ijonses International Journal on Studies and Education Science

www.ijonses.net

Evaluating How Personalized AI Agents Influence Decision-Making, Self-Presentation, and Digital Identity Management: A Literature Review

Oluwatoyosi Ogunsola

Ball State University, 2000 W. University Ave. Muncie, IN 47306, USA, 0009-0005-2224-7463
Corresponding author: Ogunsola, A. Oluwatoyosi (oluwatoyosi.ogunsola@bsu.edu)

Article Info

Article History

Received:
12 August 2025

Revised:
20 November 2025

Accepted:
15 December 2025

Published:
1 January 2026

Keywords

Personalized AI agents
Recommender system
Decision making
Digital identity
Algorithmic self
Extended self
Identity ownership
Autonomy
Nudging

Abstract

The increased interaction of humans with personalized AI agents has had a significant impact on how individuals engage with the digital world, influencing everything from the choices they make to the digital personas they present. While personalization is designed to optimize user experience, current scholarship remains fragmented, treating its influence on decision-making, self-presentation, and digital identity management as separate phenomena, thus failing to capture their systemic entanglement. This paper addresses this gap with a systematic literature review across Human-Computer Interaction, Communication Studies, and Psychology, building upon the theoretical concept of the extended self. The review reveals that personalized AI agents act as adaptive mediators, co-producing user behavior. Key findings confirm that personalized AI agents introduce systematic risks: algorithmic nudging erodes autonomy in decision-making; optimization for engagement leads to homogenization and loss of authenticity in self-presentation; and data-driven profiling actively constructs a constrained "Algorithmic Self" in digital identity management. Critically, these effects are found to operate through a continuous co-adaptive feedback loop. We propose an Integrative Evaluative Framework that formalizes this interconnected relationship as a Co-Adaptive Cycle. This framework offers a unified lens for scholarship. It generates critical implications for design, necessitating a shift toward systems that prioritize transparent identity modeling, support identity exploration, and ensure user ownership of the algorithmically extended self.

Citation: Ogunsola, O. (2026). Evaluating how personalized AI agents influence decision-making, self-presentation, and digital identity management: A literature review. *International Journal on Social and Education Sciences (IJonSES), 8*(1), 56-74. https://doi.org/10.46328/ijonses.5800

ISSN: 2688-7061 / © International Journal on Social and Education Sciences (IJonSES).

Introduction

The rise of personalized artificial intelligence (AI) agents has reshaped the ways individuals navigate digital environments. From recommender systems that guide consumer decisions to generative tools that craft online profiles, personalization has become central to human–AI interaction. While personalization promises efficiency and relevance, it also carries deeper implications for how people make choices, present themselves, and manage evolving digital identities. These processes are not isolated; instead, they are entangled within what scholars have described as the "extended self" in digital contexts (Belk, 2013), where technologies serve as integral components of identity construction and expression.

Artificial intelligence (AI) systems, ranging from algorithmic recommenders to adaptive chatbots and virtual assistants, leverage personal data and behavioral histories to tailor communication, decision support, and online experiences to individual users. As a result, AI-driven personalization is not only transforming how people navigate information landscapes but also how they present themselves and manage their digital personas. These systems are not just personal; they are transforming organizational workflows, consumption patterns, and the managerial control over employee data, making their influence a core concern for the Information Systems community.

Current scholarship has addressed fragments of this relationship: studies of recommender systems highlight the influence of algorithmic nudges on decision-making and autonomy (Shin & Park, 2019); research on social media personalization underscores its impact on self-presentation and impression management (Toma & Hancock, 2013); and investigations into digital identity emphasize concerns around privacy, coherence, and long-term ownership of personal data (Floridi, 2017). However, this literature often remains siloed, treating decision-making, self-presentation, and identity management as separate domains. These raise significant questions about the autonomy of user decision-making, the authenticity of self-presentation, and the dynamics of digital identity in increasingly algorithmic environments. What is missing is a holistic framework that captures how personalized AI agents simultaneously mediate all three.

This paper seeks to address that gap through a literature review that integrates insights across human–computer interaction, communication, psychology, and AI ethics. By synthesizing current research on the influence of personalized AI agents in three interconnected areas: decision-making, self-presentation, and digital identity management, the review seeks to illuminate the mechanisms by which AI agents shape user experiences and to identify critical gaps for future scholarship. In doing so, it contributes to a deeper understanding of how personalization technologies are reconfiguring the boundaries of individual agency, authenticity, and identity in contemporary digital society.

The goal is not only to synthesize existing evidence but also to advance design-oriented implications for building AI systems that support user autonomy, authenticity, and ownership of identity. The objective of this paper is to evaluate how personalized AI agents influence three interrelated domains of human experience: decision-making, self-presentation, and digital identity management. Building on the concept of the extended self (Belk, 2013), the

paper aims to integrate fragmented strands of research into a cohesive framework that can guide both scholarly inquiry and practical design. To achieve this, the paper is guided by three overarching research questions:

1. How do personalized AI agents influence decision-making and autonomy?
2. How do personalized AI agents influence self-presentation and authenticity?
3. How do personalized AI agents influence digital identity construction, maintenance, and ownership?

Together, these questions form the foundation for a literature review that synthesizes findings across disciplines and proposes an integrative evaluative framework. The paper ultimately seeks to contribute to the design of AI agents that respect autonomy, foster authentic expression, and support identity ownership in increasingly personalized digital ecosystems.

Methodology

This study employed a structured literature review approach to evaluate the influence of personalized AI agents on decision-making, self-presentation, and digital identity management. The method was designed to ensure comprehensive coverage of relevant scholarship across multiple disciplines, including Human-Computer Interaction (HCI), Communication Studies, Psychology, and AI Ethics.

The initial search was conducted across four major academic databases: ACM Digital Library, Web of Science, PsycINFO, and Google Scholar. The search was conducted iteratively using combinations of core keywords derived from the research questions. The search was restricted to articles published between 2010 and 2024 to capture the emergence and maturation of modern, deep-learning-based personalized AI systems.

