Article

Not peer-reviewed version

The Pseudo-Confidence Paradox: The Epistemic Gap in Everyday AI Use

Lyazzat Tulbayevna Kurmanbayeva , Anar Saduakasovna Tanabayeva * , Akmaral Ivanovna Doszhanova , Gulbakyt Kadyrzhanovna Shashayeva , Denis Bakarassov , Adilbek Knarovich Bisenbaev

Posted Date: 19 May 2026

doi: 10.20944/preprints202605.1211.v1

Keywords: artificial intelligence; pseudoconfidence; knowledge; epistemology; literacy; verification

Article

The Pseudo-Confidence Paradox: The Epistemic Gap in Everyday AI Use

Lyazzat Tulbayevna Kurmanbayeva [1], Anar Saduakasovna Tanabayeva [2,*], Akmaral Ivanovna Doszhanova [3], Gulbakyt Kadyrzhanovna Shashayeva [4], Denis Bakarassov [5] and Adilbek Knarovich Bisenbaev [6]

1 Mukhametzhan Tynyshbayev ALT University, Almaty, Kazakhstan
2 International Information Technologies University, Almaty, Kazakhstan
3 Al-Farabi Kazakh National University, Almaty, Kazakhstan
4 Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
5 Institute of Philosophy, Political Science and Religious Studies, Almaty, Kazakhstan
6 Ministry of Science and Higher Education of the Republic of Kazakhstan, Astana, Kazakhstan
* Correspondence: a.tanabayeva@iitu.edu.kz

Abstract

This study examines the phenomenon of pseudoconfident knowledge in the context of the everyday use of generative artificial intelligence. By pseudoconfident knowledge, we mean a response that is substantively plausible, rhetorically coherent, and outwardly persuasive but is treated and understood as knowledge before its actual reliability has been established**.** Of course, we do not use the term "pseudoconfident knowledge" to denote knowledge in the strict epistemological sense. Rather, it denotes a special form of AI-generated content that acquires the status of knowledge in the user's perception before its reliability, source-based justification, or factual correctness have been established. The problem here is not that such an answer is already knowledge but that it is prematurely accepted as knowledge because of its coherence, completeness, and rhetorical confidence. The aim of the study is to identify the epistemic gap between the everyday operational integration of artificial intelligence and the user's critical ability to distinguish between persuasiveness and justification. The theoretical framework combines approaches to AI literacy, epistemic vigilance, and contemporary forms of digital mediation in the circulation of knowledge. The empirical basis of the study is an online survey of AI users. The analysis was conducted using descriptive statistics, contingency tables, and methods for testing associations between categorical variables. The results show that the key differentiating factor is not the frequency of AI use but the strategy used in handling its responses. More epistemically robust positions are associated with practices of comparison, editing, and verification, whereas uncritical acceptance of the answer is associated with greater vulnerability to pseudoconfident knowledge. We conclude that the spread of generative artificial intelligence is producing a new socioepistemic problem that calls for a shift in emphasis from simple instrumental literacy toward a culture of verification, doubt, and epistemic responsibility.

Keywords: artificial intelligence; pseudoconfidence; knowledge; epistemology; literacy; verification

1. Introduction

Humanity has clearly entered an age of mirrors that answer back. However, unlike an ordinary mirror, generative artificial intelligence reflects not only the question but also its most plausible verbal form. We argue that the new danger lies not in the machine's silence or even simply in its error but in its ability to speak as although knowledge had already been achieved, when what stands before us may be only its persuasive imitation.

 doi:10.20944/preprints202605.1211.v1

The spread of generative AI has radically transformed the very mode of everyday engagement with knowledge [1–3]. If the early digital era centered on the problem of access to information, the current phase is increasingly organized around access to the ready-made answer. That answer arrives instantly, takes the form of a coherent and confident text, and thereby reduces the cognitive costs of searching, interpreting, and initially understanding information [4,5]. It is precisely here that a new epistemic problem emerges. The user receives not only information but also synthesized, rhetorically integrated, and communicatively convenient content that can easily be mistaken for knowledge before its grounds have been checked. This is especially important because international discussions of AI literacy and the use of generative AI increasingly emphasize that the key challenge lies not only in mastering new tools but also in developing the human capacity to critically evaluate the answers they produce [6–9].

In this context, we introduce and develop the concept of pseudoconfident knowledge. By this, we mean content that possesses all the outward signs of cognitive adequacy: clarity, logical coherence, structural completeness, and rhetorical confidence. However, such content has not undergone sufficient epistemic verification. Pseudoconfident knowledge is not necessarily identical to falsehood. Its essence lies elsewhere: it is perceived as reliable knowledge before its epistemic reliability or warranted acceptance has been established.

In this sense, generative AI creates not only a new form of access to answers but also a distinctive environment in which plausibility increasingly begins to function as a surrogate for reliability. In this way, research on generative AI has converged with the broader tradition of social epistemology, which has long shown that human beings depend on external sources of knowledge and therefore require mechanisms of epistemic vigilance to protect them from mistaken or premature trust [10].

The urgency of this problem is reinforced by the very nature of large language models. Contemporary research convincingly shows that even the most advanced systems can produce coherent yet incorrect answers-called hallucinations or confabulations [11]. This is not a random peripheral malfunction but a structural feature of models optimized for plausible continuations of text rather than for guarantees of truth. The central question, therefore, is no longer whether AI can err but whether users recognize the limits of its epistemic reliability. In other words, the problem shifts from the level of system accuracy to the level of the human mode of dealing with outputs.

The literature on hallucinations [12] and confabulations [13] is important for our argument but does not exhaust it. Hallucination denotes an erroneous or fabricated output of the model. Pseudoconfident knowledge describes a broader social-epistemic mechanism: the user accepts rhetorically confident output as knowledge before sufficient assessment of its grounds.

Therefore, the problem of generative AI lies not only in erroneous answers but also in the speech form in which an uncertain, probabilistic, or unverified answer is often presented as a confident assertion.

This gives rise to the central contradiction of our study: the everyday use of AI is growing rapidly, yet an increase in the frequency of interaction with the technology is not necessarily accompanied by a proportional increase in critical epistemic literacy. A user may be fully effective in an operational sense-quickly obtaining the needed text, explanation, or solution—while still remaining vulnerable to pseudoconfident knowledge.

We call this discrepancy between operational access to an answer and the epistemic ability to assess its status the epistemic gap in everyday AI use. In doing so, we shift the focus from how often people use AI to the more fundamental question of what they do with the confidence the machine produces.

Let us clarify that we call this phenomenon a paradox not in a formal-logical sense but in a social-epistemic sense. The paradox consists of the fact that the growth of practical proficiency in using generative AI is not necessarily accompanied by a growth in the ability to evaluate the epistemic status of its answers. A user may become increasingly experienced in obtaining fast, useful, and stylistically convincing answers while not becoming more capable of distinguishing rhetorical completeness from justified knowledge. It is precisely this divergence between the operational

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

integration of AI and critical epistemic orientation that forms the core of our pseudoconfidence paradox.

In this sense, our question is not whether people use generative AI or even how often they use it. A different question matters more to us: how does the user handle the machine answer after receiving it? If the answer is compared with other sources, edited, checked, or treated as a preliminary hypothesis, the user preserves epistemic distance. If, however, the answer is accepted as finished knowledge because of its confident form, the risk of pseudoconfident attribution of knowledge arises.

The empirical part of our work does not claim to provide exhaustive proof of the entire philosophical structure of pseudoconfident knowledge. Its task is to test one diagnostic sign of the presumed epistemic gap. If operational habituation to AI by itself led to epistemic maturity, we would expect more frequent AI use to be associated with a greater ability to distinguish rhetorical confidence from justified acceptance of an answer. If, however, the decisive factor is not the frequency of use but the way the answer is handled, then the strategy of checking, comparison, editing, or critical distancing should prove more significant. It is precisely this distinction that we examine in the empirical part of the work.

In this context, the empirical task of our study is to examine how everyday AI use relates to the user's readiness to distinguish persuasiveness from reliability and the comfort of an answer from its verified status. We seek to show that generative AI should be understood not only as a technology for accelerated access to information but also as a new infrastructure for the production of pseudoconfident knowledge. This definition relies directly on research into AI literacy, the digital sociology of everyday life, and the contemporary epistemology of trust.

Our central thesis is that in the age of generative AI, society confronts not only an excess of information but also the normalization of forms of knowledge that appear complete before they become justified. For this reason, the problem of the future lies not only in expanding access to AI but also in cultivating a culture of resistance to pseudoconfident knowledge.

