Bauhaus-Universität Weimar

Faculty of Media

Building Confidence in AI: Investigating User Perceptions of Trust in ChatGPT Results within Academia

By Kartik Bist

MSc. Human computer interaction,
Bauhaus-Universität Weimar,
kartik.bist@uni-weimar.de

eTeach
Netzwerk
Thüringen

The Conception

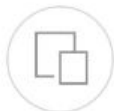# Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities

Authors:

Roman Lukyanenko
University of Virginia

Wolfgang Maass

Veda C. Storey
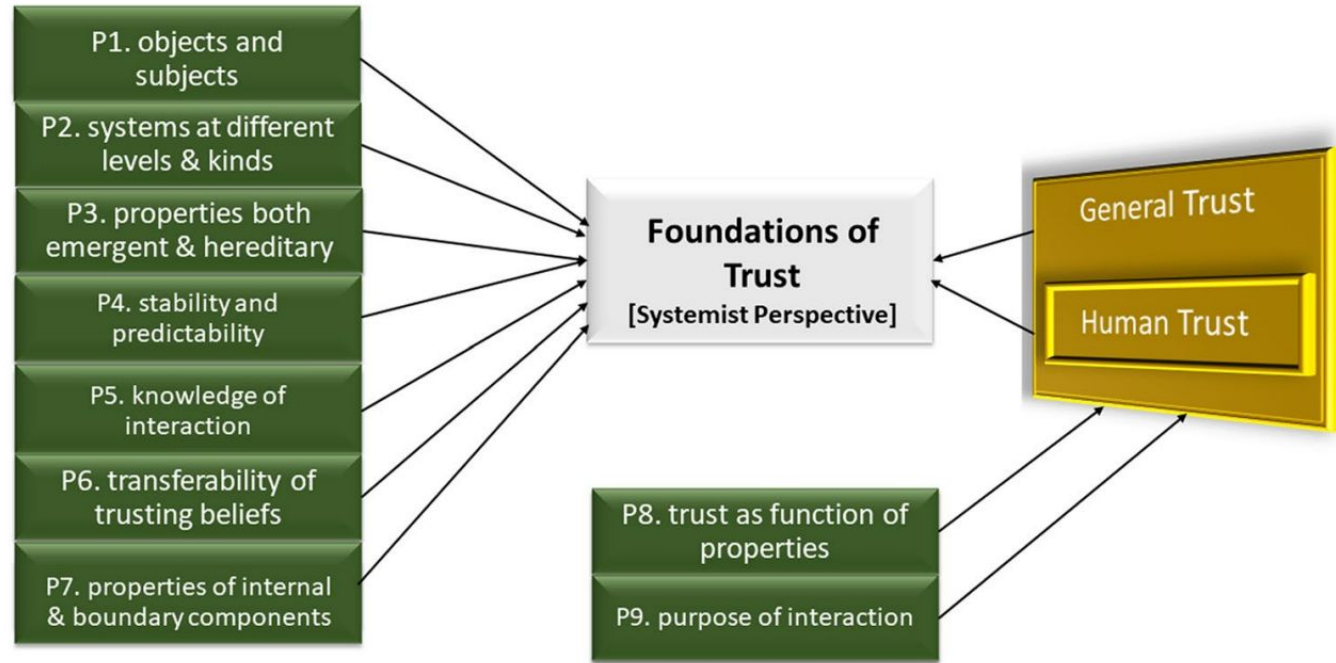
Fig. 1 Foundational Trust Framework

(Lukyanenko et al., 2022)

input (Goodfellow et al., 2016). The opacity of such models undermines the ability to understand and explain how and why such models make their decisions (Adadi & Berrada, 2018; Castelvecchi, 2016; D Gunning, 2016; Mueller et al., 2019; Storey et al., 2022). Likewise, more research is needed in psychology, neuroscience, cognitive science, human-computer interaction on understanding the inner workings of human mind, and trust formation, especially within the context of artificial intelligence (consistent with Proposition 8).

(Lukyanenko et al., 2022)

Objective

1. Examine the general perception of users on ChatGPT in academia.
2. Examine what influences people's confidence in ChatGPT's precision and dependability.
3. Determine the challenges in utilizing ChatGPT in academia.

