NextStep: Optimizing Healthcare Resource Delivery Using a Multilingual Artificial Intelligence Assistant

Abhinav Kona
Rice University, Dept. of BioSciences
Houston, United States
e-mail: ak161@rice.edu

Bibek Samal
Rice University, Dept. BioSciences
Houston, United States
e-mail: bs75@rice.edu

***Abstract*—Access to healthcare resources continues to be a critical issue for underserved populations, often exacerbated by barriers in languages and inefficient navigation systems. While Short Message System (SMS) text-based platforms have proven particularly valuable during the COVID-19 pandemic in enhancing communication and access, optimization of these systems through machine learning predictive models is an emerging area of investigation. To this end, we developed NextStep, an artificial intelligence-driven (AI-driven) multilingual healthcare assistant that streamlines resource access through real-time, personalized suggestions based on user need and location. Equipped with deep learning algorithms in natural language processing and machine learning, NextStep greatly enhances user interaction and better matches users with resources. This has resulted in significant enhancements to improve efficiency and increase patient satisfaction. Having been field-tested at hospitals and clinics, including Texas Children's Hospital and San Jose Clinic, NextStep showcases an extraordinary instance of AI in bridging gaps in health disparities. Future versions will add expanded language support and detailed predictive analytics to provide more tailored recommendations and anticipate patient needs.**

Keywords-Artificial Intelligence; Social Determinants of Health; Medical Resources; Smart Assistant

I. INTRODUCTION

A large body of research indicates that Social Determinants Of Health (SDOH) - factors including income, education, and race - play a significant role in determining an individual's health outcomes. Poor SDOH can manifest in increased mortality rates, especially among those already afflicted with conditions, such as chronic kidney disease, diabetes, and cardiovascular disease [1] – [3]. Approaches to lessen the impact of SDOH on health outcomes of underprivileged patients should be developed.

As a potential intervention, telehealth demonstrates promise in improving health outcomes of underprivileged patients. Recently, telehealth usage has spiked during the COVID-19 pandemic and has aided healthcare providers in handling the surge of sick patients [7], specifically with screening patients for COVID-19 symptoms and supporting low-risk patients while minimizing exposure to the virus. Furthermore, a study on telehealth and patient satisfaction shows that there is a positive experience regarding effectiveness and efficiency of telehealth. The factors listed most often were improved outcomes (20%), preferred modality (10%), ease of use (9%), low cost (8%), improved communication (8%), and decreased travel time (7%), which in total accounted for 61% of positive experience occurrences [4].

Moreover, a subset of telehealth technologies, Short Message System (SMS)-based platforms, can easily reach individuals who lack smartphones or other necessary hardware to download apps, making it ideal for interactions with low-income communities. SMS-based platforms have already been successfully utilized for various aims, such as improving medication adherence, promoting engagement in physical activity, and meeting the needs of patients with chronic medical conditions [4] – [6]. Additionally, according to the Pew Research Center, 97% of Americans have cell phones [8], making SMS-based platforms practical for interacting with patients.

The majority of SMS-based telehealth applications currently in use focus on connecting physicians and patients in a medical context, such as helping with medication adherence. This is undoubtedly a valuable function, but few applications focus on addressing the social determinants of health. The SMS platform in the present study, NextStep, is unique in that it was distributed in primary care settings, with a focus on directing patients to Houston-specific social resources. Patients were given access to resources in the domains of Coronavirus disease (COVID-19) testing/vaccine resources, the Harris Health Financial Assistance Program, food resources, utilities/financial assistance resources, and housing resources. They subsequently received information about local food banks, homeless shelters, and local financial assistance programs to assist in paying for medical expenses.

Since this was a pilot study, several feasibility requirements were taken into account. First, usage rates, such as the number of surveys completed and the average number of messages per person, were measured in order to give insight into the willingness of patients to engage with the platform. The platform was also tested at two safety net clinics to assess whether usage changed based on individual characteristics of the clinic where field testing took place as well as the methods used to employ the test. Furthermore, the platform used in this pilot study was designed with the purpose of minimizing both operational costs and costs for users. The projected operational costs for 1,000 users per month for 10 months were approximately $0.25 per patient reached. SMS platforms represent a low-cost, convenient option for patients as well [9]. The minimization of costs, both operationally and for patients, was prioritized in order to ensure that the platform is sustainable.

