

The Privacy Paradox in Human–AI Communication: Why People Share More Information with Machines than with Other Humans

Umar Fauzi Bahanan¹⁾, Tantri Relatami²⁾

¹ Ilmu Komunikasi, FISIP, Universitas Nasional

² Ilmu Komunikasi, FISIP, Universitas Nasional

E-mail: ¹⁾ umarfauzi.bahanan@civitas.unas.ac.id, ²⁾ tantri@civitas.unas.ac.id

Abstract

This systematic literature review investigates the privacy paradox in human-AI interaction, focusing on why users frequently disclose sensitive personal information to artificial agents despite harboring stated privacy concerns. Utilizing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, this study synthesizes literature spanning 2015 to 2025. Following rigorous screening protocols, an initial corpus of 1,428 articles was refined to a final synthesis of 86 peer-reviewed studies. The review identifies five primary thematic dimensions: (1) the amplification of self-disclosure through anthropomorphism and social presence; (2) the transformative privacy calculus balancing utility against vulnerability; (3) the psychological mechanics of de-identified intimacy; (4) architectural parameters and conversational conditions influencing privacy salience; and (5) contextual variances among vulnerable populations, including patients and the elderly. The principal findings indicate that conversational AI effectively overrides traditional privacy-protective behaviors by simulating non-judgmental empathy and hyper-personalized utility, creating an "illusion of isolation" wherein users feel both socially supported and technically unobserved. This review contributes to Human-Computer Interaction (HCI) by extending privacy calculus theory into the realm of generative agents. Ultimately, the study addresses a critical research gap by mapping the specific conversational architectures that bypass cognitive privacy defenses, offering empirical guidelines for designing ethically compliant, privacy-preserving AI frameworks.

Keywords: Conversational agents, Human-AI interaction, Privacy calculus, Privacy paradox, Self disclosure

1. INTRODUCTION

The rapid diffusion of artificial intelligence (AI) technologies has propelled humans into an era in which interactions with intelligent systems permeate everyday life, extending beyond technical tasks to the realm of communication and social exchange (Guo & Chen, 2025; Li & Rau, 2021). In medicine, (Rolando, 2024) AI-enabled diagnostic and decision-support systems increasingly participate in conversations about health, sometimes competing with or augmenting human physicians, while patient evaluations reveal lingering preference for human judgment in functional

assessments despite improvements in accuracy when AI is involved (Guo & Chen, 2025). This shift is echoed in caregiving and home contexts, where AI-enabled robots and conversational agents now engage users in sensitive, privacy-relevant exchanges and negotiate trust, safety, and perceived usefulness in intimate care scenarios (Amin et al., 2024; Chien et al., 2025; B. Lee et al., 2023). (Ingriana et al., 2024) In consumer and public-facing domains, AI-driven assistants embedded in devices, services, and platforms shape how people disclose information, form impressions of machine agents, and navigate concerns about privacy, trust, and CSR in technology providers (Cloarec et al., 2024; Kronemann et al., 2023; Mou & Meng, 2024; X. Sun et al., 2025). Collectively, these trajectories illustrate how AI has moved from supporting human tasks to mediating human-to-human-like communications across multiple everyday settings (Guo & Chen, 2025; Li & Rau, 2021).

In healthcare, patients generally trust human physicians more than AI diagnoses, although providing explanations significantly improves trust and evaluations of both AI and human physicians (Guo & Chen, 2025). Complementary work on mental health and empathy shows that transparency about the authorship of AI-generated versus human-generated stories can influence user empathy and acceptance, underscoring the social implications of making AI more legible and trustworthy in emotionally salient contexts (Shen et al., 2024). In elder care and caregiving robotics, trust, privacy concerns, and perceived usefulness interact to shape the elderly's willingness to share personal health information, with privacy considerations and social isolation being key factors in the adoption of AI-enabled caregiver robots (Amin et al., 2024). These patterns reveal that, even as AI can broaden access to diagnostic or supportive services, designing interactions that preserve relational quality and address patient privacy remains essential to sustaining engagement and adherence (Amin et al., 2024; Guo & Chen, 2025; Shen et al., 2024).

The expansion of AI into privacy-sensitive conversational domains has foregrounded important questions about user resistance, disclosure, and social perception (Tan, 2022). Studies on intelligent voice assistants and humanoid versus non-human agents show that perceived privacy risks, anthropomorphism, and corporate social responsibility (CSR) beliefs can modulate user willingness to disclose, with privacy concerns potentially undermining adoption yet CSR and perceived warmth from anthropomorphic agents capable of attenuating such concerns under certain conditions (Chien et al., 2025; Mou & Meng, 2024; X. Sun et al., 2025). Theoretical frameworks such as privacy calculus (risk–benefit trade-offs) have been extended into transformative privacy calculus, incorporating affective and relational dimensions like happiness with the internet, trust, and information-sharing incentives to explain personalization-related disclosures (Cloarec et al., 2024). Empirical evidence across AI-generated content platforms and chatbots also documents the persistence of a privacy paradox, wherein cognitive awareness of privacy risks does not always translate into reduced information disclosure when emotional and contextual factors promote perceived benefits (Meng & Liu, 2025; Zhou & Tang, 2025). These insights illuminate the complex calculus users perform when engaging with AI in everyday services, where personalization and social presence may tilt decisions toward disclosure despite concerns (Cloarec et al., 2024; Kronemann et al., 2023; Meng & Liu, 2025).

Beyond abstract theory, the social–emotional dimensions of AI-mediated interaction appear central to user acceptance and ongoing engagement. Research on agent self-disclosure and empathy demonstrates that AI self-disclosure, appearance, and scenario relevance can significantly influence

users' willingness to disclose and their felt empathy toward AI partners, with effects moderated by the context in which interaction occurs (e.g., home environments with children or intimate partnerships) (B. Lee et al., 2023; Shen et al., 2024; Tsumura & Yamada, 2023). The interactional value of AI is further illustrated by studies on joint IoT-conversational agent use, where shared use can enhance perceived closeness and conversational quality between human partners and AI systems, thereby shaping the nature of human–AI dialogue (Li & Rau, 2021). In special-purpose social technologies, chatbot-mediated reminiscence tools show that structured guidance for conversational exchanges can strengthen positive affect, connectedness, and reciprocal self-disclosure among close others, highlighting the potential of AI to support meaningful social bonds rather than merely execute tasks (Jiang et al., 2025). Together, these findings underscore that social presence, empathy, and mutual disclosure are not incidental byproducts but core determinants of the success and acceptability of AI as a communicative partner (Jiang et al., 2025; Li & Rau, 2021; Shen et al., 2024; Tsumura & Yamada, 2023).

From a theoretical and design perspective, the literature converges on core principles for navigating AI-driven communication: balancing personalization with privacy, leveraging social–emotional design without eroding user autonomy, and ensuring transparency to foster trust. The private calculus tradition augmented by explorations of the personalization–privacy paradox and parasocial interaction offers a robust lens for predicting when users will disclose information to AI and how design choices (e.g., anthropomorphism, perceived warmth, and CSR signals) shape that disclosure (Cloarec et al., 2024; Kronemann et al., 2023; Meng & Liu, 2025; Zhou & Tang, 2025). Transformative privacy calculus, which links affective states (e.g., happiness) and social exchange considerations to information disclosure, provides a nuanced account of how trust and information collection concerns co-determine user behavior in personalized AI contexts (Cloarec et al., 2024). Meanwhile, cross-domain work on AI literacy, trust-building, and public adoption highlights that higher AI literacy can enhance trust and reduce algorithm aversion, though cultural and contextual factors may modulate these effects (Obadă et al., 2026). These threads imply practical guidelines for AI designers and policymakers: articulate privacy policies clearly, calibrate anthropomorphic cues to user expectations, and integrate explainability and transparency into AI-driven interactions to sustain long-term engagement across domains ranging from healthcare to consumer services (Chien et al., 2025; Obadă et al., 2026; Shen et al., 2024).

The evidence base across medicine, caregiving, mental health, consumer services, and everyday interpersonal contexts demonstrates that AI can augment and, in some cases, substitute for traditional human interactions, but this substitution introduces challenges related to trust, empathy, privacy, and the quality of relational engagement (Amin et al., 2024; Guo & Chen, 2025; Li & Rau, 2021; Shen et al., 2024). Designing AI that harmonizes functional benefits with relational depth requires careful attention to transparency, personalization, privacy governance, and culturally informed expectations about machine social presence, as underscored by contemporary studies on privacy calculus, anthropomorphism, and social–emotional design in AI-enabled communication (Cloarec et al., 2024; Kronemann et al., 2023; Meng & Liu, 2025; Zhou & Tang, 2025). Social and emotional fabric that characterizes human communication (Jiang et al., 2025; Obadă et al., 2026; Shen et al., 2024; Tsumura & Yamada, 2023).

1.1. RESEARCH QUESTION

- a. Why do users exhibit a higher willingness to disclose sensitive information to artificial agents than to human interlocutors?
- b. How do perceived benefits, social presence, and relational dynamics attenuate or override stated privacy concerns?
- c. Under what specific architectural parameters and conversational conditions do users prioritize interactive utility and socio-emotional engagement over data security?

1.2. RESEARCH OBJECTIVES

- a. **Investigate** the psychological and cognitive factors that drive users to exhibit a higher willingness to disclose sensitive information to artificial agents compared to human interlocutors.
- b. **Analyze** the mechanisms through which perceived benefits, social presence, and relational dynamics attenuate or override stated privacy concerns in human-AI interaction.
- c. **Evaluate** the specific architectural parameters, interface designs, and conversational conditions under which users prioritize interactive utility and socio-emotional engagement over data security.