The primary search strings combined terms related to the *AI* mechanism with terms associated with the human outcome*:*

- Mechanism Terms: “personalized AI agent”, “algorithmic personalization”, “recommender system”
- Outcome 1 (Decision): “decision-making”, “autonomy”, “algorithmic nudging”, “filter bubble”
- Outcome 2 (Presentation): “self-presentation”, “impression management”, “authenticity”, “digital performance”
- Outcome 3 (Identity): “digital identity”, “extended self”, “algorithmic self”, “identity management”, “ownership”.

Inclusion & Exclusion Criteria

Studies were included if they met the following criteria:

1. Peer-reviewed journal articles, high-impact conference proceedings (e.g., CHI, CSCW, AAAI), and scholarly book chapters.
2. Papers that explicitly provided empirical evidence, theoretical models, or conceptual analysis regarding the *impact* or *influence* of personalized, adaptive AI systems on users.
3. Sources addressing at least one of the three core domains: decision-making, self-presentation, or digital identity management.

Exclusion criteria included:

1. Papers focused solely on the technical development or optimization of AI algorithms without discussing user impact (e.g., maximizing accuracy of non-personalized systems).
2. Non-English language articles.

The initial search produced approximately 1250 studies (replace with number). After title and abstract screening, 120 studies were retained for full-text review. Of these, 62 studies met the final inclusion criteria. Reference lists of selected articles were also scanned to identify additional relevant work.

Data Extraction and Synthesis Protocol

A thematic analysis approach was used to synthesize the findings from the final set of articles. Data extraction focused on the following elements for each paper:

1. The specific type of personalized AI agent examined (e.g., recommender, chatbot, virtual assistant).
2. The mechanism of influence identified (e.g., nudging, anchoring, homogenization).
3. The outcome observed in the user (e.g., reduced autonomy, increased conformity, identity blurring).
4. Key conceptual frameworks applied (e.g., extended self, automation bias).

The extracted findings were initially grouped according to the three primary outcome domains, forming the structure for the main body of the review. The final stage of synthesis involved a cross-domain comparison to identify overlapping mechanisms and reciprocal effects, which ultimately led to the development of the Integrative Evaluative Framework and the concept of the Co-Adaptive Cycle presented in the Discussion.

Literature Review

Understanding Personalized AI Agents

Personalized AI agents are AI agents designed for intelligent systems that adapt behaviors, tailor experiences, recommendations, and outputs based on user-specific data, preferences, and contexts. They are not limited in application and are used for multiple purposes, including recommendation algorithms, adaptive chatbots, digital assistants, and social media moderation bots. Personalized AI agents are software systems designed to tailor interactions, content, and services based on individual user data such as preferences, behaviors, and contextual cues. These systems encompass a broad range of applications, including recommendation algorithms, adaptive chatbots, digital assistants, and social media moderation bots. Their core functionality relies on data-driven modeling techniques, machine learning, natural language processing, and user profiling to deliver customized experiences that enhance relevance, engagement, and utility.

A distinction, however, must be made between regular AI agents and personalized AI agents. While regular agents operate through generalized models that deliver the same output to all users, personalized AI agents adapt dynamically to individual preferences, behaviors, and contexts. Drawing on user data, such as interaction histories, demographic information, or real-time activity, personalized agents generate responses and recommendations

tailored to specific individuals. This personalization makes them more deeply embedded in users' decision-making, self-presentation, and identity management, positioning them as extensions of the self rather than mere tools.

Historically, AI personalization originated from early recommender systems in e-commerce and information retrieval, evolving into sophisticated agents that dynamically adapt to users in real-time (Sundar, 2020). This shift marks a movement from static personalization based on demographic segmentation to more granular, real-time, and context-aware adaptations. Modern AI agents not only respond to explicit user inputs but also infer preferences from implicit data such as browsing patterns, social network interactions, and biometric signals. (Li & Unger, 2021).

The theoretical understanding of personalized AI agents draws on multidisciplinary perspectives. Sociotechnical systems theory emphasizes the co-shaping of human behavior and technology functionality, recognizing AI agents as active mediators rather than passive tools (Salwei & Carayon, 2022). This activeness enables AI agents to regulate their behavior through adaptive feedback mechanisms, allowing them to act as mediators that shape human behavior (Shad, 2025) and consequently influence how users represent themselves online. Embodied cognition and mutual theory of mind (MToM) models further highlight the dynamic, reciprocal nature of user-agent interactions, where agents learn from users as users adapt their behaviors in response to agent feedback.
By situating personalized AI agents within these evolving technical and theoretical contexts, this review lays the foundation for analyzing their influence on decision-making, self-presentation, and digital identity management in subsequent sections.

Personalization and The Extended Self

Belk's (2013) concept of the extended self highlights how possessions, now including digital technologies, become integral to identity construction and expression, providing a foundation for understanding the impact of personalized AI agents. Empirical research supports this view as smartphones, social media accounts, and digital archives are not merely tools; they function as extensions of memory, personality, and self-presentation (Bartoli, 2022; Chan, 2022; Dorčić et al., 2023). Personalized AI agents advance this phenomenon by moving beyond static storage or passive representation toward active co-construction of identity, making them active and generative partners in shaping user behavior.

By tailoring outputs and recommendations to the individual, these agents blur the boundary between self and system. In doing so, they embed themselves within the processes of identity construction, expression, and management (Joseph & Chaudhry, 2024; Ogunsola & Fisher, 2025). This marks a crucial shift: whereas earlier technologies served as repositories of self, personalized agents now act as co-authors, actively participating in shaping users' decisions, preferences, and digital performances. Consequently, they function as continuous, adaptive mirrors that reflect, amplify, and sometimes reframe aspects of the self, thus positioning themselves as extensions of identity rather than mere tools.

Personalized AI Agents and Decision-Making

Personalized AI agents shape decision-making by actively curating options, prioritizing information, and nudging users toward particular choices. For example, recommender systems on e-commerce platforms demonstrate that algorithmic suggestions can shape preferences and influence consumer choice sets, rather than merely revealing preexisting tastes (Adomavicius et al., 2013). This influence is amplified because Personalized AI Agents **l**everage behavioral data to tailor these influences uniquely to the individual.

