EU★BRIDGE

EU-BRIDGE
Technology Catalogue

The EU-BRIDGE Consortium

EU-BRIDGE Technology Catalogue

2nd edition, January 2015

The work leading to these results has received funding from the European Union under grant agreement n°287658.

EU-BRIDGE

Bridges Across the Language Divide

www.eu-bridge.eu

Use case University Lecture Translation

Use case European Parliament Translation

EU-BRIDGE develops automatic transcription and translation services that facilitate the development of innovative products and applications that require language technologies.

The project proves the usability of the services by implementing four applications as **use cases:**

- **Captioning Translation for TV Broadcasts:** Speech processing technologies improve the work flow in captioning TV shows and translating the captions into multiple languages. The use case demonstrates that by using speech translation technology it is possible to reduce the costs for producing subtitles for media content in many languages. Thus, it will be possible to subtitle and translate more media content and make it available to Europe's citizens.

- **University Lecture Translation:** Spoken content of university lectures is translated in real time and provided as subtitles. The use case demonstrates how speech translation technology can be applied in situations in which the use of human simultaneous translators would be too expensive and, in our case, making university lectures available to international students. The system is currently deployed in lecture halls at the Karlsruhe Institute of Technology.

- **European Parliament Translation:** EU-BRIDGE identifies opportunities for applying language technology in the interpretation process in the European Parliament. By automatically identifying unusual terminology and named entities and suggesting appropriate translations we support the parliament's interpreters in their preparation for meetings.

- **Unified Communication Translation:** A web-based speech translation service within a unified communication platform is introduced to support multilingual webinars. The lecturer's speech is translated into another language and output as text chat. The use case demonstrates how the user experience, when using different devices in a communication setting, can be improved by using speech translation technology.

EU-BRIDGE results: A high performing speech translation service infrastructure which leads to more cost effective media transcription and translation systems developed throughout the project in a commercial setting.

Content

Natural Language Processing and Supportive Technologies

The services offered by EU-BRIDGE to application developers are all centered around two natural language processing technologies: Automatic speech recognition (ASR) and machine translation (MT).

Automatic speech recognition is the art of automatically converting human speech into a written sequence of words. Within EU-BRIDGE we are especially dealing with Large Vocabulary Continuous Speech Recognition (LVCSR).

Machine Translation systems automatically translate text from a source language into a target language. Many machine translation systems are tailored towards written text that includes punctuation and other text structuring elements, such as paragraphs. However, within EU-BRIDGE, we are very often dealing with spoken language translation (SLT). For this, ASR and MT components are combined in sequence in order to be able to translate human speech, not only text. The translation of automatic transcripts of human speech adds further challenges to the machine translation component.

Automatic Transcription: Audio to Text in Real Time

Modern Large Vocabulary Continuous Speech Recognition (LVCSR) systems are learning systems that make use of statistical models. These models are trained on large amounts of transcribed audio data as well as large amounts of text in the target language. LVCSR systems work best when they are tailored to the specific domain that they are intended for. Also, LVCSR systems have a vocabulary that must contain all words that the system is supposed to recognize.

EU-BRIDGE has also researched and advanced techniques that allow such systems to automatically tune themselves to a new domain and speaker in an unsupervised manner, making use of auxiliary materials, such as slides of speakers or texts from the World Wide Web.

Statistical Machine Translation

Just like speech recognition, modern machine translation (MT) systems are learning systems that make use of statistical models. These models make use of large amounts of parallel texts that contain sentences in the source language and their translation into the target language. Also, large amounts of monolingual texts in the target language are needed.

And, just like speech recognition, MT systems work best if tailored to the exact application domain. Therefore, semi- and unsupervised adaptation techniques were also intensively researched for machine translation.

Supportive Technologies

When using the output of speech recognition or machine translation technology in real-life applications, the output often needs to be post-processed in order to meet the application's needs.

For instance, the output of a speech recognition system is usually an unprocessed sequence of words. Often enough, even numbers are not written as digits but merely as their sequences of spoken words. Human readers, however, expect additional markup to be included in the output, such as punctuation or text structuring via paragraphs, and the like.

Also, when dealing with natural language, humans perform many side tasks which they take as given, e.g. identifying the current language. Core ASR and MT systems lack these abilities.

Instead, additional supportive components are used to enrich both input and output of these systems to meet the user's expectations.

Technology Support for High-Quality
Automatic Speech Recognition Engines

Description and Exploitable Knowledge

Ongoing research provides the basis for state-of-the-art and high-quality automatic speech recognition systems. By applying newly developed technology, the recognition quality of existing engines is improved. Further, the engines become more robust and stable for a broad range of languages.

The challenge is to developed methods for the needs of different applications (e.g. recognition of lectures, TV shows etc.). Special focus is given to deep neural network techniques for acoustic and language modelling which improve speech robustness, domain adaptation and recognition accuracy in general.

The ultimate goal is to improve the quality of speech transcription in order to increase the recognition accuracy for transcription services like subtitling or interpreter support of the European Union. To reach this objective, on-going research for automatic speech recognition is very important.

Infrastructure

An essential part of the technology support is the actual transfer of new methods into the existing engines. However, before a newly developed feature is included into a running system, the impact will be verified in internal evaluations. Besides automatic evaluation metrics, tests based on the word error rate are employed. This verification pipeline ensures a continuous improvement of the used systems and a high-quality automatic speech recognition.

Terms of Availability

Can be inquired at RWTH Aachen University (Hermann Ney)

IPR Protection

RWTH Aachen University (Hermann Ney)

EU-BRIDGE Partners

RWTH, KIT, HKUST, UEDIN, FBK, PJIIT, PEV

Contacts

RWTH Aachen University
Templergraben 55
D-52056 Aachen
www-i6.informatik.rwth-aachen.de
Prof. Hermann Ney

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

EU★BRIDGE

The work leading to these results has received funding from the European Union under grant agreement n°287658

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

EU-BRIDGE Partners

Karlsruhe Institute of Technology (Germany), PEV

Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
Alexander.waibel@kit.edu

Automatic Transcription:
Audio into Text in Real Time

Description and Exploitable Knowledge

We convert audio into text in real time with a low latency. Each speech transcription worker contains two models, an acoustic model and a language model. 300+ hours of transcribed audio (Euronews, TED, podcasts, EPPS, etc.) are used for training the neural network or GMM acoustic model. The language model is trained on over 1,000 million words (from newspapers, transcripts, web dumps, etc.). These models can be optimized towards different tasks (Weatherview / Euronews / lectures / etc.) and/or different speakers resulting in many different possible workers.

