The Influence of Framing, Domain and Task Type on Trust in AI

Stefano Guidi
DISPOC
University of Siena
Siena, Italy
stefano.guidi@unisi.it

Lorenzo Reina
DISPOC
University of Siena
Siena, Italy
lorenzo.reina@unisi.it

Francesco Currò
DISPOC
University of Siena
Siena, Italy
francesco.curro2@unisi.it

Enrico Cipriani
DII
University of Pisa
Pisa, Italy
enrico.cipriani@phd.unipi.it

Simone Grassini
DPS
University of Bergen
Bergen, Norway
simone.grassini@uib.no

Oronzo Parlangeli
DISPOC
University of Siena
Siena, Italy
oronzo.parlangeli@unisi.it

Abstract

As artificial intelligence (AI) technologies become increasingly integrated into decision-making processes, understanding how users develop trust in these systems is essential. While prior research has examined isolated factors such as transparency or system performance, less is known about how trust is shaped by the combined influence of narrative framing, task type, application domain, and individual user characteristics. This study investigates these dimensions through a preregistered online experiment (N = 280) employing a 3 (narrative: human-centred, technical, control) x 2 (task type: generative, recommendation) x 3 (domain: healthcare, insurance, relationships) mixed factorial design. Participants were presented with a narrative introduction to a hypothetical AI system and asked to evaluate its trustworthiness, perceived utility, and likelihood of adoption across multiple scenarios. Results showed that narrative framing alone had minimal impact on trust, suggesting that brief descriptions may be insufficient to shape user attitudes. In contrast, both domain and task type significantly influenced trust: AI systems used in insurance and healthcare were trusted more than those in personal relationships, and recommendation systems consistently outperformed generative ones in user evaluations. These findings highlight the importance of contextual and individual factors in fostering appropriate trust in AI.

Keywords

Trust, Artificial Intelligence, Human-AI interaction, Narrative framing, Human-centred AI design

ACM Reference Format:
Stefano Guidi, Lorenzo Reina, Francesco Currò, Enrico Cipriani, Simone Grassini, and Oronzo Parlangeli. 2025. The Influence of Framing, Domain and Task Type on Trust in AI. In *36th Annual Conference of the European Association of Cognitive Ergonomics (EACE) (ECCE 2025), October 07–10, 2025, Tallinn, Estonia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3746175.3747196

ECCE 2025, Tallinn, Estonia

ACM ISBN 979-8-4007-2033-8/25/10
https://doi.org/10.1145/3746175.3747196

1 Introduction

As AI systems become more advanced, they also grow more opaque, making their decision-making processes harder to interpret and trust. This "black box problem" limits transparency and raises ethical concerns [1, 8, 44], even as explainability tools continue to evolve [5, 26]. However, trust in AI is shaped by more than just transparency. It also depends on the context of use, the type of system, and the user's background. For instance, healthcare and education applications often inspire greater trust than domains like surveillance or online dating [15, 40]. Likewise, generative AI may evoke empathy or creativity, while recommendation systems are typically judged by their accuracy and reliability [6, 12]. Individual differences—such as AI literacy, prior experience, and attitudes toward automation—further influence how users assess AI systems [22, 37]. These perceptions are also shaped by broader cultural narratives and the ways people imagine and interpret the role of AI in their lives [9, 33].

Despite increasing scholarly interest in trust in AI, previous research has often examined its key determinants in isolation. This study addresses these gaps by systematically comparing user trust between AI types, generative and recommendation, while accounting for the effects of narrative framing (human-centred vs. technical) and application domain (healthcare, insurance, or relationships). Specifically, it investigates how these factors interact to shape user perceptions of trustworthiness, comfort in disclosing personal information, and willingness to recommend the AI system.

By investigating these interactions, this research contributes to both theoretical understanding of trust formation in human-AI interaction and practical knowledge for designing AI systems that engender appropriate levels of trust across diverse contexts and user populations. As AI continues to permeate critical domains, understanding the nuanced dynamics of trust becomes increasingly essential for responsible development and implementation.

2 Related Work

2.1 AI as a Black Box

AI systems are becoming more complicated, making them challenging to understand [8]. This lack of transparency makes it difficult to assess whether AI systems align with human values and interests, as trust fundamentally requires understanding how and why decisions are made, not merely system reliability [38].

