



BOLOGNA
Translation Service

Annual Public Report

2011



KOÇ
UNIVERSITY



Contents

Executive Summary.....	3
Project Overview.....	3
Introduction	3
Consortium.....	4
Progress to Date.....	5
Achievements.....	5
Advisory Board.....	5
Project Meetings.....	5
Dissemination Activities.....	6
BTS Website	6
Publications.....	7
Conferences and Presentations	7
Press Releases.....	8
Collaterals	8
Future Work.....	8
Contact.....	8

Executive Summary

The Bologna Translation Service (BTS) project aims at promoting the implementation of the Bologna process by providing a high-quality low-cost Machine Translation (MT) service for higher education institutes to provide course syllabi and study programmes in English (and in one case, Chinese). Given the continuing and increasing need in this area, access to translated course syllabi and study programmes plays a crucial role regarding the level at which universities are able to attract good-quality foreign students and researchers. The regulatory environments in the context of the Bologna treaty combined with budget constraints and limited human resources make it very difficult for universities to deliver English (and Chinese) documentation, which impacts on the degree to which they can effectively promote their services locally in non-English-speaking countries, regionally, nationally and internationally.

The BTS approach will be to integrate existing MT components into a web-based collaboration framework. The basis will be statistical MT (SMT) engines for all language pairs. The baseline SMT systems created using the Moses toolkit will be further elaborated and improved by sophisticated techniques such as domain adaptation and automated post-editing. In order to complement these SMT engines, rule-based methods will be used. For Spanish and Portuguese we will use Apertium, and for Turkish and Finnish an adapted version of Convertus Hybrid Translator. Finally, a system combination strategy will be applied in order to further improve translation quality.

To date the activity in the project has mostly focused on collecting in-domain data for all languages in the project and attracting the attention of potential users by attending various events in the higher education space. By the end of this year the consortium partners will have baseline systems ready for all language pairs that will serve as a benchmark against which progress can be measured.

Project Overview

Introduction

Access to translated course syllabi and degree programmes plays a crucial role in the degree to which Universities and Higher Education Institutes (HEIs) are able to attract foreign students and profile themselves internationally. However, presenting all education information in English has proven to be a major challenge for a lot of Universities and HEIs.

The Bologna Translation Service (BTS) aims to help Universities and HEIs overcome this problem by offering a low-cost, web-based, high-quality machine translation (MT) service to translate course materials and degree programmes in English (and in one case, Chinese). The project will cover 9 languages (Chinese, Dutch, English, Finnish, French, German, Portuguese, Spanish and Turkish).

The work plan for BTS project consists of nine interconnecting work packages. Figure 1 gives an overview of the interaction between work packages.

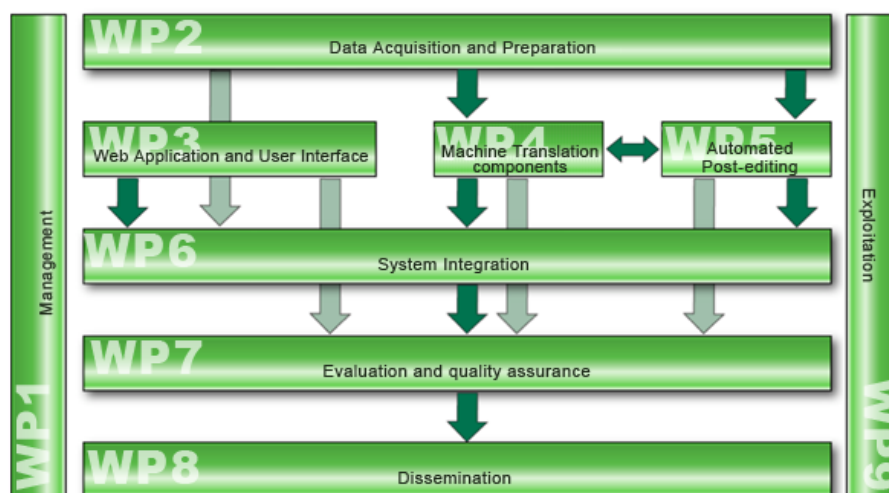


Figure 1. Bologna Work Package Structure

During the first months of the project, foundations for development work have been put in place, cooperation between the project partners has been established and the first results of the work of the consortium have been achieved. Management procedures were established and a project management plan created. Project partners established a common understanding and approach for the project implementation.

