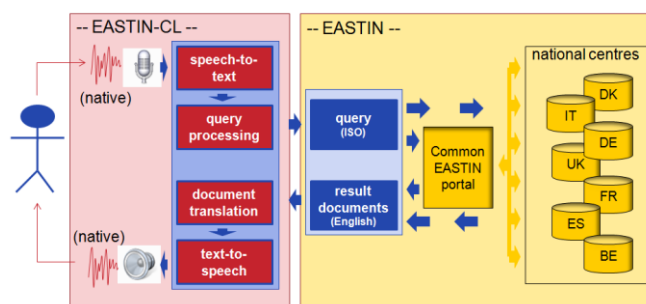
The languages supported by at least one of the EASTIN-CL tools are: Danish, English, Estonian, German, Italian, Latvian, and Lithuanian. Other languages can be added as well.

Expected Results:

The result of the project will be a dedicated language server inside the EASTIN portal. Its task is to act as an interpreter, transforming voice into text and text into voice, and translating queries and documents form and into the users' native language. The language technology components will lead to an improved access to assistive technology in Europe for all citizens.

A by-product will be a multilingual terminological glossary of the assistive domain, covering all relevant search terms and their translations. This glossary will be made available for online access.

Partners:

The EASTIN-CL consortium consists of two profiles of project partners:

On the technology development site, partners are:

- **Linguatec GmbH** (Germany), providing speech technology and translation for German, English, and Italian
- **Tilde SIA** (Latvia), providing speech and translation support for the Baltic languages

On the side of domain expertise, partners are:

- **Institut der deutschen Wirtschaft Köln / Rehadat** (Germany), provider of the largest information base on assistive technology in the German speaking countries
- **Fondazione Don Carlo Gnocchi Onlus / SIVA** (Italy), provider of the largest Italian collection, and host of the EASTIN portal
- **Hjælpemiddelinstituttet** (Denmark), responsible for the Danish centre of assistive technology

2. Project Progress

2.1 Scientific and technical progress

The first project phase focused on the specification of the relevant components, resources, and integration tasks. It resulted in a system specification document, delivered at M6 (August 2010).

2.1.1 Requirements analysis

The first task was to collect the requirements for the different system components and languages, and to define which support is required for which language, and which portal (EASTIN, national portals) would be the target of an integration effort.

On the language and resource side, the following distribution of languages was required:

Tool	languages	Portals
Terminology for translation	da, de, en, et, it, lt, lv	EASTIN, SIVA, Rehadat
Query Processing (end user)	da, de, en, et, it, lt, lv	EASTIN, SIVA (it), Rehadat (de)
Document Translation	en>de, et, it, lt, lv	EASTIN
Text-to-Speech	da, de, en, it, lv	EASTIN, HMI(da), SIVA (it), Rehadat (de)

On the system side, the main requirement was that the existing EASTIN portal should not be affected negatively by the new components, and a seamless integration of new and existing functionality should be achieved.

2.1.2 Resources

In order to be able to create a multilingual terminological resource of the domain, partners collected existing material, in all EASTIN languages. It consisted of terminological resources, mainly centred around the ISO 9999 classification, but extended by key terms, additional index terms etc. These term lists were updated, completed, and structured to be used in EASTIN-CL.

Also textual resources were collected, however there was not too much parallel data (if so, then they were machine-translated). Mainly English and German texts were collected, and the key terms were extracted from them.

At the project meeting in Belluno (cf. appendix) the material was discussed, and the project decided to base the multilingual terminology of the assistive domain on existing index and key terms, instead of additional general purpose terms as found in the texts.

So these terms were collected, normalised and cleaned, resulting in an English Master list of about 13.000 terms of the Assistive Domain, with links to the ISO codes which in turn represent the relevant documents for the given term.

An example is given in fig. 1-1

4613	four-wheeled	four-wheeled	122721;121609;122718;243615	Rehadat
3004	cycles (four-wheeled)	four-wheeled cycle	121815	ISO;HMI
4614	four-wheeled cycles	four-wheeled cycle	121815	ISO;HMI
4615	four-wheeled electric stacker	four-wheeled electric stacker	243909	Rehadat
8200	moped, four-wheeled	four-wheeled moped	121609	Rehadat
8203	mopeds (four-wheeled)	four-wheeled moped	121609	ISO;HMI
8242	motorcycles (four-wheeled)	four-wheeled motorcycle	121609	ISO;HMI
4616	four-wheeled scooter	four-wheeled scooter	121812	Rehadat
4617	four-wheeled stacker	four-wheeled stacker	243909	Rehadat
4618	four-wheeled tandem	four-wheeled tandem	121815	Rehadat
4619	fractional arithmetics	fractional arithmetics	51815	Rehadat

Fig. 1-1: Example of the term master list

This list is subject to completion and improvement in the WP 2 and the deliverable D2.1 (due Feb. 2011).

