



From human to AI: University students' trust in generative AI as reference librarians

Di Wang, Xizhou Deng, Xinyu Lu, and Jianting Guo

DOI: <https://doi.org/10.47989/ir31iConf64203>

Abstract

Introduction. We report an investigation to compare students' trust in GAI versus human librarians in the scenario of university libraries' reference services and explore factors affecting students' trust in GAI.

Method. A within-subject design experiment with five tasks covering five disciplines was carried out based on a research model of antecedents for trust in AI, including AI (trustee) characteristics, human (trustor) characteristics, and decision situation (context) characteristics. Two versions of feedback (ChatGPT versus human) were generated for each task and randomly presented to participants without indicating feedback providers' identities.

Analysis. Quantitative analyses were carried out on the data of 146 participants, including descriptive statistics and reliability tests, bivariate correlations, and hierarchical multiple linear regression.

Results. Students' trust in GAI versus human librarians shows no significant difference. AI characteristics emerged as the strongest predictors for students' trust in GAI as reference librarians, especially the perceived knowledge of AI. Human and decision situation characteristics have limited effects on students' trust in GAI.

Conclusion. The findings refine AI trust antecedents, emphasising system-centric factors over user-centric ones. It supports a shift toward designing 'trustworthy by design' AI systems and prioritises verifiable knowledge bases and clarified knowledge presentation ability in developing GAI-embedded reference service systems.

Introduction

The rapid integration of generative artificial intelligence (GAI) into library services is reshaping the way library services are provided (Balnaves et al., 2025). GAI refers to '*computational techniques that are capable of generating seemingly new, meaningful content*' including text, images, audio, and video from training data (Feuerriegel et al., 2024, p. 111). With examples such as GPT-4, DALL-E, Copilot, and Gemini, GAI systems can not only be used to create new content mimicking the provided materials, but can assist humans as intelligent question-answering systems (Feuerriegel et al., 2024), which naturally fits to be applied to the scenario of reference services in libraries.

As one of the primary functions of libraries, reference services traditionally refer to assisting patrons in using or accessing library resources (GYANG, 2020). Reference librarians aim to establish contact between a user and the right information at the right time (Onifade & Sowole, 2011). Accuracy, thoroughness, timeliness, and authority are the common values associated with successful reference answers (Tyckoson, 2001). Reference services have a long-time adoption of AI technology like chatbots dating back to the mid-2000s (McNeal & Newyear, 2013). With the pervasive use of GAI and its instant feedback, many patrons now expect reference librarians to provide direct answers to their requests instead of a list of reference sources (Kaushal & Yadav, 2022). Drawing by the requirements of patrons, libraries try to shift the paradigm of reference services by implementing the latest technology solutions to enhance user experiences (Jha, 2023), such as bringing in GAI as virtual reference librarians around the clock (Vijayakumar & Sheshadri, 2019).

On the other hand, implementation of AI technology in libraries may not be as optimistic as imagined (Wheatley & Hervieux, 2019). Besides financial challenges and inadequate infrastructure (Hussain, 2023), patrons' awareness, adoption, and readiness to use AI-embedded library services are the main constraints (Ajani et al., 2022). Insufficient empirical studies on AI technologies in libraries further affect libraries' engagement in the current AI conversation (Wheatley & Hervieux, 2019). Yan et al. (2024) recently reviewed 53 AI application scenarios involving university students since 2017, yet none of these specifically examined AI implementation within library services.

While GAI promises efficiency gains for libraries, its adoption hinges on trust, a foundational element in human-AI collaboration that influences reliance, satisfaction, and long-term engagement (Shin, 2021). Trust in this context can be defined as the perception of confidence in librarians' reliability and integrity (Belanger et al., 2002). In reference services, trust has long been tied to librarians' empathy and contextual expertise (Phelps & Campbell, 2012), but GAI as reference librarians introduces unique challenges, including perceived opacity and algorithmic biases that may erode user confidence (Rudin, 2019).

Trust in AI has drawn increasing attention in human-AI interaction studies (Wang et al., 2024), due to its close connection to users' intention to accept or use AI systems. However, no well-examined models have been identified in current literature to predict users' trust in AI (Wang et al., 2024). Early studies focus on revealing antecedents of trust in non-human trustees, such as robotics, virtual, and embedded AI (Glikson & Woolley, 2020). Saßmannshausen et al. (2021) classified the antecedents into three groups: the trustor, the trustee, and the context. Antecedents of these three groups have been examined in various fields, such as production management (Saßmannshausen et al., 2021), customer service (Pelau et al., 2021), and legal decision-making (Kahr et al., 2023). Scenarios related to AI-embedded library services seldom appear in studies related to trust in AI. Some recent studies have started to investigate patrons' perspectives and experiences of AI implementation in libraries (Kaushal & Yadav, 2022). However, as one of the key stakeholders of university libraries, students' trust in GAI as librarians has not drawn enough attention in the current literature.