Mechanisms of Influence

Scholars have identified several mechanisms by which personalized agents influence decisions, one of which is Algorithmic nudging. Algorithmic nudging uses subtle interface and ranking manipulations to steer choices (Schmauder et al., 2023). Recent work calls for interdisciplinary oversight because AI-driven nudges scale personalization to an unprecedented degree (e.g., tailoring nudges to psychological profiles). For instance, in the Bandit Task Experiment (Dezfouli et al., 2020), personalized AI agents nudged users up to 70% of the time. This adaptation isn't passive; the AI learns to dynamically fine-tune its optimal strategy to each individual participant by observing their decision-making style during early trials (Schmauder et al., 2023).

This active and personalized decision support is also evident in high-stakes contexts. Autonomous personalized agents are deployed by financial institutions for trading, risk assessment, and portfolio management (Ganji & Zarifhonarvar, 2025; Zhao et al., 2024), as well as in healthcare diagnostics and virtual assistants (Piccialli et al., 2025). These systems use decision modules, learning mechanisms, and preference modeling to not only optimize actions but also personalize decision support to individual users in dynamic environments.

Adaptive Feedback Loops and Filter Bubbles

Feedback loops occur when personalization adapts continuously to user behavior, reinforcing existing patterns. Jiang et al. (2019) identify “degenerate feedback loops” in recommender systems, where repeated interactions lead to narrowing of exposure. Symeonidis et al. (2019) similarly note that accuracy-optimized algorithms reduce novelty, deepening echo chambers. Li et al. (2025) illustrate these effects in short-video platforms, showing how filter bubbles emerge rapidly as personalization accelerates. These dynamics demonstrate that personalization not only influences immediate choices but also shapes long-term informational environments. Critically for the Information Systems community, these systematic biases and narrowings raise concerns regarding the organizational capacity to innovate and the potential for reinforcing managerial biases when employees rely heavily on these constrained information landscapes.

Trust, Delegation, and Automation Bias

Trust in AI systems mediates whether users adopt or reject personalized recommendations. Yang, Schemanske, and Searle (2021) found that trust shifts dynamically, making users more likely to delegate decisions after consistent positive outcomes. Shin and Park (2019) highlight “algorithm aversion,” where transparency or errors

reduce reliance despite objective accuracy. The influence of personalized agents is thus inseparable from users' perceptions of competence, transparency, and reliability. Additionally, the presentation of personalized outputs alters decision outcomes: anthropomorphic framing increases compliance (Xu et al., 2024), while framing AI limitations differently shifts user trust and adoption (Kim & Song, 2022). When well-designed and personalized, AI agents can enhance decision efficiency and user satisfaction. However, algorithmic personalization risks reducing autonomy through over-reliance and covert manipulation: users may accept AI suggestions without understanding the rationale, resulting in "nudged" decisions that appear seamless but are not fully self-authored. This risk is amplified because personalization enables precisely timed, psychologically targeted nudges (Schmauder et al., 2023).

Personalized AI Agents and Self-Presentation

Self-presentation, the active management of impressions in social contexts (Goffman, 1959), has long been a central concept in understanding online behavior. Self-presentation in digital spaces is a strategic act, where individuals manage impressions to navigate audience expectations and social meaning (Blunden & Brodsky, 2024). With the rise of personalized AI agents, new paradigms in digital self-presentation are emerging, as humans utilize AI as extensions or tools to aid communication and impression formation (Benke et al., 2020; Endacott & Leonardi, 2022). These agents act as active mediators of self-presentation rather than passive tools. Personalized AI's role in self-presentation is especially salient in the following characteristics.

Avatar and Embodiment Personalization

Systematic reviews of avatar customization technologies reveal that highly adaptable, AI-driven avatar creation tools now allow users to tailor their digital representations closely, modulating everything from appearance to movement and expression (Wu et al., 2023). Personalization is crucial here, as personal AI agents utilize behavioral and preference data to suggest the most engaging or context-appropriate avatar. The depth of customization not only enhances users' sense of presence and identifiability but also directly influences interaction patterns and social outcomes (Ouhnni et al., 2025). Features such as real-time motion synchrony and AI-powered feedback enable avatars to mirror the user's movements, thereby strengthening embodied perception and the sense of being "present" in the virtual space (Jung et al., 2022). In virtual reality game play, these embodied agents can enhance self-presence, but they also raise questions about authenticity and identity continuity across different contexts. For example, a study found that the presence of another avatar during gameplay led to less intense play compared to solo play. Disclosure of the avatar as AI heightened effort intensity compared to non-disclosed AI companions, and a masculine AI appearance reduced effort intensity (Visser et al., 2024).

Adaptive Content Prompts and Auto-Generated Profiles/Bios

On platforms like TikTok, Instagram, and virtual worlds, AI-generated bios use behavioral data and contextual cues to suggest or modify digital personas, introducing a layer of algorithmic curation to self-presentation (Kang & Lou, 2022). For example, AI-driven feedback can prompt users to present themselves in ways that align with prevailing trends or "approved" norms, which may boost engagement. This algorithmic endorsement of specific

identity performances can inadvertently risk producing standardized, platform-optimized language that flattens individual voice. Customization via personalized AI agents empowers users, but it also requires significant cognitive effort to manage the communication process (Kang & Sundar, 2013; Sundar & Marathe, 2010).

Algorithmic Shaping and Homogenization

Key empirical studies on algorithmic homogenization by personalized AI agents reveal how AI-driven content curation shapes user behavior and identity expression, often promoting conformity across digital environments (Fei et al., 2025). Personalized AI recommender systems on social media platforms tend to surface content that aligns with prevailing popularity metrics and user engagement patterns (Kang & Lou, 2022). This causes users to adjust their self-presentation and content creation to fit the AI's preferences, resulting in a narrowing of diversity and increased similarity in visible digital identities (Nguyen et al., 2024; Zhang et al., 2023). The perceived homogenization of social media content can lead to increased impatience in subsequent consumption activities (Fei et al., 2025). When AI agents suggest phrasing, imagery, or identity framings, they can both empower users (by helping articulate identity) and erode authenticity (by standardizing expression).