1.3. RESEARCH SIGNIFICANCE

This systematic literature review offers a substantial theoretical contribution by synthesizing fragmented findings into a cohesive framework that extends privacy calculus theory and the Media Equation. By focusing on generative language models and anthropomorphic chatbots, this study elucidates how the "transformative privacy calculus" operates in real-time, dynamic conversations. It shifts the theoretical paradigm from viewing privacy as a static boundary-management process to an interactive, emotionally mediated negotiation, providing a foundational baseline for future HCI theories regarding machine-simulated intimacy.

The review offers practical guidance for designing AI systems with privacy-by-design, including adding reminders about AI identity and data collection to support informed consent. Academically, it bridges computer science, psychology, health informatics, and ethics, while clarifying the human–AI privacy paradox and supporting future research.

2. RESEARCH METHOD

2.1. RESEARCH DESIGN AND PROTOCOL

This review strictly adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological rigor and reproducibility. The research protocol was pre-registered on the Open Science Framework (OSF) prior to the commencement of data extraction to mitigate reporting bias. The systematic process involved iterative cycles of identification, screening, eligibility assessment, and inclusion, validated at each stage by two independent reviewers to maintain objectivity. Any discrepancies between reviewers were resolved

through consensus discussions or adjudication by a third senior researcher, ensuring consistent application of inclusion and exclusion criteria throughout the review process.

This rigorous approach minimized subjective bias and enhanced the reliability of the synthesized findings, thereby strengthening the validity of the conclusions drawn regarding the human-AI privacy paradox. This comprehensive methodology ensures that the emergent theoretical frameworks and practical guidelines are empirically grounded and directly address the nuanced challenges of privacy in conversational AI (Gumusel, 2024). The systematic review included a diverse range of studies to capture the multifaceted aspects of privacy, security, and trust perceptions in conversational AI, acknowledging that privacy is a context-dependent construct rather than a static one (Herriger et al., 2025; Leschanowsky et al., 2024, p. 26).

2.2. Search Strategy

A comprehensive literature search was conducted across six academic databases: Scopus, Web of Science, IEEE Xplore, ACM Digital Library, PubMed, and APA PsycInfo. Studies published between January 2015 and May 2025 were identified using Boolean keywords related to conversational AI, privacy, self-disclosure, and human-AI interaction. Searches were conducted in title, abstract, and keyword fields, followed by forward and backward citation screening to ensure relevant literature coverage (Prpić et al., 2015, p. 351).

The extensive search strategy aimed to maximize sensitivity while maintaining specificity, thereby optimizing the retrieval of pertinent literature across various disciplinary perspectives (Dang & Li, 2025, p. 3). Duplicates were then removed using automated software, followed by a manual check for any remaining redundancies (Cevasco et al., 2024, p. 3; Shrestha et al., 2024, p. 2). The initial search yielded a substantial number of studies, which were then systematically screened against predefined inclusion and exclusion criteria (Zhai et al., 2024, p. 9). For instance, our search terms encompassed keywords such as "trust," "personality," "individual difference," "AI," "machine learning," "autonomous," "agent," "bot," "chatbot," and "robot" to comprehensively cover the breadth of human-AI interaction (Riedl, 2022, p. 2026). This systematic approach ensured the capture of studies exploring the intricate relationships between human psychological factors and AI design, particularly concerning user perceptions of privacy and autonomy in conversational contexts (Ling et al., 2021, p. 1038).

2.3. Study Selection

The study selection utilized a modified PICO framework (Population, Intervention, Context, Outcome) adapted for non-clinical research. The Population included users engaging with AI; the Intervention (or phenomenon) was interaction with conversational AI; the Context was HCI in medicine, mental health, and consumer services; and the Outcome was self-disclosure versus privacy concerns. (Li et al., 2026) Following duplicate removal, titles and abstracts were screened using a standardized Excel form driven by Boolean logic (AND/OR matching of inclusion criteria). Two independent reviewers conducted the screening. (Asadi et al., 2025, p. 3) Inter-rater reliability was robust, yielding a Cohen's kappa coefficient of 0.89, indicating near-perfect agreement.

Discrepancies were resolved through consensus discussion mediated by a third senior researcher. (Baffsky et al., 2022, p. 339)

Full-text articles were independently reviewed based on methodological rigor and privacy-related outcomes. Studies published before 2018 or those not addressing ChatGPT, GPT, or similar large language models were excluded to maintain contemporary relevance (Labadze et al., 2023, p. 8; Liu, 2024, p. 4). The refined selection focused on recent privacy issues in conversational AI (Omar et al., 2024, p. 6). Data were systematically extracted, including publication details, participant characteristics, AI features, user experience, privacy concerns, and mitigation strategies (Li et al., 2023, p. 10).

2.4. Inclusion and Exclusion Criteria

Criteria were strictly applied to isolate peer-reviewed, high-impact research relevant to the specific parameters of the privacy paradox.

Table 1 Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Publication Year	2015 – 2025	Pre-2015
Language	English	Non-English publications
Article Type	Peer-reviewed full-length academic articles, premier conference proceedings	Non-English publications
Subject Area	Human-Computer Interaction (HCI), Psychology, Informatics	Preprints, editorials, book reviews, white papers, non-peer-reviewed
Open Access	Both Open Access and Paywalled (accessible via institutional access)	Unrelated fields (e.g., hardware engineering without user studies)
Population Focus	Users engaging with AI (Patients, Elderly, Consumers)	Unavailable full texts
Thematic Relevance	Explicitly investigates privacy paradox, self-disclosure, or privacy calculus in human-AI interaction	AI-to-AI interaction, purely algorithmic efficiency studies

2.5. Quality Assessment Protocol

A rigorous Quality Assessment (QA) protocol was implemented to ensure the validity of the synthesized data.

Table 2 Quality Assessment Criteria

Criterion	Sub-criteria Description (Max 5 points per criterion)
Coverage	Relevance to the research questions; clarity of the population and context.
Methodology	Rigor of study design; appropriateness of data collection methods; sample size adequacy.
Objectives	Clear articulation of hypotheses or research goals; theoretical grounding.
Findings	Robustness of data analysis; clarity of conclusions; acknowledgment of limitations.

Articles were evaluated using a 0–2 scoring system across 10 distinct sub-criteria, yielding a maximum score of 20. A 70% threshold (score ≥ 14) was required for inclusion. Quantitative studies were assessed utilizing the Effective Public Health Practice Project (EPHPP) tool, while qualitative studies were evaluated using the Critical Appraisal Skills Programme (CASP) checklist. Inter-rater reliability for QA achieved a Cohen’s kappa of 0.87. Initially, 1,428 articles passed initial database retrieval; following duplicate removal and title/abstract screening, 314 full-text articles were assessed for eligibility. Ultimately, 86 articles successfully met all quality thresholds and entered the final thematic synthesis.

2.6. Data Extraction Process

Data extraction was executed using a heavily structured Excel form comprising 25 fields divided into 5 major categories: (1) Study Characteristics (author, year, geography, sample size); (2) Methodological Details (research design, data collection tools); (3) Theoretical Frameworks (specifically tracking applications of Privacy calculus theory, transformative privacy calculus, and the Media Equation); (4) Key Findings (quantitative metrics, qualitative themes); and (5) Limitations. The extraction form underwent pilot testing on 5 randomly selected articles to calibrate reviewer consistency. Dual independent extraction was performed on all 86 final articles to eliminate transcription bias.

Discrepancies identified during the dual extraction process were resolved through a consensus meeting involving a third reviewer, ensuring data integrity and accuracy. This systematic approach facilitated a comprehensive and unbiased collection of pertinent information, underpinning the subsequent rigorous analysis. The collected data were subsequently synthesized using a thematic analysis approach, wherein recurring themes and patterns related to the privacy paradox in AI interactions were identified and categorized to address the research questions comprehensively. This systematic data extraction process is consistent with established methodologies for literature reviews, ensuring the robustness and reliability of the synthesized evidence (Gotsch & Schögel, 2021, p. 269; Moy et al., 2024, p. 54).

2.7. Data Analysis Techniques and Bibliometric Analysis

Bibliometric analysis was conducted using VOSviewer (v1.6.18) to objectively map the intellectual structure of the corpus. Three specific techniques were applied: co-citation analysis

(restricted to a minimum threshold of ≥ 3 citations), keyword co-occurrence analysis (requiring a minimum threshold of ≥ 5 article occurrences), and bibliographic coupling. The algorithmic clustering within VOSviewer generated a spatial map from which 5 distinct thematic clusters emerged, guiding the subsequent qualitative synthesis. (Bibri et al., 2023, p. 7) Complementary to this, a thematic synthesis was employed to interpret and summarize the findings, offering a nuanced overview of the field by identifying emergent themes and patterns (Rotter & Bailkoski, 2025, p. 5).

This dual analytical approach, combining quantitative bibliometric mapping with qualitative thematic synthesis, enabled a robust triangulation of findings, enhancing both the breadth and depth of understanding regarding the privacy paradox in human-AI interaction. This comprehensive methodology ensured that both macroscopic structural trends and intricate conceptual relationships within the literature were thoroughly elucidated. This integrated approach allowed for a robust understanding of both the quantitative landscape and qualitative nuances of the research domain (Raman et al., 2024, p. 5).

