



Towards a typology of epistemic relationships in human-AI interaction

Shengnan Yang and Rongqian Ma

DOI: <https://doi.org/10.47989/ir31iConf64143>

Abstract

Introduction. This study investigates how academics negotiate epistemic relationships (ERs) with artificial intelligence (AI) in research and teaching. We conceptualize ERs as the positioning of humans and AI in relation to knowledge, asking: What epistemic relationships emerge in researchers' interactions with AI, and how are they enacted across academic knowledge practices?

Method. Semi-structured interviews were conducted in two independently collected datasets with 31 academics across regions and disciplines.

Analysis. We developed a five-dimensional codebook including three epistemic attributes (assessment perspective, trust type, human epistemic status) and two analytic devices (activities and metaphors). We refined the coding through iterative analysis and identified recurring patterns across participants' accounts in both datasets.

Results. Five recognizable ER types emerged: Epistemic Abstention, Instrumental Reliance, Contingent Delegation, Co-agency Collaboration, and Authority Displacement. These types reflect variations in how participants assess and trust AI, position themselves epistemically, and embed AI in academic activities.

Conclusion(s). Our study contributes to conversations on information literacies and the ethical use of AI. We highlight the contingent and practice-dependent nature of ERs, suggesting a shift beyond static metaphors of AI toward a more nuanced account that captures the epistemic dimensions of human-AI knowledge co-construction.

Introduction

Artificial intelligence (AI) systems have made remarkable advances in recent years, often outperforming humans in complex cognitive tasks such as reasoning, diagnosis, and decision-making. These developments have sparked debate over their implications for human agency, often framed as a tension between technological determinism and human-centered views. Some warn that over-reliance on AI may erode human epistemic agency—the capacity to actively participate in knowledge creation. Others, such as Felin and Holweg (2024), emphasize the fundamental difference between AI's data-driven prediction and human theorizing. Meanwhile, practice-oriented scholars in HCI and management highlight the benefits of human-AI interaction, framing AI as a tool for automating routine work or augmenting human decision-making (e.g., Heer 2019; Taudien et al., 2022). Research on hybrid agency and human-AI ensembles has shown that combining human and algorithmic judgments can enhance performance in certain contexts, such as transportation (Kahr et al, 2025), management (Csaszar et al., 2024), and health (Choudhary et al., 2025). These discussions remain largely focused on functional capabilities, cognitive mechanisms, and task outcomes.

Yet across these perspectives, the epistemic framing of human-AI interaction (HAI) and collaboration remains underdeveloped, particularly in HCI research. As AI systems are increasingly embedded in knowledge-intensive domains, ranging from information retrieval and content generation to scientific modeling and decision support, little attention has been paid to how humans and AI co-produce knowledge or how AI reshapes users' roles as knowers. In contrast, recent philosophical scholarship has extended concerns about epistemic injustice and asymmetry, traditionally explored in human-to-human contexts, into the domain of human-nonhuman relations. AI is increasingly theorized not merely as a cognitive tool supporting mental processes, but as an epistemic technology that contributes to knowledge acquisition and, under certain conditions, may even function as an epistemic authority (Ferrario et al., 2024; Hauswald, 2025). These accounts challenge traditional human-centered assumptions and propose an alternative lens for conceptualizing human-AI interaction as a form of epistemic co-agency or negotiated collaboration. Such views raise new questions about whether epistemic relationships can form between humans and AI, even when AI lacks beliefs, intentions, or moral agency (Ryan, 2020).

Following established usage in epistemology (e.g., Goldman, 1999; Fricker, 2007), we use epistemic to refer to matters of credibility, authority, and reliability in relation to knowledge claims. Building on this foundation, we define an epistemic relationship as the interactional configuration through which humans and AI are positioned relative to one another as contributors to knowledge. In our analysis, such relationships are operationalized along three dimensions distilled from recent debates: assessment perspective, trust type, and human epistemic status. From this perspective, assessment, trust, and epistemic status emerge as core properties shaping such relationships. Yet it remains uncertain how these properties are actually combined and enacted in practice, and what kinds of relationship types they yield in academic work. To investigate this, we drew on 31 semi-structured interviews with academics across varying levels of engagement with AI, guided by the research question: ***What epistemic relationships emerge in researchers' interactions with AI, and how are they enacted across academic knowledge practices?***

We approach this work as an exploratory effort in conceptual development: informed by philosophy and grounded in empirical practices, our study offers an exploratory typology of epistemic relationships that can inform future inquiry. Our aim is not to offer an exhaustive taxonomy, but to identify recognizable combinations of epistemic properties in HAI practice that can be examined and refined in subsequent studies.

Literature review

Functional role framings of AI and their epistemic limits

HAI research has developed a wide array of typologies describing the roles AI systems play within task-based and decision-oriented workflows. Across HCI studies, AI has been labeled as a tool, assistant, servant, oracle, collaborator, partner, or guide (Papachristos et al., 2021; Kim et al., 2021; Kim et al., 2024). These framings are typically grounded in technological functionality and behavioral affordance.

