



Priority Action Report

Speaker Recognition

Digital Multimedia SAC

Dr. Joseph P. Campbell, Chair

Prof. John H.L. Hansen, Vice Chair and Presenter

13 Feb 2017





Subcommittee Leadership

Position	Name	Organization	Term	Email
Chair	Joseph P. Campbell	MIT Lincoln Laboratory		jpc@ll.mit.edu
Vice Chair	John H.L. Hansen	University of Texas at Dallas		john.hansen@utdallas.edu
Executive Secretary	Patrick E. Gibbs	Leidos		Patrick.E.Gibbs@leidos.com





#	Name	Organization	Term
1	Aaron Lawson	SRI	2
2	Alysha Jeans	FBI	3
3	Chris Cieri	Linguistic Data Consortium, U Penn	4
4	David Farris	U.S. Govt	2
5	David Marks	Sandia National Laboratories	3
6	Dick Reeve (Legal Resource Committee)*	Denver District Attorney's Office	3
7	Dorothy Glancy	Santa Clara University School of Law	4
8	Douglas Reynolds	MIT Lincoln Laboratory	2
9	Jack Godfrey	JHU	2
10	Jim Wayman (DMSAC)*	San Jose State University	3
11	Joe Campbell (Chair + DMSAC)*	MIT Lincoln Laboratory	2
12	John Hansen (Vice Chair)	University of Texas at Dallas	3
13	Ken Marr	FBI	3
14	Kevin Farrell	Nuance Communications	3
15	Larry Solan (DMSAC)*	Brooklyn Law School	3
16	Oscar Morales	U.S. Govt	3
17	Patrick Gibbs	Leidos	3
18	Pedro Torres-Carrasquillo	MIT Lincoln Laboratory	4
19	Ray Slyh	Air Force Research Laboratory	4
20	Stephen Gibbs	U.S. Govt	2
21	Walter Andrews	Sierra Nevada Corporation	3

* Does not count against OSAC-SR's Membership quota of 20





#	Name	Organization	Term
1	Adam Dale	BAH	
2	Alvin Martin	NIST (Ret)	
3	Anil Alexander	Oxford Wave Research, Oxford	
4	Daniel Garcia-Romero	JHU	
5	Darren Haddad	Air Force Research Laboratory	
6	Doug Smith	US Army	
7	Finnian Kelly	UT Dallas; Oxford Wave Research, Oxford	
8	Geoffrey Stewart Morrison	Independent Forensic Consultant, Canada	
9	Hiro Nakasone	FBI (Ret, First Chair)	
10	Jim Loudermilk	FBI	
11	Jean-François Bonastre	Université d'Avignon, France	
12	Jennifer Stathakis	FBI	
13	Jim Horan	NIST	
14	Jodie Linger	FBI	
15	Johanna Morley	Met Police, London	
16	Jos Bouten	Ministry of the Interior, The Hague	
17	Kathleen Egan	CTTSO (Ret)	
18	Michael Jessen	Bundeskriminalamt (BKA), Germany	
19	Mitchell McLaren	SRI, Australia	
20	Niko Brümmer	Agnitio, South Africa	
21	Phil Rose	ANU, Australia (Ret)	
22	Reva Schwartz	NIST	
23	Roger Campbell/Dave Gebhard	FBI	
24	Vince Stanford	NIST	





Members, Affiliates, Representatives and Guests

- **OSAC Speaker Recognition Subcommittee** currently is
 - 17 Members
 - 3 **Digital/Multimedia Scientific Area Committee** representatives
 - 1 **Legal Resource Committee** representative
 - 24 Affiliates (including a few Invited Guests for each meeting)
- Hailing from the U.S. Government, U.S. National Labs, academia, industry, and international partner organizations
- International guests are invited from organizations in a variety of countries including Canada, France, Germany, The Netherlands, South Africa, and the United Kingdom
- Interested and qualified individuals may **apply to the OSAC** for full membership or to become affiliates in designated subcommittees