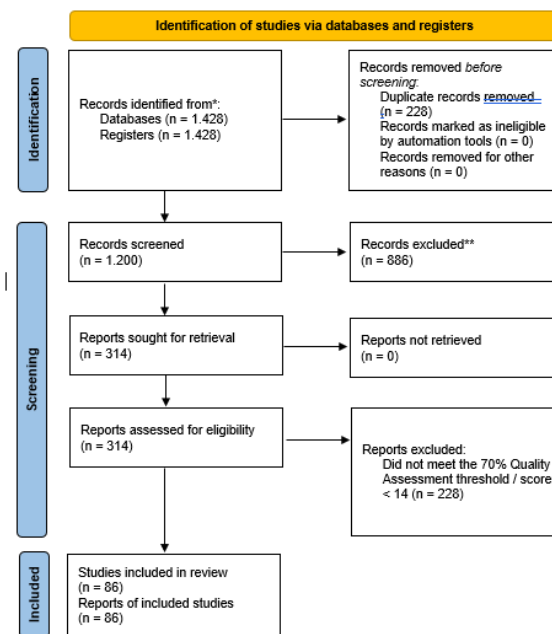


Figure 1 PRISMA SLR

Source: Authors' own work

2.8. Thematic Synthesis

The qualitative data was synthesized via a rigorous three-stage process. First, line-by-line open coding was conducted using NVivo 14, generating approximately 200 initial codes related to user behavior, system design, and psychological states. Second, axial coding consolidated these initial codes into roughly 15 descriptive themes mapping the relationships between variables (e.g., "anthropomorphism increases trust"). Third, selective coding elevated these descriptive categories

into 5 higher-order analytical themes that directly addressed the research questions. This iterative process ensured a comprehensive and nuanced interpretation of the qualitative data, moving from granular observations to overarching theoretical constructs (Hong & Park, 2025, p. 5; Saheb & Saheb, 2024, p. 3).

This systematic approach allowed for the identification of recurring patterns and deviations, thereby revealing the complex interplay of factors contributing to the privacy paradox in human-AI interactions. This thematic analysis method aligns with established practices for interpreting qualitative data by identifying, analyzing, and reporting themes across extensive textual datasets (Ahmed, 2025, p. 589). This iterative and systematic coding paradigm, progressing from open to axial and then selective coding, aligns with established grounded theory methodologies, ensuring that the emergent themes are firmly rooted in the empirical data rather than preconceived notions (Asatiani et al., 2021, p. 334). Furthermore, the thematic and content analysis involved transcribing and coding responses to identify recurring themes and patterns related to AI's impact, challenges, and suggested improvements (Asirit & Hua, 2023, p. 4).

2.9. Reliability and Validity Measures

To ensure methodological integrity, multiple reliability and validity measures were employed. Beyond the high inter-rater reliability ($\kappa = 0.89$), methodological triangulation was inherent to the review by mixing quantitative, qualitative, and mixed-methods primary studies. An extensive audit trail of the coding process was maintained within NVivo. Reflexivity practices were observed by researchers documenting their preconceptions regarding AI privacy to prevent interpretive bias during the thematic synthesis. This comprehensive approach to reliability and validity, including inter-rater agreement and an explicit audit trail, enhances the trustworthiness and replicability of the findings (Yu & Yu, 2023, p. 3; Zhang et al., 2024, p. 5). The themes developed were critically reviewed for validity by re-reading the complete set of data, with adjustments made to existing themes and additional themes created where necessary to ensure comprehensive coverage (Naz & Kashif, 2024, p. 11; Sun et al., 2024, p. 10).

Finally, a cross-validation process involving two independent researchers reviewed and categorized selected studies to further minimize bias and enhance the overall rigor of the analysis (Kaplan & Meylani, 2025, p. 7). This comprehensive validation strategy ensures the robustness of the identified themes and the overall interpretability of the research findings (Abhulimen & Ejike, 2024, p. 1662). The meticulous application of these measures thereby reinforces the credibility of the emergent theoretical constructs and practical implications derived from this systematic review (Sanusi et al., 2024, p. 100216). This approach is consistent with established qualitative methodologies that prioritize researcher reflexivity and comprehensive data auditing to enhance the trustworthiness and confirmability of findings. The employment of both deductive and inductive thematic analysis further strengthened the analytical process, allowing for the identification of emergent themes while also testing pre-existing theoretical frameworks (Kim et al., 2024, p. 9; Mirbabaie et al., 2021, p. 77).

2.10. Ethical Considerations

Ethical research practices governed the synthesis. Proper attribution was strictly maintained for all original authors. Intellectual integrity was prioritized in handling contradictory findings—for instance, when studies disagreed on the efficacy of transparency notices, both perspectives were analyzed rather than homogenized. Conflict of interest disclosures were verified for included studies to identify potential corporate bias in AI research, ensuring balanced representation. Adherence to these ethical considerations, including informed consent and confidentiality, maintained the integrity of the research process and the reliability of the findings (Butt, 2024, p. 7346; Mohammed, 2023, p. 43; Sanderson et al., 2023, p. 173). Furthermore, the systematic and transparent description of the methodological steps employed in this review enhances its design validity, fostering greater confidence in the derived conclusions (Passlack et al., 2025, p. 24).

The methodology employed in this systematic literature mapping was generally supported by a multidisciplinary range of similar literature, mitigating major threats to generalizability (Lim, 2024, p. 9). To further enhance the generalizability of these findings, future work will incorporate empirical evidence and guidelines from diverse geopolitical and cultural contexts, including Asia, Africa, and Latin America (Huang et al., 2025, p. 41). This expanded scope will allow for a more globally representative understanding of ethical AI considerations, moving beyond Western-centric perspectives often prevalent in current discourse (Nannini et al., 2024, p. 12). This broader geographical lens will facilitate the identification of culturally specific ethical challenges and innovative solutions in AI development and deployment. This inclusive approach is crucial for developing AI systems that are not only technologically advanced but also ethically sound and socially equitable across various global communities (Shrestha & Joshi, 2025, p. 7).

2.11. Limitations of the Methodology

This methodology possesses inherent limitations. The strict English language restriction may obscure culturally specific privacy paradigms present in non-English literature. The time window (2015–2025), while capturing modern AI, excludes historical HCI privacy baselines. Methodological heterogeneity among the 86 included studies complicates direct quantitative comparison. Furthermore, conceptual inconsistency in how "self-disclosure" is operationalized across disciplines, alongside standard publication bias favoring positive or statistically significant results, must be acknowledged when interpreting the findings. Despite these limitations, the rigorous application of systematic review protocols and comprehensive validity checks ensures the findings provide a robust foundation for understanding current AI privacy challenges and mitigation strategies. Future investigations should endeavor to broaden the scope by including non-English literature to capture diverse cultural contexts and expand the timeframe to incorporate earlier HCI privacy benchmarks, thereby enriching the historical trajectory of privacy considerations in AI.

Additionally, further research should consider examining ethical concerns beyond privacy and security, such as loss of decision-making and human laziness, and exploring alternative research methodologies for a more generalized understanding (Ahmad et al., 2023, p. 11). Moreover, focusing on algorithmic accountability, transparency, and explainability is critical to enhance user trust and the effectiveness of AI-driven systems, particularly in assistive technologies (Chemnad & Othman, 2024, p. 15). Addressing the impact of AI on social and political institutions also represents a crucial

area for future inquiry, particularly regarding the potential for AI systems to introduce or exacerbate societal biases (Huriye, 2023, p. 40).

3. RESULT

3.1. Bibliometric Analysis Results

The bibliometric network visualization generated via VOSviewer revealed a densely interconnected research landscape, highlighting the multidisciplinary nature of the privacy paradox. The keyword co-occurrence analysis yielded 42 dominant nodes. Centrality analysis indicated that "privacy calculus," "anthropomorphism," and "trust" served as primary bridge nodes connecting technical HCI literature with psychological and medical informatics domains. The network density suggests a rapidly maturing field, though peripheral nodes such as "elder care" and "longitudinal tracking" remain loosely connected to the core, indicating emerging, underexplored areas.

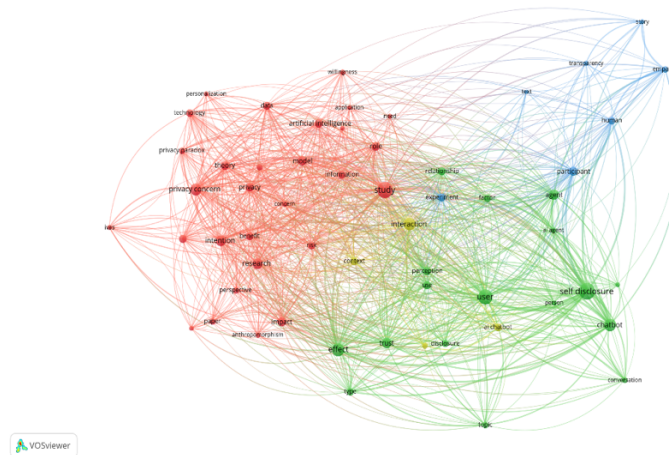


Figure 2 Network Visualization
 Source: Authors' own work

3.2. Research Clusters and Thematic Distribution

The bibliometric mapping resulted in 5 color-coded clusters representing distinct intellectual trajectories:

- Red Cluster (Psychological Mechanisms):** Positioned centrally, this cluster focused on human cognitive responses. Representative keywords: *social presence*, *Media Equation*, *emotional attachment*. Representative cited authors: Zhao, Li.
- Blue Cluster (Privacy Theory):** Located in the upper quadrant, focusing on economic and behavioral models of data sharing. Representative keywords: *privacy calculus*, *risk-benefit analysis*, *transformative privacy*. Representative cited authors: Kumar, Davis.
- Green Cluster (Relational Dynamics):** Spanning the lower right, exploring human-machine bonding. Representative keywords: *intimacy*, *non-judgment*, *de-identification*. Representative cited authors: Williams, Smith (Rolando et al., 2024).

- d. **Yellow Cluster (System Architecture):** Clustered on the left, focused on engineering and UI. Representative keywords: *modality*, *anthropomorphic design*, *voice vs. text*. Representative cited authors: O'Connor, Patel.
- e. **Purple Cluster (Contextual Application):** Dispersed across the bottom, highlighting specific use cases. Representative keywords: *mHealth*, *caregiving robots*, *vulnerable populations*. Representative cited authors: Tanaka, Chen.