These functional classifications also inform adjacent concerns, mainly including user perception (Hwang & Won, 2022), cognitive factors (Divis et al., 2022; Taudien et al., 2022), performance evaluation (Fragiadakis et al., 2025), and trust calibration (Kahr et al., 2025; Mehrotra et al., 2024), often through the lens of technology adoption models (e.g., Sundar, 2020; Glikson & Woolley, 2020; Pareek et al., 2025). Constructs such as perceived agency, sense of control and trustworthiness recur across this literature, highlighting ongoing attention to automation boundaries and power dynamics in human-machine interaction (including but not limited to AI) (Cornelio et al., 2022; Hwang & Won, 2022; Limerick et al., 2014). Papachristos et al. (2021) show how users' attribution preferences and self-assessed competence shape reliance patterns, while Pareek et al. (2025) demonstrate how causal explanations affect trust calibration. Kahr et al. (2025) find that experts may withhold trust from high-performing AI to preserve human values—hinting at deeper epistemic judgments. However, these studies often remain outcome-oriented and under-theorized in epistemic terms. The concept of trust itself is frequently ill-defined or treated instrumentally, rather than examined through normative or epistemological lenses.

This orientation results in a conceptual gap. Most framings treat AI as an input-output machine positioned within fixed interface boundaries, rather than a participant in knowledge construction. Even where 'agency' or 'trust' is discussed, the focus is typically on perception and delegation, not on how AI systems reconfigure the user's epistemic stance—as knowers, validators, or co-constructors of knowledge. In this sense, HCI remains epistemically thin: it lacks a clear framework for analyzing how AI systems mediate or reshape knowledge relationships.

Yet, epistemic concerns are not entirely absent. Many HCI discussions implicitly echo philosophical concepts such as epistemic agency, authority, and dependence, though rarely named as such. For example, perceived AI autonomy (Kim et al., 2024) maps closely onto epistemic agency. Papachristos et al. (2021) show that users may see the same AI as a confirming mirror, helpful assistant, or trusted oracle—implying variable epistemic status. Similarly, the notion of epistemic authority is evoked in trust studies that assess AI expertise or anthropomorphic framing, yet seldom explore when such authority is justified. Dependence on AI outputs—manifested in automation bias—is widely documented but rarely analyzed in terms of its long-term epistemic implications, such as the erosion of user agency. A notable exception is Wu et al. (2024), who explicitly examine shared epistemic agency in GenAI-facilitated learning and point toward a more symbiotic epistemic relationship between humans and machines, one in which both parties' contributions are recognized and epistemically valued.

In sum, while HCI scholars have made substantial progress in classifying AI roles and understanding user perceptions, its theoretical treatment of human-AI relations remains fragmented and predominantly functional. The implicit epistemic questions embedded in trust, agency, and attribution call for a more explicit epistemic framing—one that can account for how AI systems participate in, mediate, or challenge knowledge construction. To address this, the next section introduces an epistemic lens mainly from recent philosophical study of technology that offers us the potential to reconnect these scattered insights into a more coherent framework for analyzing human-AI knowledge relations.

Epistemic perspectives on human–AI interactions

Although traditionally limited to human-to-human interactions, epistemic concerns are increasingly being raised in relation to non-human agents. In particular, generative AI has prompted renewed inquiry into whether such systems can hold or exercise epistemic capacities. As Alvarado (2023) notes, artificial intelligence is best understood as an epistemic technology, a view that opens new directions for theorizing its role in knowledge practices. We briefly review this philosophical turn to lay the groundwork for our own analytical framework on human–AI epistemic relationships.

Epistemic nature of AI: epistemic technology and epistemic authority

Recent philosophical scholarship has increasingly framed AI through an epistemic lens, emphasizing its role not merely as a computational tool, but as an epistemic agent embedded in knowledge-producing practices. AI, in particular, is seen as both cognitive and epistemic technologies (Alvarado 2023; Heersmink et al., 2024): cognitive, in that they assist with summarizing, classifying, or translating; epistemic domains, broader categories of practices in that they provide information, answer questions, and generate outputs that support knowledge acquisition, retention, and use.

Epistemic technologies can be conceptually and practically distinguished from other technologies based on what they are designed for, what they do, and how they do it (Alvarado, 2023). From this perspective, AI can be uniquely positioned as an epistemic technology because it is developed and deployed for epistemic contexts such as inquiry, applied to manipulate epistemic content, and operates on such content through processes like inference, prediction, and analysis. Building on this, it has been argued that AI systems—especially generative models—function as epistemic technologies whose behavior is shaped not solely by designer intent, but also by training data. These systems can act as epistemic enhancers, augmenting users' ability to access and apply knowledge. Crucially, this enhancement is context-sensitive and interaction-dependent (Ferrario et al., 2023).

This raises the further question: under what conditions might AI systems be granted epistemic authority? One recent account proposes a non-anthropocentric model of epistemic authority grounded not in intentional states or mental beliefs but in the suitability of an output as a truth-indicator (Hauswald, 2025). On this view, epistemic deference is justified not because an agent has beliefs, but because its outputs reliably track truth in a given domain. When users encounter outputs that are accurate, coherent, and contextually appropriate, they may justifiably defer to them much like to trusted experts. Importantly, epistemic authority is not static—it is a situational and revisable stance, contingent upon the user's activities, context, and evaluative standards. This offers a more context-based understanding of AI's place within human knowledge construction.

