

## Einführung

**Warum:** Die Herausforderung besteht darin, Data Science und Python-Programmierung innerhalb von 4 SWS so zu erlernen, dass Studierende in der Lage sind, eigenständig einen Datensatz auf Kaggle für eine ML-Trainingspipeline zu entwickeln. Die Studiengruppe ist in Bezug auf Wissensstand und Motivation sehr heterogen. Die Motivation soll gehalten werden können.

**Wie:** Die Studierenden werden von einem LLM-Agenten als Tutor unterstützt. Dieses Modell ist ein feinabgestimmter lokaler MixOfExperts und Multi-Agenten-Chatbot mit RAG-Support.

**Wer:** Die Zielgruppe sind Vertiefungsstudierende der Betriebswirtschaftslehre, ohne Programmier- und Vorkenntnisse in KI, ML und Data Science. Es wird eine heterogene Kenntnisbasis unter den Teilnehmenden angenommen.

**Wann:** Wintersemester 23/24 im Rahmen der Vertiefung "Data Science" im Bachelorstudium der BWL mit 4 SWS statt.

**Ziel:** Ziel ist es, den Studierenden das individuelle Erlernen der Programmiersprache Python zu ermöglichen (primär durch MixOfExperts) und sie individuell bei ihren ML-Experimenten zu unterstützen (primär durch den Multi-Agenten-Chatbot). Die Studierenden sollen am Ende der Veranstaltung erweitertes Wissen/Transfer im Umgang mit ML/KI/Data Science-Konzepten/genKI (LLM) erlangen und anwenden können.

