# WELCOME

to

The 7th International Conference

on

# Speech Technology

# and

# Human-Computer Dialogue

## "SpeD 2013"

Cluj-Napoca, October 16-19, 2013

**The International Conference on Speech Technology and Human-Computer Dialogue ("SpeD") is now at the 7th 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).**

**The research in the following areas seemed to be strongly encouraged:**

- **Spoken language understanding**
- **Dialog management**
- **Spoken language generation**
- **Spoken document retrieval**
- **Information extraction from speech**
- **Question answering from speech**
- **Spoken document summarization**
- **Machine translation of spoken language**
- **Speech data mining and search**
- **Voice-based human computer interfaces**
- **Spoken dialog systems, applications and standards**
- **Multimodal processing, systems and standards**
- **Machine learning for spoken language processing**
- **Speech and language processing in the world wide web.**

**Another interesting dimension of the domain: "Multilingual Language Processing".**

**The trend for multilingual applications appeared to be most probably in spoken language understanding and generation, in information retrieval (legal/medical/education), in Internet businesses, search, spam filters, user generated content. "With more than 6.900 languages in the world and the current trend of globalization, one of the most important challenges in spoken language technologies today is 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.**

**The spoken document search over the Web** was and still is an important field of interest. Web statistics show the huge increase in domain names, archived Web pages, video streams, blogs, Internet users, mobile Internet users, broadband subscribers, instant messaging users, voice over IP subscribers, blog readers, searches per month.

**Spoken language translation** was and still is an important target, both from a social and economic point of view. This task implies achievements in automatic speech recognition (ASR) but also in machine translation (MT) and in the field of the human-computer dialogue management.

**In 2011, a new milestone could be the IEEE Workshop on Spoken Language Technology "SLT 2010", December 12-15, 2010, Berkeley, California.**

**The areas of interest for the conference paper submission were:**

- **spoken language understanding,**
- **speech data mining,**
- **spoken document summarization,**
- **machine translation for speech,**
- **spoken language based systems,**
- **spoken language generation,**
- **question answering from speech,**
- **human-computer interaction,**
- **spoken information extraction,**
- **spoken document retrieval,**
- **multimodal processing,**
- **spoken dialog systems,**
- **spoken language systems,**
- **spoken language databases,**
- **natural language processing.**

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 were related with medical issues:

- "Dialogue Strategy Optimization to Assist User's Decision for Spoken Consulting Dialogue Systems",
- "Automatic Identification of Qualitative Characteristics in Infant Cry",
- "Improvement of a Speech Recognizer for Standardized Medical Assessment of Children's Speech by Integration of Prior Knowledge",
- "Towards Accurate Recognition for Children's Oral Reading Fluency".

**For the 7th edition, we may consider also some tendencies pointed out at the IEEE Workshop on Spoken Language Technology "SLT 2012", December 2-5, 2012, Miami, Florida, USA:**

- **speech recognition and synthesis,**
- **spoken language understanding,**
- **spoken dialog systems,**
- **spoken document summarization,**
- **machine translation for speech,**
- **question answering from speech,**
- **speech data mining,**
- **spoken document retrieval,**
- **spoken language databases,**
- **speaker /language recognition,**
- **multimodal processing,**
- **human /computer interaction,**
- **educational and healthcare applications,**
- **assistive technologies,**
- **natural language processing.**

**Some of the invited papers are, in our opinion, relevant for the domain evolution:**

- **"The Conversational Web",**
- **"Structure Transformation for Machine Translation: Strings, Trees, and Graphs",**
- **"Language as influence",**
- **"Deep Neural Networks for Large Vocabulary Speech Recognition".**

**Very interesting debates were organized on some hot topics:**

- **"The Future of SLT Academic Research",**
- **"Beyond Siri: Next-Generation Information Services".**

**We are now in a new era of spoken language technology mobile applications: Apple's Siri personal assistant on the iPhone 4S last year became a new milestone.**

**In March 2013, the Forum for Europe's Language Technology Industry published the report on the "Status and Potential of the European Language Technology Markets".**