Considering the above research gaps, this study selects the scenario of reference services in university libraries to carry out empirical research regarding university students' trust in AI. It aims to compare university students' trust in GAI versus human librarians in providing reference services and reveal factors influencing their trust. A user test covering various discipline areas was carried out on university students. Two research questions guided the design of this study:

RQ1: Is there a difference between university students' trust in a human librarian and a GAI system to provide reference services?

RQ2: What are the factors affecting university students' trust in GAI providing reference services?

Method

To compare university students' trust in GAI versus human reference librarians, this study carried out a within-subject design experiment. To answer RQ1, five tasks from five discipline areas were included in the experiment. A reference librarian and a GAI system provided two versions of feedback to each task. Participants were asked to evaluate their trust in the feedback after completing each task without knowing the identity of the feedback provider. To answer RQ2, a research model (Figure 1) was developed based on the framework of antecedents for trust in AI by Saßmannshausen et al. (2021). This framework has been tested for its ability to explain AI trust within the scenario of production management. Since our study aims to investigate factors affecting university students' trust in GAI in reference services, this framework was thus employed as a structural foundation for the design of the study.

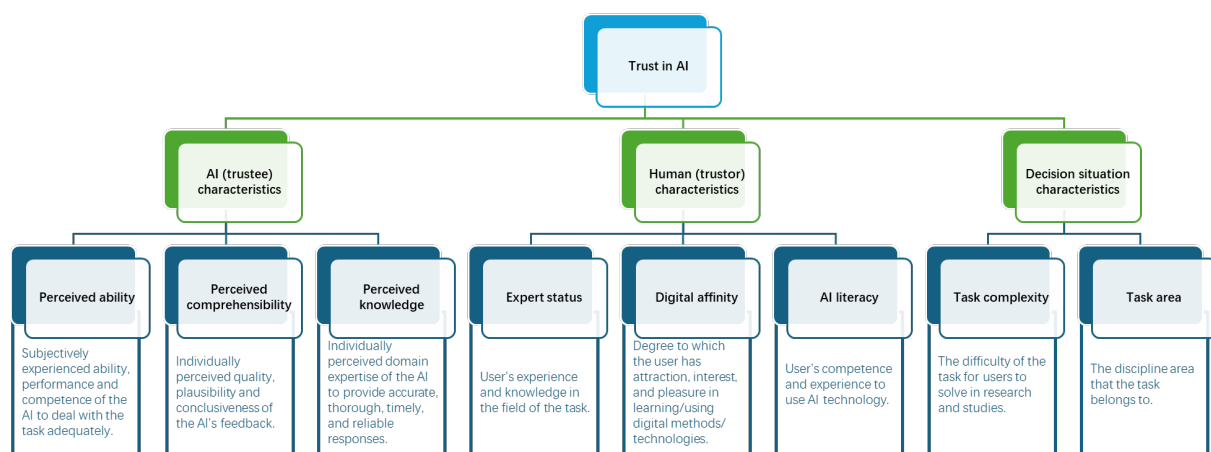


Figure 1. Research model of antecedents for trust in AI within reference services

The research model includes three aspects: AI (trustee) characteristics, human (trustor) characteristics, and decision situation (context) characteristics. AI characteristics originally include two elements (*perceived comprehensibility* and *perceived ability*). Because this study focuses on the scenario of library reference services and AI's domain knowledge about a task can affect users' trust in AI and its recommendations (Gomez et al., 2023), we added *perceived knowledge* as a third AI characteristic. Human characteristics originally include *expert status* and *digital affinity*. Because users' AI literacy reflects their competence in using AI technology (Wang et al., 2023), which includes judging the reliability of AI feedback, we included it as a third human characteristic. Decision situation characteristics originally consisted of predictability of the situation and error costs, reflecting the scenario of production management. In this study, we modified the predictability of the situation as *task complexity*, indicating the difficulty of the task for patrons to solve in their research or studies. Since reference services do not include costs for the company of a wrong decision by AI, we deleted error costs from the framework and added a new element, 'task area', to reflect the discipline area of the task for students to make decisions.

