This document is being made available at this stage of the process so that the forensic science community and interested stakeholders can be more fully aware of the efforts and work products of the Organization of Scientific Area Committees for Forensic Science (OSAC). The documents were prepared with input from OSAC Legal Resource Committee, Quality Infrastructure Committee, and Human Factors Committees, as well as the relevant Scientific Area Committee. The content of the document listed below is subject to change during the standards development process and may not represent the contents of the final published standard. All stakeholder groups or individuals are strongly encouraged to submit technical comments on this draft document during the SDO's open comment period. Technical comments will not be accepted if submitted to the OSAC Scientific Area Committee or Subcommittees.

Guidelines for the Exchange of Speech Files for Use in Speaker Recognition Using ANSI/NIST-ITL Transactions

Prepared by

The Speaker Recognition Subcommittee

of the

Digital/Multimedia Committee

of the

Organization of Scientific Area Committees

Version 1.0

6 March 2016

Table of Contents

1 Introduction

Until the mid 1990s the exchange of fingerprint data between police departments and identification bureaus was through mailing paper files. As the FBI developed the Integrated Automated Fingerprint Identification System (IAFIS) the program office worked with the National Institute of Standards and Technology (NIST) to update the ANSI/NIST-ITL Fingerprint Identification – Data Format for Information Interchange (ANSI/NBS-ICST 1-1986 to support the submittal of tenprint and latent impression records digitally to IAFIS. By the spring of 2012 the FBI stopped accepting paper fingerprint submittals – only supporting the electronic submittal using the ANSI/NIST-ITL Standard (by then on the 2011 version) American National Standard for Information Systems—Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information (ANSI/NIST-ITL 1-2011).

The Scientific Working Group for Forensic and Investigative Speaker Recognition (SWG-Speaker) collaborated with the NIST/ITL manager of the ANSI/NIST-ITL Standard for the exchange of forensic information to create the ability to exchange digital speech files and the associated metadata. Jim Wayman was the editor of the effort, which culminated in a digital speech file exchange record being added as a Type-11 Record in ANSI/NIST-ITL 1-2011 Update: 2013.

To understand the purpose and advantages of the Type-11 record it is important to have some understanding of the context in which the ANSI/NIST-ITL Standard is used. The rest of this Introduction provides that context. You will note that the ANSI/NIST-ITL Standard and related documents have changed continuously since they were first introduced in 1986. Since the 1993 version of the Standard the changes have been basically backward compatible to the extent possible.

This Guideline is the first of a series that the Organization of Scientific Area Committees (OSAC); Digital-Multi Media Scientific Area Committee; Subcommittee on Speaker Recognition (SR) will publish. An overview of the role of this Guideline and its relationship to the anticipated additional Guidelines is explained in Section 2, Scope, below.

This Guideline is primarily oriented toward system developers (software designers and programmers) as well as operational personnel that develop procedures for the submittal of digital recordings of speech for processing in a laboratory setting or storage in a exemplar repository.

1.1 History of the ANSI/NIST-ITL Standard

The ANSI/NIST-ITL Standard was originally approved in 1986. It has evolved considerably since then: Updates, Supplements, and Corrections. The following history quoted directly from the NIST website provides the highlights. The details are provided in Appendix 1.

> *The first version of the standard, ANSI/NBS-ICST 1-1986 was published by NIST (formerly the National Bureau of Standards) in 1986. It was a minutiae-based standard. Revisions to the standard were made in 1993, 1997, 2000, and 2007. Updates to the standard are designed to be backward compatible, with new versions including additional biometric modalities and associated data. All of these versions use "Traditional" encoding.*

In 2008, XML encoding of the standard was introduced, based upon the 2007 version. The 2007 and 2008 versions of the standard were designed to be the same except for the encoding. The XML encoding was developed using the naming conventions of the National Information Exchange Model (NIEM). Thus, this encoding is referred to as "NIEM-conformant XML."

In 2009, a minor supplement to the 2007 and 2008 versions was approved that extended the codes for friction ridge images to include multiple finger capture.

In November 2011 a new version of the standard was approved. This version is focused on the content of the transmission, with encoding rules handled as annexes. Thus, there is no need for separate versions of the standard as was the case for 2007 (Traditional format) and 2008 (NIEM-conformant XML). The 2011 version of the standard also includes additional modalities (DNA and plantar) as well as the extended feature set (EFS) for Type-9 record; forensic image markups for face and iris; new metadata fields such as geoposition of sample collection; biometric data hashing; an information assurance record; associated context record; original source record and data handling logs; images of all body parts in the Type-10 record and more. Current plans are to extend the standard to include records types for voice recognition and dental forensics / bitemark analysis as well as an annex for conformity assessment.

In 2013, the standard was updated to ANSI/NIST-ITL 1-2011:Update 2013. *The Forensic Dental and the Forensic and Investigatory Voice Supplements were approved.*

1.2 History of the Implementation of the ANSI/NIST-ITL Standard

The ANSI/NIST-ITL Standard is flexible to permit implementation by agencies in different languages, appropriate character sets, and various physical metrics (e.g., Metric and UK foot and pound). The first such implementation was developed by the FBI's CJIS Division in 1994 – the Electronic Fingerprint Transmission Specification (EFTS). Over time other organizations have produced their own EFTS documents. When additional modalities were added – such as faces and irises – the title was morphed into Electronic Biometric Transmission Specification (EBTS).

