Datathon Protocol Title:

Co-creating AI Solutions for Douala General Hospital

16 June, 2025

Co-creating AI Solutions for Douala General Hospital

1. Background

The healthcare system in Cameroon, like many in Sub-Saharan Africa, faces challenges that hinder both patient experience and operational efficiency. Limited adoption of digital technologies constrains its ability to collect and analyse patient feedback, automate critical functions such as appointments, medication reminders, and support data-driven decision-making. Studies by (1,2) revealed substantial gaps in digital health governance, which limit the potential of digital innovations.

These issues are exacerbated by broad resource constraints, including budget constraints, inadequate infrastructure, and unreliable IT connectivity, making it critical to develop relatively affordable and adaptable health technology solutions (3,4). Cameroon's vast linguistic diversity, which includes French, English, and other indigenous languages, requires culturally sensitive and multilingual systems that provide equal access and user engagement (5,6).

Healthcare worker shortages further compound the problem. With a ratio of just one physician per 1,000 people in Cameroon (7). Physicians often cannot provide comprehensive, personalised explanations about diagnoses, treatment plans, or medications, leading to gaps in patient understanding and adherence (8). Similarly, traditional, manual approaches to blood inventory management often result in inefficient forecasting and stock imbalances, risking shortages and the wastage of critical blood products (9,10).

Emerging technologies such as Artificial Intelligence (AI), Natural Language Processing (NLP), and Generative AI (GenAI), including Large Language Models (LLMs), present significant opportunities to address these issues. These tools can automate communication, personalise patient education, forecast resource needs, and extract actionable insights from unstructured data (11,12). Healthcare systems can create scalable and adaptable digital solutions that meet specific operational requirements using an adaptable structure approach.

A Datathon is an ideal platform for innovation. Datathons, which bring together local experts in data science, engineering, and IT, promote rapid prototyping of AI-enabled health technologies adapted to local requirements (13). They also promote long-term capability by providing hands-on, interdisciplinary experience and encouraging collaboration among hospitals, universities, NGOs, and the commercial sector.

This Datathon aligns closely with Douala General Hospital's (DGH) strategic targets, which include increasing patient-centered care, optimising hospital operations, and strengthening communication via digital platforms. It aims to develop functional, real-world digital prototypes that can be immediately valuable to DGH while achieving the Data Science Without Borders (DSWB) objective. Through focused challenge tracks, we aim to co-create AI-driven solutions that address critical healthcare challenges in DGH.

This two-month intensive Datathon will focus on developing:

1. Patient feedback and reminder management system
2. Large language model for enhanced patient education and diagnostic support
3. AI-enhanced blood bank stock monitoring and forecasting system

Each solution is designed for independent deployment or seamless integration into a cohesive, scalable digital platform tailored to DGH's needs.

2. Track 1: Patient feedback and reminder management system

Challenge description

Douala General Hospital (DGH) lacks a unified, multilingual platform to capture and act on patient feedback or send automated reminders for appointments and medications. This track addresses the critical need for systematic patient feedback collection and automated communication systems. This challenge aims to design and prototype an integrated system that functions in DGH, supporting French, English, and indigenous languages such as Douala, Bassa, and Ewondo.

2.1 Multilingual patient feedback interface

This interface will allow patients to easily provide feedback, regardless of their language or literacy level. This interface must support multiple input methods, including text and voice, and provide visual rating options such as stars or emojis for users with limited literacy. To build this, participants can use React.js or Flutter for an intuitive and responsive web/mobile frontend, while voice inputs can be captured and transcribed using Google speech-to-text. Translations between languages can be handled using open-source models such as the Google Translate API, with the option to fine-tune these tools using local language datasets to ensure accuracy. The interface must be lightweight enough to function in low-bandwidth settings.

2.2 Feedback analysis engine

Once feedback is collected, it must be processed and transformed into actionable insights through a feedback analysis engine. This component will use natural language processing (NLP) and machine learning algorithms to automatically classify feedback as positive, negative, or neutral, extract recurring topics such as long wait times and flag urgent issues for administrative review. This allows DGH staff to focus on making data-driven improvements without manually sifting through thousands of comments. Participants can implement this functionality for French and English analysis using libraries like spaCy, TextBlob, or Hugging Face Transformers, such as CamemBERT. The backend for this service can be developed using Python frameworks like FastAPI or Flask, which enable efficient, scalable API development and integration.

2.3 Automated patient reminder system

A critical third component is the automated patient reminder system, designed to reduce missed appointments and improve adherence to treatment regimens. This module should send personalised appointment and medication reminders via SMS or voice call, depending on patient preferences. The reminders should be available in multiple languages and tailored to specific patient contexts. Messaging services such as Twilio, RapidPro, or Vonage can be integrated. Reminders should be stored and tracked in a secure Google Cloud.