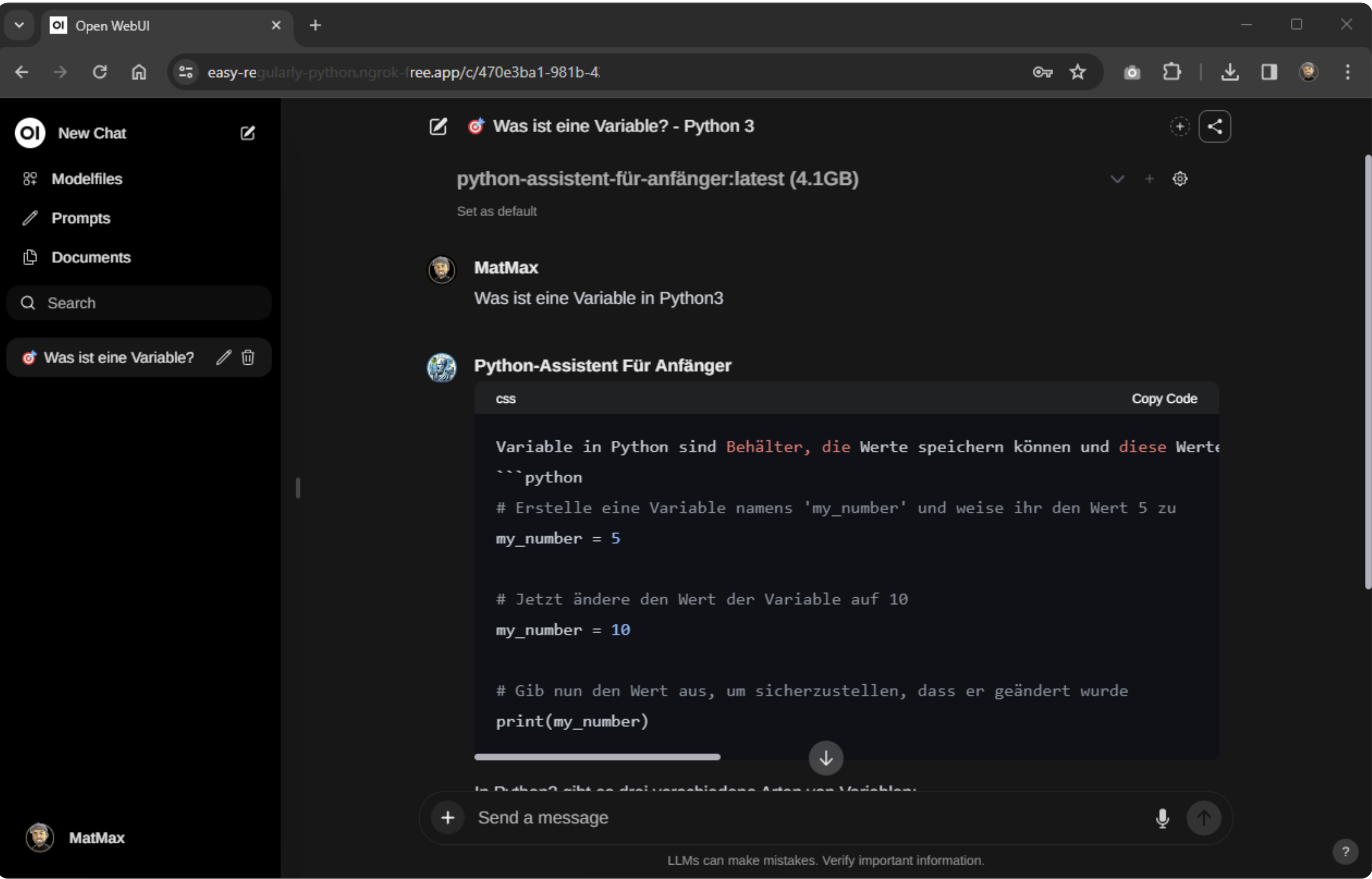


Abb.1: Screenshot der WebUI

## Iterative Fragen an die Lehrmethode

- (1) Anpassungen am Backend (Embedding, RAG; LLM):** Während des Semesters wurden gezielte Verbesserungen an dem Embedding-Modell, dem RAG-Framework und dem LLM durchgeführt, insbesondere durch die Einführung spezifischer Prompt-Dateien.
- (2) Anpassungen am Frontend (UI):** Parallel dazu erfolgten Optimierungen der Benutzeroberfläche, um die Benutzerfreundlichkeit zu steigern.

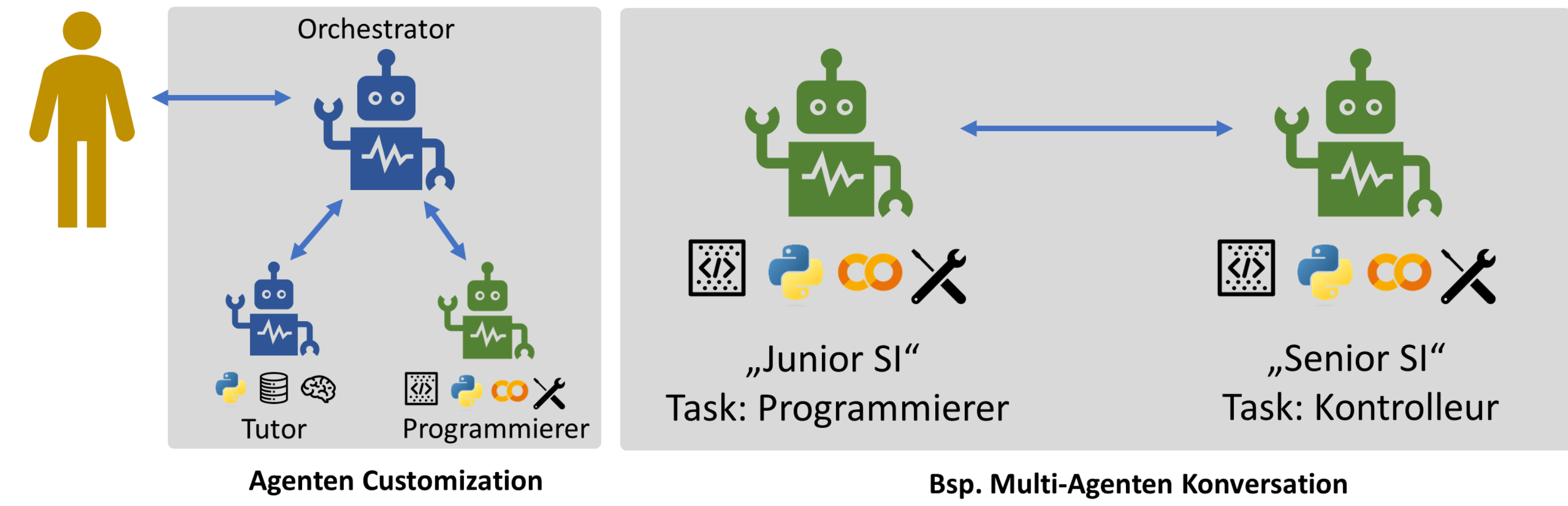
Durch diese Anpassungen wurde der Tutor intuitiver, präziser und anpassfähiger, was den Studierenden die Interaktion mit dem System erleichterte und den Lernprozess unterstützte.

## Under the hood

**Merge-Modell** aus zwei "leo-mistral-7b-Modellen".

→ **Modell 1 = Moderator:** Das erste ist speziell auf Q&A im Sokratesstil für Allgemeinwissen in KI, Programmierung, Data Science, trainiert mit meinem eigenen Data Science-Skript und diversen Lehrveröffentlichungen von mir.

→ **Modell 2= Programmier-Assistent:** Das zweite Modell fokussiert sich auf Python Grundlagen/ML und AutoML, trainiert mit Jupyter Notebooks von GitHub.