Initially these community-of-interest based implementation documents were referred to as being *domain specific*. Over time the term was changed to *profiles* as in the DoD profile. In the United States there are many versions (profiles) of EFTS and EBTS specifications. Many states have their own versions – particularly so they can control the types of transactions (known as TOTs) and the related fields – such as state ID numbers and driver's license numbers. The U.S. Federal government has at least three separate EBTS documents – one for each domain:

- The FBI EBTS[1]
- The DoD EBTS (supports Type-11 records)
- The DHS IXM [1]

In addition there are international versions of what are commonly referred to as EBTS specifications. Two that are possible of interest are:

- NATO's STANAG [Standardization Agreement] 4715 (supports Type-11 records)
- Interpol's INT-I [1]

[1] Does not support Type-11 records as of the time of publication of this Guideline.

2 Scope

Documents that may be published by the OSAC SR Subcommittee include:

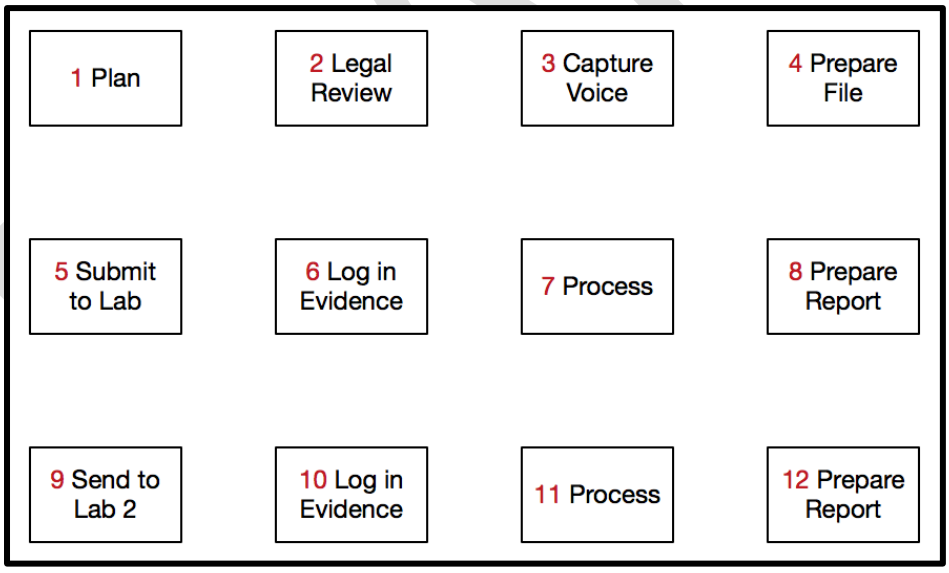
- Guidelines for the capture and recording of speech for use in speaker recognition
- Guidelines for the proper legal steps to ensure that recorded speech is admissible under the Daubert Standards
- Guidelines for the electronic transmission of speech files (this document)
- Guidelines for the processing of speech files in investigative and forensic settings
- Guidelines for the reporting of the results of the comparison of speech samples to one or more other speech samples
- …

2.1 Role in the Speaker Recognition Workflow

The speech capture through reporting the results of analysis and / or comparison to other speech samples can be shown to fit into the following 12 steps, as shown in Figure 1. Not all cases need to employ or even consider all 12 steps.

The primary assumption for the applicability of this Guideline is that the activities in Steps 5 and 9 involve the creation and transmission of an ANSI/NIST-ITL Transaction containing one or more Type-11 Record(s).

Figure 1 The Speaker Recognition Workflow



For instance, if a Lab (see Steps 5-8) determines that there is no need to forward the speech sample(s) to another lab then steps 9-12 are not necessary in that situation. The 12 steps are shown for completeness and will be addressed in other portions of this Guideline, as appropriate.