Although Explainable AI (XAI) offers potential pathways toward greater transparency [5, 26], neural networks models have been criticized for their black-box nature, where only inputs and outputs are immediately accessible, creating ethical challenges that must be addressed due to their inherent lack of transparency [39]. The opacity of “black box” AI systems and algorithms raises ethical and trust challenges.

2.2 Foundations and Contextual Determinants of Trust in AI

Trust in AI systems represents a complex, multidimensional construct [16]. Mayer et al. [25] define trust as the willingness to accept vulnerability based on expectations of another’s actions. This definition emphasises vulnerability and expectation under uncertainty. Lee and See [24] describe trust as the belief that an agent will support one’s goals amid uncertainty and vulnerability.

Trust in AI is highly context-dependent. Fast and Horvitz [15] show that public perceptions differ by domain—positive in healthcare and education, but more fearful in warfare and surveillance. Studies highlight that trust hinges more on application context than technical accuracy [27, 30]. In healthcare, transparency and accountability are key due to ethical concerns [37], while in education, chatbots seen as caring foster greater disclosure [32]. Yan et al. [41] developed a definition drawing upon trustworthy characteristics, conceptualising trust as "a party’s belief in an application’s ability to fulfill a task as expected", emphasising dependability, security, and usability. This definition aligns with [12] identification of two core dimensions of AI trust: human-like trust (empathy and intention) and functional trust (competence and reliability).

Narrative framing plays a crucial role in shaping these dimensions, with Pataranutaporn et al. [31] demonstrating that priming users with different beliefs about an AI’s inner motives significantly alters perception of trustworthiness, empathy, and effectiveness. In interpersonal domains like dating, these effects become particularly pronounced: Wu and Kelly [40] AI-generated dating profiles were seen as equally attractive but notably less trustworthy. Similarly, Dekkal et al. [13] show in the insurance sector that perceived usefulness and ease of interaction drive trust in chatbots, while technology-related anxiety can significantly lower adoption intentions. These findings underscore trust as a key factor in interactions across human-technology contexts, not just individuals and organisations [34], robots [11, 21].

2.3 Recommendation versus Generative AI Systems

AI technology development has evolved along two fundamentally different paths, with Generative AI and Recommendation AI systems embodying distinct and contrasting approaches to artificial intelligence [2]. Recommendation AI systems utilise data analysis and machine learning techniques to suggest relevant information to users by analysing their past behaviour, preferences, and interests through algorithms such as clustering, collaborative filtering, and deep neural networks to generate personalised recommendations [3]. Generative AI, on the other hand, focuses on developing algorithms and models that can generate synthetic data resembling real-world data [7]. Unlike traditional AI approaches in recommendation systems, generative AI offers significant advantages by addressing data sparsity issues through modeling implicit feedback and creating personalized content tailored to meet diverse user needs [14].

Different types of AI systems evoke varying levels and types of trust, with generative systems often appearing more autonomous and capable of human-like interaction that can foster perceived empathy (chatbots or voice assistants’ emotional support) and effectiveness (information-seeking) [6, 35]. By contrast, recommender systems tend to elicit trust based on their perceived rationality and accuracy [12]. Studies indicate that perceived novelty and human-likeness of generative systems can intensify emotional responses, potentially either strengthening or weakening trust depending on users’ existing beliefs and AI literacy [19].

2.4 Individual Differences: AI Literacy and Attitudes

Trust evolves dynamically over time through increased familiarity and knowledge acquisition [17]. Familiarity contributes significantly to the development of experiential trust [20] and enables users to anticipate and interpret the other’s behaviour [10]. Moreover, familiarity fosters trust through previous interactions [18].

This evolving nature of trust is particularly relevant in AI, where user trust is far from uniform [29]. Zhang and Li’s [45] framework conceptualises human-technology interaction as a phenomenon influenced by system characteristics, user traits, tasks, and contextual factors—elements.

Furthermore, trust is also shaped by the broader cultural and societal narratives surrounding AI [9], which influence how individuals interpret, evaluate, and emotionally respond to these technologies [33]. Researchers identified that expectations and imagined affordances of AI influence users’ perceptions and understandings [28].

3 The study

Before performing our study we formulated four operative hypotheses. These hypotheses were pre-registered on the Open Science Framework Platform (OSF).

H1: A narration of AI will significantly increase the trustworthiness of AI tools in all applicative contexts with respect to the control group (absence of narrative introduction to AI). Although we expect narrations to prime participants to give different ratings, we did not anticipate a specific direction for this effect.