- Für den Multi-Agenten-Einsatz werden die Modelle unter differierenden Aufgabenstellungen mehrfach aufgerufen. (Bsp.: Assistant ist einmal der Code-Erzeuger und in einer weiteren Rolle der Code-Kontrolleur) (dies verläuft iterativ, bis es klappt)
- Bereitgestellt im Backend (ollama.ai, ChromaDB, LangChain) Frontend (OpenWebUI, TaskWeaver), Proxmox Server LXC-Container, Hardware (Zotac Magnus One, i7-13, NVIDIA RTX 4070)

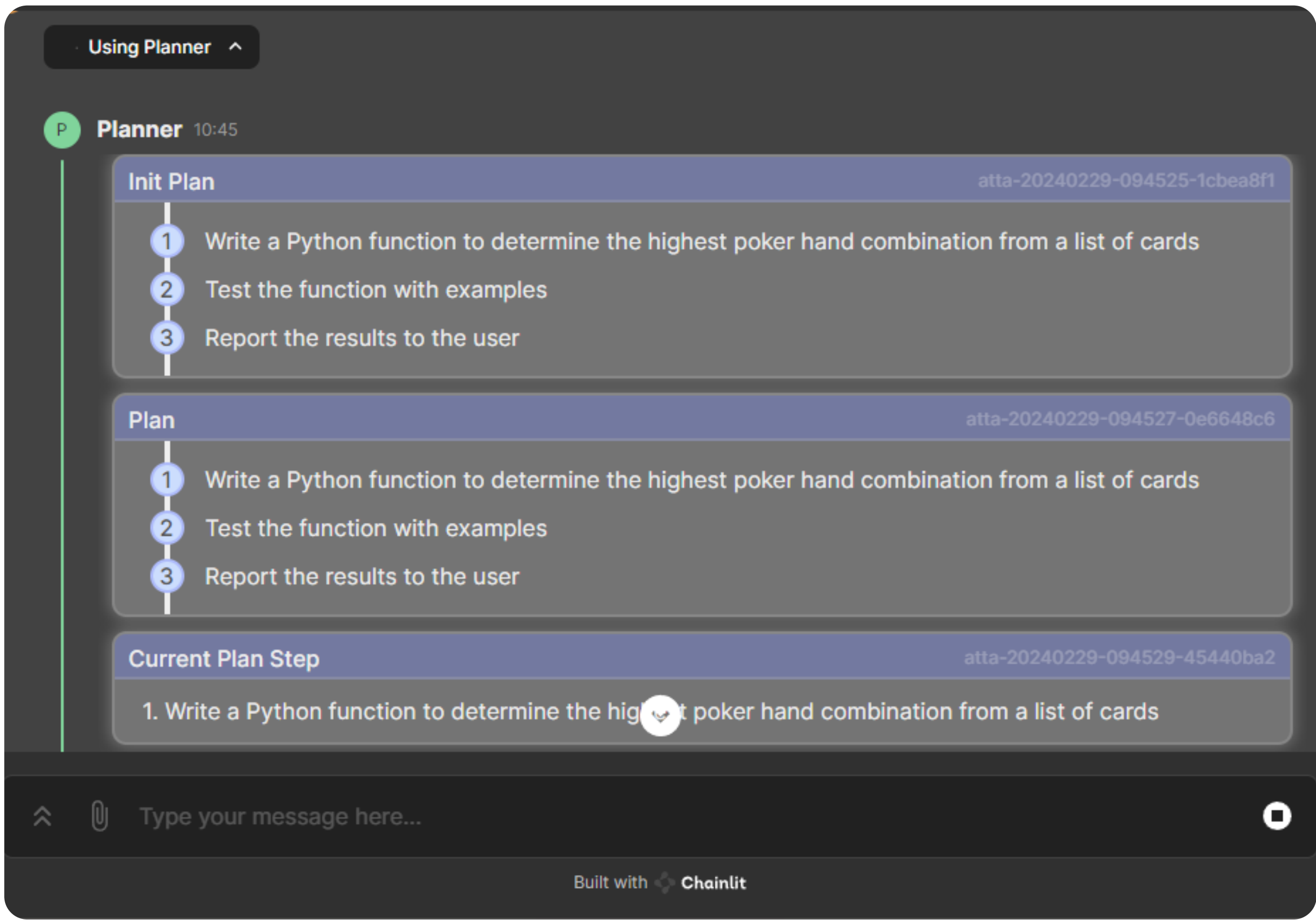


Abb.2: Screenshot eines Tasks

## Fazit und Ausblick

- Erkenntnisse:**
- Alle Studierenden konnten die Grundlagen der Programmiersprache Python erfolgreich anwenden.
  - Neben Python wurden Kompetenzen im Umgang mit Assistenzsystemen erworben.
  - Die Qualität der Arbeit aller Studierenden war überragend.
- Impacts:**
- Durchgehend hohe Motivation über das Semester.
  - Überdurchschnittliche Lernerfolge durch individuellen Lernrhythmus und Einsatz maßgeschneiderter KI-Modelle.
  - Signifikante Steigerung der Lerneffizienz und Qualitätsstandards in der Data Science Ausbildung.
- Ausblick:**
- Prüfen der Anpassung dieses Systems für weitere Veranstaltungen.