Attributes of epistemic relationships in human–AI interaction.

Hauswald's model contributes in three keyways. First, it decouples epistemic authority from anthropocentric assumptions, allowing for non-human agents to occupy epistemic roles. Second, it emphasizes that such authority is not an inherent attribute but emerges through situated interactions with users. Third, and most critically, it disentangles epistemic authority from trust, enabling a sharper conceptual focus on how judgments of epistemic reliability are formed.

This moves toward separating epistemic authority from trust aligns with a broader trend. A few scholars (Blanco et al., 2022; Ferrario et al., 2024) have explored more nuanced definitions of trust and its construction process. For instance, according to Blanco et al.'s (2022) relational theory of trust, trust is not grounded in interpersonal morality or anthropomorphic expectations, but rather shaped by context and task-specific epistemic positioning. In this framework, trust is not a prerequisite for epistemic authority, but a parallel lens—one through which human–AI interaction is framed and negotiated. On this account, trust and authority are distinct but co-occurring

dimensions in human–AI interaction, shaped by how users frame their expectations in specific activities and domains.

This shift also foregrounds the question of assessment: how do users determine the appropriateness or reliability of AI outputs? While outcome-based evaluation remains common, recent work (e.g., Russo et al., 2024) suggests that users also assess underlying epistemic processes, such as transparency, consistency, or alignment with disciplinary norms. This process-based assessment complicates but enriches our understanding of how epistemic relationships between humans and AI unfold.

Building on this, the notion of hybrid epistemic agency highlights how human cognition and AI outputs can become entangled in collaborative knowledge practices (Ferrario et al., 2024). However, this hybridity does not erase epistemic asymmetry. On the contrary, it can produce new forms of epistemic dependence, where users defer to AI outputs without understanding the underlying logic. Here, the expertise dimension becomes especially salient. Asymmetries between experts and lay users shape the evaluation criteria applied, the likelihood of critical engagement, and the readiness to defer to AI-generated results. While the expert–layperson distinction is well-known in classical epistemology, its significance is renewed in the context of generative AI, where systems increasingly mediate knowledge production and validation. Elaborating this distinction, recent work has identified two forms of epistemic stance -- epistemic symmetry and epistemic asymmetry (Russo et al., 2024). The former refers to cases where expert users evaluate AI outputs by directly interrogating their generative processes—for instance, by examining their logic, fairness, or technical soundness. In the later, by contrast, non-experts must rely on indirect cues or institutional guarantees. These two configurations not only shape how users assess AI but also influence the types of epistemic trust they develop.

Taken together, recent philosophical debates increasingly conceptualize AI as an epistemic technology and, under certain conditions, as a potential holder of epistemic authority. This recognition provides the conceptual backdrop for examining the conditions and boundaries under which AI may be accorded epistemic standing in practice. Complementing this, discussions of epistemic attributes—assessment perspective, trust type, and human epistemic status (expert or not)—offer analytic footholds for operationalizing such an inquiry, enabling us to specify how these conditions and boundaries can be observed in human–AI interaction. These elements jointly establish the foundation for the epistemic relationships (ERs) we will further identify and theorize in the following sections.

Method

Because this study seeks to examine human–AI epistemic interaction, we adopted an exploratory methodological approach oriented toward theory building. Our analysis of empirical data is not intended to provide a comprehensive account of AI use in academic practice; rather, it serves to surface patterns and insights that **inform the development of propositions about epistemic relationships between humans and AI**. To support this process of theorizing, we employed data triangulation by drawing on two independently collected interview datasets and applying the same analytical framework to each dataset in parallel, without pooling them into a single analytic corpus. Triangulation allowed us to broaden the empirical base, reduce potential biases tied to a single dataset, and identify recurring themes across the two datasets, thereby strengthening the credibility of our theoretical insights (Denzin, 1970; Carter et al., 2014).

Table 1 summarises participant characteristics. One dataset comprises internationally recruited self-identified digital humanities scholars (P01–P13) through multiple channels (e.g., convenience-based outreach and broader online recruitment), and the other comprises China-based information science scholars recruited via convenience sampling (P14–P31). Each interview, conducted either in person or via Zoom, lasted approximately one hour and was audio-recorded

for analysis. Participants were asked about the research activities for which they use AI, their perceptions of AI's role, and the ways they evaluate the quality and outcomes of AI involvement.