Infrastructure

- Server - mediator - client setup
- The mediator receives transcription requests and audio from the client and forwards the audio to the corresponding ASR worker which then returns the transcription to the mediator.
- ASR workers run constantly on a server waiting for audio. They use up no CPU time when not receiving audio or when the audio only contains silence.
- As soon as a worker is selected by the mediator it starts to receive packets of audio data which it then decodes and returns text fragments to the mediator.
- These text fragments can then be combined into sentences by a separate segmentation/punctuation prediction component and then (if required) passed onto an MT worker.

Areas of Application

- Any situation which requires turning speech into text
- News (e.g., channels like Skynews and Euronews)
- Webinars
- Lectures
- Parliamentary speeches
- Weather reports
- Captioning TV programs
- Speech translation

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Technical Requirements

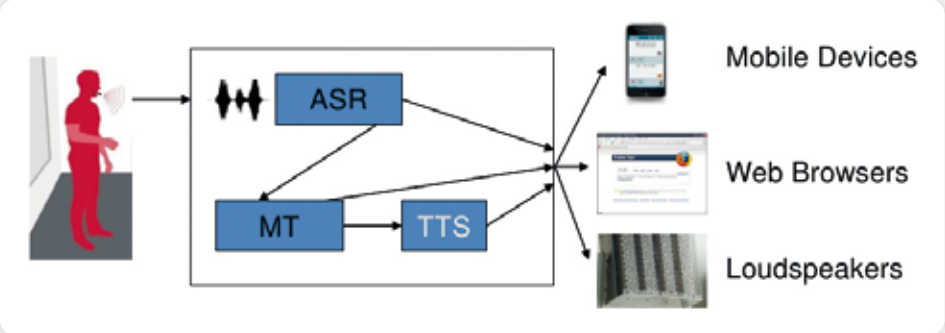
- Server with 8GB of RAM and 8 cores, 8 active systems can run on a modern server with 4 AMD Opterons (16 cores per CPU) as well as >>100 non-active but ready workers
- OS: Linux Ubuntu LTS 12.04 (Precise Pangolin) or similar
- Mediator connection: The workers can be accessed through the EU-BRIDGE Mediator service infrastructure

Terms of Availability

Can be inquired at the Karlsruhe Institute of Technology (Prof. Alex Waibel)

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel)



Lecture Translation

LT-Client

Display Server

EU-BRIDGE Partner

Fondazione Bruno Kessler (FBK)

Contacts

Fondazione Bruno Kessler (FBK)
Via Sommarive, 18
I-38123 Trento
http://hlt.fbk.eu
Roberto Gretter
gretter@fbk.eu

Euronews:
a Multilingual ASR Benchmark

Description and Exploitable Knowledge

Recording data from TV and monitoring the Web to look for new audio resources is a fundamental activity in Automatic Speech Recognition (ASR). The TV channel Euronews, which broadcasts news in several languages, is an attractive source of comparable data, which were used to design a multilingual speech corpus for ASR purposes, made of recordings from TV and downloads from the Web.

The corpus includes data in 10 languages: Arabic, English, French, German, Italian, Polish, Portuguese, Russian, Spanish and Turkish; it was designed both to train Acoustic Models (AMs) and to evaluate ASR as well as Language Identification (LID) performance. For each language, the corpus is composed of about 100 hours of speech for training (60 for Polish) and about 4 hours manually transcribed for testing. Training data include the audio, some reference text coming from the Euronews portal, which is sometimes a partial orthographic transcription, the ASR output and their alignment. Thanks to the light supervision technique, about 60 hours per language can be considered correctly transcribed.

This data is used inside the EU-BRIDGE consortium as a multilingual benchmark to evaluate ASR progress in the 10 languages, using similar amount of training data and comparable evaluation data. More details about this corpus can be found in the paper “Euronews: a multilingual speech corpus for ASR” by Roberto Gretter, in Proceedings of LREC, Reykjavik, Iceland, 2014.

Euronews Portal

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Areas of Application

ASR and LID training and evaluation

Terms of Availability

In 2013 the EU-BRIDGE consortium signed an agreement with Euronews, which gives to EU-BRIDGE partners the right to use Euronews material for research purposes, and to exchange it within the project.

Concerning the availability of this data for the whole research community, in 2014 Euronews agreed to make it available for research purposes. At present, part of the data is available as AM training data for the ASR multilingual evaluation benchmark for IWSLT 2014. We plan to make available more data for the next IWSLT evaluations.

language	#videos	speech duration	#ref words	#rec words	#common words	aligned speech
Arabic	4406	107:22:58	650,146	756,100	379,000	49:54:03
English	4512	112:18:29	973,210	1,032,727	699,850	63:02:34
French	4434	108:56:37	954,242	1,123,709	796,997	62:13:46
German	4438	108:33:23	809,289	896,387	653,372	61:46:27
Italian	4464	110:35:51	900,291	1,012,521	765,559	61:31:36
Polish	2626	58:32:03	350,729	454,977	278,854	27:42:35
Portuguese	4431	108:03:27	841,148	966,586	699,681	59:29:38
Russian	4418	107:42:24	714,363	828,060	611,347	60:49:15
Spanish	4465	109:16:50	939,408	1,053,255	797,698	63:29:23
Turkish	4387	106:30:31	683,041	764,329	556,760	60:52:55

EU-BRIDGE Partners

RWTH, KIT, HKUST, UEDIN, FBK, PJIIT, PEV

Contacts

RWTH Aachen University
Templergraben 55
D-52056 Aachen
www-i6.informatik.rwth-aachen.de
Prof. Hermann Ney

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

EU★BRIDGE

The work leading to these results has received funding from the European Union under grant agreement n°287658

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Technology Support for
High-Quality Speech Translation Engines

Description and Exploitable Knowledge

Ongoing research provides the basis for state-of-the-art and high-quality speech translation systems. By applying newly developed technology, the translation quality of existing engines is improved. Furthermore, the engines become more robust and stable for a broad range of language pairs.

The challenge is to developed methods for the needs of different applications (e.g. translation of lectures, TV shows etc.). Special focus is given to adaptation methods in order to take into account the fast changing domains of the application tasks.