H2: Participants will give significantly different trust ratings according to the specific domain in which AI is applied: AI tools in insurance will be rated significantly more trustworthy than AI tools applied for medical purposes, which in turn will be rated significantly more trustworthy than AI tools applied to relationships. AI applications raise different user concerns depending on the domain [15, 40]. We expected lower trust in emotionally sensitive contexts (e.g., personal relationships) compared to domains like insurance, which involve primarily monetary risk. Computational systems may be seen as better suited for analytical tasks—such as risk assessment—than for handling personal or emotional matters.

H3: There will be a significant interaction between application context and narration: human-centered narration will increase the trustworthiness of AI for medical and relationship applications. At the same time, a technical narration will increase the trustworthiness of AI for insurance applications. A human-centered narrative could prime characteristics that are still generally considered as missing in AI systems, and thus could positively steer the judgments of the system in the domains in which we predict (H2) that it would be trusted less (personal relationships and medical). Conversely, a technical narrative, in which performance is emphasized using numbers and figures (e.g. number of parameters, size of training dataset), might positively affect trust in domains in which the use of statistics and numbers are important, such as the insurance domain [31].

H4: Generative AI tools will receive higher trustworthiness ratings than recommendation AI tools. Given the recent hype surrounding generative AI tools and the pervasive narrative that portrays them as possessing superior reasoning and precision, and their more "human-like" interactions, we expected users to view them as more trustworthy [19].

4 Methods

4.1 Participants

We recruited 280 Italian participants for the study using Prolific (https://www.prolific.com/), who received a fee of 1.5 GBP for their participation (approximately 1.79 EURO). The age of participants ranged between 21 and 67 ($M = 33.3, SD = 9.78$). The sample was approximately balanced for gender, with 130 females (46.4%), 148 males (52.9%) and 2 participants who reported "other" or chose not to disclose their gender. Concerning education, 38.2% of participants have a high-school or vocational diploma ($N = 107$), 25.4% a bachelor's degree ($N = 71$), 27.1% a master's degree ($N = 76$), and 8.2% a PhD or a postgraduate degree ($N = 23$). Three participants have only a middle school diploma (1.1%). After being informed about the study procedure, data conservation/treatment, and right to withdrawal, participants were directed to an online questionnaire implemented in the PsyToolkit platform (https://www.psytoolkit.org/). Before completing the questionnaire, they had to provide explicit consent to participate in the study. The study was approved by the Ethical Review Board of the University of Siena (act n. 12/2025).

4.2 Design and procedure

An experiment was conducted with a mixed 3x2x3 factorial design. We manipulated between-subjects the text presented to three groups of participants (Narrative about AI: human-centered, technical and control), and collected ratings about different AI applications in 6 scenarios deriving from the combinations of the levels of the within-subject factors Task type (two-level - generative and recommendation) and Domain (three-level - medicine, insurance and intimate relationships).

Participants were randomly assigned to one of the three Narrative conditions. The presentation order of the scenarios was counterbalanced across participants with a 6x6 Latin square. Every scenario of the Narrative variable was accurately and equally informative, with the same length and number of arguments proposed. A pretest of the three narratives was conducted before the experiment, with a different sample of participants (students). The text of the 3 Narratives scenario and Task-Domain scenario descriptions are reported in Supplementary Material 1.

The questionnaire consisted of three stages. In the first stage, we requested socio-demographic information (Prolific ID, gender, age, degree), interest in AI ("How interested are you in learning more about AI?") and self-assessed self-efficacy in AI text comprehension ("How well do you believe you can understand a text on AI?").

In the second stage, participants were randomly assigned to one of three scenario descriptions of a hypothetical AI system: history and development, technical characteristics and abilities, or control. In the history and development narrative, we shared the story and progression of AI, from the beginning until today, highlighting the key stages and innovations. In the narrative of characteristics and abilities, we provided an overview of AI performance in linguistic and cognitive tasks. Lastly, in the control narrative, we presented the recipe for the famous dessert "Tiramisu, " including detailed explanations (with no mention of AI). The length and complexity are balanced across the scenarios. Three questions follow each scenario to assess comprehension level, the length, and the interest aroused by the text on a 5-point Likert scale ranging from "Very few" to "Very much". A comprehension check followed, requiring at least four correct answers out of six; otherwise, participants repeated the second stage.

In the third stage, participants read six text-based scenarios combining task and domain variables (Task × Domain), presented in counterbalanced order. Each scenario began identically and asked participants to imagine interacting with the AI. They then rated trust, compliance, likelihood of recommending, and willingness to disclose using 7-point Likert scales (1 = "Not at all"; 7 = "Very much"), each on separate trials.