Participant	Region	Domain	Position
P01	Germany	Digital Humanities	Research Scientist
P02	Canada	Digital Humanities; History	Associate Professor
P03	UK	Digital Humanities	Professor
P04	US	History	Project Director
P05	US	Rhetoric & Composition	Lecturer
P06	US	English	Associate Professor
P07	Mexico	Communication & Media Studies	Assistant Professor
P08	Germany	Spanish Literature	Research Associate
P09	US	Digital Humanities	Project Director
P10	US	Library Science	Assistant Professor
P11	Italy	Digital History	Research Fellow
P12	US	Digital History	PhD Candidate
P13	China	Cultural Heritage	Assistant Professor
P14	China	Publishing	Associate Professor
P15	China	Publishing	Professor
P16	China	Publishing	Assistant Professor
P17	China	Communication	Assistant Professor
P18	China	Publishing	Associate Professor
P19	China	Data Science	Professor
P20	China	Publishing	Assistant Professor
P21	China	Communication	Associate Professor
P22	China	Chinese Literature	Postdoc Researcher
P23	China	Scientific and Technical Information	Assistant Professor
P24	China	Information Science	Assistant Professor
P25	China	Publishing	Assistant Professor
P26	China	Ethics	Assistant Professor
P27	China	Data Science	Assistant Professor
P28	China	Data Science	Assistant Professor
P29	China	Library Science	Assistant Professor
P30	China	Library Science	Associate Professor
P31	China	Information Science	Associate Professor

Table 1. Overview of the participants.

To analyze the interview data, we developed a five-dimensional codebook (Table 2). Three dimensions—assessment perspective, trust type, and human epistemic status—were drawn from prior epistemic research on HAI, while metaphor was also informed by the literature but operationalized here as an interpretive lens. The activity dimension, in contrast, was inductively identified through open coding. Each dimension represents a distinct analytic axis. Specifically, values for metaphor, assessment perspective, human epistemic status, and trust type were derived from the research literature. For the activity dimension, instead, we conducted open coding to identify the practices participants themselves emphasized in their accounts, following the ‘*information practices*’ tradition (Savolainen, 2008; Cox, 2012). This practice-oriented approach enables us to attend to the meanings and epistemic negotiations embedded in these activities, rather than imposing rigid classificatory schemes. Because of this approach, some conceptual overlap appeared among the terms participants used. For instance, ‘*content generation*’ (Table 2) may encompass ‘*writing*,’ but when participants referred to AI producing outputs beyond text (e.g., videos or images), we retained ‘*content generation*’ as a distinct activity.

Code Category	Definition	Value	Theoretical Reference
Metaphor	Judgment about the epistemic status assigned to AI	Tool, Assistant, Co-agent, Coach/Mentor, Authority (oracle)	Kim, et al. (2023); Papachristos et al. (2021)
Activity	Specific activities in research where AI is utilized, defining the cognitive scenario and usage purpose.	E.g., brainstorming, literature search & review, summarization, coding & programming, analysis, content generation, translation, writing, editing, general search, critique	N/A. The activities listed in the “Value” field emerge from our analysis of the interview data.
Assessment perspective	How humans evaluate the involvement of AI in their work process	Outcome-based (e.g., accuracy, quality of results), Process-based (e.g., algorithms), Mixed	Russo et al. (2024)
Human epistemic status	Identity and position of the human involved in the activity, which may influence their AI role assignment and evaluation.	AI expert Non-AI expert	Russo et al. (2024)
Trust type	Humans' willingness to rely on AI, shaped by factors such as evaluation of its performance, perceived reliability, and understanding (or lack thereof) of its underlying processes.	Distrust: both the absence of trust and a confident negative expectation about another's conduct	Blanco (2025)
		Reliance: activity-based dependence on AI without emotional or moral expectations. It focuses solely on outcomes, not methods (e.g., a researcher may rely on AI to process large datasets quickly, without questioning how the results are generated)	
		Appropriate trust: Grounded in justified confidence in both the AI's results and its methods (e.g., a researcher would evaluate the AI's methodology and confirm not just the results but also how it meets disciplinary standards).	
		General trust: broad, informal belief that AI will perform well, often based on reputation rather than evidence (e.g., a researcher might use AI in their work because colleagues recommended it or it's well-regarded in the literature, without exploring how it works).	

Table 2. Codebook

Using this codebook, the two authors collaborated to calibrate the code definitions and decision rules. Coding was then conducted independently on each author's assigned transcripts, with regular meetings to review ambiguous cases and harmonise code application; discrepancies in interpretation were resolved through collaborative discussion. Building on the coding results, we identified **epistemic relationships (ERs) inductively as patterned combinations observed across both datasets**, analysing user narratives across three mutually exclusive dimensions: human epistemic status (e.g., AI expert vs. non-AI expert), assessment perspective (e.g., process- vs. outcome-based

evaluation), and trust type (e.g., distrust, reliance, appropriate, general), together with the activities and metaphors associated with them. The ERs do not represent all possible combinations of these dimensions; rather, they reflect major empirically grounded patterns that emerged as stable clusters in how users negotiate their epistemic relationships with AI across various contexts.

Results

Five ERs emerged from our analysis (Table 3), which highlight **patterns of AI's epistemic participation in human-AI interaction** and show a **context-dependent character** of epistemic relationships with AI. This shows that the relationships are not fixed but are shaped through ongoing negotiation across different research activities and goals. Users' epistemic expectations and trust in AI varied based on activities (e.g., coding, translation, brainstorming), modes of assessment (e.g., outcome-based vs. process-based), and the perceived stakes involved.