Consortium

The Bologna consortium consists of a mix of industry and academia partners, each with their own area of expertise.

CrossLang contributes its web-based collaboration framework – the CL Gateway – to the project as well as its MT evaluation expertise. Convertus brings its experience and technology from a high-quality MT service – the Convertus Syllabus Translator – in use by several Swedish universities. ALS provides world-leading experience with SMT including hierarchical techniques. The project takes advantage of Koç University’s competence in Turkish and relies on Eleka’s experience with rule-based MT systems to provide these important components in the project.

To ensure the outcomes of the project meet the expectations of potential users, a lot of effort has been put in building up a strong User Group that will assist in evaluating the service. To date, the project boasts 46 active user group members.

Progress to Date

Achievements

Taking into consideration that the project has been running for just eight months significant progress has been made towards achieving some critical milestones and goals:

- User study: A survey was conducted with registered user group members to collect feedback on service requirements.
- Data collection: A large amount of time was spent on collecting in-domain data for training the domain-specific SMT engines. Data was obtained through contributions from user group members and by crawling data from the web.
- Baseline engines: Strong SMT baseline systems were built for three key language pairs in the project (German into English, French into English, Spanish into English). These baseline systems will serve as benchmarks against which the quality of the BTS engines will be measured.

Advisory Board

Following members have agreed to join our Advisory board representing both experts in the field of MT and representative expert users:

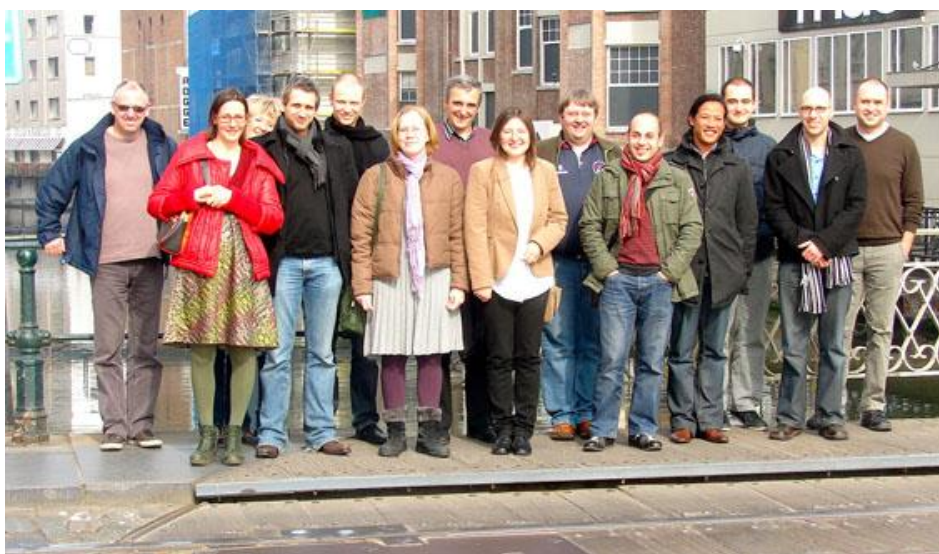
- Dr. Fred Hollowood, Symantec
- Prof. Björn Gambäck, Swedish Institute of Computer Science

Project Meetings

Two consortium meetings were held over the course of the first six months of the project:

- Project Kick-off meeting on March 14th/15th, Gent, Belgium

In addition to representatives from the BTS consortium, our kick-off meeting was attended by the Bologna project officer and two User group members



- Project General Assembly, August 1st/2nd, Uppsala, Sweden

Progress in all work packages was discussed and evaluated in view of key six month project deliverables. During the second day small work groups were formed to work through specific topics including the integration framework of the Bologna service.