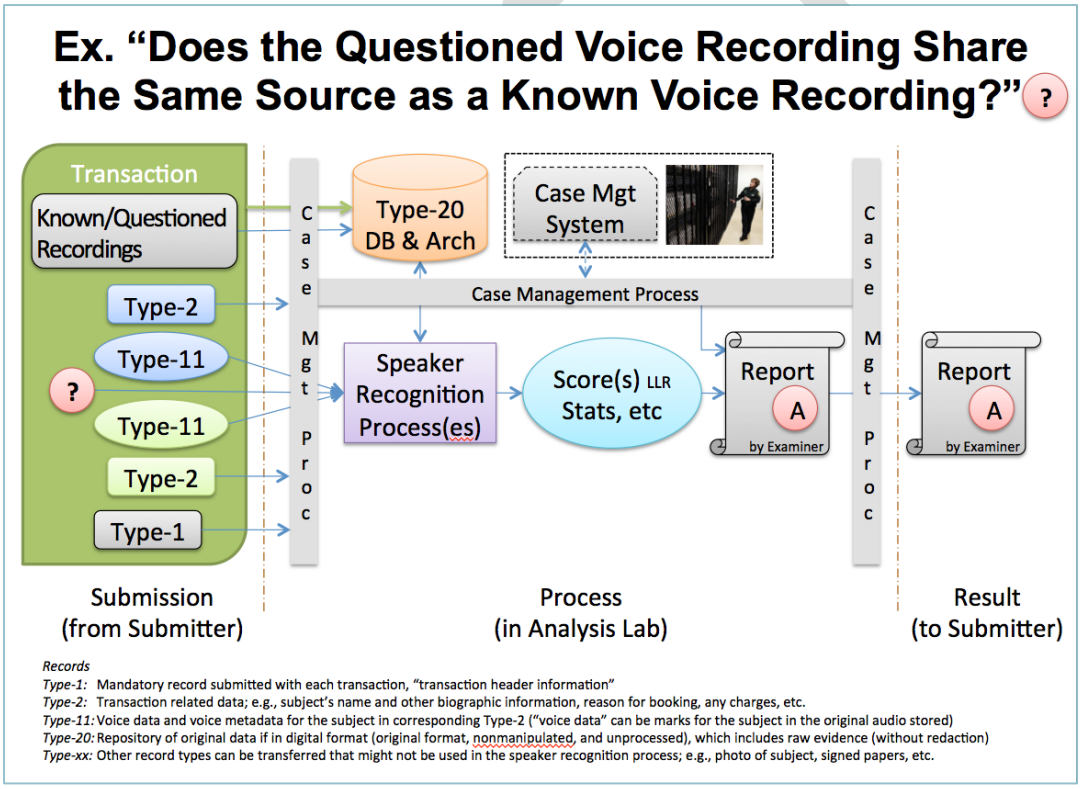
Examples of use of Type-11 Records, as envisioned by the SR Subcommittee, include submissions as part of ANSI/NIST-ITL Transactions:

- From a police station, interview room, or other controlled environment to a crime lab that can either process the voice data or forward it to a more appropriate forensic lab
- From a mobile collection device (possibly equipped with some local speaker recognition capability) used in criminal justice, border control, or military operations to an appropriate forensic lab or to automated/semi-automated speaker recognition system.
- As part of a booking record where the Type-11 Record is being legally captured for storage in a repository of biometric information for possible future use

2.2 Use of Type-11 in the Comparison of Two Voice Recordings

The following figure shows the use of ANSI/NIST records in determining if two voice-recording samples are from the same speaker. The Type-1 header data and the Type-2 subject (known or unknown) and case related data would be part of the transaction.

Figure 2 Example of ANSI/NIST Transaction in a SR Comparison



2.3 Other Possible Uses

While the Type-11 Records were designed to be used as part of an ANSI/NIST-ITL Transaction, they can also be used in various creative ways. The possible uses could include but are not recommended one way or the other by this document or by the OSAC SR Subcommittee:

- Exchange of voice records (as a set of Type-11 records with or without any other ANSI/NIST-ITL Record Types) between forensic labs – for collaboration or comparison with other known exemplars
- Storage in a system as background files for searching against – not dissimilar to latent or known fingerprint images being stored as Type-13 or -14 records to support fingerprint examiners in comparisons
- Transmission of speech samples from a central site to fixed or mobile sites for use in future tactical comparisons – extraction of the model(s) could occur either centrally prior to transmission or at the receiving end

3 References

The following list of documents are mentioned in this Guideline:

- ANSI/NIST-ITL 1-2011 Update: 2015 Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information; NIST Special Publication 500-290 Rev2 (2015) (in Final Draft for review)
- CRIMINAL JUSTICE INFORMATION SERVICES (CJIS) ELECTRONIC BIOMETRIC TRANSMISSION SPECIFICATION (EBTS); dated July 2, 2013; NGI-DOC-01078-10.0
- Criminal Justice Information Services (CJIS) Security Policy; Version 5.3; dated 8/4/2014; CJISD-ITS-DOC-08140-5.3
- Mobile ID Device Best Practice Recommendation Version 2.0; NIST Special Publication 500-280 V2
- NATO – STANAG 4715 BIOMETRICS DATA, INTERCHANGE, WATCHLISTING AND REPORTING

4 Terms and Definitions

4.1 Abbreviations

The following table provides the full terms associated with abbreviations used in this Guideline.

Table 1 Abbreviations

Abbreviation	Full Term
ANSI/NIST-ITL	American National Standards Institute / National Institute of Standards and Technology – Information Technology Laboratory
CAR	A type of TOT used in the criminal justice community – short for Criminal Tenprint Submission (Answer Required)
CJI	Criminal Justice Information
CJIS	Criminal Justice Information Services Division of the FBI

Abbreviation	Full Term
CJIS WAN	A Wide Area, secure network operated by the CJIS Division of the FBI
CPS	The Interpol equivalent to the CAR, a Criminal Print-to-Print Search transaction
DoD	Department of Defense
EBTS	Electronic Biometric Transmission Specification
EFTS	Electronic Fingerprint Transmission Specification
FBI	Federal Bureau of Investigation
FTP	File Transfer Protocol
ISO	International Organization for Standardization
IXM	IDENT Exchange Messages Specification
LIMS	Laboratory information management system
NGI	The Next Generation Identification System – a multi-modal biometric and forensic information processing system run by the FBI's CJIS Division
NIST	National Institute of Standards and Technology
RCMP	Royal Canadian Mounted Police
STANAG	A NATO abbreviation for Standardization Agreement
TOT	Type of Transaction – a mandatory field in ANSI/NIST-ITL Type-1 Records
XML	Extensible Markup Language

4.2 Terms and Definitions

To Be Provided in an update to this document as the Subcommittee's vocabulary document evolves.