2.4 Real time hospital performance dashboard

The solution must include a real time hospital performance dashboard that offers DGH administrators a clear overview of key metrics. This dashboard should present visual summaries of patient satisfaction trends, most frequent feedback themes, and the effectiveness of reminders. A clean, responsive interface can be built using React.js and data visualisation libraries like D3.js or Plotly Dash. The dashboard should support data filtering by period, language, or department, and allow exports for reporting purposes. It should include basic role-based authentication for secure access, ensuring sensitive patient data is only visible to authorised users.

The system must be designed with data privacy and security in mind. APIs developed using RESTful design principles will allow each module to function independently. However, each module will be integrated within one system.

3. Track 2: Large Language Model for Enhanced Patient Education and Support

Challenge description

High patient load and limited physician time result in inadequate patient education about their diagnoses, treatments, and medications. Hence, this challenge invites participants to design and prototype a conversational AI system powered by large language models (LLMs) that can function as a virtual patient assistant.

3.1 Patient support chatbot

The core of this solution is a patient support chatbot powered by an open-source large language model such as LLaMA, Mistral, or Google Gemini. The chatbot must be able to interpret patient queries written or spoken in natural language and return medically accurate, culturally appropriate responses that are easy to understand. To achieve this, participants can use LangChain to manage conversation flow and retrieval-augmented generation (RAG) logic, enabling the chatbot to access and explain static clinical summaries or public health information. The frontend interface of the chatbot should be developed using React.js (for web) or Flutter (for mobile), ensuring accessibility across devices. It should offer a clean, intuitive user experience and support both typed and spoken input, enabling patients of varying literacy and digital fluency levels to engage with the tool. This chatbot must be equipped with multilingual capabilities. For backend deployment, cloud infrastructure such as Google Cloud running on local servers can be used to ensure scalability and uptime.

3.2 Diagnostic and therapeutic explanation

This module equips the chatbot with the ability to explain patient diagnoses, treatments, and prescribed medications in layperson language. It will not generate clinical decisions or diagnoses, but interpret the clinician's written summaries to answer patient questions, such as *what does my diagnosis mean?* or *how should I take this medication?* The system should also explain lifestyle recommendations, follow-up care plans, and common side effects in a compassionate, personalised tone.

In order to support this, participants can fine-tune or prompt-engineer LLMs using synthetic medical data relevant to the DGH context. Models like Mistral or LLaMA-2 can be enhanced through retrieval-based methods using tools like LangChain, allowing the chatbot to pull from curated medical content.

3.3 Platform testing, evaluation, and clinical review

The chatbot must undergo usability testing, accuracy evaluation, and clinical content validation in order to validate the tool's effectiveness. Participants should compare chatbot-generated answers to those provided by clinicians using structured checklists or questionnaires. This assessment should focus on whether the chatbot's responses are medically accurate, clearly understandable, and empathetically delivered.

4. Track 3: AI-Enhanced Blood Bank Stock Monitoring and Forecasting System

Challenge description

This track aimed at developing and validating a real-time AI-enhanced monitoring dashboard that enables DGH blood banks to track inventory levels accurately, forecast future demand, and optimise stock management, thereby improving efficiency, reducing waste and shortages, and potentially improving patient outcomes.

4.1 System integration and data ingestion layer

The foundation of this system lies in its ability to integrate seamlessly with existing hospital information systems, particularly the District Health Information System 2 (DHIS2). The data pipeline must support real-time ingestion of key variables, including donor demographics such as age, sex, occupation, clinical data such as blood type, screening results, and operational metrics such as donation dates, stock movements, and expiry dates. For secure, scalable, and reliable performance, participants can build this layer using RESTful APIs developed with Java Spring Boot or Python FastAPI, ensuring compatibility with other systems.

4.2 Forecasting and AI modelling framework

The forecasting engine predicts future blood demand based on clinical and operational indicators. This module will combine classical time series techniques, such as Seasonal-Trend Decomposition using Loess (STL) and ARIMA, with advanced machine learning models including XGBoost, Random Forests, and Neural Networks built using libraries like TensorFlow and scikit-learn.

Feature engineering should be guided by exploratory data analysis and domain expertise, with predictive variables such as blood type usage patterns, patient diagnoses, department needs, day-of-week effects, and consumption trends. Variable selection methods like Lasso regression and feature importance rankings from tree-based models can be used to prioritise inputs and optimise model performance. All trained models should be deployed in an API accessible format to support integration with the frontend dashboard.