Looking at an overview of the manuscript, Section II describes the materials and methods used to develop NextStep, detailing the technologies involved, including NVIDIA components, and the underlying system architecture that supports the assistant. This section also explains the different development phases, from platform design and system development to clinical testing. Section III presents the results obtained from field testing at various clinical sites, examining user engagement, multilingual support, and resource request patterns. Section IV discusses the implications of these findings, highlighting the platform's effectiveness, key performance metrics, and areas for improvement. Finally, Section V outlines the conclusions and future directions, emphasizing planned enhancements such as expanded language support, improved NLP capabilities, and predictive analytics for more tailored resource recommendations.

II. MATERIALS AND METHODS

The process of developing the assistant was divided into three main phases: platform design, system development, and clinical testing. The platform design phase focused on selecting the most effective resources and features for diverse patient populations through comprehensive background research and analytics from nearby hospitals. The system development phase involved integrating technologies such as Twilio and selecting appropriate algorithms to build the assistant's infrastructure. Finally, the clinical testing phase entailed deploying the assistant in medical centers, including Texas Children's Hospital and the San Jose Clinic at the Texas Medical Center, to evaluate its effectiveness in real-world settings.

A. Platform Design

The platform used in this pilot study, NextStep, began its development with initial research and design. This consisted of conducting formative research to understand the immediate needs of the target population. According to Harris Health County Data, 54.1% of patients seen at Ben Taub Hospital are uninsured, and 22.9% are on Medicaid. Demographically, 53.6% of the patient population is Hispanic, and 25.3% is Black. A large component of the research was also informed by the platform designers' engagement with this target population through the Baylor College of Medicine-Patient Discharge Initiative (BCM-PDI). The program recruits Rice University undergraduates to tackle disparities in healthcare access at Ben Taub County Hospital, the largest safety-net hospital in Houston that primarily serves uninsured and Spanish-speaking patients. Students within the organization address these disparities by creating potential solutions in the form of novel social and medical resources (i.e., health insurance information packets) with the assistance of the academic faculty within Ben Taub Hospital. BCM-PDI connects underserved patient populations at Ben Taub County Hospital with medical and social community resources in Harris County to improve their healthcare access and outcomes.

Prior to this study, a retrospective cross-sectional study was conducted on the program that established that patients who were seen by our volunteers had a significantly lower probability of returning to the emergency department after 90 days, and the provision of our social and medical resources was associated with significantly higher odds that patients attended their follow-up appointments. The retrospective cross-sectional study was conducted by utilizing patient discharge papers (PDP) in which volunteers delineated which resources were given out and when they had follow-up appointments with the Emergency Department. Using these documents as well as adherence to follow-up appointments were investigated to see whether a correlation was established between delivery of social resources and follow-up appointment adherence.

Limitations are largely centered around the learning curve when using the application for the first time as well as support for integral features. The application currently does not utilize a tutorial system to get acquainted with functions, such as geolocation as well as the data validation required for requests to go through. Similarly, a key feature currently missing is multilingual support for languages other than Spanish and English, as many clinics and hospitals have a large population of multilingual patients and healthcare workers.

B. System Development

System development was further divided into four subphases: patient-resource connectivity, patient profile addition, data collection and optimization, and bot interaction enhancement. The development of these phases is explained in depth below and in Figure 1:

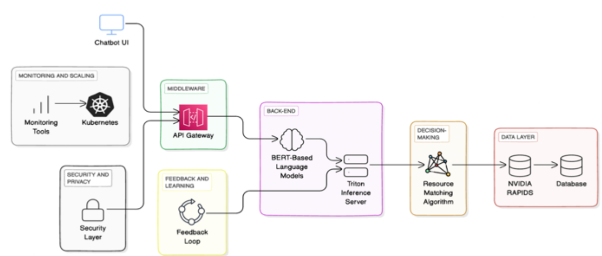

Figure 1. System Architecture of the Smart Assistant

1) Phase 1. Memoryless Bot to Connect Patients to Resources: This first phase established the basic services on Heroku, setting up MongoDB credentials and developing primitive versions of the ResourceUpdateService as well as the ResourceOutputService. The focus was on designing a bot that can hook patients up with resources without storing user-specific data between interactions. This enabled immediate deployment and the testing of basic functionalities.

