Designation: E2364 – 04 (Reapproved 2010)

An American National Standard

Standard Guide to Speech Recognition Technology Products in Health Care[1]

This standard is issued under the fixed designation E2364; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide identifies system types and describes various features of speech recognition technology (SRT) products used to create the healthcare record. This will assist users (health information professionals, medical report originators, administrators, medical transcriptionists, speech recognition medical transcription editors (SRMTEs), system integrators, support personnel, trainers, and others) to make informed decisions relating to the design and utilization of SRT systems.

1.2 This guide does not address the following items:

1.2.1 System and data (voice and text) security.

1.2.2 Administrative processes such as authentication of the document, productivity measurements, etc.

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E1902 Specification for Management of the Confidentiality and Security of Dictation, Transcription, and Transcribed Health Records (Withdrawn 2011)[3]

E1985 Guide for User Authentication and Authorization

E2084 Specification for Authentication of Healthcare Information Using Digital Signatures (Withdrawn 2009)[3]

E2184 Specification for Healthcare Document Formats (Withdrawn 2011)[3]

E2185 Specification for Transferring Digital Voice Data Between Independent Digital Dictation Systems and Workstations (Withdrawn 2010)[3]

E2344 Guide for Data Capture through the Dictation Process

2.2 *Other Documents:*

Resource Interchange File Format (RIFF) Standard

3. Terminology

3.1 *Definitions:*

3.1.1 *acoustic model, n*—phoneme map of user.

3.1.2 *authentication, n*—the process of confirming authorship of an entry or of a document, for example, by verifying with a written signature, identifiable initials, computer key, or other methods.

3.1.3 *author, n*—person responsible for content of text file.

3.1.4 *back-end system, n*—delayed processing for document completion.

3.1.5 *compound file, n*—a file containing recorded voice with its transcribed text.

3.1.6 *context, n*—a long list of vocabulary words and phrases used for the particular subject matter, with their spellings and pronunciations, statistical information about usage of each word alone and in combination. For example, the context may include the number of times that "right," "Wright," "turn right," "right turn," "right hand," and "Mr. Wright" occur in a body of text. It also includes grammar and style information. Language model, lexicon, topic, and vocabulary are terms that are all used synonymously with context.

3.1.7 *digital signature, n*—data associated with, or a cryptographic transformation of, a data unit that allows a recipient to prove the source and integrity of the data unit and protect against forgery, for example, by the recipient.

3.1.8 *edit, v*—to review the document while listening to the originator's recorded voice and reading the associated transcribed text (compound file), checking for recognition errors and correcting document formatting and other inconsistencies. When the SRMTE is not the originator, the SRMTE may need to flag the document for originator/author clarification of unclear content or intent.

3.1.9 *encryption, n*—the process of transforming plain text (readable) into cipher text (unreadable) for the purpose of security and privacy.

3.1.10 *front-end system, n*—a system incorporating real-time recognition and may include real-time self-editing by the originator.

3.1.11 *language model, n*—context specific to medical specialty, user, or practice setting.

[1] This guide is under the jurisdiction of ASTM Committee E31 on Healthcare Informatics and is the direct responsibility of Subcommittee E31.15 on Healthcare Information Capture and Documentation.

Current edition approved March 1, 2010. Published August 2010. Originally approved in 2004. Last previous edition approved in 2004 as E2364–04. DOI: 10.1520/E2364-04R10.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] The last approved version of this historical standard is referenced on www.astm.org.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959. United States

Documents
Document Sharing Hub

Full document details: ASTM E2364-04(2010)

3.1.12 *lossless compression, n*—a lossless compression reduces the amount of data required to represent the original voice file but has no impact on sound quality. The original file can be replicated precisely at any time.

3.1.13 *lossy compression, n*—a lossy compression loses some information, resulting in degradation of the sound quality inherent in the original voice file and an inability to precisely regenerate that original file.

3.1.14 *microphone, n*—an instrument whereby sound waves are caused to generate or modulate an electric current usually for the purpose of transmitting or recording sound (as speech or music).