Furthermore, the translation of speech is a more challenging task than the written text translation. In contrast to the translation of written text, speech translation has to deal with missing punctuation, speech disfluency and the recognition of errors (as the input is automatically transcribed speech). Thus, the aim of developed techniques is to fix theses errors and to reintroduce punctuation marks.

The ultimate goal is to improve the quality of speech translation in order to increase the user acceptance of the produced output. To reach this objective, ongoing research for speech translation is very important.

Infrastructure

An essential part of the technology support is the actual transfer of new methods into the existing engines. However, before a newly developed feature is included into a running system, the impact will be verified in internal evaluations. Besides automatic evaluation metrics, tests based on human scores are employed. This verification ensures a continuous improvement of the used systems and a high-quality speech translation.

Terms of Availability

Can be inquired at RWTH Aachen University (Hermann Ney)

IPR Protection

RWTH Aachen University (Hermann Ney)

Statistical Machine Translation

EU-BRIDGE Partner

Karlsruhe Institute of Technology (Germany)

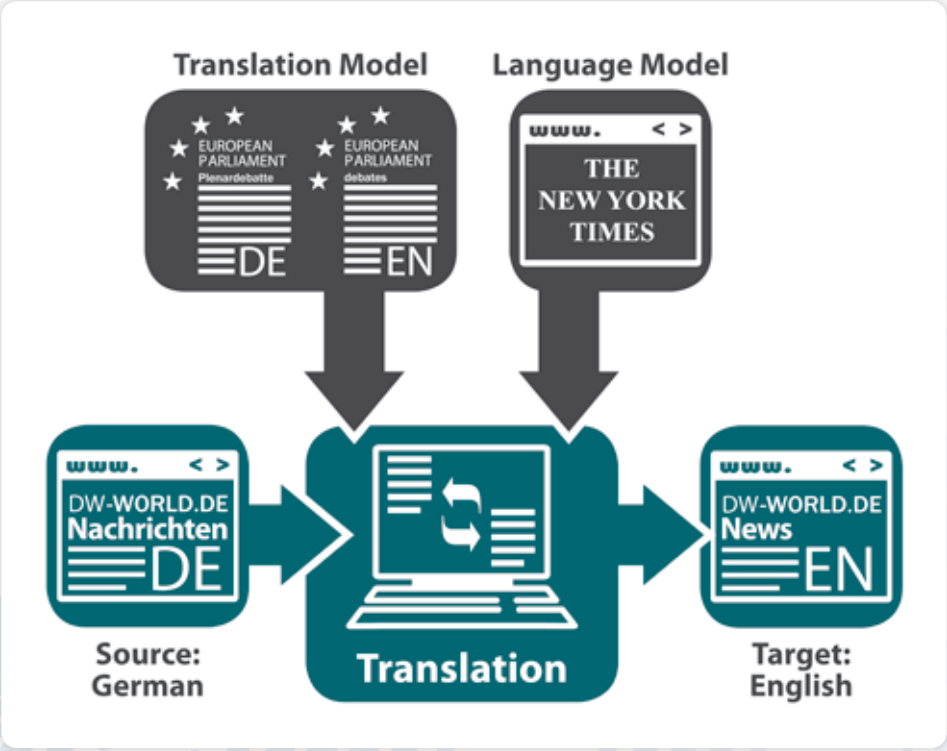
Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

Description and Exploitable Knowledge

A statistical machine translation system translates text and transcribed speech from one language into another language.

The systems are trained on large amounts of parallel and monolingual data using publicly available collections such as the EPPS or News Commentary corpora as well as data collected specifically for a given task. In order to achieve the best performance, the systems can be adapted to this task using small amounts of perfectly matching data (lectures, Sky News,…).



Translation Model

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Infrastructure

- MT engine runs as a worker in the EU-BRIDGE infrastructure
- Worker is constantly on the server waiting for text
- Worker receives segmented text from other workers in the source language and translates this text into the target language
- If needed, different workers can be combined to translate via a pivot language (e.g. French – English – German)

Application Sectors

- Sky News
- Lecture translation
- Webinars

Technical Requirements

- Server:
  - 8GB RAM and 1 core
  - Memory requirements can be reduced depending on the task
  - OS: Linux Ubuntu LTS 12.04

Terms of Availability

Can be inquired at the Karlsruhe Institute of Technology (Prof. Alex Waibel)

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel)

EU-BRIDGE Partner

University of Edinburgh

Contacts

The University of Edinburgh
South Bridge
UK-Edinburgh EH8 9YL
www.ilcc.inf.ed.ac.uk
Philipp Koehn
pkoehn@inf.ed.ac.uk

Open Source Statistical Machine Translation

Description and Exploitable Knowledge

Moses is an open source statistical machine translation (SMT) project, started in 2005 at the University of Edinburgh. Since then, hundreds of researchers have contributed to the project, and it has become the most widely adopted translation engine both as a baseline research system and for commercial use in industry.

There is extensive online documentation (http://www.statmt.org/moses) and there is an active mailing list (http://mailman.mit.edu/mailman/listinfo/moses-support) for support.

Technology / Application Sectors:

SMT models are the dominant technology for automatic machine translation due to their ability to leverage large amounts of human translated text. SMT systems can be trained within a few days, for any language pair, delivering state-of-the-art performance when enough training data is available. SMT models win annual machine translation competitions, and are deployed by the likes of Google and Microsoft. SMT models are trained on parallel corpora, where sentences, translated by humans, are aligned in the source and the target language. By using parallel corpora, Moses is able to discover which words or phrases are well translated from each other by looking at how often they co-occur. SMT models also use large amounts of monolingual text in the target language to learn what sentences in the target language should look like.

Components

The two main components in Moses are the training pipeline and the decoder. There are also a variety of contributed tools and utilities. The training pipeline is really a collection of tools (mainly written in perl, with some in C++) which take the raw data (parallel and monolingual) and turn it into a machine translation model and a model of the target language. The decoder is a single C++ application which, given a trained machine translation model and a source sentence, will translate the source sentence into the target language.

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Development

Moses is an open-source project, licensed under the LGPL, which incorporates contributions from many sources. There is no formal management structure in Moses, so anyone is welcome to contribute. For those interested in getting involved, there is a list of possible projects on the Moses website. The annual MT Marathon is also a good way to learn about the Moses project.

In general, the Moses administrators are fairly open about giving out push access to the git repository. This means that trunk occasionally breaks, but given the active Moses user community, it does not stay broken for long. The nightly builds and tests of trunk are reported on the cruise control web page, but for more stable versions, official releases are also available.