2. Theoretical Framework

We build this study around four interrelated concepts: pseudoconfident knowledge, the epistemic gap, AI literacy, and epistemic vigilance. Their conjunction is necessary because generative AI changes not only the speed of access to information but also the very form in which knowledge is presented to the everyday user. In the classical digital environment, individuals usually dealt with a plurality of sources and had to compare, select, and interpret fragments of information themselves. Generative AI radically alters this configuration by delivering an answer that is already assembled, stylistically smoothed, and rhetorically complete. The key problem, therefore, is no longer a shortage of data but a transformation in the regime of trust toward content that looks like knowledge before it has been tested. We designate this type of content as pseudoconfident knowledge.

We should emphasize that by pseudoconfident knowledge, we do not simply mean false information or just any erroneous answers produced by a generative system. We refer to a more complex cognitive form: content that bears the external marks of epistemic adequacy-coherence, logical continuity, explanatory clarity, structural completeness, and rhetorical confidence-yet still lacks the status of sufficiently verified knowledge.

Let us clarify the status of the concept we use. We emphasize once again that the term "pseudoconfident knowledge" does not mean that the content under consideration is knowledge in the strict sense of classical epistemology. Here, the prefix "pseudo" indicates precisely the absence of a fully established epistemic status. We are speaking of content that occupies the place of knowledge in the user's practical consciousness before its grounds have been established. Therefore, the object of our analysis is not knowledge as such but the situation of premature attribution of knowledge: When the user encounters a text that speaks like knowledge, looks like knowledge, and functions like knowledge, it has not yet passed through the procedures required for recognizing it as reliable.

We use the term "pseudoconfident" deliberately because the problem lies not in reliability as such but in the effect of confidence created by the rhetorical form of the AI-generated answer.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

Accordingly, the problem lies not only in possible inaccuracy but also in premature epistemic legitimation. Users tend to perceive such an answer as reliable not because its validity has been established but because its form already resembles finished knowledge. In most cases, users approach an AI answer with an a priori readiness to trust it; otherwise, they would not have posed the question in the first place. Current research on large language models heightens the significance of this distinction by showing that such systems can generate not only useful and correct answers but also plausible confabulations-substantively persuasive yet false statements. The issue, therefore, concerns not only the model's technical accuracy but also the social regime in which machine persuasiveness is received as a surrogate for truth.

It is from this point that the concept of the epistemic gap emerges naturally. By this, we mean a structural mismatch between two processes: on the one hand, the accelerating everyday integration of generative AI into practices of search, writing, learning, and explanation; and on the other hand, the much slower formation of skills for critically evaluating its answers.

Importantly, this gap cannot be reduced to a simple lack of digital literacy. It is deeper than that. Users may be fully competent in an operational sense: they may know how to turn to AI, obtain useful answers from it, and effectively incorporate those answers into their activity. However, they remain epistemically unprotected in the face of pseudoconfident knowledge. In this respect, generative AI does not merely expand access to answers, as it is often assumed to do. Rather, it redistributes the boundary between plausibility and reliability, making that boundary blurrier in everyday experience.

To remove possible conceptual ambiguity, we must distinguish four levels of analysis that are often conflated in everyday interactions with generative AI.

The first level is the rhetorical persuasiveness of the answer. It concerns the form of the utterance: its coherence, confidence, completeness, logical sequence, and stylistic smoothness. A rhetorically persuasive answer may create an impression of knowledge, but by itself, it is not yet a sign of truth or reliability.

The second level is source reliability. Here, the question is not how convincingly a particular answer is formulated but what grounds exist for trusting the source as such. In the case of generative AI, this question is especially complex because the system may produce plausible answers without a stable connection to a verifiable source, without its own responsibility for the assertion, and without an explicit distinction between knowledge, probabilistic assumption, and stylistically completed reconstruction.

The third level is the verification of a specific claim. It concerns not trust in the system as a whole but the status of an individual claim: a fact, date, reference, numerical value, causal relation, or interpretation. In this sense, verification does not mean that the user must independently prove every claim from the beginning. Rather, it presupposes proportionate epistemic checking: consulting independent sources, comparing with authoritative data, identifying questionable points, and assessing the degree of risk of error.

The fourth level is the epistemic reception of the answer. We note that the central problem of our work arises precisely here. The same AI-generated answer may be received as a draft, a hypothesis, a heuristic prompt, a probabilistic interpretation, or as already completed knowledge. Pseudoconfident knowledge arises in the last case: when a rhetorically complete and outwardly confident answer receives the status of knowledge in the user's consciousness before the source, grounds, and verifiability of its claims have been assessed.

Therefore, we do not reduce the problem of pseudoconfidence to answer errors, technical unreliability of the model, or subjective gullibility of the user. It emerges at the intersection of three elements: the machine form of confident assertion, the limited transparency of the source basis, and the human tendency to take a cognitively convenient answer as already justified knowledge.

The concept of AI literacy plays a decisive role here. Recent reviews show that AI literacy can no longer be understood solely as technical knowledge of how algorithms work or as the ability to formulate prompts effectively [14–16]. In contrast, AI literacy is increasingly defined as a

 doi:10.20944/preprints202605.1211.v1

multicomponent competence that includes understanding AI's capabilities and limitations, interpreting its responses, assessing risks, noticing distortions, and using the technology responsibly and critically. However, under conditions of generative AI, even this is insufficient if literacy remains merely functional. Users may know how to obtain a good result while failing to distinguish between a truly reliable answer and one that merely imitates cognitive completeness. For the purposes of this article, therefore, AI literacy is treated not as a general digital skill but as a transitional zone between operational competence and epistemic maturity. It is precisely here that knowing how to use AI and how to resist its pseudoconfidence are not the same.

The contemporary literature on AI literacy [17] is gradually shifting from instrumental mastery of the system to critical and postdigital literacy. However, we show that even broad AI literacy must be supplemented by a narrower but fundamental dimension: the ability to resist the pseudoconfident form of the answer.

This idea takes on a more rigorous form through the concept of epistemic vigilance developed in social epistemology and cognitive science [10,18,19]. Sperber and his coauthors showed that humans cannot exist without trusting external sources of information, but for that very reason, they must develop special mechanisms of epistemic vigilance: they must evaluate the reliability of a source, the plausibility of a message, and the degree of trust that can legitimately be granted to it. In the case of generative AI, the problem becomes sharper because the user encounters a new kind of "speaking source." The system speaks smoothly, quickly, and confidently but possesses neither personal responsibility nor a human experience of truth nor an inward relation to its own utterance. Epistemic vigilance under conditions of AI must therefore be directed not only toward the content of the answer but also toward the very effect of its rhetorical persuasiveness. The user must learn to recognize not only a possible error but also the subtler situation in which the form of the answer produces trust faster than its grounds warrant.

2.1. AI Answers as Quasitestimony and Assertive Speech Acts

For a more precise understanding of pseudoconfidence, we need to consider generative AI answers not only as informational objects but also as specific speech acts. In human communication, an assertion usually involves not only content but also the pragmatic commitment of the speaker. In ordinary cases, the person who asserts something presents himself or herself as someone who has grounds for what is said or who is prepared to bear responsibility for its justification. This is why an assertion differs from a supposition, a guess, a question, a hypothesis, or a conditional interpretation.

Generative AI disrupts this familiar connection between the form of assertion and epistemic responsibility. Its answer often has the grammar and style of confident assertion, but it does not have a speaking subject in the full sense of the word. The system does not testify to lived experience; does not possess its own responsibility for the truth of what is said; has no personal history of reliability; and does not internally distinguish, as a human does, between knowledge, assumption, reconstruction, and a stylistically probable continuation of text. For this reason, we can describe AI outputs as quasitestimony. That is, it is functionally similar to testimony because it communicates something to the user in the form of a ready-made assertion, but ontologically and epistemically, it differs from human testimony.

In terms of speech act theory [20], the problem consists of a mismatch between speech form and epistemic status. AI output often takes the form of an assertive speech act [21], although its content should often be presented as a hypothesis, a probabilistic judgment, a preliminary interpretation, or an answer indicating uncertainty.

This is precisely where we see the connection between pseudoconfidence and speech act theory. The problem is not only that generative AI may be wrong. The subtler risk is that the system often formulates its answer in the form of a confident assertion even when the epistemically more honest form would be a hypothesis, a probability, an assumption, a preliminary interpretation, or an answer that indicates the limits of uncertainty. In general, pseudoconfidence arises when a conditional or probabilistic epistemic status receives the form of unconditional assertion.

