

*English version below*

## **Ausgangssituation**

In modernen Produktionsumgebungen werden Menschen zunehmend mit anspruchsvollen Aufgaben konfrontiert, die Fähigkeiten wie das Navigieren in großen Informationsbeständen oder die Bedienung komplexer Maschinen erfordern. Um dieser Herausforderung zu begegnen, wird im Rahmen eines Forschungsprojekts ein modularer Chatbot entwickelt, der Menschen bei verschiedenen Produktionsaufgaben unterstützt, indem er Daten abrufen und Aktionen ausführt. Um den Chatbot in einer realen Produktionsumgebung zu bewerten, ist eine strukturierte Testmethodik erforderlich.

## **Zielsetzung**

Ziel dieses Projekts ist es, eine Methode zum strukturierten Testen von Chatbots in Produktionsumgebungen zu entwickeln. Die notwendigen Schritte zur Erstellung von Testanfragen, Auswertung der Antworten und Analyse der Ergebnisse sollen beschrieben werden. Relevante Leistungsmetriken sollen definiert werden. Die Methodik soll implementiert und weitestgehend automatisiert werden. Schließlich soll die Methode validiert werden, indem sie auf den bestehenden Chatbot angewendet und dessen aktuelle Leistung getestet wird.

Relevante Forschungsfragen sind:

- Welche Methoden gibt es für das Testen von Chatbots?
- Wie sollte eine Methode zum Testen von Chatbots in Produktionsumgebungen aussehen?
- Wie kann eine solche Methode implementiert werden?
- Wie ist die aktuelle Leistung des bestehenden Chatbots?

Bei sehr guten Ergebnissen besteht die Möglichkeit, als Co-Autor an einer wissenschaftlichen Veröffentlichung mitzuwirken.

## **Anforderungsprofil**

- hohes Engagement, Interesse an generativer KI und LLMs
- selbständige, methodische und strukturierte Arbeitsweise
- sehr gute Kenntnisse in Deutsch oder Englisch
- Programmiererfahrung (optional, aber von Vorteil)

Ich freue mich auf Ihre Bewerbung, idealerweise mit einem kurzen Lebenslauf und einem aktuellen Notenauszug.

Das Projekt kann ab sofort begonnen werden.

Kontakt: M. Sc. Lasse Streibel, [lasse.streibel@iwb.tum.de](mailto:lasse.streibel@iwb.tum.de)

----- English version -----

## **Initial situation**

In modern production environments, workers are increasingly faced with demanding tasks that require skills such as navigating through large information repositories or operating complex machinery. To address this challenge, a research project is developing a modular chatbot to assist workers in various production tasks by retrieving data and performing actions. To evaluate the chatbot in a real production environment, a structured testing methodology is required.

## **Objective**

The aim of this project is to develop a method for structured testing of chatbots in production environments. The necessary steps for generating test requests, evaluating responses and analyzing the results should be described. Relevant performance metrics should be defined. The methodology should be implemented and automated as much as possible. Finally, the method should be validated by applying it to the existing chatbot and testing its current performance.

Relevant research questions are:

- What methods exist for testing chatbots?
- What should a method for testing chatbots in production environments look like?
- How could such a method be implemented?
- What is the current performance of the existing chatbot?

For very good results, there is the possibility of co-authoring a scientific publication.

## **Requirements profile**

- high commitment, interest in GenAI and LLMs
- independent, methodical, and structured way of working
- very good proficiency in German or English
- programming experience (optional but beneficial)

I look forward to receiving your application, ideally with a short CV and a recent transcript of records.

The project can be started immediately.

Contact: M.Sc. Lasse Streibel, [lasse.streibel@iwb.tum.de](mailto:lasse.streibel@iwb.tum.de)