5 Body of Guideline

5.1 Design Considerations

There are at least three high-level sets of conditions that will drive how Type-11 Records will be populated and used in an ANSI/NIST-ITL Transaction.

1. Fully controlled situations

2. Partially controlled – to uncontrolled situations
3. Third party provision of recordings

In each case it is important to understand that a full ANSI/NIST-ITL Transaction file need not be generated at the collection system (fixed or mobile device) level. The sequence of collection, preparation of the Type-11 record, and the generation of a full ANSI/NIST-ITL transaction can occur on using separate software programs on one or more computers or be integrated into one application suite on a single computer. Since the ANSI/NIST-ITL standard is a data transmission standard there are no constraints in the standard as to how records are to be generated. A "full ANSI/NIST-ITL transaction" is a data file that consists of the appropriate Record Types associated with a specific application profile of the ANSI/NIST-ITL Standard. The designers will need to familiarize themselves with that information as it is outside of the scope of this Guideline.

The following subsections provide three alternative approaches to populating Type-11 records; they are not intended to limit design or operational application of these guidelines.

5.1.1 Fully controlled collection / acquisition situations

Where the collection / acquisition can be fully controlled such as in an interview room at an FBI Field Office then the "system" can generate the appropriate Type-11 fields associated with technical aspects of the collection such as:

- Source Agency
- Geographical information
- Microphone type, model, and serial number
- Recording technique (e.g., encoding scheme)
- Distance between subject and the microphone
- Recording duration
- Date recording was made

Then a technician can work with the interviewing Special Agent and use the system to populate other information fields such as names, any diarization, or comments.

5.1.2 Partially controlled to fully un-controlled / acquisition situations

Where the controls are less rigorous (e.g., an opportunistic recording out of doors) or total control is not possible (e.g., an audio intercept of a phone conversation where the telephonic equipment used by the subject can not be determined) then post-processing will be required. The speech recording can be post-processed by an appropriate application that will create the Type-11 record by soliciting data entry of Type-11 and any other [e.g., Type-2] fields required for a full ANSI/NIST-ITL transaction. The application should solicit this information from an appropriate technician or investigator when the data fields cannot be auto-generated.

5.1.3 Third party acquisition situations

Where an audiotape or disk file is acquired from a third party there are two possible scenarios:

1. The submitter is able and willing to provide appropriate parametric information – Type-11 fields can be easily filled
2. The submitter is unable or unwilling to provide required macro-information then post-processing will be required to fill in some Type-11 field information

The speech recording can be post-processed by an appropriate application that will create the ANSI/NIST-ITL Transaction to include the Type-11 record by soliciting information from an analyst or investigator where the fields cannot be auto-generated. In case 2 scenarios often the technical recording information, such as Audio Recording Device, has to be marked as *unknown*.

5.2 Data Format

Data formats are closely tied to the overall design requirements and the overall transaction purpose. The various application profiles (e.g., FBI's the EBTS and the RCMP's NIST Interface Control Document or ICD) for implementation of the ANSI/NIST-ITL standard provide agency / domain specific transactions that are acceptable. The most frequently used type of transaction (known as a TOT) is probably the criminal tenprint submission (answer required), which is abbreviated CAR, since it is used by the FBI, the DoD, most US states, and the RCMP, while Interpol uses a similar TOT – CPS for Criminal Print-to-Print Search.

Each application profile identifies mandatory and optional Record types for each TOT associated with that profile. Examination of the appropriate profile will show which ANSI/NIST-ITL Record Types must / may optionally be employed within a speech-related TOT in addition to the Type-11 Record.

Having addressed the selection of TOT and mandatory / option fields there are numerous Data Fields in the Type-11 Record Format to be addressed. The appropriate implementation profile (e.g., the DoD EBTS) will provide some guidance and limits on which data fields are mandatory, optional, and conditional, beyond what is provided in the ANSI/NIST-ITL Standard itself.

It is important to note that the Type-11 records (like the records for friction ridge images, facial images, etc.) are intended to represent data acquired from a single subject. This becomes very important in situations such as a recording of a subject that also contains the voice of the interrogator, a translator, the subject, and an attorney OR a recording of two people talking on an intercepted telephone conversation.

There are at least three ways to deal with these situations.

1. If the speech data from the speakers other than the primary subject are not required for elimination, etc. then through diarization a single Type-11 record could be made and transmitted with or without the Type-20 record.
2. The original recording can be made into a Type-20 record and then a Type-11 for each person could be made into a separate Type-11 record and multiple Type-11 records could be in the same ANSI/NIST-ITL transaction.
3. The recording can be made into a Type-20 record and then be diarized with the individual Type-11s just referencing the start and stop times (a person-centric diary) for that individual and, again, multiple Type-11 records could be in the same ANSI/NIST-ITL transactions.