This is how we distinguish generative AI from many traditional sources of knowledge. A book, teacher, expert, or witness may also be mistaken and may also be rhetorically persuasive. However, in those cases, the user normally addresses a particular author, institution, genre, reputation, disciplinary context, or responsibility of the source. In the case of generative AI, the answer appears as a personalized, immediate, stylistically complete, and often impersonal utterance. It looks like the result of knowledge, but it is not always accompanied by a clear source trail, an indication of the degree of confidence, or the possibility of reconstructing the path by which it was obtained.

Therefore, epistemic vigilance toward generative AI is not a completely new virtue detached from ordinary practices of learning and trust. It continues a broader tradition of rational engagement with testimony, authority, and source reliability. However, it acquires a new form because the user encounters a source that imitates confident testimony but lacks the human responsibility of a witness. In this sense, critical AI literacy must include not only the technical ability to use the system but also sensitivity to the speech status of its answers. In particular, the user must be able to ask whether a given answer is an assertion, a hypothesis, a probabilistic reconstruction, an interpretation, or merely a rhetorically completed form of unfinished knowledge.

2.2. Illustrative Forms of Pseudoconfident AI Output

To clarify the concept of pseudoconfident knowledge, it is useful to consider several typical forms in which it may arise in the everyday use of generative AI. These examples should not be understood as a separate empirical corpus. Rather, they perform an analytical function and show how the rhetorical form of an answer may outrun its epistemic basis.

The first form is confident factual substitution. The user asks a question that presupposes a precise fact: a date, name, source, regulatory provision, or statistical value. The system answers coherently and confidently, but instead of a verified fact, it may provide close, plausible, or partially reconstructed information. In such a situation, the risk arises not only because of a possible error but also because the form of the answer does not signal that it has not been verified. The answer looks like complete knowledge, although in fact, it requires source-based checking.

The second form is smooth explanatory overreach. Generative AI may create a logically coherent explanation of a complex phenomenon even when the field itself contains competing interpretations, insufficient data, or methodological disputes. Here, pseudoconfidence manifests not as a simple falsehood but as excessive completeness of explanation. The user receives a feeling of understanding where the more appropriate epistemic form would be a cautious hypothesis, a map of possible interpretations, or an indication of the limits of knowledge.

The third form is the imitation of academic or expert justification. An answer may be constructed in the style of a scholarly explanation: with confident terminology, structured argumentation, and a citation-like tone. However, such a form still does not guarantee the existence of real sources, the correctness of citations, or the reliability of interpretation. This is why the external academic appearance of AI outputs should not automatically be equated with source reliability.

The fourth form is practical acceptance without epistemic distance. In this case, the problem does not arise so much in the answer itself as in the way it is used. The user transfers AI outputs into a text, decision, academic assignment, business document, or personal belief without comparison, editing, checking, or clarifying the degree of uncertainty. It is precisely here that AI-generated content begins to function as knowledge before it has obtained sufficient grounds for such a status.

Through these forms, we show that pseudoconfident knowledge is not reducible to hallucination or factual error. Rather, it arises in a broader situation in which confident speech form, the cognitive convenience of the answer, and insufficient epistemic distance on the part of the user converge into a single mechanism of premature trust.

2.3. Epistemic Vigilance as Calibrated Trust, Not Total Verification

We must clarify that epistemic vigilance should not be understood as a demand for the complete independent verification of every claim received from AI. Such a standard would be not only

 doi:10.20944/preprints202605.1211.v1

practically impossible but also philosophically mistaken. Much of human knowledge is formed through trust in testimonies, experts, books, institutions, educational practices, and other external sources. Therefore, rational cognition cannot be built on the total verification of every claim.

This logic continues the tradition of epistemic vigilance and the epistemology of testimony [22,23]. To a large extent, human knowledge depends on trust in external sources, but such trust is not passive. It requires evaluation of the source, context, speech genre, track record, and degree of risk of error.

Therefore, we understand epistemic vigilance as the capacity for calibrated trust. It presupposes that the user does not accept AI outputs passively but also does not have to restart verification from zero every time. The user's task is to align the degree of trust with the type of claim, the significance of the error, the context of use, and the availability of independent grounds. The higher the risk of error, normative significance, scientific cost, or practical consequences of a claim, the higher the degree of checking should be.

In this sense, the user's critical work includes several levels. At the minimal level, the user must understand that the confident form of an answer is not equal to its justification. At the intermediate level, the user must be able to compare the answer with alternative sources, ask for clarification of grounds, identify questionable points, and distinguish fact, interpretation, and hypothesis. At the high level, the user must develop the ability for epistemic self-correction: revising the degree of trust in a source on the basis of accumulated experience, detected errors, and the situation of use.

Consequently, epistemic maturity under conditions of generative AI is neither total distrust nor complete dependence on the machine. It is a practice of rational trust: trusting enough to use the answer as an intellectual resource while maintaining enough distance not to transform its rhetorical confidence into completed knowledge without grounds.

Generative AI thus creates an environment in which answers become faster, more accessible, and more cognitively convenient. This environment produces pseudoconfident knowledge—the content is accepted as reliable because of the way it is presented. The result is an epistemic gap between the operational use of the technology and the ability to evaluate its outputs critically. AI literacy points to the need to cultivate not only technical skills but also interpretive and critical skills, while the concept of epistemic vigilance explains precisely why these skills become central in the age of generative AI.

From this, we can formulate our central hypothesis: the principal line of difference does not run between AI users and nonusers but between different regimes of trust toward the machine's answer. That is why the focus of our empirical analysis is not only the frequency of AI use but also the ways in which users handle its persuasiveness.

3. Methodology

Overall, we interpret our study within the logic of a one-shot cross-sectional survey. Within this framework, our task was not only to describe user practices of interacting with generative AI but also to empirically examine one observable manifestation of the epistemic gap—that is, the gap between everyday involvement in AI use and the ability to critically evaluate the answers it produces. A cross-sectional survey makes it possible to capture the current distribution of attitudes, practices, and differences within a sample at a specific historical moment. This is particularly important in the rapidly changing field of generative technologies, where the norms of interaction themselves have not yet stabilized [24].

We emphasize that our empirical design has a diagnostic rather than a reductionist function. It is not intended to reduce the philosophical problem of pseudoconfident knowledge to a single statistical test. Its task is narrower: to examine whether operational involvement in AI use and critical epistemic orientation toward AI answers are distinct. Therefore, the frequency of AI use is treated not as a direct indicator of maturity but as an indicator of the operational integration of the technology into everyday practices. In turn, the strategy of handling AI outputs is treated as a closer indicator of the user's epistemic position.

 doi:10.20944/preprints202605.1211.v1

This formulation allows us to avoid the false conclusion that more frequent interactions with AI should automatically form critical thinking. This study tests precisely this intuitive but questionable hypothesis: whether repeated use of the technology leads to a more mature attitude toward its answers or whether epistemic resilience depends on other practices, such as comparison, checking, editing, and maintaining distance from machine confidence.

The empirical basis of the study consisted of 216 questionnaires collected through a self-administered online survey hosted on Google Forms (https://docs.google.com/forms/d/e/1FAIpQLSfT-e5WQTN9fxOjt7dpoBEc7svpMeQwQRvwTuURyecQHHUpLQ/viewform?usp=dialog). The survey was conducted among citizens of the Republic of Kazakhstan aged 18 and older.

The choice of this instrument was determined by the nature of the object under study. Because generative AI functions within digital everyday life, the online questionnaire methodologically corresponds to the very environment in which the practice under investigation exists. Moreover, our study does not claim strict probabilistic representativeness of the general population. The sample was voluntary and nonprobabilistic and should therefore be interpreted as an analytic cross-section of a specific user environment that is suitable primarily for identifying substantive patterns rather than for mechanically extrapolating findings to the entire population. This stance corresponds to an exploratory research logic: we are interested not in the “average AI user” as a statistical abstraction but in the structure of differences between modes of everyday engagement with machine-generated answers.

The methodological core of our study lies in distinguishing between two analytical levels.

The first level was interpreted as operational involvement in AI use. It captures how deeply generative systems are embedded in the respondent’s everyday practices.

The second level was interpreted as epistemic orientation—that is, the ability to distinguish rhetorical confidence from justified epistemic acceptance, rather than merely the rhetorical persuasiveness of an answer from its actual reliability.

In making this distinction, we relied on the contemporary understanding of AI literacy as a multicomponent competence that includes not only the functional skill of interacting with AI but also the critical ability to interpret its limitations, risks, and effects.

To formalize these two levels, we introduce five key variables:

$U_i \in \{0,1\}$ - whether AI is used,

$F_i \in \{1...,5\}$ - frequency of use,

$G_i \in \{1...,k\}$ - primary purpose of use,

$S_i \in \{1...,m\}$ - strategy of handling the AI answer,

$E_i \in \{0,1\}$ - epistemic outcome on the control question.