Seven 5-point Likert scale instruments were modified based on related studies (Gomez et al., 2023; Saßmannshausen et al., 2021; Wang et al., 2023) in order to fit the scenario of reference services. For task complexity, we asked participants to evaluate how difficult it was for them to complete the task. For task area, it was pre-decided and requires no input from the participants.

Experiment materials

The experiment was conducted online and used two types of materials: experiment tasks and a digital platform to guide participants through the testing process. Five tasks were selected from existing university assignment questions of five disciplines: computer science, physics, law, economics, and history, most of which were rarely explored in prior studies (Cavalcanti et al., 2021). Two feedback versions were developed for each task: One was generated by ChatGPT, prompted to act as a reference librarian providing feedback with explanations and reference sources, while the other was provided by five subject librarians in university libraries (one for each discipline). Both feedback versions were formatted consistently: Presenting answers and explanations first, followed by a reference sources list. Figure 2 shows five tasks and one example of both versions of feedback. The online platform Credamo was employed to create the experimental environment, direct participants, and gather data. Credamo randomly assigned one of the two feedback versions for each task to participants, and all items in the instruments were displayed in random order.

T1 (History): Why did the Soviet Union sign a non-aggression pact with Germany?

Answer from AI:	Answer from Librarian:
Geopolitical Considerations: The Soviet Union sought to avoid a direct military confrontation with Germany in the short term by signing the pact. In the late 1930s, the Soviet Union's military capabilities and national defense preparations were still inadequate, making it ill-prepared for a full-scale war with Germany. The treaty thus served as a means for the Soviet Union to buy time to strengthen its armed forces and defense infrastructure. Strategic Buffer Zone: The secret protocol attached to the pact delineated spheres of influence in Eastern Europe, enabling the Soviet Union to establish a strategic buffer zone in areas such as eastern Poland, the Baltic States, and Finland. By doing so, the Soviet Union aimed to secure its western borders and delay potential enemy advances in the event of future military conflicts. Distrust of Western Powers: The Soviet Union's lack of trust in Western countries, particularly Britain and France, was another critical factor behind the treaty. The 1938 Munich Agreement led the Soviets to believe that Britain and France were indifferent to Soviet security interests and might even tolerate German expansion eastward. As a result, the Soviet Union opted to reach an agreement with Germany to safeguard its own national security. German Objectives: From Germany's perspective, the non-aggression pact was instrumental in avoiding a two-front war. Hitler intended to first deal with France and the United Kingdom in the west before turning his attention to the Soviet Union. The pact thus allowed Germany to concentrate its military efforts on the Western Front without fear of immediate Soviet intervention. Economic Cooperation: The treaty also encompassed a series of economic agreements that were mutually beneficial. The Soviet Union agreed to supply Germany with raw materials such as oil, grain, and minerals, while Germany provided industrial goods and military technology in return. This economic collaboration was attractive to both parties given the international circumstances of the time. References: Molotov-Ribbentrop Pact - Wikipedia Peter Calvocoressi, "World Politics Since 1945" Richard Overy, "The Road to War"	The Molotov-Ribbentrop Pact, formally known as the Treaty of Non-Aggression between Germany and the Union of Soviet Socialist Republics, was signed in Moscow in 1939, shortly before the outbreak of the Second World War. Under the agreement, both parties pledged not to attack one another. Furthermore, if either party became the target of hostile actions by a third country, the other would not support that third country "in any form." Additionally, Germany and the Soviet Union agreed not to participate in any international alliances directed—either directly or indirectly—against the other. The primary aim of the treaty was to establish a preliminary basis for cooperation and mutual understanding between the two powers amidst their expansionist pursuits. For Germany, the pact marked a diplomatic triumph of its blitzkrieg diplomacy, while for the Soviet Union, it represented a strategic choice to maximize national interests under a rapidly shifting geopolitical landscape. Subjective Motivation: A profound distrust toward Britain and France shaped Soviet decision-making, coupled with a suspicion that the Western powers intended to redirect Nazi aggression eastward. Following the rise of fascist regimes in Germany and Italy, the Soviet Union proposed a policy of collective security to counter the expansionist ambitions of Hitler and Mussolini. However, the lack of a substantive response from Britain and France, who continued their appeasement policy toward fascist aggression, further deepened Soviet skepticism. Objective Conditions: The collapse of the Anglo-Franco-Soviet negotiations and the simultaneous covert talks between Britain and Germany were pivotal. Germany's annexation of Czechoslovakia posed a direct threat to Soviet western borders. Confronted with a deteriorating security environment, Stalin sought to either form a united anti-fascist front to deter further war or, at minimum, avoid Soviet entanglement in a looming conflict. While Britain initiated negotiations with the Soviet Union under mounting domestic and international pressure—partly to signal strength and compel German concessions—it also engaged in four months of secret bilateral talks with Germany. Though these secret talks ultimately failed due to irreconcilable differences, they intensified Soviet mistrust of British intentions. Strategic Calculations: In order to avoid premature involvement in war and the threat of fighting on multiple fronts, the Soviet leadership opted to "ally with the thief to guard against the bandit". From Stalin's perspective, Britain and France were demanding the Soviet Union risk a devastating war with Germany without offering substantial security guarantees in return. In contrast, Germany not only recognized Soviet freedom of action along the eastern Baltic coast, but also agreed to a mutual partition of Poland—offering the Soviet Union tangible strategic gains. Reference: https://www.163.com/dy/article/CRRDDH2R0519DF4V.html https://baike.baidu.com/item/%E8%BD%A7%E5%AD%A7%E4%B7%B1%E4%B7%B5%E7%BB%A7%E6%96%B4%E7%B4%A6/1631588 GENG Jin-ang On the Reason Why Stalin signed the Soviet German Non-aggression Treaty [J]. Journal of North China University of Water Resources and Electric Power (Social Sciences Edition), 2014, 36(03): 155-160. DOI: 10.13790/j.cnki.issn1008-4444.2014.03.044. Li Shuang, Cao Shengqiang, Research on the Molotov-Ribbentrop Pact in China (1979-1998) [J]. World History, 2001, (01): 96-97.