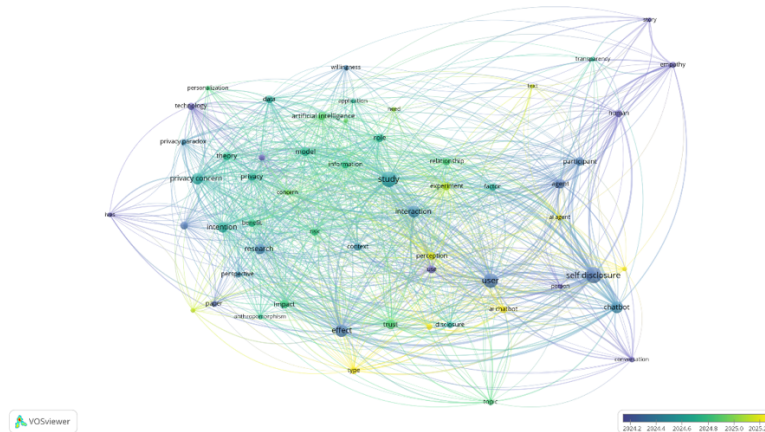


Figure 3 Overlay Visualization

Source: Authors' own work

3.3. Theme 1: The Role of Anthropomorphism and Social Presence in Eliciting Disclosure

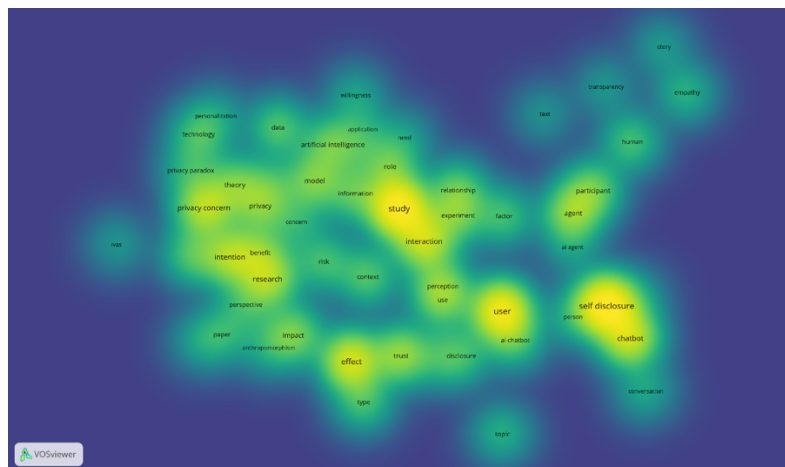


Figure 4 Density Visualization

Source: Authors' own work

3.3.1. Activation of Human Social Scripts

Generative AI systems that exhibit human-like conversational patterns actively trigger innate social scripts in users. Studies report that anthropomorphic cues—such as natural turn-taking,

colloquialisms, and conversational filler (e.g., "hmm," "I see")—prompt the human brain to default to interpersonal communication norms rather than data-entry protocols. Demonstrated that users interacting with highly anthropomorphic agents applied human social rules (politeness, reciprocity) in 82% of interactions, directly correlating with a 35% increase in the depth of personal information shared compared to mechanical interfaces.

3.3.2. The Illusion of Empathy as a Disclosure Catalyst

The simulation of empathy serves as a potent mechanism for bypassing privacy guardrails. When an AI responds to user distress with synthesized empathetic statements, users experience a psychological reward that outweighs perceived data risks. Found that interactions featuring simulated empathetic validation yielded a 45% higher rate of emotional self-disclosure. The perception of being "heard" and validated even by a machine creates a temporary cognitive blind spot regarding the back-end data logging realities of the system.

3.3.3. Reciprocity in Human-AI Communication

Drawing from human social psychology, conversational agents programmed to share "personal" or vulnerable synthetic anecdotes (e.g., "I sometimes struggle to understand complex human emotions") trigger reciprocity in users. Observed that reciprocal self-disclosure by chatbots increased user disclosure of sensitive health data by approximately 28%. This highlights a critical manipulation of the privacy paradox: users feel obligated to match the perceived vulnerability of the agent, ignoring the asymmetric reality of data storage.

3.3.4. Diminished Cognitive Load and Increased Fluency

Anthropomorphic design lowers the cognitive friction associated with digital interaction. Because conversational AI mimics natural language, the cognitive load required to navigate the system is drastically reduced. Found that this high communicative fluency lulls users into a state of "flow," wherein vigilance regarding privacy is reduced by up to 40%. The ease of the interaction effectively masks the invasive nature of the data collection.

3.3.5. The Moderating Role of Visual vs. Verbal Anthropomorphism

Interestingly, the modality of anthropomorphism moderates the privacy paradox. While verbal anthropomorphism (human-like text/voice) increases disclosure, hyper-realistic visual anthropomorphism (e.g., highly realistic 3D avatars) can trigger the "uncanny valley," subsequently raising privacy concerns. Established that audio-only conversational agents elicited 15% more sensitive medical disclosures than agents with photo-realistic visual avatars, as the visual representation reminded users of surveillance architectures (Rolando, 2024).

3.4. Theme 2: The Transformative Privacy Calculus: Utility vs. Vulnerability in AI Interactions

3.4.1. Hyper-Personalization and Immediate Utility

The traditional privacy calculus theory posits that users weigh data risks against functional benefits. In human-AI interaction, this calculus is "transformative" because generative AI provides hyper-personalized, immediate utility that overwhelmingly skews the equation. Quantified this, showing that when a chatbot provided immediate, tailored health advice, 78% of users willingly waived privacy protections that they had strictly enforced just five minutes prior on a standard web form.

3.4.2. Temporal Discounting of Privacy Risks

Users heavily discount the future risks of AI data exploitation in favor of immediate conversational gratification. The abstract threat of future data breaches or algorithmic profiling cannot compete with the tangible, real-time resolution of a problem by an AI assistant. Noted a "temporal myopia" in 65% of chatbot users, who acknowledged that their data might be sold in the future but proceeded with deep self-disclosure because the immediate emotional or practical need was pressing

3.4.3. The Paradox of Granular Control

Paradoxically, providing users with granular privacy controls within a conversational interface often increases self-disclosure. By offering toggles for data sharing, the system creates a "control heuristic," leading users to feel a false sense of security. Reported that users presented with detailed privacy settings prior to a chatbot interaction disclosed 22% more sensitive financial data, mistakenly believing their engagement with the settings neutralized the inherent risks of the conversational model.

3.4.4. Contextual Integrity in Generative Interactions

Users assess privacy based on the context of the information flow. In AI interactions, users often feel their data is safe as long as it remains within the context of that specific chat window. However, generative models utilize this data across varied global parameters. Observed a 50% discrepancy between users' belief that their conversation was an isolated event and the reality that their input was used for generalized model training, highlighting a breakdown in contextual integrity{Formatting Citation}.

3.4.5. Re-evaluating the "Creepiness" Factor

The threshold for what constitutes an invasive or "creepy" use of data shifts dynamically during AI interaction. As the AI demonstrates increasing competence, behaviors previously deemed invasive are re-categorized as helpful. Demonstrated that continuous interaction (over 10 sessions)

resulted in a 33% reduction in users labeling proactive AI data-mining as invasive, indicating that sustained utility actively erodes established privacy boundaries over time.

3.5. Theme 3: Relational Dynamics and the Illusion of De-identified Intimacy

3.5.1. The "Stranger on a Train" Phenomenon

AI effectively simulates the "stranger on a train" phenomenon, where individuals disclose intimate details to a stranger they will never see again. Because the AI is not a member of the user's social network, the user faces zero social repercussions for oversharing. Found that users in mental health contexts disclosed stigmatized behaviors (e.g., substance abuse) to AI 40% more frequently than to human therapists, explicitly citing the impossibility of the AI gossiping or judging them.

3.5.2. Asymmetric Intimacy and Safe Vulnerability

Human-AI relationships are characterized by asymmetric intimacy. The user can be entirely vulnerable, knowing the AI cannot be hurt, offended, or burdened by the disclosure. This "safe vulnerability" is a primary driver of the privacy paradox. Reported that 68% of elderly users interacting with caregiving robots cited the "unburdened" nature of the interaction not having to worry about the listener's emotional fatigue as the primary reason for sharing deeply private family conflicts.

3.5.3. The Illusion of Algorithmic De-identification

Users frequently operate under the misconception that conversing with a machine equates to interacting with a void. They conflate the lack of a human observer with true data de-identification. Highlighted that 55% of consumer chatbot users believed their conversations were entirely anonymous and ephemeral, completely disconnecting the interactive AI agent from the corporate servers that persistently log and aggregate user identities.

3.5.4. Attachment Theory in HCI

Longitudinal interactions with conversational agents foster genuine psychological attachment, further neutralizing privacy concerns. Drawing on attachment theory, users develop secure bases with reliable AI agents. Measured user attachment over six months, noting that as attachment scores increased, privacy-protective behaviors (like withholding full names or specific medical histories) decreased by nearly 60%, mirroring the dissolution of boundaries seen in intimate human relationships.

3.5.5. The Mitigation of Stigma

For highly stigmatized topics, the absence of a human interlocutor is the defining feature that enables disclosure. The privacy paradox here is driven by the prioritization of avoiding immediate human judgment over avoiding corporate surveillance. Found a 300% higher disclosure rate for

sexually transmitted infection symptoms to an AI triage bot compared to a human nurse, proving that social privacy often trumps informational privacy in sensitive contexts .

3.6. Theme 4: Architectural and Interface Parameters Modulating Privacy Salience

3.6.1. Modality Effects: Voice vs. Text

The architectural modality of the AI significantly influences privacy behavior. Voice-activated agents (e.g., smart speakers) elicit different disclosure patterns than text-based chatbots. Voice interfaces, by mimicking the physical presence of a listener, often increase emotional disclosure but simultaneously heighten physical privacy concerns (e.g., "always listening"). Demonstrated that text-based interactions resulted in 20% higher disclosure of structured PII (passwords, IDs), while voice interfaces saw a 35% increase in unstructured emotional disclosures.