5 Results

5.1 Design and group equivalence checks

We initially analysed participants' characteristics across the between-subjects condition. Neither age nor any of the AI attitude and literacy variables differed significantly across groups (all ps between .16 and .87), nor for interest in the narrative topic (AI or making desserts) and self-efficacy for understanding the narrative scenarios. The results showed that there was no difference in the reported interest in learning about AI in the two experimental groups ($M_{human} = 5.49, SD = 1.26; M_{tech} = 5.65, SD = 1.18; p = .71$) and that the interest in learning about AI for both groups was significantly higher than the self-reported interest for learning about of making dessert ($M_{control} = 4.85, SD = 1.56$). No difference was found between the groups in the self-reported efficacy ($p = .133$). No differences were found in the ratings about the length of the texts ($p = .28$), but both the comprehensibility ($p < .001$) and the interestingness ($p < .001$) judgments varied significantly across narratives. Comprehensibility was lowest for the technical narrative ($M = 3.22, SD = 1.07$), significantly lower than for the human narrative ($M = 4.18, SD = 0.82, p < .001$) and for the control narrative ($M = 4.58, SD = 0.67, p < .001$), which, in turn, was rated as significantly more comprehensible than the human narrative ($p = .005$). Concerning the ratings for interestingness, the results showed that they were significantly lower for the technical narrative (M = 3.35, SD = 1.09) than for the other narratives ($M_{human} = 3.88, SD = 0.98, p < .001; M_{control} = 3.8, SD = 0.93, p = .007$), which were not rated as significantly different from each other ($p = .84$).

5.2 Tests of the research hypotheses

We analysed the data using linear mixed-effects models (LMM), fitting a different model for each dimension of AI perception (trust, willingness to comply with recommendations, likelihood of recommending the system, willingness to disclose the use of the system). In all the models, we included Narrative, Task type and Domain and their interactions as fixed effects. Sum coding was used for the categorical variables.

In Table 1 are reported the results of the tests of the main effects and interactions for all the fitted models. The full set of parameters estimated in the analyses is available as Supplementary Materials 2. For all the dependent variables, the results showed significant main effects of Domain and Task type (all ps < .001). A significant Task x Domain interaction was found for willingness to disclose the use of AI, and, marginally, also for trust and compliance. A marginally significant effect of Narrative was found for compliance only. Among the covariates, attitude toward AI (AIAS-4) was significantly and positively associated with all the dependent variables ($B_{trust} = 0.50, B_{compl} = 0.47, B_{recom} = 0.49, B_{discl} = 0.41, all ps <$ .001). The level of AI awareness (AILS-A) was significantly and negatively associated with trust ($B = -0.15$) and, marginally, also to compliance ($B = -0.14$) and likelihood of recommending the system to others ($B = -0.15$). No other main effects or interactions were significant in the analyses.

Regarding the marginally significant effect of Narrative on the willingness to comply and use the systems outcomes, pairwise comparisons of the marginal means showed no significant differences between the control condition ($M = 4.3, SE = 0.089$) and the human narrative ($M = 4.25, SE = 0.088, p = .898$) or the technical narrative ($M = 4.51, SE = 0.087, p = .238$) conditions. The mean scores in the technical narrative condition tended to be higher than in the human narrative condition, but the difference was only marginally significant in the pairwise comparison ($p = .097$).

Pairwise comparisons of the marginal means as a function of Domain (averaging across task types) revealed a similar pattern for all the variables. The lowest mean scores were found in the domain of personal relationships, and were significantly lower than in the other two domains (all $ps < .001$), and highest in the insurance domain, significantly higher than in the medical domain ($p < .05$). When the analyses were repeated separately for the different types of tasks, the results showed that this pattern was significant for all the dependent variables for recommendation tasks (all $ps < .001$), and also for generative tasks with the only exception of willingness to disclose the use of AI. For this variable, in fact, the scores for the relationships domain were significantly lower than for the other domains ($p < .001$), and no significant difference was found between the scores for the insurance and medical domains ($p = .338$).