The first four variables describe the everyday mode of using the technology. The fifth serves as a minimal indicator of epistemic vigilance. We deliberately designed a control item to test whether a respondent could recognize a fundamental substitution error, since a confident and detailed form of presentation does not guarantee epistemic reliability or warranted acceptance. The variable E_i therefore captured not an attitude toward AI as such but a more fundamental cognitive orientation toward plausibility and reliability. Conceptually, this approach aligns with the idea of epistemic vigilance, according to which human cognition inevitably depends on external sources, and for precisely that reason, special mechanisms of critical control over communicated content are needed.

Data processing included cleaning the dataset, standardizing the categorical responses, and analytically recoding them. We then used two complementary levels of analysis.

At the first level, we applied descriptive statistics and calculated absolute frequencies, shares, and percentage distributions for all key variables. Formally, the percentage distribution for category j was defined as follows:

$$p_j = \frac{n_j}{N} \times 100 \qquad (1)$$

where n_j is the number of observations in category j and N is the total sample size.

 doi:10.20944/preprints202605.1211.v1

At the second level, we used contingency tables and a comparative analysis of categorical variables. This choice was crucial because our research hypothesis concerned not mean values but the structure of differences. Specifically, we were interested in how an epistemically correct answer was related to the frequency of AI use, its purposes, and especially the strategy used to handle the machine's answer. Accordingly, we analyzed conditional distributions of the following form:

$$P(E=1\mid F=f), P(E=1\mid S=s), P(E=1\mid G=g) \tag{2}$$

To test the statistical significance of differences between categorical variables, we used Pearson's chi-square test, the standard procedure for detecting associations between nominal features:

$$\chi^2=\sum_{r=1}^{R}\sum_{c=1}^{C}\frac{(O_{rc}-E_{rc})^2}{E_{rc}} \tag{3}$$

where O_{rc} is the observed frequency in a cell of the contingency table and E_{rc} is the expected frequency under the hypothesis of independence.

This test enabled us to determine whether differences between categories were statistically significant or could be explained by a random distribution. To assess the strength of the observed association, we used Cramer's V coefficient, which is especially appropriate for categorical data because it allows one to move from the mere fact of significance to an estimate of the intensity of association:

$$V=\sqrt{\frac{\chi^2}{N\cdot min(R-1,C-1)}} \tag{4}$$

Within the logic of this study, this was especially important: even when statistical significance is present, we are interested not simply in association as such but in which associations are epistemically meaningful. In other words, the task was to determine what better explains a user's epistemic resilience—the intensity of AI use or the strategy of handling its answer.

Accordingly, we designed the methodology as a test of one key hypothesis: the everyday integration of AI into a respondent's practices is not identical to the epistemic maturity of their interaction with machine-generated answers.

In terms of the analytical model, this assumption can be expressed as follows:

$$O_i\uparrow \;\not\Rightarrow\; C_i\uparrow \tag{5}$$

where O_i denotes operational involvement in AI practices and C_i denotes critical epistemic literacy. All the results of our study support, in effect, an empirical test of this inequality.

4. Results

The empirical material obtained supports the central thesis of the study: In the everyday use of generative AI, what proves decisive is not the level of technological involvement as such but the mode of epistemic engagement with the machine's answer.

Our data therefore capture not only the growth of AI use but also a more consequential stratification—specifically, between operational adaptation to the technology and the ability to resist pseudoconfident knowledge.

In the first block of results, we show that generative AI has already become embedded in the respondents' everyday cognitive infrastructure.

In the sample of 216 participants, 90.4% use AI, while only 9.6% do not (Table 1). The frequency distribution indicates not episodic but stable incorporation of AI into everyday practice: 36.6% use it daily, 22.8% use it 3–5 times a week, and 20.8% use it 1–2 times a week. Thus, no fewer than four-fifths of the sample use AI at least weekly. At this level alone, it becomes clear that the object of study is not a peripheral innovation but a normalized form of digital mediation in working with knowledge.

doi:10.20944/preprints202605.1211.v1

Table 1. Baseline distributions: AI use, frequency, purposes, and epistemic indicators.

Table 1A. AI use

Category	n	%
Yes	191	90.5
No	20	9.5

Valid $N = 211$

Table 1B. Frequency of AI use

Category	n	%
Daily	75	36.8
3–5 times a week	47	23.0
1–2 times a week	42	20.6
1–3 times a month	21	10.3
Less than once a month	19	9.3

Valid $N = 204$

Table 1C. Main purpose of AI use

Category	n	%
Searching for information/simple explanations	86	41.3
Writing/editing text (messages, documents, resumes, etc.)	39	18.8
Study/self-learning	38	18.3
Work/business (plans, reports, ideas, analysis)	26	12.5
Entertainment (content, ideas, games)	10	4.8
Other	9	4.3

Valid $N = 208$

Table 1D. Response to the epistemic control item

Response option	Epistemic classification	n	%
B) Confidence and detail do not guarantee truth; sources and proportionate verification are needed	Epistemically correct response	124	60.2
D) Cannot say	Noncommittal response	30	14.6
C) True, if the text is "logical"	Epistemically incorrect response	26	12.6
A) A confident and detailed text is usually true	Epistemically incorrect response	26	12.6

Note. Percentages were calculated on the basis of valid responses for each item; therefore, valid N differed across subgroups. The epistemic control item was designed to test whether respondents distinguished rhetorical confidence and textual coherence from warranted epistemic acceptance. Option B was coded as the epistemically correct response. Options A and C were coded as epistemically incorrect, whereas option D was coded as a noncommittal response.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

The baseline descriptive distributions of the sample, including AI use, frequency of use, primary purposes of use, and responses to the epistemology control item, are presented in Embedded Table 1.

The social meaning of normalization-that is, the normalization of digital mediation in work with knowledge-is revealed not through frequency alone but through the functional orientation of use. The most common purpose is to search for information and obtain "simple explanations" (41.7% of respondents). This is followed by writing and editing text (18.9%), study and self-learning (17.5%), and work/business (12.6%) (Table 1).

Generative AI is therefore used first and foremost as a mechanism for reducing cognitive complexity. It does not simply provide information; it converts uncertainty into a rapidly digestible form of answer. This is precisely where the central problem of our study emerges: the more strongly AI mediates between question and understanding, the greater the likelihood that rhetorically convenient content will be accepted as already verified knowledge.

Our assumption finds direct empirical support in the results of the control epistemic indicator. Only 59.8% of the respondents gave epistemically correct answers to the question, distinguishing rhetorical confidence from warranted epistemic acceptance. Accordingly, approximately 40% either failed to recognize the error involved in such a substitution or took an indeterminate position (Table 1). For us, this result is fundamental because it means that high everyday integration of AI is not accompanied by an equivalent level of epistemic consolidation. AI has already been mastered as a convenient tool, but it has not yet been mastered to the same extent as a source that requires risk-sensitive checking and proportionate verification.

The decisive line of difference emerges when we compare the epistemic indicator with the strategy used to handle AI answers (Table 2).

Table 2. Answer-handling strategy and epistemic correctness.

Strategy for handling AI-generated answers	Incorrect or noncommittal response, n	Epistemically correct response, n	Correct response, %
I compare the answers with those of other sources (search engines, official websites, people).	28	60	68.2
I edit the answer before using it.	20	38	65.5
I verify the answer in practice (through an example, calculation, or additional test).	12	19	61.3
I use the answer as it is.	15	6	28.6
I rarely use AI answers because I do not trust them.	7	1	12.5

Valid *N* = 206.

Table 2 compares strategies for handling AI answers with the results of the epistemic control item. Its purpose is to show whether epistemic correctness is associated with a passive or active mode of handling AI output.

Note. "Epistemically correct response" means option B in the epistemic control item. "Incorrect or noncommittal response" combines incorrect answers and the "Cannot say" option. The percentages show the share of epistemically correct answers within each strategy for handling AI answers. Valid *N* = 206.

The table shows that epistemic correctness is higher among respondents who compare, edit, or practically verify AI answers and lower among those who use answers as they are. These findings support our conclusion that epistemic resilience is not strongly associated with AI use but with the way in which AI output is handled.

Among the respondents who compared their answers with those of other sources, the share of epistemically correct responses reached 67.8%. Among those who slightly edited the answer and then used it, 64.9% were correct. Among those who verify it in practice, it is 61.3%. In contrast, in the group that uses the answer as is, the share of correct responses falls to 28.6%. Among respondents who rarely use AI because they distrust it, it decreases to 12.5%.