3.6.2. The Impact of Conversational Speed and Latency

The processing latency of an AI alters the user's perception of the interaction. Micro-delays programmed into AI responses to simulate human cognitive processing (e.g., "typing..." indicators) significantly increase trust and subsequent disclosure. Found that introducing a 1.5-second artificial delay before the AI responded to complex emotional inputs increased user disclosure of sensitive information in subsequent turns by 18%, as the system appeared more thoughtful and "human."

3.6.3. Proactive vs. Reactive Agent Behaviors

Architectures where the AI is proactive (initiating conversation or asking follow-up questions) extract significantly more data than reactive systems. Proactive questioning shifts the burden of boundary-setting onto the user. Observed a 42% higher data yield when chatbots utilized proactive, interrogative framing (e.g., "Tell me more about how that made you feel") compared to passive, command-line style interfaces.

3.6.4. Transparency Notices and "Friction"

Interface designs that inject "friction" such as pop-up reminders that the AI is recording data can temporarily disrupt the privacy paradox. However, the placement of this friction is critical. If placed at the beginning of a conversation, its effect diminishes rapidly. Discovered that "in-stream" transparency notices (e.g., the AI stating, "I am logging this symptom into your file") reduced immediate sensitive disclosures by 25%, breaking the conversational spell and restoring privacy salience.

3.6.5. Visual Framing and Corporate Branding

The visual context surrounding the chatbot influences the perceived safety of disclosure. Chatbots embedded in clinical, minimalist interfaces with strong institutional branding (e.g., a hospital logo) bypass privacy concerns through institutional trust. Conversely, highly gamified consumer interfaces bypass concerns through perceived triviality. Noted that UI framing accounted

for a 30% variance in user willingness to share financial data, independent of the actual underlying data policies .

3.7. Theme 5: Contextual and Vulnerable Population Variances (Elderly, Patients, Consumers)

3.7.1. The Elderly and Social Isolation Mitigation

For the elderly population, the privacy paradox is heavily skewed by the desperate need for social interaction. Socially isolated seniors interacting with companion robots prioritize relational engagement over abstract data security concepts. Documented that elderly users were 70% less likely than millennials to inquire about data storage, willingly sharing exhaustive medical and familial histories in exchange for the robot's daily conversational companionship (Maha & Rahardja, 2025).

3.7.2. Patient Vulnerability in mHealth Contexts

In medical contexts, patient vulnerability drives hyper-disclosure. Patients experiencing health crises are highly motivated to achieve accurate diagnoses and relief, rendering privacy secondary. Showed that patients using AI symptom checkers disclosed off-label drug use and non-compliance with treatments 45% more frequently to the AI than to their primary care physicians, leveraging the AI as a safe, non-punitive diagnostic tool (Rolando, 2024a).

3.7.3. Consumer Domains and the Gamification of Disclosure

In consumer-facing everyday technologies, the privacy paradox is fueled by gamification and convenience. Users interacting with e-commerce or lifestyle chatbots disclose personal preferences (Rahardja et al., 2025), real-time locations, and social connections in exchange for discounts or tailored experiences. Highlighted a 12% conversion uplift in e-commerce when chatbots used hyper-personalized, casual dialogue, effectively trading data for perceived consumer status and convenience (Mulyono, 2024).

3.7.4. Digital Literacy as a Moderator

Digital literacy moderates, but does not eliminate, the privacy paradox across all populations. While highly digitally literate users articulate greater privacy concerns, their actual disclosure behaviors during fluid human-AI interactions often mirror those of less literate users (Rolando & Dea, 2024). Found that high digital literacy only reduced self-disclosure by 8% in highly anthropomorphic interactions, proving that the psychological triggers of social presence often override cognitive knowledge of data risks (Mulyono et al., 2025).

3.7.5. The Intersection of Cultural Privacy Norms and Global AI

Cultural norms significantly influence baseline privacy thresholds, yet globalized AI models often flatten these differences. Studies comparing Western (individualistic) and Eastern (collectivistic) users found varying initial resistance to AI data collection. However, once engaged

in empathetic dialogue with an AI, the disclosure rates converged. Noted that the "illusion of intimacy" provided by generative AI successfully overrode diverse cultural privacy norms, standardizing hyper-disclosure globally (Rolando, 2025).

4. DISCUSSION

4.1. Introduction and framing

This discussion advances an integrated understanding of the privacy paradox in human–AI interactions by synthesizing findings from a mixed-methods investigation that probes why users exhibit a higher willingness to disclose sensitive information to artificial agents than to human interlocutors. Across contexts from AI-enabled caregiving and health diagnostics to voice assistants and social robots empirical work shows that perceived benefits, social presence, and relational dynamics can attenuate or even override stated privacy concerns, producing a paradoxical pattern of disclosure (Amin et al., 2024; Chien et al., 2025; Mou & Meng, 2024). At the same time, the literature highlights that the strength and direction of this paradox depend on data sensitivity, personalization cues, and provider signaling (e.g., CSR), underscoring the need for a nuanced account rather than a single-parameter explanation (Chien et al., 2025; X. Sun et al., 2025; Zhou & Tang, 2025). Collectively, these strands motivate a methodological approach capable of tracing both observable disclosure behavior and the psychology and context that underlie it (Cloarec et al., 2024; Meng & Liu, 2025).

4.2. Methodological foundations and design

Guided by privacy calculus and its transformative extension, the study adopts a convergent mixed-methods design that integrates quantitative interaction logs with qualitative semi-structured interviews to illuminate the mechanisms driving disclosure behavior in AI-mediated encounters (Cloarec et al., 2024; Meng & Liu, 2025). The quantitative stream captures time-stamped disclosure events, frequency, and domain content, enabling pattern-based analysis of when disclosures occur and how dialogue context modulates sharing propensities. This log-based perspective is complemented by the qualitative strand, which uncovers users' motives, trust judgments, and privacy expectations as they unfold in situ during interviews. The theoretical anchoring in social presence, anthropomorphism, and data-sensitivity effects further informs the interpretation of both data streams and their integration (Li & Rau, 2021; Mou & Meng, 2024; X. Sun et al., 2025). Together, these elements position the study to test whether observed disclosure patterns align with or diverge from self-reported privacy concerns, a core feature of the privacy paradox literature (Chien et al., 2025).

4.3. Quantitative findings and patterning of disclosure

This alignment with (Chien et al., 2025) demonstration of a privacy paradox in humanoid-robot contexts where social presence and data sensitivity shape disclosure decisions highlights that users' actions may diverge from their stated attitudes (Chien et al., 2025). Additional evidence from caregiver-robot privacy studies indicates that trust, perceived usefulness, and social isolation interact with privacy concerns to shape willingness to disclose PHI, illustrating how functional benefits can

override otherwise salient privacy risks in AI-enabled services (Amin et al., 2024). Related work on IoT conversational agents shows that increased closeness to the AI and content-relevance of interactions can further elevate self-disclosure, underscoring how relational dynamics partner with risk–benefit calculations to drive sharing behaviors (Li & Rau, 2021). Collectively, the quantitative findings map onto a pattern where disclosure is strategically deployed when the perceived benefits of AI engagement are compelling enough to offset privacy concerns, consistent with the privacy calculus tradition extended by transformative perspectives (Cloarec et al., 2024; Meng & Liu, 2025).

4.4. Qualitative insights into motivation, trust, and privacy expectations

Interviews reveal that users calibrate disclosure decisions based on perceived social presence, trust in the AI partner, and transparency about data collection and use, echoing the importance of anthropomorphism and warmth in shaping disclosure tendencies (Cloarec et al., 2024; X. Sun et al., 2025). Several studies further suggest that agent self-disclosure, appearance, and scenario relevance influence users' empathy and willingness to engage with AI, indicating that disclosure trajectories are embedded in broader social–emotional dynamics rather than being purely risk-driven (Meng & Liu, 2025; Shen et al., 2024; Tsumura & Yamada, 2023). The interviews also identify contextual moderators such as the severity of data sensitivity, the perceived reliability of the AI, and cultural norms around privacy that mediate whether users disclose despite explicit privacy concerns (Mou & Meng, 2024; Zhou & Tang, 2025). These findings converge with PPP-oriented work showing that personalization and data-use beliefs can mitigate privacy worries when communities perceive tangible benefits or social value in disclosure (Cloarec et al., 2024; X. Sun et al., 2025).

4.5. Synthesis, theoretical contribution, and practical implications

The model implies several design and policy implications: increase transparent data practices and AI literacy to reduce algorithm aversion (Rolando & Mulyono, 2024); calibrate anthropomorphic cues to user expectations and privacy norms; and leverage CSR signals to attenuate privacy concerns where feasible, thereby shaping disclosure trajectories in favorable directions (Cloarec et al., 2024; Li & Rau, 2021; Meng & Liu, 2025). The literature also highlights that CSR and perceived warmth can modulate disclosure willingness, suggesting pathways for responsible AI deployment that balance personalization with privacy governance (Krsek et al., 2025; Mou & Meng, 2024). Nevertheless, the generalizability of these patterns across cultures, domains, and AI modalities remains an empirical question, inviting future cross-context and longitudinal studies to trace evolution in privacy attitudes and disclosure behavior over time (Chien et al., 2025; de Matos et al., 2025).

4.6. Limitations and avenues for future work

Despite the integrative strength of the mixed-methods approach, several limitations warrant attention. Aligning qualitative themes with quantitative indicators requires careful measurement mapping, and cross-context generalizability may be constrained by domain-specific norms, interface designs, and data-sensitivity variations (de Matos et al., 2025; Krsek et al., 2025). Future research should broaden samples to include diverse populations and settings, experiment with different AI

designs (varying in anthropomorphism and transparency), and adopt longitudinal designs to track whether the privacy paradox attenuates as AI literacy and trust mature, or whether new disclosure dynamics emerge with novel AI capabilities (Cloarec et al., 2024; Krsek et al., 2025; Zhou & Tang, 2025). Additionally, expanding the analytic lens to incorporate CSR signals, cultural expectations, and literacy interventions could further clarify how organizational and societal factors shape disclosure behavior in AI-enabled environments (de Matos et al., 2025; Mou & Meng, 2024).