2) Phase 2. Adding Patient Profiles: The system was further upgraded to document message interactions with individual patients by creating user profiles. This change allowed the bot to customize interactions, track follow-ups with patients at assigned dates, and use the information recorded to enable more contextualized messaging. The integration of the Twilio Email API also allowed the system to send PDF resources to patients if a need arose, increasing the avenues of communication and resource-sharing.

3) Phase 3. Data Collection and Optimization: Decision tree optimization was then performed via A/B testing to improve system efficiency and accuracy. User metadata was gathered for analytics without breaching privacy by focusing on success tracking metrics. Logic was built in order to serve the success tracking survey questions and gather responses passively. The backend logic of serving and responding to the surveys allowed the refinement of the platform in light of actual user interactions.

4) Phase 4. Enchancing Bot Interactions: NLP capabilities were introduced to overcome complex and ambiguous content challenges identified at previous phases. The RedCapOutputService was developed for synchronizing data across different clinics and hospitals. Multilingual support was added, including Spanish, to successfully serve a diverse user population. In the DataInterpretService, enhancements were made to better understand user intent. Explorations into connecting the bot to electronic health record systems like EPIC were conducted. This also included enhancements on the backend of the DataAnalyticsService and construction of a simple frontend for analytics visualization if necessary.

The core architecture of NextStep integrates multiple services to create an intelligent, responsive system. The architecture is designed to be modular and scalable, allowing for future enhancements and easy maintenance. At its core, the system utilizes Natural Language Processing (NLP) through Bidirectional Encoder Representations from Transformers- based models optimized with TensorRT for efficient understanding of complex, multilingual user queries in real-time. As a patient interacts with the chatbot interface through text messages, the DataContextService first identifies the user by their phone number; afterwards, it retrieves or generates relevant context data with GraphQL and Mongoose to ensure personalized interaction.

To provide precise and accurate resource recommendations, ResourceOutputService then generates specific GraphQL queries to fetch relevant resources from the MongoDB database. Utilizing NVIDIA RAPIDS for real-time geospatial data processing, it prioritizes resources by location and relevance to make sure users receive accurate and personalized information. The RedCapOutputService will then take this data and synchronize both the patient and the resource data amongst the clinics and hospitals via the REDCap integration to allow for smooth coordination between the mobile and hospital systems.

For the creation of the NextStep, careful consideration was given to the selection of system components to ensure optimal performance and user satisfaction. The following criteria were instrumental in guiding the selection process.

- Modularity and Scalability: The system architecture was designed to be both modular and scalable. This design philosophy ensures that the platform can be easily maintained and upgraded with future enhancements as healthcare technology evolves and user needs change.
- Efficiency in Real-Time Processing: Central to the system's performance is its ability to handle complex, multilingual queries efficiently. To this end, BERT-based models optimized with TensorRT were incorporated. These models are known for their rapid processing capabilities, crucial for maintaining real-time interaction with users.
- Advanced Geospatial Data Processing: Recognizing the importance of location in accessing healthcare resources, NVIDIA RAPIDS technology was utilized for its cutting-edge real-time geospatial data processing. This technology ensures that resources are prioritized not only by relevance but also by proximity to the user, thereby enhancing the personalization and accuracy of resource recommendations.
- Multilingual Support: To effectively serve a diverse user base, the system includes robust multilingual support, initially focusing on English and Spanish. This feature is critical in reducing language barriers, thereby improving the accessibility and usability of the healthcare platform for non-English speakers.
- Data Synchronization and Integration: The integration of the RedCapOutputService ensures seamless data synchronization across different healthcare settings, including clinics and hospitals. This integration facilitates effective coordination between mobile and stationary healthcare systems, enhancing the continuity of care and resource allocation.