**One of the most important directions pointed out by this Forum was "*Where language is the very stuff of our digital system – customer interactions, employee conversations, technical and scientific knowledge, cultural and social objects of all kinds – the era of the Lingua Franca is over. Interacting across the many languages of the digital world is no longer optional.*"**

**So, the "multilingual language processing" is and will be an important issue.**

**Some major markets for Speech are identified in the cited report:**

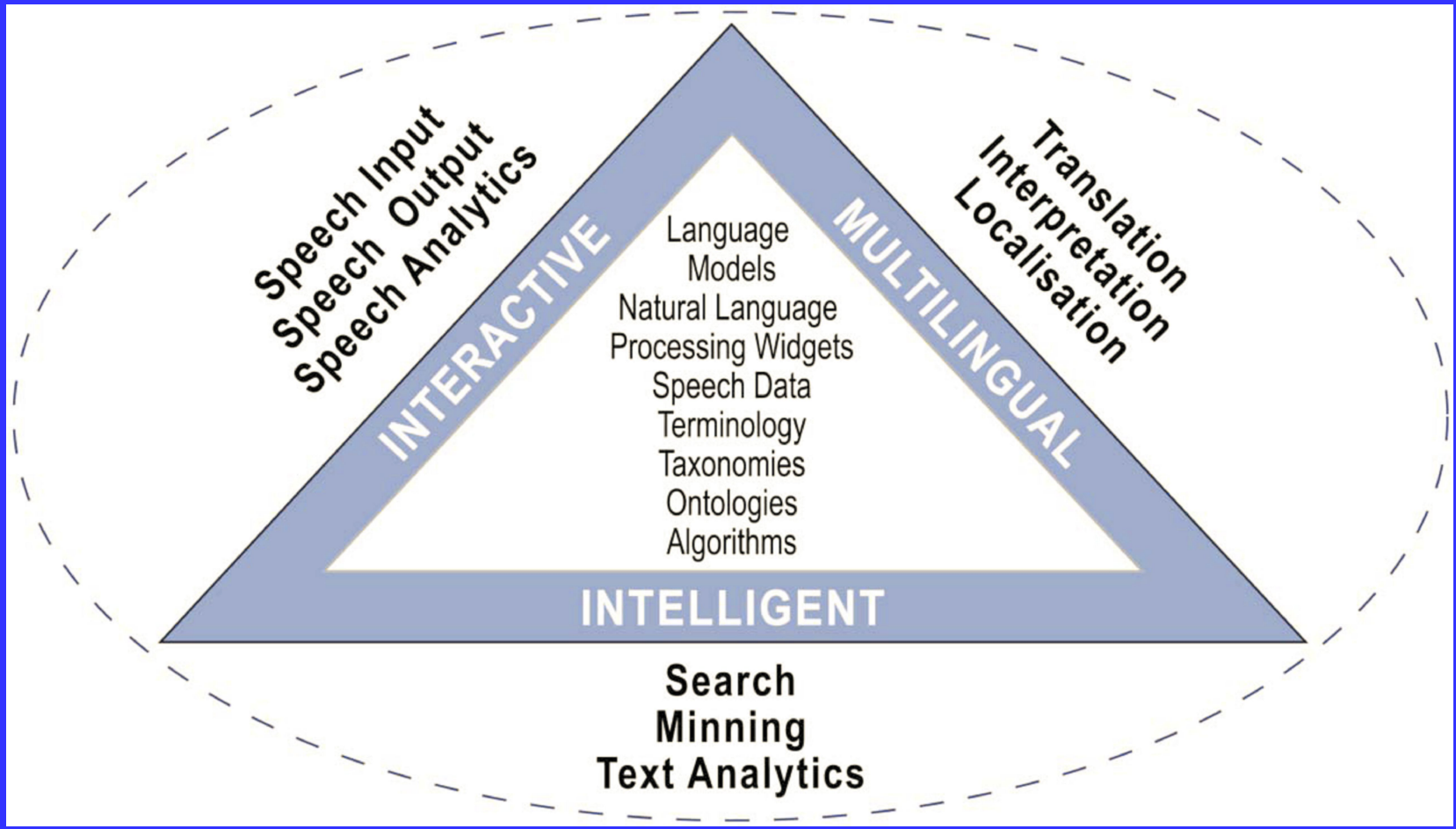
- **Call Centre is a core global market.**
- **Medical reporting and transcription is growing (especially in the USA for compliance with new Electronic Health Records regulations).**
- **Large and stable government customer base (including specialized defense applications).**
- **Speedy growth in consumer markets on devices and social platforms.**

