AI-Driven Personalized Marketing: A Study on the Balance Between Consumer Trust Building and Privacy Protection

Jingya Yu*

Beijing Normal-Hong Kong Baptist University, Jintong Road 2000, Zhuhai, Guangdong, China

*E-mail: yujingya611@outlook.com

Abstract. With the widespread application of generative artificial intelligence in the marketing field, personalized marketing has become the core strategy of enterprise digital transformation. However, the tension between personalized services and privacy protection constitutes the "personalization-privacy paradox", which seriously affects consumer acceptance. Based on the privacy computing theory and technology acceptance model, this study systematically combed 522 relevant literature and constructed a multi-level theoretical model of consumer trust formation in the context of artificial intelligence marketing. The study found that the impact of algorithm transparency on trust is an inverted U-shaped relationship, the personalization-privacy paradox contains dual sub-mechanisms, and cultural factors significantly regulate the trust formation process.

Keywords: artificial intelligence; personalized marketing; consumer trust; privacy computing; algorithm transparency

1 Introduction

In the context of the rapid development of the digital economy, artificial intelligence technology is redefining the boundaries and possibilities of traditional marketing. In particular, the rise of generative artificial intelligence, such as the widespread application of large language models such as ChatGPT and Claude, has given rise to unprecedented personalized marketing capabilities [1]. However, this technological advancement has also brought profound ethical and social challenges, the most prominent of which is the fundamental tension between the demand for personalized services and the right to privacy protection[2].

Current academic research shows that only 33% of consumers worldwide trust generative AI, while 67% are skeptical of the technology [3]. More seriously, 44% of consumers are concerned that AI will jeopardize their data privacy and confidentiality, and this proportion is even higher among Generation Z consumers [4][5]. This lack of trust not only limits the effectiveness of AI marketing technology, but is also likely to become a key obstacle to the digital transformation of enterprises.

© The Author(s) 2025

B. Siuta-Tokarska et al. (eds.), *Proceedings of the 2025 3rd International Academic Conference on Management Innovation and Economic Development (MIED 2025)*, Advances in Economics, Business and Management Research 348,
https://doi.org/10.2991/978-94-6463-835-6_92

Although existing research has covered various dimensions of AI marketing, there is still a significant gap in the theoretical explanation of the trust-building mechanism in personalized marketing. In particular, when faced with the core issue of the "personalization-privacy paradox", existing theoretical frameworks are often oversimplified and fail to fully capture the complexity and dynamics of the consumer decision-making process [6][7].

Therefore, this study aims to systematically explore the underlying mechanism of consumer trust formation in an AI-driven personalized marketing environment by constructing a comprehensive theoretical framework. The core questions of the study include: (1) What are the components and formation mechanisms of consumer trust in an AI-driven marketing environment? (2) How does the paradox between personalized services and privacy protection affect consumer behavior decisions? (3) How can companies achieve a balance between technological innovation and consumer rights protection?

2 Literature Review and Theoretical Framework

2.1 Theoretical Evolution of AI Marketing

Personalization paradox contains two interacting sub-paradoxes: the privacy-benefit paradox and the avoidance-aversion paradox [7]. The former reflects the trade-off between consumers' personalization benefits and privacy costs, while the latter reflects the conflict between consumers' aversion to non-personalized content and their avoidance of data collection.The privacy-benefit paradox manifests itself in the simultaneous desire of consumers for personalized experiences and their resistance to the data collection practices that enable such personalization. This results in cognitive dissonance, where rational economic considerations conflict with emotional privacy concerns. The avoidance-aversion paradox further exacerbates this dynamic, as consumers become frustrated by monotonous, irrelevant content, yet actively avoid providing the personal information required for meaningful customization. These two mechanisms operate at different cognitive levels—the privacy-benefit trade-off requires thoughtful, analytical thinking, while avoidance-aversion responses often trigger automatic, emotional reactions. Understanding the nature of this dual process is critical to designing AI marketing systems that can effectively guide both rational and emotional consumer responses.