4.3 Real time monitoring dashboard

This module involves building an interactive, user-friendly dashboard that allows blood bank staff and hospital administrators to monitor inventory levels in real time. The dashboard should present stock levels segmented by blood type and status, such as available, reserved, near expiry, and time to expiry, using colour-coded visual indicators such as red for expired, green for good and yellow for new expiry.

The participants should use React.js to build a responsive and accessible interface and integrate visual libraries like D3.js, Plotly, or Chart.js to create intuitive visualisations, including histograms, trend lines, and shelf-life indicators. The dashboard should support data filtering by date, blood type, and location and provide exportable reports for planning and audits.

4.4 Inventory optimisation

The inventory optimisation module synthesises forecasting outputs and real-time stock data to generate recommendations for optimal ordering quantities. It should account for routine delivery cycles, safety stock levels, wastage rates, and the cost implications of emergency replenishments. The goal is to enable decision-makers to balance clinical preparedness and financial constraints. This module can be built using linear programming or reinforcement learning frameworks, with optimisation models implemented in Python using libraries such as PuLP or SciPy. It should generate dynamic suggestions based on forecasted shortages, expiration risks, and usage trends, enabling proactive planning and reducing last-minute stockouts.

4.5 Evaluation

An evaluation strategy must be implemented to ensure the system's usability, accuracy, and relevance to DGH's operational workflows. Participants should conduct usability testing sessions with blood bank staff using interactive prototypes, followed by structured feedback collection through surveys, interviews, and walkthroughs.

5. Expected Outcomes

The expected outcome of Datathon will be the successful development and deployment of a cost-effective, multilingual patient feedback system, an LLM-powered patient support chatbot that enhances patient education, improves understanding of diagnoses and therapeutic plans, and increases adherence to treatment through accessible, personalised, and empathetic communication. A software application that will enhance accuracy in forecasting blood demand using AI models.

6. Implementation framework

6.1 Datathon structure and timeline

The Datathon will be organised as a comprehensive three-phase initiative to maximise immediate outcomes and long-term impact (Table 1). The pre-event/preparatory phase will span four weeks and focus on participant recruitment, team formation, and resource preparation. This phase will involve active outreach to Cameroonian universities, technology hubs, and healthcare networks to identify and engage talented individuals across multiple disciplines. Participants will be provided with the necessary resources.

The Datathon sprint will span one week and focus on collaborative work. To maximise accessibility, it will combine in-person activities at DGH or partner institutions with virtual participation options. This phase will feature expert mentorship, regular progress check-ins, and prototype demonstrations to a distinguished panel of judges, including the DGH, technical, and other partners.

The post-Datathon will extend for three weeks, focusing on prototype refinement, integration planning, and pilot deployment preparation. This phase is critical for translating Datathon outputs into real-world impact, with selected teams receiving continued support and resources to develop their prototypes toward production readiness.

6.2 Participant engagement and team formation

The success of this datathon depends on assembling diverse, interdisciplinary teams that combine technical expertise with a deep understanding of healthcare challenges and local contexts. Recruitment efforts will target data scientists, software engineers, developers, and designers. Special emphasis will be placed on ensuring gender diversity to bring varied perspectives to solution development.

Team formation will be facilitated through pre-event workshops and networking sessions that allow participants to identify complementary skills and shared interests. Each team should ideally include members with technical development capabilities, healthcare domain knowledge, and an understanding of local cultural and linguistic contexts. The DSWB's technical team will provide mentorship and guidance on technical implementation and challenges.

6.3 Resource Provision and Technical Infrastructure

Comprehensive resource provision is essential for enabling participants to develop and deploy prototypes within the Datathon timeframe. Computing resources will include access to cloud platforms such as Google Cloud. Data resources will include carefully prepared synthetic datasets that protect patient privacy while providing realistic scenarios for solution development.

6.4 Evaluation criteria and success metrics

Evaluation of Datathon outcomes will employ both quantitative and qualitative measures to assess immediate results and long-term potential. Technical evaluation criteria will include prototype functionality, scalability, user interface design, and integration capabilities. Healthcare relevance will be assessed based on alignment with DGH operational needs, potential for improving patient outcomes, and feasibility of implementation within existing hospital systems. Innovation and creativity will be evaluated based on the novelty of approaches, effective use of AI and machine learning techniques, and potential for broader application beyond DGH.

5.5 Post-Datathon integration and deployment strategy

The post-Datathon phase represents a critical transition from prototype development to real-world implementation, requiring careful planning and continued engagement with successful teams. A comprehensive integration workshop will combine teams with complementary modules to ensure technical compatibility and seamless data flow between system components. This workshop will address both technical integration challenges and operational workflow considerations.