**Some foremost gaps and opportunities for speech technology may be emphasized:**

- **Language coverage, need to develop a competitive market for languages beyond English.**
- **Asian markets, high digital growth and low penetration outside Japan.**
- **Speech data for developers, competitive choices on open platforms.**

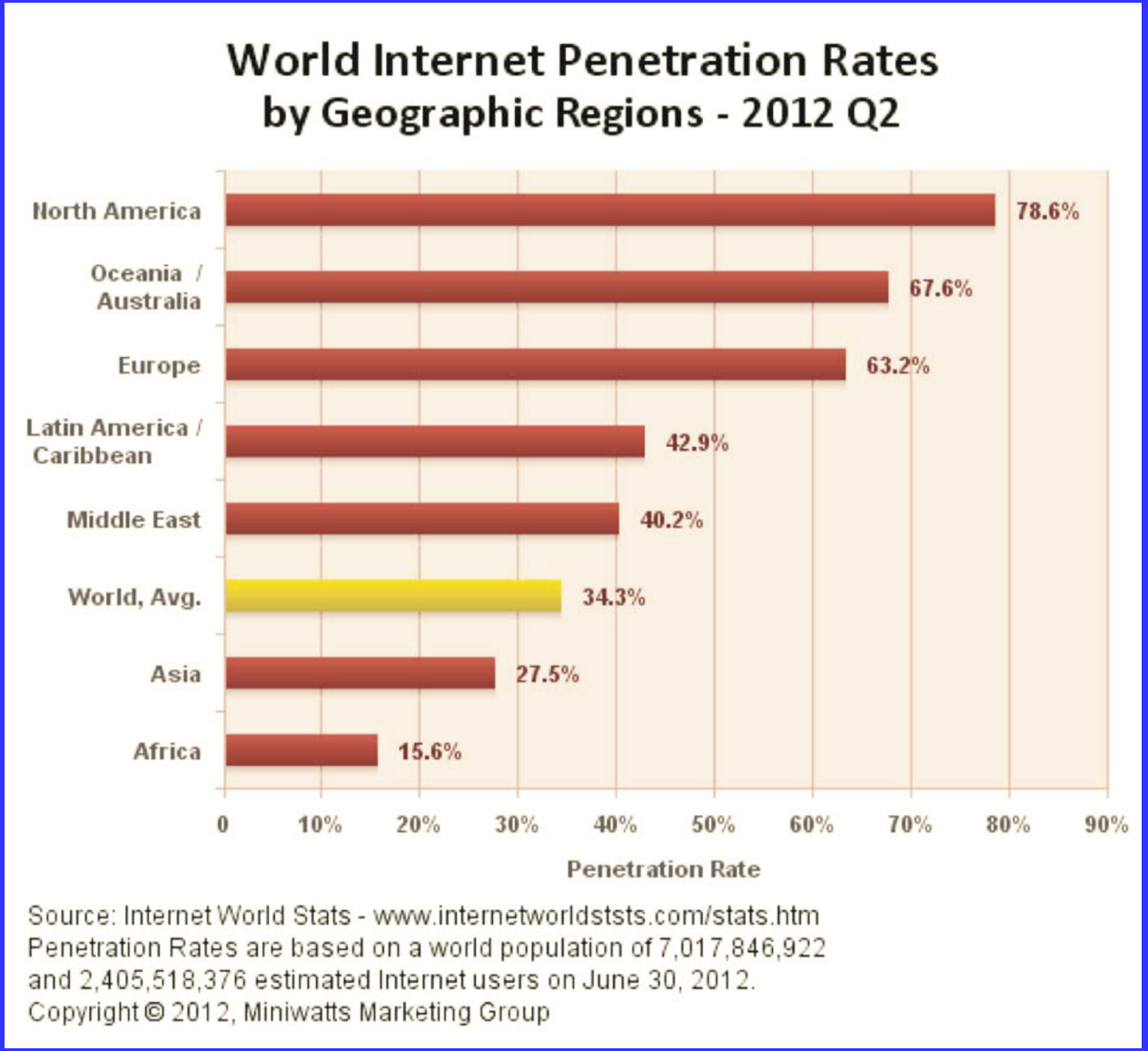
**Milestones in the history of Speech Technology:**