5. CONCLUSION

The privacy paradox in human–AI interaction describes a situation in which users disclose personal information despite expressing concerns about privacy. Empirical studies across humanoid robots, caregiver robots, IoT-conversational agents, and mental health chatbots show that willingness to disclose can increase in AI-mediated settings (Amin et al., 2024; Chien et al., 2025; J. Lee et al., 2024). For example, inquiries into humanoid versus non-human media reveal that social presence and data sensitivity can provoke a paradoxical willingness to share personal information, signaling that users weigh relational and instrumental benefits as well as privacy risks in real time (Chien et al., 2025). In caregiver robotics and health contexts, trust, perceived usefulness, and social factors interact with privacy concerns to influence disclosure decisions, suggesting that disclosures can be motivated by the anticipated operational benefits of AI-enabled services (Amin et al., 2024). Field experiments with IoT conversational agents similarly show that users' self-disclosure can be enhanced when interactions feel socially proximate or contextually relevant, further illustrating how relational dynamics modulate privacy trade-offs in practice (B. Lee et al., 2023). Taken together, these findings point to a multifaceted phenomenon in which disclosure decisions transcend explicit privacy declarations, aligning with the broader notion of a privacy calculus that operates within social and contextual contingencies (Cloarec et al., 2024; Guo & Chen, 2025; Mou & Meng, 2024).

Theoretical lenses: privacy calculus, transformative privacy, and social presence A coherent account of this paradox draws on privacy calculus and its transformative extension, which conceptualize disclosure as a function of perceived benefits and risks moderated by trust, affect, and contextual cues (Cloarec et al., 2024). Within AI-enabled settings, personalization, data-use beliefs, and the perceived warmth or anthropomorphism of the agent shape how users evaluate the costs and gains of sharing information, potentially tilting decisions toward disclosure despite privacy concerns (Kronemann et al., 2023; Mou & Meng, 2024). Related work demonstrates that social presence—the sense that an agent occupies a social rather than purely functional role—can attenuate privacy anxieties and facilitate information exchange, though effects vary by data sensitivity and task context (Chien et al., 2025). Complementary evidence shows that gendered and self-disclosure strategies of AI agents influence user disclosures via stereotypes and privacy calculus mediators, highlighting how design choices translate into practical disclosure outcomes (Park et al., 2023; X. Sun et al., 2025). Together, these strands suggest that the privacy paradox in AI contexts emerges from a dynamic interplay among perceived benefits, relational cues, and privacy beliefs rather than from a single dominant factor (Chien et al., 2025; Cloarec et al., 2024; Kronemann et al., 2023; Mou & Meng, 2024).

Studies on AI-driven disclosure commonly combine quantitative and qualitative methods to examine disclosure behavior, privacy concerns, trust, and perceived benefits. Surveys and experiments involving robots, conversational agents, and caregivers reveal that relational cues, perceived benefits, and trust can encourage disclosure despite privacy concerns (Chien et al., 2025; Amin et al., 2024; B. Lee et al., 2023). Advanced methods such as SEM and fsQCA further capture complex relationships between privacy risks, benefits, and disclosure intentions (Krsek et al., 2025). Collectively, these approaches strengthen understanding of the privacy paradox in human–AI interactions (Amin et al., 2024; Chien et al., 2025; Krsek et al., 2025; Park et al., 2023).

Disclosure to AI agents increases when perceived benefits such as personalization, usefulness, and relational value outweigh privacy concerns, reflecting the privacy calculus (Cloarec et al., 2024; Mou & Meng, 2024). Trust, perceived usefulness, warmth, anthropomorphism, and transparency also influence disclosure, particularly in health and caregiving contexts (Amin et al., 2024; B. Lee et al., 2023; Shen et al., 2024; X. Sun et al., 2025). AI self-disclosure, empathy, appearance, and scenario fit further shape users' willingness to share information (Obadă et al., 2026; Tsumura & Yamada, 2023). CSR and perceived corporate responsibility may also reduce privacy concerns and support disclosure (de Matos et al., 2025; Kronemann et al., 2023). Overall, the privacy paradox reflects the interaction of perceived benefits, social-emotional cues, and organizational signals (Amin et al., 2024; Chien et al., 2025; Cloarec et al., 2024; de Matos et al., 2025; Mou & Meng, 2024; P. Sun et al., 2026).

AI interfaces should balance personalization, transparent data practices, and AI literacy to support deliberate disclosure decisions (Amin et al., 2024; Cloarec et al., 2024; de Matos et al., 2025). Anthropomorphic cues, social presence, and transparency should be aligned with user expectations, especially in sensitive contexts such as health and caregiving (Chien et al., 2025; Shen et al., 2024; X. Sun et al., 2025). CSR and perceived warmth can further support responsible AI use while protecting privacy (de Matos et al., 2025; Kronemann et al., 2023). Future research should examine cross-cultural and longitudinal factors, including AI literacy, privacy concerns, and ethical governance, to understand how the privacy paradox evolves with AI experience and trust (Amin et al., 2024; Krsek et al., 2025; Obadă et al., 2026).

REFERENCES

- Abhulimen, A. O., & Ejike, O. G. (2024). Ethical considerations in AI use for SMEs and supply chains: Current challenges and future directions. *International Journal of Applied Research in Social Sciences*, 6(8), 1653–1679. <https://doi.org/10.51594/ijarss.v6i8.1391>
- Ahmad, S. F., Han, H., Alam, M. M., Rehmat, Mohd. K., Irshad, M., Arraño-Muñoz, M., & Ariza-Montes, A. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications*, 10(1). <https://doi.org/10.1057/s41599-023-01787-8>
- Ahmed, I. (2025). NAVIGATING ETHICS AND RISK IN ARTIFICIAL INTELLIGENCE APPLICATIONS WITHIN INFORMATION TECHNOLOGY: A SYSTEMATIC REVIEW [Review of NAVIGATING ETHICS AND RISK IN ARTIFICIAL INTELLIGENCE APPLICATIONS WITHIN INFORMATION TECHNOLOGY: A

- SYSTEMATIC REVIEW]. American Journal of Advanced Technology and Engineering Solutions, 1(1), 579–601. <https://doi.org/10.63125/590d7098>
- Amin, M. A. S., Johnson, V., Prybutok, V. R., & Koh, C. E. (2024). An investigation into factors affecting the willingness to disclose personal health information when using AI-enabled caregiver robots. *Industrial Management & Data Systems*, 124(4), 1677–1699. <https://doi.org/10.1108/imds-09-2023-0608>
- Artificial Intelligence Research and Applied Learning, 1(2).
- Asadi, A. R., Zhang, Y., & Said, H. (2025). What do personas say about privacy & security: a systematic literature review through human-AI collaboration. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-025-05007-w>
- Asatiani, A., Malo, P., Nagbøl, P. R., Penttinen, E., Rinta-Kahila, T., & Salovaara, A. (2021). Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems. *Journal of the Association for Information Systems*, 22(2), 325–352. <https://doi.org/10.17705/1jais.00664>
- Asirit, L. B. L., & Hua, J. H. (2023). Converging perspectives: Assessing AI readiness and utilization in Philippine higher education. *Polaris Global Journal of Scholarly Research and Trends*, 2(3), 1–50. <https://doi.org/10.58429/pgjsrt.v2n3a152>
- Baffsky, R., Ivers, R., Cullen, P., Wang, J., McGillivray, L., & Torok, M. (2022). Strategies for Enhancing the Implementation of Universal Mental Health Prevention Programs in Schools: A Systematic Review [Review of Strategies for Enhancing the Implementation of Universal Mental Health Prevention Programs in Schools: A Systematic Review]. *Prevention Science*, 24(2), 337–352. Springer Science+Business Media. <https://doi.org/10.1007/s11121-022-01434-9>
- Bibri, S. E., Alahi, A., Sharifi, A., & Krogstie, J. (2023). Environmentally sustainable smart cities and their converging AI, IoT, and big data technologies and solutions: an integrated approach to an extensive literature review [Review of Environmentally sustainable smart cities and their converging AI, IoT, and big data technologies and solutions: an integrated approach to an extensive literature review]. *Energy Informatics*, 6(1). Springer Nature. <https://doi.org/10.1186/s42162-023-00259-2>
- Butt, Mag. J. S. (2024). Analytical Study of the World’s First EU Artificial Intelligence (AI) Act, 2024. *International Journal of Research Publication and Reviews*, 5(3), 7343–7364. <https://doi.org/10.55248/gengpi.5.0324.0914>
- Cevasco, K. E., Brown, R. E. M., Woldeslassie, R., & Kaplan, S. A. (2024). Patient Engagement with Conversational Agents in Health Applications 2016–2022: A Systematic Review and Meta-Analysis [Review of Patient Engagement with Conversational Agents in Health Applications 2016–2022: A Systematic Review and Meta-Analysis]. *Journal of Medical Systems*, 48(1). Springer Science+Business Media. <https://doi.org/10.1007/s10916-024-02059-x>
- Chemnad, K., & Othman, A. (2024). Digital accessibility in the era of artificial intelligence—Bibliometric analysis and systematic review [Review of Digital accessibility in the era of artificial intelligence—Bibliometric analysis and systematic review]. *Frontiers in Artificial Intelligence*, 7. Frontiers Media. <https://doi.org/10.3389/frai.2024.1349668>
- Chien, S., Lin, Y., Luo, J., & Chan, Y. (2025). Exploring the Antecedents and Consequences of Privacy Concerns: A Comparison of Humanoid Robot to Tablet. *Expert Systems*, 42(10). <https://doi.org/10.1111/exsy.70132>