3.1.15 *microphone element, n*—the component within the microphone that does the actual conversion from sound waves to electrical signals.

3.1.16 *natural language processing, n*—method used in artificial intelligence to process and derive interpretation of human language.

3.1.17 *networked system, n*—system connected to a network.

3.1.18 *"normal" dictation, n*—routine phrases or paragraphs.

3.1.19 *originator, n*—person who provides oral input or dictation, not necessarily the person responsible for the content.

3.1.20 *phoneme, n*—smallest unit of sound in a spoken language.

3.1.21 *prompts, n*—reminders provided in order to complete a task.

3.1.22 *real-time recognition, n*—simultaneous speech-to-text transcription.

3.1.23 *RecO*—speech recognition error

3.1.24 *RIFF file, n*—Resource Interchange File Format (RIFF) is self-descriptive; that is, the voice file format is defined within the file.

3.1.25 *speech recognition, n*—computerized transcription of speech to text.

3.1.26 *speech recognition medical transcription editor, n*—medical transcriptionist who edits compound files and/or the SRT language model.

3.1.27 *SRT engine, n*—speech recognition processor.

3.1.28 *standalone system, n*—system not connected to a network.

3.1.29 *synchronization, v*—having voice and text matched such as in a point-and-play manner.

3.1.30 *text file, n*—a file that contains text message.

3.1.31 *voice enrollment, n*—the process whereby a user reads aloud selected text so the SRT software can map or record the user's speech sound pattern (phonemes).

3.1.32 *voice file, n*—digitalized audio message representing voice input.

3.1.33 *voice macros, n*—stored keystrokes that are activated by a voice command.

3.1.34 *WAV, n*—voice file format.

3.2 *Acronyms:*

3.2.1 *MT*—medical transcriptionist

3.2.2 *SRMTE*—speech recognition medical transcription editor

3.2.3 *RIFF*—resource interchange file format

3.2.4 *SRT*—speech recognition technology

4. Significance and Use

4.1 This guide is intended to provide general guidelines toward the design and utilization of SRT products used for healthcare documentation. It is intended to recommend the essential elements required of SRT systems in healthcare.

4.2 This guide will not identify specific products or make recommendations regarding specific vendors or their products or services.

4.3 A well-edited SRT document may result in improved quality over current methods of documentation, that is, handwritten notes and improved productivity over traditional dictation and transcription.

4.3.1 Faster turnaround times.

4.3.2 Legible documentation over handwriting has many advantages:

4.3.2.1 Improved patient care communication.

4.3.2.2 Enhanced patient safety.

4.3.2.3 Reduced malpractice risks.

4.3.2.4 Facilitation of appropriate reimbursement.

4.3.3 For the medical transcriptionist and/or SRMTE, decreased repetitive stress injuries, such as neck, arm, wrist, and heel pain.

4.3.4 Facilitation of cost controls related to document completion.

4.3.5 Better utilization of medical language skills of MTs as productivity is not limited by keyboarding skills.

5. Speech Recognition Technology Systems

5.1 Speech recognition technology (SRT) is designed to capture voice and transcribe that speech into text. This can be done by a single user working at a standalone computer or by a large group of users working on a network. Another method is processing a pre-recorded digital voice file through an SRT system, with the resulting text and/or SRT engine being edited by the MTE.

5.2 Speech recognition technology system workflow.

5.2.1 Front-end speech recognition process involves:

5.2.1.1 Recording the voice.

5.2.1.2 SRT transcription of the voice file to text.

5.2.1.3 Editing may be done by the originator and/or SRMTE.

5.2.1.4 Compound file may be saved as an option.

5.2.1.5 Text file can be printed, archived, transmitted, or integrated into an electronic health record.

5.2.1.6 Update the SRT context for RecOs and new terminology.

5.2.2 Back-end speech recognition process involves:

5.2.2.1 Recording the voice.

AVA Documents
Document Sharing Hub

Full document details: ASTM E2364-04(2010)