Moses in Use

The liberal licensing policy in Moses, in combination with its wide coverage of current SMT technology and complete tool chain, turn it into the probably most widely used open-source SMT system. It is used in teaching, research, and, increasingly, in commercial settings.

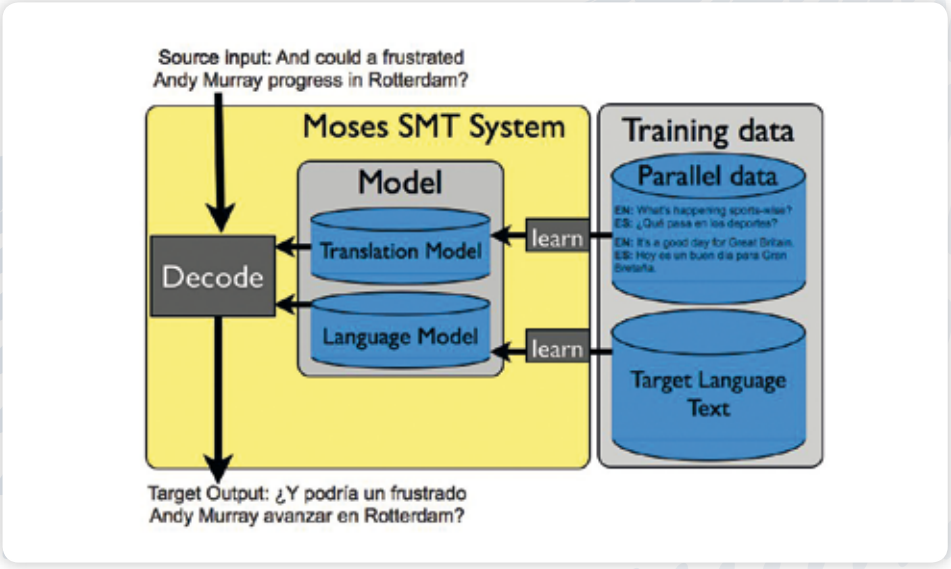
Commercial use of Moses is promoted and tracked by TAUS. Currently, the most common use for SMT in commercial settings is post-editing where machine translation is used as a first-pass, with the results then being edited by human translators. This can often reduce the time (and hence total cost) of translation. There is also work on using SMT in computer-aided translation, which is the research topic of two current EU projects, Casmacat and MateCat. In the EU-Bridge project, Moses is used for spoken language translation and for punctuating and segmenting ASR output.

Terms of Availability

Can be inquired at the University of Edinburgh (Philipp Koehn)

IPR Protection

University of Edinburgh (Philipp Koehn)



Moses diagram

Polish Spoken Language Translation

EU-BRIDGE Partners

KIT, Pervoice, FBK, UEDIN, PJIIT

Contacts

PJIIT
Koszykowa 86
02-008 Warszawa
Poland
Krzysztof Marasek
kmarasek@pjwstk.edu.pl

Description and Exploitable Knowledge

Polish, one of the West-Slavic languages, presents a challenge for both automatic speech recognition (ASR) and spoken language translation (SLT) due to its complex inflection and free word order. Seven cases, three genders, animate and inanimate nouns, adjectives agreed with nouns in terms of gender, case and number and a lot of words borrowed from other languages which are sometimes inflected like those of Polish origin, cause problems in establishing vocabularies of manageable size for translation to/from other languages and sparseness of data for statistical model training. Despite the existente of about 60 mio. Polish speakers worldwide, a number of publicly available resources for preparation of SMT system is rather limited, making the progress slower compared to other languages.

The EU-BRIDGE project helped us develop our skills and collect more resources necessary to accomplish the task of Polish SMT, specifically for Polish/English language pair. Our SMT systems have since been successfully applied to various domains, including lecture and broadcast news, parliament speeches, tourist assistance, movie subtitles and medicine-related documents. We developed our own BLSTM/DNN ASR decoder, Polish speech synthesis, mobile clients (Android) and specialized tools for the Polish language (text normalizer, corpus alignment, phonetic transcription, speech detection).

Android Client

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Infrastructure

Depending on the application, our systems work either on dedicated servers (e.g. in telecommunication) or on a cloud (with mobile applications). We currently support MCloud service architecture and the system developed by the U-STAR consortium compatible with ITU F.745 and H.625 recommendations.

Areas of Application

Telecommunication, tourism, medicine, entertainment

Technical Requirements

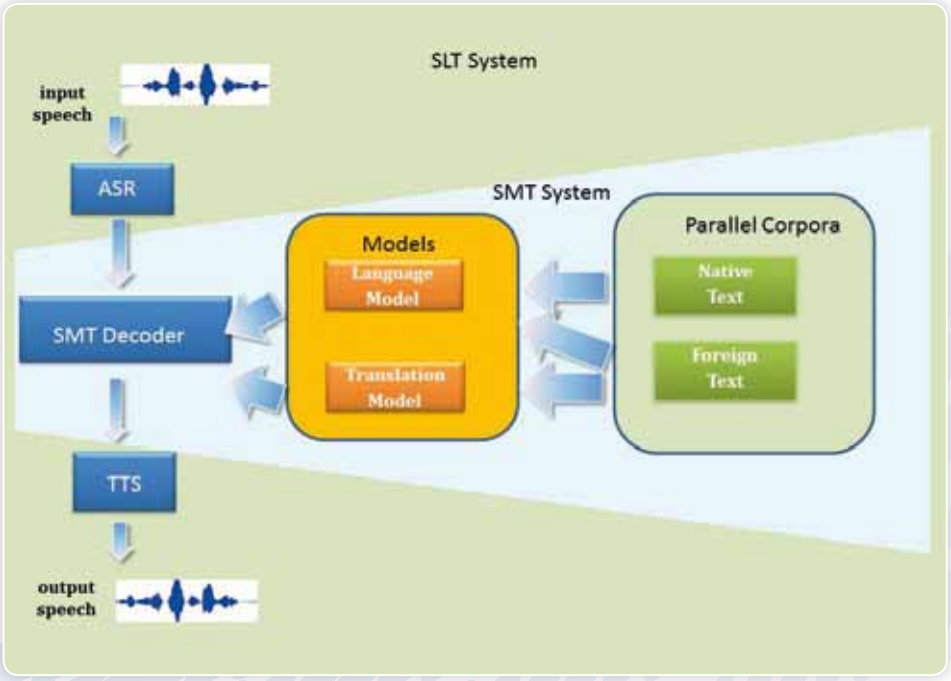
Fast internet connection or a dedicated SLU server

Terms of Availability

Can be obtained from the Polish-Japanese Institute of Information Technology

IPR Protection

Polish-Japanese Institute of Information Technology, EU-BRIDGE Consortium



SLT-Architecture

EU-BRIDGE Partners

Karlsruhe Institute of Technology (Germany), PEV

Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
Alexander.waibel@kit.edu

Segmentation and Punctuation

Description and Exploitable Knowledge

This Technology Catalogue entry describes the punctuation prediction worker that adds punctuation marks to unpunctuated text. This can be done either in an online mode, which only uses a 4-word window and a simple language model to decide which, if any, punctuation marks are to be added to a stream of text, or in an offline mode using a 20-word context window and multiple models.