A major challenge for the creation of the multilingual resource consists in the relation of the index terms to the ISO codes they point to, and the identification of the right level of granularity for a given term in its relation to the assistive products structured according to the ISO hierarchy.

This problem, and other translation-related issues, were subject of a special workshop on terminology, held in Cologne beginning of November 2010.

2.1.3 Software integration

Integration

In the system software field, it was necessary to have a better understanding of the current structure of the EASTIN portal, and how the EASTIN-CL components could be integrated.

A workshop in Milano served the purpose to discuss the current software, the change requests, and a possible integration of the new components. For the interfaces, it was decided to use web services as the basic communication means to access the components hosted by the technology partners. The specification document provides the API descriptions for the Machine Translation and the Text-to-Speech services. The overall schema of integration is shown in fig. 1-2:
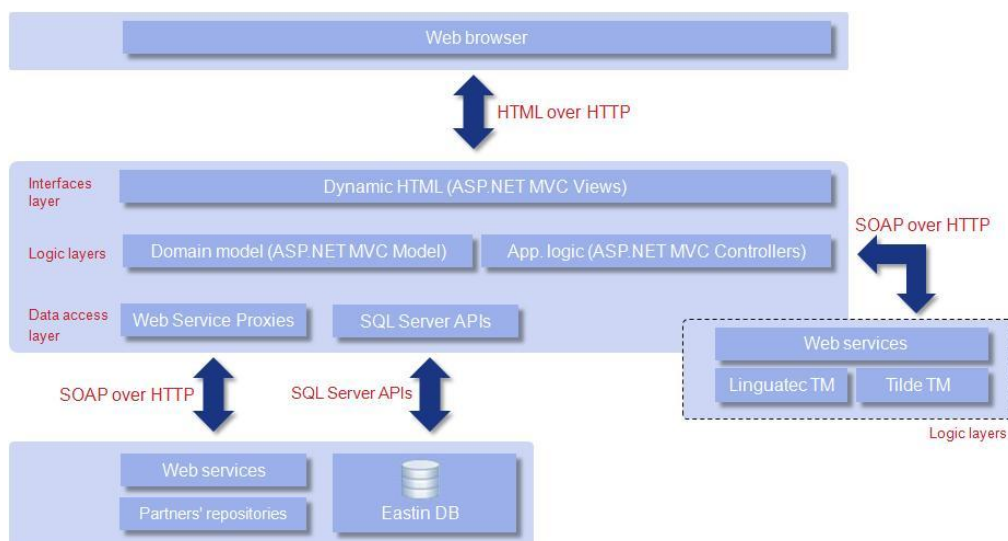

Fig. 1-2: EASTIN-CL architecture

The original EASTIN software must be adapted in several areas; these areas were identified (e.g. upgrade to new component versions, better structuring of the business layers of the application, etc.); parts of them are subject to a subcontract with Don Gnocchi Sistemi, starting in September.

User Interface

To make the portal more user-friendly, it was decided to offer a free-text field for users not so familiar with expert knowledge of the assistive domain. The best way of integrating a free text search was discussed; the current proposal is shown in fig. 1-3; it extends the existing EASTIN interface.

Fig. 1-3: EASTIN-CL User interface for search

Single components

While there is a complete implementation of some of the EASTIN-CL backend components, like Machine Translation or Text-to-Speech, this is not the case for other components.

Query Translation was given special attention, in order to cover the broadest possible variance in search vocabulary usage by end users, in particular for multiword units. Fig. 1-4 gives an impression on the architecture of the component.
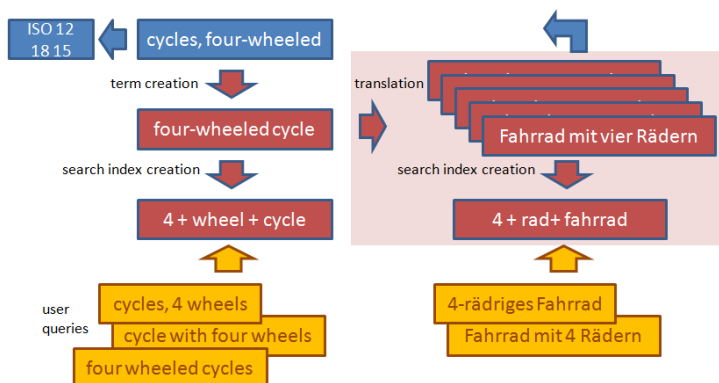
Fig. 1-4: Query Processing