T2 (Economics): Why are some economic activities reasonable and effective from a micro perspective, but unreasonable and ineffective from a macro perspective?

T3 (Law): Why is the principle of equitable liability applicable to the High Torts Act?

T4 (Physics): Why does time change with speed in the theory of relativity?

T5 (Computer science): Why does deep learning outperform traditional computer vision methods in image recognition?

Figure 2. Experiment tasks and sample feedback from GAI and human librarians

Data collection and analysis

Figure 3 shows the procedure of the experiment. Participants first read the information and consent form to agree to participate, then they get the first task with feedback provided by human librarian or ChatGPT. Participants do not know the identity of the feedback provider, and they are asked to evaluate their trust in the provider based on the feedback. They will then answer three AI (trustee) instruments and evaluate the difficulty of the task. The process will repeat until all five tasks are completed. Then participants answer three human (trustor) instruments and one social-

demographic questionnaire about their gender, age, educational background, major, and trust tendency.

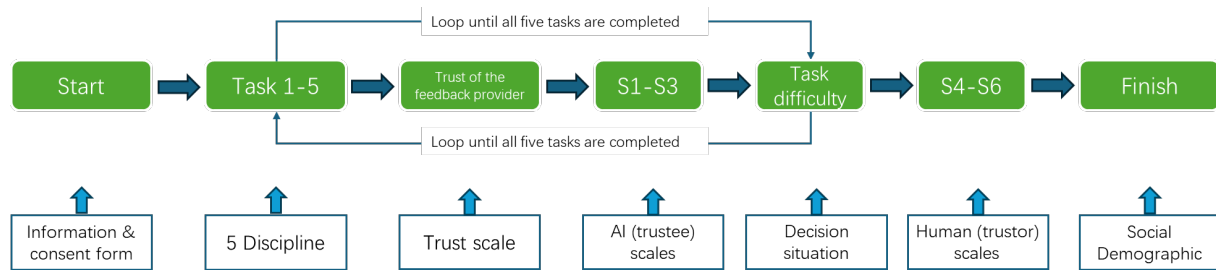


Figure 3. Experiment procedure

The participant recruitment was carried out via Credamo from July 28th to August 11th, 2024, targeting university students in China, including undergraduates, master's, and doctoral students. Each participant received compensation of 15 RMB upon completing the entire procedure.

Collected data was assessed through descriptive statistics and reliability tests, and bivariate correlations were used to identify preliminary associations. To evaluate the relative contributions of the three groups of characteristics, hierarchical regression models were employed, with predictors entered in blocks sequentially. This approach enabled a comparison of the explanatory power of AI, human, and situational factors while accounting for socio-demographic controls. Multicollinearity was assessed using variance inflation factors, and all analyses were conducted in SPSS, with significance set at $p < 0.05$.