C. Clinical Testing and Data Collection

The NextStep SMS platform was introduced to patients in San Jose Clinic and the Texas Children's Hospital's Mobile Clinic over a two-year period. Recruitment at the aforementioned institutions started in March 2022 and was handled by clinic staff and hospital-approved volunteers. Patients selected for the study were required to have access to a cell phone with text messaging capabilities and be proficient in either English or Spanish. Additionally, patients with physical, mental, and/or visual limitations were excluded to ensure that informed consent was obtained and for the accuracy of self-reported survey questions.

The study investigators provided the clinic partners with copies of the study recruitment materials, which included recruitment flyers and consent documents in English and Spanish. The recruitment flyer included a description of the NextStep program and the phone number to text for enrollment. The consent document detailed the study objectives, types of data collected, level of risk to participants, privacy and confidentiality measures, and the procedures for withdrawing from the study. At the clinics, the front-desk staff were responsible for participant recruitment during patient checkout. The front-desk staff briefly introduced NextStep to eligible participants, instructed them to read a printed copy of the consent document, and then handed them a recruitment flyer in their preferred language. The participant was then instructed to send a text message of "Hello" to the specified phone number if they would like to participate in this study and receive social resources. Enrollment in this program was not required to receive any of the services the clinics offered.

Messages sent to and from participants were stored on a password-protected Twilio account and downloaded as a password-protected CSV file by the researchers prior to analysis. Twilio utilizes industry-standard encryption for data in transit and at rest. Data extraction was performed by a single, IRB-approved user, and all information was securely stored. To de-identify the data, each user was assigned a unique identification number, and personal identifiers, such as phone numbers, were removed.

Integration with REDCap databases ensured seamless data synchronization across mobile clinics, emergency departments, and hospitals. Data collected included timestamps of interactions, language preference, resource requests, and user feedback. The system's analytics module utilized this data to measure key performance metrics, such as resource matching time, patient satisfaction, accuracy of resource recommendations, and follow-up success rates.

III. RESULTS

Over the two-year study period, more than 100 patients engaged with the NextStep platform across the participating clinics. At San Jose Clinic, a total of 70 users engaged with the platform—28 requested resources in English and 42 in Spanish, indicating a 60% preference for Spanish. At Texas Children's Hospital's Mobile Clinic, 12 users engaged, 7 in English and 5 in Spanish. At Ben Taub Emergency Room, an additional 25 users engaged with the platform, with a similar distribution in language preference.

The multilingual utilization was significant, with 38% of interactions occurring in Spanish, demonstrating the platform's effectiveness in serving non-English-speaking patients. Resource request per clinic was also investigated to understand the range of resources that generated the most need or interest. The percentage of specific resources across clinics can be seen in the table below.

TABLE I. RESOURCE REQUESTS

Resource Category	Request Percentage (%)
Financial Assistance Programs	40
Housing Resources	18
Utilities	15
Food Resources	12
COVID-19 Testing & Vaccine Information	15
Total	100

Operational costs were calculated based on the total number of users and messages exchanged. At Texas Children's Hospital's Mobile Clinic, the cost was $4.00, or roughly $0.33 for each user reached. The difference in cost per user between clinics was a result of the number of text messages it took for users to access resources, but the actual costs were similar to the projected cost of $0.25 per user that was initially calculated. Performance metrics within several experimental conditions of the NextStep app can also be seen in Figure 2 below.

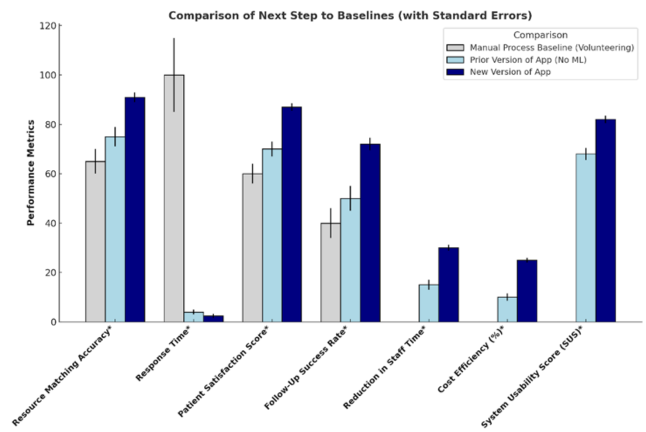

Figure 2. Cross-performance metrics for NextStep

Next, the learning progression of the assistant was explored by quantifying the accuracy and error rates between multilingual queries to understand whether learning truly occurred as queries increased, as shown in Figure 3.