Thus, the relevant differences do not run between AI "users" and "nonusers" but between different regimes of trust toward the machine's answer. When an external procedure of source-sensitive checking and proportionate verification is retained, epistemic resilience is higher. When the answer is assimilated without distance, the vulnerability to pseudoconfident knowledge increases.

Statistically, this relationship is confirmed by Pearson's criterion. Specifically, the association between the answer-handling strategy and epistemic correctness is significant and of moderate strength.

$$\chi^2(4) = 18.939, p = 0.0008, V = 0.305 \quad (6)$$

We should stress that these results constitute the empirical core of the article. Here, we show that critical literacy in the environment of generative AI is determined less by involvement in the technology itself than by the cognitive mode in which its products are handled. Pseudoconfident knowledge, therefore, can be interpreted not only as a property of the answer but also as a property of one's relation to the answer, arising where the form of persuasiveness begins to replace the practice of proportionate verification.

Moreover, the relationship between the frequency of AI use and epistemic correctness does not prove statistically significant (Table 3).

Table 3. AI use frequency, purpose of use, and epistemic correctness.

Table 3A. Frequency of AI use and epistemic correctness

Frequency of AI use	Incorrect or noncommittal response, n	Epistemically correct response, n	Correct response, %
1–2 times a week	16	26	61.9
1–3 times a month	8	12	60.0
3–5 times a week	12	35	74.5
Daily	32	43	57.3
Less than once a month	10	8	44.4

Valid $N = 202$.

 doi:10.20944/preprints202605.1211.v1

Table 3B. Main purpose of AI use and epistemic correctness

Main purpose of AI use	Incorrect or noncommittal response, n	Epistemically correct response, n	Correct response, %
Entertainment (content, ideas, games)	2	8	80.0
Other	5	4	44.4
Searching for information/simple explanations	42	42	50.0
Study/self-learning	14	24	63.2
Work/business (plans, reports, ideas, analysis)	7	19	73.1
Writing/editing text (messages, documents, resumes, etc.)	12	27	69.2

Valid *N* = 206.

Note. "Epistemically correct response" means option B in the epistemic control item. "Incorrect or noncommittal response" combines incorrect answers and the "Cannot say" option. Table 3A shows the distribution by frequency of AI use, while Table 3B shows the distribution by the main purpose of use. Percentages are calculated within each row.

Unlike the strategy for handling the answer, the frequency of AI use does not show a statistically significant association with epistemic correctness. This difference is central to our argument: operational habituation to AI and epistemic vigilance toward AI answers should not be treated as the same phenomenon.

Thus, although the highest level of correct responses is observed among respondents who use AI 3–5 times a week (73.9%), the figure is lower among daily users (56.8%), and the overall test does not confirm a stable direct relationship:

$$\chi^2(4) = 5.881, p = 0.208, V = 0.171 \quad (7)$$

Analytically, this is especially important because it undermines the intuitive but false model according to which more frequent use of a technology automatically leads to a more mature understanding of it.

Our data point in the opposite direction. Operational intensity and epistemic reflexivity belong to different orders. A user may be deeply integrated into AI practices while still remaining vulnerable to cognitively persuasive yet insufficiently verified answers.

We also found an additional, albeit weaker, difference by purpose of AI use. The highest shares of epistemically correct responses appear among those using AI for work/business (73.1%) and for writing/editing text (69.2%). Among those who turn to AI primarily for information search and "simple explanations," by contrast, the correct response rate is only 50.0% (Table 3).

Formally, this relationship does not reach the conventional threshold of significance,

$$\chi^2(5) = 9.311, p = 0.097, V = 0.214 \quad (8)$$

Nevertheless, the substantive tendency is evident. Specifically, in the zone where AI is used as a tool for rapid cognitive relief, epistemic resilience is weaker. This is particularly important because that scenario is also the most widespread.

The integrated meaning of these results is summarized in Figure 1, which shows that the central fault line runs between two analytical levels: operational AI competence and critical epistemic literacy. The main conclusion of the study is presented in Figure 1: increased everyday involvement in AI practices does not automatically lead to greater critical ability to assess the status of an answer. In formal terms, this is expressed as follows (Formula 5):

$$O_i \uparrow \;\not\Rightarrow\; C_i \uparrow$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

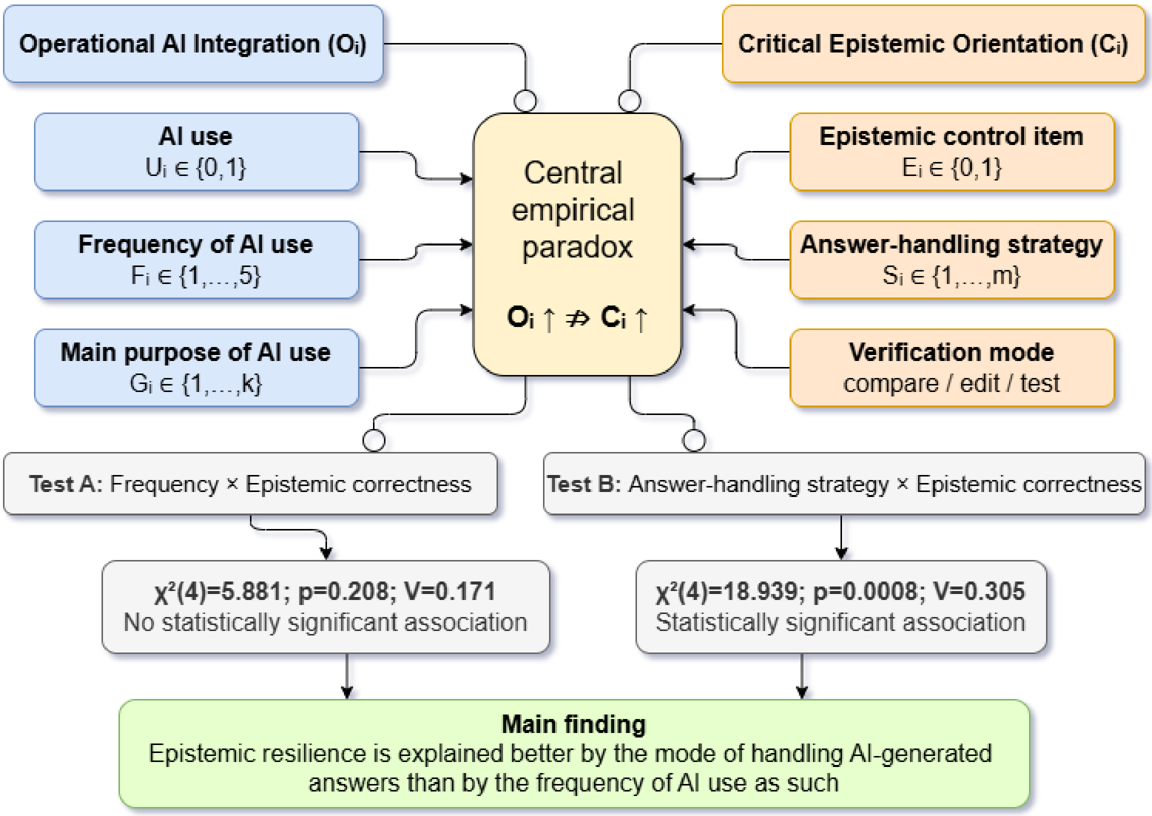

Figure 1. Integrated block diagram of the empirical model of the pseudoconfident knowledge paradox.

Figure 1 presents an integrated conceptual-empirical schema of the pseudoconfident knowledge paradox.

The upper-left block of Figure 1, Operational AI Integration (O_i), represents the behavioral dimension of everyday AI use and is broken down into three subordinate components:

AI use ($U_i \in \{0,1\}$),

Frequency of AI use ($F_i \in \{1...,5\}$), *and*

Primary purpose of AI use ($G_i \in \{1...,k\}$).

Taken together, these elements capture the degree to which generative AI is embedded in routine cognitive practice.

The upper-right block, critical epistemic orientation (C_i), represents the epistemic dimension of the study. It is organized around three interrelated components: the epistemic control item ($E_i \in \{0,1\}$), the answer-handling strategy ($S_i \in \{1...,m\}$), and the mode of verification (comparison, editing, testing). These elements make it possible to determine whether respondents consume AI answers as a finished product or subject them to procedures of checking, editing, and epistemic distancing.

At the center of the figure is the principal theoretical and empirical formula: the central empirical paradox, $O_i \uparrow \nRightarrow C_i \uparrow$. That is, high everyday integration of AI does not imply a proportional increase in the ability to distinguish rhetorical confidence from warranted epistemic acceptance. This central node functions as the interpretive hinge of the entire model because it shows that behavioral immersion in AI does not automatically generate epistemic vigilance.