The Implementation

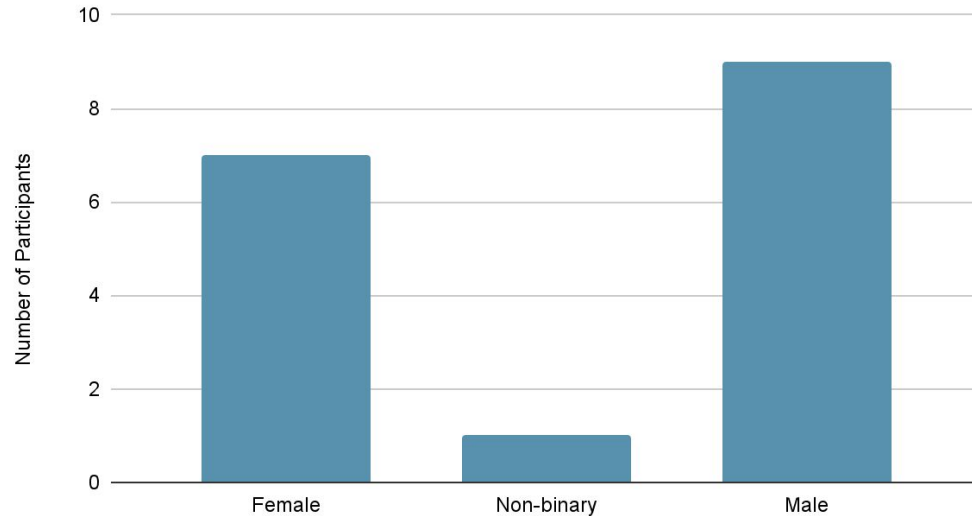# Research Methodology

18
Participants
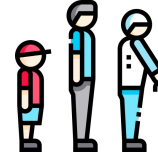
~20 Min
Online
Interview

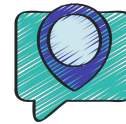Thematic
Analysis

Fact sheet about sample

Participants based on gender

Education: Masters students with Comp. sci. background

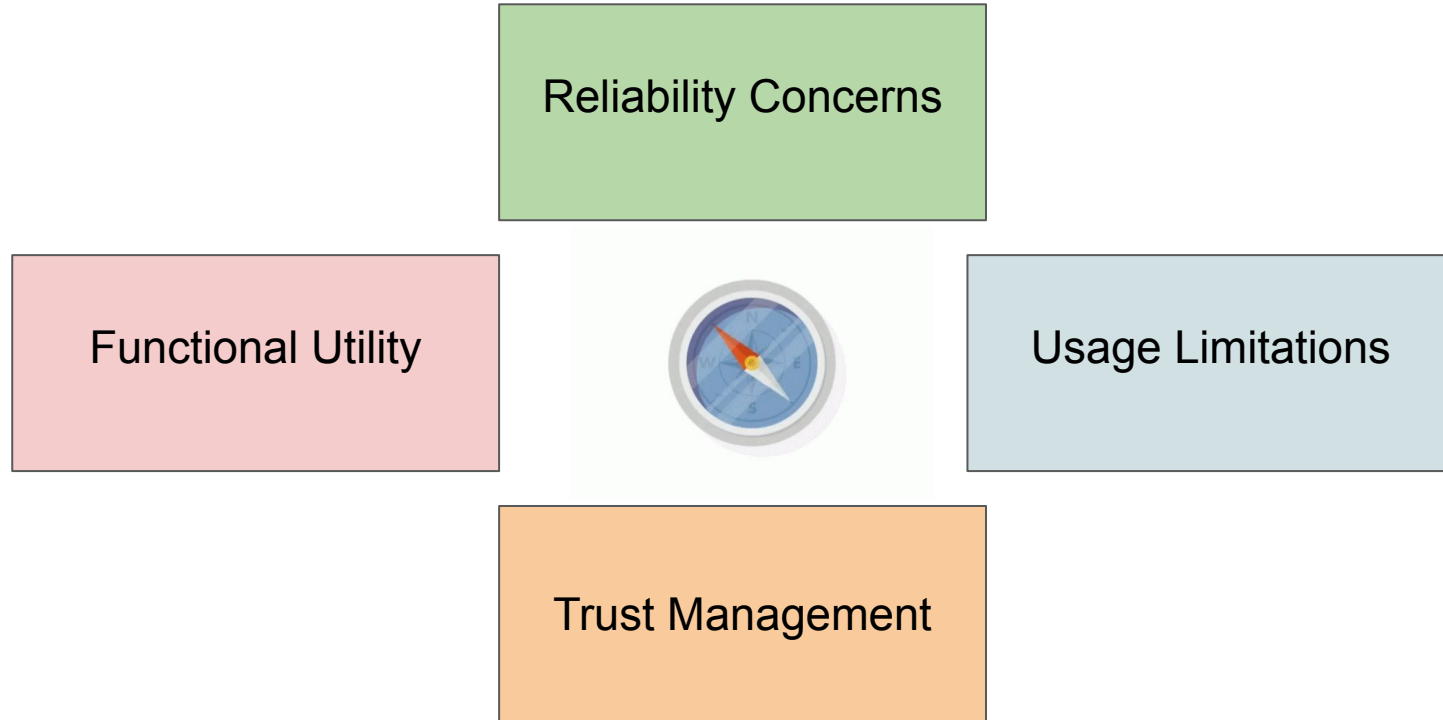Age: Between 23-34 with median age of 24

Location: Weimar

Proficiency : Medium or higher expertise with 3.5 model

Questions

1. Can you describe your experience with utilizing ChatGPT in your academic work?
2. How would you describe your level of confidence in the accuracy and reliability of ChatGPT-generated results?
3. What factors do you consider when determining the trustworthiness of ChatGPT outcomes?
4. Have you encountered any obstacles or constraints when using ChatGPT for academic purposes?
5. From your perspective, what role does ChatGPT play in supporting academic research?
6. Can you discuss any instances where you had to verify or cross-reference ChatGPT results to ensure their accuracy?
7. Do you believe that your familiarity with AI technologies influences your confidence in the outputs produced by ChatGPT?
8. What methods or approaches do you use to evaluate the reliability of content generated by ChatGPT?
9. If given 2 text (not known topic , not known person,same quality) what would you trust more in academic contexts?
10. In your opinion, what steps can the academic community take to cultivate greater confidence in AI technologies like ChatGPT among researchers and scholars?

The Results

4 Theme areas around Trust

Reliability Concerns

Functional Utility

Usage Limitations

Trust Management

Theme	Codes
Reliability Concerns	"Verification needed" "Source transparency" "Human-like errors" "Contextual reliability"
Functional Utility	"Tool not source" "Improves workflow" "Efficiency over accuracy"
Usage Limitations	"Helpful but limited" "Creative limitations"
Trust Management	"Trust with caution"

Table 1: Codes from interviews across categories

1. Reliability Concerns

Participants expressed significant concerns about the reliability of ChatGPT, particularly highlighting the need for verification of its outputs against credible sources.

Reliability often contextual

Participant 16: "For like code I guess it's pretty easy to check if it works and just paste it in your <u>IDE.</u>"

Participant 8: "When it comes to like things like paraphrasing and shortening paragraphs then I have to like look over the material and <u>compare it to the original.</u>"

Participant 12: "But <u>finding is usually restricted</u> to the recency of the papers."

2. Functional Utility

ChatGPT is perceived primarily as a supplementary tool rather than a standalone source. It enhances workflow efficiency by aiding in tasks like summarization and information retrieval, which are less dependent on depth of understanding and creativity.

Participant 17: "But if I write something on my own and then ask ChatGPT to help me rephrase it better or like write it academically, that way it really helps us."

