

Impressionen zum Workshop – 21. November 2023 KI-Szenarien im Zeitalter von ChatGPT & Co



Eröffnung des hybrid gestalteten Workshops durch:

Dr. Andreas Jedlitschka (*Fraunhofer IESE bzw. Sprecher GI FG Data Science*)
Prof. Dr. Andreas Schmietendorf (*HWR Berlin/Uni Magdeburg*)



Prof. Dr. Claudia Nass
(*HS Mainz/Fraunhofer IESE*)

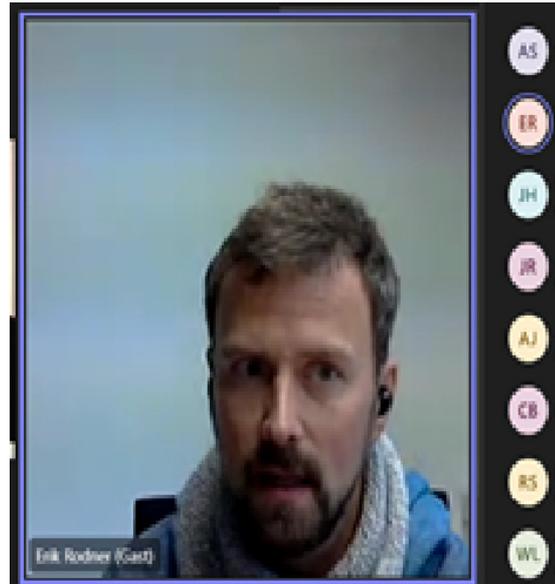
Möglichkeiten des Design Thinking im
Diskurs von KI-Szenarien

- Bedürfnisse der Menschen beim KI-Design in den Mittelpunkt stellen.
- Sinnvoller Einsatz von Design Frameworks (Design Thinking, User Centered-, Human Centered-Design)
- Ethische Aspekte berücksichtigen, was sollten wir mit dieser Technologie machen, Werte und Akzeptanz

Prof. Dr. Erik Rodner
(HTW Berlin)

KI-Robustheit evaluieren sowie Stärken und Schwächen aufzeigen (Online-Beitrag)

- Vertrauen in KI-Lösungen (Sicherheit, Robustheit, Reputation, ...) kann mit Hilfe der Evaluierung bzw. Benchmarking eingesetzter KI-Modelle erreicht werden.
- Sprachmodelle implizieren typischerweise einen systematischen BIAS (z.B. Verzerrungen durch geschlechterspezifische oder auch ortsspezifische Besonderheiten).



Dr. Peter Münte
(Universität Innsbruck)

Fachliche Daten transkribierter
Gesprächsverläufe im Mediationskontext

- Reale Mediationsverfahren mit ggf. mehreren Sitzungen implizieren sehr umfangreiche Transkripte (typ. 1000 Seiten) – eine manuelle Auswertung ist eher für Mikroanalysen realistisch.
- KI könnte einen Beitrag leisten, um Emotionen, Kompromissfähigkeiten oder auch mediatorische Haltungen direkt oder indirekt zu erfassen.

Sandro Hartenstein
(HWR Berlin)

Prototypische Analyse von KI-APIs für
LLMs (Cloud vs. OnPremise)

- Einsatz verschiedener LLMs (GPT 3.5, GPT 4 und LLAMA 2) zur Analyse von exemplarischer Transskripte.
- Die Interpretation der Analyseergebnisse muss in Zusammenarbeit mit den fachlichen Experten erfolgen (Hypothesen – bestätigen/wiederlegen).





Walter Letzel
(TU Berlin)

Konzeptioneller Ansatz zur Analyse und Bewertung von Mediationssitzungen (Online-Beitrag)

- Erfassen von quantitativen (Häufigkeiten – z.B. Gesprächsanteile) und qualitative Merkmalen (Mediationsphasen, ...)
- Einsatz von strukturprüfenden (z.B. Intervention durch Doppeln) und strukturendeckenden Verfahren (z.B. Sentimentanalysen)

Prof. Dr. Ralf Schnieders
(HTW Berlin)

Compliance-Aspekte im Diskurs der Datenüberlassung - Datenschutz, Urheberrecht und Haftung

- Der Umgang mit personenbezogenen Daten im Forschungsdiskurs stellt eine ressourcenmäßig zumeist nicht abbildbare Herausforderung dar.
- Entsprechende Daten sollten für die Verarbeitung dringend einer Anonymisierung unterzogen werden.