- 1952 Creation of a small vocabulary recognizer for digits over the phone, by Bell Labs.
- 1962 IBM developed Shoebox – a 16‑word speech 16‑word speech recognizer, used to interface with a calculator.
- 1972 Introduction of stochastic processing with Hidden Markov models to speech recognition.
- Mid‑1970s Creation of a small vocabulary recognizer for hands‑free industrial applications, by NEC and Threshold Technology.
- Late‑1970s Verbex launches a speech application based on a small vocabulary recognizer, which is useful for telephone toll management and financial services.
- 1990 Introduction of a general‑purpose dictation application, by Dragon Systems.
- 1992 AT&T unveiled automated operator system, which is capable to understand spoken utterances that include ‘operator’ and ‘collect call’.
- Late‑1990s Large‑scale deployment of commercial speech recognition solutions.



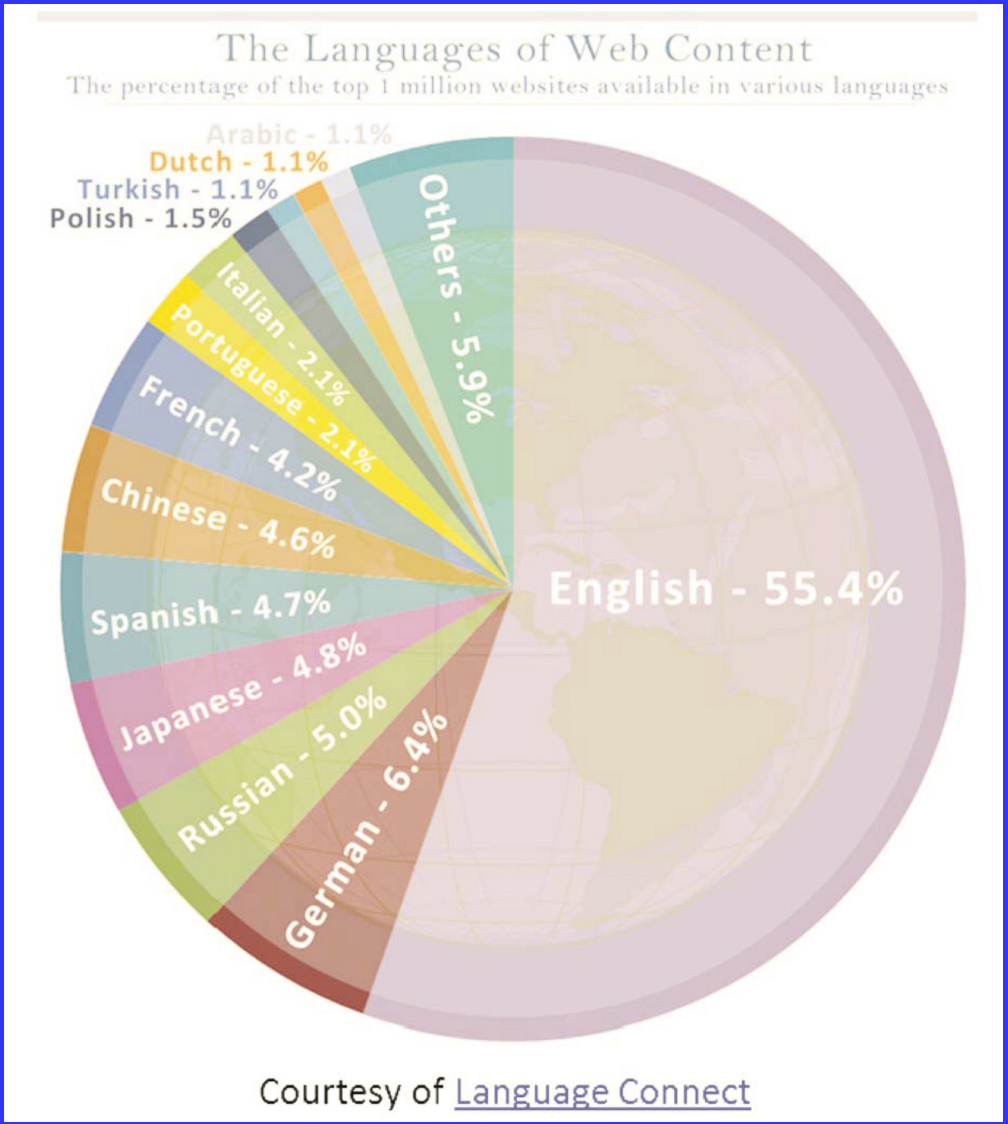

**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

**World Internet Penetration Rates (2012 Q2)**



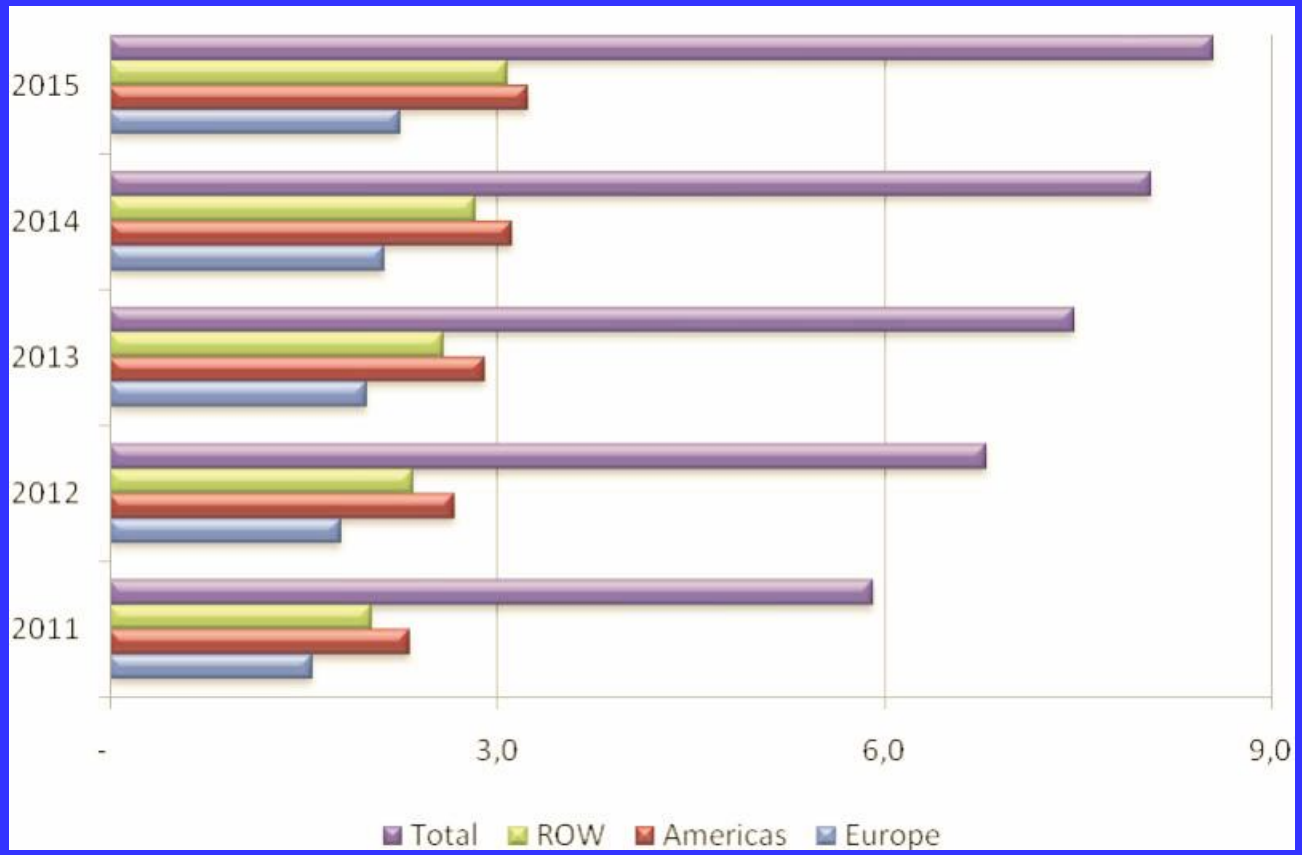

**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

## The Languages of the Web Content




**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

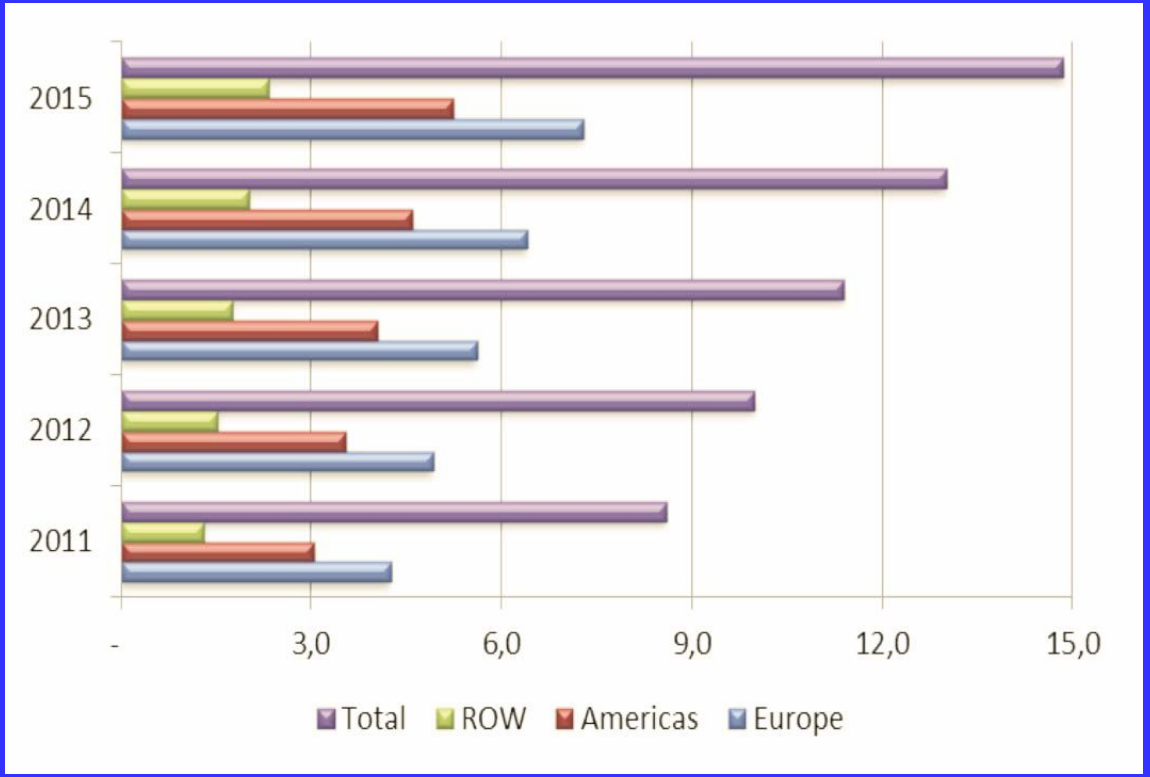
## Worldwide Speech Technology Software & Services Market 2011 - 2015 (€B)