Participants and reliability

A total of 150 participants volunteered for the experiment, 146 (97.3%) of whom completed the entire procedure. Socio-demographic distribution of the participants is shown in Table 1.

Topic	Dimension	Total	%	Topic	Dimension	Total	%
Gender	Male	43	29.50%	Educational background	Bachelor's degree	112	76.70%
	Female	103	70.50%		Master's degree	28	19.20%
Major	Arts	4	2.70%		PhD.	6	4.10%
	Business	45	30.80%	Age	<19	0	0.00%
	Education	15	10.30%		19	6	4.10%
	Engineering	20	13.70%		20	23	15.80%
	Health	3	2.10%		21	34	23.30%
	Humanities	18	12.30%		22	38	26.00%
	Law	6	4.10%		23	18	12.30%
	Science	12	8.20%		24	13	8.90%
	Social Science	14	9.60%		25	6	4.10%
	Technology	9	6.20%		>25	8	5.50%
Total		146	100%	Total		146	100%

Table 1. Participants' socio-demographics

Reliability of the seven instruments was tested with Cronbach's alpha. The values are all above 0.7 (Table 2), indicating high internal consistency. The Kaiser-Meyer-Olkin (KMO) measures also exceeded 0.7, indicating satisfactory sampling adequacy.

Instrument	Valid	%	Cronbach's Alpha	KMO
Trust in the reference librarian	730	100%	0.816	0.815
Perceived ability	730	100%	0.763	0.747
Perceived comprehensibility	730	100%	0.745	0.753
Perceived knowledge	730	100%	0.761	0.777
Digital affinity	146	100%	0.738	0.756
AI literacy	146	100%	0.709	0.865
Trust tendency	146	100%	0.845	0.722

Table 2. Reliability and adequacy test of instruments

Results

Students' trust in GAI as reference librarians

We compared students' trust in a human versus a GAI librarian in the scenario of reference services. The descriptive values for students' trust are shown in Table 3. The two independent samples T-test shows that though the average score (GAI=4.04, Librarian=4.02) of trust for AI is higher than for the human librarian, the differences were not significant with $p=0.65$.

Type	Variables	No	Mean	Median	Minimum	Maximum	Standard deviation
Dependent variable	Trust in librarian	368	4.02	4.20	1.0	5.0	0.59
	Trust in GAI	362	4.04	4.20	1.4	5.0	0.61
Independent variables	Perceived ability	362	3.93	4.00	1.3	5.0	0.71
	Perceived comprehensibility	362	3.82	4.00	1.3	5.0	0.81
	Perceived Knowledge	362	4.11	4.25	1.5	5.0	0.60
	Digital affinity	362	3.81	4.00	1.5	5.0	0.70
	AI literacy	362	3.96	4.00	2.2	4.7	0.33
	Expert status: History	362	3.16	3.00	1.0	5.0	1.12
	Expert status: Economics	362	2.98	3.00	1.0	5.0	1.09
	Expert status: Law	362	2.49	2.00	1.0	5.0	1.01
	Expert status: Physics	362	2.33	2.00	1.0	5.0	1.17
	Expert status: Computer science	362	2.83	3.00	1.0	5.0	1.18
	Task complexity	362	3.62	4.00	1.0	5.0	1.18
	Trust tendency	362	3.74	4.00	1.0	5.0	0.92

Table 3. Key figures of the dependent and independent variables

We selected data with GAI providing feedback and compared students' trust in GAI with different socio-demographic backgrounds. The two independent samples T-test shows no significant differences in the average and standard deviation of male and female students' trust in GAI. The one-way ANOVA shows that students' educational background ($F=2.657$, $p=0.072$) and major ($F=1.276$, $p=0.248$) have no significant effect on their trust in GAI, while students of different ages show different trust in GAI ($F=2.116$, $p=0.041$). Students older than 24 have significantly lower trust in GAI compared to students aged 21 and 22. Students' trust tendency shows a significant positive correlation with their trust in GAI, with the Spearman rank-order correlation coefficient $r=0.23$.

Correlation analyses

We carried out the Spearman rank-order correlation analysis of AI characteristics, human characteristics, decision situation characteristics, socio-demographic characteristics, and students' trust in GAI providing reference services. Results show that six variables have strong significant positive correlations with students' trust in GAI, which are perceived ability, perceived knowledge, perceived comprehensibility, digital affinity, AI literacy, and trust tendency ($r_s=0.50\sim0.68$, $p_s<0.001$). Perceived knowledge has the highest coefficient ($r=0.68$). Students' expert status in physics has a weak significant correlation with students' trust in GAI. All other variables have insignificant correlations with students' trust in GAI.

