

Code: FP7-ICT-2011-7, Nr. 287678

Local coordinator; Prof.dr.ing. Mircea Giurgiu

Team: Dr.ing. Adriana Stan, Dr.ing. Jozef Domokos, Dr.ing. Cristian Contan, Drd.ing. Andrei Homodi, Msc.eng. Andrei Barbos, Msc.eng. Ioana Muresan.

eMail: [Mircea.Giurgiu@com.utcluj.ro](mailto:Mircea.Giurgiu@com.utcluj.ro)

Web: <http://simple4all.org>

## OBJECTIVES:

- to create a portable and adaptable speech synthesis technology suitable for any domain or language;
- to provide a complete framework for automatic learning and self-improvement of the system;
- to enable the generation of highly expressive synthetic speech;
- to automate the creation of a new speech synthesiser from scratch.

## ACHIVEMENTS:

### Core Techniques:

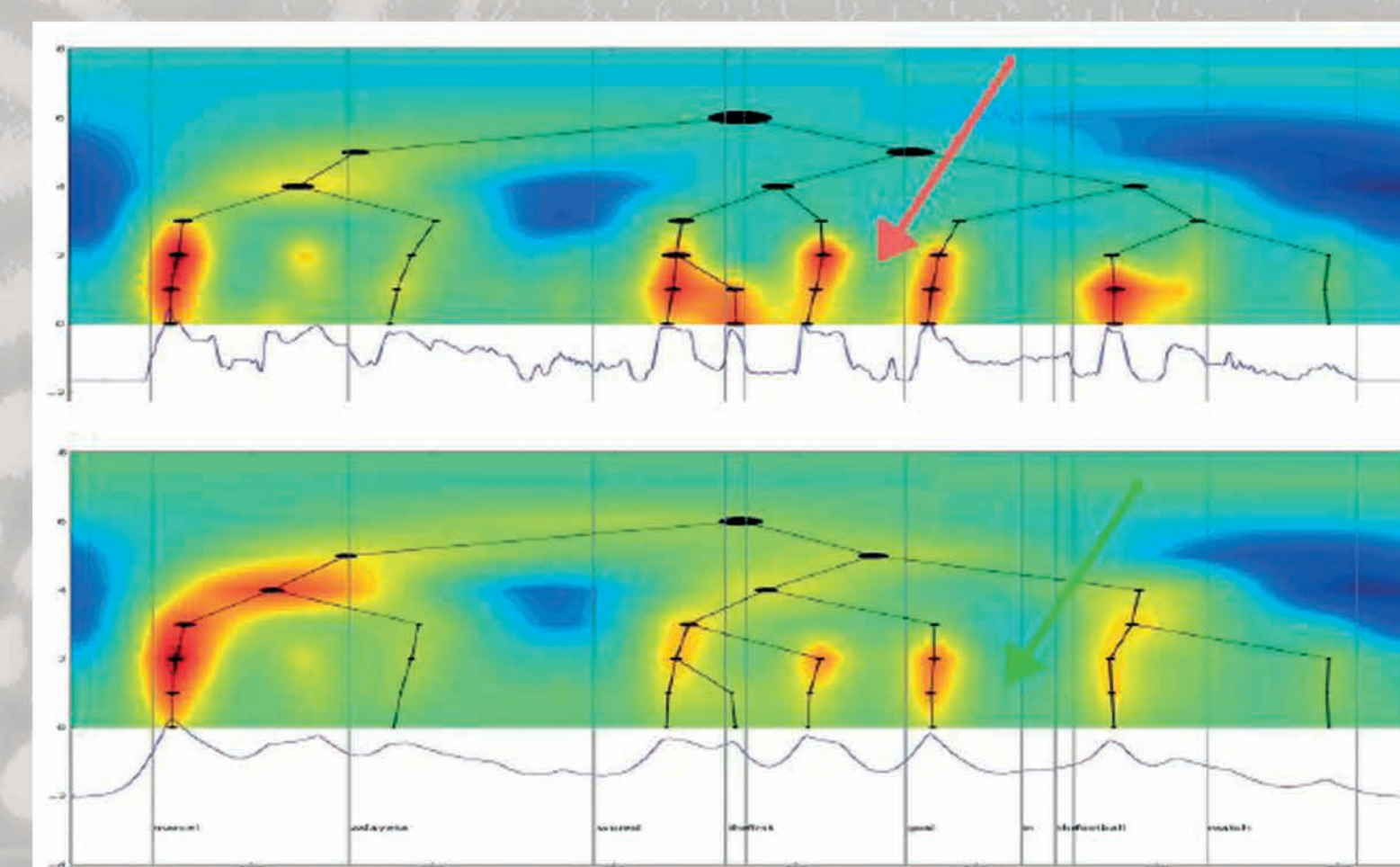
- Unsupervised linguistic representation using vector space models
- Text Normalization using statistical machine translation
- Vocoding using physiologically-motivated approaches.

