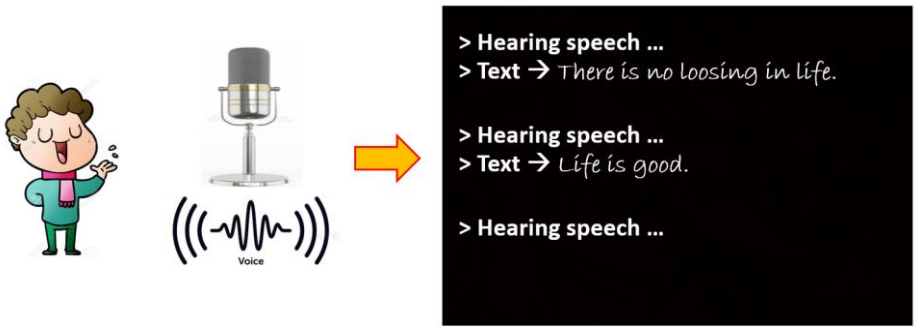
SOFTWARE REQUIREMENT SPECIFICATION (Version 0.9)

Speech-to-Handwritten Text, (SThT)

Application that generates a Personalized Handwritten Text for a given Speech



Nov. 2022

DUE DATES

	Scope of AD	Due Date	Length Limit
Interim Report	Chapters 1~4	11/10(Th), 9pm	30± Pages
Pre-Final Report	Chapter 5	11/20(Sun), 9pm	50± Pages
Final Report	Chapters 6~7	12/05(Mon), 9pm	70± Pages

CEP INSTRUCTOR

Soo Dong Kim, Ph.D.
Professor, School of Software
Soongsil University, Seoul, Korea
010-7392-2220, sdkim777@gmail.com, https://soodongkim.net

삼성전자 첨단기술연수소

\

TABLE OF CONTENTS

Speech-to- Handwritten Text, (SThT)

1. Purpose of the Document

The purpose of this document is to specify the requirement for developing the target system in this CEP. The requirement will become the basis for designing the software architecture of a target system, which is required as a fulfillment to acquire the Samsung Associate Architect (AA) certification.

2. Comprehensive Evaluation Project (CEP)

2.1. About Associate Architect (AA) Program

Associate Architect Program of Samsung Electronics is to provide participants with two sets of software architecture design proficiency.

- **Body of Knowledge (BoK) on Software Architecture**

  This set includes the fundamental theories and methods for designing SW architecture

- **Skillset for Designing SW Architecture**

  This set includes the practical skill for applying architecture design methods to a given SRS.

CEP is designed to fulfill the second set of AA program through an individual design project.

2.2. How is the CEP problem prepared?

The CEP problem is prepared by the instructor, based on the following principles.

- **Principle 1. Utilizing the Whole BOK of Software Architecture Design**
  - Utilizing Architecture Styles
  - Designing Architecture for multiple Views
  - Designing Architecture for Non-Functional Requirements
- **Principle 2. Handling the Complexity of Industrial Systems**

  The target system to design in CEP is an industry-level complex software system, i.e., not an academic problem appeared in books or literature.
- **Principle 3. Solution Not Available in Public**

  CEP problem is not a reproduction of already existing exercise problem in books, and hence the architecture design solution for the CEP problem is not available in public.

2.3. Architecture Design Reports in CEP

Each participant designs and submits the design of software architecture for the target system in incremental manner.

- Interim Report
  This report includes the context analysis model and the skeleton architecture design of the target system.
- Prefinal Report
  This report includes the architecture design for multiple views of the system: functional view, information view, behavior view, and deployment view.
- Final Report
  This report includes the architecture design for non-functional requirements and the validation of the architecture design.

❑ **Weight Distribution of CEP Reports**

Interim Report	Prefinal Report	Final Report	TOTAL
30 points	30 points	40 points	**100 points**

2.4. Template for Architecture Design

A template for designing the software architecture in CEP is provided and hence participants can utilize the template in specifying the architecture design. The template is devised to be consistent with the architecture design methodology provided by the instructor.