The Type-11 Record has 36 Fields listed and defined in Table 76 of the 2013 Update to the 2011 version of the ANSI/NIST-ITL Standard. A subset of the information is shown below in Table 2. Note that in the actual ANSI/NIST-ITL Table 76 there are additional critical piece of information such as subfields, condition codes (optional, mandatory, or conditional), number of occurrences, etc. In addition to the 36 defined fields there are 801 optional fields (Fields 100-900) simply designated as *User Defined* and approximately 150 fields designated as *Reserved for Future Use Only By ANSI/NIST-ITL*.

There are Machine Readable Tables (MRT) available for the ANSI/NIST-ITL 2013 revision Type-11 fields that describe the related Record and Field information – all downloadable for use in conformance checking, software development, etc. They are available on the Web at http://www.nist.gov/itl/iad/ig/ansi_references.cfm.

Table 2 Tyle-11 Record Fields

Field Number	Mnemonic	Content Description
11.001	N/A	Record Header
11.002	IDC	Information Designation Character
11.003	AOD	Audio Object Descriptor Code
11.004	SRC	Source Agency
11.005	VRSO	Voice Recording Source Organization
11.006	VRC	Voice Recording Content Descriptor
11.007	AREC	Audio Recording Device
11.008	AQS	Acquisition Source
11.009	RCD	Record Creation Date
11.010	VRD	Voice Recording Date
11.011	TRD	Total Recording Duration
11.012	PMO	Physical Media Object
11.013	CONT	Container
11.014	CDC	CODEC
11.021	RED	Redaction
11.022	RDD	Redaction Diary
11.023	DIS	Discontinuities
11.024	DSD	Discontinuities Diary
11.025	VOC	Vocal Content

Field Number	Mnemonic	Content Description
11.026	VCD	Vocal Content Diary
11.027	OCON	Other Content
11.028	OCD	Other Content Diary
11.032	SGEO	Vocal Segment Geographical Information
11.033	SQV	Vocal Segment Quality Values
11.034	VCI	Vocal Collision Identifier
11.035	PPY	Vocal Segment Processing Priority
11.036	VSCT	Vocal Segment Content Description
11.037	SCC	Vocal Segment Speaker Characteristics
11.038	SCH	Vocal Segment Channel
11.051	COM	Comments
11.902	ANN	Annotation Information
11.993	SAN	Source Agency Name
11.994	EFR	External File Reference
11.995	ASC	Associated Context
11.996	HAS	Hash
11.997	SOR	Source Representation
11.999	Data	Voice Data

It is important to note that while the default way to transmit a speech sample in a Type-11 record is to place it in field 11.999, a perfectly acceptable alternative is to use field 11.994, External File Reference. This is particularly useful if the recording is of substantial size, which could disrupt other network traffic if transmitted electronically. The following are some of the reasonable alternative data elements that could be provided via field 11.994.

- The URL of a web site where the file is available
- The FedEx tracking number of a shipment that contains a portable media version of the voice data
- A reference number for a evidence or property room stored recording
- A laboratory information management system (LIMS) case file number where further location information is available

5.3 Transmission of ANSI/NIST-ITL Transactions

There are two aspects of transmission that must be addressed:

1. Security considerations
2. Network considerations

They are not to be considered separately as the transmission of criminal justice information (CJI) can drive network selection and controls. It is not the intent of this Guideline to spell out the various network and security policies but rather to alert those considering, planning for, or implementing the electronic transmission of speech data for investigatory and forensic purposes be aware of these transmission factors: network and security.

5.3.1 Security Considerations

Typically the transmission of speech records will be for investigative / forensic use. In these cases there are usually some constraints on the handling and exchange of this data for reasons that can include: classification of the speech contents, privacy laws, confidentiality (not in the classification sense) of the associated purpose for the exchange of the records. In the US, the FBI runs the largest criminal justice network for the exchange of information and they have set rules. The DoD runs worldwide networks and they also have security policies for the use of the networks and the exchange of data. It is the responsibility of the groups (police departments, forensic laboratories, military sites, etc.) implementing the Type-111 Record in an ANSI/NIST–ITL transaction to conform to the appropriate security policies.

The rest of this section provides some insight into the current, relevant FBI policies. The Criminal Justice Information Services (CJIS) Division of the FBI provides a Security Policy document (CJISD-ITS-DOC-08140-5.3). The introduction section of this CJIS Security Plan states that it provides "*a minimum set of security requirements for access to Federal Bureau of Investigation (FBI) Criminal Justice Information Services (CJIS) Division systems and information and to protect and safeguard Criminal Justice Information (CJI). This minimum standard of security requirements ensures continuity of information protection. The essential premise of the CJIS Security Policy is to provide the appropriate controls to protect CJI, from creation through dissemination; whether at rest or in transit."*

This CJIS policy is a baseline that can be supplemented by or integrated with local departmental polices and forensic laboratory standards. It is the obligation of the person transmitting speech records to determine if they are CJI records, contain personally identifiable information (PII), or are being sent to a controlled system environment such as the CJIS systems and networks or certain DoD systems and networks. Having made such a determination it is incumbent on the sending authority to determine and implement the appropriate security policies.