**ROW: "Rest of the World"**

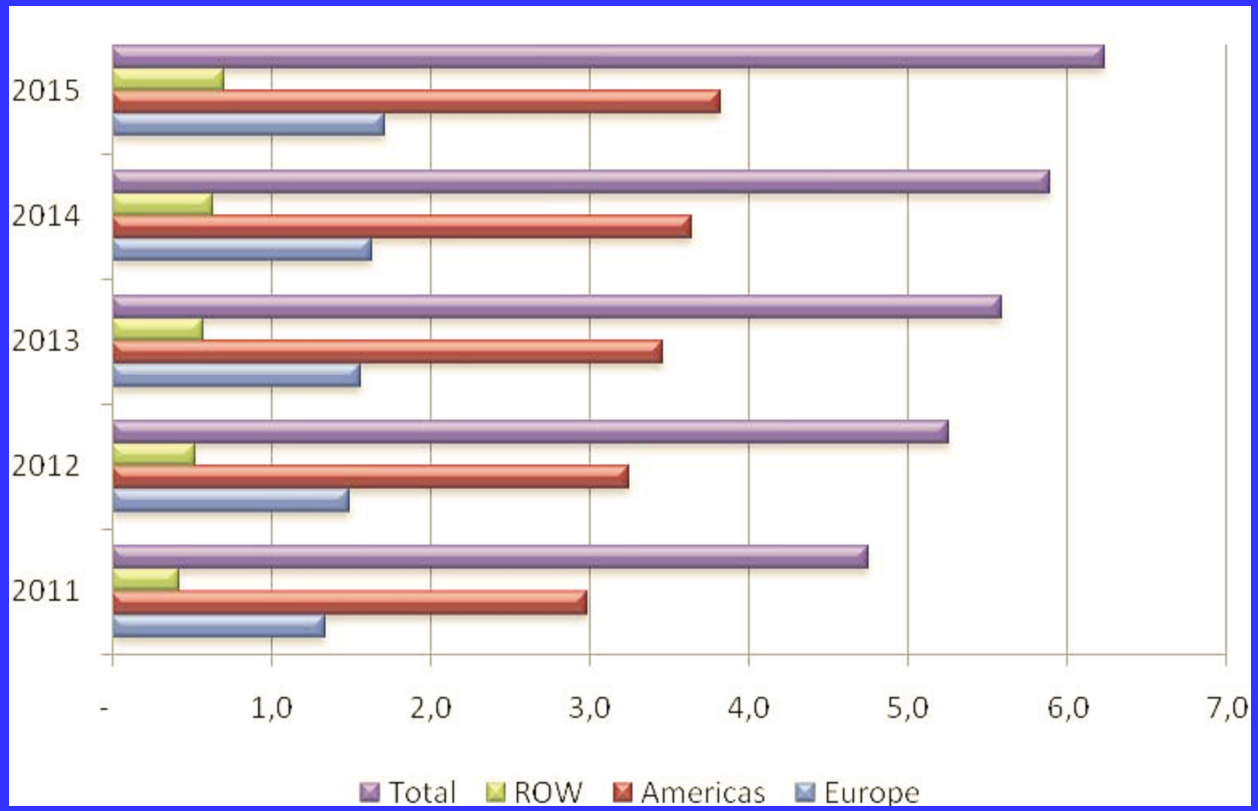
**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

