

Six Thinking Chatbots -Workshop und Vortrag-

Agenda

- 13:10 Teil 1: Sechs Denkhüte Workshop
- 13:40 Teil 2: Erkenntnisse aus Experimenten mit den Six Thinking Chatbots

Sechs Denkhüte Workshop

Sechs Denkhüte

- Die „Sechs Denkhüte“ sind eine Kreativtechnik.
- Anhand der Technik sollen verschiedene Blickwinkel / Denkweisen auf ein Thema zusammengeführt werden.
- Jeder Hut vertritt eine Denkweise; jeder Hut wird durch eine Farbe repräsentiert.
- Es gibt mehrere Ausprägungen der Technik; wir führen heute eine sehr einfache und strukturierte durch.



Daten



Kreativität



Positivität



Negativität



Gefühlsbetont



Prozessorientiert

Sechs Denkhüte – Anleitungen zum Workshop

- Wir bilden Gruppen a 6 Personen; falls es nicht genau aufgeht, vergeben wir eine Farbe mehrere male.
- Der Moderator (ich) vergibt eine Aufgabe; die Teilnehmer (ihr) denken individuell 2 Minuten über das Thema nach.
- Die Teilnehmer erklären kurz in der untenstehenden Reihenfolge ihre Sichtweise auf die Aufgabe.
- Der „blaue Hut“ stellt das Ergebnis der anderen Teilnehmer im Plenum kurz vor.

Gruppenarbeit

Plenum



Daten



Kreativität



Positivität



Negativität



Gefühlsbetont



Prozessorientiert

Was müssen Trainingsdaten bieten, damit KI im Requirements Engineering nützlich ist?



Daten



Kreativität



Positivität



Negativität



Gefühlsbetont



Prozessorientiert

Erkenntnisse aus Experimenten mit den Six Thinking Chatbots

Agenda

- Scope des Vortrags
- Six Thinking Chatbots
- Kreativität
- Erste Experimentenserie im Winter 2023: Forschungsfragen, Setting und Ergebnisse
- Zweite Experimentenserie im Sommer 2024: Forschungsfragen, Setting und Ergebnisse
- Vorläufige Erkenntnis

Überschriften aus den Jahren 2022/23/24

Zukunftsforschung

So verändert KI die Arbeitswelt

Stand: 30.04.2023 08:07 Uhr

Künstliche Intelligenz verändert die Welt: Sie schreibt Referate, Werbetexte oder Programmcodes in Sekundenschnelle. Viele Menschen fürchten, KI könnte ihnen bald den Job wegnehmen. Ist das realistisch?

<https://www.tagesschau.de/wissen/forschung/berufe-kuenstliche-intelligenz-100.html>

ChatGPT, Aleph Alpha und Co.

In diesen Bereichen kann KI uns die Kreativarbeit abnehmen

Künstliche Intelligenz kann mittlerweile Texte schreiben, Bilder erzeugen oder Fakten sammeln. Wo Computer

<https://www.handelsblatt.com/karriere/chatgpt-aleph-alpha-und-co-in-diesen-bereichen-kann-ki-uns-die-kreativarbeit-abnehmen/28989050.html>

Articles > 20 Berufe, die schon bald durch KI ersetzt werden könnten

20 Berufe, die schon bald durch KI ersetzt werden könnten

Die Kraft von KI, Machine Learning & Big Data

Daniel Schwarz 10 Min. 12406 ★★★★★ Aktualisiert: 12. August 2024

<https://www.bitrix24.de/articles/20-berufe-die-schon-bald-durch-ki-ersetzt-werden-koennen.php>

Künstliche Intelligenz

Welcher Arbeitsplatz ist vor der KI noch sicher?

Künstliche Intelligenz ist spätestens mit ChatGPT im Alltag angekommen und wird auch den Arbeitsmarkt verändern. Manche Jobs werden sich durch KI stark wandeln, andere vielleicht sogar verschwinden. Wie sieht die Arbeitswelt der Zukunft aus?

25.05.2023

<https://www.deutschlandfunk.de/kuenstliche-intelligenz-ki-arbeitsplatz-arbeitsmarkt-chatbot-102.html>

Six Thinking Chatbots

- „Six Thinking Chatbots“ (STC) ist ein ChatGPT-basiertes Programm, welches den Ablauf eines Six Thinking Hats Workshops simuliert / an ChatGPT delegiert.
- STC wurde als nicht-produktive Demonstration der Kreativität von LLMs erdacht.
 - Nicht-produktiv: STC soll kein marktfähiges Produkt sein; es wird nicht mehr weiterentwickelt; es ist verfügbar unter <https://github.com/rachmann-alexander/sixthinkingchatbots>
 - Kreativität: **Kann eine LLM überhaupt kreativ sein? Kann eine LLM kreativer sein als Menschen? Was bedeutet das für das Requirements Engineering?**

Kreativität

- Nützlichkeit (usefulness)
- Neuheit (originality)

Beides ist letztlich nie objektiv und allgemeingültig bewertbar, sondern muss in einem bestimmten Kontext bewertet werden: Wie nützlich ist eine Idee für eine bestimmte Anwendung? Wie neu ist eine Idee für ein bestimmtes Unternehmen? Etc.

