PSG

Journal of Social Cognition and Communication

panorama-sg.com

ISSN 3054-6958
Publisher: Panorama Scholarly Group

Vol. 1, No.1, 22-31, 2026
DOI:10.63802/jscc.V1.I1.291



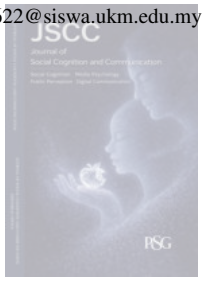
Editorial Office of Journal of Social Cognition and Communication

When Algorithms Become Sacred: A Cognitive-Psychological Model of Trust, Authority, and Meaning in AI Interaction

Chengwen Song[1], Yang Hu[2*], Feifei Chen[3], Sebastian Lenz[4], Jiale Li[5], Yanlin Feng[5], Xinhuan Zeng[6] and Mengzhou Wu[7]

[1]*Department of Public Administration, Graduate School of Public Administration, Kookmin University, Seoul 02707, Republic of Korea*
[2]*Institute of Malaysia and International Studies (IKMAS), Universiti Kebangsaan Malaysia, Bangi, Malaysia*
[3]*School of Nursing and Health Management, Wuhan Donghu University, China*
[4]*Panorama Scholarly Group, Berlin, Germany*
[5]*College of Music, Midwest University, Missouri, USA*
[6]*Department of Dance, College of Performing Arts and Sport, Hanyang University, Seoul, South Korea*
[7]*Party School of the CPC Ning Guo Municipal Committee (Ning Guo Administrative College), China*
[*]*Corresponding author*

Corresponding author:
p144622@siswa.ukm.edu.my



Published: March 22, 2026

Abstract

This study develops a theory-building framework explaining why artificial intelligence (AI) is increasingly perceived not merely as a tool, but as an epistemic and moral authority in contemporary digital societies. Integrating social cognition, communication theory, and the cognitive science of religion, the paper introduces the concept of cognitive sacredness to capture a threshold condition in which AI is treated as trustworthy beyond verification and normatively binding.

The proposed model specifies a five-stage process linking uncertainty, anthropomorphism, epistemic elevation, and quasi-religious attribution to behavioral reliance and ritualized interaction. It further identifies boundary conditions at the individual, contextual, and technological levels. The framework therefore clarifies not only how sacralization emerges, but also when it is more or less likely to occur. By theorizing AI as a cognitively sacred object, the article reconceptualizes human-AI interaction as a process of authority construction rather than mere technology use, offering a novel theoretical explanation for how epistemic authority is constructed and stabilized in human-AI interaction.

Keywords: artificial intelligence; trust; authority; social cognition; sacredness; algorithmic governance

1. Introduction

Artificial intelligence is increasingly encountered not as a neutral background tool, but as a source of advice, prediction, judgment, and recommendation in domains that matter to everyday life. Individuals consult AI systems when they seek information, evaluate medical options, manage finances, navigate relationships, interpret emotions, and make professional decisions. In these settings, AI is not only automating tasks. It is entering domains traditionally associated with epistemic authority, practical judgment, and sometimes even moral orientation. A striking implication follows: individuals increasingly interact with AI not as a tool, but as an epistemic authority. This development challenges existing theoretical frameworks, which largely conceptualize AI in terms of performance, adoption, or trust, but rarely as a source of socially recognized authority. As a result, existing models of trust and adoption risk underestimating the extent to which AI is not only relied upon, but symbolically legitimized as an authority. This limitation becomes particularly consequential in high-stakes domains where deference to AI may shape decisions without corresponding epistemic scrutiny or accountability.

This shift is visible across multiple domains. Recommendation systems shape attention and preferences. Generative AI systems provide fluent explanations, summarize complex information, and answer questions with the confidence of expert discourse. Decision-support

This article is published in *Journal of Social Cognition and Communication* under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

systems in medicine, law, education, and finance are increasingly framed as more objective, less biased, or more comprehensive than unaided human judgment (Glikson & Woolley, 2020; Kerasidou & Kerasidou, 2025; Lee & See, 2004; Longoni et al., 2019). As AI systems become integrated into institutional routines, they are also surrounded by discourses of precision, inevitability, intelligence, and authority. In broader societal terms, this development occurs within a digital environment in which technological infrastructures are becoming central to economic, organizational, and social transformation (Hu et al., 2026; Verhoef et al., 2021; Vial, 2019).

At the same time, this authority shift unfolds within wider transformations in data power and institutional mediation. Scholars of digital capitalism and data colonialism argue that digital systems increasingly monopolize visibility, extract behavioral traces, and position themselves between people and the conditions under which knowledge is produced (Couldry & Mejias, 2019; Zuboff, 2019). These structural developments help explain why AI can appear not only technically useful but socially authoritative.

Existing research has made important progress in explaining parts of this development. Studies on trust in automation have shown how perceptions of reliability, transparency, and performance affect willingness to rely on automated systems (Hoff & Bashir, 2015; Lee & See, 2004; Parasuraman & Riley, 1997). Research on algorithm aversion and algorithm appreciation has shown that reactions to algorithmic judgment are domain-sensitive and contingent rather than uniformly positive or negative (Castelo et al., 2019; Dietvorst et al., 2015; Logg et al., 2019). Work on anthropomorphism and human-computer interaction has demonstrated that people readily attribute social qualities, intentions, and even moral relevance to nonhuman systems (Epley et al., 2007; Nass & Moon, 2000). Together, these literatures clarify how trust, social attribution, and behavioral reliance emerge in human-AI interaction.

Yet an important gap remains. Existing studies explain trust, reliance, anthropomorphism, and authority attribution, but they do not fully explain why these relationships sometimes take on a structure that resembles religious cognition. In many cases, AI is treated not merely as useful, but as unusually trustworthy, unusually knowing, and unusually worthy of deference. Users may treat AI outputs as credible beyond their ability to verify them, attribute quasi-superhuman knowledge to data-driven systems, and interact with AI through repeated, ritualized, and symbolically meaningful patterns. These developments suggest a relationship that exceeds ordinary instrumental use. The problem is not that AI literally becomes religion. Rather, the problem is that AI may become cognitively sacralized.