Task Groups and Ad Hoc Committees

Task Groups

- **Research, Development, Test & Engineering (RDT&E), Chairs**
 - Dr. Pedro Torres-Carrasquillo
 - Mr. Stephen Gibbs
- **Best Practices (BP), Chairs**
 - Mr. Kenneth Marr
 - Dr. Raymond Slyh
- **Legal Aspect of Speaker Recognition (LASR), Chair**
 - Prof. Dorothy Glancy
- **Proposed Task Group**
 - Verification, Validation & Evaluation

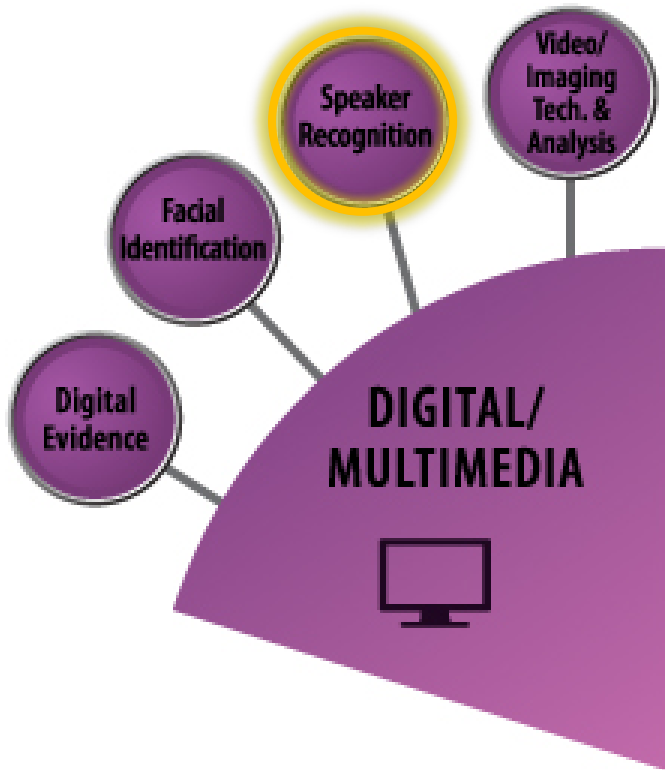
Ad Hoc Committees

- **Membership**
- **Vocabulary I**
- **Vocabulary II**





Discipline Description



“Speaker recognition is the process of comparing human speech samples to determine if they were produced by the same speaker.”

– Adopted at the OSAC Kick-Off Meeting, Norman, OK, 20-22 Jan 2015





Subcommittee Purpose and Objectives

Purpose: *To support and promote the scientific foundations and practice of speaker recognition, voice data collection, measurement, transmission, and retrieval, through the development and dissemination of consensus-based standards, guidelines, best practices, and recommendations for forensic and investigatory applications*

Objectives: Speaker Recognition Subcommittee shall engage in:

- Identifying the *scope, theories, and practice areas* of the disciplines of voice biometrics, speaker identification, voice data collection, measurement, transmission, and retrieval
- Recommending *guidelines* for conducting speaker recognition examinations, including data preparation, data collection, quality assessment, compression, decompression, transmission protocols, results, reports, and terminology
- Proposing standard procedures for data interpretation and wording of *conclusions* presented in court
- Suggesting *education, training, and continuing education* requirements for practitioners
- Promulgating and disseminating *research and development priorities* to the community
- Collecting and distributing discipline-specific information on *scientific foundations*
- Seeking *international recognition* and harmonization of appropriate work products





Summary of Priority Projects

Priority	OSAC Process	Working Title of Document
1	SDO	“Forensic Speaker Recognition Should-Be Process Map,” Best Practices, 2017
2	RA or SDO, parent ANSI	“Guidelines for Electronic Transmission of Speech Files,” Peter H. & Best Practices, 2016 (with DMSAC)
3	RA or SDO	“Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location,” RDT&E, Kavi, 1/1/17
4	RA (reference)	“Concise Glossary of Speaker Recognition Terms,” Ad Hoc Vocabulary Committee and Best Practices
5	n/a (reference)	“Annotated List of Important Speaker Recognition Cases,” LASR





Standards/Guidelines Development Priority 1 Document

Document Title: “Forensic Speaker Recognition As-Is Process Map”

Scope: Map of as-is practices for forensic speaker recognition

Objective/rationale: Create forensic speaker recognition maps to structure our activities. Develop a should-be process map to inform our guidelines and standards development.

