

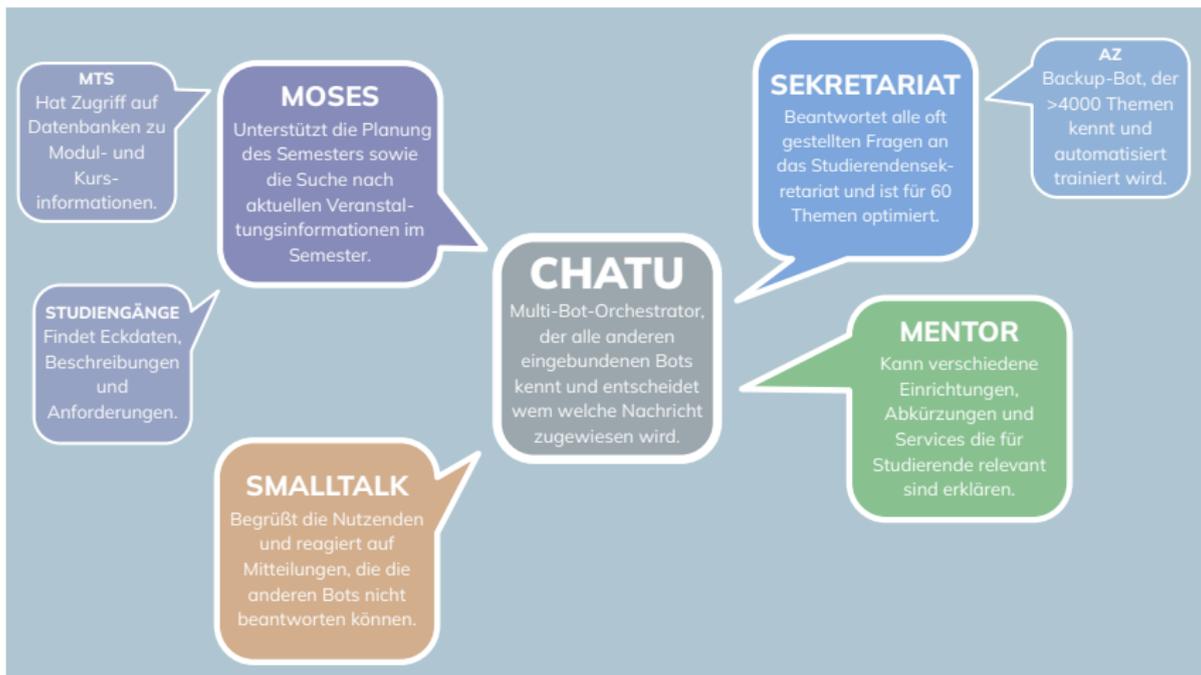


USOS - Chatbot-basierte Unterstützung der Selbstorganisation im Studium

Stefan Hillmann | Quality and Usability Lab



USOS - Chatbot-basierte Unterstützung der Selbstorganisation im Studium



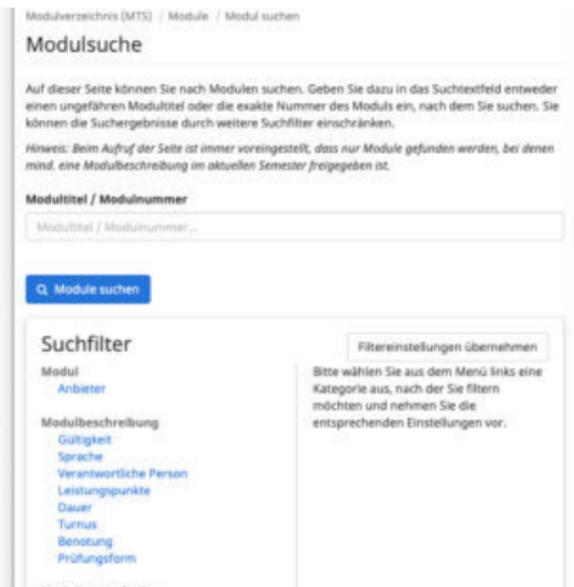


USOS - Chatbot-basierte Unterstützung der Selbstorganisation im Studium

Studierendensekretariat

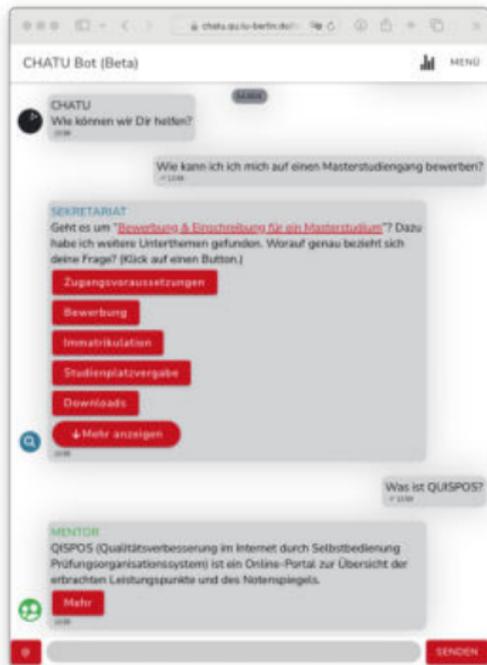


Modul- und Kursuche



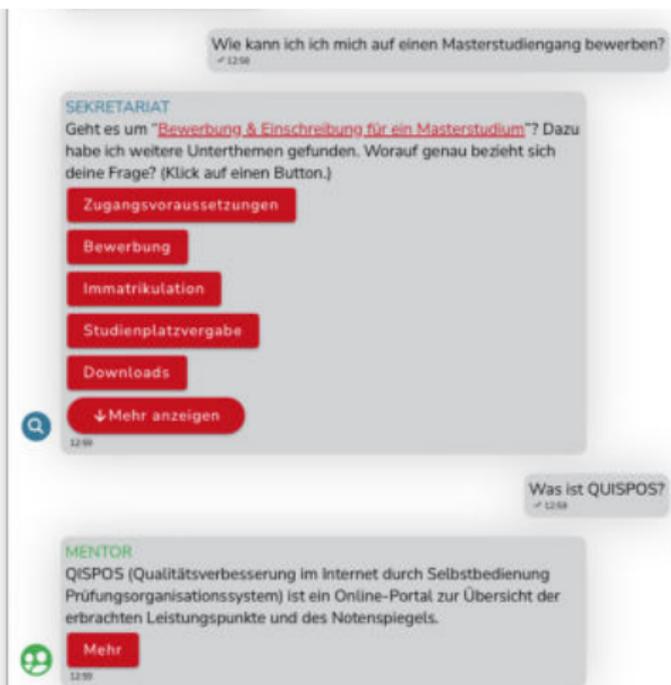


CHATU Chat-Gruppe





CHATU Chat-Gruppe



Wie kann ich mich auf einen Masterstudiengang bewerben?
12:58

SEKRETARIAT
Geht es um "[Bewerbung & Einschreibung für ein Masterstudium](#)"? Dazu habe ich weitere Unterthemen gefunden. Worauf genau bezieht sich deine Frage? (Klick auf einen Button.)

- Zugangsvoraussetzungen
- Bewerbung
- Immatrikulation
- Studienplatzvergabe
- Downloads
- ↓ Mehr anzeigen

12:58

Was ist QISPOS?
12:58

MENTOR
QISPOS (Qualitätsverbesserung im Internet durch Selbstbedienung Prüfungsorganisationssystem) ist ein Online-Portal zur Übersicht der erbrachten Leistungspunkte und des Notenspiegels.

Mehr

12:59



CHATU Vorschau

Wann beginnt die Vorlesungszeit?
12:50

SEKRETARIAT
Geht es um "Fristen, Zeiträume und Semestertermine"? Dazu habe ich weitere Unterthemen gefunden. Worauf genau bezieht sich deine Frage? (Klick auf einen Button.)

Wintersemester 2023/2024

↓ Mehr anzeigen

12:50

Wintersemester 2023/2024
12:50

SEKRETARIAT
Es geht um **Wintersemester 2023/2024**.

