
Source:	Siemens AG¹
Title:	Comments on the time scale of proposed WI “Speech Enabled Services Based on Distributed Speech Recognition (DSR)”
Document for:	Discussion
Agenda Item:	7.1

SA1 proposes a new work item called “Speech Enabled Services Based on Distributed Speech Recognition (DSR)” intended for Release 5. The goal of the WI is to define the necessary components for speech enabled services based on Speech Recognition, for example automatic speech access to information. Furthermore, it aims to identify the necessary changes and additions required in the current SA1 specifications. An invitation was sent to SA2, SA4, TSG-CN and T2 to participate in the work.

We appreciate the proposal towards the work on speech enabled services as part of the 3G specifications. With this contribution attention shall be drawn to the fact that as a Rel-5 work item two months (in case of December 2001 being Rel-5 issuing date) or five months only (in case of March 2002 being Rel-5 issuing date) are left for the specification of an efficient speech enabled service solution.

The ETSI STQ-Aurora DSR working group (Aurora) is developing a new DSR front-end that is planned to be decided for publication in Q2 2002.

Given the Rel-5 time frame, only the adoption of the new ETSI Aurora standard “as is” for a distributed speech recognition service seems to be possible. However, due to the planned publication of the ETSI Aurora Standard in 2002 this approach does neither guarantee timely adoption of the ETSI standard, nor the consistency with existing 3GPP specifications.

Speech recognition for speech enabled services might work also based on coded (e.g. AMR) speech with the full front-end and back-end part of the speech recogniser located in the network. In response to a UMTS Forum request from TSG-SA#12 (expressed in Tdoc SP-010294) SA4 states that (see Tdoc SA4 (01)0538):

“SA4 has not carried out any studies on the compatibility of speech codecs with voice recognition systems. There is no specific information available within SA4 which confirms or casts doubt on their suitability for use with network based voice recognition.”

i.e. a proof-of-concept still needs to be produced in order to ensure quality and usability of the speech enabled service.

As a consequence, we feel that the time scale of the WI should therefore be broadened, in order to ensure the specification of a consistent, correct and complete speech enabled service within 3GPP. For this purpose, it seems more feasible and relevant to open the respective WI for the time frame beyond Rel-5.

¹ Dirk Didascalou, Siemens AG
Dirk.Didascalou@mch.siemens.de
Phone: +49 89 722 58574