The online system is trained on text which has been double-checked for correct punctuation. For the offline system, a phrase table trained on a large amount of text data as well as both maximum entropy and n-gram-based language models are combined log-linearly.

Infrastructure

- Server – mediator – client setup
- The mediator receives transcription requests and audio from the client and forwards the audio to the corresponding ASR worker which then returns the transcription to the mediator.
- ASR workers run constantly on a server waiting for audio. They use up no CPU time when not receiving audio or when the audio only contains silence.
- As soon as a worker is selected by the mediator, it starts to receive packets of audio data, then decodes the data and afterwards returns text fragments to the mediator.
- These text fragments can then be assembled into sentences by a separate segmentation/punctuation prediction component and then, if required, passed on to an MT worker.

Areas of Application

Wherever there is text without punctuation, e.g.:

- News (Sky and Euro)
- Webinars
- Lectures
- EPPS stuff
- Weather reports
- Subtitles
- 2nd step in speech translation
- and many others

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Technical Requirements

- Server with 8GB of RAM and 8 cores, 8 active systems can run in parallel on a modern server with 4 AMD Opterons (16 cores per CPU) as well as >>100 non-active but ready workers
- OS: Linux Ubuntu LTS 12.04 (Precise Pangolin) or similar
- Mediator connection: The workers can be accessed through the EU-BRIDGE Mediator service infrastructure

Terms of Availability

Can be obtained from the Karlsruhe Institute of Technology (Prof. Alex Waibel)

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel)

Service Architecture

Natural language processing technologies are sensitive to misconfiguration and handling errors. In the past, these have been serious barriers making the technologies' use and proper integration into complex applications difficult to impossible.

In order to facilitate the integration of speech recognition and machine translation technology into application, EU-BRIDGE offers its technologies via an easy-to-use, network-based service infrastructure.

The infrastructure makes use of a simple and easy-to-handle API which hides most of the complexity of the core technology from the application developer, thus fostering the use of natural language technologies in new, innovative applications.

EU-BRIDGE Partners

PerVoice (Italy),
Karlsruhe Institute of Technology (Germany)

Contacts

PerVoice SpA
Viale Verona 190/1
I-38123 Trento
www.pervoice.com
Dario Franceschini
dario.franceschini@pervoice.it
info@pervoice.it

Service Architecture

Description and Exploitable Knowledge

The service architecture is a platform to abstract the integration of client applications and transcription/translation service providers. Based on a lightweight library, the service architecture makes it easy for application developers to create complex transcription and translation workflows without having any knowledge of the underlying transcription and translation engines.

The service architecture decouples service providers and clients by providing a simple, XML-based protocol as well as a reference implementation library, available for the major platforms, to connect both end-user application and service engines to it.

To create transcription and translation workflows, the audio needs to be processed in a specific sequence by multiple engines. For instance, to produce a translated and punctuated text in Italian out of an English speech, one has to invoke the following engines:

- English speech to English text transcription engine;
- English phrase segmentation engine;
- English to Italian translation engine;
- Italian punctuation engine.

The service architecture simplifies the creation of this workflow by providing automatic workflow creation given the input and output language pairs (called fingerprints).

The service architecture provides APIs for both batch and real-time processing, supporting all transcription and translation needs. To simplify the integration of batch processes, it also provides a set of web-based REST APIs.

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

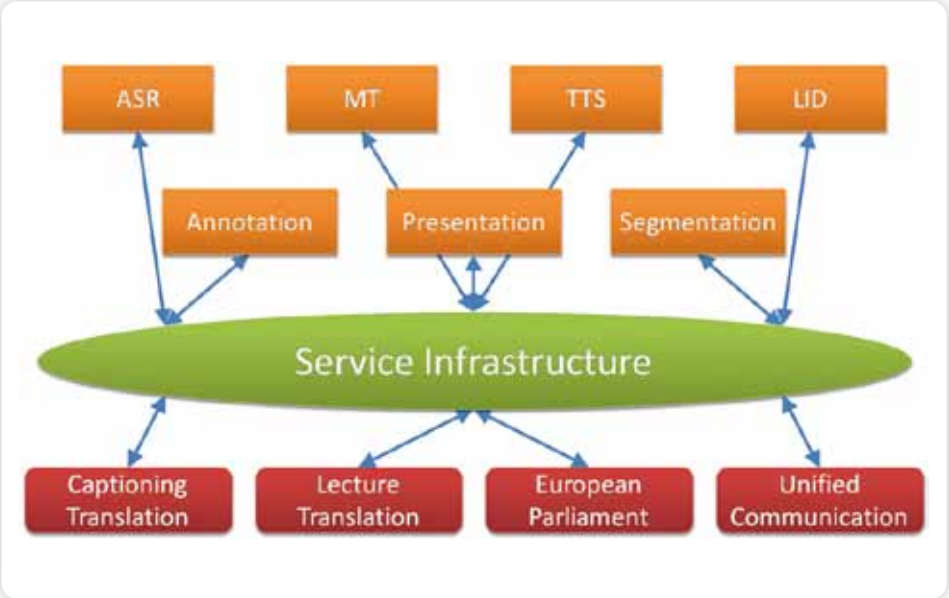
EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Infrastructure

The service architecture is a centralized service hosted at Karlsruhe Institute of Technology and PerVoice. It consists of two major components: the Service Mediator, which handles all the clients' and services' connections, and the REST API module.

The service architecture was designed to provide multiple, concurrent transcription and translation workflows by invoking the correct sequence of the required service provider engines (workers).



Service Infrastructure

Areas of Application

Transcription services, translation services

Technical Requirements

Fast internet connection

Terms of Availability

Terms and conditions can be obtained from PerVoice and KIT upon request.