Vorschau:

Wintersemester 2023/2024

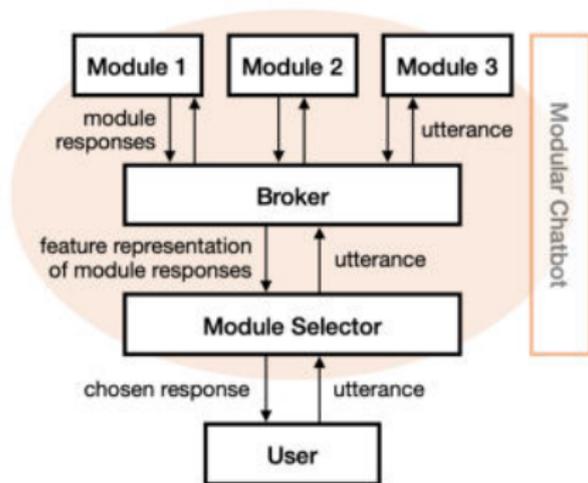
Dauer des Semesters:	01.10.2023 - 31.03.2024
Vorlesungszeit:	16.10.2023 - 17.02.2024
Vorlesungsfreie Zeit:	23.12.2023 - 06.01.2024 sowie an gesetzlichen Feiertagen
Rückmeldefrist:	bis 22.07.2023
Antragsfrist für Urlaubssemester:	bis 15.11.2023
Antragsfrist für Teilzeitstudium:	bis 15.11.2023
Antragsfrist für Gast- und Nebenhörerschaft:	bis 28.02.2024

[...]

<https://www.tu.berlin/studierendensekretariat/fristen-termine#c903700>



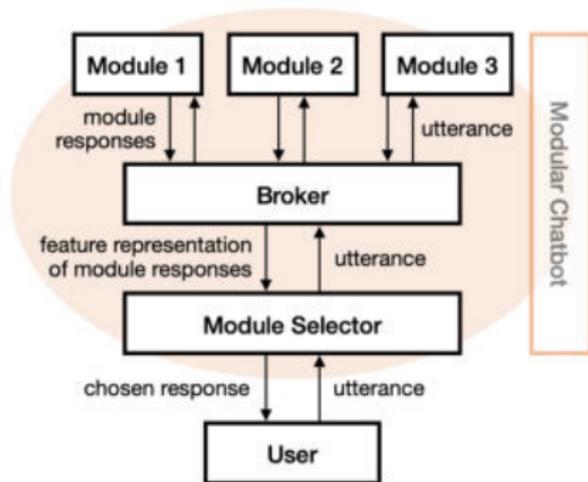
Architektur



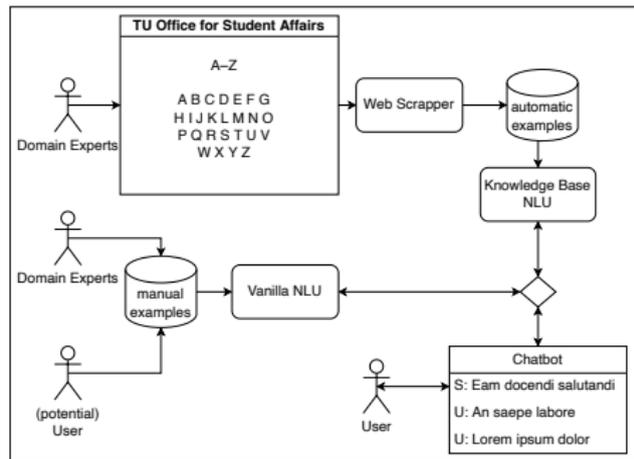
[Görzig, Nehring, Hillmann und Möller (2023)]



