

# **Language Data Space (LDS)**

**workshop in Romania**

**Language Data and Language Technologies  
in Romania and for the Romanian language**

**Panel Session**

# Emotion Recognition from Speech

## Deceptive Speech Detection (DSD)

### Discrete categories (classification):

- **simple binary classification: deceptive (untruthful) vs. non-deceptive (truthful) speech;**
- **main interest: global untruthfulness (annotations available for subjects lying at the utterance / phrase / turn level, even if parts of their speech are truthful);**
- **alternative: local untruthfulness (annotations available for subjects lying at the segment / short phrase level).**

### Databases

#### Main problems:

- **simulated data, using actors or amateur speakers (e.g. students) => often little relevance for real-life situations;**
- **controlled environment / scenario => predictable, lacking spontaneity;**
- **low-stakes context => little incentive for participants to “do their best”;**
- **subjective data annotation => relatively unreliable data.**

## **Romanian:**

- **19 speakers;**
- **25 very long recs. (~13.500 utterances);**
- **authentic and reliable data (non-subjective annotation, high-stakes contexts).**

# DNN Approach to Romanian Speech Recognition

## Experimental setup. Speech Corpora

- **Read Speech Corpus (RSC):**  
Read & clean speech utterances in silent environment.
- **Spontaneous Speech Corpus (SSC):**  
Spontaneous utterances from talk shows and news broadcasts.
- **Unlabeled corpus:**  
Speech crawled from Romanian online media (two news websites and one radio station), during a 9 months period.

| Purpose    | Set       | Size        |
|------------|-----------|-------------|
| Training   | RSC-train | 94 h , 46 m |
|            | SSC-train | 130 h, 44 m |
| Evaluation | RSC-eval  | 5 h, 29 m   |
|            | SSC-eval  | 3 h, 29 m   |
| Annotation | Source #1 | 367 h, 57 m |
|            | Source #2 | 331 h, 44 m |
|            | Source #3 | 78 h, 12 m  |

# **GMM-UBM modeling for speaker recognition on a Romanian large speech corpora**

**Speaker recognition (verification & identification)**

**Objectives:**

- **Comprehensive experiments on a 5 times larger corpus than other previous Romanian language attempts**
- **Usage of RoDogits corpus: over 120 speakers, Romanian connected-digits**
- **Usage of the GMM-UBM framework**
- **Getting error trends related to variation of some of the most important parameters of the system**

## Experimental setup. Speech Corpora

- **RoDigits v0.9**
- **Connected-digits corpus in the Romanian language**
- **129 speakers, 100 utterances each one**
- **Each utterance contains 12 random digits**
- **Corpus total size: 31 hours**

| <b>Speech set</b>            | <b># of utterances per speaker</b> | <b># of speakers</b> |
|------------------------------|------------------------------------|----------------------|
| Enrollment corpus            | 80                                 | 100                  |
| Closed set evaluation corpus | 20                                 | 100                  |
| Open set evaluation corpus   | 100                                | 29                   |

**The International Conference  
on  
Speech Technology  
and  
Human-Computer Dialogue**

**“SpeD”**

**sped.pub.ro**

**The International Conference on Speech Technology and Human-Computer Dialogue (“SpeD”) is now at the 13<sup>th</sup> edition.**

**After a first edition in the 80’s years, we decided to resume the events in 2003, under the Aegis of the Romanian Academy - Section of Information Science and Technology and with the invaluable help of Academician Mihai Drăgănescu, former President of Romanian Academy.**

At the beginning (**the first edition in 1986**), our Conference was focused more on **the applications of digital signal processing**.

But after 2000, The Conference was more and more connected to the tendencies in the domains of **spoken language technology and human-computer dialog**.

We also intended to evaluate and encourage the Romanian achievements in these fields and to invite representative specialists in the domain from abroad.

**Every 2 years we had new editions of the Conference that took the name SpeD (Speech and Dialogue) and we wanted to involve university centers in the country with achievements in the field: București, Iași, Cluj, Timișoara, Brașov, Constanța.**

**The Conference has become, we hope, a meeting place between specialists in the field from different parts of the world, with the conviction that they can share ideas and achievements with each other.**

**The Conference has expanded in the meantime; we are trying to cover a wider area of topics related in one way or another to the initial purpose.**