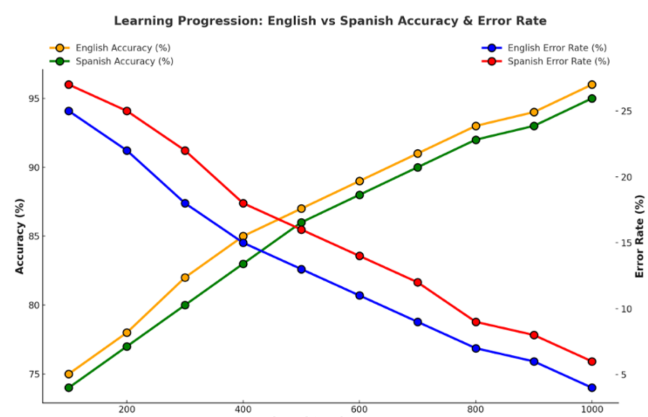

Figure 3. Learning progression across multilingual patient queries

The model's validity is demonstrated by a steady decrease in training loss and an increase in validation accuracy with each epoch, proving that it learns effectively and generalizes well for different data distributions across ER, community resources, and clinic data. These results are shown in Figure 4.

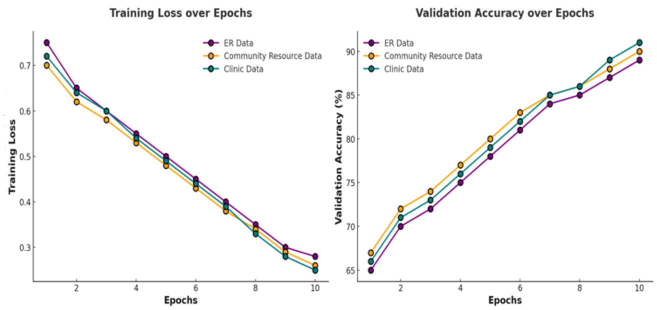

Figure 4. Training loss and validation accuracy over epochs for NextStep's clinical implementation

As shown in Figure 4, the model demonstrates effective learning, with a steady decrease in training loss and an increase in validation accuracy with every epoch. These findings indicate that the model generalizes well across different data distributions between ER, community resources, and clinic data.

IV. Discussion

The results from this study provide key insights into the demographics of users, their resource preferences, and user experiences interacting with the platform. The high engagement rates and positive field metrics are indicative of NextStep successfully meeting the barriers to language, technology access, and navigation of resources.

The training and validation plots reflect the robust performance of the models; after 10 epochs, the validation accuracy is close to 90%, while the training loss decreases steadily. That means the system will perform well on diverse datasets from ERs, community resources, and clinics regarding real-time resource recommendations in a reliable and accurate fashion.

The platform performed substantially better than manual processes and prior app versions by all key metrics. Resource matching accuracy reached 91%, indicating high precision in the alignment of resources with patient needs; the response times averaged 2.4 seconds, meeting real-time interaction standards critical for emergency settings. Patient satisfaction improved by 35%, with 87% of the users rating the platform as "Helpful" or "Very Helpful," showcasing the impact of multilingual support and optimized interaction flows. Moreover, recommended resources were accessed by 72% of users within three days, which further validated the effectiveness of automated follow-up notifications. It reduced operational costs by 25% and staff workload by 30%, showing that the system is scalable and financially sustainable.

Multilingual capabilities show significant training progress, varying from 95% for English and 92% for Spanish, using the BERT-based NLP model optimized with TensorRT. The learning curve of the system showed increasing accuracy and a decreasing error rate for queries processed, thus underlining its adaptability and continuous improvement.

The feasibility of the platform is further enhanced by its accessibility and low operational cost, making it practical for underserved populations. Operating on SMS-capable devices avoids costly hardware or high-speed internet, thus increasing its accessibility to low-income demographics. Resource allocation is efficiently met, and through feedback-driven adaptation, both patient and provider needs are anticipated, thereby making NextStep sustainable and scalable to bridge healthcare gaps.