### Enhancing Technologies:

- Speaking style diarization and speaking style transplantation
- Automatic use of user interaction and feedback to automatically improve systems.

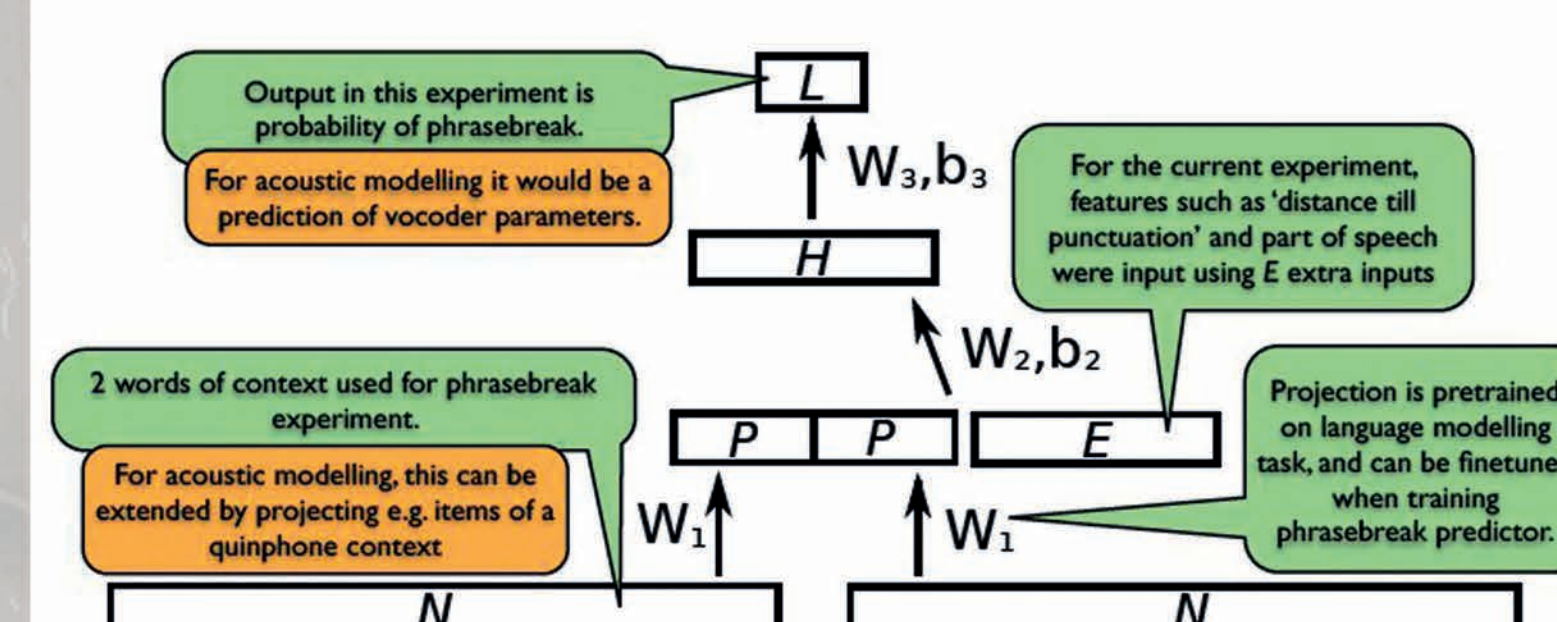
### Main results:

- DEXTER - a complete diarisation system that uses unsupervised machine learning and being able to detect who is speaking when.
- NORMA - a language independent text normalization toolkit using data driven machine translation models.
- OSSIAN - a tool to create the text processing front end in a speech synthesis system.
- ALISA - a tool for automatic alignment of imperfect speech and text data sets.
- TUNDRA - a multilingual corpus of 14 languages created from automatic labeling of audiobook data using ALISA.