Issues/Concerns: Proper evaluation of paths through process map using realistic data and use cases with relevant subsystems

Task Group Name: Best Practices

Task Group Chair Name: Kenneth Marr and Raymond Slyh

Task Group Chair Contact Information:

Kenneth.Marr@ic.fbi.gov and raymond.slyh@us.af.mil

Date of Last Task Group Meeting: 2017





Standards/Guidelines Development Priority 1 Document

Key Components of Standard:

- **Building blocks representing atomic units of speaker recognition processes with descriptions of each block's inputs, outputs, parameters, states, and internal processes**
- **Groups of building blocks representing speaker recognition subsystems**
- **Path through the blocks representing an instance of an end-to-end speaker recognition process**
- **A single process map can support many instances**
- **Attempt to document all practices and note where individual labs differ in their processes**





Priority 1: “Forensic Speaker Recognition As-Is Process Map” Task Group/Subcommittee Action Plan

Planned Actions	OSAC Process Stage (e.g., SDO 100)	Assignee	Estimated Completion Date
Create high-level <u>as-is</u> Process Map (PM) containing main sections and block descriptions.		Best Practices	Begin 2/6/2017
Adding sub-task processes (such as “processing,” etc.) to the <u>as-is</u> PM	n/a (reference)	Best Practices	May 2017
Technical and research review of <u>as-is</u> document to identify tasks/processes that can move directly to the <u>should-be</u> map, extract tasks/processes that either don’t move on at all, or need research and/or development to move on to the what <u>should be</u> process map	n/a (reference)	Best Practices w/ NIST support	Fall 2017
Development of what <u>should-be</u> process map. The Process Stage for the Process map <u>as-is</u> is a reference only. The Process Stage for the Process map of “what should be” to become a standard	SDO	Best Practices w/ NIST support	Spring 2017





Standards/Guidelines Development Priority 2 Document

Document Title: “Guidelines for Electronic Transmission of Speech Files”

Scope: Guidelines for the use of ANSI/NIST-ITL Type 11 Voice Records to electronically transmit speech files

Objective/rationale: Guidelines and recommendations are needed for Type 11 Voice Records implementation to enable standardization of voice data transactions by using the “[ANSI/NIST-ITL 1-2011: Update 2015](#)”

Issues/Concerns: OSAC-SR approved the draft submitted in April 2016 and elevated to DMSAC, waiting for action

Task Group Name: Best Practices

Task Group Chair Name: Kenneth Marr and Raymond Slyh

Task Group Chair Contact Information:

Kenneth.Marr@ic.fbi.gov and raymond.slyh@us.af.mil

Date of Last Task Group Meeting: 2017





Standards/Guidelines Development Priority 2 Document

Key Components of Guideline/Standard:

- **History of the ANSI/NIST ITL Standard and its use in criminal justice systems and forensic case work**
- **Role of Type-11 voice Records in Speaker Recognition casework with an example in the case of the comparison of two voice recordings**
- **Design and implementation considerations to include data formats and the transmission of transactions**





Priority 2: “Guidelines for the use of ANSI/NIST-ITL Type 11 Voice Records”

Task Group/Subcommittee Action Plan

Planned Actions	OSAC Process Stage (e.g., SDO 100)	Assignee	Estimated Completion Date
A designated team within Best Practice Task Group will work with the Department of Defense, Defense Forensic and Biometric Agency and the FBI Criminal Justice Information Services Division to create 1 st draft	RA or SDO (parent is an ANSI/NIST-ITL standard)	Peter Higgins	6/1/2015
2 nd draft		Peter Higgins	9/30/2015
Final draft		Peter Higgins	2/1/2016
Approved draft by OSAC-SR		Peter Higgins	4/2016
Elevated to DMSAC	RA or SDO	Peter Higgins	2016





Standards/Guidelines Development Priority 3 Document

Document Title: “Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location”

Scope: This guideline is intended to be one of a series, and covers only one scenario - the collection of speech samples for speaker recognition at a temporary, non-laboratory location.