Pilot deployment planning will involve close collaboration with the DGH-DSWB and IT teams to select the most promising prototypes for real-world testing. Continuous monitoring and evaluation during pilot phases will provide feedback for iterative improvement and refinement of the solutions.

Long-term sustainability planning will focus on transitioning maintenance and development responsibilities to local teams and establishing partnerships with other hospitals. The goal is to create a sustaining innovation ecosystem that can continue to develop and deploy health technology solutions beyond the initial Datathon. This Datathon will provide immediate value to DGH and build lasting capacity for health technology innovation in Cameroon.

6. References

1. Bediang G. Implementing Clinical Information Systems in Sub-Saharan Africa: Report and Lessons Learned From the MatLook Project in Cameroon. JMIR Med Inform. 2023 Oct 18;11:e48256.

2. Ngwakongnwi E, Atanga MBS, Quan H. Challenges to implementing a National Health Information System in Cameroon: Perspectives of stakeholders. J Public Health Afr. 2014 Mar 17;5(1).

3. World Health Organization. Global Strategy on Digital Health 2020-2025. World Health Organization; 2021.

4. Cossy-Gantner A, Germann S, Schwalbe NR, Wahl B. Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? BMJ Glob Health. 2018 Jul 1;3(4).

5. Agboryah BENM, Ndip VA, Ngomba AV, Tazinya AA, Adiogo D. Factors associated with the use of digital health among healthcare workers in the Buea and Tiko health districts of Cameroon: a cross-sectional study. Pan African Medical Journal. 2024 Feb 8;47.

6. Mbah M, Bang H, Ndi H, Ndzo JA. Community Health Education for Health Crisis Management: The Case of COVID-19 in Cameroon. Community Health Equity Research and Policy. 2023 Jul 1;43(4):443–52.

7. World Bank Group. https://data.worldbank.org/indicator/SH.MED.PHYS.ZS?locations=CM. 2025. Physicians (per 1,000 people) - Cameroon.

8. Biswas A, Talukdar W. Intelligent Clinical Documentation: Harnessing Generative AI for Patient-Centric Clinical Note Generation. International Journal of Innovative Science and Research Technology (IJISRT). 2024 May 28;994–1008.

9. Stanger SHW, Yates N, Wilding R, Cotton S. Blood Inventory Management: Hospital Best Practice. Transfus Med Rev. 2012 Apr;26(2):153–63.

10. Shih H, Rajendran S. Stochastic Inventory Model for Minimizing Blood Shortage and Outdating in a Blood Supply Chain under Supply and Demand Uncertainty. J Healthc Eng. 2020;2020.

11. Bhuyan SS, Sateesh V, Mukul N, Galvankar A, Mahmood A, Nauman M, et al. Generative Artificial Intelligence Use in Healthcare: Opportunities for Clinical Excellence and Administrative Efficiency. J Med Syst. 2025 Dec 1;49(1):10.

12. Marey A, Saad AM, Killeen BD, Gomez C, Tregubova M, Unberath M, et al. Generative Artificial Intelligence: Enhancing Patient Education in Cardiovascular Imaging. BJR|Open. 2023 Dec 12;6(1).

13. Peñaloza C, Gutierrez AP, Eöry L, Wang S, Guo X, Archibald AL, et al. A chromosome-level genome assembly for the Pacific oyster Crassostrea gigas. Gigascience. 2021 Mar 1;10(3).

Table 1: Timeline

		Month							
Activities		**July**				**August**			
		Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 1**	**Week 2**	**Week 3**	**Week 4**
Phase 1: Preparatory phase									
Launch call for participants	DGH								
Finalise datasets and challenge briefs	Data Managers								
Onboard mentors	Technical team								
Deliver Pre-Datathon Workshops	DGH, Technical team								
Phase 2: Datathon sprint									
Opening ceremony and team formation	All participants, DGH, Technical team								
Intensive prototyping with mentor check-ins									
Team presentations of prototypes	Participants, Jury Panel								
Judging and awards									
Begin documentation of winning solutions									
Select the top prototypes for pilot	DGH, Technical team								
Start refinement of selected prototypes: Patient Feedback Module, LLM Chatbot, Blood Stock Dashboard	DGH, IT, Technical team								
Phase 3: Post Datathon									
Continue development and validation	DGH								
Align data flows and system architecture for LLM chatbot integration, Dashboard UI review and secure APIs	DGH								
Deployments	DGH, IT								
Write Datathon report	DGH								