5.3.2 Network Considerations

When transmitting speech records electronically the security policies and the relative sensitivity of the data will determine and possibly limit the choice of networks available. It is well beyond the scope of this document to attempt to delineate or recommend specific networks. This section, therefore, provides a brief outline on some alternatives.

There are several networks available in the US for potential use by those sharing speech records electronically. A basic consideration, beyond security constraints, is the level of network connectivity available between the sites intending to share transactions. The FBI CJIS Division runs a secure, CJI-compatible, wide area network between their facilities in Clarksburg, WV, the FBI Lab in Quantico, VA, all 50 states and several federal agencies. While that network, known as the CJIS WAN, is able to handle the sharing of transactions that include Type-11 records the Next Generation Identification (NGI) System is not yet prepared to process these records. The DoD has various networks over which they are currently transmitting speech records as Type-11 records between and among some sites where they are stored and processed.

In fact, there are numerous networks – all purpose-built as opposed to just using the Internet to exchange files via email, the File Transfer Protocol (FTP), etc. Examples of purpose-built / dedicated networks include:

- The CJIS WAN
- The DoD data networks at various security levels across the globe
- Interpol connectivity to the National Police Bureau offices in each member country
- Many US states have secure networks suitable for exchanging CJI, which are tied into the CJIS WAN

5.3.3 Best Practices for Mobile ID Use of Voice Data

NIST has been working on an update to the Mobile ID Device Best Practices Recommendation; Version 2 of these recommendations was published in November 2015. This is a useful document for designers of mobile applications / devices. Section 4 of that DRAFT document provides System Design Considerations that go beyond these Guidelines.

Section 10.0 specifically addresses Voice Signals. There are three subsections to Section 10 of the Mobile ID Best Practices document:

- Voice Signal Capture Devices (Section 10.1)
- Voice Signal Capture (Section 10.2) – this addresses container and CODEC codes
- Voice Signal Transmission (Section 10.3)

6 Appendix 1 Detailed History of the ANSI/NIST-ITL Standard

The standard started as a result of discussions at a fingerprint conference held in London in 1985. The participants, including the FBI, were looking for a way to exchange fingerprint data between agencies and even with the UK Home Office. This effort led the then National Bureau of Standards (NBS) to host a workshop in 1985 and to canvass 43 organizations. The result was that they gained ANSI approval of a standard entitled: ANSI/NBS-ICST 1-1986: American National Standard for Information Systems - Fingerprint Identification – Data Format for Information Exchange.

According to those involved at the NBS (now NIST) in the development of the 1986 standard and later versions – *while the 1986 version was never used or tested, it set the stage and provided the model for the development of the 1993 ANSI/NIST-ITL standard.*

The 1993 version was funded by the FBI to permit agencies to submit fingerprints electronically to the Integrated Automated Fingerprint identification System (IAFIS) that was under development at the time. There were three workshops held between 1990 and 1992 that led to this standard being approved by the canvassees.

The 1993 version was entitled: ANSI/NIST-CSL 1-1993: American National Standard for Information Systems - Fingerprint Identification – Data Format for Information Exchange. At that point in time there were nine record types:

Table 3 Record Type Definitions

Record Type	Record Name	Purpose
Type-1	Transaction record	Header information necessary for transmission addressing, determining the purpose for the transmission, and parsing of the transaction
Type-2	User-defined text record	Exchange of information about the subject of the fingerprint data to include demographic and biographic information
Type-3	Low-resolution grayscale image record	Fingerprint images – never used operationally
Type-4	High-resolution grayscale image record	Fingerprint images
Type-5	Low-resolution binary image record	Fingerprint images – never used operationally
Type-6	High-resolution binary image record	Fingerprint images – never used operationally
Type-7	User-defined image record	Latent palm prints and any other images of interest
Type-8	Signature image record	Signature of subject fingerprinted and person capturing the fingerprints
Type-9	Minutiae record	Typically used to latents to avoid sending a high resolution image, which in 1993 was expensive to store as well as transmit

The move from the 1986 version to the 1993 version saw the size of the Standard shrink from 64 pages to 30 pages as many FBI-specific field values were removed and made optional fields in the Type-2 record. These included such fields as hair color codes. They were removed to permit the UK and others to use the Standard without having to change their codes and abbreviations.

By 1997 NIST issued, through one workshop held in 1995, an addendum to the 1993 Standard entitled: ANSI/NIST-ITL 1a-1997 Addendum to ANSI/NIST-CSL 1-1993 American National Standard for Information Systems – Data Format for the Interchange of Fingerprint, Facial & SMT Information. This addendum was 32 pages and it added the following Record Type:

Table 4 1997 Addendum Record Type Definitions

Record Type	Record Name	Purpose
Type-10	Facial & SMT image record	Mugshots and scars (needle) marks, and tattoos

During a workshop held in September 1998, these two standards (1993 and the 1997 Addendum) were reviewed and discussed. Several agreements were reached affecting various aspects of the standard. These included the addition of new fields, new record structures, and the merging of the two standards into this single standard. This 2000 version of the standard is the combination of the two previous standards plus additional record type descriptions, the incorporation of agreements reached during the September workshop, and subsequent proposals offered for improving the standard. It is formally entitled: ANSI/NIST-ITL 1-2000 Revision of ANSI/NIST-CSL 1-1993 & ANSI/NIST-ITL 1a-1997. This version of the Standard grew to 81 pages and it added the following Record Types:

Table 5 2000 Version Record Type Definitions

Record Type	Record Name	Purpose
Type-11	Reserved for future use	
Type-12	Reserved for future use	
Type-13	Variable-resolution latent image record	Exchange of latent images – preferably scanned at 1,000 ppi; to be used in lieu of Type-7 Records.
Type-14	Variable-resolution tenprint image record	Exchange of fingerprint images at a higher resolution than supported by the Type-4 record; scanned and transmitted at 500 ppi times powers of 2.
Type-15	Variable-resolution print [2]image record	Exchange of palm print images scanned and transmitted at 500 ppi times powers of 2; to be used in lieu of Type-7 Records.
Type-16	User-defined testing image record	Exchange of other images – such as footprints, crime scenes, and iris images; eventually new Record Types were added to support each of these image sets.

The 2007 version of the standard [ANSI/NIST-ITL 1-2007 Data Format for the Interchange of Fingerprint, Facial, & Other Biometric information – Part 1] was the result of agreements reached during two workshops held in April and December of 2005 to review the ANSI/NIST-ITL 1-2000 standard. During the first workshop, proposals to update the standard were introduced. The proposals were refined and presented for discussion and approval during the second workshop. The major enhancements in this revision include:

- Image quality and segmentation data to support the processing of the "flat" images

[2] While there are two alternative spelling of friction ridges printed from palms (palm print and palmprint) this document uses the alternative found in the ANSI/NIST-ITL Standard: palm print.

- Definition of a new block of minutiae fields to harmonize with the INCITS M1 minutiae standard
- Best practice application levels for the capture of facial images
- A new record type for the exchange of iris information
- A new record type to contain biometric information not described in this standard but conformant to other registered biometric data format standards; and
- An XML alternative representation for this standard.

This version of the Standard grew to 164 pages without Part 2 – the XML description and it added the following Record Types:

Table 6 2007 Version Record Type Definitions

Record Type	Record Name	Purpose
Type-17	Iris image	The exchange of iris images.
Type-18-98	Reserved for future use	
Type-99	CBEFF Biometric data record	Type-99 logical records shall contain and be used to exchange biometric data that is not supported by other ANSI/NIST-ITL logical record types. This provides a basic level of interoperability and harmonization with the ANSI INCITS and other biometric interchange formats.

In 2008, 'NIEM-conformant encoding' using Extensible Markup Language (XML) was adopted. NIEM, the National Information Exchange Model, is a partnership of the U.S. Department of Justice and Department of Homeland Security. NIEM is designed to provide a common semantic approach in XML applications. With some minor exceptions, the 2007 and 2008 versions of the standard are equivalent except for the encoding format. In 2009, an amendment to the 2007 and 2008 versions was approved that extended codes to handle multiple finger capture. These were supplemented in 2009 [ANSI/NIST-ITL 1a-2009] with the ability to exchange fingerprint images with multiple fingers.

The 2011 version of the standard [ANSI/NIST-ITL 1-2011 Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information] replaces ANSI/NIST-ITL 1-2007 and ANSI/NIST-ITL 2-2008 standards and the amendment ANSI/NIST-ITL 1a-2009.

This version of the Standard grew to 472 pages and it deprecated three existing Record Types, specified two records to be added in the future, and added five new Record Types. The changes are shown below:

Table 7 2011 Version Record Type Definitions

Record Type	Record Name	Purpose
Type-3	Low-resolution grayscale image record	Deprecated

Record Type	Record Name	Purpose
Type-5	Low-resolution binary image record	Deprecated
Type-6	High-resolution binary image record	Deprecated
Type-11	Reserved for voice data	
Type-12	Reserved for dental records	
Type-18	DNA data image record	Exchange DNA and related data. It was developed to provide a basic level of interoperability with the draft format of the ISO/IEC 19794-14 DNA data interchange format.
Type-19	Variable-resolution plantar image record	Exchange of variable-resolution plantar print image data together with fixed and user-defined textual information fields pertinent to the digitized image.
Type-20	Source representation image record	Some possible uses of the Type-20 record are: • From a group photo stored in a Type-20 record, a subject's face is segmented and stored in a Type-10 record. • From a high-resolution color image in a Type-20 record, two latent fingerprint images are segmented, rescaled and gray-scaled for storage in separate Type-13 records. • From a series of off-angle face images stored in separate Type-20 records, a single 2D face image is generated (using fusion) that is stored in a Type-10 record.
Type-21	Associated context image record	The Type-21 record shall contain an associated context image, audio / visual recording or other related data. This record type does NOT contain information used to derive biometric information contained in other records. Record Type-20 serves that function. Record Type-21 may be used to convey contextual information, such as an image of the area where latent fingerprints were captured.
Type-22-97	Reserved for future use	
Type-98	Information assurance data record	Exchange of security information that allows for the assurance of the authenticity and/or integrity of the transaction, including such information as binary data hashes, attributes for audit or identification purposes, and digital signatures.

