

Gaia

1. BASIC INFORMATION

1.1. Tool Name

GAIA: Common Framework for the Development of Speech Translation Technologies

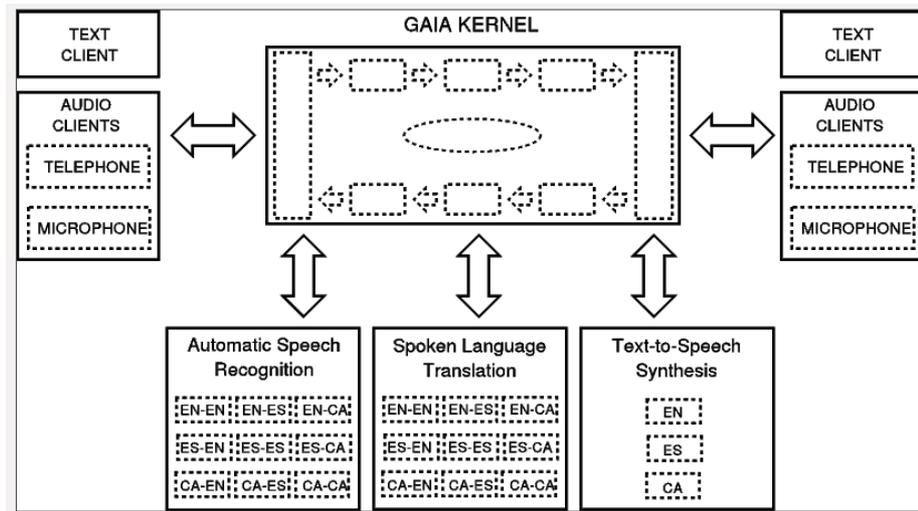
1.2. Overview and purpose of the tool

GAIA is an open-source software platform for the integration of speech translation components. The tool can be used to facilitate and record person-to-person and machine-to-person communication.

This tool is useful to integrate into a common framework different automatic speech recognition, spoken language translation and text-to-speech synthesis solutions. Gaia operates with great flexibility, and it has been used to obtain the text and speech corpora needed when performing speech translation. The platform follows a modular distributed approach, with a specifically designed extensible network protocol handling the communication with the different modules. A well defined and publicly available API facilitates the integration of existing solutions into the architecture. Completely functional audio and text interfaces together with remote monitoring tools are provided. The speech and audio engines are not included but the interface is defined and skeleton engines are provide to help the integration of engines.

The configuration file allows to use GAIA in several scenarios, for facilitating the communication between two people or to provide service to one person.

- Speech-to-speech translation using telephone lines. Two users participate in a dialog and GAIA translate the turns. Speech recognition, spoken translation and speech synthesis is required. This was the scenario of the demonstration prototype of the LC-STAR EU Project.
- Speech/Text translation system. In this scenario one of the dialog participants uses speech and the other uses text.
- Speech collection: GAIA can manage data collection either from the desktop or from the telephone. The user can be prompt either using text or using pre-recorded or synthetic speech.



1.3. A short description of the algorithm

GAIA structure follows a modular design, where different parts can be launched in separate machines. The architecture uses a client/server approach, where each part is conceived as an individual entity, with the ability to run independently of the platform. Standard sockets and a predefined network protocol are used to interconnect the different modules of the system. The speech processing tasks (recognition, translation and synthesis) are implemented in independent servers (not provided), communicating with the platform kernel via standard network sockets. The kernel of the platform communicates with different types of servers: terminals (telephone, speech console and text console), speech servers (ASR, TTS and SLT). The Gaia kernel handles the communication among the terminals (platform clients) and the different speech technology servers. The kernel keeps a database with the network location (IP address and port) of the different ASR, SLT and TTS servers, and establishes the appropriate connections based on the language, preferences and characteristics of the client. Several options are supported to select the pair of languages for the source and target terminal: either the source client sets both, or the destination client selects its language. Predefined pairs of languages are also available during the configuration stage. Depending on whether the terminal servers use audio or text, the platform uses the speech recognizers and synthesizes accordingly.

2. TECHNICAL INFORMATION

2.1. Software dependencies and system requirements

Portability among different platforms is guaranteed by using standard C++ for programming the different parts. The remote monitoring and configuration applets are programmed in Java and can be used from several Java compatible browsers (e.g. Mozilla, Galeon or Opera). Since the software is provided together with full sources and technical documentation, developers will find it easy to adapt any related technology to Gaia.

To compile in Linux environments (last tested on Ubuntu 12.04 precise) the development tools make, gcc, g++ and javac compilers are required. Furthermore, the libncurses5 is needed for the console terminal and ALSA libraries are needed for the audio support (libasound2 package on Debian and derivatives).

2.2. Installation

```
# Decompress gaia-1.2.tgz
tar xzf gaia-1.2.tgz
# Compile the sources
cd gaia/gaia_prj
make all
```

2.3. Execution instructions

The directory bin/release and bin/debug should contain a set of executable programs to be executed from the terminal:

```
kernel platform: gaia
servers: audio_user cmd_bridge cmd_tester logclient_txt
prompt_client text_user wav_client
```

The source of some template servers are provided to include speech recognition, speech synthesis and machine translation engines.

2.4. Input/Output data formats

2.5. Input data formats

The kernel (gaia) and the servers require configuration files to indicate server/client ip, port, and other configuration aspects. Configuration files are provided in the documentation.

2.6. Output data formats

Depending of the configuration different text and or audio log files are generated. The audio files are stored using RIFF/WAV files. In case that the telephone interfaz is used, the signal is sampled at 8kHz and 8bits/sample. If the speech console is used, the signal is sampled at 16kHz and 16 bits/sample.

4. ADMINISTRATIVE INFORMATION

4.1. Contact person

Name: Antonio Bonafonte
Address: Jordi Girona 1-3 Edifici D5, 08034 Barcelona, Spain
Affiliation: TALP research center. Universitat Politècnica de Catalunya
Position: Professor
Telephone: +34 93 401 6437
Fax: +34 93 401 6447
e-mail: antonio.bonafonte@upc.edu

4.2 Delivery medium (if relevant; description of the content of each piece of medium)

The resource will be uploaded on the MetaShare platform as an archive.

4.3 . Copyright statement and information on IPR

The resource has copyright. The Copyright belongs to Universitat Politècnica de Catalunya. The resource is free, license-based, for research purposes and free license-based for commercial purposes.

5. RELEVANT REFERENCES AND OTHER INFORMATION

Javier Pérez and Antonio Bonafonte.
GAIA: A Common Framework for the Development of Speech
Translation Technologies
5th Int. Conf. on Language Resources and Evaluation, LREC 2006
Genoa, Italy. May, 2006.

Javier Pérez, 2004. Gaia technical report. TALP Research Center.
Included in this package.