Regression model for students' trust in GAI as reference librarians

To further investigate factors affecting students' trust in GAI providing reference services (RQ2), a hierarchical multiple linear regression analysis was performed, with trust in GAI as the dependent variable. Predictors were entered in four blocks, following the research model: Block 1 includes AI characteristics (perceived ability, perceived comprehensibility, and perceived knowledge). Block 2 includes human characteristics (digital affinity, AI literacy, and expert status of five discipline areas). Block 3 includes decision situation characteristics (task complexity and task area). Block 4 includes socio-demographic characteristics (gender, age, educational background, major, and trust tendency).

The overall model was significant, explaining 67.5% of the variance in trust. Table 4 summarises the unstandardised (B) and standardised (β) coefficients, standard errors (SE), t-values, p-values, and VIFs from the final model. AI characteristics accounted for the majority of the explained variance. All three variables were significant positive contributors, with perceived knowledge emerging as the strongest ($\beta = 0.524$, $p < .001$). Human characteristics added small incremental variance ($\Delta R^2 = 0.003$, $\Delta F(7, 351) = 0.42$, $p = 0.882$), and no individual predictor was significant. Decision situation characteristics contributed a small but significant increase ($\Delta R^2 = 0.005$, $\Delta F(2, 349) = 3.67$, $p = 0.027$), though neither task area nor task complexity reached individual significance. Social-demographic characteristics explained an additional 0.4% of variance ($\Delta R^2 = 0.004$, $\Delta F(1, 348) = 4.58$, $p = 0.033$), with educational background as the only significant predictor.

Predictors	B	SE	β	t	p	VIF
Block 1: AI Characteristics						
Perceived ability	0.182	0.040	0.212***	4.52	<0.001	2.36
Perceived comprehensibility	0.162	0.033	0.217***	4.95	<0.001	2.07
Perceived knowledge	0.532	0.050	0.524***	10.71	<0.001	2.57
Block 2: Human Characteristics						
Digital affinity	0.011	0.035	0.013	0.32	0.752	1.70
AI literacy	-0.048	0.070	-0.026	-0.69	0.492	1.50
Expert status: History	0.005	0.019	0.010	0.27	0.784	1.35
Expert status: Economics	-0.005	0.020	-0.009	-0.26	0.792	1.39
Expert status: Law	0.013	0.021	0.022	0.62	0.535	1.30
Expert status: Physics	0.015	0.020	0.028	0.73	0.464	1.59
Expert status: Computer science	-0.027	0.021	-0.053	-1.30	0.196	1.78
Block 3: Decision Situation Characteristics						
Task area	0.022	0.014	0.051	1.53	0.126	1.17
Task complexity	0.030	0.018	0.058	1.64	0.102	1.36
Block 4: Social-Demographic						
Education background	0.091	0.042	0.067*	2.14	0.033	1.04
Model Summary	R ² =0.675					
	Adj. R ² =0.663					
	F (13.348) =55.68					

Note: Coefficients from the final model. B=unstandardised coefficient, β =standardised coefficient, SE=standard errors. * $p < .05$, *** $p < .001$. VIFs indicate no multicollinearity concerns.

Table 4. Hierarchical multiple regression results predicting students' trust in GAI providing reference services

Discussion

This study provides empirical insights into the factors shaping university students' trust in GAI providing reference services. Based on the analysis, students' trust in GAI versus human reference librarians shows no significant difference. In this study, participants did not know the identity of the reference feedback provider. They evaluated their trust in the provider based on the feedback for each task. On one hand, this indicates that students may think the quality of the feedback provided by human librarians has no significant differences compared to the one provided by GAI, which agrees with previous studies that manual feedback may not be more efficient than AI-generated feedback (Cavalcanti et al., 2021). On the other hand, our finding supports the view that when participants were deceived about the identity of AI, they tended to report the AI as more competent and helpful (Zhang et al., 2023). This indicates that merging GAI in reference services to provide high-quality feedback is a practical way for university libraries.