The lower analytical layer contains two empirical tests.

The left block, Test A: Frequency × Epistemic Correctness, leads to the statistical result $\chi^2(4)=5.881$; $p=0.208$; $V=0.171$, which we interpret as indicating that no statistically significant association was found.

The right block, Test B: Answer-Handling Strategy × Epistemic Correctness, leads to the result $\chi^2(4)=18.939$; $p=0.0008$; $V=0.305$, which is interpreted as indicating a statistically significant association.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

At the bottom of the figure, the integrated conclusion is formulated as follows.

Main finding: Endemic resilience is better explained by the way in which AI-generated answers are handled than by the frequency of AI use.

Two additional interpretive blocks reinforce the general logic of the scheme. On the left, the descriptive layer is reduced to the formula: AI already functions as a routine cognitive infrastructure. On the right, the epistemic layer is expressed through the formula: pseudoconfident knowledge arises where plausibility replaces verification.

In summary, Figure 1 visualizes the difference between operational AI integration (O_i) and critical epistemic orientation (C_i). It also shows that the central paradox of the study lies in the nonidentity between everyday AI use and epistemic resilience. Whereas the relationship between AI use frequency and epistemic correctness is statistically insignificant, the relationship between the answer-handling strategy and epistemic correctness is statistically significant.

Therefore, Figure 1 synthesizes the main empirical conclusion of our study: epistemic resilience is explained better by the way AI-generated answers are handled than by the frequency of AI use as such.

Overall, the scientific value of Figure 1 lies in the fact that it translates the results of the study from a set of separate descriptive and statistical observations into an integrated analytical model. The figure shows that the empirical problem addressed in the article cannot be reduced to technological acceptance or rejection of AI. The broad and routine use of generative AI is already an accomplished social fact. What remains unresolved is something else: the epistemic regime through which users relate to machine-generated answers. In this sense, the figure marks a shift from a sociology of access to a sociology of epistemic mediation.

Its principal contribution is that it dismantles the naive linear assumption that more frequent AI use must automatically lead to a more mature and critical mode of engagement. The left side of the figure shows the expansion of operational integration, whereas the right side demonstrates that epistemic orientation follows a different logic.

For this reason, the statistical asymmetry between Test A and Test B should be understood as the empirical diagnostic of our philosophical argument. The absence of a significant relationship between the frequency of use and epistemic correctness means that simply increasing contact with AI does not necessarily produce epistemic competence. In contrast, the presence of a significant relationship between the answer-handling strategy and epistemic correctness shows that what matters is not how often users turn to AI but how they cognitively process the result they receive from it.

We believe that this distinction is of fundamental importance for the concept of pseudoconfident knowledge. The figure shows that pseudoconfident knowledge is not merely an internal property of an AI-generated answer. Rather, it arises at the boundary between machine plausibility and the human mode of reception. An answer becomes pseudoconfident knowledge when its rhetorical smoothness, coherence, and convenience begin to be perceived as a substitute for warranted epistemic acceptance. In this sense, the figure shifts the analytical focus from the ontology of machine error to the sociology of epistemic trust.

We also believe that the figure has important methodological significance. These findings show that the most powerful analytical variable is not AI use as such or even its frequency but precisely the strategy of handling the answer. It follows that future studies of AI literacy must move beyond crude indicators of familiarity with the technology and the intensity of its use and instead focus on practices of verification, comparative checking, editorial intervention, and practical testing of answers.

Accordingly, the decisive line of difference in the age of generative AI may run not between users and nonusers but between different regimes of trust.

Finally, the figure also has practical significance. It shows that educational and policy measures aimed merely at increasing AI literacy in the narrow instrumental sense may prove insufficient or even misleading if they enhance operational dexterity without strengthening epistemic vigilance at

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 May 2026 doi:10.20944/preprints202605.1211.v1

the same time. The figure therefore supports the broader conclusion of the article: the central problem of the generative AI era lies not only in expanding access to AI systems but also in cultivating a culture capable of resisting pseudoconfident knowledge.

In other words, society is mastering generative AI faster than it is developing the skills needed to resist pseudoconfident knowledge. That is why the key explanatory factor is not intensity of use but the presence or absence of a verification regime-comparison, editing, and external checking.

Our results therefore allow us to formulate three interrelated empirical conclusions on the basis of the authors' analysis of the sociological survey dataset ($N = 216$).

First, generative AI has already become a normal part of everyday cognitive infrastructure.

Second, its principal function in the environment under study is the reduction of complexity and the rapid production of an intelligible answer.

Third, this is precisely where the main epistemic risk arises, which we formulate as follows: the habit of using AI is not identical to the habit of checking it.

Consequently, the principal divide does not run between use and nonuse of the technology but between different regimes of trust toward machine persuasiveness.

This is, in fact, the empirically confirmed paradox of pseudoconfident knowledge.

5. Discussion

The principal result of our study is that in the everyday use of generative AI, the decisive factor is not the frequency of contact with the technology itself but the strategy by which its answer is handled.

This result is important to us because it shifts the problem of AI literacy from the field of simple instrumental competence to the field of epistemic responsibility. A user may be technically experienced, obtain useful answers quickly, and effectively integrate AI into everyday work. However, this does not mean that the user has developed the ability to evaluate the status of the answer received. Mature use of generative AI therefore begins not with the frequency of turning to the system but with the mode of working with its output: comparison, clarification, checking, editing, and the preservation of methodological doubt.

We regard this conclusion as fundamental because it overturns one of the most widespread yet methodologically naive intuitions of the digital age, namely, that the intensity of technology use will gradually produce maturity of understanding on its own.

Our data suggest a different picture. The frequent use of AI does testify to its social normalization, but it does not guarantee a critical epistemic stance. Operational adaptation and epistemic maturity therefore belong to different orders. The former is associated with convenience, speed, and routinization of use; the latter is associated with a regime of checking, distance, and the capacity to preserve the question of the answer's truth status. This distinction constitutes the core of the pseudoconfident knowledge paradox.

Why does the answer-handling strategy matter more than the frequency of use?

Because generative AI does not simply deliver information, it produces a cognitively finished text that already has a plausibility effect. Users are confronted not with fragments of data that must be assembled and interpreted but with a ready-made discursive object already optimized to be perceived as a meaningful answer. In such a situation, what matters is not how often a person turns to the system but whether they preserve an external epistemic distance from the answer.

When users compare, edit, test, and proportionately verify, they preserve the distinction between the answer received and the knowledge actually attained. When the answer is accepted as given, the rhetorical confidence of the text begins to replace its justification. Current research on AI literacy confirms precisely this shift: the issue is not merely the ability to use the system but also the ability to recognize its limitations, interpret its outputs critically, and understand that a persuasive answer is not identical to reliable knowledge.

This problem is also connected with the broader literature on automation bias [25–27]. Automated systems can reduce the cognitive load, but at the same time, they can provoke excessive

 doi:10.20944/preprints202605.1211.v1

trust in automatically produced outputs. In the context of generative AI, automation bias acquires a linguistic form: the user trusts not only the machine as an instrument but also the smoothness of the text as if it were already complete knowledge.

In this sense, our findings allow us to refine the very concept of AI literacy. A substantial portion of the literature has already moved beyond a narrow understanding of AI literacy as knowledge about algorithms or a set of technical skills. The prevailing view is that genuine AI literacy includes critical judgment, interpretive caution, and an awareness of the socioepistemic effects of generative systems.

Our results, however, make it possible to take one step further. They show that even this broader thesis is insufficient unless one distinguishes between operational literacy and epistemic maturity or verification-oriented AI literacy. The former answers the question of whether a person knows how to use AI. The latter answers the question of whether that person has the ability to treat AI output as an epistemic proposal rather than as finished knowledge. It is the second of these that has now become critically important.

5.1. Prompt-Based Triangulation as an Epistemic Strategy

Prompting deserves separate attention as an independent epistemic factor. In everyday understanding, a prompt is often treated as a technical instruction that allows the user to obtain a more convenient or more complete answer. Under conditions of generative AI, however, the prompt also performs a deeper function. It frames how exactly the system will produce, limit, and format the answer.

Therefore, critical AI literacy should include not only checking the answer already obtained but also the ability to ask questions that reduce the risk of a pseudoconfident form. The user can reformulate the same question in several ways, ask the system to indicate possible limitations of the answer, propose alternative explanations, distinguish fact from interpretation, indicate the degree of uncertainty, or show which claims require external verification.