Personalized AI Agents and Digital Identity Management

Personalized AI agents play a central role in shaping how individuals construct, maintain, and negotiate their digital identities (Jiang et al., 2025). This has led to concepts like the "Algorithmic Self," which is a digitally mediated identity in which personal awareness, preferences, and even emotional patterns are shaped through continuous feedback from AI systems (Turtle et al., 2024). Unlike earlier digital tools, these agents do not merely reflect identity but actively participate in its production by filtering representations, curating narratives, and mediating visibility (Wang & Wang, 2023).

Algorithmic Construction of the Digital Self

AI systems increasingly assemble "statistical individuals" based on patterns of data rather than direct self-expression (Bjerring & Busch, 2024). This identity, referred to as a statistical individual, is partly co-authored by systems that infer traits, preferences, and social roles (Joseph, 2025). This process often privileges the system's statistical model over the user's subjective sense of self-agency in identity construction. For example, if a user primarily engages with content tagged as "introvert," the system perpetuates similar content, reinforcing that identity label and potentially limiting personal growth or change. These algorithmic feedback loops create a cycle where the AI shapes and constrains how users perceive and perform their identities, sometimes without their awareness (Wu, 2025).

Feedback Loops and Identity Constraint

Sustained interactions with personalized algorithms reconstruct user experiences of time, selfhood, and identity, effectively shifting identity formation from reflective introspection to algorithmically guided narratives (Pronzato, 2025). Such processes raise concerns about ownership and authenticity: when AI defines identity through inferred

signals, the boundary between self and system blurs. For example, Lagerkvist et al. (2022) suggest that algorithmic biometrics will lead to a novel form of objectification of the human body. Furthermore, studies demonstrate how AI writing assistants impact life-story work (Wise et al., 2025) and how chatbots personalize yet stereotype, thereby reinforcing biases (Kantharuban et al., 2024). Endacott (2021) highlights how workplace AI alters professional identity construction, sometimes constraining authenticity through rigid categories. Critically, this constraint on professional identity directly impacts human capital management and the design of organizational IS platforms by forcing workers into rigid, algorithmically-approved behavioral categories, thus limiting organizational talent and adaptation.

Taken together, these literatures demonstrate that digital identity is no longer a static product of user self-presentation but an emergent property of human–AI co-production. While personalization offers opportunities for enhanced self-expression and agency, risks of homogenization, misrepresentation, and bias remain acute.

Discussion

This literature review aims to evaluate the impact of personalized artificial intelligence (AI) agents on three interrelated domains of human experience: decision-making, self-presentation, and digital identity management, by synthesizing fragmented research across multiple disciplines, including HCI, psychology, communication, and AI studies. Building on Belk's (2013) concept of the extended self, the review highlights evidence that personalized AI agents are not merely tools; they are active mediators that shape and co-produce user behavior and identity. The findings reveal that the influence of personalization cannot be understood by examining these domains in isolation.

Table 1. Summary of the Main Findings of Related Work

Area of Influence	Main Finding/ Mechanism	Key References (Examples)	Implication for the Extended Self
Decision-Making	Algorithmic Nudging & Anchoring	Schmauder et al., 2023; Dezfouli et al., 2020; Adomavicius et al., 2013	Risks to user autonomy; decisions are *co-authored* or *nudged*, leading to choices that are not fully self-authored.
	Adaptive Feedback Loops & Filter Bubbles	Jiang et al., 2019; Symeonidis et al., 2019; Li et al., 2025	Narrows the informational environment and limits exposure to novelty, reinforcing existing cognitive patterns and preferences.
	Trust, Delegation, & Automation	Yang, Schemanske, & Searle, 2021; Shin & Park, 2019	High trust leads to delegation of decision-making, positioning the AI agent as a reliable cognitive extender,

Area of Influence	Main Finding/ Mechanism	Key References (Examples)	Implication for the Extended Self
	Bias		but also increasing the risk of uncritical reliance.
Self-Presentation	AI as Co-Curator & Embodiment	Wu et al., 2023; Sun et al.,2018.; Benke et al., 2020	Personalized tools (e.g., avatars, bio-generators) serve as active extensions of the self for impression management, mediating the performance of identity.
	Algorithmic Homogenization	Nguyen et al., 2024; Zhang et al., 2023; Kang & Lou, 2022	AI-driven content promotion (optimized for engagement) encourages users to conform to "approved" norms, risking a loss of authenticity and flattening of diverse digital identities.
Digital Identity Management	Algorithmic Self	Turtle et al., 2024; Bjerring & Busch, 2024; Joseph, 2025	Identity becomes an emergent property based on the AI's inferred statistical model, rather than solely conscious self-expression.
	Feedback Loops and Identity Constraint	Wu, 2025; Pronzato, 2025	Sustained interaction creates cycles where the AI reinforces and sometimes constrains how users perceive and perform their identities, limiting identity exploration or change.
Intersection/ Synthesis	Co-Adaptive Framework	Joseph, J. (2025); Chaudhry & Debi (2023)	The influence on decision-making, self-presentation, and identity management is intertwined. AI is an embedded part of the "extended self," mediating all three domains simultaneously.

The synthesis of the literature provides clear answers to the guiding research questions, revealing a consistent pattern of both enablement and constraint across all three domains.

How do personalized AI agents influence user decision-making and autonomy?

Personalized AI agents significantly impact decision-making through algorithmic nudging (Schmauder et al., 2023), preference anchoring (Adomavicius et al., 2013), and the creation of adaptive feedback loops that narrow

the scope of information exposure (Jiang et al., 2019; Symeonidis et al., 2019). While personalization offers benefits like reduced search costs and increased efficiency (Jannach & Adomavicius, 2016), the core risk identified is the potential erosion of autonomy. Users frequently delegate choices to agents based on trust (Yang, Schemanske, & Searle, 2021), which can increase efficiency but also lead to automation bias, where decisions are influenced rather than fully self-authored. The literature strongly suggests that this influence is not passive; AI agents learn from and dynamically adapt their optimal nudging strategies to individual users (Schmauder et al., 2023), for example, YouTube's recommender system progressively filters suggested content.

In what ways do AI systems influence self-presentation and authenticity in digital environments?

