**TERMS OF REFERENCE FOR TECHNICAL ASSESSMENT**

**OF THE MANDUKHAI CHATBOT**

**Type of Contract:** Professional Services/Information and Communication Technology

**Required service:** Mandukhai chatbot performance evaluation and Scalability testing

**Location:** Ulaanbaatar, Mongolia

**Language Required:** English/Mongolian

**Expected Start Date:** January 20, 2025

**Timeline:** 60 non-consecutive days from January 20 to April 2, 2025

**Background**

With the support of UNFPA Mongolia, a Facebook-based Mandukhai chatbot was launched in 2021, to deliver accurate and confidential information on sexual and reproductive health (SRH) to adolescents. This AI-powered bot was designed to increase access to SRH information and services and empower its users with the correct knowledge and informed decision-making. Since its installation at the National Center for Maternal and Child Health (NCMCH) and the appointment of a live agent for troubleshooting, Mandukhai demonstrated promising results, achieving an accuracy level of 62% in June 2021, and through ongoing enhancements to its functionality, the chatbot's accuracy improved to 73% by the end of 2022.

Since its launch, there has been limited information regarding its performance and level of accuracy in the response of the Mandukhai chatbot including user experience. Additionally, no formal evaluation has been conducted on the bot’s compliance with the data security requirements as per the Mongolia ICT Sector Medium-Term Development Policy (2022-2027), and health-related ethical considerations. Recent advancements in AI technology have the potential to significantly enhance the performance of AI-powered chatbots, even in complex interactions involving sensitive information. Given this potential, the evaluation will also focus on the aspects the chatbot has “learned” since its launch.

Therefore, UNFPA Mongolia is seeking a potential IT service provider with the appropriate expertise to perform a comprehensive technological assessment of the Mandukhai chatbot and identify areas for further improvement in the AI function of the bot. The evaluation will utilize specific criteria designed for chatbot performance metrics, focusing on accuracy in response rates, overall user experience, and responsiveness to user needs. To capture a broad spectrum of users’ experiences, the assessment will include a diverse group of users, particularly adolescents from remote areas, ensuring the dataset is representative and inclusive.

**Objective**

The purpose of the professional service is to conduct a comprehensive technical assessment of the Mandukhai chatbot by evaluating its functionality, performance, and scalability, while also identifying opportunities for enhancements to its current capabilities.

**2. Scope of consultancy work**

**1. Develop the evaluation design:** The selected company will be responsible for conducting a comprehensive desk review of existing documentation and researching best practices and standards for chatbot evaluation metrics. This includes studying use cases of advanced AI-powered chatbots in the medical field. By gathering background information and understanding the objectives, design, and any previous evaluations of the chatbot, the selected company will develop an evaluation matrix. This matrix will feature defined performance metrics derived from the past conversation logs, including accuracy or effectiveness, engagement, user interaction, operationalities such as fallback and escalation, and error and context awareness metrics.

2. **Usability Testing**: Usability testing is essential for evaluating the chatbot, aimed at ensuring that it is intuitive, efficient, and satisfying for users. The selected vendor will conduct one-on-one sessions to observe participants as they interact with the chatbot, completing specific tasks.Immediate feedback will be collected through observation, post-session questionnaires, or brief interviews. It is important to recruit adolescents from varied demographics, including differences in age, gender, education level, and geographic location, to participate in the usability testing to ensure comprehensive and relevant results.

3. **Scalability and performance testing** are crucial for evaluating how well the chatbot manages numerous queries or simultaneous conversations, providing insights into its scalability. Performance testing should encompass the chatbot's depth of understanding of the Mongolian language, AI algorithms, and conversation flow and ability to handle tones, proficiency in handling context, multi-turn conversations, and variations in user input, memory-related functions such as recalling the previous conversation and generating of dashboard and statistics. These tests will help determine the chatbot's strength and readiness to function effectively under varied and demanding usage scenarios.