Prof. Dr. Jens Heidrich
(HS Mainz/Fraunhofer IESE)

Auswertung und Abschlussdiskussion

- Möglichkeiten und Grenzen von KI-Sprachmodellen (speziell "Large Language Models").
- Gefahr des Übergehens zum Teil unbekannter Stakeholder.
- Bedarfs interdisziplinär zusammengesetzter KI-Teams.
- Konsequenzen des LLM-Einsatzes lassen aktuell schwer abschätzen.
- ...

Fazit und Ausblick

Auf der Grundlage der vorgenannten Impulsvorträge entstand eine lebhafte und vor allem interdisziplinär geprägte Diskussion zu den Möglichkeiten und Grenzen von KI-Sprachmodellen (speziell "Large Language Models"). Zur Geltung kamen dabei insbesondere die Sichtweisen der Informatik, der Soziologie, der Psychologie aber auch der Rechtswissenschaften. Aus fachlicher Sicht galt das Interesse insbesondere der KI-basierten Verarbeitung anonymisierter Transskripte, alternativ wäre auch die direkte und abermals KI-basierte Verarbeitung mitgezeichneter Audiosequenzen denkbar. Die manuell erstellten Transskripte wurden in Anlehnung an den GAT 2 Standard (Gesprächsanalytisches Transkriptionssystem) erstellt.

Als besondere Herausforderung stellte sich dabei die Aufgabe der Anonymisierung heraus. Insbesondere bei umfangreichen Transskripten erscheint eine manuell durchgeführte Anonymisierung, ob des damit einhergehenden Aufwands nur noch bedingt zielführend. Dem entsprechend wäre hier ebenfalls eine mit Hilfe von Algorithmen bzw. Programmen gesteuerte Anonymisierung sinnvoll. Dabei sollten sowohl die rollenbasierte Identifikation der Gesprächspartner als auch ggf. innerhalb des Textes referenzierte Personen berücksichtigt werden. Aus Sicht des Autors bleibt dabei allerdings ein Restrisiko, d.h. die Anonymisierung kann ggf. nur mit einer Wahrscheinlichkeit erfüllt werden. Im Zusammenhang mit den Herausforderungen der Compliance (z.B. EU DSGVO) entstehen für den wissenschaftlichen Bereich dabei immense und ggf. auch unkalkulierbare Herausforderungen.

Konkrete Tests bezogen sich auf Quantifizierungs- und Klassifizierungsaspekte aber auch Sentimentanalysen von Transskripten. Dafür wurden die KI-Modelle ChatGPT 3.5, ChatGPT 4 sowie LLAMA2 zum Einsatz gebracht. Im Zusammenhang mit der Ergebnisinterpretation zeigten sich hier massive Herausforderungen hinsichtlich der Ergebnisinterpretation bzw. -validierung aber auch der Erklärbarkeit der eingesetzten Large Language Models. Weitere Untersuchungen sollen sich auf ein mediationsspezifisches Training der LLM-Ansätze fokussieren.

Neben den hier zur Verfügung gestellten Informationen ist die Publikation eines Tagungsbandes zum 1. Quartal 2024 vorgesehen. Titel: Diskussionsbeiträge zum Projekt TAHAI (TrustAdHocAI) – Ergebnisse eines öffentlichen Expertenworkshops am Fraunhofer IESE – November 2023, in Berliner Schriften zu modernen Integrationsarchitekturen, Shaker-Verlag, Düren, Band 28, ISBN 978-3-8440-9244-8 

Dank

Ein solcher Workshop wäre ohne professionelle Referenten nicht möglich, denen gilt ein herzliches Dankeschön. Ein besonderer Dank gilt den Gastgebern des Workshops vom Fraunhofer IESE (Herrn Dr. Andreas Jedlitschka und Herrn Prof. Dr. Jens Heidrich)! Ein Dank geht ebenfalls an die Unterstützer der Veranstaltung, d.h. die Gesellschaft für Informatik ([GI e.V.](#)) sowie die Central Europe Computer Measurement Group ([ceCMG e.V.](#)).