WELCOME

to

The 9th International Conference
on

Speech Technology
and
Human-Computer Dialogue

"SpeD 2017"

Bucharest, July 6-9, 2017

The International Conference on Speech Technology and Human-Computer Dialogue ("SpeD") is now at the 9th 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, 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.

In the 2nd edition (2003), we noticed the evolution from *speech technology* to *spoken language technology*.

The big step from small vocabulary recognition machines to medium size vocabulary voice command and control systems for information retrieval and electronic commerce, till large vocabulary speech dictation, spontaneous speech understanding, and limited domain speech translation.

It was the moment to point out the listening and talking machines large impact on the society.

In 2005, we appreciated that no dramatic changes occurred in the domain, since our past Conference; however, some trends were more and more obvious and some new fields of interest appeared to be a promise for the future.

We were able to identify a constant development of what is called "*speech interface technology*" which includes

- **automatic speech recognition,**
- **synthetic speech,**
- **natural language processing.**

For the edition in 2007:

The evolution of speech and language technologies over the past decade has spawned a stimulating new research area: Spoken Language Technology (SLT). The important issues pointed out at this workshop could be summarized as:

- Advances in natural language understanding and large vocabulary continuous speech recognition.

- Advances in machine translation technology.

-Advances in information search and data mining.

One of the main issue emphasized is about the interaction between speech and NLP (natural language processing).

Another interesting dimension of the domain: "Multilingual Language Processing"; more than 6.900 languages in the world.

trends:

- **spoken language understanding and generation,**
- **information retrieval (legal/medical/education),**
- **Internet businesses,**
- **search,**
- **spam filters,**
- **user generated content,**
- **the need to support multiple input and output languages, especially if applications are intended for international markets, linguistically diverse user communities, and nonnative speakers.**

The 2009 Conference edition was noticed some possible expansions of the trends already mentioned:

- **spoken language understanding,**
- **spoken language generation,**
- **machine translation of spoken language,**
- **spoken document retrieval,**
- **information extraction from speech,**
- **spoken document summarization,**
- **speech data mining and search,**
- **Web search,**
- **voice-based human computer interfaces,**
- **spoken dialogue systems,**
- **dialog management,**
- **applications and standards,**
- **multimodal processing,**
- **machine learning for spoken language processing,**
- **security, bioinformatic/genomic signals.**

In 2011, we are already far from the simple digital processing of speech. The spoken language is now analyzed in the context of NLP, emotion profile investigations, multimodality dialogues using even haptic devices, spoken dialog system evaluation etc. A very promising field of interest seems to be related with medical issues.

For the 7th edition (2013), we considered also some important tendencies: • speech recognition and synthesis, • spoken language understanding, • spoken dialog systems, • spoken language databases, • speaker /language recognition, • multimodal processing, • human /computer interaction, • educational and healthcare applications, • assistive technologies, • natural language processing.

The 8th Conference on Speech Technology and Human-Computer Dialogue "SpeD 2015" had several main topics:

- **Speech Analysis, Representation and Modeling. Spoken Language Recognition and Understanding.**
- **Speech and Text Analysis for Prevention and Mitigation of Emergencies and Acute Situations.**
- **Audio Analysis for Event Detection. Speaker Verification in Mobile Biometry.**
- **Natural Language Processing. Speech Analysis for Linguistics and Phonetics.**
- **Audio Signal Processing. Speech Analysis and Synthesis.**
- **Speech Databases, Technologies and Systems for Smart Homes and Assistive Applications.**

We hope that the 9th Conference on Speech Technology and Human-Computer Dialogue "SpeD 2017" will continue to be connected to the new tendencies but is also emphasizing the problems we still have to obtain better performances for the Romanian language.

The Conference includes six sections:

- ***Automatic Speech Recognition and Speech Databases.***
- ***Audio Signal Processing***
- ***Natural Language Processing***
- ***Speech analysis and applications***

A Poster section is also available.

There will e also to Project presentations:

- **Automatic Infant Crying Recognition System - "SPLANN".**
- **Romanian Language Phonetic Analysis: Study and applications - "AFLR".**

HiPEAC Vision 2017

HiPEAC's domain (High Performance and Embedded Architecture and Compilation) – computers, programmable systems, processors, microcontrollers etc – is evolving very rapidly and has a far-reaching impact on our everyday life. The aim of this 2017 HiPEAC Vision is to highlight some of the ongoing evolutions in this domain and to outline a number of recommendations to steer it. The main findings of the Vision can be summarized as follows:

The computer is disappearing from view, yet is becoming embedded in the very fabric of everyday life. Change can be seen at every level: not only in the content of our interactions with computers, which are evolving from answers to computational problems to interactions through social networks, but also in the forms in which we interact, which are shifting from letters and digits to sounds, gestures, images and movies.