Monthly EMB calls via skype are held to discuss progress and plan for upcoming deliverables and dissemination activities.

Dissemination Activities

We believe that our BTS logo used in all dissemination outlets helps people to recognise and identify with the service we intend to promote.

A dissemination plan was created as a deliverable at M3 detailing dissemination channels, target groups and a schedule of planned activities.

We intend to promote our project through a number of channels including:

- BTS website
- Publications
- Conference attendance with presentations, publications, BTS stand and demos
- Press releases
- Collaterals

BTS Website

The BTS website (at <http://www.bologna-translation.eu>) provides project information including a news section and links to downloadable content. It also displays the members of our user group.

The screenshot shows the homepage of the Bologna Translation Service website. The browser is Firefox, and the address bar shows bologna-translation.eu. The website has a navigation menu with links for HOME, PROJECT, CONSORTIUM, REGISTRATION, DOWNLOAD, and NEWS. The main content area features the Bologna Translation Service logo, a brief description of the service, and a section for User Group Registration. The registration section invites universities and other higher education institutions to join the Bologna User Group, offering free access to the pilot service in exchange for bilingual data and user feedback. It also provides a link to register and a link to the Bologna User Group. A sidebar on the right contains a 'Latest News' section with two articles: 'Bologna attended EAIE 2011' and 'Our First non-European User Group member'. The 'Coordinator Contact' section provides contact information for Heidi Depraetere, including telephone, fax, and email addresses.

Publications

- **Heidi Depraetere, Joachim Van den Bogaert and Joeri Van de Walle. 2011.** *Bologna Translation Service: Online translation of course syllabi and study programmes in English.* In Proceedings of the 15th International Conference of the European Association for Machine Translation 2011, Leuven, Belgium.

Conferences and Presentations

- Quality and Renewal 2011, May 19th, Uppsala, Sweden

Convertus presented a poster on the Bologna project as a contribution to the international evaluation of research at Uppsala University.

- EAMT 2011, May 30th/31st, Leuven, Belgium

Crosslang and ALS attended EAMT 2011. A paper was accepted in the User track and BTS was presented during the first day in the User track session.

BTS was also introduced in the short “poster booster” talks. As part of the main programme of the conference attendees were invited to come and see BTS at the poster sessions.



- Metanet, June 27th/28th, Budapest, Hungary

BTS was presented as a poster in the exhibition hall at META-NET.

- EAIE, (September 13th/15th, Copenhagen, Denmark)



BTS was represented with a booth at this three-day conference where more than 4,200 international higher education professionals from 87 countries met to exchange ideas and discover new trends and best practices. Our BTS stand provided us with excellent opportunities to talk to interested prospective users: in total, over 150 universities and organisations visited our stand and this resulted in signing up nineteen additional user group members from nine countries.

Press Releases

- BTS received press coverage in the Basque Country including press releases in the local newspapers [Gara](#) and [Noticias de Gipuzkoa](#), Information on the project was also posted on Lagune's web site, an association of Basque Country companies belonging to the Language Industry sector. Ibon Aizpurua from Eleka was also invited by local radio station Euskadi Irratia to talk about the Bologna project. The [interview](#) can be found on the web site of the Faktoria show.

Collaterals

A fact sheet and a poster have been created for distributions and display at our dissemination events.

Future Work

Before the end of this year, the Bologna consortium partners will introduce:

- A proof-of-concept system: This is being built for one language pair, German into English. The proof-of-concept system will demonstrate the first version capable of translating from German to English including automated post-editing. It is based on the baseline SMT engine, trained on available data and fine-tuned by means of automated post-editing.
- Additional SMT baseline systems: Baseline SMT systems for the remaining Bologna language pairs (Portuguese into English, Turkish into English, Finnish into English, Dutch into English and English into Chinese) are being built and will be ready by the end of the year.
- Initial RBMT systems: Rule-based components for Spanish into English and Portuguese into English are in the pipeline are being developed.

Contact

Please contact the project coordinator for additional information.

Name:	Heidi Depraetere
E-mail:	heidi.depraetere@crosslang.com
Phone:	+32 498 90 65 91