We may call this practice prompt-based triangulation. Its meaning is not to obtain one maximally smooth answer from AI but to compare several response trajectories and check their stability. If different formulations of the question lead to substantially divergent answers, this itself is an epistemic signal: the user is dealing not with finished knowledge but with a zone of uncertainty that requires additional checking.

Prompt-based triangulation differs from ordinary prompt engineering [28]. Its goal is not only to improve the quality of a single answer but also to test the stability of the result through different formulations of the question, requests for grounds, alternatives, and indications of uncertainty.

Prompt-based triangulation is especially important for us because pseudoconfidence often arises from an isolated answer. One coherent text may look complete and sufficient. Several answers obtained through different formulations, requests for grounds, and requests for alternatives disrupt the illusion of finality. We therefore interpret prompting not only as a tool of productivity but also as a practice of epistemic distancing.

A direct conclusion for education follows from this. If educational systems understand preparation for the AI age primarily as training in the effective use of generative tools-writing prompts, accelerating search, and automating text production-they will inevitably strengthen only the operational side of interaction with AI.

However, this does not solve the principal problem because pseudoconfident knowledge arises not from a lack of access but from a lack of epistemic discipline. Educational policy must therefore shift its emphasis from mere tool acquisition to the formation of habits of proportionate verification, rational trust, linguistic sensitivity, and epistemic self-correction: comparison with independent sources, identification of unstated grounds behind an answer, recognition of spurious logicality, and the ability to distinguish explanatory smoothness from evidential reliability. International recommendations move in precisely this direction. UNESCO, for example, stresses that working with

 doi:10.20944/preprints202605.1211.v1

generative AI in education requires not only functional efficiency but also critical thinking, human responsibility, and the ability to question the ready-made result.

From this, we must draw important educational clarification. The aim of AI literacy cannot be the formation of total distrust toward generative AI. Such a position would be as poor as the uncritical acceptance of its answers. Education should cultivate not the refusal of trust but the ability to calibrate it. The learner must understand when AI output may be used as a preliminary hint, when it must be compared with independent sources, when alternative explanations should be requested, and when strict checking of a fact, reference, number, or normative claim is necessary.

Therefore, epistemic maturity in the age of AI includes not only technical competence but also deeper intellectual virtues: conscientiousness toward truth, cautious skepticism, epistemic humility, linguistic sensitivity, the ability to distinguish assertion from assumption, and self-trust as the capacity not to dissolve one's own judgment in machine confidence while also correcting it when confronted with more reliable grounds.

5.2. Educational Implications: Epistemic Maturity, Judgment, and Linguistic Sensitivity

We do not reduce the educational significance of these results to the need for separate instructions on the "*proper use of AI*." Such an interpretation would be too narrow. The problem of pseudoconfident knowledge shows that preparation for the era of generative AI must be interpreted as part of a broader task: the formation of epistemic maturity.

By epistemic maturity, therefore, we mean not a special skill of working with one technology but the subject's ability to deal rationally with external sources of knowledge under conditions of uncertainty. This capacity includes critical judgment, epistemic conscientiousness, cautious skepticism, readiness to verify, sensitivity to source authority, and the ability to adjust one's trust on the basis of experience. In relation to generative AI, such maturity acquires a special form because the user encounters a source that quickly produces coherent and confident answers but does not always display the grounds, degree of uncertainty, or source of the assertion.

Education must therefore develop not only AI literacy but also a deeper culture of epistemic judgment. Learners should be taught to distinguish not only correct and incorrect answers but also different speech statuses of utterances: assertion, assumption, hypothesis, interpretation, speculation, generalization, and probabilistic reconstruction. This linguistic sensitivity becomes one of the central competencies under conditions of generative AI. The user must notice when a text speaks too confidently, when it conceals uncertainty, and when an explanation appears complete before the grounds for it have actually been presented.

In this sense, critical AI literacy should be connected with virtue epistemology. We are not speaking of cultivating total distrust toward AI but of forming intellectual virtues: concern for truth and accuracy, fair-mindedness, epistemic humility, the capacity for doubt, the willingness to seek independent grounds, and the ability to preserve one's own judgment in the face of machine confidence. We believe that self-trust is especially important: not self-assurance but the ability to maintain one's own responsibility for accepting or not accepting an answer.

Consequently, the educational task is not to teach learners merely to obtain answers from AI more quickly. It is to teach them to live in an environment where a ready-made answer appears before an understanding of its status has been formed. In such an environment, the mature user is not the one who uses AI more often but the one who can transform a machine answer from an object of passive trust into an object of critical, linguistically sensitive, and rationally calibrated judgment.

The same applies to digital literacy more broadly. Traditional digital literacy was formed under conditions of information overload and therefore taught people to search for, filter, and compare sources. Generative AI changes the very object of that literacy. Users increasingly receive not a plurality of sources but one smooth answer that removes the need for initial selection. Digital literacy must therefore evolve toward an epistemic literacy appropriate to the generative age. Its central question is no longer "*Where can I find information*?" but rather "*In what sense does this persuasive answer* *deserve trust*?" Here, a new societal scale emerges: how to teach users not only to work with AI but also to resist the excessive ease of trusting it.

Finally, our results are significant for the epistemology of AI itself. They show that the problem of generative systems is not exhausted by questions of errors, hallucinations, or the model's technical limitations; however, those limitations undoubtedly are fundamental. Contemporary research convincingly demonstrates that large language models can produce semantically coherent yet false answers and that this problem is structural rather than accidental. The socioepistemic novelty of the present situation lies elsewhere: even when users know that error is possible, they may still succumb to the effect of the persuasive form. The principal philosophical question therefore shifts from "*Can AI be wrong*?" to "*What regimes of trust does AI make possible as a rhetorical machine*?" In this sense, generative AI should be understood not only as a tool for the computational production of text but also as an environment that reorganizes the human relation to justification, source, and truth.

That is precisely why the concept of pseudoconfident knowledge seems methodologically productive to us. It makes it possible to describe not only erroneous content but also a special cognitive situation in which persuasiveness precedes verification and plausibility begins to function as a substitute for reliability. Our study shows that this situation is socially heterogeneous because it depends less on the intensity of AI use than on the regime by which users organize their trust. Future AI research should therefore be concerned not only with measuring adoption, usage frequency, or perceived usefulness but also with developing a typology of strategies for epistemically handling answers. We believe that this is where the principal line of difference lies between the enhancement of thought and its subtle erosion.

5.3. Systemic Feedback and Long-Term Epistemic Degradation

We note that pseudoconfident knowledge has not only an individual but also a systemic dimension. At the individual level, the risk consists of the user accepting rhetorically persuasive AI outputs as already justified knowledge. In the digital environment, however, such answers rarely involve only private cognitive events. They may be transferred into academic assignments, reports, blogs, social media, reference materials, presentations, and other forms of public or semipublic text.

In this case, the pseudoconfident answer begins to participate in a wider cycle of information circulation. If factually inaccurate or weakly justified AI-generated outputs are disseminated at scale and then indexed, cited, retold, or reused, they may gradually change the very information environment in which future users and systems search for grounds for answers. We therefore believe that the epistemic risk ceases to be only an error of an individual user and becomes a cumulative environmental risk.

This problem is connected with broader discussions of synthetic data and model collapse. Studies have shown that the recursive use of generated data in model training may lead to quality degradation, loss of rare elements of the original distribution, and narrowing of informational diversity [29]. For our argument, this matters not because every erroneous AI output will automatically enter a future training corpus but because the mass circulation of unverified machine content creates the possibility of epistemic feedback. That is, once weak knowledge is accepted as reliable, it may return to the system as part of a new informational norm.

It is important not to exaggerate this risk: not every erroneous AI output automatically contaminates future training datasets. However, the mass circulation of unverified machine content may create conditions for epistemic feedback, especially if such texts are indexed, copied, and reused in future information systems [30].

Therefore, we understand the epistemic gap not only as a difference between AI use and the critical capacity of an individual user. It also has a structural dimension. When pseudoconfident answers are disseminated, copied, and incorporated into new texts, they may amplify informational drift and reduce the quality of the collective epistemic environment. In this sense, a culture of checking, calibrated trust, and responsible handling of AI outputs is important not only for individual literacy but also for preserving the quality of the shared space of knowledge.

 doi:10.20944/preprints202605.1211.v1

The significance of our findings thus extends beyond the local survey. They point to a broader transformation of digital modernity: society is entering a phase in which access to answers grows faster than the capacity to distinguish their epistemic status. Under these conditions, the central task is not only to democratize access to AI but also to institutionally, educationally, and culturally cultivate practices of resistance to pseudoconfident knowledge.

In other words, the future of AI literacy is decided not by how quickly people learn to write the right prompts but by whether they learn to preserve the right to doubt in the face of an answer that is too smooth.