Erste Experimentenserie im Winter 2023: Forschungsfragen, vor allem zur Neuheit der Responses

- Unterscheiden sich die Responses der Hüte?
- Unterscheidet sich das Ergebnis von STC zu einem normalen Prompt?
- Antwortet STC immer das gleiche auf einen Prompt?

Erste Experimentenserie im Winter 2023: Setting

STC wird geprompted mit der Aufgabenstellung:

The goal of the workshop is to evaluate three alternatives to a problem. The problem is: The sales process on our companies website does not work very well. Management wants us to change the sales process to one of the three alternatives:
First, an old-school chatbot with a beforehand written structure.
Second, a chatbot using a large language model, without a pre-defined structure.
Third, an old-school contact form.
Evaluate which solution one would prefer. Give at most 15 arguments. After listing the arguments, write a conclusion including a selection on a solution.
Use markdown syntax.

Erste Experimentenserie im Winter 2023: Unterscheiden sich die Responses der Hüte? Unterscheidet sich das Ergebnis von STC zu einem normalen Prompt?

Argument	W	G	Y	B	R	WP
Factual Analysis	X					
Objective Comparison	X					
Long-Term Impact and Scalability	X					
Innovation & AI-Driven		X	X			
Enhanced User Experience		X			X	X
Limitation of Alternatives		X				
Adaptability / Limited Flexibility of old school chatbot			X		X	X
Potential for Growth / Scalability			X			X
Non-comprehensiveness				X		
Potential for confusion				X		
Outdated				X	X	X
Consistency / Reliability						X
Costs						X
Speed						X
Ease of Implementation of old school chatbot / form						X
Advanced Capabilities of LLM						X
Complexity of LLM						X
Data Privacy of LLM						X
Data Collection of old school form						X
Limited information via old school form						X

Erste Experimentenserie im Winter 2023: Wie ähnlich sind Antworten von STC, wenn mehrfach gepromptet wird? Beantwortet anhand der Jaccard-Ähnlichkeit, berechnet mit dem Shingle-Algorithmus

	1	2	3	4	5
1	1.00	0.61	0.59	0.60	0.58
2		1.00	0.63	0.62	0.61
3			1.00	0.60	0.61
4				1.00	0.63
5					1.00

Zweite Experimentenserie im Sommer 2024: Forschungsfrage

- Sind die Responses von STC origineller als die Ergebnisse von Menschen?
- Sind die Responses von STC nützlicher als die Ergebnisse von Menschen?

Zweite Experimentenserie im Sommer 2024: Setting

On June 3, 2024, three groups of second-semester business informatics students were given tasks as part of two workshops. The first workshop took place in the morning in Cologne. Two groups attended. The second workshop took place in the afternoon in Solingen. One group attended.

First, the students were explained the concept of a Six Thinking Hats workshop. The number of students was not a multiple of six, so some hats were filled twice. The order of the hats was defined as: white, green, yellow, black, red (as programmed in STC in "Idea" mode).

Due to time constraints, the first two groups (located in Cologne) were given only one task. The third group (located in Solingen) was given two tasks, which were completed one after the other.

The tasks were:

- First task: Describe an app that can be used to make the personal carbon footprint clear. ("Beschreibt eine App, mit der der persönliche CO2-Fußabdruck deutlich gemacht werden kann.")
- Second task (only group 2): Describe how the person's data should be entered into the app. ("Beschreibt, wie die Daten der Person in die App eingegeben werden sollten.")