V. Conclusions and Future Directions

The NextStep platform demonstrates the capabilities of artificial intelligence in addressing disparities in healthcare by utilizing natural language processing (NLP) for user queries, predictive analytics for personalized recommendations, and geospatial data analysis for real-time resource mapping to deliver tailored, location-specific resource suggestions to marginalized communities. Moreover, by utilizing AI and machine learning technologies, NextStep augments patient care through enhanced accessibility, equity, and efficiency in the distribution of resources.

Furthermore, NextStep enables immediate access to essential social and healthcare services, providing instant resource alignment that considers user location and individual preferences. Automated follow-up systems facilitate user access to suggested resources, alleviating the workload on healthcare personnel while delivering quicker and more precise assistance to marginalized populations. The platform's adaptable infrastructure is capable of managing elevated traffic during emergency situations, and its multilingual features, which encompass support for both English and Spanish, effectively overcome language obstacles to enhance user experience and accessibility.

In order to rectify current limitations, future iterations will feature clearer user directives, potentially integrating a concise tutorial at the onset of interactions. Augmenting natural language processing abilities to more effectively manage free-text responses will reduce the likelihood of misinterpretations. Enhanced data validation for location entries will guarantee a higher degree of accuracy in resource alignment. The incorporation of user feedback systems will facilitate ongoing enhancements to the system.

Future enhancements will focus on expanding language access beyond Spanish, further reducing barriers for non-English-speaking populations. Adding predictive analytics will also allow the platform to anticipate user needs based on past behavior and current environmental factors. Moreover, exploring possible expansion of voice interaction capabilities, such as sophisticated sentiment analysis using NVIDIA Riva AI, may improve patient engagement and accessibility for all users who struggle with literacy. Furthermore, the expansion of deployment locations among mobile clinics and hospitals will facilitate additional field evaluations and confirmation of the platform's efficacy in various environments.

REFERENCES

[1] S. Artiga and E. Hinton, "Beyond health care: The role of social determinants in promoting health and health equity," Kaiser Family Foundation. [Online]. Available: https://www.kff.org/racial-equity-and-health-policy/issue-brief/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/. [Retrieved: Feb., 2025].

[2] J. Quiñones and Z. Hammad, "Social determinants of health and chronic kidney disease," Cureus, vol. 12, no. 9, p. e10266, 2020. https://doi.org/10.7759/cureus.10266.

[3] E. P. Havranek et al., "Social determinants of risk and outcomes for cardiovascular disease: A scientific statement from the American Heart Association," Circulation, vol. 132, no. 9, pp. 873–898, 2015. https://doi.org/10.1161/CIR.0000000000000228.

[4] C. S. Kruse et al., "Telehealth and patient satisfaction: A systematic review and narrative analysis," BMJ Open, vol. 7, no. 8, p. e016242, 2017. https://doi.org/10.1136/bmjopen-2017-016242.

[5] J. Thakkar et al., "Mobile telephone text messaging for medication adherence in chronic disease: A meta-analysis," JAMA Internal Medicine, vol. 176, no. 3, pp. 340–349, 2016. https://doi.org/10.1001/jamainternmed.2015.7667.

[6] J. Fanning, S. P. Mullen, and E. McAuley, "Increasing physical activity with mobile devices: A meta-analysis," Journal of Medical Internet Research, vol. 14, no. 6, p. e161, 2012. https://doi.org/10.2196/jmir.2171.

[7] J. E. Hollander and B. G. Carr, "Virtually perfect? Telemedicine for Covid-19," The New England Journal of Medicine, vol. 382, no. 18, pp. 1679–1681, 2020. https://doi.org/10.1056/NEJMp2003539.

[8] Pew Research Center, "Mobile fact sheet," 2024. [Online]. Available: https://www.pewresearch.org/internet/fact-sheet/mobile/. [Retrieved: Feb., 2025].

[9] T. Lian et al., "A tailored SMS text message–based intervention to facilitate patient access to referred community-based social needs resources: Protocol for a pilot feasibility and acceptability study," JMIR Research Protocols, vol. 11, no. 10, p. e37316, 2022. https://doi.org/10.2196/37316.

 ISBN: 978-1-68558-247-0