It should be noted that the search in the EASTIN portal is not a free-text search but based on ISO codes; i.e. terms are linked to ISO codes which in turn point to documents. Moreover, in EASTIN there is the agreement to have an English version for each document (mainly machine-translated). As a result, there is no need for crosslingual search, instead a multilingual mapping of the search terms to the ISO codes is required, and a re-translation of the found documents from English into the query language.

Speech Recognition is only available in German and in the Rehadat portal; the specifications describe which adaptations will be required to support spoken input and output.

2.1.4 System Specifications (WP 1) – Milestone M1

The different aspects of work were collected into the specification document (Deliverable D2.1). In addition to some fine-tuning and corrections, the specifications also contain an updated task list and work plan. This work plan can serve as an update of the deliverable D8-4 (work plan).

Project milestone 1, defined by the availability of the system specifications (D1.1), has been reached in time, and the WP 1 can be closed.

2.1.5 Technical Work packages (WP 2 – WP 5)

The technical work packages (WP 2 terminology, WP 3 query processing, WP 4 document re-translation, WP 5 Audio channel) started in M4 (June 2010) as planned, and essentially delivered the input of their respective components to the specification deliverable. No special deliverable was foreseen for them in this phase.

Two internal memos were circulated:

- WP 2 for the preparation of the terminology master list.
- WP 3 for the design of the query translation.

They were used to collect feedback for the respective section is the specification document.

2.2 Communication, dissemination and awareness activities

Work in this field is bundled in WP 7 (Dissemination and Sustainability), which started at M1 (March 2010).

2.2.1 Dissemination Plan

The dissemination plan (D7.1) was set up for month M3 (May 2010).

The main goals of the dissemination activities within the EASTIN-CL project are:

- to inform end users (disabled and elderly people), professionals and their organisations about the existence of the easy to use search possibilities of the EASTIN-CL project: especially non-professionals in the field of assistive technology, people in the Baltic countries and non-English-speaking users;
- to promote understanding of European co-operation in the area of assistive technology and in the area of speech technology;
- to present a showcase for good ontology-based crosslingual and multimodal search in a relevant social domain.

The main target groups for the dissemination activities are end-users, professionals, their organisations at European and national level and the language technology community including language engineering experts and users of language technology products.

EASTIN CL will be presented in relevant trade fairs, conferences and exhibitions, the main target event being the Rehacare fair (largest fair on assistive technology worldwide; Düsseldorf, Oct. 2011) where EASTIN-CL is planned to be presented at a special workshop). "Traditional" information materials will be produced such as brochures and posters including a purposely designed logo. The

EASTIN-CL Web Site plays a key role in dissemination, as up-to-date information source about the project progress and the new services it will generate.

The realisation of the specific dissemination activities will be adjusted to the local situation of each partner's country. However, such activities will be monitored centrally and recorded in the course of the project.

2.2.2 Dissemination Material

The following dissemination material has been prepared

- **The EASTIN logo**

Fig.2-1: EASTIN-CL logo

- **The EASTIN-CL web site** (www.eastin-cl.eu)

It contains all relevant information, the project related dissemination material, the fact sheet, the presentation etc. It is localised into all EASTIN-CL languages (de, de, en, et, it, lt, lv).

The web site is the deliverable D8-2.
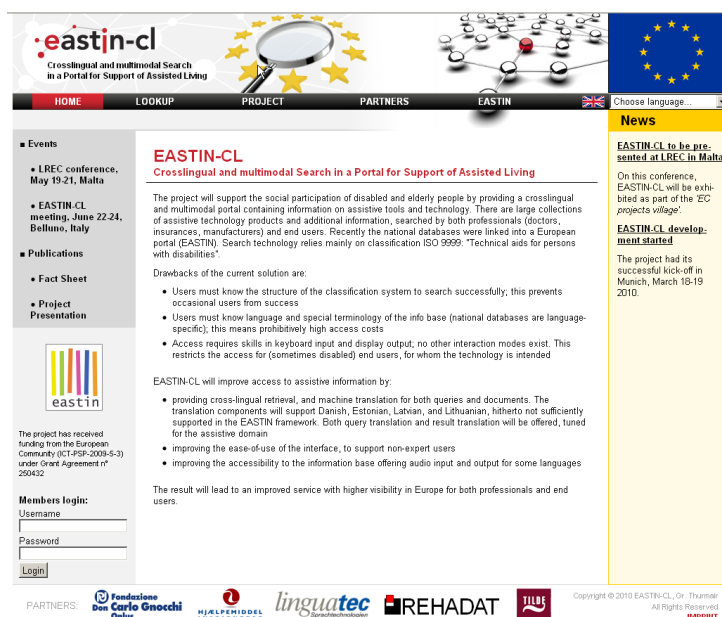
Fig. 2-2: EASTIN-CL Web Site

