



OSAC Research Needs Assessment Form

Title of research need: DATASETS: For community research activities*

Keywords: forensic speaker recognition data collection

Submitting subcommittee(s): Speaker Recognition

Date Approved: 2/1/16

(If SAC review identifies subcommittees, add them to the box above.)

additional

Background information:

1. Description of research need:

Automatic speaker recognition (ASR) technology is developed and evaluated on available data to improve and measure performance. As ASR matures, researchers have realized that the data available for development is not representative of the data received by forensic and investigative examiners. We therefore propose the collection of realistic datasets. This realistic data is the key enabler for all the work of OSAC-Speaker and is the top priority!

2. Key bibliographic references relating to this research need:

- [1] Cieri, C., Andrews, W., Campbell, J., et al., The Mixer and Transcript Reading Corpora: Resources for Multilingual, Crosschannel Speaker Recognition Research, International Conference on Language Resources and Evaluation (LREC), ELRA, Genoa, Italy, pp. 117-120, 22-28, May 2006.
- [2] Campbell, J. P., Reynolds, D. A., Corpora for the Evaluation of Speaker Recognition Systems. In Proc. International Conference on Acoustics, Speech, and Signal Processing in Phoenix, Arizona, IEEE, pp. 829-832, 15-19 May 1999

3a. In what ways would the research results improve current laboratory capabilities?

Evaluating the existing ASR methods using data more representative of what the labs receive for analysis will (1) help labs better understand the capabilities and limitations of the ASR tools they use and (2) help identify capability gaps that research can be focused on. The requirement for multiple datasets arises out of the wide variation in the types of data received by the various users of ASR in the United States.

3b. In what ways would the research results improve understanding of the scientific basis for the subcommittee(s)?

The more realistic testing of the existing tools will lead to additional detailed studies of the phenomena causing the existing ASR methods to produce erroneous results. The results of those studies can be expected to improve not only our understanding of ASR, but also of other areas of speech processing and forensics.

3c. In what ways would the research results improve services to the criminal justice system?

A better understanding of our current ASR capabilities will help ensure that investigators, judges and juries are getting the best possible information when ASR is involved. The expected improvement of those capabilities by developments based on those research results will also allow the more effective use of ASR by the justice system and other USG users.

4. Status assessment (I, II, III, or IV):

I

	Major gap in current knowledge	Minor gap in current knowledge
No or limited current research is being conducted	I	III
Existing current research is being conducted	II	IV

This research need has been identified by one or more subcommittees of OSAC and is being provided as an informational resource to the community.

The RDT&E Task Group understands that there is a critical need to compile a set of corpora that supports the evaluation of systems on data that closely resembles the conditions seen by forensic speech examiners. Therefore:

* We endorse the initiative proposed by Reva Schwartz (NIST, 301-975-2175) to fund a data collection that produces data that is realistic and is able to support research and evaluation of automatic systems for the foreseeable future. The need for realistic data is likely the main enabler for follow-on activities critical to this OSAC committee. These activities include validation of methods, evaluation of current and new methods across the spectrum and research into new forms of conducting the identification.

Subcommittee

Approval date: 7/1/16

(Approval is by majority vote of subcommittee. Once approved, forward to SAC.)

SAC

1. Does the SAC agree with the research need? Yes No

2. Does the SAC agree with the status assessment? Yes No

If no, what is the status assessment of the SAC:

Approval date: 6/15/16

(Approval is by majority vote of SAC. Once approved, forward to NIST for posting.)