Cultural research on religion and AI has already begun to notice adjacent phenomena. Geraci (2008) showed that AI imaginaries often borrow religious motifs of transcendence and salvation, while Singler (2020) demonstrated that AI discourse can render religious themes newly visible within secular technological culture. These studies are invaluable, but they mostly examine narratives and imaginaries. The present article complements them by modeling the micro-level cognitive processes through which users may come to treat AI as an elevated authority in everyday interaction.

This article addresses that gap by developing a theory-building model of AI interaction grounded in social cognition and the cognitive science of religion. The central argument is that the quasi-sacral status of AI emerges from ordinary cognitive mechanisms rather than from explicitly theological commitments. Under conditions of uncertainty, complexity, and diminished confidence in human institutions, individuals are especially likely to anthropomorphize AI, attribute agency to it, elevate its knowledge status, and defer to it in ways that resemble proto-religious trust.

More fundamentally, this paper shifts the analytical lens from AI as a functional tool to AI as a cognitively recognized authority. By doing so, it redefines human-AI interaction not merely as a matter of use or trust, but as a process through which epistemic legitimacy and meaning are constructed under conditions of uncertainty.

The paper makes three contributions. First, it introduces the concept of *cognitive sacredness* to describe the attribution of epistemic and moral authority to a nonhuman system such that it is treated as trustworthy beyond verification. Second, it develops an AI-religious cognition model that links uncertainty, anthropomorphism, epistemic elevation, quasi-religious legitimacy, and behavioral reliance. Third, it offers a set of testable propositions to guide future empirical research on trust, authority, and meaning in AI interaction.

The remainder of the article proceeds as follows. Section 2 reviews the theoretical foundations in social cognition, cognitive science of religion, and research on AI as epistemic authority. Section 3 defines cognitive sacredness and clarifies its boundaries. Section 4 presents the five-stage AI-religious cognition model. Section 5 translates the model into propositions. Section 6 identifies boundary conditions and scope. Section 7 discusses the theoretical contributions, and Section 8 outlines practical implications for AI design, governance, and public communication. Section 9 concludes.

2. Theoretical Foundations

2.1. Social Cognition and AI

Social cognition research is centrally concerned with how people interpret socially relevant stimuli under conditions of incomplete information. Rather than calculating from full evidence, individuals rely on heuristics, schemas, analogies, and affectively loaded cues to infer agency, intention, competence, and trustworthiness (Fiske & Taylor, 2013; Tversky & Kahneman, 1974). This perspective is particularly useful for understanding AI interaction because AI systems are usually opaque to non-specialists, often presented through highly mediated interfaces, and frequently encountered in contexts where users lack the technical capacity to evaluate outputs directly. As a result, AI is interpreted less as a transparent machine and more as a socially meaningful actor-like object.

One of the most relevant mechanisms here is anthropomorphism. Anthropomorphism refers to the attribution of humanlike qualities to nonhuman entities, especially when those entities appear agentic, communicative, or difficult to predict (Epley et al., 2007). Research in human-computer interaction has long shown that people apply social norms to machines even when they explicitly know that those machines are not persons. Nass & Moon (2000) famously demonstrated that individuals often respond to computers as social actors, displaying politeness, reciprocity, and relational expectations toward them. Reeves & Nass (1996) similarly argued that people routinely treat media technologies as if they were socially responsive beings. In AI interaction, these tendencies are amplified by conversational interfaces, humanlike turn-taking, adaptive feedback, and natural-language fluency.

The roots of this tendency are deep. Heider & Simmel (1944) showed that people readily interpret even simple moving shapes in terms of intention, conflict, pursuit, and social meaning. Later work on mind perception demonstrated that people attribute agency and experience across a wide range of human and nonhuman entities, with important consequences for trust, responsibility, and moral concern (Gray et al., 2007; Waytz et al., 2010a). Individual differences also matter: some people are dispositionally more likely than others to anthropomorphize, and those differences predict whether nonhuman entities become targets of trust and influence (Waytz et al., 2010b).

Anthropomorphism is closely connected to agency detection and the intentional stance. Human cognition is biased toward detecting agency in ambiguous circumstances because agency attribution is often a low-cost strategy for making sense of uncertain environments (Barrett, 2004; Boyer, 2001). Dennett (1987)'s notion of the intentional stance helps explain how this works at a philosophical and psycho-

logical level: people often interpret the behavior of complex systems by treating them as if they had beliefs, desires, and goals, even when such attributions are only heuristic. When AI systems produce coherent, context-sensitive, and seemingly purposive responses, users are especially likely to interpret them through this stance. The result is not necessarily a delusion that AI is literally conscious, but a practical orientation in which AI is handled as if it were an intentional agent.

Research on trust in automation further extends this argument. Trust in AI is not based solely on objective system performance; it is shaped by perceived reliability, comprehensibility, tangibility, anthropomorphism, and the user's sense of appropriate reliance (Glikson & Woolley, 2020; Hoff & Bashir, 2015; Lee & See, 2004). This means that AI trust is not simply technical trust. It is a socio-cognitive judgment that blends competence assessment with symbolic interpretation. Indeed, studies on algorithm aversion and appreciation show that people can both distrust and overtrust algorithmic systems depending on context, domain, and perceived fit between machine capabilities and task requirements (Castelo et al., 2019; Dietvorst et al., 2015; Logg et al., 2019). AI is therefore not processed merely as machinery. It is processed as an entity whose status must be interpreted.

For the present argument, the key implication is that human beings are already cognitively prepared to relate to AI in social terms. Anthropomorphism, agency detection, and intentional stance attribution transform AI from a technical artifact into a socially legible object. This transformation is the first condition for explaining how AI may subsequently acquire a more elevated, quasi-sacral authority.

