Building AI Solutions Based on Essential Needs, Not Just Hype (Part 2)

This is the second blog post in a three-part series exploring the development and deployment of large language models for community health workers in low-resource settings.

"Is this much nausea normal during my pregnancy?" This is one of many questions Sunita, a community health worker (CHW), tries to answer. She relies on her experience but remembers this lady had a previous miscarriage. With limited time, she sends a message to her WhatsApp group of fellow CHWs, receiving different answers from everyone. This scenario is typical for these workers, who often depend on informal methods beyond their scheduled training sessions, which adhere to a fixed curriculum and may thus not adequately address their most urgent needs.

A typical day for Sunita involves walking to multiple field visits in her community. As part of the world's largest female community health workforce — Accredited Social Health Activists (ASHAs) — she is known as a trusted frontline health worker. "ASHA" means "hope" in Hindi, and this army of one million women supports the maternal health needs of 25 million women in India. Sunita is the first to know when a woman misses her period, becomes pregnant, or is about to deliver. Many women call her "*didi*," which means "older sister" in Hindi.

Despite only being educated up to Grade 10, Sunita is the go-to person for maternal health questions. She often answers based on her experience or the input of other ASHAs, rather than having real-time, evidence-based guidance. This is the challenge we aimed to solve with a ChatGPT-based chatbot.

As ChatGPT was all over the news, we were excited by the possibilities. Our multidisciplinary team was eager to incorporate ChatGPT into our existing <u>SMARThealth Pregnancy</u> app, which is part of a trial running in Telangana and Haryana, India. We thought the 144 ASHAs involved in this program could surely benefit from it. However, as we soon learned, there were many challenges and lessons ahead.

Communicating as an Interdisciplinary Team

As someone with a clinical background, I was initially lost during one of our early meetings with the software development team. They were using terms like 'temperature' and 'zero-shot,' which sounded foreign to me. I soon learned that 'temperature' in natural language processing was not about body heat.

I was in charge of leading the fieldwork and clinical validation in India, but I needed to understand how we could bridge the gap between the capabilities of conversational AI and the real needs of ASHAs. I reached out to our data scientists, and one of them, who had just completed a course on large language models, eagerly shared her notes with me. She created a presentation to help non-programmers like me understand concepts like 'chunk size' and 'fine-tuning.'

This experience was not unique. While the software team didn't fully understand what an ASHA worker does, the field team in India did not grasp the technical programming concepts. Bridging this gap early in our development process was crucial for creating a functional chatbot for ASHA workers. The next challenge was figuring out how we would gather the frequently asked questions ASHAs needed answers for.

Getting Our Priorities Right

Our team was excited to use Retrieval-Augmented Generation (RAG) to develop a custom chatbot for community health workers. This would ensure that the chatbot only provided responses based on Indian guidelines. But what type of questions would it need to answer? The possibilities seemed endless.

One exciting capability of ChatGPT is its ability to generate potential user questions from a given document. We termed these "synthetic" user questions. While this was exciting, we realised that we needed to focus not on what was exciting, but on what was really needed. So, we went into the field.

We had some formative understanding of ASHA needs from previous discussions in the SMARThealth Pregnancy study. Using this, we created a topic guide and frequently asked questions. We then organised a "Pile Sorting Exercise," a qualitative research method that helps participants sort information into categories. We translated 120 questions, printed them on small cards, and distributed them to groups of ASHAs. They sorted the questions into "often asked," "sometimes asked," and "never asked."

We also demonstrated the chatbot prototype through a short video in their language. During group discussions, ASHAs shared more questions they found difficult to answer. This workshop confirmed that our chatbot was needed, and we gathered nearly 200 real-world questions from ASHAs across two states. Was it easy? Not really. Was it worth it? Absolutely. Now, all we had to do was get the ASHA workers to enter their questions into the prototype. We thought this would be the easy part, but we were wrong.

Challenging Our Assumptions

Initially, our team built a text-based chatbot for ASHAs in the SMARThealth Pregnancy study. This seemed simple—an ASHA would type her question and receive a written response. However, during our first field test, we encountered an unexpected challenge. When we asked the first ASHA to type a question, she gave us a puzzled look. The second ASHA said they never typed anything on their phones. They showed us how they used voice dictation and voice messages on WhatsApp instead. Our assumption that a text-based chatbot would work was wrong.

We quickly pivoted to developing a chatbot with voice input and a read-aloud feature. We tested this twice in the field, and in both Haryana and Telangana, ASHAs took much longer to type questions compared to using voice input. We also began collecting voice recordings to better understand how ASHAs structured their questions. Interestingly, in Siddipet (Telangana), ASHAs structured their questions better with voice, whereas in Haryana, they needed more help.

This initiative was not about using cool technology to build a chatbot — it was about making it functional and useful. The chatbot had to understand the terms ASHAs used and function well with voice input. Our goal was to use AI to enhance the existing workflow of ASHA workers, not replace them. Therefore, our development team integrated voice recognition and question-reframing features based on the field recordings.

Understanding that technology is there to serve our users and being open to challenging our own assumptions about what technology can do brought in a work culture within our team that was key to setting the stage for our prototype development.

The author(s) would like to thank Dr. D Praveen, Prof. Jane Hirst, Dr Huiqi Yvonne Lu, Uday Adavi Neelanjana Das, Sreya Majumdar and Sridevi Gara for their invaluable support and contributions throughout the development of this blog.