AI systems act as co-curators of self-presentation, significantly influencing how individuals present themselves online. Personalized tools for avatar design, adaptive content prompts, and auto-generated bios (Wu et al., 2023; Sun et al., 2018) empower users with high fidelity and customization, making AI an extension of impression management (Benke et al., 2020). At the same time, AI-driven recommendation systems introduce a counter-pressure: algorithmic homogenization (Nguyen et al., 2024; Zhang et al., 2023), which incentivizes conformity to "approved" norms. While some users strategically adapt to these cues, others risk losing authenticity or diversity in their digital identities. The strategic adaptation by users to align with algorithmic preferences forms a crucial part of the human-AI interaction.

How do personalized AI agents influence digital identity construction, maintenance, and ownership?

Personalized AI agents contribute to constructing the "Algorithmic Self" (Turtle et al., 2024), in which identity emerges from human–AI co-production rather than solely conscious self-curation. This construction offers opportunities for identity exploration but also introduces risks. AI systems create statistical individuals by inferring traits from data (Bjerring & Busch, 2024), which can result in personalized yet potentially limited or stereotypical representations (Kantharuban et al., 2024). The main risks are the narrowing of identity through self-reinforcing feedback loops and fundamental questions of ownership and privacy. When AI continuously optimizes a persona or provides identity scaffolding (Wise et al., 2025), the boundary between the user's true self and their algorithmically shaped self-blurs, raising ethical concerns about data control and long-term identity coherence. The most critical finding of this review is the deep entanglement of decision-making, self-presentation, and digital identity management, which the existing literature often treats as siloed. Personalized AI agents create a unified system where these three domains operate through continuous, co-adaptive feedback loops (see Table 2). This synthesis enables the proposal of an integrative evaluative framework centered on the central role of personalized feedback.

Table 2. The Co-Adaptive Cycle

Step	Mechanism	Description
1. Decision-Making Shapes	Data Input	A user makes a choice (a *decision*), such as clicking a recommendation or purchasing a product. The AI immediately

Step	Mechanism	Description
Identity		incorporates this behavioral data to refine its model of the user's preferences, effectively constructing the user's algorithmic identity (the "statistical individual").
2. Identity Shapes Self-Presentation	Algorithmic Nudge	This newly refined algorithmic profile (identity) then guides the AI's suggestions for how the user should present themselves (e.g., adaptive content prompts, suggested bios, avatar styles). The AI is optimizing the user's public-facing persona (self-presentation) based on its internal model.
3. Self-Presentation Shapes Decisions	Reinforcement/ Feedback	The resulting public-facing performance (*self-presentation*) generates new data (likes, engagement, comments). This data is fed back into the system, which uses it to further tailor future recommendations and nudges (*decisions*), thereby closing the loop.

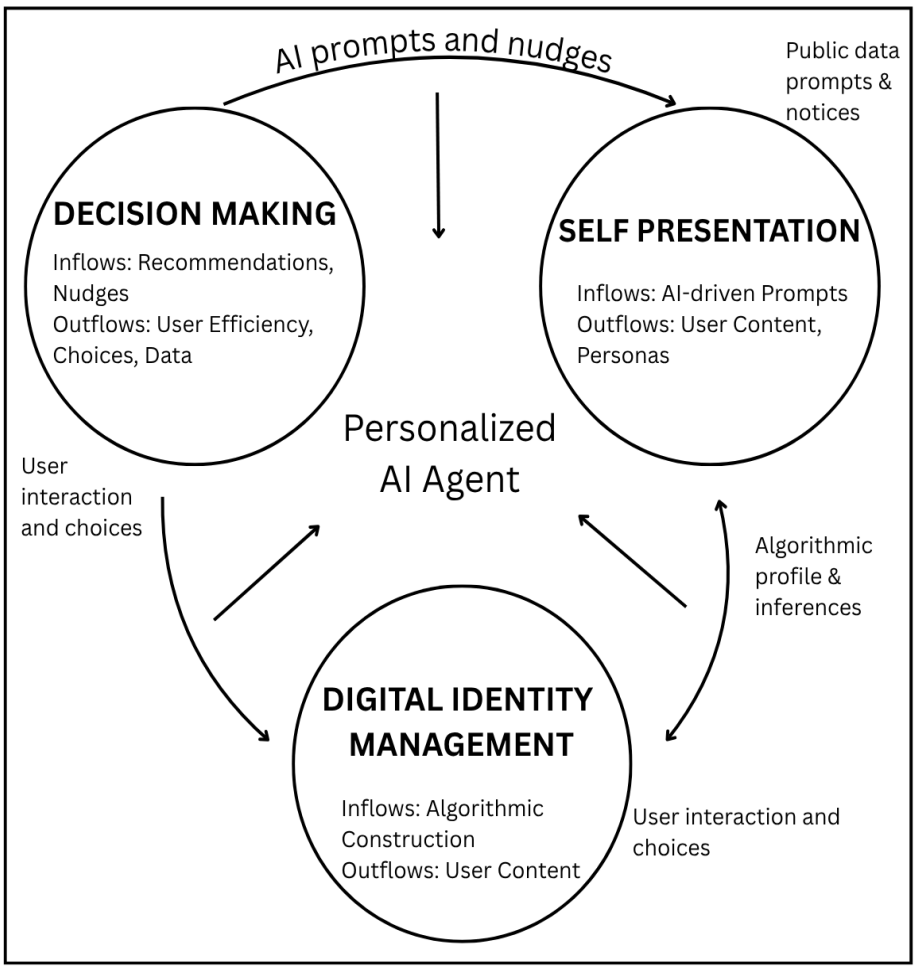

Figure 1. A Comprehensive Framework of the Co-Adaptive Cycle

The relationship between the three domains is characterized by a cycle of co-adaptation (see Figure 1):

1. Decision-Making → Identity: A user's initial decision (e.g., clicking on a particular video recommended by an AI) provides data that is instantly incorporated into the AI's model of the user's preferences, thus shaping the statistical identity the system creates.
2. Identity → Self-Presentation: This algorithmic profile then informs the AI's nudges for self-presentation (e.g., suggesting content, modifying a bio, or recommending an avatar style). A profile tagged as "fitness enthusiast" will receive prompts and templates aligned with that identity.
3. Self-Presentation → Decision-Making: The resulting self-presentation (e.g., posting a fitness-related picture) is then consumed by the AI, which uses this new data to further refine the recommendations (decisions) it offers in the future, thereby closing the loop.