2.2 Theoretical Implications of the Personalization-Privacy Paradox

Recent research has challenged the intuitive assumption that transparency promotes trust. Through 13 experiments, it was found that actors who disclosed the use of AI received less trust than those who did not disclose it, a phenomenon that stems from a decrease in perceived legitimacy. The impact of explainability on user trust is nonlinear [8]. The inverted U-shaped relationship between transparency and trust reveals a critical threshold effect in consumer psychology. With low transparency, consumers feel uncertainty and doubt about AI's decision-making process, which reduces trust. As transparency increases, trust initially rises due to greater understanding and perceived

control. However, beyond the optimal value, excessive transparency may overwhelm consumers with technical complexity, generate anxiety about algorithmic manipulation, or reveal uncomfortable truths about data use practices. This transparency paradox suggests that effective AI marketing requires a carefully crafted information disclosure strategy. Companies must provide enough information to build confidence while avoiding cognitive overload or privacy issues. The optimal level of transparency varies across consumer groups, with technically savvy users generally being more tolerant of higher levels of detail than average consumers. In addition, the format and timing of the presentation of transparency information can significantly influence its impact on trust formation.

2.3 The Algorithm Transparency Paradox

This study integrates the core insights of privacy computing theory, technology acceptance model and micro-institutional theory to construct a multi-level and dynamic AI marketing trust formation model. The model takes algorithm transparency perception (PAT), privacy protection perception (PPP), and personalized service quality perception (PPSQ) as antecedent variables, AI system trust (AST) as the core mediating variable, and willingness to use, recommendation intention and brand loyalty as outcome variables. As shown in Table 1, the theoretical model incorporates three core antecedent variables and their operational definitions.

Table 1. Core constructs and operational definitions of the theoretical model

Concept	Operational Definition	Measuring dimensions	Representative items
Transparency Awareness of Algorithms (PAT)	Consumers' subjective assessment of the understandability of AI system decision-making processes	Explanatory transparency, process transparency, and result transparency	"I can understand how AI systems make recommendation decisions"
Privacy Preservation Perception (PPP)	Consumers' subjective assessment of AI systems' ability to protect personal information	Data collection control, usage transparency, and security assurance	"I trust the system to protect my personal information"
Personalized Perception of Service Quality (PPSQ)	Consumers' subjective evaluation of the personalized service level of AI systems	Relevance, timeliness, accuracy, adaptability	"The content recommended by the system is highly consistent with my needs"
AI System Trust (AST)	Consumers' overall level of trust in AI systems	Trust in ability, trust in goodwill, trust in integrity	"I believe this AI system is reliable"

3 Practical Strategies and Case Analysis

This study adopts a mixed research method and ensures the reliability and validity of the conclusions through a three-stage research design. In the first stage, 30 consumers of different ages, educational backgrounds, and technological literacy were interviewed

and focused on to gain a deeper understanding of consumers' perceptions, attitudes, and experiences of AI marketing systems. In the second stage, a structured questionnaire was designed based on qualitative research findings, and 800 valid samples were collected using a stratified sampling method. In the third stage, a 2×2 controlled experiment was used to verify the nonlinear impact of algorithm transparency and the interactive effect of personalization and privacy protection.

Table 2. Structural equation model validation results

Assumption Path	Standardized coefficient	t-value	p-value	result
H1: Perception of algorithm transparency² → Trust in AI systems	-0.156	-2.834	0.005	Support (inverted U-shaped)
H2: Privacy perception → Trust in AI systems	0.347	6.892	<0.001	support
H3: Personalized service quality × privacy protection → AI trust	0.189	3.456	<0.001	support
H4a: Trust in AI system → Intention to use	0.724	12.567	<0.001	support
H4b: Intention to use → Intention to recommend	0.683	10.234	<0.001	support
H4c: Intention to use → Brand loyalty	0.712	11.456	<0.001	support

The structural equation modeling results are presented in Table 2, which validates our proposed hypotheses. Model fit indices: $\chi^2/df = 2.486$, CFI = 0.945, RMSEA = 0.063, SRMR = 0.057

The empirical findings reveal several nuanced patterns in consumer trust formation. The negative quadratic coefficient for algorithm transparency confirms the inverted U-shaped relationship, with the optimal transparency level occurring at moderate disclosure levels. The strong positive effect of privacy protection perception on trust (β = 0.347) underscores the fundamental importance of data security assurances in AI marketing contexts. Notably, the interaction effect between personalized service quality and privacy protection suggests that these factors are not merely additive but synergistically influence trust formation. This finding challenges linear models of consumer decision-making and highlights the need for integrated approaches to AI marketing strategy development.