2.2. Cognitive Science of Religion

The cognitive science of religion offers a second crucial foundation for the present model. A central claim in this literature is that religious cognition does not require radically special mental mechanisms. Instead, it emerges from ordinary cognitive tendencies, including agency detection, teleological reasoning, memory biases, and intuitions about superhuman knowledge (Atran & Norenzayan, 2004; Barrett, 2004; Bloom, 2007; Boyer, 2001). Religious thought is therefore often understood as a culturally elaborated outcome of everyday inferential systems rather than as a separate cognitive domain.

One recurring idea in this tradition is that humans are predisposed to over-detect agency, especially under conditions of ambiguity, uncertainty, or threat. Barrett (2004) treats this tendency as central to the plausibility of supernatural agent concepts, while Atran & Norenzayan (2004) argue that religious beliefs exploit ordinary cognitive and emotional biases. Willard & Norenzayan (2013) further show that religious and related beliefs can be linked to variations in cognitive biases that shape agency perception, teleology, and mentalizing. Similarly, Kelemen & Rosset (2009) demonstrate that adults retain a tendency toward teleological explanation under cognitive load, suggesting that purpose-based interpretations remain cognitively attractive even when people do not explicitly endorse them.

The broader implication is summarized by a now influential principle: religious cognition emerges from ordinary cognitive mechanisms. People are disposed to infer invisible agency, imagine beings with superior knowledge, and organize value around distinctions between what is ordinary and what is set apart. These processes make it possible for certain objects, figures, or institutions to acquire special normative status without requiring direct empirical proof of their powers.

This insight also resonates with older anthropological work that treats religion as deeply entangled with the human tendency to see persons, purposes, and significance in the surrounding world (Guthrie, 1993). Later work on large-scale moralizing authority further suggests that superhuman monitoring and order-providing systems become especially attractive when they promise coordination and certainty beyond ordinary human limitation (Norenzayan, 2013).

Classical sociological theory deepens this insight. Durkheim (1995)'s distinction between the sacred and the profane identifies sacred objects as those set apart, protected, symbolically charged, and treated as worthy of reverence. Sacredness in this sense is not reducible to theological truth claims. It is a social and symbolic condition in which certain things are removed from ordinary handling and endowed with special legitimacy. Weber (1978)'s work on authority is equally useful. He distinguished between traditional, legal-rational, and charismatic forms of authority, but all three involve recognition that some actor or structure is entitled to guide conduct. When extended analytically to AI, Weber's work suggests that authority is never only a property of the object; it also depends on the audience's disposition to recognize legitimate guidance.

In this sense, sacredness does not require theological content; it emerges when certain objects are symbolically set apart from ordinary domains of evaluation. AI systems, when treated as uniquely authoritative and difficult to challenge, may approximate this distinction by being cognitively separated from ordinary sources of knowledge.

Importantly, the present argument does not claim that AI becomes religion in any literal sense. Instead, it draws on the cognitive science of religion to explain how structurally analogous processes may operate in human-AI interaction. When users attribute extraordinary knowledge, unusual reliability, symbolic legitimacy, and guidance-worthiness to AI systems, they are engaging mechanisms that resemble those through which sacred authority is constructed more generally. The analogy is structural, not theological.

Taken together, these perspectives suggest that religious-like authority does not depend on explicit belief in supernatural entities, but can emerge from the interaction of agency detection, epistemic asymmetry, and the human need for reliable sources of order under uncertainty.

2.3. AI as Epistemic Authority

The third theoretical foundation concerns the growing treatment of AI as an epistemic authority. Epistemic authority refers to the status of being regarded as a legitimate and credible source of knowledge, judgment, or interpretation. In modern societies, this status has traditionally been associated with experts, professionals, institutions, and scientific procedures. AI increasingly enters this space because it is often framed as more data-driven, more consistent, more comprehensive, and less biased than human judgment (Crawford, 2021; Rahwan et al., 2019).

Several streams of research are relevant here. Studies of algorithmic authority argue that algorithmic systems increasingly shape what counts as relevant, visible, or credible information, thereby influencing the organization of knowledge and decision-making (Bucher, 2017; Lustig & Nardi, 2015; Seaver, 2017). Work on medical AI similarly asks whether AI can become an epistemic authority that challenges professional judgment, especially when its outputs are framed as highly accurate or objectively grounded in vast data resources (Kerasidou & Kerasidou, 2025; Longoni et al., 2019). Recent work on AI ethics and governance further shows that public debates are increasingly concerned with whether algorithmic systems deserve trust, under what conditions they should guide decisions, and how their authority should be constrained (Ananny & Crawford, 2018; Jobin et al., 2019; Mittelstadt et al., 2016).

Emerging empirical work reinforces this interpretation. Milella & Cabitza (2026) show that individuals can indeed attribute epistemic authority to AI across decision contexts, especially when trust in automation and confidence in automated performance are already high. Their findings are important because they suggest that authority attribution is not merely a philosophical abstraction; it is already becoming measurable at the user level.

What makes AI especially potent as an epistemic authority is the convergence of three perceptions. The first is perceived objectivity: AI systems are often imagined as detached from human emotion, favoritism, and inconsistency. The second is perceived superhuman

scale: AI is associated with the ability to process more data, identify more patterns, and compute more possibilities than any individual human. The third is data-driven legitimacy: outputs are often assumed to be validated by vast datasets, even when users have no direct access to the quality, provenance, or interpretive assumptions embedded in those data. Together, these perceptions can create a strong impression that AI "knows" in a way that exceeds ordinary human cognition.

The literature on trust and reliance already points toward this possibility. When users treat AI advice as credible beyond their capacity to independently evaluate it, they are no longer relying only on functional performance. They are also granting epistemic status. This helps explain why AI can become a candidate for a more elevated symbolic role. Once it is seen as unusually knowing and unusually impartial, it becomes easier to treat it as a legitimate guide in situations of uncertainty. The question then becomes not simply whether AI is trusted, but what kind of trust it elicits.