Participant 1: "I use it for like getting information sometimes. So long text and paste it in there and I'll get a summary of it."

Participant 13: "I like that it automatically breaks down your topic into the most important points where I can then start thinking of my own way to complete it."

3. Usage Limitations

The limitations of ChatGPT are particularly evident in tasks requiring deep understanding or creativity. Participants noted that while the tool is helpful for generating starting points and ideas, For participant it may not work best for the work requiring deeper understanding and originality required for high-level academic work.

Participant 15: "Sometimes the _language also is like okay what is this_? So sometimes I even though ChatGPT will rephrase it for me

Participant 11: "We can't rely on ChatGPT with 100%. It's for sure because it will _give the most vague answers_ but sometimes it can be unnecessary for us."

Participant 14: "Sometimes the _ideas just aren't really useful_. And then I just feel like okay it's just wasting space."

4. Trust Management

Managing trust in ChatGPT involves cautious use, with users emphasizing the importance of verifying outputs and recognizing the tool's limitations.

Managing trust involves recognizing and adapting to the tool's limitations.

Participant 7: "Yeah, I mean <u>I don't really rely on it</u> because I don't know I've you know it's trained on so much stuff that of course there's a <u>lot of good stuff in there but there's also bad stuff</u>. In there. So I think as a concept that cannot really be trusted anyways."

Participant 2: "Accuracy <u>I feel it's pretty accurate.</u> Like the data it uses because the data chargeability has is before 2021 I guess. So there are certain things that are not updated as per now like if you will ask what is entry chip will add up it won't know. <u>So these cases we can ignore and rest if you are ask freely</u>

Key takeaways

1. **Reliability Concerns**: Needs for verification, transparency, and accuracy impact ChatGPT's perceived usefulness. Users find it beneficial but not a standalone tool, often requiring prompt refinement.
2. **Limitations in Creativity**: ChatGPT supports efficiency but shows limitations in creativity, with some users noting it occasionally produces unhelpful ideas.
3. **Trust Issues**: Many users hesitate to fully trust AI, preferring human expertise for personalized, reliable feedback. ChatGPT is sometimes seen as a "black box" lacking transparency.
4. **Human vs. AI Trust**: When choosing between AI and human-generated content, mostly trusted human expertise over AI, which is valued mainly for reliability of its data access.

Possible Practical implications for teaching

Using Chat GPT

1. Ideation
2. Breaking down ideas
3. Summarising

Not so much

1. Using it as data source
2. Knowledge on expert topic

eTeach
Netzwerk
Thüringen

Thank you !

Danke !

Reference

1. Pictures Designed by Freepik
2. Research paper on trust : Link