Concerning the effect of Task type, the results showed that for all the variables, the scores (averaged across domains) were significantly higher for recommendation tasks than for generative tasks. However, when we followed-up the significant interactions by analysing the simple effects of Task type for the different domains, the results showed that for the personal relationships domain generative systems received significantly lower scores than recommender systems for all the variables (all $ps < .001$), while no task differences were found in the medical domain ($p > .117$). For the insurance domain, instead, the results showed for generative tasks lower trust in the app ($p < .001$), lower likelihood of recommending its use to others ($p = .002$) and willingness to disclose its use ($p < .001$) than for recommendation task, and no differences across task types for compliance ($p = .159$).

6 Discussion

The study tested whether narrative descriptions of an AI system influenced trust, use, recommendation, and disclosure. First of all (H1), we predicted that a narrative description could increase the trust in AI application and its outcomes, as compared to the control group [12, 31, 33]. Data did not support this hypothesis. No narrative significantly outperformed the control. While cognitive load from technical texts might explain slight differences, overall, narratives had minimal impact on trust and use.

Hypothesis 2 (H2) predicted that trust would vary by domain: highest in insurance, followed by medical, then personal relationships. This pattern was confirmed: trust, compliance, and recommendation ratings significantly differed across domains, with insurance and medical scoring above the midpoint, and relationships below. Only willingness to disclose did not differ significantly between insurance and medical. These results align with prior studies showing higher trust in pragmatic domains like healthcare and insurance, and lower trust in emotionally sensitive areas like dating or personal relationships [13, 15, 40].

We also predicted (H3) a significant interaction between application context and narrative description, and specifically that a human narrative of the AI would increase trust in the relationship and medical domain, while a technical narrative would increase it in the insurance domain. These predictions were not confirmed in the analyses, which did not find a significant interaction between domain and narrative for trust or the other variables.

Hypothesis 4 (H4) proposed that generative AI tasks elicit more trust than recommendation tasks, given recent hype around systems like ChatGPT and their success in various domains [4, 36]. However, results showed the opposite in both the insurance and personal relationships domains. This may reflect users' greater familiarity with recommender systems, which are widely embedded in everyday tools like streaming platforms and social media. According to the mere exposure effect [43], familiarity fosters positive attitudes. Moreover, while generative AI dominates media narratives, it is also associated with fears about job loss and existential risks [23]. Supporting our findings, a recent study found that trust was highest for recommendation tasks across 25 AI applications [42]. Interestingly, in the medical domain, no task-type differences emerged—possibly because the generative task (creating a diet plan) was perceived as similar to a recommendation.

6.1 Limitations

This study has several limitations. First, the number of tasks and domains was small —participants rated only one recommendation and one generative task per domain— limiting generalisability and raising the possibility that results reflect stimulus-specific effects. Future studies should include a broader range of tasks and domains to clarify effects, especially in the medical domain where task differences were absent. Second, the sample size may have been insufficient to detect interactions between narrative and domain, particularly given the minimal main effects. Although the

Table 1: Results of the tests of main effects and interactions in the LMMs for each dependent variable

	Trust		Compliance		Recommend		Disclosure	
Effect	F	p	F	p	F	p	F	p
Task	24.15	**<.001**	13.10	**<.001**	25.89	**<.001**	51.85	**<.001**
Domain	208.38	**<.001**	213.31	**<.001**	196.83	**<.001**	245.48	**<.001**
Narrative	1.60	.203	2.39	*.093*	1.49	.226	0.49	.611
Task × Domain	2.41	*.090*	2.91	*.055*	1.71	.182	11.71	**<.001**
Task × Narrative	1.37	.254	0.24	.786	1.81	.164	0.24	.784
Domain × Narrative	0.25	.912	0.58	.679	1.16	.330	0.45	.775
Task × Domain × Narrative	0.70	.590	0.54	.703	1.72	.144	0.66	.623