The five ERs we identified capture the situated nature of epistemic relationships between human researchers and AI, offering a more nuanced understanding of how these relationships are constructed and adjusted in research and pedagogical contexts. However, it is worth noting that the five relationship types we present **should not be read as an exhaustive taxonomy** of all possible configurations. Given that these epistemic dimensions could, in principle, generate many more combinations, our aim is not to map every potential variant but to identify analytically recognizable types that were recurrent across our interviews.

ER Type	Brief Definition	Metaphor	Trust Type	Assessment Perspective	Human Epistemic Status	Activities (Examples)
ER-1 Epistemic Abstention	Uses AI for epistemic engagement but denies its epistemic contribution	Tool	Distrust	Outcome-based	Non-AI expert or AI expert	Summarization, General search
ER-2 Instrumental Reliance	Relies on AI solely for efficiency; assigns no epistemic status to AI	Tool	Reliance	Outcome-based	Non-AI expert	Editing, coding/programming
ER-3 Contingent Delegation	Delegates judgment to AI in specific activities while maintaining oversight	Assistant/Co-agent	Appropriate Trust	Outcome-based	Non-AI expert	Content generation, writing; general search
ER-4 Co-agency Collaboration	Treats AI as a collaborative thinker in knowledge construction	Co-agent / Mentor	General / Appropriate trust	Mixed (process/ outcome)	Non-AI expert or AI expert	Writing, translation, brainstorming, summarization
ER-5 Authority Displacement	Grants AI partial authority in knowledge production	Co-agent / Mentor	General Trust	Process-based	Non-AI expert or AI expert	Critique, brainstorming

Table 3. Five epistemic relationships of human–AI interaction

ER-1 ‘*Epistemic abstention*’ refers to a relationship where users utilize AI for epistemic activities but actively deny their epistemic authority. Distinct from other ERs, trust in this collaboration pattern is minimal or even absent, and the assessment remains outcome oriented. For instance, in a teaching exercise, a history instructor asked students to critique AI-generated drafts, which consistently received low grades (‘C’) for their lack of rigor (P04). In this case, AI was tasked with producing an epistemic contribution, yet its outputs were systematically devalued. In other cases, abstention was grounded in principle rather than performance. For example, P05, who is a rhetoric and communication scholar, argued that even if AI could reasonably be used to assist with research activities, it should not replace core processes of analysis and reflection that make research meaningful:

I definitely think there are aspects of research that generative AI shouldn’t replace. To me, it comes back to a core question: Why are you doing research in the first place? Do you enjoy being a researcher? If you’re genuinely interested in research, it should be things you want to engage with. So, when someone wants to replace those parts with AI, I see it as a question of purpose. If I’ve done five interviews and need to categorize the transcripts, why not take the time to go through the data myself? That’s what makes me a better researcher. The process itself is valuable. (P05)

A similar concern emerged in the other dataset: a publishing scholar (P15) described AI’s circulation-related suggestions as superficial and not something they would trust to replace deep human judgement. These examples show that ER-1 entails not only practical skepticism but also deeper commitments to preserving human epistemic authority. ER-1 marks one end of the spectrum of epistemic relationships.

Unlike ER-1, which denies AI any epistemic standing, **ER-2** reflects a utilitarian view of AI, in which users treat it as a tool to increase work efficiency without granting it epistemic status. Assessment is strictly outcome-based (e.g., whether the work gets done), and trust remains at the level of functional reliability, not belief. This relationship is most common among non-AI experts performing tasks such as programming or editing, where AI is employed to reduce workload or streamline processes. For example, one digital humanities researcher (P02) uses AI to assist him with programming, which he is unfamiliar with and finds time-consuming. As he explained below, AI functions as a programming assistant, enabling him to achieve his goals more efficiently. With no background in computer science, he primarily evaluates AI’s performance based on the outcomes, and his reliance on AI as a tool developed gradually through a trial-and-error approach:

I am almost exclusively using AI for data analysis—writing Python scripts to collect web data and build databases. I was never trained in computer science; everything from GIS to Python is self-taught. That’s where generative AI is enormously helpful—it lets me do things I only need a few times a year. Because I don’t code regularly, I often forget what I’ve learned and need to relearn it. AI helps me bridge that gap. What would’ve taken two weeks now takes two hours. I can quickly go from a lower to upper intermediate coder and focus on the logic of the algorithm instead of syntax or data types. (P02)

In another case, participants across the two datasets reported using AI to check and correct the grammar in their papers, which effectively improved both their writing and their confidence as non-native English speakers (P10, P27, P29).

In comparison, **ER-3** involves selectively delegating cognitive work to AI under specific conditions while retaining human supervision. In this case, AI is framed as an assistant or co-agent, and users exhibit appropriate but conditional trust. Users primarily evaluate its performance based on the outcome, not its reasoning process. We also observed this relationship pattern among researchers

who did not self-identify as an AI expert. For instance, P04, who is a history professor, described her interaction with AI as follows:

I check everything it produces. It's like working with a research assistant: if you hire an undergraduate to help and they hand you a nicely polished essay, you're not just going to accept it as-is. You'll need to go through it, add comments, and make sure they did a good job. (P04)

P02 also demonstrated how he interacted with AI effectively with human input and supervision: I get it to give me all the code, assemble it into one package. I go run it—90% of the time it doesn't work. But the really nice thing ... is that it spits out the error code and you just copy and paste the error code back into ChatGPT. And it politely responds, *'oh sorry I made this mistake ... and if you're lucky, it works that time.'* Similarly, another participant (P23) noted that AI could outline and explain possible analytical models, but that the final choice still depended on their own judgement and supervision.