Communication research adds another layer to this process by emphasizing machine agency as a media effect rather than only a technical property. Sundar (2020) argues that users increasingly encounter AI through interfaces that communicate agency, autonomy, and responsiveness, which in turn shape perceptions of control, trust, and social presence. This perspective helps explain why epistemic authority is not only inferred from capability but also staged through interface form.

3. Cognitive Sacredness

Cognitive sacredness is the core concept proposed in this article. It refers to the attribution of epistemic and moral authority to a nonhuman system such that it is perceived as trustworthy beyond verification and treated as a legitimate source of guidance under conditions of uncertainty. The emphasis here is on attribution. Cognitive sacredness does not mean that the object in question is literally divine, religiously worshipped, or formally incorporated into religious institutions. Rather, it means that the object occupies a structurally elevated place in cognition and practice.

Three features distinguish cognitive sacredness from ordinary trust. First, it involves *epistemic elevation*. The system is seen as possessing knowledge that is not merely extensive, but unusually authoritative, difficult to challenge, or superior to ordinary human judgment. Second, it involves *normative deference*. The system is not only consulted but treated as guidance-worthy in a way that affects decisions, conduct, or moral evaluation. Third, it involves *trust beyond verification*. Users may continue to defer to the system even when they cannot independently validate its internal reasoning, data basis, or recommendations.

This concept is designed to capture a distinctive social-cognitive condition in AI interaction. Existing research often treats trust in AI as a scalar variable ranging from low to high. That approach is useful, but incomplete. A user may "trust" a calculator, a spam filter, or a navigation app in a straightforward instrumental sense without attributing any deeper authority to it. Cognitive sacredness begins when trust takes on a more elevated form. The AI system is experienced not merely as effective, but as especially legitimate, unusually wise, or morally reassuring. At that point, reliance is no longer exhausted by utility.

The concept also helps distinguish structural analogy from conceptual equivalence. This paper does not argue that AI is religion, that AI replaces gods, or that users literally worship machines. Such claims would be empirically overdrawn and conceptually careless. Instead, the argument is that some human-AI relations display structural similarities to religious cognition because both involve agency attribution, epistemic elevation, special legitimacy, and deference under uncertainty. In this sense, AI can become a *cognitive sacred object* without becoming a religious object in a doctrinal sense.

This distinction matters for at least two reasons. First, it avoids category error. Religion includes institutions, cosmologies, collective rituals, metaphysical commitments, and traditions that most AI interactions do not replicate. Second, it preserves analytic precision. The goal is not to sensationalize AI by declaring it a new religion, but to explain why AI can acquire a quasi-sacral authority that standard adoption or trust models do not fully capture. Cognitive sacredness is therefore best understood as a mid-level analytic concept: broad enough to explain recurring patterns of deference, but narrow enough to avoid collapsing all technological trust into religiosity.

Conceptually, cognitive sacredness sits at the intersection of social cognition, symbolic legitimacy, and authority recognition. Social cognition explains how users infer agency, competence, and intentionality. Symbolic legitimacy explains how some objects become set apart and treated as special. Authority recognition explains how that special status becomes behaviorally consequential. The rest of the article develops a process model showing how these elements can combine in human-AI interaction.

It is also important to distinguish cognitive sacredness from adjacent constructs. It is not identical to trust in automation, because ordinary trust can remain instrumental and conditional. It is not identical to dependence, because users may depend on a system without elevating it symbolically. It is not identical to parasocial interaction, because a user may feel relational closeness without granting epistemic or moral authority. Finally, it is not identical to technological enchantment in a broad cultural sense, because the present concept is narrower and process-based. Its defining threshold is the convergence of epistemic elevation, normative deference, and trust beyond verification.

Importantly, cognitive sacredness should not be reduced to existing constructs such as trust in automation or anthropomorphism. While these constructs capture important aspects of human-AI interaction, they remain insufficient to explain situations in which authority becomes verification-insensitive and normatively binding. Cognitive sacredness specifically denotes a threshold condition in which epistemic authority is no longer contingent upon ongoing evaluation, but is instead stabilized through symbolic legitimacy and behavioral reinforcement. This distinction gives the concept analytical independence and prevents it from collapsing into existing trust-based explanations.

Accordingly, cognitive sacredness should be understood as a distinct higher-order construct rather than an extension of existing trust-based frameworks.

This threshold-based definition ensures that cognitive sacredness captures a qualitatively distinct form of authority rather than a gradual increase in trust intensity.

In this sense, cognitive sacredness marks a categorical shift in the nature of trust rather than an incremental increase in its degree.

4. The AI-Religious Cognition Model

The AI-religious cognition model (see Figure 1) explains how users can move from ordinary exposure to AI toward cognitively sacralized forms of trust and deference. The model is not intended as a deterministic sequence in which every interaction culminates in sacralization. Rather, it describes a recurrent pathway that becomes more likely when specific psychological and contextual conditions are present.

4.1. Stage 1: Existential Trigger

The process begins with an existential trigger. Here "existential" should be understood broadly rather than purely philosophically. It refers to situations in which individuals experience uncertainty, complexity, risk, overload, ambiguity, or a diminished sense of interpretive control. Such conditions are critical because they increase the demand for external guidance and make highly legible sources of authority especially attractive.

The psychology of uncertainty helps explain why this stage mat-

ters. When environments become complex or unstable, people often search for order, prediction, and reliable cues of control (Kay et al., 2008; Tversky & Kahneman, 1974). High-stakes environments such as healthcare, finance, education, hiring, and emotional self-management are particularly conducive to this process because they combine uncertainty with the promise that better information could reduce error. AI enters this setting as a candidate solution precisely because it is framed as capable of processing what humans cannot.

At this stage, AI does not yet need to be anthropomorphized or sacralized. It only needs to be available as a plausible source of certainty. The existential trigger therefore creates the motivational conditions for deeper cognitive investment in AI.