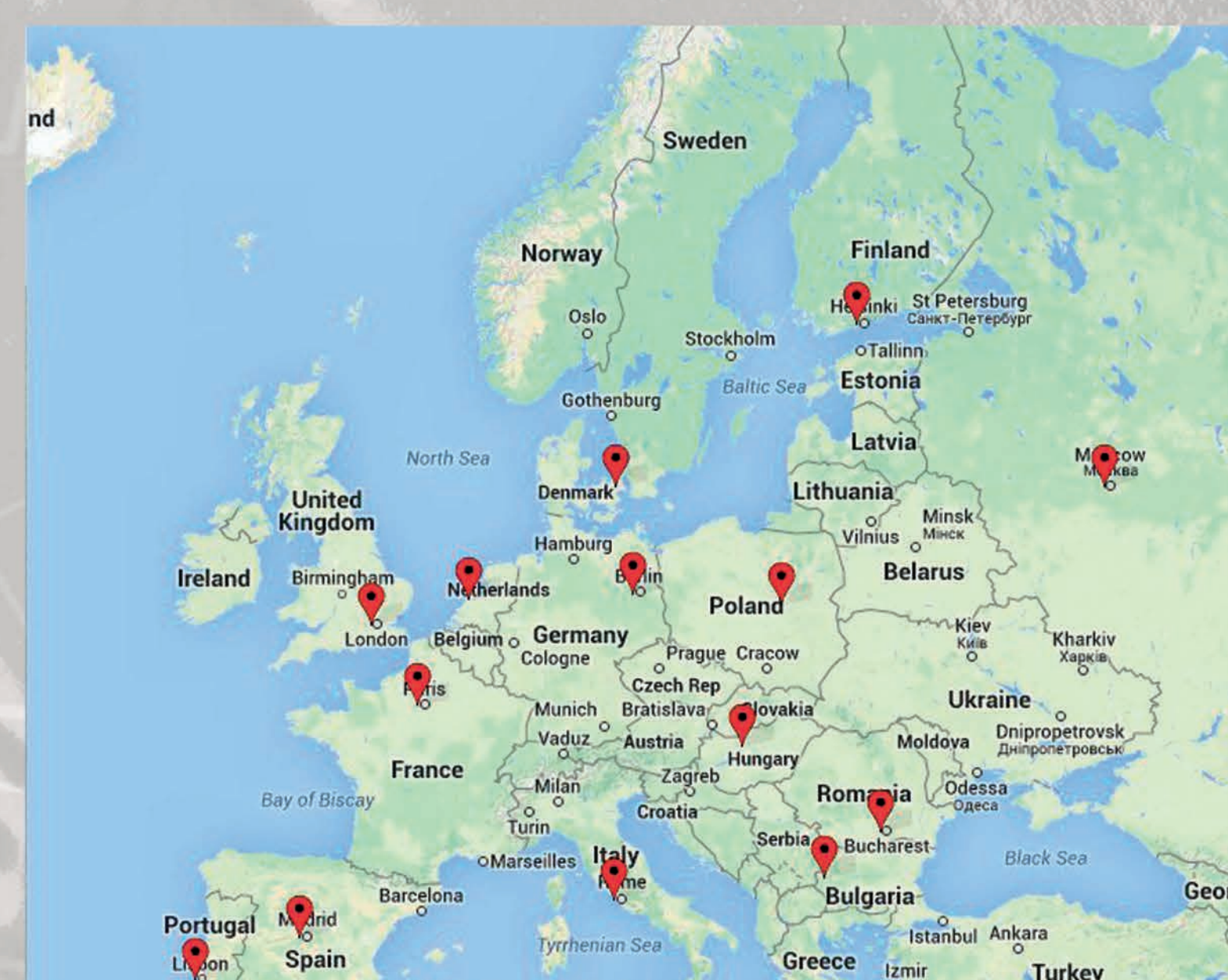


Automatic Syllable Segmentation

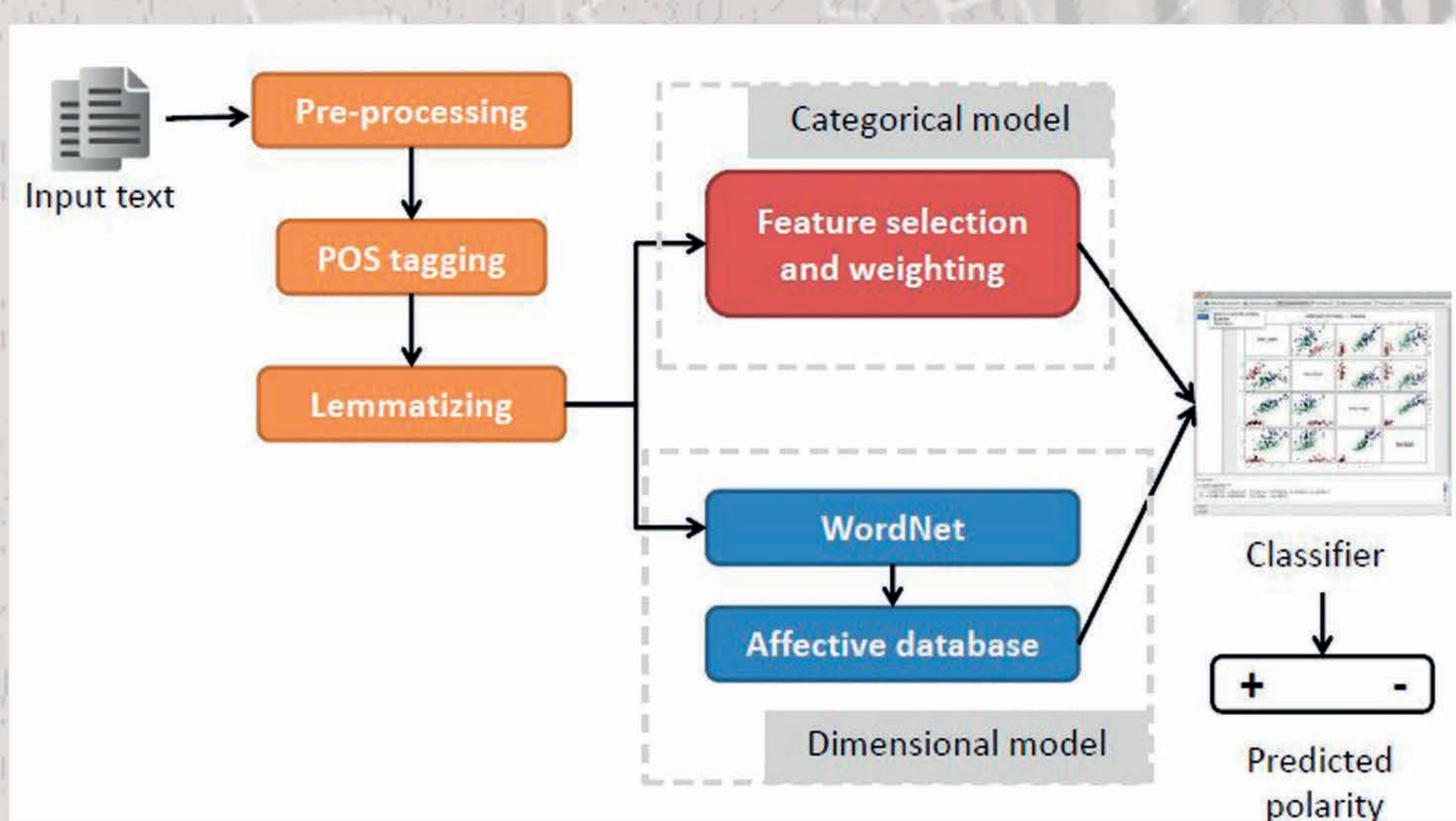
## General shared projection predictor



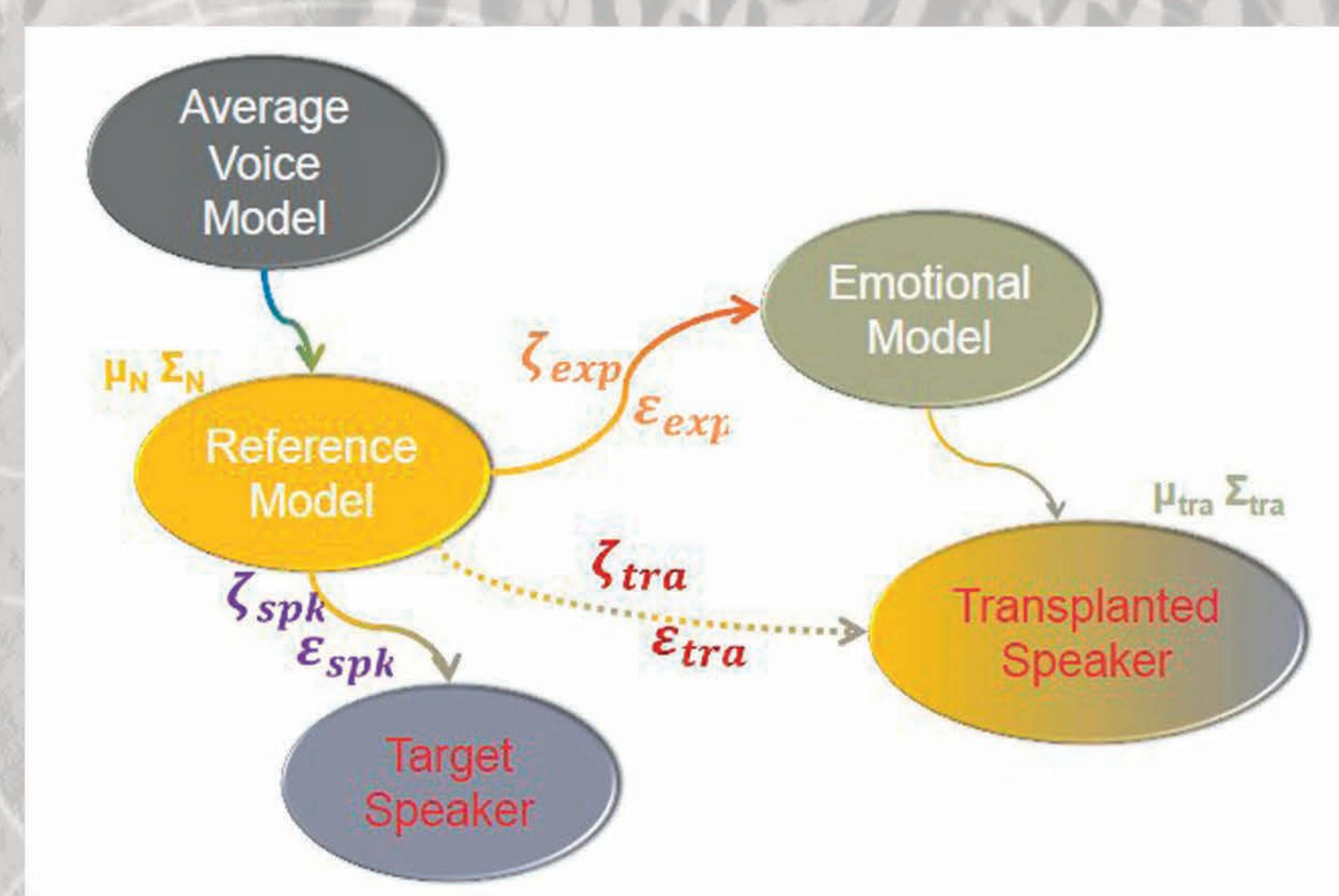
Deep Neural Network Predictor