- **Project Presentation**
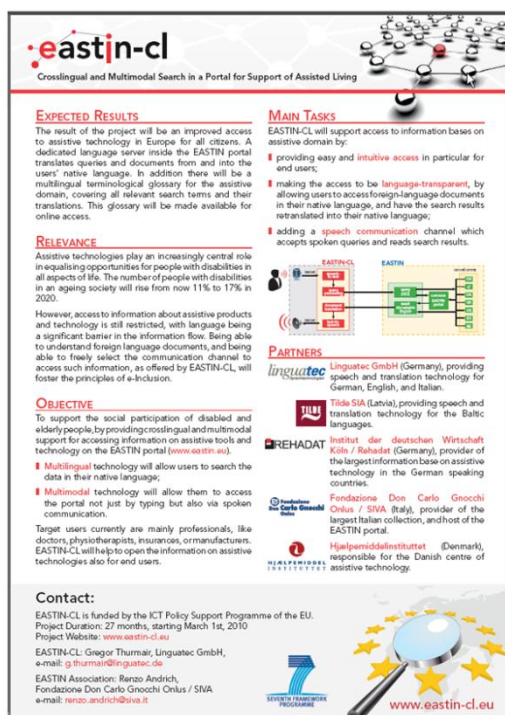
EASTIN-CL produced a project presentation (cf. deliverable D8.1). It has been localised into some of the other EASTIN-CL languages in the meantime.

- **Project Fact Sheet**

A Fact Sheet was also produced (deliverable D8.3) and distributed e.g. at the LREC conference. Again, it has been localised into several EASTIN-CL languages in the meantime.

- **Poster**

A special poster was made for EASTIN-CL in the context of the LREC 2010 in Malta where EASTIN-CL was represented with its own booth.

Fig. 2-3: EASTIN-CL poster for LREC

- **Brochure**

For this event, a brochure was designed and distributed, as a 6 page leporello.

All the dissemination material is documented in the dissemination plan (deliverable D7.1).

2.2.3 Dissemination Activities

The following dissemination activities were carried out in the reporting period (it should be mentioned that it is just the first period, and there are no project results yet which could be presented)

Activities in the reporting period

a. Announcement of EASTIN-CL in the Rehadat Newsletter

There was an announcement of EASTIN-CL in the March 2010 issue of the Rehadat Newsletter, adapted to the German audience.

b. LREC Conference, May, Malta

EASTIN-CL was represented with a booth at this conference, where a 'EC project village' was initiated by the organisers. It was a three-day conference.

Fig. 2-4: EASTIN-CL booth at LREC

c. Baltic countries, May 2010

With the help of Tilde, a tour of Mr. Andrich (EASTIN-CL member and EASTIN president) through the Baltic countries was organised, in order to coordinate activities in the area of assistive support, and broaden the basis of EASTIN in the Baltic countries.

The outcome was very encouraging:

- EASTIN national contacts have been officially established in Lithuania, Latvia, and Estonia
- Key actors (users, professionals, policy makers) are supportive and will collaborate to disseminate awareness at national level
- Key persons have been identified that can help validate terminology in national languages for the EASTIN-CL project

Fig. 2-5: Presentation of EASTIN-CL in the Baltics

Meanwhile, the result of this effort was that the **first national contact point for EASTIN** in the Baltic countries could be established, with the Ministry of Health in Latvia. A copy of the letter can be found in Annex 1.

d. Announcement of EASTIN-CL on the EASTIN website

A short presentation of the EASTIN-CL project in (English, Italian, German and Danish) has been published in the News section of the EASTIN website.

e. BICC-Net, July 2010, Munich

BICC-NET, the Bavarian Information and Communication Technology Cluster, organised an event for representatives of the Bavarian IT industry, focusing on 'Cloud Computing'. EASTIN-CL was presented as an example of 'Multilingual eServices' by Mr. Busch (Linguatec).