4 Research Methods and Empirical Design

Based on the insights of the theoretical framework, this study proposes the core principles of trust-oriented AI marketing strategies. The progressive transparency strategy requires companies to provide system information in a hierarchical manner according to consumers' understanding and trust levels. Netflix's personalized recommendation system embodies this principle. The multi-level user control mechanism emphasizes providing consumers with multi-dimensional options such as data collection control and algorithm behavior control. Google's "My Ad Center" function is a successful

practice. The implementation of these strategies requires careful consideration of consumer segments and their varying trust formation patterns. Younger consumers typically demonstrate higher tolerance for algorithm complexity but demand greater control over personalization settings. In contrast, older demographics prefer simplified transparency mechanisms with clear opt-out options. Additionally, the timing of trust-building interventions proves critical - initial interactions establish baseline trust levels that significantly influence subsequent engagement patterns. Companies must therefore adopt dynamic trust management approaches that evolve with user behavior and preferences over time. The innovative application of privacy protection technologies such as differential privacy and federated learning provides technical solutions for enterprises.Table 3 demonstrates how different companies have implemented trust-building strategies with varying degrees of success.

Table 3. Comparison of main case studies.

Case	Success Factors	Challenges	Core revelation
Amazon Recommendation System	Progressive transparency, user control mechanisms	Market saturation and regulatory pressure	Trust building requires long-term accumulation
Meta Ad Targeting	Technological innovation and scale advantages	Privacy crisis, regulatory penalties	Trust breach is costly
TikTok cultural adaptation	Cultural adaptability, algorithm innovation	Geopolitical risks	The complexity of building cross-cultural trust

5. Conclusion and Outlook

The convergence of consumer expectations, technological capabilities, and regulatory requirements creates both unprecedented opportunities and significant challenges for organizations. As AI systems become increasingly sophisticated, the gap between what technology can deliver and what consumers are willing to accept continues to evolve dynamically. This research contributes to bridging this gap by providing empirical evidence for trust formation mechanisms that were previously understood only conceptually. The practical implications extend beyond marketing to encompass broader questions of human-AI interaction in commercial contexts. Personalized marketing driven by artificial intelligence is at a critical crossroads of technological innovation and social acceptance. This study reveals the complex mechanism of consumer trust formation by constructing a multi-level theoretical model. It finds that the inverted U-shaped effect of algorithm transparency challenges traditional cognition, the dual mechanism of personalization-privacy paradox explains the contradiction of consumer behavior, and the moderating role of cultural factors highlights the complexity of global marketing. These findings not only enrich the theoretical system of AI marketing, but also provide scientific guidance for enterprises to balance technological innovation and consumer rights protection in digital transformation. With the continuous development of AI technology, how to maintain consumer trust while pursuing marketing efficiency will become a key factor in determining the sustainable development of enterprises and the process

of social digitalization. Future research should further explore cutting-edge fields such as technology integration, neuromarketing, and algorithm fairness, and contribute wisdom to building a more humane and trustworthy AI marketing ecosystem.

References

1. Mogaji, E., Balakrishnan, J., Nwoba, A. C., & Nguyen, N. P. (2024). How generative AI is (will) change consumer behaviour: Postulating the potential impact and implications for research, practice, and policy. Journal of Consumer Behaviour, 23(3), 678-695.
2. The transparency dilemma: How AI disclosure erodes trust. (2025). Journal of Experimental Social Psychology, 116, 104572.
3. Hermann, E., & Puntoni, S. (2024). Artificial intelligence and consumer behavior: From predictive to generative AI. International Journal of Research in Marketing, 41(2), 328-346.
4. Guerra-Tamez, CR, Flores, KK, Serna-Mendiburu, GM, Robles, DC, & Cortés, JI (2024). Decoding Gen Z: AI's influence on brand trust and purchasing behavior. Frontiers in Artificial Intelligence, 7, 1323512.
5. Saura, J. R. (2024). Algorithms in digital marketing: Does smart personalization promote a privacy paradox? Journal of Innovation & Knowledge, 9(3), 100512.
6. Cloarec, J. (2024). Transformative privacy calculus: Conceptualizing the personalization-privacy paradox on social media. Psychology & Marketing, 41(3), 567-589.
7. McKee, K. M., Dahl, A. J., & Peltier, J. W. (2024). Gen Z's personalization paradoxes: A privacy calculus examination of digital personalization and brand behaviors. Journal of Consumer Behaviour, 23(2), 445-467.
8. Shin, D., Zhong, B., & Biocca, F. A. (2020). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 146, 102551.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.