Sample Applications

In order to prove the benefit of its service architecture, EU-BRIDGE has implemented five sample use cases which apply the technologies and service infrastructure developed in the project.

These use cases revolve around the transcription and translation of spoken content, such as TV broadcasts, lectures, webinars or translation services in the European Parliament.

The use cases were implemented with industrial leaders in the respective areas and are evaluated in real-life to measure the advantages stemming from the use of EU-BRIDGE technology.

EU-BRIDGE Partner

Karlsruhe Institute of Technology, Germany

Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
Alexander.waibel@kit.edu

Automatic Simultaneous Translation

Service for University Lectures

Bridging the language barrier with automatic simultaneous lecture translation systems

Spoken language translation: A combination of automatic speech recognition and machine translation

Description and Exploitable Knowledge

Academic lectures and technical talks provide high-quality content which is of value to audiences with many different mother tongues. However, many lectures only reach a certain part of the potentially interested, multilingual audience due to the limits imposed by the language barrier between lecturer and some of the listeners.

Lectures at the Karlsruhe Institute of Technology are mainly held in German. Therefore, foreign students who intend to study at the KIT need to learn German not only on a conversational level, in fact, they must be proficient enough to follow highly scientific and technical lectures carrying complex content. While students from abroad often take a one-year preparatory course to learn the language, experience shows that even after one year of studying, their level of proficiency in German is not high enough in order to be able to follow German lectures and thus perform well.

Since the use of human interpreters for bridging the language barrier in lectures is too expensive, we intend to solve this issue with the help of our automatic simultaneous lecture translation system. We use the spoken language translation technology (SLT), which combines automatic speech recognition (ASR) and machine translation (MT), together with other auxiliary components to build a system that simultaneously translates continuous speech from German to English.

The system works with the help of a cloud-based service infrastructure. The lecturer's speech is recorded via a local client and then sent to the service infrastructure. Afterwards, another service controls the flow of data through ASR, MT, and other components. The final result is then made available as a website which continuously displays the outcome of the recognition and translation process.

In addition, the system equally offers the possibility to file the lectures so that they can be searched afterwards in the archives.

Areas of Application

Universities, higher education, conference organizers

Technical Requirements

Fast internet connection to communicate with the server-based service architecture

Terms of Availability

Can be obtained from the Karlsruhe Institute of Technology (Prof. Alex Waibel)

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel)

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

EU-BRIDGE Partners

Red Bee Media Limited,
The University of Edinburgh

Contacts

The University of Edinburgh
South Bridge
UK-Edinburgh EH8 9YL
www.ilcc.inf.ed.ac.uk
Steve Renals
s.renals@ed.ac.uk

Red Bee Media Limited
58 Uxbridge Road
UK-W5 2ST London
www.redbeemedia.com
Matt Simpson
matt.simpson@redbeemedia.com

Automated Captioning
of Multimedia Content

Description and Exploitable Knowledge

One essential part of multimedia content is the use of subtitles in order to make it accessible to people suffering from hearing impairment. Although the process of captioning such content is already highly optimized with regard to accuracy and efficiency, it remains a rather costly and challenging part of the overall production process.

This can be attributed to the fact that captioning demands a significant amount of manual labour by trained experts. Primarily, subtitles are created through the act of 're-speaking'. The subtitler listens to the audio of the television programme and repeats it into a highly optimized, personalized ASR (Automatic Speech Recognition) engine, within an idealized acoustic environment, free of noise and reverberation.

Re-speaking

Re-speaking

The subtitles are usually created live, meaning there are stringent time constraints. Since there is no time to correct potential errors, accuracy is of paramount importance as well. Subtitling is therefore an extremely demanding and exhausting task for the human subtitler.

With the goal of producing subtitles more easily and less costly, a system was designed to produce captions in an automated way, reducing the need for human intervention.

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

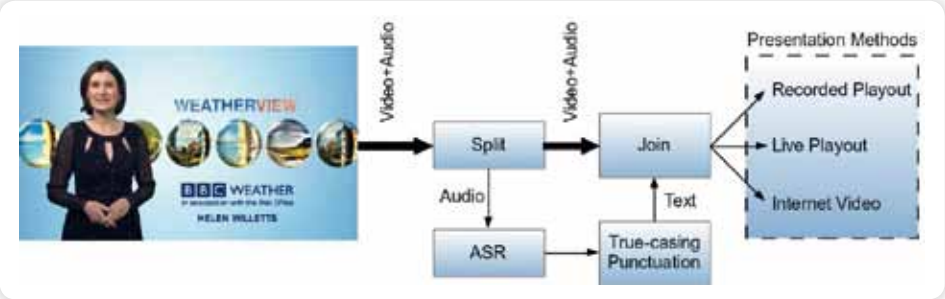
The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

A simplified schematic overview of the system is shown below. It is an end-to-end modular captioning system, that uses state-of-the-art speech and text processing technology. It generates highly accurate transcriptions of the spoken content and converts the resulting transcriptions into readable text by breaking it in sentences, inserting punctuation, and changing words to uppercase where appropriate. Assigning a time stamp to each subtitle is automatically done by the ASR module.



Schematic Overview

Areas of Application

Currently, two versions of the system have been set up and are in the process of being tested at Red Bee Media®:

1. Weather reports
2. Broadcast news

Although quality and accuracy of the generated subtitles are not yet on a par with manually produced ones, the system provides a valuable and dependable backup, should the manual production of subtitles fail.

IPR Protection

Red Bee Media®

EU-BRIDGE Partners

Karlsruhe Institute of Technology (Germany), RWTH Aachen University (Germany), Fondazione Bruno Kessler (Italy)

Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
Alexander.waibel@kit.edu

European Parliament Interpreter Support

Web Service on Terminology and Named Entities

Description and Exploitable Knowledge

A web-based service aiming to support the European Parliament (EP) interpreters to reduce their preparatory work and find key information easily has provided easy-to-use and user-friendly terminology extraction and named entity tagging.

As part of the preparatory work, interpreters need to go through the EP documents provided for each new session and look for terminology which, in many cases, is difficult to translate. Based on the same (PDF) documents, our web-based service allows interpreters to retrieve terminology lists and translate them automatically.