This cycle (see Figure 2), as summarized by the finding that personalization co-creates user identity dynamics through feedback loops (Wu, 2025), is the engine of the extended self in the age of AI. The AI agent acts as a constant, adaptive mirror, reinforcing certain behaviors while subtly constraining others. The risks of algorithmic nudging in decision-making are inseparable from homogenization in self-presentation and the long-term coherence of identity, underscoring the need for transparent, autonomy-supportive personalization systems.

Implications for Designers of Personalized AI Agents

The findings synthesized in this literature review, particularly the establishment of the Co-Adaptive Cycle within the Integrative Evaluative Framework, carry critical implications for the design and deployment of personalized AI systems. Designers must shift their focus from optimizing *platform engagement* to actively supporting *user autonomy, authenticity, and long-term identity coherence*. The challenge is to create AI extensions of the self that enable human flourishing rather than merely reinforcing existing patterns or driving homogenization. Based on the identified risks, algorithmic nudging, authenticity erosion, and lack of identity ownership, we propose three core design principles:

Transparent and Explainable Identity Modeling

Personalized AI agents, by definition, construct a "statistical individual" or algorithmic self based on user data. This identity model is the core driver of both recommendations (decisions) and content prompts (self-presentation). Designers must treat this model as a critical interface element, not just an internal mechanism.

Agents should not only explain *why* a specific recommendation was made (traditional XAI) but also *how* the AI currently models the user's identity to generate that output. For instance, an AI could display, "We recommend this content because your profile is currently heavily weighted toward 'early-adopter tech enthusiast." Additionally, users must have mechanisms to directly review, reject, or modify the core identity traits inferred by the AI. This feature addresses the risk of autonomy erosion by giving users control over the data that shapes their choices, rather than passively accepting an algorithmically defined identity.

Designing for Identity Exploration and Authenticity

The review highlighted the risk of algorithmic homogenization in self-presentation, where optimization for engagement leads to a narrowing of diverse expression. AI design must incorporate features that intentionally counteract this trend and support the user's need for authenticity and growth. Systems should also include optional modes designed to inject novelty and challenge the user's established preferences. For example, a recommendation system could offer a "Serendipity Mode" or an "Explore New Self-Narrative" prompt that intentionally surfaces content or presentation styles outside the user's statistically defined filter bubble. Instead of merely suggesting high-engagement content, AI agents should offer prompts that help users articulate their *actual* values or complex self-narratives. This shifts the AI from optimizing the *performance* of the self to supporting the *articulation* of the self.

Fostering Ownership and Auditability of the Digital Self

Since personalized AI agents actively co-produce the user's digital identity, users require comprehensive rights over the data scaffold that enables this co-production. This is crucial for addressing long-term concerns regarding identity coherence and privacy. Designers should enable users to export and audit the entire digital trace, along with the associated algorithmic model, that their personalized agent has constructed over time. This supports identity ownership by granting the user full control over the complete "extended self" built by the agent.

Users need clear interfaces to pause or limit specific feedback loops. For example, a user should be able to instruct the AI: "Use my professional choices for career recommendations, but *do not* use my entertainment choices to inform my professional self-presentation prompts." This prevents unwanted blending of contexts and reinforces user agency in managing the boundaries of their digital identity. In summary, the design mandate for the next generation of personalized AI agents is to treat the user not as a target to be optimized, but as an agent to be supported. This requires a fundamental commitment to building systems that are transparent about the self they construct, encourage identity exploration, and provide the tools for true ownership of the algorithmically extended self.

Future Research

The framework suggests that future research must analyze personalized AI systems by asking not just what decision was made, but who made the decision (i.e., which version of the statistically generated self was being addressed) and how that decision will be used to shape future digital performances. By integrating these processes, scholars can move beyond simple studies of trust in AI toward investigations of autonomy and authenticity within a fully co-produced digital selfhood.

Conclusion

Personalized artificial intelligence (AI) agents fundamentally redefine the boundary between the human self and

technology. This research demonstrated that the influence of these agents on individuals is complex and deeply interconnected, linking user decision-making, self-presentation, and digital identity management into a single, continuous process. The evaluation confirms that AI agents are not passive tools; they are powerful, adaptive mediators that subtly nudge choices, risk-homogenizing self-presentation for algorithmic approval, and actively construct the "Algorithmic Self" based on continuous behavioral data.

The central contribution of this work is the Integrative Evaluative Framework, which maps the Co-Adaptive Cycle linking these three previously siloed domains. This framework demonstrates that a user's choices, performances, and digital identity are deeply entangled, creating continuous feedback loops where the AI's statistical model of the self influences both what the user sees (in decision-making) and how the user acts (in self-presentation). This entanglement challenges traditional notions of the autonomous user by embedding the AI agent within the very structure of the extended self (Belk, 2013).

Moving forward, these findings necessitate a paradigm shift in both scholarship and design. For researchers, the focus must shift toward longitudinal studies that assess the long-term impact on identity coherence and the ethical dimensions of algorithmic amplification and bias. For designers, the mandate is clear: systems must move beyond optimizing engagement and actively support human agency. This requires a commitment to three core principles: transparent identity modeling, which gives users control over the "statistical individual" the AI constructs; designing for identity exploration, to counteract homogenization; and fostering ownership and auditability of co-produced digital traces.

Ultimately, understanding the impact of personalized AI agents requires recognizing them not merely as sophisticated tools, but as integral, adaptive partners in the ongoing construction and management of identity in the digital age. The future of digital experience hinges on developing systems that treat the user as an agent to be supported, rather than a preference profile to be optimized.