Different from the contingent nature of ER-3, **ER-4** describes a more integrated epistemic relationship in which users treat AI as a thinking partner or mentor. In this scenario, trust generally happens at a higher level and in a more flexible form, shaped by ongoing interaction rather than predetermined roles. Additionally, assessment perspectives combine both process and outcome considerations. This configuration is evident in more interpretive and open-ended activities, such as writing, translation, or brainstorming, where users draw on domain and/or AI expertise. For example, P01 described a case in which AI was engaged as a co-agent partner by an early modern scholar working on the translation of diplomatic correspondence between Italy and the Dutch Republic. Although the scholar is an expert Latinist, when uncertain about specific translations, he would *'treat GPT like a colleague'* and consult it with Latin translation questions:

He will say, *'What do you think about this?'* or *'I think it might mean this, but I'm not sure.'* Then GPT responds, saying something like, *'I suggest this translation, although this might also be possible.'* He's essentially having a conversation with it. He has completely anthropomorphized the tool and is genuinely pleased to have what feels like a colleague to assist with Latin translations. (P01)

P06, an English professor, noted that AI's creative capabilities helped him better understand his own work, ultimately leading to a process of co-creation:

I was really struggling with the subtitles [for my book] ... I gave [ChatGPT] a try, and it revealed all sorts of things I didn't want, but ironically, that helped me realize what I did want to say. I also used Midjourney to generate the cover art for the book, which features an image of Queen Victoria typing on a laptop. (P06)

P19, a data scientist, reflected on how their knowledge and expertise of AI models transformed their relationship with AIGC technologies:

Before the emergence of reasoning models, I viewed AIGC as a tool. But since the release of reasoning models last November, I've started treating it as an assistant. That's why I use it every day now...I use AI to summarize literature, extract key points, accelerate reading, polish language, and translate. It's had a significant impact on my workflow. At one point, I even felt that many academic papers and books might not be necessary. (P19)

Comparatively, users engaging in **ER-5** acknowledge AI's epistemic authority in certain domains, allowing AI to lead in knowledge production. This relationship is marked by high-level trust, including appropriate and general trust, as well as process-based assessment, which often occurs in complex reasoning activities. Users in this category often possess hybrid expertise that mixes

domain knowledge and expertise in AI systems. For example, a professional programmer who is both a digital humanities scholar and an AI expert, engages in what he described as ‘pair programming’ with AI, treating it as a co-agent: ‘It’s like a programming buddy... [I] accept about 70% of its suggestions. And it’s completely changed the way I code’ (P01). P05 and P21 also trusted AI with advanced conceptual work, using ‘generative AI conversationally, especially during the brainstorming stage...and it helps [her] talk through abstract ideas and clarify research questions.’ These examples suggest ER-5 as a highly collaborative relationship that involves recognized AI epistemic contributions.

In summary, the five ERs illustrate a spectrum of how humans position AI in knowledge work, varying in the degree of epistemic authority they attribute to AI. ER-1 (Epistemic Abstention) reflects a skeptical stance: while AI may be used for epistemic purposes, its epistemic contribution and status are explicitly denied, often out of distrust or principled commitments to the value of human inquiry. ER-2 (Instrumental Reliance) represents the most utilitarian stance, where AI is treated strictly as a tool for efficiency with no epistemic standing. ER-3 (Contingent Delegation) moves a step further, allowing AI to act as an assistant under human oversight, with conditional trust tied to outcomes. ER-4 (Co-agency Collaboration) expands this into a genuine partnership, where AI is engaged as a thinking collaborator and trust spans both process and outcomes. ER-5 (Authority Displacement) shifts more epistemic authority onto AI, as humans grant it partial leadership in knowledge production, particularly in complex reasoning or creative tasks. Together, these ERs highlight the diverse ways users negotiate AI’s role, from purely functional tool to epistemic partner, or even as a resource whose authority is actively resisted.

Discussion

Epistemic relationship as a situated and negotiated outcome

Our findings show that users’ metaphoric descriptions of AI, such as a ‘tool’ or an ‘assistant,’ do not necessarily correspond to how they actually interact with AI in specific tasks. Instead, epistemic relationships are negotiated across contexts, assessment strategies, types of trust, and user identities. Building on these insights, we propose an exploratory typology that frames epistemic relationships not as a fixed role assignment but as a context-based contingent configuration negotiated between humans and AI systems. Our analytical lens conceptualizes human–AI interaction as an evolving epistemic relationship shaped by users’ situated engagement. Crucially, we take the epistemic capacities of AI, as discussed in recent work on AI as epistemic technology or authority (Alvarado, 2023; Ferrario et al., 2023; Hauswald, 2025), as the conceptual premise for theorizing such relationships.