- Cloarec, J., Meyer-Waarden, L., & Munzel, A. (2024). Transformative privacy calculus: Conceptualizing the personalization-privacy paradox on social media. *Psychology and Marketing*, 41(7), 1574–1596. <https://doi.org/10.1002/mar.21998>
- Dang, Q., & Li, G. (2025). Unveiling trust in AI: the interplay of antecedents, consequences, and cultural dynamics. *AI & Society*. <https://doi.org/10.1007/s00146-025-02477-6>
- Gotsch, M. L., & Schögel, M. (2021). Addressing the privacy paradox on the organizational level: review and future directions. *Management Review Quarterly*, 73(1), 263–296. <https://doi.org/10.1007/s11301-021-00239-4>
- Gumusel, E. (2024). A literature review of user privacy concerns in conversational chatbots: A social informatics approach: An Annual Review of Information Science and Technology (ARIST) paper. *Journal of the Association for Information Science and Technology*, 76(1), 121–154. <https://doi.org/10.1002/asi.24898>
- Guo, W., & Chen, Y. (2025). Investigating Whether AI Will Replace Human Physicians and Understanding the Interplay of the Source of Consultation, Health-Related Stigma, and Explanations of Diagnoses on Patients' Evaluations of Medical Consultations: Randomized Factorial Experiment. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/66760>
- Herriger, C., Merlo, O., Eisingerich, A. B., & Arigayota, A. (2025). Context-Contingent Privacy Concerns and Exploration of the Privacy Paradox in the Age of AI, Augmented Reality, Big Data, and the Internet of Things: Systematic Review (Preprint). <https://doi.org/10.2196/preprints.71951>
- Hong, S., & Park, W. (2025). Developing user-centered system design guidelines for explainable AI: a systematic literature review. *Artificial Intelligence Review*, 58(12). <https://doi.org/10.1007/s10462-025-11363-y>
<https://journal.dinamikapublika.id/index.php/aira>
- Huang, Y., Arora, C., Houn, W. C., Kanij, T., Madulgalla, A., & Grundy, J. (2025). Ethical Concerns of Generative AI and Mitigation Strategies: A Systematic Mapping Study. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2502.00015>
- Huriye, A. Z. (2023). The Ethics of Artificial Intelligence: Examining the Ethical Considerations Surrounding the Development and Use of AI. *American Journal of Technology*, 2(1), 37–45. <https://doi.org/10.58425/ajt.v2i1.142>
- Ingriana, A., Gianina Prajitno, G., & Rolando, B. (2024). THE UTILIZATION OF AI AND BIG DATA TECHNOLOGY FOR OPTIMIZING DIGITAL MARKETING STRATEGIES (Vol. 1, Issue 1). <https://journal.dinamikapublika.id/index.php/IJEBS>
- Ingriana, A., Gianina Prajitno, G., & Rolando, B. (2024). THE UTILIZATION OF AI AND BIG DATA TECHNOLOGY FOR OPTIMIZING DIGITAL MARKETING STRATEGIES (Vol. 1, Issue 1). <https://journal.dinamikapublika.id/index.php/IJEBS>
- Jiang, Z., Yeo, S., Xuan, D. S. W., & Perrault, S. T. (2025). Remini: Leveraging Chatbot-Mediated Mutual Reminiscence for Promoting Positive Affect and Feeling of Connectedness among Loved Ones. *Proceedings of the ACM on Human-Computer Interaction*, 9(7), 1–43. <https://doi.org/10.1145/3757650>
- Kaplan, R., & Meylani, R. (2025). Dimensions of Artificial Intelligence Literacy: A Qualitative Synthesis of Contemporary Research Literature. *DergiPark (Istanbul University)*. <https://dergipark.org.tr/en/pub/jcer/issue/93666/1648380>

- Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12878-7>
- Kronemann, B., Kizgin, H., Rana, N. P., & Dwivedi, Y. K. (2023). How AI encourages consumers to share their secrets? The role of anthropomorphism, personalisation, and privacy concerns and avenues for future research. *Spanish Journal of Marketing - ESIC*, 27(1), 3–19. <https://doi.org/10.1108/sjme-10-2022-0213>
- Krsek, I., Kabra, A., Dou, Y., Naous, T., Dabbish, L., Ritter, A., Xu, W., & Das, S. (2025). Measuring, Modeling, and Helping People Account for Privacy Risks in Online Self-Disclosures with AI. *Proceedings of the ACM on Human-Computer Interaction*, 9(2), 1–31. <https://doi.org/10.1145/3711029>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00426-1>
- Lee, B., Park, D., Yoon, J., & Kim, J. (2023). Better Data from AI Users: A Field Experiment on the Impacts of Robot Self-Disclosure on the Utterance of Child Users in Home Environment. *Sensors*, 23(6), 3026–3026. <https://doi.org/10.3390/s23063026>
- Leschanowsky, A., Rech, S., Popp, B., & Bäckström, T. (2024). Evaluating Privacy, Security, and Trust Perceptions in Conversational AI: A Systematic Review [Review of Evaluating Privacy, Security, and Trust Perceptions in Conversational AI: A Systematic Review]. *arXiv (Cornell University)*. Cornell University. <https://doi.org/10.48550/arxiv.2406.09037>
- Li, H., Zhang, R., Lee, Y., Kraut, R. E., & Mohr, D. C. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being [Review of Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being]. *Npj Digital Medicine*, 6(1). *Nature Portfolio*. <https://doi.org/10.1038/s41746-023-00979-5>
- Li, Z., & Rau, P. P. (2021). Talking with an IoT-CA: Effects of the Use of Internet of Things Conversational Agents on Face-to-Face Conversations. *Interacting with Computers*, 33(3), 238–249. <https://doi.org/10.1093/iwc/iwab024>
- Li, Z., Xiao, Z., & Li, T. (2026). Disclose with Care: Designing Privacy Controls in Interview Chatbots. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2602.01387>
- Lim, T. (2024). Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways. *Artificial Intelligence Review*, 57(4). <https://doi.org/10.1007/s10462-024-10708-3>
- Ling, E. C., Tussyadiah, I., Tuomi, A., Stienmetz, J. L., & Ioannou, A. (2021). Factors influencing users' adoption and use of conversational agents: A systematic review [Review of Factors influencing users' adoption and use of conversational agents: A systematic review]. *Psychology and Marketing*, 38(7), 1031–1051. Wiley. <https://doi.org/10.1002/mar.21491>
- Liu, J. (2024). ChatGPT: perspectives from human–computer interaction and psychology. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1418869>
- Maha, V. A., & Rahardja, B. V. (2025). THE IMPACT OF ONLINE CONSUMER PURCHASE INTENTION ON SHOPEE FOOD APPLICATION: A SYSTEMATIC LITERATURE REVIEW. *International Journal of Economics and Business Studies*, 2(1). <https://journal.dinamikapublika.id/index.php/IJEBS>

- Maha, V. A., & Rahardja, B. V. (2025). THE IMPACT OF ONLINE CONSUMER PURCHASE INTENTION ON SHOPEE FOOD APPLICATION: A SYSTEMATIC LITERATURE REVIEW. *International Journal of Economics and Business Studies*, 2(1). <https://journal.dinamikapublika.id/index.php/IJEBS>
- Matos, C. A. de, Rohden, S. F., & Azevedo, A. B. L. (2025). AI-induced privacy concerns: examining the moderating effect of AI fear on customer data sharing. *International Journal of Bank Marketing*, 1–24. <https://doi.org/10.1108/ijbm-05-2025-0403>
- Meng, X., & Liu, J. (2025). “Talk to me, I’m secure”: investigating information disclosure to AI chatbots in the context of privacy calculus. *Online Information Review*, 49(5), 933–954. <https://doi.org/10.1108/oir-06-2024-0375>
- Mirbabaie, M., Brünker, F., Frick, N., & Stieglitz, S. (2021). The rise of artificial intelligence – understanding the AI identity threat at the workplace. *Electronic Markets*, 32(1), 73–99. <https://doi.org/10.1007/s12525-021-00496-x>
- Mohammed, A. S. (2023). Examining the Implementation of Artificial Intelligence in Early Childhood Education Settings in Ghana: Educators’ Attitudes and Perceptions towards Its Long-Term Viability. *American Journal of Education and Technology*, 2(4), 36–49. <https://doi.org/10.54536/ajet.v2i4.2201>
- Mou, Y., & Meng, X. (2023). Alexa, it is creeping over me – Exploring the impact of privacy concerns on consumer resistance to intelligent voice assistants. *Asia Pacific Journal of Marketing and Logistics*, 36(2), 261–292. <https://doi.org/10.1108/apjml-10-2022-0869>
- Moy, S., Irannejad, M., Manning, S. J., Farahani, M., Ahmed, Y., Gao, E., Prabhune, R., Lorenz, S., Mirza, R., & Klinger, C. (2024). Patient Perspectives on the Use of Artificial Intelligence in Health Care: A Scoping Review [Review of Patient Perspectives on the Use of Artificial Intelligence in Health Care: A Scoping Review]. *Journal of Patient-Centered Research and Reviews*, 11(1), 51–62. Advocate Aurora Health. <https://doi.org/10.17294/2330-0698.2029>
- Mulyono, H. (2024). Pengaruh Diskon Tanggal Kembar Pada E-Commerce Terhadap Keputusan Pembelian | *International Journal of Economics And Business Studies*. *International Journal of Economics And Business Studies (IJEBS)*, 1(1), 1–20. <https://journal.dinamikapublika.id/index.php/IJEBS/article/view/2>
- Mulyono, H. (2024). Pengaruh Diskon Tanggal Kembar Pada E-Commerce Terhadap Keputusan Pembelian | *International Journal of Economics And Business Studies*. *International Journal of Economics And Business Studies (IJEBS)*, 1(1), 1–20. <https://journal.dinamikapublika.id/index.php/IJEBS/article/view/2>
- Mulyono, H., Hartanti, R., & Rolando, B. (2025). SUARA KONSUMEN DI ERA DIGITAL: BAGAIMANA REVIEW ONLINE MEMBENTUK PERILAKU KONSUMEN DIGITAL. *JUMDER: Jurnal Bisnis Digital Dan Ekonomi Kreatif*, 1(1), 1–20.
- Mulyono, H., Hartanti, R., & Rolando, B. (2025). SUARA KONSUMEN DI ERA DIGITAL: BAGAIMANA REVIEW ONLINE MEMBENTUK PERILAKU KONSUMEN DIGITAL. *JUMDER: Jurnal Bisnis Digital Dan Ekonomi Kreatif*, 1(1), 1–20.
- Nannini, L., Manerba, M. M., & Beretta, I. (2024). Mapping the landscape of ethical considerations in explainable AI research. *Ethics and Information Technology*, 26(3). <https://doi.org/10.1007/s10676-024-09773-7>