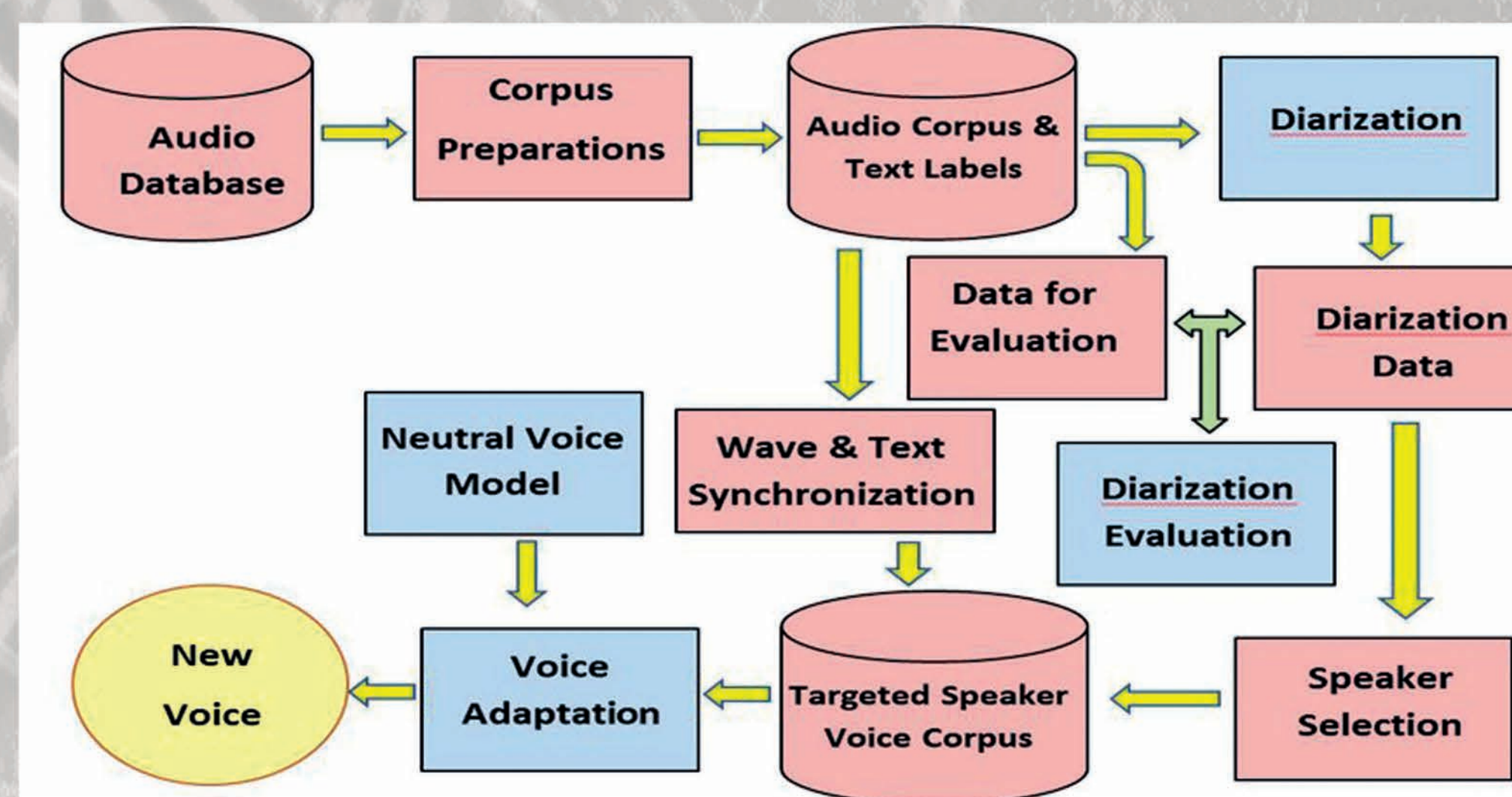
Multilingual Audiobook Data



Sentiment Analysis



Speaking Style Transplantation



Voice Cloning

## APPLICATION FIELDS:

- Mobile applications for assistive technologies
- Interactive Voice Response Systems (IVRs),
- Creation of new synthetic voices in new languages using small amount of training data,
- Voice Cloning applications,
- Machine reading applications.
- Creation of expressive voices according to the text genre and sentiment in the text.

## Conclusion:

SIMPLE4ALL enables the building of speech synthesis systems automatically from data, with little or no expert supervision.