f. EASTIN Assoc., June 2010, Belluno

EASTIN-CL was presented to the members of the EASTIN association (not all of them are partners in EASTIN-CL) at the EASTIN meeting in Belluno, with very positive feedback. Issues of common interest are testing, the use of some components (like speech technology) in the national portals, and integration options.

g. AAATE Seminar, October 4-5, Sheffield

A presentation of the project at the AAATE seminar in Sheffield (4-5 October 2010) was given by R. Andrich.

h. Rehacare, October 4-7, Düsseldorf

Rehacare is the biggest fair for assistive technology worldwide. EASTIN-CL was presented by IWKöln. A special flyer was created for this event, as well as an updated poster. The project itself was also presented there.

Information on EASTIN-CL is also given in an article in "Die Rehabilitation".

2.3 User and market related aspects

Due to dissemination activities just mentioned, the awareness of assistive technology has grown, most notably in the Baltic countries.

4. Future Work

The next steps in EASTIN-CL will consist in the implementation of the main technical components, and the preparation of the testing phase.

- The **terminology resource** (WP 2) will be produced early 2011, and will be made available for the public, among others by an online lookup facility on the EASTIN-CL web site. It is the basis for the speech related and translation activities.

- The **query processing** will be specified (by Q1 / 2011) and implemented.

- The **speech components** (speech recogniser for German / IW Köln, and text to speech for all portals will be provided and integrated; both components' vocabularies will be checked against the EASTIN-CL terminology resource.

- The **machine translation** components will be adapted towards the EASTIN-CL terminology, and integrated into the EASTIN portal

- The **portal** itself will be upgraded and adapted to the AESTIN-CL requirements, based on a cooperation with Don Gnocchi Sistemi (DGS), which is the host of this portal.
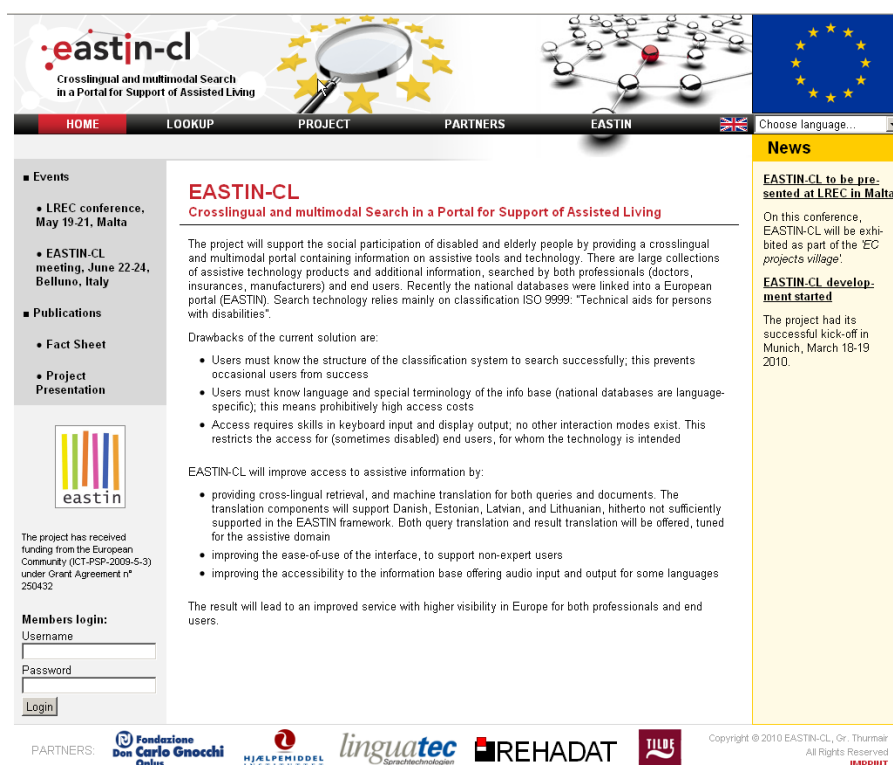
There is a target for the development: EASTIN-CL plans to give a presentation and a specific workshop at the next **Rehacare in Düsseldorf in September 2011**, This will be a big event, with invited speakers and guests. By then, a demonstrator version will be available.

After Rehacare, the focus of the project will be on the preparations for the system test phase.