The 2013 version of the Standard [ANSI/NIST-ITL 1-2011 Update: 2013] includes the full integration of supplemental information, added after the 2011 version was approved, defining Voice [ANSI/NIST-ITL 1-2011 Sup:Voice**)**] and Dental [ANSI/NIST-ITL 1-2011

Sup:Dental] records. It also corrects miscellaneous errors that were noticed after publication of the 2011 version. Some additional capabilities have been added to the standard that were not part of the Supplements, but have been requested by users of the standard.

The introduction to this version spends a full page describing the genesis of the Type-11 Record and it is worth repeating here:

> *Editor for Forensic and Investigatory Voice: Jim Wayman*
>
> *Investigatory Voice Biometric Committee (IVBC)*
>
> *Joseph Campbell, Carson Dayley, Craig Greenberg, Peter Higgins, Alysha Jeans, Ryan Lewis, Jim Loudermilk, Kenneth Marr, Alvin Martin, Hirotaka Nakasone, Mark Przybocki (Chair), Vince Stanford, Pedro Torres-Carrasquillo, James Wayman, Brad Wing.*
>
> *ANSI/NIST-ITL Voice Working Group*
>
> *Kristin Allen, Jeremiah Bruce, Lynn Clelland, Lindsey Crookshanks, Lisa Fritts, Brian Finegold, Patrick Gibbs, Craig Greenberg, Martin Herman, Peter Higgins, Kathy Higgins, Brendan Klare, Ken Marr, Bill Michael, Larry Nadel, Hiro Nakasone, Marina Nastasenro, Anastasia Pavlovic, Mark Przybocki (Chair), Mary Rendel, John Roberts, Bonnie* [sic] *Scheier (Original Chair), Vince Stanford, Jennifer Stathakis, Reva Schwartz, Carolyn Taborini, Cathy Tilton, An (Mike) Tran, Ryan Triplett, Jim Wayman, Cathy Wimer, Brad Wing.*
>
> *Work on a proposed Supplement for voice data evolved out of the NIST / FBI co-sponsored 'US Government Interagency Symposium for Investigatory Voice Biometrics' which occurred in 2009. The Symposium created four working groups, one of which dealt with interoperability. That working group's report stated: "One of the fundamental goals of the Symposium on Investigatory Voice Biometrics was to initiate 'a multi-year program to develop investigatory voice biometric collection and interoperability standards'". That Working Group recommended the further examination of adding a voice data capability to the ANSI/NIST-ITL standard. Subsequent to that Symposium, the FBI's Biometric Center of Excellence (BCOE) initiated a multi-year sponsorship to form the Investigatory Voice Biometric Committee (IVBC) in collaboration with the NIST to define the U.S. government agency's technical requirements for voice data collection, transmission format, and analysis. One of its key recommendations was to develop a standardized approach for voice data exchange.*
>
> *Concurrently, certain stakeholders involved in the ANSI/NIST-ITL process had recognized the need for a standardized way to exchange voice data. An ANSI/NIST-ITL working group was formed in 2010 to develop text for inclusion in the standard, under the leadership of Bonny Scheier. That working group agreed that the 2011 update to the standard would proceed without a new voice record, since there was still substantial work to be done*

in creating a new record type. Note that the record for the 2010 workshop recognized the need for a voice record by establishing a placeholder for a new voice record type. The Workshop summary notes that a working group would further pursue development of a new Voice data record type.

The IVBC developed the first draft of the Type-11 record. In August 2012, the IVBC turned over its work to a newly re-formed ANSI/NIST-ITL Voice Working Group, in order to allow participation of a wider group in the development process. It is the result of the combined work of all of these groups that this Supplement was produced. The development of Record Type-11 owes much to the dedicated effort of Jim Wayman, who served as editor throughout the process, and to Hiro Nakasone, who directed the IVBC activity and championed the development of the Supplement.

This version of the Standard grew to 608 pages and it added two Record Types (Types 11 and 12), previously indicated as *to be added in the future*, and added one new Record Type (Type-22). The changes are shown below:

Table 8 2013 Version Record Type Definitions

Record Type	Record Name	Purpose
Type-11	Forensic and investigatory voice data records	Support the transmission and / or descriptions of audio recordings containing vocalizations by one or more speakers. Type-11 records support transactions related to detecting and recognizing speakers, extracting speech segments from an audio recording that are attributable to a single speaker, and linking speech segments by speaker. Type-11 does not define the transmission of features or models extracted from voice data, but does allow the user to use specific fields to contain such information.
Type-12	Forensic dental and oral data records	Exchange information concerning an individual's dental or oral characteristics. It is designed to use the lexicon of ANSI/ADA Standard No. 1058 – Forensic Dental Data Set. A Type-12 record is typically used in conjunction with a Type-22 record, which can convey dental radiographs and other related imagery useful in forensic dental procedures. Type-10 intra-oral and extra-oral photographs are also used in conjunction with a Type-12 record.
Type-22	Non-photographic imagery records	Transmit forensic imagery and associated data that is not included in the Type-10 record – which contains 2D visible light (380-780 nm) images.
Type-23-97	Reserved for future use	

There is a 2015 Update to the ANSI/NIST-ITL 2011 2013 Update in the review cycle.

There are no substantive changes to the Type-11 record in that version.