4. **Organize stakeholders’ consultation** meetings to validate the findings of the chatbot evaluation and improvement plan with recommended actions for further enhancement of the bot level of AI functions

5. **Provide a final consultancy report:** The selected vendor will provide the final consultancy report of the evaluation for further improvements and required updates to the chatbot based on user interaction and scalability testing results.

**Timeline:**

The selected institution will be contracted for 60 non-consecutive days starting from January 15, 2025.

**Institutional arrangement**

The service provider (contractor) should work in close consultation with the UNFPA SRHR team and provide progress updates through weekly email as per the schedule mutually confirmed upon the start of the consultancy.

**Key Deliverables and Timeline**

|  |  |  |  |
| --- | --- | --- | --- |
| **Key deliverables** | **Format** | **Deadline** | **Installments** |
| 1. Develop an evaluation matrix based on the desk review | MS Word document, English | By January 30, 2025 | 1st instalment 30% |
| 2. A draft findings of the bot usability and performance testing analysis | MS Word document, English | By February 17, 2025. |
| 3. Scalability and performance testing results of theevaluation along with further improvement plan for enhancement of the bot AI levels | MS Word document, English | By March 10, 2025 | 2nd instalment  70 % |
| 4. Organize stakeholders’ consultation meetings to validate the chatbot evaluation and improvement plan findings with recommended actions to further enhance the bot level of AI functions. | MS Word document, English | By March 23, 2025. |
| 5. Final consultancy service reports for the UNFPA approval | MS Word document, English | By March 28, 2025 |

**COA: UNFPA.MNG.FPA90. 62300/ 71 405, MNG07SRH-23ASSEMAND74.PU0074**

UNFPA will process payments upon completion of all deliverables and acceptance by the CO.

**Requirements and Qualifications**

1. A well-organized local and international legal entity that consistently delivers high-quality work and meets deadlines. Collaboration with local counterparts should be incorporated into the team structure for international entities.
2. A minimum of two years of experience developing chatbots equipped with artificial intelligence modules.
3. Demonstrated expertise in completing similar assignments, with verifiable references or case studies.
4. Provide detailed info on the dedicated project team with the following minimum composition and qualifications:

* *Team Leader:* Minimum of 3 years of experience managing IT projects, with demonstrated leadership and communications skills.
* *AI/Data Specialist*: Bachelor’s degree in computer science, Data Science, or a related field and at least 2 years of experience in AI/ML model development, including experience with Natural Language Processing (NLP) and chatbot development.
* *Software Engineer*: Bachelor’s degree in software engineering or a related field. Minimum of 2 years of experience in software development and experience with chatbot deployment and integration.

**Documentation to be submitted to UNFPA:**

* 1. The service provider is expected to present a detailed information demonstrating their expertise in developing chatbots utilizing advanced AI tools for enhanced functionality.
  2. A comprehensive technical proposal outlining the consultancy approach, including an evaluation matrix, design framework, user research strategy for chatbot performance, proposed methodologies, scalability analysis, improvement plan, and an estimated project timeframe.
  3. A financial proposal with a detailed budget breakdown, including professional fees, out-of-pocket expenses, taxes, and other relevant costs.
  4. Supporting documents to verify the organization's experience and capacity, such as a company profile, reference letters, a list and copies of similar contracts, and other pertinent materials.
  5. CVs of team members, highlighting relevant experience with similar projects.
  6. A scanned copy of the organization’s state registration certificate.

**Criteria for Selection of the Best Offer**

The selection of a service (offer) will follow the combined scoring method, with the technical proposal weighted at a maximum of 70% and the financial proposal (price offer) weighted at a maximum of 30%.

Technical Proposal Scoring Breakdown (70% of total scoring, equivalent to 100 points):

* Expertise of the organization (NGO, company, or academic institution): 30 points
* Proposed methodology and timeliness of the implementation plan: 30 points
* Proposed team management structure and qualifications of key personnel: 40 points

The financial score will be evaluated in comparison with the lowest and reasonably-priced, technically compliant, and highest-quality offer to ensure value for money.