

Join us!

Partners

The **SAVAS End-User panel** accommodates advisors and experts from both industrial and scientific communities. It brings together Broadcasters, Subtitling Companies, Universities and organisations involved in the industries of Media, Subtitling, Accessibility and LVCSR.

The main goal of the End-User Panel is to give advice to development, thus supporting dissemination and contributing to bridge the project results into the market. The role of participants is to provide recommendations and feedback, to prioritise language and domain development, to take part in testing and evaluation together with SAVAS End-Users Partners.

The **Members of the End-User Panel will receive funds to participate in 4 End-User Workshops**, that are scheduled as follows:

- 1st End-User Workshop
November 2012 - it will aim at introducing the objectives and expected results of the project.
- 2nd End-User Workshop
May 2013 - it will be focused on the presentation of preliminary results of the project.
- 3rd End-User Workshop
September 2013 - linked to IBC 2013, it will be an Open End-User Workshop, for discussing possible uses of SAVAS results.
- 4th End-User workshop
may 2014 - it will be focused on the demonstration of the final project results.

**Are you interested in participating?
Contact us for any detail.**



Project coordinator: Dr. Arantza Dal Pozo

Head of the Speech and Language Technology Group
Vicomtech - Visual Interaction and Communication
Technologies Center
Mikeletegi Pasealekua, 57 - Parque Tecnológico E-20009
Donostia-San Sebastián, Spain
Phone: +34 943 30 92 30
E-mail: adelpozo@vicomtech.org

Dissemination Manager: Dr. Carlo Aliprandi

Language and Speech Solutions Manager
Synthema S.r.l.
Via Malasoma 24, I-56121 Ospedaletto (PI), Italy
Phone: +39 050 9656411
E-mail: carlo.aliprandi@synthema.it



www.fp7-savas.eu



**LIVE SUBTITLING
MADE EASY**

Respeaking and Automatic Speech
Transcription services for top quality
Subtitling and Captioning





The Savas Project

3 R&D companies and 5 user companies are working at European level **to develop a brand new generation of products and services for better production of Multilingual Subtitles** for the **Broadcasting and New Media Industry** and its stakeholders.

SAVAS collects audiovisual resources to build Large Vocabulary Continuous Speech Recognizers (LVCSR).

More specifically, SAVAS objectives are:

1. To make more effective the **acquisition and annotation of audiovisual language resources** produced by Broadcasters and Subtitling Companies
2. To deploy a **platform for audiovisual language resources sharing** through the most suitable legal and business data trading approaches
3. To **develop a brand new LVCSR technology for Multilingual Live Subtitling**
4. To show the impact of feeding **LVCSR technology** with existing audiovisual language resources for **Automated Subtitling**

Technologies

Automatic Speech Recognition is the technology underlying Respeaking and it is proving to be partially useful for Subtitling. Most Respeaking and Assisted Subtitling applications currently rely on commercial Speech Recognition systems.

Commercial systems have several limitations, because they have been designed for dictation applications, thus performing too poorly in specific tasks, like Live Subtitling.

Moreover their language coverage may be limited in specific domains, as they have been developed to target the main language markets (i.e. English, Spanish, French, German, ..) and they are not available for many EU languages and dialects.

SAVAS aims to develop a brand new Automatic Speech Recognition (ASR) technology that will **bridge the gap of current commercial systems**.

An initial set of 7 languages will be developed: **Basque, Spanish, Italian, French, German, Portuguese** and **English**.

A wide set of new Dictionaries will be created for the following domains:

- **Broadcast News**
- **Sports**
- **Interviews and Debates**

With SAVAS Data Collection and Acquisition methodology, based on the combination of automatic and collaborative approaches, the creation of new Languages or new Dictionaries will be automated as much as possible, thus reducing the creation effort and time-to-market.

Products and Services

The main goal of SAVAS is **to develop a brand new Speech Recognition (ASR) technology**, specifically tuned to the needs of the **Broadcasting and New Media Industries**.

The ASR Engines will allow the release and tuning of new subtitling products and services, such as:

1. a Speaker Independent **Transcription System** for live and batch production of multilingual subtitles
2. a collaborative **Respeaking System**, for live and batch production of multilingual subtitles

As a result, Broadcasters and Subtitling Companies will benefit of more productive and less expensive subtitling alternatives to the traditional process.

The main SAVAS results include:

TOOLS

- Transcription toolkit (SDK)
- Respeaking toolkit (SDK)

APPLICATIONS

- Transcription of news for Spanish
- Transcription of interview/debates for Portuguese
- Transcription of news for Italian
- Transcription of news for French
- Transcription of news for German
- Dictation of news for Italian
- Dictation of sports for Basque

The main results of SAVAS are available either as Products and as Services.