

Development of speech technologies for the South Slavic languages

AlfaNum team

AlfaNum Ltd. was founded in 2003, originating from an existing project at the Faculty of Technical Sciences (FTN), Novi Sad, related to speech technologies. The co-founders of AlfaNum are FTN project leader PhD Vlado Deli and key researcher Darko Pekar as CEO. This team of remarkable experts has been carefully expanded for 15 years and now has 20 people. AlfaNum still has tight cooperation with FTN, co-finances the development at FTN and engages associates employed at FTN. To date, AlfaNum has developed software components for speech recognition and synthesis in Serbian, Croatian, Macedonian and Hebrew. In the field of speech technologies AlfaNum is practically the only player in the region of western Balkans.

Speech technologies, automatic speech recognition (ASR) and text-to-speech (TTS) synthesis have the potential of introducing essential changes to human interaction with their environment. If a person is able to speak to a computer, and if the computer is able to speak back, then similar communication with other devices can be realized, starting from home equipment through industrial machines, cars, robots and toys, all the way to remote computers which can retrieve the required information and deliver it by speech.

The development of speech technologies is especially significant for the disabled:

- computers can read books, news from the Internet, e-mail and SMS messages to the visually impaired,
- computers can convert what a speech impaired person writes into speech,
- the physically disabled can communicate with the devices in their environment,
- automatically recognized speech is easily translated into text which can then be read by people with hearing disabilities.

Speech technologies help the disabled to overcome their disabilities to a certain degree, enabling them to become more independent and perform tasks they were not able to do before. The first application in Serbian-speaking community was created for the visually impaired, for whom a number of aids have been developed. That is our speech synthesizer, and Reader which was declared as the official aid for the blind and visually impaired people by the Republic Fund of Health Insurance. A Reader has given our visually impaired citizens the possibility to use computers unaided, as many of their counterparts in more developed countries already do. It is clear that, with the development of speech technologies, conditions are being met for development of resources that can redefine some aspects of social policy related to the disabled.

Besides a Reader there are other applications for the disabled based on speech technologies, like speech portals, Audio library, adding speech features to web sites or subtitling TV shows. These applications are intended for the blind but also for the people with other disabilities like people with multiple sclerosis, people with infantile cerebral palsy, dyslexic people etc.

On the other side AlfaNum also develops call centres, interactive voice response (IVR) systems and other speech-enabled solutions. It offers both speech-enabled modules to be embedded into other systems, as well as end solutions. All applications of speech technologies in the western Balkans, related to Serbian and kindred languages, are the result of AlfaNum. Another significant competitive advantage

of AlfaNum lies in the existing language and speech resources, whose development required years of work.

The barriers and drivers during the development of ASR and TTS were very complex multidisciplinary problems, and their successful treatment required not only technical skills, but also expert knowledge in areas such as linguistics, psychoacoustics and speech perception, acoustics and digital signal processing. All these elements had to be combined and implemented into available hardware and software resources in order for a computer to be enabled to understand human utterances and to generate speech of its own.

The main objective of AlfaNum team is development of speech technologies that will be used in different spheres of social activity, and that will ultimately enable intuitive dialogue between man and machine in the Serbian and other South Slavic languages.

Contact:

jovana.vukov@alfanum.co.rs