Objective/rationale: Instruct an interviewer, who need not be technical, to create recordings in the field that are likely to be appropriate for speaker recognition applications.

Issues/Concerns: Proper evaluation using realistic data and use cases with relevant subsystems

Task Group Name: RDT&E

Task Group Chair Name: Pedro Torres and Stephen Gibbs

Task Group Chair Contact Information:

ptorres@ll.mit.edu and segibbs.work@gmail.com

Date of Last Task Group Meeting: 2017





Standards/Guidelines Development Priority 3 Document

Key Components of Guideline/Standard:

- Introduction, scope, references, terms
- **Speech Collection Scenario: Audio Collection at a Temporary Location**
 - Collection Environment
 - Collection Equipment
 - Speech Collection
- **Appendices**
 - Possible Known-Text Phrases for the Subject
 - Examples of Images to Elicit Speech Samples





Priority 3: “Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location” Task Group/Subcommittee Action Plan



Planned Actions	OSAC Process Stage (e.g., SDO 100)	Assignee	Estimated Completion Date
“Audio Collection at a Temporary Location” draft document ready for Subcommittee approval	RA or SDO	RDT&E Task Group Co-Chairs	July 2016 OSAC meeting in Phoenix
Ballot on "Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location," Version 1.9, 31 July 2016	RA or SDO	RDT&E Task Group Co-Chairs	11 Nov 2016
Revise document with comments from ballot	RA or SDO	RDT&E Task Group Co-Chairs	1 Jan 2017
Accepted revised document version 1.10, 1 Jan 2017	RA or SDO	OSAC-SR	11 Jan 2017 VM#3
Elevate to DMSAC	RA or SDO	OSAC-SR Chair	Feb 2017





Standards/Guidelines Development Priority 4 Document

Document Title: “Concise Glossary of Speaker Recognition Terms”

Scope: Subcommittee terminology reference

Objective/rationale: Support the development of consistent guidelines and standards within the Subcommittee and improve intra and inter Subcommittee communications

Issues/Concerns: Focus on the terminology used within OSAC SR documents, resist grander and international scope

Task Group Name: Ad Hoc Vocabulary Committee

Task Group Chair Name: Joe Campbell and Chris Cieri

Task Group Chair Contact Information: jpc@ll.mit.edu

Date of Last Task Group Meeting: 2017





Standards/Guidelines Development Priority 4 Document

Key Components of Guideline:

- **Selection**

- Vocabulary from OSAC SR documents beginning with the process map
- Terms having definitions within the discipline that do not match their generic sense

- **Scope**

- Construct definitions as they apply to the (growing body) of OSAC SR documents
- Reference external definitions, uses only where necessary to avoid ambiguity

- **Target audience: educated readers who are no expert in speaker recognition, linguistics or standard processes**





Priority 4: “Concise Glossary of Speaker Recognition Terms” Task Group/Subcommittee Action Plan

Planned Actions	OSAC Process Stage (e.g., SDO 100)	Assignee	Estimated Completion Date
“Vocabulary Terms for Speaker Recognition”	RA	Vocabulary Ad Hoc Committee I	2016, but died in committee
Restart, “Concise Glossary of Speaker Recognition Terms”	RA	Joe Campbell Vocabulary Ad Hoc Committee II	26 July 2016, OSAC Phoenix
“Concise Glossary of Speaker Recognition Terms” draft 2	RA	Joe Campbell	July & Aug 2016
“Concise Glossary of Speaker Recognition Terms” draft 3	RA	Joe Campbell	14 Nov 2016
“Concise Glossary of Speaker Recognition Terms” draft 3	RA	Chris Cieri, begin committee review	11 Jan 2017





Standards/Guidelines Development Priority 5 Document

Document Title: “Annotated List of Important Speaker Recognition Cases”

Scope: Analysis of US federal and state reported cases in which Speaker Recognition plays a role. Identify relevant legal rulings with regard to forensic application of speaker recognition techniques in US courts. Annotate types of cases.