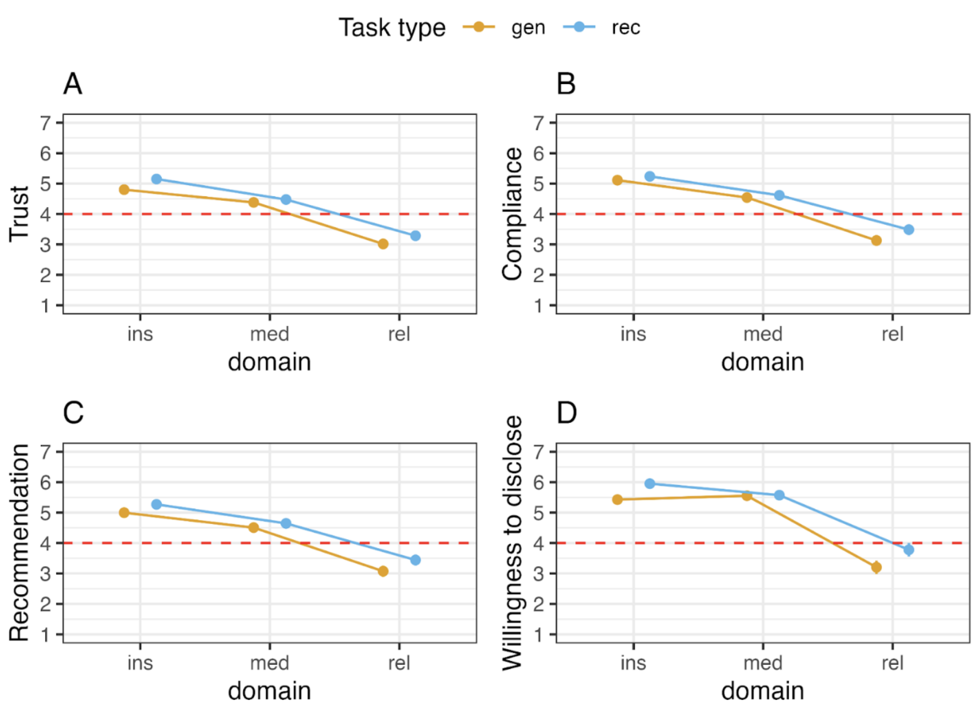

Figure 1: Plots of the marginal means for the ratings of (A) Trust, (B) willingness to to adopt the AI solutions, (C) likelihood of recommending the AI to others, and, (D) willingness to disclose use of the AI to others, as a function of Domain and Task type.

design targeted 90% power for small-to-moderate effects, assumptions about correlations may have led to reduced power. Lastly, all participants were Italian, limiting cross-cultural generalisability, especially given varying levels of AI exposure across populations.

6.2 Conclusion

This study examined how narrative framing, task type, and application domain influence trust in AI. Results showed that brief descriptions—whether technical or human-centered—did not significantly affect trust or use compared to a control, suggesting limited impact of surface-level framing. However, trust clearly varied by domain (highest in insurance, then medical, then relationships) and was greater for recommendation tasks. These findings highlight the importance of context in shaping AI trust. Promoting AI understanding through targeted education and training could enhance user confidence and adoption across domains.

Acknowledgments

To assist with improving the language in the paper, we used the Grammarly software. The paper accurately represents the authors' underlying work and novel intellectual contributions.

References

[1] Amina Adadi and Mohammed Berrada. 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6 (2018), 52138–52160. doi:10.1109/ACCESS.2018.2870052

[2] Ankur Aggarwal. 2025. Evolution of recommendation systems in the age of Generative AI. *International Journal of Science and Research Archive* 14, 1 (Jan. 2025), 485–492. doi:10.30574/ijsra.2025.14.1.0061

[3] Matthew O. Ayemowa, Roliana Ibrahim, and Muhammad Murad Khan. 2024. Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review. *IEEE Access* 12 (2024), 87742–87766. doi:10.1109/ACCESS.2024.3416962

[4] Ajay Bandi, Pydi Venkata Satya Ramesh Adapa, and Yudu Eswar Vinay Pratap Kumar Kuchi. 2023. The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet* 15, 8 (July 2023), 260. doi:10.3390/fi15080260

[5] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58 (June 2020), 82–115. doi:10.1016/j.inffus.2019.12.012

[6] Petter Bae Brandtzaeg and Asbjørn Følstad. 2017. Why People Use Chatbots. In *Internet Science*, Ioannis Kompatsiaris, Jonathan Cave, Anna Satsiou, Georg Carle, Antonella Passani, Efstratios Kontopoulos, Sotiris Diplaris, and Donald McMillan (Eds.). Vol. 10673. Springer International Publishing, Cham, 377–392. doi:10.1007/978-3-319-70284-1_30 Series Title: Lecture Notes in Computer Science.

[7] Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, and Lichao Sun. 2023. A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT. doi:10.48550/ARXIV.2303.04226 Version Number: 1.