# Worldwide Translation Technology Software & Services Market 2011 - 2015 (€B)



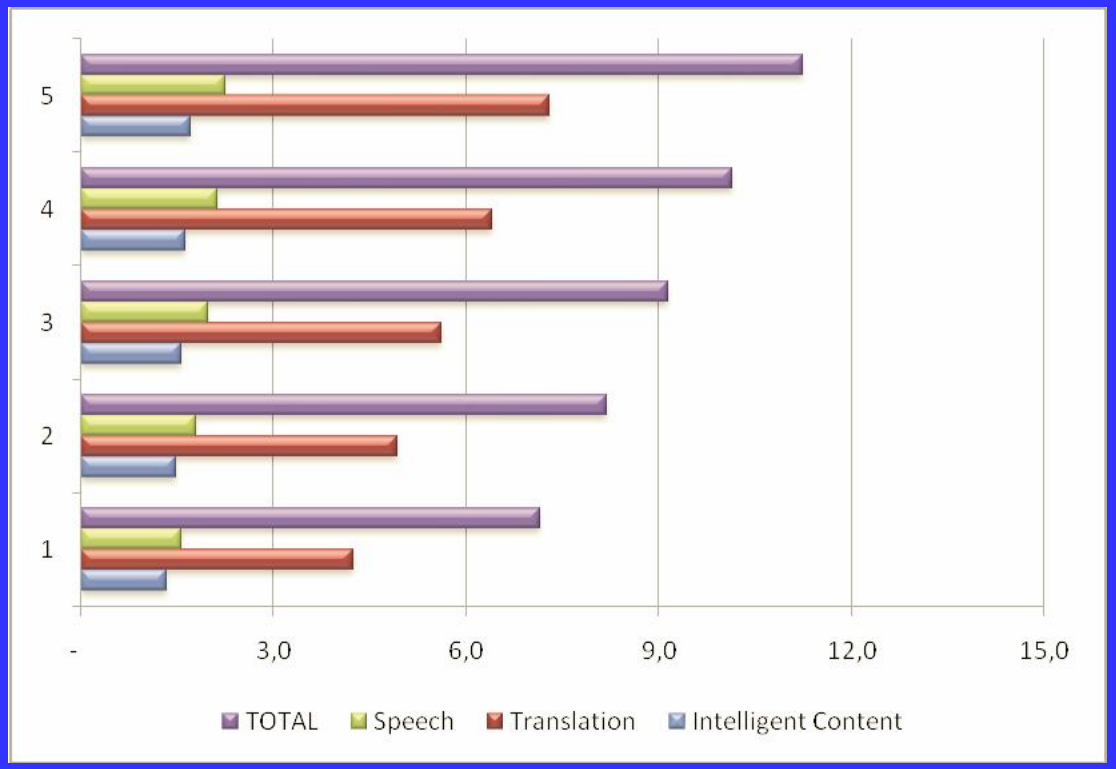

**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

## Worldwide Intelligent Content Market 2011 - 2015 (€B)




**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

# European Language Technology Market 2011 - 2015 (€B)




**"Status and Potential of the European Language Technology Markets"**
**Forum for Europe's Language Technology Industry, March 2013**

**Global players:**

**Speech**

- **Microsoft – speech embedded in its software platforms.**
- **Google – speech‐enabled search.**
- **Nuance – Enterprise and packaged solutions provider in the US and Europe.**
- **iFlytek – 5,000 partners, owns 70% of the Chinese‐language speech recognition market.**

**Leading technology suppliers in the global market:**

**Translation Technology**

- **Google – free online translation and API for developers.**
- **Microsoft – free online translation and API for developers.**
- **Youdau – free translation in Chinese search engine.**

**We hope that the 7th Conference on Speech Technology and Human-Computer Dialogue "SpeD 2013" 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 four sections:**

- ***Speech Analysis, Representations and Models. Audio Signal Processing,***
- ***Spoken Language Recognition. Keyword Spotting and Information Retrieval. Human-Computer Interfaces,***
- ***Natural Language Processing. Speech Translation,***
- ***Text-to-Speech Synthesis. Speaker Recognition. Assistive Technologies,***

**A Poster section is also available.**

## Acknowledgements

**The organizers of this event are honored to consider the "SpeD 2013" Conference as a humble tribute to the memory of Acad. Mihai Drăgănescu who was our mentor for the functional electronics 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.**

**The "SpeD 2013" Organizing Committee acknowledges the sponsoring of the Conference by the following companies: Infineon Technologies Romania, Softwin Group, Intel Romania, Microchip Romania, but also University "Politehnica" of Bucharest, Technical University of Cluj-Napoca 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, Vice-Rector of the Technical University of Iaşi, for his valuable effort to be co-editor of our Conference Proceedings and in all aspects of the Conference.**