The named entities are important, since interpreters find it difficult to remember numbers, names, and so on, which complicates an accurate translation of these terms. Our web service provides a 13-type named entity tagging. Right now, it works on pre-selected documents, yet, it aims to work on written text as well as automatic speech transcriptions. Furthermore, the 13 types of named entities are highlighted in different colours.

Our web service supports the user in the uploading process of their own files. In addition, the tool also provides access to parliamentary documents enabling the interpreter to prepare for upcoming sessions even at home.

Finally, we invite you to access the web service under:
https://www.interpreter-support.eu

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

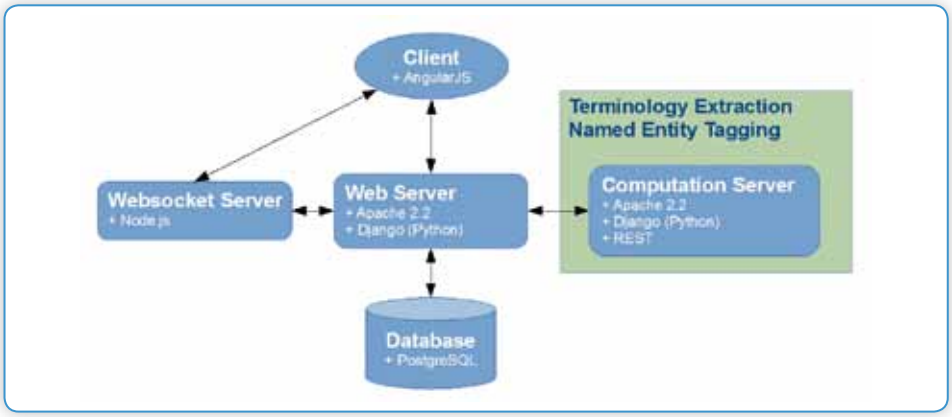
EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Infrastructure

The website is a single-page application (SPA) using the JavaScript framework AngularJS. On the backend, it is written in Python 2.7 and uses the Django framework. It runs on an Apache 2.2 web server with PostgreSQL as database. To support real-time events, we use a separate Node.js-driven WebSocket server. The hard work (terminology extraction and named entity tagging) is done by a dedicated computation server, with Celery as task runner and RabbitMQ as message queue.

The following diagram illustrates the setup currently used.



Areas of Application

Translators, interpreters, document editors

Technical Requirements

Fast internet connection

Terms of Availability

https://www.interpreter-support.eu

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel), European Parliament

Help interpreters to work faster and more easily.

Highlight the named entities in a document.

EU-BRIDGE Partner

Karlsruhe Institute of Technology, Germany

Contacts

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
Alexander.waibel@kit.edu

Automatic Simultaneous Translation

Service for Voting Sessions

Description and Exploitable Knowledge

Voting sessions generally take place in a multilingual setting. Each contribution by a member of parliament is made in the speaker's respective native language. To overcome language barriers between the members of parliament, interpreters render one language into another in real time. While, in principle, this is an effective solution to the problem, it is still both costly and extremely demanding for the interpreters, as voting sessions can take a considerable amount of time and speaker changes occur rapidly.

Our automatic simultaneous translation system for voting sessions is tailored to solve this issue by providing important aid to both audience and interpreters. In this system, we use state-of-the-art spoken language translation technology (SLT), which combines automatic speech recognition (ASR) and machine translation (MT), along with auxiliary components to build a system that is able to simultaneously translate speech produced in voting sessions. Speech recognition as well as translation into the target languages take place in real time, thus meeting the requirements of a voting session in parliament.

The system works with the help of a cloud-based service infrastructure. The speech held in the parliament is recorded via a local client and then sent to the service infrastructure. Afterwards, the flow of data is controlled automatically and passes through ASR, MT, and other components. The final translation result is made available in the form of a text flow which can be accessed by the audience.

In order to adjust the system to the needs of a voting session environment, it is trained on large amounts of parallel and monolingual data. The incorporated subsystems use publicly available data collections, such as the EPPS or News commentary corpora, as well as data collected specifically for this task.

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

Infrastructure

The EU-BRIDGE service architecture allows a server-based recognition and translation of an audio stream by providing a well-defined and lightweight API. Voting sessions can be monitored in real time via a client that is connecting to the service. ASR, MT, and other auxiliary components run as individual workers in the EU-BRIDGE infrastructure.

The service architecture enables a connection-based communication with multiple service requests at the same time. A client establishes a connection to the mediator on the server side and the mediator establishes a workflow of the client's output media stream through suitable components in order to accomplish a specific service request. The client's task is to capture the voiced audio of the voting sessions and present it to the server-side service architecture. The individual workers, such as speech recognition and machine translation, will be requested and allocated in order to complete the task.

Areas of Application

Voting sessions, parliament talks

Technical Requirements

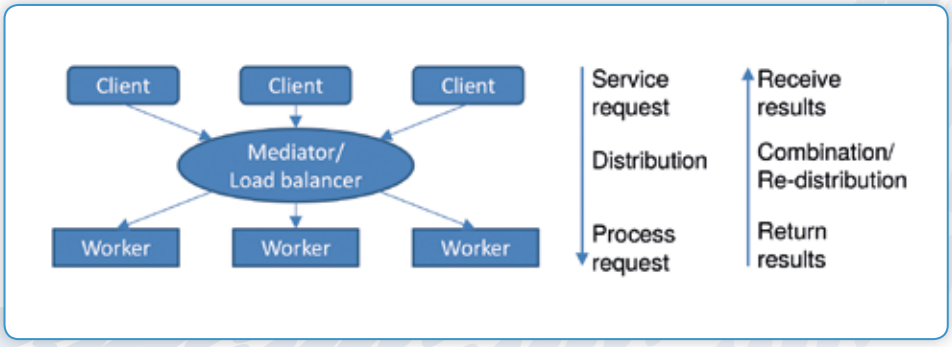
Fast internet connection to communicate with the server-based service infrastructure, recording equipment on client side

Terms of Availability

Can be obtained from the Karlsruhe Institute of Technology (Prof. Alex Waibel)

IPR Protection

Karlsruhe Institute of Technology (Prof. Alex Waibel), EU-BRIDGE Consortium



Schematic overview of the service architecture

EU-BRIDGE Partners

KIT for English ASR and English-French translation engines;
KIT for online webinar language model adaptation;
PerVoice for infrastructure;
PJIIT for English-Polish translation engine

Contacts

Karolyn Favreau
17 Rue Louise Michel
92300 LEVALLOIS-PERRET
France
karolyn.favreau@andrexen.com

The Serenty Webinar Platform for
Enhanced Multilingual Business Communication

Description and Exploitable Knowledge

Serenty aims at providing the end-user with real-time webinar captioning and webinar translation. This product addresses corporate clients for their cross-border webinars as well as professional users interested in foreign-language webinars.