[8] Davide Castelvecchi. 2016. Can we open the black box of AI? *Nature* 538, 7623 (Oct. 2016), 20–23. doi:10.1038/538020a

[9] Stephen Cave and Kanta Dihal. 2019. Hopes and fears for intelligent machines in fiction and reality. *Nature Machine Intelligence* 1, 2 (Feb. 2019), 74–78. doi:10.1038/s42256-019-0020-9

[10] Sandy C. Chen and Gurpreet S. Dhillon. 2003. Interpreting dimensions of consumer trust in e-commerce. *Information Technology and Management* 4, 2/3 (2003), 303–318. doi:10.1023/A:1022962631249

[11] Oscar Hengxuan Chi, Shizhen Jia, Yafang Li, and Dogan Gursoy. 2021. Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior* 118 (May 2021), 106700. doi:10.1016/j.chb.2021.106700

[12] Hyesun Choung, Prabu David, and Arun Ross. 2023. Trust in AI and Its Role in the Acceptance of AI Technologies. *International Journal of Human–Computer Interaction* 39, 9 (May 2023), 1727–1739. doi:10.1080/10447318.2022.2050543

[13] Massilva Dekkal, Manon Arcand, Sandrine Prom Tep, Lova Rajaobelina, and Line Ricard. 2024. Factors affecting user trust and intention in adopting chatbots: the moderating role of technology anxiety in insurtech. *Journal of Financial Services Marketing* 29, 3 (Sept. 2024), 699–728. doi:10.1057/s41264-023-00230-y

[14] Biao Du, Lin Tang, and Lin Liu. 2020. An Overview of Recommendation System Based on Generative Adversarial Networks. In *Proceedings of the 2020 3rd International Conference on E-Business, Information Management and Computer Science*. ACM, Wuhan China, 79–82. doi:10.1145/3453187.3453316

[15] Ethan Fast and Eric Horvitz. 2017. Long-Term Trends in the Public Perception of Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence* 31, 1 (Feb. 2017). doi:10.1609/aaai.v31i1.10635

[16] Andrea Ferrario, Michele Loi, and Eleonora Viganò. 2020. In AI We Trust Incrementally: a Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions. *Philosophy & Technology* 33, 3 (Sept. 2020), 523–539. doi:10.1007/s13347-019-00378-3

[17] Gefen, Karahanna, and Straub. 2003. Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly* 27, 1 (2003), 51. doi:10.2307/30036519

[18] David Gefen. 2000. E-commerce: the role of familiarity and trust. *Omega* 28, 6 (Dec. 2000), 725–737. doi:10.1016/S0305-0483(00)00021-9

[19] Ashish Ghosh, Debasrita Chakraborty, and Anwesha Law. 2018. Artificial intelligence in Internet of things. *CAAI Transactions on Intelligence Technology* 3, 4 (Dec. 2018), 208–218. doi:10.1049/trit.2018.1008

[20] R. Gulati. 1995. DOES FAMILIARITY BREED TRUST? THE IMPLICATIONS OF REPEATED TIES FOR CONTRACTUAL CHOICE IN ALLIANCES. *Academy of Management Journal* 38, 1 (Feb. 1995), 85–112. doi:10.2307/256729

[21] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 3 (May 2015), 407–434. doi:10.1177/0018720814547570

[22] Davinder Kaur, Suleyman Uslu, Kaley J. Rittichier, and Arjan Durresi. 2023. Trustworthy Artificial Intelligence: A Review. *Comput. Surveys* 55, 2 (Feb. 2023), 1–38. doi:10.1145/3491209

[23] Hisham O. Khogali and Samir Mekid. 2023. The blended future of automation and AI: Examining some long-term societal and ethical impact features. *Technology in Society* 73 (May 2023), 102232. doi:10.1016/j.techsoc.2023.102232

[24] J. D. Lee and K. A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, 1 (Jan. 2004), 50–80. doi:10.1518/hfes.46.1.50_30392

[25] Roger C. Mayer, James H. Davis, and F. David Schoorman. 1995. An Integrative Model of Organizational Trust. *The Academy of Management Review* 20, 3 (July 1995), 709. doi:10.2307/258792

[26] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (Feb. 2019), 1–38. doi:10.1016/j.artint.2018.07.007

[27] Rania Badr Mostafa and Tamara Kasamani. 2022. Antecedents and consequences of chatbot initial trust. *European Journal of Marketing* 56, 6 (June 2022), 1748–1771. doi:10.1108/EJM-02-2020-0084

[28] Gina Neff. 2016. Talking to Bots: Symbiotic Agency and the Case of Tay. (2016).