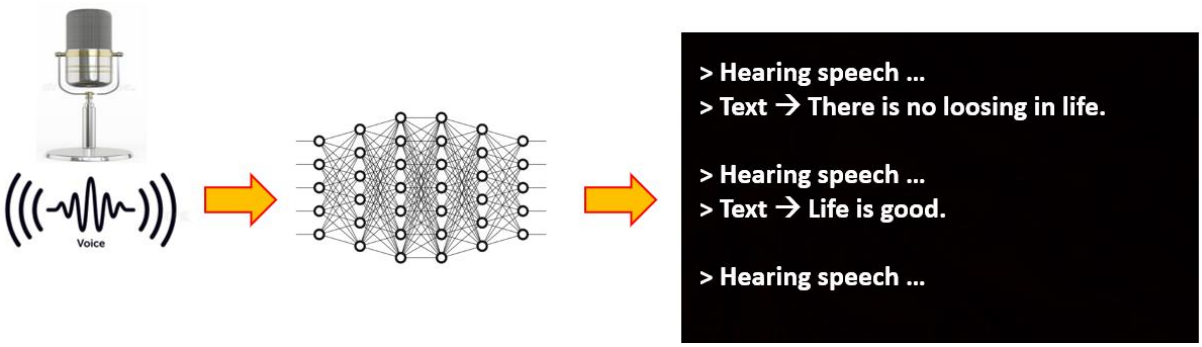
❑ **Filename**

AA-2022-A5, CEP, AD Template.docx

3. Overview of the Target System

3.1. What is Speech to Text (STT)?

Speech-to-Text or speech recognition is a technology of transcribing spoken words or audio content into text. It is accomplished using applications, APIs, tools, and other software solutions.



STT performs the speech recognition to transcribe voice into written text by utilizing machine learning models to detect patterns in sound waves for accurate transcription.

- **Applications of STT Services**

  STT provides an alternative way to enter information. Instead of typing words manually, we can dictate words, and the STT service will produce the written text for the given speech. Hence, it be utilized in various types of applications, including the following.

  - Voice commanding
  - Smart assistant
  - Chatbot
  - Translation
  - Mixed Language Detection
  - Transcriptions for Call Centers

- **Providers of STT Service**

  There are a number of STT service providers, including the followings.

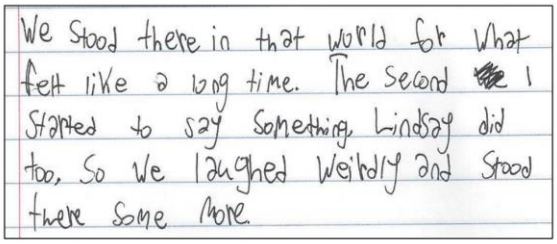
  - Google Cloud's Speech-to-Text
  - IBM Watson Speech to Text
  - Microsoft Cognitive Services
  - Microsoft Azure Speech to Text
  - Amberscript
  - AssemblyAI
  - Rev.ai

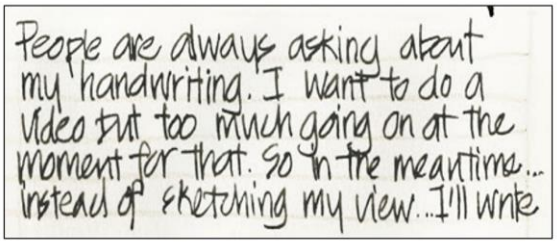
- Scriptix
- Speechmatics

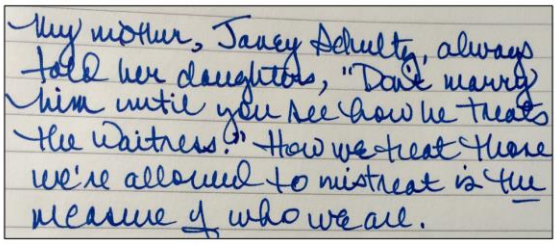
3.2. Handwritten Texts

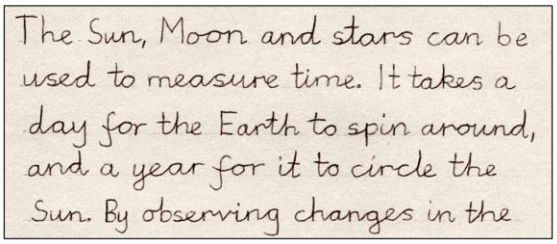
A handwritten text is a form of text that is specific to each writer. Hence, a handwriteen text represents the write's unique and distinguishable patterns of writing with a pen.

Examples of handwritten texts are shown in the following figure. Each handwriting is unique and distinct from others' handwritten texts, and hence a handwritten text is also used a means to identify the writer of the text.