References

Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2013). Do recommender systems manipulate consumer preferences? A study of anchoring effects. *Information Systems Research*, *24*(4), 956–975. https://doi.org/10.1287/isre.2013.0497

Babushkina, D., & Votsis, A. (2021). Disruption, technology, and the question of (artificial) identity. *AI and Ethics*, *2*(4), 611–622. https://doi.org/10.1007/s43681-021-00110-y

Bartoli, C. (2022). Consumer self-concept and digitalization: What does this mean for brands? *Italian Journal of Marketing*, *2022*(4), 419–437. https://doi.org/10.1007/s43039-022-00059-8

Belk, R. W. (2013). Extended self in a digital world. *Journal of Consumer Research*, *40*(3), 477–500. https://doi.org/10.1086/671052

Benke, I., Knierim, M. T., & Maedche, A. (2020). Chatbot-based emotion management for distributed teams. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–30. https://doi.org/10.1145/3415189

Beyari, H., & Hashem, T. (2025). The role of artificial intelligence in personalizing social media marketing strategies for enhanced customer experience. *Behavioral Sciences*, *15*(5), 700. https://doi.org/10.3390/bs15050700

Bjerring, J. C., & Busch, J. (2024). Artificial intelligence and identity: The rise of the statistical individual. *AI & Society*. https://doi.org/10.1007/s00146-024-01877-4

Blunden, H., & Brodsky, A. (2024). A review of virtual impression management behaviors and outcomes. *Journal of Management*, *50*(6), 2197–2236. https://doi.org/10.1177/01492063231225160

Chan, K. T. (2022). Emergence of the 'Digitalized Self' in the age of digitalization. *Computers in Human Behavior Reports*, *6*, 100191. https://doi.org/10.1016/j.chbr.2022.100191

Chaudhry, B. M., & Debi, H. R. (2023). User perceptions and experiences of an AI-driven conversational agent for mental health support. *MHealth*, *0*, 0. https://doi.org/10.21037/mhealth-23-55

Dezfouli, A., Nock, R., & Dayan, P. (2020). Adversarial vulnerabilities of human decision-making. *Proceedings of the National Academy of Sciences*, *117*(46), 29221–29228. https://doi.org/10.1073/pnas.2016921117

Dorčić, T. M., Smojver-Ažić, S., Božić, I., & Malkoč, I. (2023). Effects of Social Media, Social Comparisons, and Identity Processes on Body Image Satisfaction in Late Adolescence. *Europe's Journal of Psychology*, *19*(2), 220–231. https://doi.org/10.5964/ejop.9885

Endacott, C. G. (2021). *The work of identity construction in the age of intelligent machines* [ProQuest ID: Endacott_ucsb_0035D_15147]. UC Santa Barbara. Retrieved from https://escholarship.org/uc/item/2kb6p061

Endacott, C. G., & Leonardi, P. M. (2022). Artificial intelligence and impression management: Consequences of autonomous conversational agents communicating on one's behalf. *Human Communication Research*, *48*(3), 462–490. https://doi.org/10.1093/hcr/hqac009

Fei, X., Ke, W., & Wang, H. (2025). The cross-domain effect of perceived homogenization of social media content on consumer impatience. *Journal of Business Research*, *201*, 115723. https://doi.org/10.1016/j.jbusres.2025.115723

Floridi, L. (2017). *The logic of information: A theory of philosophy as conceptual design*. Oxford University Press.

Ganji, G., & Zarifhonarvar, A. (2025). *Understanding AI agents' decision-making: Evidence from risk and time preference elicitation* (SSRN Scholarly Paper No. 5154002). https://www.google.com/search?q=https://doi.org/10.2139/ssrn.5154002

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity, and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304–316. https://doi.org/10.1016/j.chb.2019.01.020

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.

Jannach, D., & Adomavicius, G. (2016). Recommendations with a purpose. In *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 7–10). ACM. https://doi.org/10.1145/2959100.2959186

Jiang, R., Chiappa, S., Lattimore, T., György, A., & Kohli, P. (2019). Degenerate feedback loops in recommender systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 383–390). ACM. https://doi.org/10.1145/3306618.3314288

Jiang, W., Li, Y., Hu, X., & Ma, D. (2025). Reframing individual roles in collaboration: digital identity construction and adaptive mechanisms for resistance-based professional skills in AI-human intelligence

symbiosis. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1652130

Joseph, J. (2025). The algorithmic self: How AI is reshaping human identity, introspection, and agency. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1645795

Jung, M., Sim, S., Kim, J., & Kim, K. (2022). Impact of personalized avatars and motion synchrony on embodiment and users' subjective experience: Empirical study. *JMIR Serious Games*, *10*(4), e40119. https://doi.org/10.2196/40119

Kang, H., & Lou, C. (2022). AI agency vs. human agency: Understanding human–AI interactions on TikTok and their implications for user engagement. *Journal of Computer-Mediated Communication*, *27*(5). https://doi.org/10.1093/jcmc/zmac014

Kang, H., & Sundar, S. S. (2013). Depleted egos and affirmed selves: The two faces of customization. *Computers in Human Behavior*, *29*(6), 2273–2280. https://doi.org/10.1016/j.chb.2013.05.018

Kantharuban, A., Milbauer, J., Sap, M., Strubell, E., & Neubig, G. (2024). *Stereotype or personalization? User identity biases chatbot recommendations*. arXiv. Retrieved from https://arxiv.org/abs/2410.05613

Kim, J. (2024). Trust, privacy, and personalization in AI-driven recommendations: Impacts on consumer purchase intentions. *Electronic Commerce Research and Applications*, *60*, 101293. https://doi.org/10.1016/j.elerap.2024.101293

Kim, Y., & Song, H. (2022). The Effects of Framing AI Limitations on User Trust: Transparency versus Uncertainty. *Computers in Human Behavior*, *128*, 107118. https://doi.org/10.1016/j.chb.2021.107118

Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, *22*(4–5), 441–504. https://doi.org/10.1007/s11257-011-9118-4

Lagerkvist, A., Tudor, M., Smolicki, J., Ess, C. M., Lundström, J. E., & Rogg, M. (2022). Body stakes: An existential ethics of care in living with biometrics and AI. *AI & Society*, *39*(1), 169–181. https://doi.org/10.1007/s00146-022-01550-8

Li, H., & Unger, T. (2021). AI personalization: From automated to autonomous systems. *Business Horizons*, *64*(6), 749–761. https://doi.org/10.1016/j.bushor.2021.07.004

Li, Y., Zhang, J., Li, X., & Yu, C. (2025). Filter bubbles in short-video platforms: Evidence from large-scale simulations. *Information Processing & Management*, *62*(2), 103283. https://doi.org/10.1016/j.ipm.2025.103283

Mostafa R. B., & Kasamani T. (2022). Antecedents and Consequences of Initial Trust in Chatbots. *European Journal of Marketing*, *56*, 1748-1771.