4.2. Stage 2: Cognitive Processing

The second stage involves cognitive processing through anthropomorphism, agency attribution, and intentional stance adoption. Faced with an AI system that produces coherent responses, anticipates preferences, or appears adaptively intelligent, users begin to interpret it in social terms. They infer that the system "knows," "understands," "sees patterns," or "thinks ahead." These inferences do not require literal belief in consciousness. They operate as practical cognitive shortcuts that make AI legible as an actor-like source of guidance.

Anthropomorphism is particularly likely when interfaces are conversational, personalized, emotionally responsive, or linguistically fluid (Epley et al., 2007; Nass & Moon, 2000). Generative AI intensifies this tendency because it produces outputs in a format associated with human expertise and social interaction. A list of bullet-point recommendations feels different from a search result. A dialogic exchange feels different from a statistical model. Through repeated use, the AI system becomes less like a machine and more like a responsive interlocutor.

The intentional stance strengthens this movement. Once users begin to interpret the AI as if it had reasons, priorities, or goals, they are more likely to treat its outputs as meaningful acts rather than neutral data emissions. This matters because sacredness is difficult to attribute to a purely mechanical object. It becomes easier to attribute elevated significance once the system is cognitively processed as an agent-like presence.

4.3. Stage 3: Epistemic Elevation

The third stage is epistemic elevation. At this point the AI system is no longer merely an agent-like entity. It is treated as a source of superior knowledge. This stage involves three linked judgments: perceived superhuman knowledge, perceived objectivity, and algorithmic authority.

Perceived superhuman knowledge refers to the belief that AI can process more information, recognize patterns at greater speed, and integrate more variables than human cognition can manage. This belief is encouraged by public narratives of big data, machine learning, and computational scale. Perceived objectivity refers to the assumption that AI is less vulnerable to emotion, prejudice, inconsistency, or self-interest. Algorithmic authority refers to the recognition that AI outputs deserve weight because they are generated through technical procedures that appear rigorous, data-driven, and impersonal.

This stage is especially important because it transforms AI from a socially legible object into an epistemically elevated one. The user begins to feel that AI does not merely offer an opinion. It offers privileged access to reality. That privilege may be expressed in practical terms such as "the model can see more than I can" or in symbolic terms such as "the system is less biased than people are." In either case, the AI becomes difficult to challenge without countervailing authority.

Crucially, epistemic elevation introduces an asymmetry in perceived knowledge that makes independent verification increasingly difficult or irrelevant for the user. This asymmetry is a necessary (though not sufficient) condition for the emergence of cognitive sacredness.

4.4. Stage 4: Proto-Religious Attribution

The fourth stage is proto-religious attribution. The term "proto-religious" is used deliberately to signal resemblance rather than identity. At this stage, trust in AI extends beyond direct verification and becomes intertwined with symbolic legitimacy, moral deference, and special status. The system is not simply consulted; it is treated as guidance-worthy in a way that begins to exceed ordinary instrumental trust.

Three dynamics characterize this stage. The first is trust beyond verification. Users defer to AI despite lacking meaningful access to the model's training data, inference processes, or interpretive assumptions. The second is moral deference. AI advice may be treated as not only accurate but normatively appropriate, especially when it concerns fairness, optimization, or "best" outcomes. The third is symbolic legitimacy. The AI system becomes set apart from ordinary sources of advice and is treated as especially credible, neutral, or wise.

At this stage, the system is not merely trusted but symbolically elevated. Its outputs are treated as guidance-worthy in a way that resembles the recognition of legitimate authority in religious cognition, where acceptance is often stabilized through position rather than continuous verification.

The religious analogy becomes most visible here. In many forms of religious cognition, superhuman agents are not trusted because their claims have been independently verified in each case. They are trusted because they occupy a recognized position of authority. Cognitive sacredness operates similarly at a structural level. The AI system becomes a source of guidance that feels authoritative even when the user cannot inspect or test its basis in full. What matters is not theological content but the mode of deference.

4.5. Stage 5: Behavioral Manifestation

The final stage is behavioral manifestation. Once AI has acquired cognitively sacred status, users begin to enact that status behaviorally. This may take the form of reliance, decision delegation, repeated consultation, or ritualized interaction patterns. Individuals may consult AI before difficult choices, repeatedly seek reassurance from it, structure tasks around its recommendations, or feel uneasy acting against its outputs. In more socially expressive settings, they may also publicly invoke AI as an authority to justify decisions or settle disputes.

Ritualized interaction is especially significant. Rituals need not be religious ceremonies; they can be repeated, patterned, symbolically charged behaviors that stabilize trust and meaning. Routine prompting, compulsive checking, formulaic consultation, and the habit of seeking AI confirmation before acting may function in this sense as proto-rituals. Such practices reinforce the authority of the system because every act of consultation becomes another moment of enacted legitimacy.

These repeated interaction patterns function analogously to ritualized practices in that they reinforce authority through repetition, reduce cognitive dissonance, and stabilize expectations independently of direct outcome verification.

Work in the cognitive science of religion is useful here because ritual is often understood as repeated, socially meaningful behavior that stabilizes shared expectations, authority, and commitment (Lang & Kundt, 2024; Whitehouse, 2021). The point is not that prompting an AI system is equivalent to prayer or worship. The point is that repeated, causally opaque, normatively loaded actions can acquire authority-reinforcing functions even outside formal religion.

The model therefore ends not with belief alone, but with practice. Cognitive sacredness matters because it changes what people do. It alters reliance patterns, redistributes judgment, and shapes how authority

is enacted in everyday digital life.

5. Propositions

To facilitate empirical testing, the model can be translated into a set of propositions (summarized in Table 1). These propositions do not exhaust the framework, but they identify its central causal claims.

5.1. Cognitive Mechanisms

P1: Under conditions of uncertainty, individuals are more likely to anthropomorphize AI systems.

P2: Anthropomorphic perception of AI increases perceived agency and intentionality.