The results of the regression model show that AI characteristics emerged as the strongest predictors for students' trust in GAI providing reference services, especially the perceived knowledge of AI. This aligns with previous studies that AI characteristics are decisive to explain and build users' trust (Glikson & Woolley, 2020; Saßmannshausen et al., 2021), since they indicate the competence of AI systems (Glikson & Woolley, 2020). However, perceived ability is traditionally recognised as a strong influencing factor for trust in AI (Schaefer et al., 2016). In this study, the prominent effects of GAI's perceived knowledge may reflect the context of reference services, and at the same time, indicate that with GAI's ability in knowledge work becoming more remarkable

(Ritala et al., 2023), its domain knowledge related to working tasks has become more decisive in affecting users' trust.

Although correlation analysis shows digital affinity and AI literacy are positively related to students' GAI trust, the regression model shows that human (trustor) and decision situation characteristics have a limited effect on predicting GAI trust. This finding supports previous studies that decision situations may contribute to the explanation of GAI trust through mediated indirect effects (Safmannshausen et al., 2021; Schaefer et al., 2016). On the other hand, our findings are against the previous study's view that people with advanced AI literacy show higher trust in AI (Huang & Ball, 2024). This may be due to our experiment did not reveal feedback providers' identities to participants. Another explanation is that AI literacy does not directly build users' trust in AI but facilitates users with the ability to critically appraise AI outputs (Bećirović et al., 2025).

Conclusion

Despite the wide implementation of AI technology in library services, studies about patrons' trust in GAI as virtual librarians are limited. This study chooses university libraries' reference services, one of the primary functions of libraries and the earliest adoption of AI expert systems, as the scenario to investigate students' trust in GAI. The empirical results show no significant differences in students' trust in GAI versus human reference librarians. AI characteristics, including perceived ability, comprehensibility, and knowledge, strongly affect students' GAI trust, while users' characteristics and decision situation characteristics have no significant direct effects. Theoretically, these findings refine AI trust antecedents for information science, emphasising system-centric factors over user-centric ones in reference services and bringing in AI's perceived knowledge as a new antecedent to predict trust. It also supports a shift toward designing 'trustworthy by design' AI systems, integrating transparency, interpretability, and explainability features to enhance comprehensibility (Abdul et al., 2018). Practically, academic librarians could prioritise GAI-embedded reference service systems with verifiable knowledge bases, clarified knowledge presentation ability, and user-friendly interfaces to boost adoption. Future research could employ longitudinal designs to track students' trust evolution and investigate GAI librarians' effects on different dimensions of patrons' trust.

Acknowledgements

This work was supported by the National Natural Science Foundation of China Project No. 72104188 and Fundamental Research Funds for the Central Universities (Renmin University of China) 24XNF041.

About the authors

Di Wang is an Associate Professor in the School of Information Resource Management, Renmin University of China, China. She received her double PhDs from Macquarie University in Australia and Wuhan University in China. Her research interests are in big data management, human-AI interaction, open data utilisation, and smart services. She can be contacted at di.wang@ruc.edu.cn

Xizhou Deng is an undergraduate student in the School of Information Resource Management, Renmin University of China, China. His principal research interests lie in the field of human-AI interaction. He can be contacted at 2024201408@ruc.edu.cn

Xinyu Lu is an undergraduate student at Wuhan University. Her primary research interests focus on human-computer interaction and the application of artificial intelligence in libraries. She can be reached at Luxinyu0728@163.com

Jianting Guo is an undergraduate student in the School of Information Resource Management, Renmin University of China, China. His principal research interests lie in the field of human-AI interaction and AI's contribution in library services. He can be contacted at 2023202341@ruc.edu.cn

References

- Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and Trajectories for Explainable, Accountable and Intelligent Systems: An HCI Research Agenda Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada. <https://doi.org/10.1145/3173574.3174156>
- Ajani, Y. A., Tella, A., Salawu, K. Y., & Abdullahi, F. (2022). Perspectives of librarians on awareness and readiness of academic libraries to integrate artificial intelligence for library operations and services in Nigeria. *Internet Reference Services Quarterly*, 26(4), 213-230.
- Balnaves, E., Bultrini, L., Cox, A., & Uzwyshyn, R. (2025). *New Horizons in Artificial Intelligence in Libraries*. De Gruyter Saur. <https://doi.org/10.1515/9783111336435>
- Bećirović, S., Polz, E., & Tinkel, I. (2025). Exploring students' AI literacy and its effects on their AI output quality, self-efficacy, and academic performance. *Smart Learning Environments*, 12(1), 29. <https://doi.org/10.1186/s40561-025-00384-3>
- Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *The journal of strategic Information Systems*, 11(3-4), 245-270.
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, 100027. <https://doi.org/10.1016/j.caeai.2021.100027>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111-126. <https://doi.org/10.1007/s12599-023-00834-7>
- Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, 14(2), 627-660. <https://doi.org/10.5465/annals.2018.0057>
- Gomez, C., Unberath, M., & Huang, C.-M. (2023). Mitigating knowledge imbalance in AI-advised decision-making through collaborative user involvement. *International journal of human-computer studies*, 172, 102977.
- GYANG, C. H. (2020). Library reference services based on artificial intelligence. *Villanova Journal of Science, Technology and Management*, 2(1).
- Huang, K.-T., & Ball, C. (2024). The Influence of AI Literacy on User's Trust in AI in Practical Scenarios: A Digital Divide Pilot Study. *Proceedings of the Association for Information Science and Technology*, 61(1), 937-939. <https://doi.org/10.1002/pra2.1146>
- Hussain, A. (2023). Use of artificial intelligence in the library services: prospects and challenges. *Library Hi Tech News*, 40(2), 15-17.

