Harnessing Speech-To-Text (STT) technology to translate clinical conversations into medical records.

A. Problem Statement (Current State)

Clinical documentation is an integral part of the practice. However, the documentation process is often burdensome and time consuming for clinicians and staff. For instance, during the patient's health assessment session, the therapist will gather information such as patient history, existing medication, lifestyle changes, exercises, diet, symptoms, and examination. This is then manually input into the EMR system prior to formulating an evaluation or finding. After which, a treatment plan is recommended.

Another typical scenario is the capturing of conversation amongst the care team. Multi-disciplinary rounds are held where doctors from different disciplines will discuss about the patients' condition and the clinical follow ups required. The various doctors will then complete their documentation in the EMR separately after the discussion.

Other similar scenario which involves a team of clinicians is during resuscitation; a scribe nurse receiving information from the rest of the team providing intervention to patient, will write the intervention on the whiteboard and read back the information to the team to confirm and then document them in EMR.

For front-line staff at the contact centre who receives phone calls from patients, next-of-kin or members of the public, they too need to summarise the content of the phone call. Depending on the nature of the call, staff will either enter the summary into a post-call log in the CRM system or in the EMR.

B. Challenge Statement

How might we improve healthcare through harnessing Speech-to-Text Technology (STT) to translate and summarise clinician-patient or/and clinicians' conversation/dictation accurately into medical records to increase clinician efficiency, improve overall patient care and enhance the overall healthcare delivery process?

C. What are we looking for? (Desired Outcome)

The proposed STT solutioning for this Proof of Value (POV) trial should cover the following 3 types of use cases (refer to Table 1 below);

- Type A will generally be in a clinical setting, in which the solution will capture the voice commands (primarily English language) from the user to navigate the system hands free.
- Type B will generally be in a clinical setting, in which the Ambient STT solution will capture the conversation (primarily in English or Mandarin languages) among the clinicians or/and nurses, and create a summary note in near real-time for review and edit.
- Type C will generally be in a clinical / non-clinical environment (e.g. patient's home, call centre etc.), in which the Ambient STT solution will dictate and capture the speech content (ideally multi-language -

English, Mandarin or Malay) involving clinicians/staff and non-clinical personnel (patient, next-of-kin, etc.), and create a summary note for review and edits before update into EMR.

Technology	Voice Control AI & Speech to Text Dictation	Ambient Speech-to-Text AI	
Use Case Types	**Type A**	**Type B**	**Type C**
Use Case Description	**Voice Control (Navigation)**	**Clinician & Clinician**	**Clinician & Patient**
Environment Type	Clinical Setting	Clinical Setting	Clinical or Non-Clinical Setting
Capability	• Voice command to navigate and execute software functions • Speech to text dictation	• Ambience Speech Recognition & Transcription + Summarization • Convert audio conversation into text transcriptions	
User type	Clinician(s) only		Clinician and patient
Language	Monolingual (English)		Multilingual (min. English, Mandarin or Malay)
Use Cases	1. Rehab Centre Hand Occupational Therapy	1. Emergency Department resuscitation 2. Multi-disciplinary discussion	1. Rehab Centre Hand Occupational Therapy 2. Contact Centre

Table 1: Use Case Types - For detail of the use cases, please refer to Annex A- Use Cases Summary.

1. The following key capabilities of the proposed solutions are the key shortlisting criteria:

i. **Ambience and Speech to Text Dictation** –
Ability to accurately generate medical notes from clinician/patient conversations and almost instantly convert voice conversation data into draft medical notes. In addition, ability to translate pure speech to text dictation.

ii. **Clinical Summary** –
Ability to summarise healthcare worker's/patient's conversation's captured text and create a summary note in near real-time with appropriate tagging to the clinical notes. It allows clinicians to review and finalise the notes before they are updated into the hospital's EMR system.

iii. **Voice Control Navigation** –
Ability to allow users to perform navigation and operational tasks with Epic system, or with other EMR system via voice commands.

iv. **Multi-Languages** –
Ability to support multi-languages (both English and Mandarin at the minimum). Added advantage if the solution could support other mixed languages (e.g. Malay etc).

v. **Medical Vocabularies** –
Ability to support in-built medical vocabularies to capture medical terminologies.

vi.   **Epic Integration** –
Solution providers need to have either proven track records in Epic integration or express explicit commitment to establish Epic integration by the time of actual implementation of the solution after the trial. For the latter, it is the sole responsibility of the solution provider to work directly with Epic to get their integration endorsed by Epic.

vii. **AI-Safety** –
Added advantage if there is AI safety services to detect and block sensitive health information and personal identifiable information.

viii. **User Friendliness –**
The solution is intuitive for users to use without elaborate setup or configurations.