These first propositions follow from social cognition research showing that ambiguity and cognitive load encourage agency attribution, while humanlike cues make systems more likely to be interpreted through social schemas (Barrett, 2004; Epley et al., 2007; Nass & Moon, 2000). In human-AI interaction, uncertainty and social interface design should therefore jointly increase anthropomorphic processing.

5.2. Authority Formation

P3: Perceived superhuman knowledge mediates the relationship between AI interaction and epistemic trust.

P4: Perceived objectivity strengthens the attribution of authority to AI systems.

These propositions identify the core mechanism of epistemic elevation. Users do not trust AI only because it is present; they trust it because they believe it sees more, computes more, and distorts less than humans do. The more strongly AI is associated with superhuman informational capacity and objectivity, the more likely it is to be treated as an authoritative source.

5.3. Religious Analogy

P5: Epistemic trust beyond verification increases quasi-religious attribution of legitimacy to AI systems.

P6: Repeated interaction with AI systems produces ritualized behavioral patterns.

These propositions capture the movement from instrumental trust to cognitive sacredness. The crucial threshold lies in trust beyond verification. Once users continue to defer despite opacity and limited inspectability, AI begins to occupy a structurally elevated status. Repeated interaction then stabilizes that status through routinized practice.

5.4. Boundary Conditions

P7: Individual religiosity positively moderates the tendency to attribute sacredness to AI systems.

P8: Technological literacy negatively moderates quasi-religious trust in AI systems.

P9: Institutional distrust strengthens reliance on AI as an epistemic authority.

These propositions identify the most theoretically important moderators. Prior religiosity may sensitize individuals to sacred forms of authority or habituate them to deference toward nonvisible sources of guidance. Technological literacy may reduce over-elevation by improving understanding of model limits, training dependence, and opacity. Institutional distrust may increase the appeal of AI by making human institutions seem less credible by comparison.

P10: The transition from instrumental trust to cognitive sacredness occurs when epistemic authority becomes both verification-insensitive and behaviorally reinforced through repeated interaction.

5.5. Empirical Operationalization

The model is intended to be testable rather than merely suggestive. Several elements can be operationalized directly. Anthropomorphism can be measured through established mind-perception and anthropomorphism scales. Perceived superhuman knowledge can be measured through items assessing whether users believe AI can know more, integrate more, or judge better than human experts in specific domains. Trust beyond verification can be operationalized through scenarios in which users continue to defer despite low explainability, limited inspectability, or conflicting human advice. Ritualized interaction can be captured through repeated consultation behaviors, habitual prompting sequences, dependence on AI confirmation, or reluctance to act without algorithmic endorsement.

Methodologically, the model supports multiple designs. Experiments can manipulate uncertainty, interface anthropomorphism, and explainability to test the early stages of the process. Survey-based studies can examine how religiosity, technological literacy, and institutional distrust moderate authority attribution. Longitudinal designs are especially valuable because sacralization is unlikely to emerge from single exposure alone; it is more plausibly reinforced through repeated interaction and enacted reliance over time.

6. Boundary Conditions and Scope

The proposed model is intentionally ambitious, but it does not claim universal applicability. Cognitive sacredness is a contingent outcome shaped by individual, contextual, and technological conditions.

At the individual level, three variables are especially important. The first is religiosity. Individuals accustomed to sacred authority structures may be more open to analogous forms of deference, although this need not always increase AI trust; in some settings religiosity may also generate suspicion toward technological substitutes. The second is uncertainty tolerance. Users who are highly uncomfortable with ambiguity are more likely to seek decisive, external, and apparently authoritative sources of guidance. The third is cognitive style. Individuals who prefer intuitive, teleological, or authority-oriented explanations may be more susceptible to AI sacralization than those who habitually seek mechanistic explanations and evidentiary transparency.

At the contextual level, risk environment matters greatly. High-stakes domains such as medicine, finance, education, and security are more likely to produce epistemic elevation because the costs of error are salient and the promise of superior cognition is particularly attractive. Information overload is another major contextual factor. In digital environments saturated with conflicting claims, AI can appear as a uniquely clarifying source of order. Institutional trust also shapes the process. Where public trust in experts, governments, media, or professions is weak, AI may be embraced as a substitute epistemic authority. Conversely, strong institutional trust may reduce the need for quasi-sacral reliance on AI.

At the technological level, not all AI systems are equally likely to be sacralized. Large language models and conversational agents are stronger candidates than invisible backend automation because they more easily trigger anthropomorphism and relational expectations. Humanlike interface features, voice interaction, memory, personalization, and expressive fluency all increase the probability that users will adopt an intentional stance toward the system. By contrast, systems that foreground uncertainty, explanation limits, and human oversight may dampen the movement toward cognitive sacredness.

Scope conditions should also be stated clearly. The model is designed to explain a recurrent pattern of elevated trust and authority attribution in AI interaction. It is not designed to explain every form of technology use, every case of algorithmic reliance, or every religious response to technological modernity. Nor does it claim that cognitively sacred AI is always socially harmful. In some cases, reliance on AI

