PUBLISHABLE SUMMARY

Introduction

The present project represents the collaborative effort of the major machine translation (MT) companies of Europe to establish a common translation website and provide free quality MT from any European languages to any other. All relevant stakeholders of European MT industry with translation services decided to cooperate in this project.

Project objectives

The core functionality of the website is quality MT of user input, where translation is available from any of the covered languages to any other.

Table 1: Covered languages

Official EU languages	Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish
Minority or regional languages in the EU	Albanian, Arabic, Catalan, Croatian, Galician, Macedonian, Russian, Serbian, Turkish, Ukrainian, Welsh, Yiddish
Other languages	Afrikaans, Belarusian, Chinese, Dari, Haitian Creole, Hebrew, Hindi, Icelandic, Indonesian, Japanese, Kazakh, Korean, Malay, Norwegian, Pashto, Persian, Swahili, Tagalog, Tajik, Thai, Urdu, Vietnamese

Additional user-friendly features were added by the partners in order to maximise the usability of the site in a variety of usage scenarios: spell checking, virtual keyboard, automatic language detection, domain selection, translated search, interface for feedback and user contribution. The site also offers some special features, such as offering more than one translators for the most used language pairs. In the first year the Project Coordinator provided an evaluation campaign involving machine and human evaluation, which results will be used in ordering the translated texts and translation systems (the order of translated texts per language pairs can be modified and fixed by the User in settings menu).

The website will also include links to human translation services which users are encouraged to consult if the machine translation is proven insufficient for their purposes. Importantly, the site accepts user suggestions, which are public and open to other users to make comments. This service will be done in the beginning of the second year.

The new translation platform is contributing to the development of a single European information space and addressing the language barriers that currently limit the use of the web by offering MT at a single website between any European languages. The website is supporting more than fifty languages in an equal way.

The participation in the consortium is open to providers of any new language pairs subject to quality criteria. During the first year of the project two other MT companies have joined the project – Lingenio (Germany) and GrammarSoft (Denmark).

The website is able to support cross-lingual access to content in any internet page in general and also web resources fostering language diversity and locality in content production.

The solution offers the best available quality machine translation solution for every European language. The integrated service has developed a common API – a set of language translation conventions – that is acceptable for the current as well as for the future partners. The project is utilizing existing mature technologies and methods by integrating the solutions of the leading providers of the particular language pair. The website offers cross-lingual search facility to make available Internet search in other languages.

The present consortium is open to new partners, especially providing services in less-covered languages. The focus of the present project is the integration of proven technologies. The consortium partners currently use rule based methods but statistical translation will also be used, especially for minor languages. High preference is given for adoption of new language pairs by local providers.

The platform will provide user driven community based evaluation that will govern the selection of the default translation system if more than one translator is available for a certain language pair. After launching the webpage we will be able to collect user feedbacks and these results will be used for ranking until that the webpage is using randomised ordering of translators.

The website is also sharing resources that users suggest as better ones or alternative translation for words or sentences. The site provides also place for user activity.

The specific objectives of the project are:
● to provide the best MT for all the languages at a common website;
● to secure the sustainability of the service by generating sufficient revenue;
● to advertise extra translation services offered by the partners;
● to share the task of covering many languages with other MT providers;
● to collaborate with other MT companies as this co-operation might foster the development of new technologies and industry standards;
● to harvest the resources produced by community contribution;
● to secure sufficient visibility and prominence on the web to counter the overwhelming dominance of other translation systems in spite of their lower quality.

The work performed in the first project year

WP1 (Service Integration)
Consortium members discussed and agreed on the specification of a common API for translation services that supports all features of individual services. This new interface includes specifications that address the specific problems of linked translations, in order to prevent the cumulative degeneration of translations due to the interlingua approach.

A central server application was designed and developed, which communicates with the service providers and the itranslate4.eu website. This application forwards translation queries to the appropriate partner servers, performs linking as required, and sends the translated text back to the user interface.

WP2 (Tools & Extensions)

The itranslate4.eu website offers extra services apart from text translation in order to make the site more usable, more accessible and more attractive to users with a variety of needs and expectations. A working prototype of these extra services was developed as follows: the spell checking service; the virtual keyboard facility; the language detection engine; the translated search module; the system managing online ordering of human translations; the user account management system.

WP3 (User Interface)

A user friendly GUI with a stylish design was created. The interface is ergonomic, modern and reflects the professionalism of the services behind. The GUI was localized to 13 languages. The web page indicates that the itranslate4.eu site has been made with the assistance of the EU.

WP4 (Evaluation)

Evaluation forms an integral part of the implemented solution. Consortium members discussed possible methods of evaluation and specified a standard procedure and set of tools. Three different resources were targeted: automatic evaluation, human evaluation and human feedback. A software test suite was developed to run and evaluate the translators on the requested language pairs. Human evaluation campaign based on crowd sourcing has been started which will end in the second year of the project. Collection of user feedback has started on the portal by the user evaluation tool.

WP5 (Operation)

In the first year central servers and individual partners' translation servers were installed. Analytical tool for the central server were introduced that monitors visitor statistics and geographic information. Another this tool monitors the operation of the main service, handle emergency situations, alert maintenance personnel in the case of abnormal operation. User feedback system was developed and available in Contact Us.

WP6 (Dissemination & Exploitation)

For good visibility by search engines the web page was optimised and cross referenced by positioning keywords from different domains.

   Dissemination materials (leaflet, brochure, poster) were designed and printed.

WP7 (Management)

There are efficient communication channels within and outside the consortium. An internal web site and a mailing list was designed and deployed which support interaction among consortium members. Almost every step of the project was lead by the initiative of the Management WP.

Achieved results after first year

After the 1st project year the translation page was launched. The webpage contains translation services of the founding members, as well as of associated and external partners. Beside the main achievement (a working web page of the main MT companies offering text translation, chat translation, web page translation and translated search) the web page offers special features that helps users in exploiting the service (virtual keyboard, user settings, automatic language detection, spell checker, translators' community).

Expected final results

The web site involves an innovative use of existing methods and tools, and allows all European citizens an efficient management of multilingual online content. To accomplish the "progressive abolition of language barriers in accessing multilingual content on the web", the project has developed the first European website that provides free online translation from any European language to any other. Since all the official European languages are supported, including all the under-resourced languages, the project also takes significant steps towards the objective of the equality of languages. The existence of a new online service used by several hundred thousand visitors every day will naturally stimulate business opportunities. In addition, the partners see real business opportunities as well, such as translation marketplace of human translators, possible integration to translation memories, advertising, etc.

However, there are also several positive impacts foreseen beyond mere damage limitation. The project is the first one in which most of the major stakeholders in the European MT industry collaborate together. The synergic effects of such collaboration are difficult to overestimate. The common developments of new interfaces and services, their endeavour to work out interoperable solutions will make a positive feedback onto our technology.

Furthermore, we expect additional impact on professional translators as up to now translators have not used machine translator programs.

Address of the public website
   www.itranslate4.eu

Relevant contact details:

Coordinator:	Technical Coordinator:
Tamás Váradi	László Tihanyi
Benczúr utca 33	Benczúr utca 33
1068 Budapest	1068 Budapest
HUNGARY	HUNGARY
Tel. +36 1 3214830 126	Tel. +36 1 3214830 131
Fax. +36 1 3229797	Fax. +36 1 3229797
E-mail: varadi@nytud.hu	E-mail: tihanyil@nytud.hu