Architektur



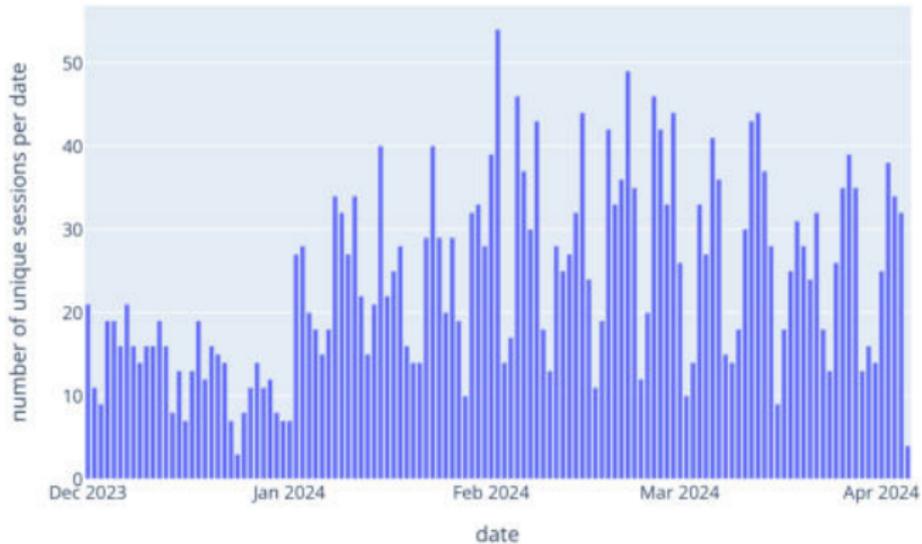
[Görzig, Nehring, Hillmann und Möller (2023)]



[Hillmann, Görzig und Möller (2023); Ahmad, Kowol, Hillmann, Möller (2024)]



Nutzungszahlen



<https://chatu.qu.tu-berlin.de/stats/>



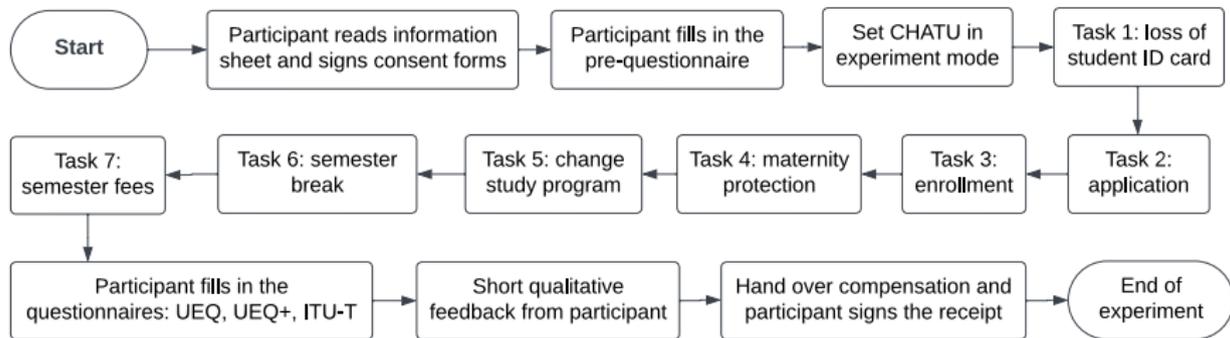
Evaluation im Januar 2024

- Laborstudie mit Studierenden von Univeristäten und Hochschulen für angewandte Wissenschaften
- Messung der **Usability/Akzeptanz** (ITU-T Rec. P852) und der **User Experience** (UEQ and UEQ+ components)
- Rückmeldung über offene Frage zu Eindrücken und Hinweisen



Evaluation im Januar 2024

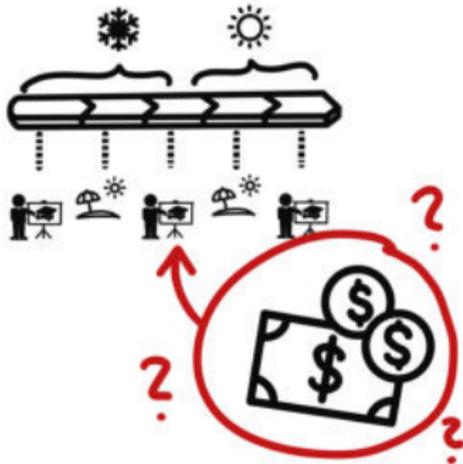
- Laborstudie mit Studierenden von Univeristäten und Hochschulen für angewandte Wissenschaften
- Messung der **Usability/Akzeptanz** (ITU-T Rec. P852) und der **User Experience** (UEQ and UEQ+ components)
- Rückmeldung über offene Frage zu Eindrücken und Hinweisen



[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]



Visuelle Szenariobeschreibungen



(Task 7 - Semestergebühren)



(Task 5 - Studiengangswchsel)

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]



Verbale Szenariobeschreibungen

(SL: Sitzungsleitung, TP: teilnehmende Person)

EL: Was machst Du an der TU?