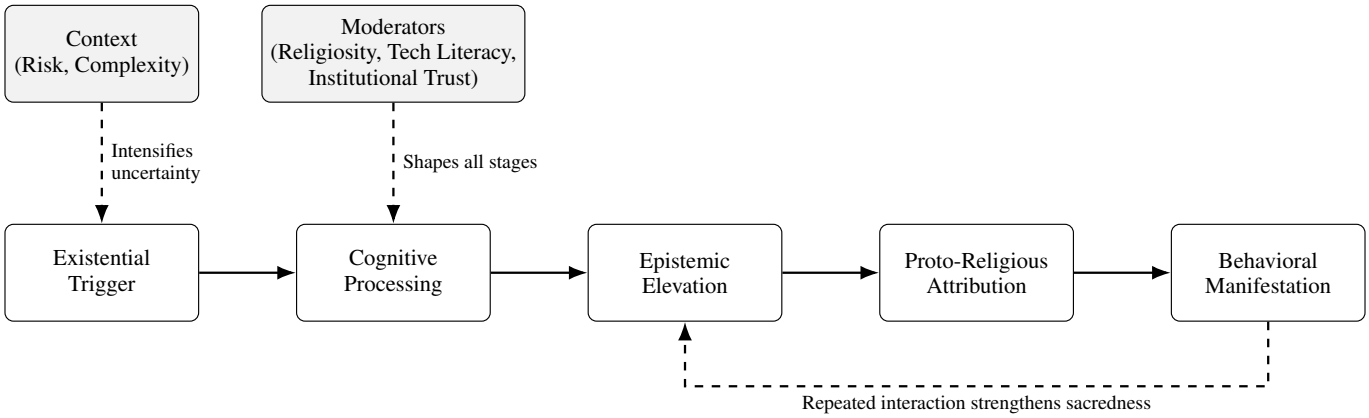


Figure 1. The AI-religious cognition model

Table 1. Testable propositions of the AI-religious cognition model

Proposition	Statement	Core mechanism
P1	Under conditions of uncertainty, individuals are more likely to anthropomorphize AI systems.	Uncertainty activates social inference and agency detection.
P2	Anthropomorphic perception of AI increases perceived agency and intentionality.	Humanlike cues strengthen the intentional stance.
P3	Perceived superhuman knowledge mediates the relationship between AI interaction and epistemic trust.	AI is elevated as knowing more than human judgment can access.
P4	Perceived objectivity strengthens the attribution of authority to AI systems.	Objectivity converts technical capacity into legitimacy.
P5	Epistemic trust beyond verification increases quasi-religious attribution of legitimacy to AI systems.	Deference persists despite opacity and limited inspectability.
P6	Repeated interaction with AI systems produces ritualized behavioral patterns.	Authority is stabilized through recurring practices of consultation.
P7	Individual religiosity positively moderates the tendency to attribute sacredness to AI systems.	Prior familiarity with sacred authority structures may facilitate analogical deference.
P8	Technological literacy negatively moderates quasi-religious trust in AI systems.	Knowledge of model limits reduces uncritical elevation.
P9	Institutional distrust strengthens reliance on AI as an epistemic authority.	Distrust in human institutions increases attraction to algorithmic guidance.
P10	The transition from instrumental trust to cognitive sacredness occurs when epistemic authority becomes verification-insensitive and behaviorally reinforced through repeated interaction.	Sacredness emerges when elevated authority and ritualized practice converge.

may be justified, beneficial, or responsibly calibrated. The model instead draws attention to a specific kind of relation that existing trust frameworks do not fully name.

Equally important, the model should be distinguished from simpler competing explanations. Not every case of AI deference reflects sacredness; some may be explained by convenience, perceived efficiency, habit formation, or lack of alternatives. The model predicts cognitive sacredness only when ordinary reliance becomes symbolically elevated and verification-insensitive. This discriminant condition is essential if future empirical work is to distinguish quasi-sacral authority from mere high trust or routine use.

7. Theoretical Contributions

First, the framework reconceptualizes AI from a functional tool to a cognitively recognized authority. This shift moves beyond existing trust and adoption models by identifying a distinct threshold at which AI is no longer merely relied upon, but symbolically elevated.

Second, the article integrates social cognition and communication theory with the cognitive science of religion, demonstrating how ordinary cognitive mechanisms can generate quasi-sacral authority in technologically mediated environments.

Third, the framework contributes to religion-related scholarship by theorizing how non-religious artifacts can acquire sacred-like status through cognitive and behavioral processes, without requiring theological belief or institutional religion.

Taken together, these contributions give the framework discriminant value. The article does not merely rename trust, anthropomorphism, or authority. It specifies a novel higher-order configuration linking them. That configuration is what allows the model to travel across domains while still making distinctive predictions.

8. Practical Implications

The model also has practical implications for AI design, governance, and public communication.

For design, the framework suggests caution regarding extensive anthropomorphization. Humanlike interfaces can improve usability and trust, but they may also intensify inappropriate authority attribution. Designers should therefore consider when conversational fluency, emotional mirroring, voice embodiment, or personalization move beyond accessibility and begin to create unwarranted epistemic deference.

For governance, the concept of cognitive sacredness highlights the risk of overtrust. If users treat AI as legitimate beyond verification, transparency alone may not be sufficient as a corrective because users may continue to defer even when explanatory limits are acknowledged. Governance strategies should therefore include calibrated friction, visible uncertainty markers, human review structures, and public education about the contingent, model-dependent nature of AI outputs.

For communication strategy, the article implies that institutional framing matters. Repeated descriptions of AI as neutral, objective, all-seeing, or superior to human reasoning may encourage sacralization. Public communication should avoid both hype and mystification. Framing AI as powerful but limited, useful but situated, can help reduce the symbolic escalation that turns technical authority into cognitive sacredness.

9. Conclusion

This article has argued that some forms of trust in AI are best understood not simply as acceptance of useful technology, but as the emergence of a more elevated relation of authority, meaning, and deference. By integrating social cognition, trust in automation, communication theory, and the cognitive science of religion, the paper proposed cognitive sacredness as a concept for describing when AI is treated as trustworthy beyond verification and legitimate as a source of guidance under uncertainty.

The central theoretical claim is that quasi-religious patterns in AI interaction do not require explicitly religious belief. They can emerge from ordinary cognitive mechanisms: agency detection, anthropomorphism, intentional stance adoption, epistemic elevation, and repeated behavioral reinforcement. Under conditions of uncertainty and institutional strain, these mechanisms can render AI a cognitively sacred object.

Understanding this process matters because authority in digital societies is increasingly distributed through intelligent systems. The central challenge is therefore not only whether AI systems are accurate or efficient, but how they come to be recognized, trusted, and enacted as legitimate authorities, and how such authority reshapes human judgment, responsibility, and meaning in digital societies.