Zweite Experimentenserie im Sommer 2024: Ergebnisse der Workshops vs. STC Responses vs. Referenztext

Co2 App:

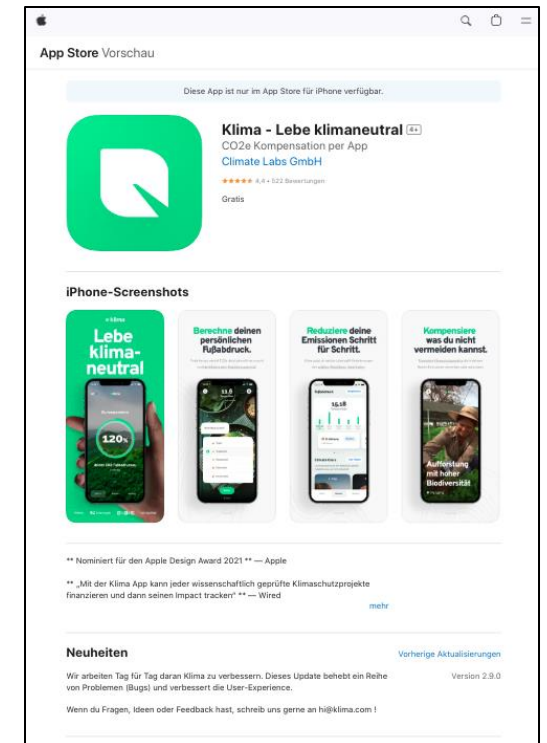
Wsp: Algorithmus?, Datenschutz, Genauigkeit und wie schnell sind die Daten kontinuierlich Daten bereitstellen
 Grün: Beispiel wie bei Fitness app tägliche Zielherausforderungen, Halbtags wärd sich gemäß des Aktivitäts verändert, fröhlich die Verlesungen, Tipps u. Tricks
 Gelb: die App schütt Benachrichtigung beim Nutzer, zur Verhaltensänderung führen, duldung über das Thema, eigenen Fortschritt drücken, Community bilden
 Rot: Super App aber Zuspätkommen zeigt wie gut aber auch wie schnell man ist.
 Schwarz: Datenschutz bedenken, wie genau sind die Daten, Kosten?, besteht Überhaupt Abhängigkeit, Schuld und Stress

Blau:

1. Nützliche Funktion
2. Nachrichten prüfen
3. Datenanalyse regelmäßig
4. App planen (z.B. Anforderungen von Grün)
5. Implementieren u. testen
6. rausbringen

Based on the contributions of the participants, the following ideas should be chosen:

1. Data Collection: The app should collect data on the user's daily activities, such as their mode of transportation, dietary habits, energy consumption, and waste production. This data is crucial in determining the individual's carbon footprint.
2. Carbon Footprint Calculation: The app should have a built-in algorithm that calculates the user's carbon footprint based on the collected data. This algorithm should be based on scientifically accepted methods of carbon footprint calculation to ensure accuracy.
3. Personalized Recommendations: The app should offer personalized recommendations on how to reduce the user's carbon footprint. These recommendations should be based on the user's specific activities and lifestyle.
4. Innovation Potential: The app could be designed to track and analyze user's



Zweite Experimentenserie im Sommer 2024: Sind die Responses von STC origineller als die Ergebnisse von Menschen? Sind die Responses von STC nützlicher als die Ergebnisse von Menschen? Subjektive Einschätzung anhand eines Referenztextes

		Text	Originality	Usefulness
Studierende	STC	B.1	No originality above reference text ("No")	Yes. Example: "However, there is also a psychological confrontation if, for example, you have a completely red app after a trip." (l. 175ff.)
		B.2	No	Yes
		B.3	No	Yes
		B.4	No	Yes
		B.5	No	Yes

Vorläufige Erkenntnis

Maschinen sind nicht immer kreativ,
aber Menschen auch nicht.

Rückblick zum Sechs Denkhüte-Workshop

„[...] Basierend auf diesen Beiträgen scheint die Alternative, die am ehesten mit den gesammelten Erkenntnissen übereinstimmt, die Erstellung eines **umfassenden** und **anpassungsfähigen** Datensatzes zu sein.

Dieser Datensatz würde reichhaltiges **domänenspezifisches Wissen**, **verschiedene Szenarien**, **hochwertige Anmerkungen** und **empfindsame Stakeholder-Engagements** enthalten und gepflegt und aktualisiert werden, um aktuelle Branchentrends und Erkenntnisse widerzuspiegeln.

Dieser Ansatz gleicht die Notwendigkeit technischer Genauigkeit mit den menschenzentrierten Nuancen aus, die für ein effektives Requirements Engineering von entscheidender Bedeutung sind.“



Wer war kreativer?

Mehr Details

- Workshopbeitrag „Six Thinking Chatbots: A Creativity Technique deployed via a Large Language Model“ bei CreaRE auf der RefSQ 2024.
- Arbeitsbericht „Evaluation Details of Six Thinking Chatbots“ auf figshare (https://figshare.com/articles/preprint/Evaluation_Details_of_Six_Thinking_Chatbots_Version_2_from_October_2024_/27249327?file=49849557).