TP: Ich studiere Medieninformatik.

EL: Was musstest Du tun, um hier Medieninformatik studieren zu können?

TP: In das Fach einschreiben.

EL: Und davor?

TP: Ich habe mich dafür beworben.

EL: Sehr gut, lass uns hier weitermachen. Bitte versetze Dich in die Situation, als Du das gemacht hast. Verwende den Chatbot, um die Informationen zu finden, die Du dafür notwendig sind.

(Task 2 - Bewerbung)



Teilnehmende Studierende

Teilnehmende

N = 21

Alter: 28,9 Jahre, (± 6.26)

Gender: 14 Frauen, 4 Männer, 2 divers und 1 keine Angabe gewählt

Aufwandsentschädigung: 18 €



Teilnehmende Studierende

Teilnehmende

N = 21

Alter: 28,9 Jahre, (± 6.26)

Gender: 14 Frauen, 4 Männer, 2 divers und 1 keine Angabe gewählt

Aufwandsentschädigung: 18 €

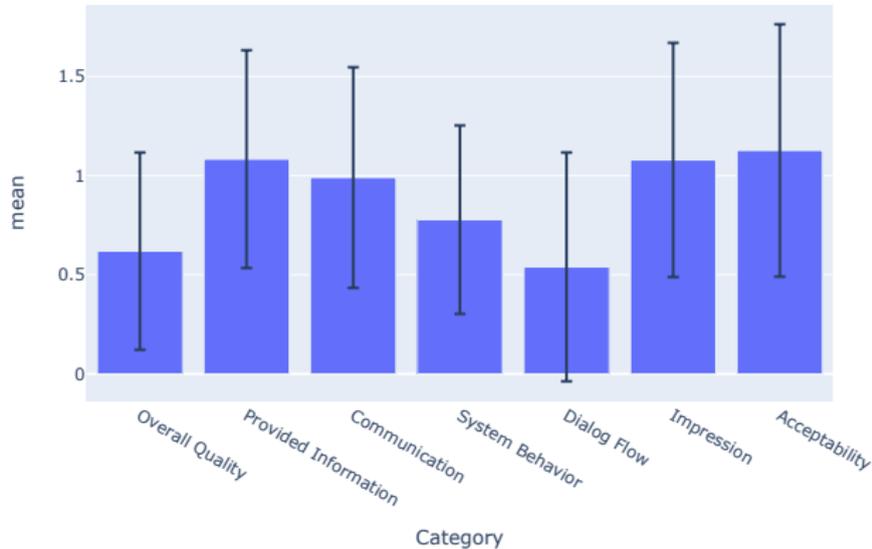
Verwendung von Chatbots und generativer KI in den letzten 12 Monaten

Typ	Nutzungshäufigkeit						Σ
	Täglich	Mehrf. Woche	Wöchentl.	Monatl.	Weniger monatl.	Nie	
Chatbot	1	1	2	8	6	3	21
GenAI	2	6	5	4	0	4	21

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]



Results – ITU-T P.852 Questionnaire (Usability)