Primary target language pairs are English-French and English-Polish.
Serenty's objective is to integrate a translation tool into an online communication system to enhance multilingual business communication. As a consequence:

- It reduces communication barriers between businesses from different countries
- It creates new markets for businesses which had formerly been inaccessible.

All conference participants can see the content of the webinar in their personal device's browser. They also receive the original audio (e.g. as presented by the speaker) and a written translation in their own language. The target language translation is provided in the form of subtitles. The transcript of the audio in original language might also appear on screen, in addition to its translated counterpart, as chosen by the user.

We shall highlight in particular the full web "no-plug-in" and "immediate-play" property of Serenty which is a key element of the system and represents a clear innovation compared to the state-of-the-art.

Infrastructure

Automatic Speech Recognition (ASR)

A cloud-based ASR engine receives a continuous audio stream sent by the unified communication platform. It returns the webinar's content in the form of a written transcript. This transcript also serves as a basis for machine translation. Prior to the webinar itself, in case the webinar content is made available to Serenty (e.g. through slides), a language model adaptation module analyses this content in order to detect specific product names and/or acronyms that are not part of the basic ASR engine's vocabulary. These words are added to the ASR vocabulary to allow their recognition.

Machine Translation (MT)

A cloud-based machine translation service takes the ASR transcript as input and returns a sentence translated into the target language as output.

Project Coordinator

Karlsruhe Institute of Technology
Adenauerring 2
D-76131 Karlsruhe
http://isl.anthropomatik.kit.edu
Prof. Alex Waibel
alexander.waibel@kit.edu

The work leading to these results has received funding from the European Union under grant agreement n°287658.

www.eu-bridge.eu

EU-BRIDGE - the Project

EU-BRIDGE is a European Integrated Project that aims at developing automatic transcription and translation technology that will permit the development of innovative multimedia captioning and translation services of audio-visual documents between European and non-European languages.

The source language is English, primary target languages are French and Polish. More language pairs involving German, Russian, etc. can be provided upon request.

Unified Communication Platform (UC)

The unified communication platform handles webinar session, slides, audio, and text inputs and outputs. It interfaces with the other required components (ASR, MT, LM adaptation). It also supports the cross-border webinar user interface on HTML 5 browsers.

Areas of Application

The suggested showcase addresses corporate clients for their cross-border webinars.
Serenty allows:

- Companies to present their products and services to a multilingual audience, thus having a maximum market impact;
- Executive managements and VPs to directly address international markets or to be part of presentations organized by local sales forces.

The technology is used as a complementary assistance tool rather than a critical communication component: It is aimed at webinars conducted in a chosen common language – English – but involving non-native speakers who will benefit from subtitles in their native tongue helping them to follow the webinar.
The domain is restricted to “sales & marketing” webinars, to “partners training & products updates” webinars, or in general to “corporate communication” webinars.

The technology’s approach as a “no-plug-in” solution allows the system to be compliant with all webinar systems available on the market (WebEx, Adobe-Connect,…)

Technical Requirements

User device:

- HTML 5 browser (e.g. Chrome)
- Device with microphone and speakers (preferably a headset)
- Link to Serenty

Fast internet connection

Terms of Availability

Can be obtained from Andrexen, End-to-End Unified Communications.

IPR Protection

No IPR protection taken into consideration when this document was published.

Project Coordinator

Karlsruhe Institute of Technology
Prof. Alexander Waibel
Adenauerring 2
76131 Karlsruhe, Germany
phone: +49 721 608 44730
e-mail: alexander.waibel@kit.edu

Contact

Karlsruhe Institute of Technology
Press and Communication
Margit Rödder
Adenauerring 2
76131 Karlsruhe, Germany
phone: +49 721 608 48676
e-mail: margit.roedder@kit.edu

www.eu-bridge.eu

EU-BRIDGE – Project Partners

Ten partners from university/research and industrial sectors form the EU-BRIDGE consortium:

- **Karlsruhe Institute of Technology (KIT)**
  Germany, Adenauerring 2, 76131 Karlsruhe
  Contact: Alexander Waibel, Web: isl.anthropomatik.kit.edu
- **Fondazione Bruno Kessler (FBK)**
  Italy, Via Santa Croce 77, 38122 Trento
  Contact: Marcello Federico, Web: hlt.fbk.eu
- **Polish Japanese Institute of Information Technology (PJIIT)**
  Poland, ul. Koszykowa 86, 02-008 Warszawa
  Contact: Krzysztof Marasek, Web: www.pjwstk.edu.pl/en/
- **RWTH Aachen University**
  Germany, Templergraben 55, 52056 Aachen
  Contact: Hermann Ney, Web: http://www-i6.informatik.rwth-aachen.de
- **The University of Edinburgh (UEDIN)**
  United Kingdom, South Bridge, Edinburgh EH8 9YL
  Contact: Philipp Koehn, Web: http://www.ilcc.inf.ed.ac.uk/
- **The Hong Kong University of Science and Technology (HKUST)**
  China, Clear Water Bay, Kowloon, 999077
  Contact: Dekai Wu, Web: http://www.cs.ust.hk/~hltc/
- **Red Bee Media Limited (RBM)**,
  United Kingdom, 58 Uxbridge Road, W5 2ST London
  Contact: Nicola Greaves, Web: www.redbeemedia.com
- **PerVoice SpA (PEV)**
  Italy, Viale Verona 190/1, 38123 Trento
  Contact: Dario Franceschini, Web: www.pervoice.it/en/
- **Accipio Projects GmbH**
  Germany, Roermonder Str. 216, 52072 Aachen
  Contact: Volker Steinbiss, Web: www.accipio-projects.eu
- **Andrexen**
  France, 17 Rue Louise Michel, 92300 Levallois-Perret
  Contact: Grégoire Boutonnet, Web: www.andrexen.com

Imprint

Editors:
Alex Waibel, KIT, Karlsruhe, Germany
Sebastian Stüker, KIT, Karlsruhe, Germany
Margit Rödder, KIT, Karlsruhe, Germany

Graphic Design: Simone Winter, Rheinstetten, Germany


The work leading to these results has received funding from the European Union under grant agreement n°287658.