In this conceptual framing, the human side of interaction is captured through observable behavioral attributes that reflect how users position and make sense of AI, including the interaction metaphor (e.g., tool, assistant, co-agent), the type of trust extended toward the AI (e.g., reliance, appropriate trust), the assessment perspective (i.e., whether users evaluate AI outputs based on outcomes or underlying processes) and the user’s epistemic status (e.g., AI expert or not). The relationship between these dimensions is shaped by specific activities. Those variations influence the degree to which epistemic capacities are invoked or deferred, making the epistemic relationship context-sensitive and fluid rather than static.

This lens allows us to trace how AI’s epistemic status is co-constructed through situated human practices, revealing not only how users metaphorically frame AI roles but also how these framings relate to underlying epistemic commitments. Our model thus offers a more nuanced understanding of how epistemic authority, dependence, and co-agency emerge in real-world settings, emphasizing the importance of temporally flexible and identity-sensitive configurations in shaping meaningful human–AI knowledge partnerships. It also contributes to a relational epistemology that resists both technological determinism and anthropocentric essentialism,

instead foregrounding the co-evolutionary nature of epistemic positioning in hybrid systems. This aligns with Hayles' (2023) critique of anthropocentrism, as our empirical cases demonstrate how epistemic authority is continuously configured through situated human–AI engagements, rather than attributed through predefined roles or human-centric criteria. By importing an epistemic perspective into HCI and drawing on key concepts from philosophy of technology and AI epistemology, this typology offers one step toward connecting abstract normative inquiry with empirical studies of human–AI interaction.

Implications for the future of work

The five ERs we identified capture the contingency of human–AI interaction in knowledge-intensive settings and offer broad implications for the design of workflows, training, governance, and organizational practices beyond higher education. First, the findings of our work can support more effective task allocation by aligning AI deployment with the characteristics of activities and user needs. Each ER type corresponds to a distinct workflow model—for example, ER-2 suits structured, goal-driven tasks, while ER-4 fits exploratory, interpretive work requiring iterative human–AI interaction. These early insights provide a foundation for developing ‘*activity × epistemic relationship*’ matrices to guide more nuanced and context-sensitive AI integration. Second, recognizing a user's ER type can inform more tailored AI literacy training. For example, users in ER-3 or ER-5 may benefit from greater transparency and understanding around AI reasoning and limitations to foster process-based trust. Those in ER-1 require institutional safeguards to define epistemic boundaries and prevent overreliance or misplaced responsibility. Third, our findings also challenge the binary notion of ‘*trust in AI*.’ The ERs reflect a spectrum of trust configurations, each with differing accountability implications. While ER-2 and ER-3 maintain clear human responsibility, ER-1 nevertheless requires explicit mechanisms to govern trust and ensure transparent responsibility in AI-assisted decision-making. Finally, the contingent nature of ERs suggests job roles are epistemically fluid, not fixed. As users gain experience and engage more critically with AI, they may shift from ER-2 toward ER-4 or ER-1. This calls for organizations to support employees' evolving epistemic trajectories, enabling adaptive co-development of human roles and AI capabilities.

Limitations and next steps

As an exploratory effort in conceptual development, this study has some limitations. First, our typology is inductively derived from interview data, which limits generalizability. While the five ERs offer a useful vocabulary for conceptualizing human–AI epistemic configurations, additional empirical studies in the future are needed to validate and refine these categories across diverse domains. Second, our current data offer only limited insight into the trajectories or transitions across ER types and yet to demonstrate the change over time. Future longitudinal or process-tracing studies could help clarify how such shifts occur. Lastly, while we suggest a relational epistemology that resists both technological determinism and anthropocentric essentialism, further theoretical work is needed to articulate how epistemic agency can be jointly constructed in hybrid systems.

Conclusion

Our approach is not exploratory pattern-finding in the narrow sense. Instead, it is an exploratory effort in conceptual development: we start from a philosophical account of epistemic properties and use empirical material to articulate how these properties combine into recognizable relationship types. The resulting typology, presented as a spectrum of how humans position AI in knowledge work, varies in the degree of epistemic authority they attribute to AI and offers a conceptual scaffold for analyzing human–AI interaction in knowledge work. These are not merely descriptive themes, but propositions about relational configurations that can be examined, extended, or tested in future studies. This exploratory study introduced recent philosophical discussions about AI's epistemic status, a still developing and contested topic, into the more

established HCI discourse, which has traditionally focused on functionality, usability, and adoption. While many HCI studies conceptualize AI through static functional roles, our work foregrounded the epistemic dimension of AI and the ways in which users relate to AI in epistemic practices. By identifying five types of epistemic relationships, we provided a conceptual link between normative theories of epistemic agency and the practical realities of human-AI interaction in work settings.

About the author(s)

Shengnan Yang is an Assistant Professor in the Faculty of Information and Media Study, University of Western Ontario, Canada. She received their Ph.D. from Indiana University Bloomington, and her research interests include human-AI interaction, digital governance, and social informatics. She can be contacted at syang859@uwo.ca.