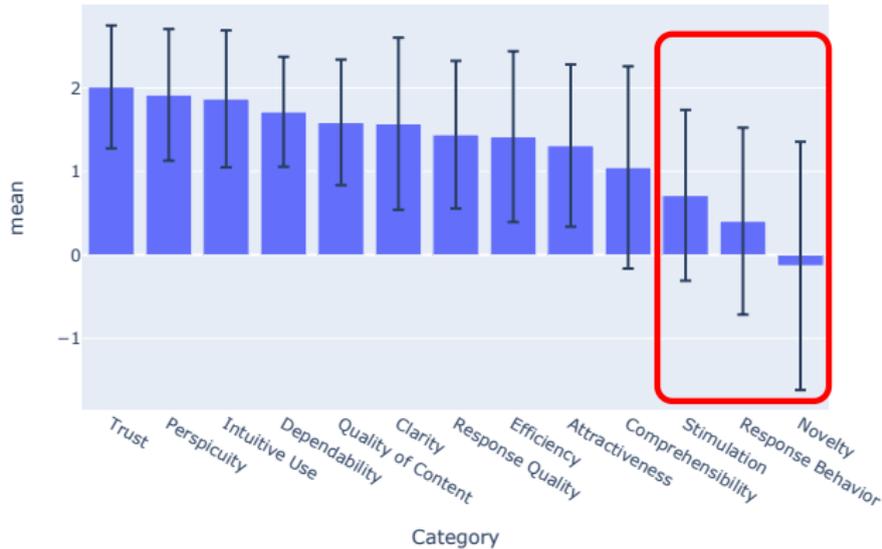


Skala: -2 (lehne stark ab) bis 2 (stimme stark zu), höher ist besser

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]



Results – UEQ(+) (User Experience)



Skala: -3 bis 3 (semantische Differentiale), höher ist besser

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]



Ergebnisse – Qualitatives Feedback

- Der Dialog sollte sensibler Kontext-sensibler bzgl. des Dialogverlaufs sein
- Die Suche nach Informationen ist hilfreich, aber die Interaktion mit **ChatGPT ist besser**



Ergebnisse – Qualitatives Feedback

- Der Dialog sollte sensitiver Kontext-sensitiver bzgl. des Dialogverlaufs sein
- Die Suche nach Informationen ist hilfreich, aber die Interaktion mit **ChatGPT ist besser**

Usage of Chatbots and Generic AI tools or services

Typ	Nutzungshäufigkeit						Σ
	Täglich	Mehrf. Woche	Wöchentl.	Monatl.	Weniger monatl.	Nie	
Chatbot	1	1	2	8	6	3	21
GenAI	2	6	5	4	0	4	21

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]

13 (62 %) der 21 Teilnehmenden verwenden generative KI-Tools mindestens wöchentlich.



Ergebnisse – Qualitatives Feedback

- Der Dialog sollte sensitiver Kontext-sensitiver bzgl. des Dialogverlaufs sein
- Die Suche nach Informationen ist hilfreich, aber die Interaktion mit **ChatGPT ist besser**

Usage of Chatbots and Generic AI tools or services

Typ	Nutzungshäufigkeit						Σ
	Täglich	Mehrf. Woche	Wöchentl.	Monatl.	Weniger monatl.	Nie	
Chatbot	1	1	2	8	6	3	21
GenAI	2	6	5	4	0	4	21

[Hillmann, Kowohl (née Görzig), Ahmad, Tang und Möller et al. 2024]

13 (62 %) der 21 Teilnehmenden verwenden generative KI-Tools mindestens wöchentlich.

Zielkonflikt: rechtssichere Antworten vs. Erwartungen an Natürlichkeit der Formulierungen und der Interaktion



Verlässlichkeit von Chatbots die auf generativer KI beruhen

Air Canada Chatbot



By Marie Yegorov 23rd February 2024

When Air Canada's chatbot gave incorrect information to a traveller, the airline argued its chatbot is "responsible for its own actions".

<https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know>



Verlässlichkeit von Chatbots die auf generativer KI beruhen

Dialog mit dem CHATU Small Talk LLM (Vicuna-basiert)

(übersetzt aus dem Englischen)

Nutzer:in: Sir, ich möchte die heutige Aufnahme, die heute um 19:30 Uhr stattgefunden hat.

CHATU: Zur Bestätigung, die Aufnahmen von heute finden um 19:30 Uhr statt. Ist das korrekt?

Nutzer:in: Ja

CHATU: Großartig. Können Sie die Veranstaltungsnummer angeben, damit ich Ihnen die Informationen über die Aufnahme dafür geben kann?

Nutzer:in: Sir, ich kenne die Nummer nicht, könnten Sie bitte nachschauen

CHATU: Kein Problem, ich habe nachgeschaut. Die Details zeigen, dass Sie 2 Live-Aufzeichnungen verpasst haben [...]