3. The proposed STT solution for this Proof of Value (POV) trial shall fulfil the following Technical Requirements:

   i.   The architecture design pattern addresses cybersecurity & data sovereignty, and data residency requirements. Solutions must comply with cybersecurity standards, especially when there is personal and clinical data involved (Refer to Annex B: Table of Security and Infrastructure Requirements and Controls for Trial & Pilot Environment).

   ii. For STT Solutions that leverage their proprietary Large Language Model (LLM) hosted at their native location for the processing of audio to text and summarising capabilities, the solution provider may have to demonstrate that the audio and transcript text file do not reside at their native location.

4. Overall collaboration requirements:

   i.   **Scalable:** The proposed solutions should consider a broader plan to scale across other healthcare institutions and settings, and consider the need to integrate and interface with the various relevant systems.

   ii. **Cost-effectiveness:** The proposed solutions must be cost-effective and beneficial to the public healthcare in Singapore.

   iii. **Implementation Timeline:** This POV collaboration is scoped to address the challenge statement and the trial completion period is to be capped at not more than 3 months (including the setup of the trial).

   iv. **Cost:** The proposed solution should include indicative cost (if applicable) for the trial.

Annex A: Use Cases Summary

S/No.	Use Case Description
1	Rehab Centre Hand Occupational Therapy (OT) – The time allocated for each hand OT session is 15 minutes. Each session involves evaluation, discussion, and implementation of treatment plans with the patient. The hand occupational therapist measures, records range of motion at relevant joints and writes detailed progress notes into the EMR system to document updates and treatments performed. Entering this information into the system can be time consuming, and hinders further thorough evaluations, treatment discussions and interventions.
2	Emergency department resuscitation care team involves a team of doctors, nurses, and respiratory therapists playing different roles to provide effective resuscitation. The role of a scribe taken up by a nurse/doctor, is to receive information from the rest of team regarding the interventions provided to patient and to retrospectively transcribe into EMR. Often, the scribe has difficulty due to multiple factors. This would include a chaotic environment with multiple ongoing conversation by many different personnels, multiple ongoing activities that need documentation and the frequency in which the interventions are being administered.
3	Multi-disciplinary rounds are held involving healthcare workers from different disciplines to discuss about the patients' condition and the clinical follow ups required. The various healthcare workers will then complete their documentations in the EMR separately after the discussion. There could be inconsistency due to different recollection or understanding. Significant time is also spent on performing the documentation in the EMR
4	Contact Centre and staff manning helplines will receive phone calls from patients, next-of-kin or members of the public. Depending on the nature of the call, the staff will either enter the summary into a post-call log in the CRM system or in the EMR. The proposed solution should be able to assist with transcription and summarising the conversation. In addition, it should ideally be able to recognize key words, classify and triage requests, and make simple, relevant recommendations to the staff.

Annex B: Table of Security and Infrastructure Requirements and Controls for Trial & Pilot Environment

	A. Oversea Hybrid SaaS solution	B. SG local SaaS solution
1. **Pilot Trial**	(a) **Added advantage** to provide SOC2 Type 2 certification and other relevant reports*	
2. **Production**	(a) **Mandatory** to provide SOC2 Type 2 certification and other relevant reports* (b) **Mandatory** to have certification and management practices for safeguarding and securing data (e.g., ISO/IEC 27001, 27017, and 27018). (c) **Added advantage** to reduce data risk (e.g., by removing non-necessary data fields from use; or de-identify PII info). (d) **Added advantage** to have Confidential Computing~ as part of the solution design.	(a) **Added advantage** to provide SOC2 Type 2 certification and other relevant reports* (b) **Added advantage** to have certification and management practices for safeguarding and securing data (e.g., ISO/IEC 27001, 27017, and 27018). (c) **Optional** to reduce data risk (e.g., by removing non-necessary data fields from use; or de-identify PII info). (d) **Optional** to have Confidential Computing~ as part of the solution design.
3. **Applicable for Pilot & Trial**	**Additional risk controls such as:** (a) All data stores are logically partitioned per-customer and no sharing of data. (b) Provision of security logs and to be made available to subscriber. (c) Activity logfiles are expected to be monitored and alerts sent to Service management team.	
	(d) Evidence to show and ensure conversation history is not logged or cached at oversea location / storage.	

* Secure Code Review; Pen Test; VA Scan reports.

~ encrypting data-in-use, transient data encrypted in memory during processing.

Participate in this challenge statement by submitting a proposal via email to healthx@synapxe.sg. Have any questions? Check out our FAQ section.