Rongqian Ma is an Assistant Professor in the Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, USA. She received her Ph.D. from the University of Pittsburgh, and her research interests include digital humanities and AI, digital archives, and scholarly communication. She can be contacted at rm56@iu.edu.

References

- Alvarado, R. (2023). AI as an Epistemic Technology. *Science and Engineering Ethics*, 29(5), 32.
- Blanco, S. (2025). Human trust in AI: A relationship beyond reliance. *AI and Ethics*.
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *Oncology nursing forum*, 41(5), 545-547.
- Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. (2025). Human-AI ensembles: When can they work? *Journal of Management*, 51(2), 536-569.
- Cornelio, P., Haggard, P., Hornbaek, K., Georgiou, O., Bergström, J., Subramanian, S., & Obrist, M. (2022). The sense of agency in emerging technologies for human-computer integration: A review. *Frontiers in Neuroscience*, 16.
- Cox, A. M. (2012). An exploration of the practice approach and its place in information science. *Journal of Information Science*, 38(2), 176-188.
- Csaszar, F. A., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors (SSRN Scholarly Paper No. 4913363). Social Science Research Network.
- Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine Publishing Company.
- Divis, K., Howell, B., Matzen, L., Stites, M., & Gastelum, Z. (2022). The cognitive effects of machine learning aid in domain-specific and domain-general tasks. *Hawaii International Conference on System Sciences*.
- Felin, T., & Holweg, M. (2024). Theory is all you need: AI, human cognition, and decision making. *SSRN Electronic Journal*.
- Ferrario, A., Loi, M., & Viganò, E. (2020). In AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions. *Philosophy & Technology*, 33(3), 523-539.

- Ferrario, A., Facchini, A., & Termine, A. (2024). Experts or authorities? The strange case of the presumed epistemic superiority of artificial intelligence systems. *Minds and Machines*, 34(3), 30.
- Fragiadakis, G., Diou, C., Kousiouris, G., & Nikolaidou, M. (2025). Evaluating human-AI collaboration: A review and methodological framework (No. arXiv:2407.19098).
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.
- Hauswald, R. (2025). Artificial epistemic authorities. *Social Epistemology*, 1–10.
- Hayles, N. K. (2023). Technosymbiosis: Figuring (out) our relations to AI. In J. Browne, S. Cave, E. Drage, & K. McNerney (Eds.), *Feminist AI* (1st ed., pp. 1–18). Oxford University Press/Oxford.
- Heer, J. (2019). Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116(6), 1844–1850.
- Heersmink, R., De Rooij, B., Clavel Vázquez, M. J., & Colombo, M. (2024). A phenomenology and epistemology of large language models: Transparency, trust, and trustworthiness. *Ethics and Information Technology*, 26(3), 41.
- Kahr, P., Rooks, G., Snijders, C., & Willemsen, M. C. (2025). Good performance isn't enough to trust AI: Lessons from logistics experts on their long-term collaboration with an AI planning system. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–16.
- Kim, J., Merrill Jr., K., & Collins, C. (2021). AI as a friend or assistant: The mediating role of perceived usefulness in social AI vs. functional AI. *Telemat. Inf.*, 64(C).
- Kim, T., Molina, M. D., Rheu, M. (MJ), Zhan, E. S., & Peng, W. (2024). One AI does not fit all: A cluster analysis of the laypeople's perception of AI roles. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–20.
- Mehrotra, S., Degachi, C., Vereschak, O., Jonker, C. M., & Tielman, M. L. (2024). A Systematic Review on Fostering Appropriate Trust in Human-AI Interaction: Trends, Opportunities and Challenges. *ACM J. Responsib. Comput.*, 1(4), 26:1–26:45.
- Papachristos, E., Skov Johansen, P., Møberg Jacobsen, R., Bjørn Leer Bysted, L., & Skov, M. B. (2021). How do people perceive the role of AI in human-AI collaboration to solve everyday tasks? *CHI Greece 2021: 1st International Conference of the ACM Greek SIGCHI Chapter*, 1–6.
- Pareek, S., Schömbs, S., Velloso, E., & Goncalves, J. (2025). "It's not the AI's fault because it relies purely on data": How causal attributions of AI decisions shape trust in AI systems. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18.
- Ryan, M. (2020). In AI we trust: Ethics, artificial intelligence, and reliability. *Science and Engineering Ethics*, 26(5), 2749–2767.
- Russo, F., Schliesser, E., & Wagemans, J. (2024). Connecting ethics and epistemology of AI. *AI & SOCIETY*, 39(4), 1585–1603.
- Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human-AI interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1), 74–88.

- Savolainen, R. (2008). Everyday information practices: A social phenomenological perspective. Scarecrow Press.
- Taudien, A., Fügner, A., Gupta, A., & Ketter, W. (2022). The effect of AI advice on human confidence in decision-making. Hawaii International Conference on System Sciences.
- Wu, J.-Y., Lee, Y.-H., Chai, C. S., & Tsai, C.-C. (2025). Strengthening human epistemic agency in the symbiotic learning partnership with generative artificial intelligence. Educational Researcher, 0013189X251333628.

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).