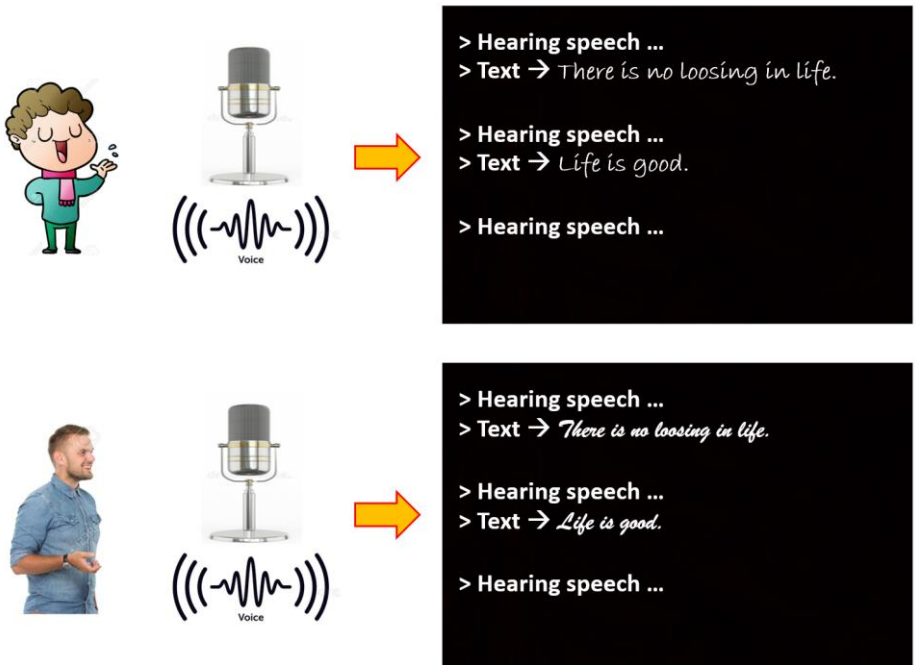


Handwritten texts also provide a feel of warmth, respect, and sincerity to the readers, which are not available in computer fonts-generated texts. For instance, a handwritten letter would provide a sense of care and consideration to the letter recipient.

3.3. What is Speech to Handwritten Text Service?

The target system, Speech to Handwritten Text (SThT) is a software application that records a short speech or reads an audio file of speech, and transforms the speech into a personalized handwritten text.

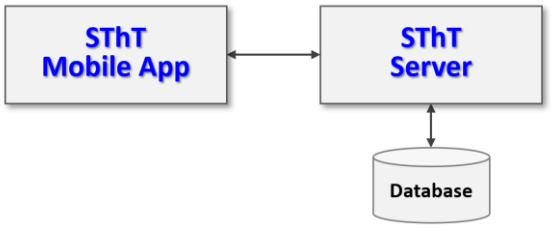
Each person has its own handwriting style as shown in the following figure.



Note that two different people have two different handwriting styles as shown in the figure. Therefore, the SThT system must learn the personal handwriting style of each user with a high accuracy. A system generated handwritten text should appear as close as the manually handwritten text.

3.4. Deployment of the System

The target system consists of a mobile app as the client node and its server maintain the master database.



- **(Native) Mobile App**

  This client node provides the functionality of user registering, training personal handwriting styles, and requesting the SThT service.

- **Server**

  The server provides all the key functionality of the system and maintains the master database.

4. Functional Requirements

The functionality of the system is organized as a set of functional categories.

4.1. User Profile Management

This functionality is to manage the profiles of users who use the SThT service. When registering, the user provides his or her identification information, contact information, login information, and credit card information.

4.2. Staff Profile Management

This functionality is to manage the profiles of company staffs who manage the system operations and the server computer system. When registering, the staff provides his or her identification information, contact information, login information, and departmental information.

4.3. Handwriting Training Management

This functionality is to train the personal handwriting style from the handwritten texts by a user. The user repeats writing pangram sentences provided by the system, which then are used as a training data.