HiPEAC Vision 2017

Yet we expect more change to come: programming will become learning for the machine, and interactions with computers will be augmented by virtual and augmented reality and modelled as interactions between humans. All this is made possible by the use of Artificial Intelligence-based techniques. We expect systems not only to observe, but also to interact with the physical world and to control it. The most visible developments by early 2017 are Intelligent Personal Assistants and Advanced Driver Assistance Systems; the latter is evolving into the autonomous driving car.

Voice controlled personal assistants

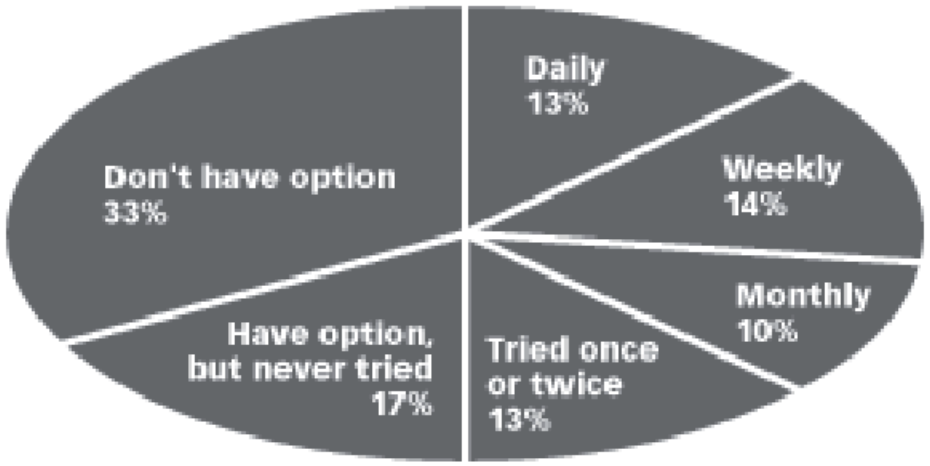
Intelligent Personal Assistants (IPA) such as Apple Siri, Google Now, Microsoft Cortana and Amazon Alexa (Echo) are now widely available on mobile devices, computers and custom devices. They use speech and natural language processing techniques, are connected to the cloud to offer different kinds of services, and use machine learning and neural network techniques to perform voice recognition. As a result, natural language might supplant the keyboard and touch screen as the new standard user interface to drive applications.

These services are provided for free on smartphones and computers. Amazon proposes a custom device, Echo, which does not have other interfaces besides an array of microphones and a speaker:
"Amazon Echo is a hands-free speaker you control with your voice. Echo connects to the Alexa Voice Service to play music, provide information, news, sports scores, weather, and more—instantly. All you have to do is to ask."

Google offers a similar device to the Amazon Echo with its 'Google home' device.

Frequency with Which US Mobile Phone Owners Use a Voice-Controlled Personal Assistant on Their Device, June 2015

% of respondents



Note: n=15,209

CAMPUS

“Research Center for Advanced Materials, Products and Processes”

- **about 18 million EUR investment;**
- **intelligent building;**
- **more that 8500 sm;**
- **41 research laboratories covering many domains (electronics, telecommunications, new materials, energy managements, artificial intelligence, aeronautics);**
- **the first trans-faculties research center gathering teams from many departments.**

Acknowledgements

The organizers of this event are honored to consider this edition of "SpeD" Conference as a tribute to the memory of Acad. Mihai Drăgănescu who was our mentor for the functional electronics domain.

We want also to pay a tribute to the memory of Prof. Eugeniu Oancea, one if the pioneer in the domain.

We are also grateful to the panel of referees for their effort in revising the papers. We thank the members of the Scientific Committee and the members of the Organizing Committee. I must also mention the helpful collaboration with the members of the Local Organizing Committee, especially my students from the Faculty of Electronics, Telecommunications and Information Technology.

The "SpeD 2015" Organizing Committee acknowledges the sponsoring of the Conference by the following companies: Infineon Technologies Romania, Softwin Group, Microchip Romania, but also University "Politehnica" of Bucharest, and "Electronica 2000" Foundation.

I want to express my gratitude to Professor Corneliu Rusu, Co - chair, Technical University of Cluj-Napoca, for his very important involvement to keep a high scientific level and to organize the Conference.

I am also especially grateful to Professor Horia-Nicolai Teodorescu, Co – chair, c.m. of the Romanian Academy, Technical University of Iaşi, for his valuable effort to be co-editor of our Conference Proceedings and in all aspects of the Conference.