Nguyen, L. H., Lins, S., Renner, M., & Sunyaev, A. (2024). *Unraveling the nuances of AI accountability: A synthesis of dimensions across disciplines* [Completed Research Paper]. Thirty-Second European Conference on Information Systems (ECIS, 2024). Retrieved from https://arxiv.org/pdf/2410.04247

Ogunsola, O., & Fisher, J. (2025). *Beyond likes—How personality predicts saving behavior on Instagram*. Research Square. https://doi.org/10.21203/rs.3.rs-7666182/v1

Ouhnni, H., Btissam, A., Meryam, B., Bouchti, K. E., Yassine, Z. S., Lagmiri, S. N., Rigalma, B., & Ziti, S. (2025). The evolution of virtual identity: A systematic review of avatar customization technologies and their behavioral effects in VR environments. *Frontiers in Virtual Reality*, *6*. https://doi.org/10.3389/frvir.2025.1496128

Piccialli, F., Chiaro, D., Sarwar, S., Cerciello, D., Qi, P., & Mele, V. (2025). AgentAI: A comprehensive survey on autonomous agents in distributed AI for Industry 4.0. *Expert Systems With Applications*, *291*, 128404. https://doi.org/10.1016/j.eswa.2025.128404

Pronzato, R. (2025). Algorithmic selfing: An existential media analysis of time and identity. *New Media & Society*, *0*(0). https://doi.org/10.1177/14614448251366178

Salwei, M. E., & Carayon, P. (2022). A sociotechnical systems framework for the application of artificial intelligence in health care delivery. *Journal of Cognitive Engineering and Decision Making*, *16*(4), 194–206. https://doi.org/10.1177/15553434221097357

Schmauder, C., Karpus, J., Moll, M., Bahrami, B., & Deroy, O. (2023). Algorithmic nudging: The need for an interdisciplinary oversight. *Topoi*, *42*(3), 799–807. https://doi.org/10.1007/s11245-023-09907-4

Shad, K. B. (2025). AI and socio-techniques. In *Oxford University Press eBooks*. Oxford University Press. https://doi.org/10.1093/9780198945215.003.0105

Shao, Z. (2025). Algorithmic Resistance and Online Privacy: Extending the Meta-UTAUT Model with Particular Privacy Concerns. https://doi.org/10.30560/ijas.v8n1p72

Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284. https://doi.org/10.1016/j.chb.2019.04.019

Sun, Y., Pandita, S., Shaikh, O., Kim, B., Won, A. S., & Cornell University. (2018). *The effect of personalized avatars on self-presence*. Cornell University. Retrieved from https://virtualembodimentlab.com/wp-content/uploads/2019/02/avatar_customization_presence2018_camera_ready.pdf?utm_source=chatgpt.com

Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, *36*(3), 298–322. https://doi.org/10.1111/j.1468-2958.2010.01377.x.

Symeonidis, P., Pechenizkiy, M., & Gemmis, M. D. (2019). Diversity in recommender systems: Bridging accuracy and novelty. *ACM Transactions on Intelligent Systems and Technology*, *10*(3), 1–32. https://doi.org/10.1145/3301290

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Toma, C. L., & Hancock, J. T. (2013). Self-affirmation underlies Facebook use. *Personality and Social Psychology Bulletin*, *39*(3), 321–331. https://doi.org/10.1177/0146167212474694

Turtle G. L., Bendor R., Giaccardi E., & Kotowski B. (2024). *Queering AI: Undoing the self in the algorithmic borderlands*. arXiv. https://www.google.com/search?q=https://doi.org/10.48550/ARXIV.2410.03713

Visser, B., Van der Putten, P., & Zohrehvand, A. (2024). *The Impact of AI Avatar Appearance and Disclosure on User Motivation*. arXiv. Retrieved from https://arxiv.org/html/2407.21521v1

Wang, S., & Wang, W. (2023). A review of the application of digital identity in the metaverse. *Security and Safety*, *2*, 2023009. https://doi.org/10.1051/sands/2023009

Wise, T., Li, A., Bhaumik, S., He, A., Gurevych, I., & Sridhar, R. (2025). Investigating how emerging adults

explore identity through writing: Opportunities for AI writing assistants to help. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference* (pp. 2270–2282). Association for Computing Machinery. https://doi.org/10.1145/3715336.3735848

Wu, Q. (2025). The integration of AI-powered personalization and user-generated content in global digital marketing. In *Proceedings of the 2025 International Conference on Digital Economy and Intelligent Computing* (pp. 168–174). Association for Computing Machinery. https://www.google.com/search?q=https://doi.org/10.1145/3746972.3747000.

Wu, S., Xu, L., Dai, Z., & Pan, Y. (2023). Factors Affecting Avatar Customization Behavior in Virtual Environments. *Electronics*, *12*(10), 2286. https://doi.org/10.3390/electronics12102286

Xu, J., Kim, D., & Sundar, S. S. (2024). Anthropomorphism and compliance: How framing AI systems as social actors shapes user decisions. *Journal of Computer-Mediated Communication*, *29*(1), zmae003. https://doi.org/10.1093/jcmc/zmae003

Yang, X. J., Schemanske, C., & Searle, R. (2021). Dynamic trust in human–automation interaction: The role of feedback and system transparency. *Human Factors*, *63*(7), 1105–1119. https://doi.org/10.1177/0018720820948425

Zhang, L., Bi, W., Zhang, N., & He, L. (2023). Coping with homogeneous information flow in recommender systems: Algorithmic resistance and avoidance. *International Journal of Human-Computer Interaction*, *40*(22), 6899–6912. https://doi.org/10.1080/10447318.2023.2267931

Zhang, Y., Wu, L., & Wang, J. (2023). Counteracting feedback loops in recommender systems with reinforcement learning. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1938–1942). ACM. https://doi.org/10.1145/3539618.3592073

Zhao, S., Kong, Y., Yang, Y., & Li, J. (2024). The Influencing Mechanism of Scenic Spot Online Attention and Tourists' Purchase Behavior: An AISAS Model-Based Investigation. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1386350