Objective/rationale: Provide a sense of trends regarding how the US legal system treats speaker recognition methods

Issues/Concerns: Far fewer cases than expected (precedents)

Task Group Name: Legal Aspect of Speaker Recognition

Task Group Chair Name: Prof. Dorothy Glancy

Task Group Chair Contact Information: dglancy@scu.edu

Date of Last Task Group Meeting: 2017





Standards/Guidelines Development Priority 5 Document

Key Components of Standard-Setting Reference:

- **Annotated list of reported court decisions and order in which speaker recognition issues arise**
- **Description of important court precedents regarding speaker recognition issues**
- **Assistance**
 - **Support other Speaker Recognition task groups developing standards and guidelines**
 - **Outreach to legal communities with regard to modern speaker recognition technologies**
 - **Coordination with OSAC Legal Resource Committee**





Priority 5: “Annotated List of Important Speaker Recognition Cases”

Task Group/Subcommittee Action Plan

Planned Actions	OSAC Process Stage (e.g., SDO 100)	Assignee	Estimated Completion Date
Task Group LASR will update the annotated list every 6 months	n/a (reference)	LASR Task Group Chair: Dorothy Glancy	Every 6 months (next meeting Feb 2017)



Summary of Priority Projects

Priority	OSAC Process	Working Title of Document
1	SDO	“Forensic Speaker Recognition Should-Be Process Map,” Best Practices, 2017
2	RA or SDO, parent ANSI	“Guidelines for Electronic Transmission of Speech Files,” Peter H. & Best Practices, 2016 (with DMSAC)
3	RA or SDO	“Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location,” RDT&E, Kavi, 1/1/17
4	RA (reference)	“Concise Glossary of Speaker Recognition Terms,” Ad Hoc Vocabulary Committee and Best Practices
5	n/a (reference)	“Annotated List of Important Speaker Recognition Cases,” LASR

Great progress on all topics, including nearing completion on 2 & 3!





Standards/Guidelines Reviewed For Technical Merit

Title	Developing Organization	Status*	OSAC Process Stage (e.g., RA 100)
“Forensic Speaker Recognition <u>Should-Be</u> Process Map,” Best Practices	TBD SDO	Go to DMSAC in 2018	SDO
“Guidelines for Electronic Transmission of Speech Files,” Peter H. & Best Practices, 2016	RA or TBD SDO (parent ANSI)	With DMSAC since mid 2016	RA or SDO
“Speech Collection Guideline for Speaker Recognition: Audio Collection at a Temporary Location,” RDT&E, Kavi, 1/1/17	RA or TBD SDO	Expected to go to DMSAC in Feb 2017	RA or SDO
“Concise Glossary of Speaker Recognition Terms,” Ad Hoc Vocabulary Committee	RA	Go to DMSAC in 2017	RA





Research & Development Needs Identified



- **Proper evaluation of speaker recognition methods**
- **Realistic data for use in evaluation (and development)**
 - **Data that addresses similarity and typicality and has relevant population statistics (or at least a start in this direction)**
 - **Development of a standard reference dataset for use by researchers and examiner community**
 - **Forensic casework involves variables not yet studied, for which we lack realistic data**
 - **Seek carefully controlled access to case data that cannot be realistically recreated in a lab (e.g., physical harm, terror)**
- **Realistic use cases be exercised in evaluations**
 - **With relevant systems/subsystems**
- **Verification/validation of the suitability of tools for a given case**
 - **Today's technologies are developed on data whose suitability must be considered in addition to the technology**





Additional Items of Interest



- **The community also needs**
 - Better performing systems
 - Knowledge about performance of different listener types
 - Comparison of performance between listeners and automated systems
- Please visit the [OSAC Speaker Recognition Subcommittee](#)



*OSAC
Kick-Off
Meeting,
Norman,
Oklahoma
20-22 Jan
2015*





Priority Action Report

Speaker Recognition

Digital Multimedia SAC

Dr. Joseph P. Campbell, Chair

Prof. John H.L. Hansen, Vice Chair and Presenter

13 Feb 2017