5. Contact

Web Site

There is the EASTIN-CL website (www.eastin-cl.eu) providing public information on the project; it will offer a term lookup facility as soon as the terminology is available.

Context and Contact

EASTIN-CL is part of the Competitiveness and Innovation Framework Programme of the European Union, in the ICT Policy Support Programme (ICT-PSP),
Theme 5: Multilingual Web Content Management

Project Number: CIP-ICT-PSP-2009-5-3-250432.

Project Duration: 27 months, starting March 1st, 2010

Project Website: www.eastin-cl.eu

Contact: EASTIN-CL: Gr. Thurmair, Linguatec GmbH,
                email: g.thurmair@linguatec.de

                EASTIN Association: R. Andrich, Fondazione Don Carlo Gnocchi Onlus / SIVA
                email: renzo.andrich@siva.it

Annex 1: New National contact for EASTIN in Latvia

LATVIJAS REPUBLIKAS
LABKLĀJĪBAS MINISTRIJA

MINISTRY OF WELFARE
OF THE REPUBLIC OF LATVIA

Skolas icla 28, Riga, LV-1331
Tālr. 67021600
Fakss 67276445
E-pasts: lm@lm.gov.lv

28 Skolas str., Riga, LV-1331, Latvia
Phone 371 67021600
Fax 371 67276445
E-mail: lm@lm.gov.lv

05 -08- 2010   Nr. *16.7 - aaja100*

EASTIN Association
Piazzale Morandi, 6
20121 Milano Italy

For the attention of
Mr. Renzo Andrich, president
renzo.andrich@siva.it

Subject: national contact for Latvia

Following to our preliminary talks, we are pleased to confirm our willingness to cooperate with EASTIN.
In particular, we confirm we are prepared:
- to answer questions posted by Latvian citizens or companies through the EASTIN website;
- to provide or supervise the EASTIN assistive technology classification, keywords and terminology in Latvian language;
- to regularly review – and update when needed – the EASTIN pages in Latvian language;
- to take part in on-line discussions and exchange of expertise within the EASTIN network on issues of common interest.

The person responsible for these tasks in Latvia is Mr Aivars Vētra – member of the board of the National center of rehabilitation "Vaivari" (email: aivars_vetra@inbox.lv, phone: +371 67766128). Therefore we kindly ask:
- to redirect any message sent through the EASTIN website in Latvian language to his email address;
- to include his email address in the EASTIN mailing list;
- to send him the password needed to edit the Latvian web pages of the EASTIN website;

- to acknowledge credit to the National center of rehabilitation "Vaivari" as EASTIN national contact in the EASTIN website.

Yours sincerely,
State Secretary deputy   I.Alliks

Annex 2: New National contact for EASTIN in Lithuania

TECHNINĖS PAGALBOS NEĮGALIESIEMS CENTRAS

PRIE SOCIALINĖS APSAUGOS IR DARBO MINISTERIJOS

CENTER OF TECHNICAL AID FOR DISABLED PEOPLE

UNDER MINISTRY OF SOCIAL SECURITY AND LABOUR

14 July, 2010 Nr. (3.6) R3- 346

EASTIN Association **Piazzale Morandi, 6** **20121 Milano Italy**	Mindaugo str. 42A 03209 Vilnius Lithuania

For the attention of
Mr. Renzo Andrich, president
renzo.andrich@siva.it

Following to our preliminary talks, we are pleased to confirm our willingness to act as EASTIN national contact for Lithuania.

In particular, we confirm we are prepared to answer questions posted by Lithuanian citizens or companies through the EASTIN website, to consult about the EASTIN assistive technology classification, keywords and terminology in Lithuanian language.

The person responsible for these tasks within our organization is Ms Dovilė Sabaliauskaitė, head of methodical department, tel. 370 85 272 11 06.

Therefore we kindly ask to redirect any message sent through the EASTIN website in Lithuanian language to her email d.sabaliauskaite@tpnc.lt and to include her email address in the EASTIN mailing list. Also we ask to send her the password needed to edit the Lithuanian web pages of the EASTIN website.

Best regards

Ms. Elena Nekrošaitė
acting deputy director
elena.nekrosaite@tpnc.lt

Dovilė Sabaliauskaitė, d.sabaliauskaite@tpnc.lt
(the original document not be send)

Kodas 190789945   Mindaugo g. 42 A,   Tel. (8~5) 2 72 11 06, tel. (8~5) 2 79 01 89   El. paštas: centras@tpnc.lt
LT-01311, Vilnius   tel./faks. (8~5) 2 73 47 96