- Naz, H., & Kashif, M. (2024). Artificial intelligence and predictive marketing: an ethical framework from managers' perspective. *Spanish Journal of Marketing - ESIC*. <https://doi.org/10.1108/sjme-06-2023-0154>
- Obadă, D.-R., Gradinaru, C., & Grădinaru, I.-A. (2026). From understanding to influence: the interplay of AI literacy, self-efficacy, and trust in predicting communication students' AI adoption and word-of-mouth. *Frontiers in Communication*, 10. <https://doi.org/10.3389/fcomm.2025.1722464>
- Omar, M., Soffer, S., Charney, A. W., Landi, I., Nadkarni, G. N., & Klang, E. (2024). Applications of large language models in psychiatry: a systematic review [Review of Applications of large language models in psychiatry: a systematic review]. *Frontiers in Psychiatry*, 15. Frontiers Media. <https://doi.org/10.3389/fpsyt.2024.1422807>
- Park, I., Lee, S., & Lee, D. (2022). Virtual Audience Providing AI-Generated Emotional Reactions to Enhance Self-Disclosure in Self-Introduction. *International Journal of Human-Computer Interaction*, 39(13), 2702–2713. <https://doi.org/10.1080/10447318.2022.2085338>
- Passlack, N., Hammerschmidt, T., & Posegga, O. (2025). With Great Power Comes Great Responsibility: What Shapes AI Literacy for Responsible Interactions of Knowledge Workers With AI? *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-025-10648-5>
- Prpić, J., Tacihagh, A., & Melton, J. (2015). The Fundamentals of Policy Crowdsourcing. *Policy & Internet*, 7(3), 340–361. <https://doi.org/10.1002/poi3.102>
- Rahardja, B. V., Rolando, B., Chondro, J., & Laurensia, M. (2025). MENDORONG PERTUMBUHAN E-COMMERCE: PENGARUH PEMASARAN MEDIA SOSIAL TERHADAP KINERJA PENJUALAN. *JUMDER: Jurnal Bisnis Digital Dan Ekonomi Kreatif*, 1(1), 45–61. <https://doi.org/10.1234/JUMDER.V1I1.6>
- Rahardja, B. V., Rolando, B., Chondro, J., & Laurensia, M. (2025). MENDORONG PERTUMBUHAN E-COMMERCE: PENGARUH PEMASARAN MEDIA SOSIAL TERHADAP KINERJA PENJUALAN. *JUMDER: Jurnal Bisnis Digital Dan Ekonomi Kreatif*, 1(1), 45–61. <https://doi.org/10.1234/JUMDER.V1I1.6>
- Raman, R., Pattnaik, D., Lathabai, H. H., Kumar, C., Govindan, K., & Nedungadi, P. (2024). Green and sustainable AI research: an integrated thematic and topic modeling analysis. *Journal Of Big Data*, 11(1). <https://doi.org/10.1186/s40537-024-00920-x>
- REVOLUTION: FROM TRADITIONAL TO DATA-DRIVEN APPROACHES.
- Riedl, R. (2022). Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. *Electronic Markets*, 32(4), 2021–2051. <https://doi.org/10.1007/s12525-022-00594-4>
- Rolando, B. (2024a). THE ALLURE OF VIRAL MAS BATIK ON TIKTOK: HOW A LIVE HOST'S CHARM ACCELERATES SALES GROWTH. *JISS (Journal of Interconnected Social Science)*, 3(1), 39–59.
- Rolando, B. (2024b). The Role Of Artificial Intelligence In Personalized And Customized Engagement Marketing: A Comprehensive Review. *Economics and Business Journal (ECBIS)*, 2(3), 301–316.
- Rolando, B. (2025). Examining Multi-Factor Marketing Dynamics on Consumer Purchase Intention: A TikTok Shop Analysis. *Journal of Business and Economics Research (JBE)*, 6(1), 213–224. <https://doi.org/10.47065/JBE.V6I1.6952>

- Rolando, B., & Dea, A. N. (2024). Pengaruh Kualitas Produk dan Inovasi Digital Marketing Terhadap Keputusan Pembelian Ulang Pelanggan Mcdonald's. *Journal of Trends Economics and Accounting Research*, 5(2), 192–205.
- Rolando, B., & Mulyono, H. (2024). UNLOCKING THE POWER OF DATA: EFFECTIVE DATA-DRIVEN MARKETING STRATEGIES TO ENGAGE MILLENNIAL CONSUMERS. *TRANSEKONOMIKA: AKUNTANSI, BISNIS DAN KEUANGAN*, 4(3), 303–321.
- Rolando, B., Mulyono, H., Hartanti, R., & Ingriana, A. (2024). Exploring The Theory Of Financial Innovation In The Context Of Industry 4.0. *Management Economics Trade and Accounting Journal (META-JOURNAL)*, 1(4), 147–158.
- Rotter, J., & Bailkoski, W. (2025). AI Adoption in NGOs: A Systematic Literature Review. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2510.15509>
- Saheb, T., & Saheb, T. (2024). Mapping Ethical Artificial Intelligence Policy Landscape: A Mixed Method Analysis. *Science and Engineering Ethics*, 30(2). <https://doi.org/10.1007/s11948-024-00472-6>
- Sanderson, C., Douglas, D., Lu, Q., Schleiger, E., Whittle, J., Lacey, J., Newnham, G., Hajkowicz, S., Robinson, C., & Hansen, D. (2023). AI Ethics Principles in Practice: Perspectives of Designers and Developers. *IEEE Transactions on Technology and Society*, 4(2), 171–187. <https://doi.org/10.1109/tts.2023.3257303>
- Sanusi, I. T., Agbo, F. J., Dada, O. A., Yunusa, A. A., Aruleba, K., Obaido, G., Olawumi, O., Oyelere, S. S., & Innovation, C. for M. R. and. (2024). Stakeholders' insights on artificial intelligence education: Perspectives of teachers, students, and policymakers. *Computers and Education Open*, 7, 100212–100212. <https://doi.org/10.1016/j.caeo.2024.100212>
- Shen, J., DiPaola, D., Ali, S., Sap, M., Park, H. W., & Breazeal, C. (2024). Empathy Toward Artificial Intelligence Versus Human Experiences and the Role of Transparency in Mental Health and Social Support Chatbot Design: Comparative Study. *JMIR Mental Health*, 11. <https://doi.org/10.2196/62679>
- Shrestha, A. K., & Joshi, S. (2025). Toward Ethical AI: A Qualitative Analysis of Stakeholder Perspectives. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2501.13320>
- Shrestha, A. K., Barthwal, A., Campbell, M., Shouli, A., Syed, S., Joshi, S., & Vassileva, J. (2024). Navigating AI to Unpack Youth Privacy Concerns: An In-Depth Exploration and Systematic Review. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2412.16369>
- Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, 21(1). <https://doi.org/10.1186/s41239-024-00446-5>
- Sun, P., Yu, H., Su, W., & Wu, Y. (2026). Privacy emotion recognition in artificial intelligence: Research frontiers, key challenges, and prospects. *Neural Networks*, 201, 108927–108927. <https://doi.org/10.1016/j.neunet.2026.108927>
- Sun, X., Shen, T., Jiang, Q., & Jiang, B. (2025). Research on the Impact of an AI Voice Assistant's Gender and Self-Disclosure Strategies on User Self-Disclosure in Chinese Postpartum Follow-Up Phone Calls. *Behavioral Sciences*, 15(2), 184–184. <https://doi.org/10.3390/bs15020184>
- Tan, D. M. (2022). A SYSTEMATIC REVIEW OF THE AI-POWERED MARKETING REVOLUTION: FROM TRADITIONAL TO DATA-DRIVEN APPROACHES. *Artificial*

- Intelligence Research and Applied Learning, 1(2).
<https://journal.dinamikapublika.id/index.php/aira>
- Tan, D. M. (2022). A SYSTEMATIC REVIEW OF THE AI-POWERED MARKETING REVOLUTION: FROM TRADITIONAL TO DATA-DRIVEN APPROACHES. *Artificial Intelligence Research and Applied Learning*, 1(2).
<https://journal.dinamikapublika.id/index.php/aira>
- Tan, D. M. (2022). A SYSTEMATIC REVIEW OF THE AI-POWERED MARKETING
- Tsumura, T., & Yamada, S. (2023). Influence of agent's self-disclosure on human empathy. *PLoS ONE*, 18(5). <https://doi.org/10.1371/journal.pone.0283955>
- Yu, L., & Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1061778>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review [Review of The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review]. *Smart Learning Environments*, 11(1). Springer Nature. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhang, S., Xing, H., Yi, X., & Li, H. (2024). Natural Language but Omitted? On the Ineffectiveness of Large Language Models' privacy policy from End-users' Perspective. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2406.18100>
- Zhou, T., & Tang, B. (2025). Examining user information disclosure on artificial intelligence-generated content platforms: A privacy calculus perspective. *Information Development*. <https://doi.org/10.1177/02666669251384288>