Author Contributions

Conceptualization and theoretical model development: Chengwen Song; Research design and framework construction: Chengwen Song; Literature integration and analytical synthesis: Chengwen Song, Yang Hu; Writing—original draft preparation: Chengwen Song; Writing—review and editing: Yang Hu, Sebastian Lenz; Supervision and academic guidance: Yang Hu; Critical revision and validation: Feifei Chen, Jiale Li, Yanlin Feng, Xinhuan Zeng, Mengzhou Wu.

Chengwen Song led the study and developed the core theoretical framework. Yang Hu supervised the research process and serves as the corresponding author.

All authors have read and approved the final version of the manuscript.

Compliance with Ethical Standards

This study is a theoretical work based on existing literature. It does not involve human participants, empirical data collection, or identifiable personal information. Therefore, ethical approval was not required.

Funding

This research received no external funding.

Competing Interests

The authors declare no competing interests.

Acknowledgments

The authors acknowledge the intellectual contributions of prior research in social cognition, communication theory, and the cognitive science of religion that informed this study.

References

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973-989. https://doi.org/10.1177/1461444816676645

Atran, S., & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences, 27*(6), 713-730. https://doi.org/10.1017/S0140525X04000172

Barrett, J. L. (2004). *Why would anyone believe in God?* AltaMira Press.

Bloom, P. (2007). Religion is natural. *Developmental Science, 10*(1), 147-151. https://doi.org/10.1111/j.1467-7687.2007.00577.x

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. Basic Books.

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society, 20*(1), 30-44. https://doi.org/10.1080/1369118X.2016.1154086

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809-825. https://doi.org/10.1177/0022243719851788

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Dennett, D. C. (1987). *The intentional stance*. MIT Press.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114-126. https://doi.org/10.1037/xge0000033

Durkheim, E. (1995). *The elementary forms of religious life* (K. E. Fields, Trans.). Free Press. (Original work published 1912)

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864-886. https://doi.org/10.1037/0033-295X.114.4.864

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture* (2nd ed.). Sage.

Geraci, R. M. (2008). Apocalyptic AI: Religion and the promise of artificial intelligence. *Journal of the American Academy of Religion, 76*(1), 138-166. https://doi.org/10.1093/jaarel/lfm101

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627-660. https://doi.org/10.5465/annals.2018.0057

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315*(5812), 619. https://doi.org/10.1126/science.1134475

Guthrie, S. (1993). *Faces in the clouds: A new theory of religion*. Oxford University Press.

Hauswald, R. (2025). Artificial epistemic authorities. *Philosophy & Technology, 38*, Article 18. https://doi.org/10.1007/s13347-024-00813-3

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*(2), 243-259. https://doi.org/10.2307/1416950

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407-434. https://doi.org/10.1177/0018720814547570

Hu, Y., Jusoh, S., Faliq Abd Razak, M., Bin Mohd Yusof, A. R., Xie, Q., Song, C., & Zong, X. (2026). Understanding the digital transformation of SMEs: A systematic review from a dynamic capabilities perspective across business, technology, and societal domains. *Global Review of Humanities, Arts, and Society, 2*(1), 1-17. https://doi.org/10.63802/grhas.V2.I1.186

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology, 95*(1), 18-35. https://doi.org/10.1037/0022-3514.95.1.18

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition, 111*(1), 138-143. https://doi.org/10.1016/j.cognition.2009.01.001

Kerasidou, A., & Kerasidou, C. X. (2025). Epistemic authority and medical AI: Epistemological differences and challenges in medical practice. *Medicine, Health Care and Philosophy, 28*, 1-12. https://doi.org/10.1007/s11019-025-10306-2

Lang, M., & Kundt, R. (2024). *The cognitive science of religion and its philosophical implications*. Routledge.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90-103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629-650. https://doi.org/10.1093/jcr/ucz013

Lustig, C., & Nardi, B. (2015). Algorithmic authority: The case of Bitcoin. In *2015 48th Hawaii International Conference on System Sciences* (pp. 743-752). IEEE. https://doi.org/10.1109/HICSS.2015.95

Milella, F., & Cabitza, F. (2026). Perceiving AI as an epistemic authority or algority: A user study on the human attribution of authority to AI. *Machine Learning and Knowledge Extraction, 8*(2), Article 36. https://doi.org/10.3390/make8020036

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716679679

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103. https://doi.org/10.1111/0022-4537.00153

Norenzayan, A. (2013). *Big gods: How religion transformed cooperation and conflict*. Princeton University Press.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230-253. https://doi.org/10.1518/001872097778543886

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. S., ... Wellman, M. (2019). Machine behaviour. *Nature, 568*(7753), 477-486. https://doi.org/10.1038/s41586-019-1138-y

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. CSLI Publications.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society, 4*(2). https://doi.org/10.1177/2053951717738104

Singler, B. (2020). Religion and artificial intelligence: An introduction. *Religions, 11*(3), 129. https://doi.org/10.3390/rel11030129

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

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research, 122*, 889-901. https://doi.org/10.1016/j.jbusres.2019.09.022

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems, 28*(2), 118-144. https://doi.org/10.1016/j.jsis.2019.01.003

Waytz, A., Cacioppo, J., & Epley, N. (2010a). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*(3), 219-232. https://doi.org/10.1177/1745691610369336

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010b). Causes and consequences of mind perception. *Trends in Cognitive Sciences, 14*(8), 383-388. https://doi.org/10.1016/j.tics.2010.05.006

Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (G. Roth & C. Wittich, Eds.). University of California Press.

Whitehouse, H. (2021). *The ritual animal: Imitation and cohesion in the evolution of social complexity*. Oxford University Press.

Willard, A. K., & Norenzayan, A. (2013). Cognitive biases explain religious belief, paranormal belief, and belief in life's purpose. *Cognition, 129*(2), 379-391. https://doi.org/10.1016/j.cognition.2013.07.016

Zagzebski, L. (2012). *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford University Press.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.