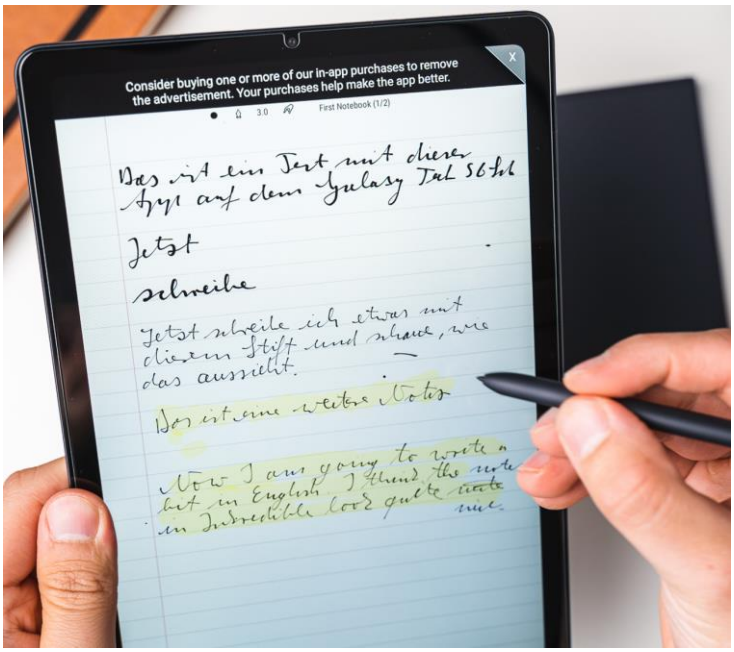
A pangram or holo-alphabetic sentence is a sentence using every letter of a given alphabet at least once. Pangrams have been used to display typefaces and develop skills in handwriting. Examples of pangram sentences are given here.

- The quick brown fox jumps over a lazy dog. (33 letters)
- The five boxing wizards jump quickly. (31letters)
- Jackdaws love my big sphinx of quartz. (31 letters)
- Pack my box with five dozen liquor jugs. (32 letters)

Each pangram sentence includes each character of 26 letters in English alphabet at least once.

- **Entering Personal Handwritings for Training**

  Each user writes the given pangram sentences for some number of times. This is to capture the variability of handwriting writings for a same sentence.



- **Learning the Personal Handwriting**

  The system must capture and learn the personal handwriting styles and store them in the system database for further applications.

  Note that the training of personal handwriting style may not be successful if the user has a high variability of handwriting, i.e., having high irregularity in writing same letters and words. In this case, the system asks the user to handwrite different pangram sentences repeatedly until the training is successfully completed.

4.4. SThT Session Management

This functionality is to acquire a speech or an audio file of speech, and transform the speech into a handwritten text.

- Input: A Segment of Speech
- Output: Handwritten Text

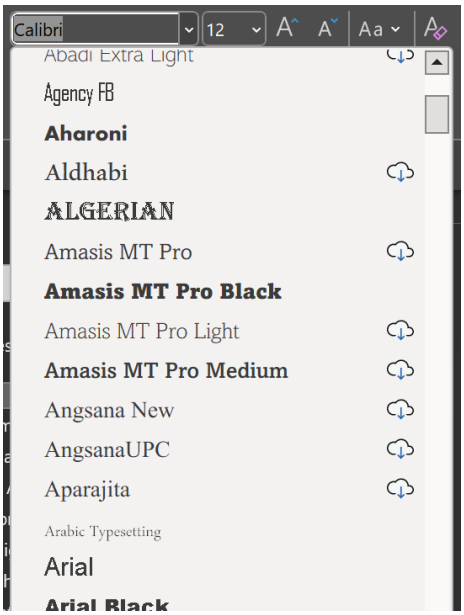
The procedure of running a SThT session is an ordered list of tasks.

- Task 1. Record a speech at runtime. Alternatively, the system reads an audio file of speech
- Task 2. Convert the speech into a text by invoking a STT microservice.
- Task 3. Convert the text into a personalized handwritten text. The system also computes the confidence value, i.e., the performance measure for the generated handwritten text.
- Task 4. Return the handwritten text to the user, as the output.

4.5. Personalized Font Generation

This functionality is to generate a software font that is specific to the user. That is, the generated handwriting font is supposed to be unique and distinguishable from handwritings of other users.

Once the personalized font is generated, it can be downloaded by the user and utilized in OA software applications such as Word, as shown in the following figure.



The system should support popular file formats of the fonts including the followings.

- FNT, TTF, and TTC

4.6. Billing and Payment Management

The system provides the SThT conversion service free of charge. However, the system chargers a fixed fee for each generated personal font. The system processes the fee payment using the registered credit card of users.

4.7. Report Generation

This functionality is to generate various reports on the system operations. A set of pre-specified reports are periodically generated. Reports can also be generated on demand.