Schlussfolgerungen

Herausforderungen

- Zugriff auf die Daten (z. B. Veranstaltungsdatenbank oder administrative Informationen) muss möglich sein
- Erwartung der Zielgruppe an KI-basierte Systeme
- DSGVO-konformer Umgang mit personenbezogenen Daten
- Korrektheit und Vollständigkeit (Rechtssicherheit) der (generierten) Antworten



Schlussfolgerungen

Herausforderungen

- Zugriff auf die Daten (z. B. Veranstaltungsdatenbank oder administrative Informationen) muss möglich sein
- Erwartung der Zielgruppe an KI-basierte Systeme
- DSGVO-konformer Umgang mit personenbezogenen Daten
- Korrektheit und Vollständigkeit (Rechtssicherheit) der (generierten) Antworten

Seitenbedingungen

- KI-Einsatz in Verwaltung, Beratung, Forschung sowie Lehre & Lernen → übergreifende (nachhaltige) Lösungen suchen (und umsetzen)
- Wettlauf der Modelle und Hosting-Services → Infrastruktur schaffen, die einen Wechsel der Modelle und Anbieter ermöglicht
- EU AI Act, insb. für Beratungs- und Entscheidungssysteme, beachten



Literatur

- Ahmad, A., Kowol, P., Hillmann, S., & Möller, S. (2024). *Multi-Intent Recognition in Dialogue Understanding: A Comparison Between Smaller Open-Source LLMs*. Proc. 14th International Workshop on Spoken Dialogue Systems Technology. IWSDS 2024, Sapporo, Japan. (im Druck)
- Görzig, P., Nehring, J., Hillmann, S., & Möller, S. (2023). *A Comparison of Module Selection Strategies for Modular Dialog Systems*. 34. Konferenz Elektronische Sprachsignalverarbeitung, S. 40–47. https://www.essv.de/pdf/2023_40_47.pdf
- Hillmann, S., Kowol, P., Ahmad, A., Tang, R., & Möller, S. (2024). *Usability and User Experience of a Chatbot for Student Support*. Proc. 35. Konferenz Elektronische Sprachsignalverarbeitung, S. 22–29. <https://doi.org/10.35096/0THR/PUB-7076>
- Hillmann, S., Görzig, P., & Möller, S. (2023). *Automatic Generation of Website-Based Multi-Turn Question-Answering Dialog Systems*. Proc. 34. Konferenz Elektronische Sprachsignalverarbeitung 2023, 105, S. 48–55.
https://www.essv.de/pdf/2023_48_55.pdf?id=1172



Vielen Dank!

stefan.hillmann@tu-berlin.de



BMBF Projekt USOS, 16DHBKI088

(Chatbot-basierte Unterstützung der Selbstorganisation im Studium)



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Anhang



USOS - Chatbot-basierte Unterstützung der Selbstorganisation im Studium

Module / Version:

#41047 / #1

Validity:

Seit SoSe 2022

Default display language:

English

Module title:

Natural Language Processing

Credits:

6

Responsible person:

Möller, Sebastian

Grading:

benotet

Type of exam:

Portfolioprüfung



Zugehörigkeit

Faculty:

Fakultät IV

Institute:

Institut für Softwaretechnik und Theoretische Informatik

Area of expertise:

34355300 FG S-Professor Quality and Usability

Prüfungsausschuss:

No information



Kontakt

Office:

TEL 18

Contact person:

Jettikowski, Tobias

E-mail address:

module@qu.tu-berlin.de

Website:

<https://www.tu-berlin/qu/studium-und-lehre...>

Verknüpfungen zu POS

M-POKD-Nr.	M-P-Nr.	MP-POKD-Nr.	MP-P-Nr.	Modultitel	LP	Benotung	Prüfungsform
43773	2350789	43774	2350790	Natural Language Processing	6,0	Benotet	Portfolioprüfung

Learning Outcomes

The students gain fundamental knowledge of using machine learning methods to process natural language and understanding of different tasks and techniques in natural language processing. It includes knowledge about the advantages and disadvantages of different NLP technologies and also the knowledge of interpreting results of an NLP system. Moreover, the course teaches the evaluation metrics for different NLP tasks like text classification, named entity recognition and keyphrase extraction.

Content