- Jha, S. K. (2023). Application of artificial intelligence in libraries and information centers services: prospects and challenges. *Library Hi Tech News*, 40(7), 1-5.
- Kahr, P. K., Rooks, G., Willemsen, M. C., & Snijders, C. C. P. (2023). It Seems Smart, but It Acts Stupid: Development of Trust in AI Advice in a Repeated Legal Decision-Making Task. *Proceedings of the 28th International Conference on Intelligent User Interfaces*, Sydney, NSW, Australia. <https://doi.org/10.1145/3581641.3584058>
- Kaushal, V., & Yadav, R. (2022). The role of chatbots in academic libraries: An experience-based perspective. *Journal of the Australian Library and Information Association*, 71(3), 215-232.
- McNeal, M. L., & Newyear, D. (2013). Introducing chatbots in libraries. *Library technology reports*, 49(8), 5-10.
- Onifade, F. N., & Sowole, A. O. (2011). Reference services in a university library: awareness and perception of undergraduate students. *Pacific Northwest Library Association in Quantity*.
- Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy, and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in human behavior*, 122, 106855.
- Phelps, S. F., & Campbell, N. (2012). Commitment and Trust in Librarian-Faculty Relationships: A Systematic Review of the Literature. *The Journal of Academic Librarianship*, 38(1), 13-19. <https://doi.org/https://doi.org/10.1016/j.acalib.2011.11.003>
- Ritala, P., Ruokonen, M., & Ramaul, L. (2023). Transforming boundaries: how does ChatGPT change knowledge work? *Journal of Business Strategy*, 45(3), 214-220. <https://doi.org/10.1108/jbs-05-2023-0094>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. <https://doi.org/10.1038/s42256-019-0048-x>
- Saßmannshausen, T., Burggräf, P., Wagner, J., Hassenzahl, M., Heupel, T., & Steinberg, F. (2021). Trust in artificial intelligence within production management—an exploration of antecedents. *Ergonomics*, 64(10), 1333-1350.
- Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Human Factors*, 58(3), 377-400. <https://doi.org/10.1177/0018720816634228>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International journal of human-computer studies*, 146, 102551.
- Tyckoson, D. A. (2001). What is the best model of reference service? *Library trends*, 50(2), 183-196.
- Vijayakumar, S., & Sheshadri, K. (2019). Applications of artificial intelligence in academic libraries. *International Journal of Computer Sciences and Engineering*, 7(16), 136-140.
- Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & information technology*, 42(9), 1324-1337.

- Wang, D., Zheng, K., Li, C., & Guo, J. (2024). Transitioning to Human-Centered AI: A Systematic Review of Theories, Scenarios, and Hypotheses in Human-AI Interactions. *Proceedings of the Association for Information Science and Technology*, 61(1), 673-678.
<https://doi.org/https://doi.org/10.1002/pr2.1078>
- Wheatley, A., & Hervieux, S. (2019). Artificial intelligence in academic libraries: An environmental scan. *Information services & use*, 39, 347-356. <https://doi.org/10.3233/ISU-190065>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112.
<https://doi.org/https://doi.org/10.1111/bjet.13370>
- Zhang, G., Chong, L., Kotovsky, K., & Cagan, J. (2023). Trust in an AI versus a Human teammate: The effects of teammate identity and performance on Human-AI cooperation. *Computers in human behavior*, 139, 107536.
<https://doi.org/https://doi.org/10.1016/j.chb.2022.107536>

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).