5. Non-Functional Requirements

There can be several non-functional requirements that are essential in the target system. For CEP, we consider only 2 NFR items.

5.1. NFR-1. Accuracy of Personalized Handwritten Texts

The system should be designed to provide a high accuracy of personalized handwritten texts. That is, the system has to learn the user's specific handwriting styles for each letter of the alphabet with a high accuracy.

Note that the handwriting of each letter may vary depending on the surrounding characters of a letter. That is, the way to writing a letter depends on the preceding letter and the subsequent letter. For example, The way to write a letter 'a' could vary depending on what the sequence of letters including 'a'.

- apple
- access
- axis
- diary
- jazz
- Kenya
- Corona

The system should capture all different writing styles of each letter from the users' handwritten texts.

In order to evaluate the accuracy, i.e., the performance of generated handwriting style, appropriate performance measures must be defined by the development team and the system should compute the accuracy by using the defined accuracy metrics.

5.2. NFR-2. High Reliability of the Server

The server side of the system should be developed and operated with high reliability since all the mobile clients access the SThT functionality through the server. Therefore, the system should provide a high level of reliability in terms of its sub-quality attributes as defined.

- High Maturity of the Server
  The system should avoid the failure in case of fault occurrences.
- High Fault Tolerance of the Server
  The system should detect various types of faults and tolerate the faults to continue provide the system functionality with the specialized level of performance.
- High Recoverability of the Server

The system should be designed to be recovered for its software and the transaction dataset in effective and efficient manner.

6. Guidelines for Conducting CEP

6.1. Guidelines for Designing the Architecture

Apply the following guidelines for writing CEP Reports.

- **Conformance to the given SRS**

  The submitted AD should conform to the given SRS.

- **Conformance to UML Standards**

  The submitted AD should conform to the notational and usage standards of UML.

- **Consistency among various Artifacts in AD**

  There should be a high consistency among various artifacts (such as diagrams) in the submitted AD.

- **Comprehensibility of Textual Description**

  The textual elaboration of the architecture design should be written in accurate, precise, and condensed way. Hence, the understandability of the AD becomes high. The textual description can be written in English, Korean, or their mixture.

- **Readability of Figures and Tables**

  The figures and tables should be easily readable by applying good formats, right font size and special effects on them. For example, a use case diagram with 100 use cases should be well structured and enlarged if needed.

- **Reasonable Details of Machine Learning design**

  The submitted AD would include a design for managing machine learning models. The description of the machine learning model generation should be written in reasonably details. The description typically includes machine learning algorithms utilized, training sets used, the details for designing the model generation components.

- **Originality of the AD**

  The submitted AD should be an individual work. Any same or highly similar solutions would get a score penalty.

6.2. Guidelines for Submitting Reports

Apply the following guidelines for writing CEP Reports.

❑ **Due Dates for Submission**

The due dates and times for each CEP report are specified. The CEP reports should be submitted by the due. Late submissions of CEP reports are not accepted.

❑ **Format of the CEP Report**

Use the word processor, MS Word, for formatting your CEP reports. Submit the word files, not the PDF files.

❑ **Submission**

Submit your CEP report to the course manager, not to the instructor.

6.3. Evaluation Form for CEP Reports

The following form is used to evaluate the CEP report.

Evaluation of CEP Report

Name: 홍길동

Criteria	Max	Earned
Ch.1, Introduction	2	2
Ch.2, (A1) Architectural Requirement Refinement	2	2
Ch.3, (A2) System Context Analysis		
System Boundary Context	3	3
Functional Context	5	5
Information Context	5	5
Behavior Context	5	5
Ch.4, (A3) Skeleton Architecture Design		
Justification of Candidate Architecture Styles	6	6
Integrating Selected Styles into Architecture	2	2
Ch.5, (A4) Architecture with Views		
Applying Functional Viewpoint	8	8
Applying Information Viewpoint	8	8
Applying Behavior Viewpoint	8	8
Applying Deployment Viewpoint	6	6
Ch.6, (A5) Architecture with Quality-driven Design		
Design for NFR #1, Applying Process	8	8
Design for NFR #1, Quality Delivered	7	7
Design for NFR #2, Applying Process	8	8
Design for NFR #2, Quality Delivered	7	7
Ch.7, (A6) Architecture Evaluation		
Intermediate Steps of applying Evaluation	6	6
Correctness of the Evaluation	4	4
CEP Score	100	100

❑ **Strength**

❍

❑ **Weakness**

❍