[29] Nessrine Omrani, Giorgia Rivieccio, Ugo Fiore, Francesco Schiavone, and Sergio Garcia Agreda. 2022. To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. *Technological Forecasting and Social Change* 181 (Aug. 2022), 121763. doi:10.1016/j.techfore.2022.121763

[30] Sangwon Park. 2020. Multifaceted trust in tourism service robots. *Annals of Tourism Research* 81 (March 2020), 102888. doi:10.1016/j.annals.2020.102888

[31] Pat Pataranutaporn, Ruby Liu, Ed Finn, and Pattie Maes. 2023. Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. *Nature Machine Intelligence* 5, 10 (Oct. 2023), 1076–1086. doi:10.1038/s42256-023-00720-7

[32] Joonas A. Pesonen. 2021. 'Are You OK?' Students' Trust in a Chatbot Providing Support Opportunities. In *Learning and Collaboration Technologies: Games and Virtual Environments for Learning*, Panayiotis Zaphiris and Andri Ioannou (Eds.). Vol. 12785. Springer International Publishing, Cham, 199–215. doi:10.1007/978-3-030-77943-6_13 Series Title: Lecture Notes in Computer Science.

[33] Laura Sartori and Giulia Bocca. 2023. Minding the gap(s): public perceptions of AI and socio-technical imaginaries. *AI & SOCIETY* 38, 2 (April 2023), 443–458. doi:10.1007/s00146-022-01422-1

[34] Katie Seaborn, Norihisa P. Miyake, Peter Pennefather, and Mihoko Otake-Matsuura. 2022. Voice in Human–Agent Interaction: A Survey. *Comput. Surveys* 54, 4 (May 2022), 1–43. doi:10.1145/3386867

[35] Chun Shao and K. Hazel Kwon. 2021. Hello Alexa! Exploring effects of motivational factors and social presence on satisfaction with artificial intelligence -enabled gadgets. *Human Behavior and Emerging Technologies* 3, 5 (Dec. 2021), 978–988. doi:10.1002/hbe2.293

[36] Christoph Treude and Margaret-Anne Storey. 2025. Generative AI and Empirical Software Engineering: A Paradigm Shift. doi:10.48550/ARXIV.2502.08108 Version Number: 1.

[37] Takane Ueno, Yuto Sawa, Yeongdae Kim, Jacqueline Urakami, Hiroki Oura, and Katie Seaborn. 2022. Trust in Human-AI Interaction: Scoping Out Models, Measures, and Methods. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7. doi:10.1145/3491101.3519772 arXiv:2205.00189 [cs].

[38] Warren J. Von Eschenbach. 2021. Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology* 34, 4 (Dec. 2021), 1607–1622. doi:10.1007/s13347-021-00477-0

[39] Weiyu Wang and Keng Siau. 2019. Artificial Intelligence, Machine Learning, Automation, Robotics, Future of Work and Future of Humanity: A Review and Research Agenda. *Journal of Database Management* 30, 1 (Jan. 2019), 61–79. doi:10.4018/JDM.2019010104

[40] Yihan Wu and Ryan M. Kelly. 2020. Online Dating Meets Artificial Intelligence: How the Perception of Algorithmically Generated Profile Text Impacts Attractiveness and Trust. In *32nd Australian Conference on Human-Computer Interaction*. ACM, Sydney NSW Australia, 444–453. doi:10.1145/3441000.3441074

[41] Zheng Yan, Yan Dong, Valtteri Niemi, and Guoliang Yu. 2013. Exploring trust of mobile applications based on user behaviors: an empirical study. *Journal of Applied Social Psychology* 43, 3 (March 2013), 638–659. doi:10.1111/j.1559-1816.2013.01044.x

[42] Ruoxin Yang, Sisheng Li, Yawei Qi, Jiali Liu, Qinghua He, and Haichao Zhao. 2025. Unveiling users' algorithm trust: The role of task objectivity, time pressure, and cognitive load. *Computers in Human Behavior Reports* 18 (May 2025), 100667. doi:10.1016/j.chbr.2025.100667

[43] Robert B. Zajonc. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology* 9, 2, Pt.2 (1968), 1–27. doi:10.1037/h0025848

[44] Carlos Zednik. 2021. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology* 34, 2 (June 2021), 265–288. doi:10.1007/s13347-019-00382-7

[45] Ping Zhang, Lina Li, and Syracuse University. 2005. The Intellectual Development of Human-Computer Interaction Research: A Critical Assessment of the MIS Literature (1990-2002). *Journal of the Association for Information Systems* 6, 11 (Nov. 2005), 227–292. doi:10.17705/1jais.00070