5.4. Limitations and Future Research

Our study has several limitations.

First, the sample was voluntary and nonrepresentative. Therefore, the results should not be interpreted as a statistical description of the entire population of AI users.

Second, the design of our study was cross-sectional. This does not allow us to make causal claims about whether particular strategies for handling AI output form a higher level of epistemic resilience or whether already more epistemically attentive users more often choose strategies of comparison, editing, and checking.

Third, the epistemic control item was a minimal indicator of the distinction between rhetorical confidence and warranted acceptance of the answer. It does not exhaust the whole structure of epistemic maturity. Future studies should include more complex experimental tasks, analyses of real AI-dialogs, measurements of prompting strategies, and separate scales for source reliability, claim verification, linguistic sensitivity, and calibrated trust.

In the survey, we did not measure prompting strategies as a separate variable. Therefore, prompt-based triangulation is introduced here as a conceptual and practical implication, not as an empirically tested predictor. Future studies should separately examine whether users who apply multiprompt comparison, uncertainty prompts, and justification prompts demonstrate a higher level of epistemic resilience.

Finally, we treat the empirical data as diagnostic material rather than as exhaustive proof of the philosophical concept. Its main contribution consists of connecting the conceptual framework with the observable distinction between operational AI integration and critical epistemic orientation. Further research may test this framework on larger samples in different educational contexts and using mixed-methods designs.

6. Conclusion

Our study allows us to conclude that generative artificial intelligence changes not only everyday practices of searching for and processing information but also the very epistemic regime through which human beings interact with answers. At the center of this change lies the phenomenon we have termed pseudoconfident knowledge: content perceived as reliable before its reliability has actually been established.

The empirical results revealed that the principal divide runs neither between AI users and nonusers nor between more and less frequent users of the technology but between different strategies for handling machine-generated answers. The cognitive mode in which an answer is processed-comparison, editing, verification, or uncritical acceptance-proves more significantly than the intensity of AI use itself does.

This confirms the main thesis of our study: the everyday integration of generative AI is not identical to the epistemic maturity of interacting with it.

Theoretically, this means that generative AI should be understood not only as a tool for accelerated access to information but also as a new infrastructure for redistributing trust, plausibility, and knowledge. In practice, this means that education and digital literacy must shift their emphasis from simple mastery of AI tools toward the formation of a culture of verification, doubt, and epistemic vigilance.

 doi:10.20944/preprints202605.1211.v1

Thus, the pseudoconfident knowledge paradox lies in society mastering AI faster than it is developing the skills needed to resist its persuasiveness. That is why the central task of the near future is not only to broaden access to AI but also to develop the ability to distinguish between a convenient answer and justified knowledge.

Our main practical conclusion is that society should not equate the spread of AI with the growth of epistemic maturity. Instrumental literacy allows the user to obtain answers more quickly. Only epistemic responsibility, however, prevents these answers from turning into prematurely recognized knowledge. Therefore, the future culture of AI literacy must be a culture of checking, rational trust, and methodological doubt.

In a broader sense, the results of the article also point to a possible connection with the logic of SRL (scientific readiness level) [31,32]. If generative AI is being increasingly integrated into scientific, educational, and analytical practice, then readiness to work with it cannot be assessed solely in terms of the technical command of the tool. The structure of scientific readiness must also include epistemic readiness: the capacity to distinguish persuasiveness from reliability, to assess source support, proportionately verify the grounds of an answer and to maintain critical distance from generated knowledge.

In this respect, the pseudoconfident knowledge paradox shows that a high level of digital or instrumental involvement does not yet imply a high level of scientific maturity. Accordingly, SRL may in the future be supplemented by a dedicated dimension capturing the readiness of a subject, project, or research environment for the epistemically responsible use of AI.

Author Contributions: Removed for peer review.

Funding: Removed for peer review.

Institutional Review Board Statement: Formal approval from an Institutional Review Board or ethics committee was not obtained for this study. The study was conducted as an anonymous, voluntary, non-interventional online survey of adult respondents. It did not involve biomedical, clinical, therapeutic, diagnostic, experimental, or risk-related procedures. No biological materials, clinical or epidemiological data, medical information, sensitive personal data, or personally identifiable information were collected. Under the applicable national biomedical research regulations of the Republic of Kazakhstan, mandatory prior review by bioethics committees is regulated primarily in relation to biomedical research. Since the present study was not biomedical research and posed no more than minimal risk to participants, formal approval by a biomedical ethics committee was not considered applicable to this type of study. The survey data were processed anonymously and reported only in aggregated form.

Informed Consent Statement: Electronic informed consent was obtained from all participants involved in the study. Prior to completing the questionnaire, participants were informed about the purpose of the study, the voluntary and anonymous nature of participation, the absence of collection of personally identifying data, and the use of the collected data for scientific purposes. Completion and submission of the questionnaire were treated as informed consent to participate. Participants were free to decline to answer any question and to discontinue participation at any time before submitting the questionnaire.

Data Availability Statement: The aggregated data supporting the findings of this study are presented in the article. Because the survey was conducted anonymously and no personally identifiable data were collected, only the aggregated results are reported.

Acknowledgments: No LLMs were used in this study. The questions for the sociological questionnaires were formulated manually. The questionnaires themselves were manually created in Google Forms templates. The survey responses were collected in a real student environment in CSV format. The survey data were processed and analyzed entirely manually in SPSS, using the majority of its technical and methodological capabilities. Thus, the structure, process, and methodology of our work neither assumed nor required the participation of LLMs.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Andrade Girón, D.; Marin Rodriguez, W.; Zúñiga Rojas, M. Inteligencia artificial generativa, paradigmas en crisis y el futuro epistémico de la investigación. In *Inteligencia artificial: usos y aplicaciones desde el contexto universitario*; Universidad Andina del Cusco/High Rate Consulting, 2025; pp. 45–53. https://doi.org/10.36881/ia2025.3
2. Mollema, W. J. T. A taxonomy of epistemic injustice in the context of AI and the case for generative hermeneutical erasure. *AI and Ethics* **2025**, *5*(5), 5535–5555. https://doi.org/10.1007/s43681-025-00801-w
3. Al-Zuhairi, T. Genognosis: AI's impact on knowledge production. *Cybrarians Journal* **2025**, *(75)*, 34–47. https://doi.org/10.70000/cj.2025.75.634
4. Yu, R., & Liu, J. (2025). Chat as Learning in Interactive Information Retrieval and Generation. *Annals of the Entomological Society of America*, *59*(1). https://doi.org/10.1145/3769733.3769735
5. Keshav Agrawal. The Future of AI in Digital Search: Toward a Fully Conversational Experience. *Journal of Computer Science and Technology Studies* **2025**, *7*(4), 298–306. https://doi.org/10.32996/jcsts.2025.7.4.34
6. Yan, W.; Nakajima, T.; Sawada, R. Beyond tool use: Tracking the evolution of generative AI literacy among university students through a process-oriented investigation. *Computers and Education: Artificial Intelligence* **2025**, *9*. https://doi.org/10.1016/j.caeai.2025.100465
7. Benjamin, J. J.; Lindley, J.; Edwards, E.; Rubegni, E.; Korjakow, T.; Grist, D.; Sharkey, R. Responding to Generative AI Technologies with Research-through-Design: The Ryelands AI Lab as an Exploratory Study. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference, DIS 2024*; Association for Computing Machinery, Inc., 2024; pp. 1823–1841. https://doi.org/10.1145/3643834.3660677
8. Pinski, M.; Benlian, A. AI literacy for users — A comprehensive review and future research directions of learning methods, components, and effects. *Computers in Human Behavior: Artificial Humans* **2024**, *2*(1), 100062. https://doi.org/10.1016/j.chbah.2024.100062
9. UNESCO. *Guidance for Generative AI in Education and Research*; UNESCO, 2023. https://doi.org/10.54675/ewzm9535
10. Sperber, D.; Clément, F.; Heintz, C.; Mascaro, O.; Mercier, H.; Origgi, G.; Wilson, D. Epistemic vigilance. *Mind and Language* **2010**, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x
11. Farquhar, S.; Kossen, J.; Kuhn, L.; et al. Detecting hallucinations in large language models using semantic entropy. *Nature* **2024**, *630*, 625–630. https://doi.org/10.1038/s41586-024-07421-0
12. Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Fung, P. Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys* **2023**. https://doi.org/10.1145/3571730
13. Sui, P.; Duede, E.; Wu, S.; Evans, J. A. Confabulation: The Surprising Value of Large Language Model Hallucinations. *Proceedings of the Annual Meeting of the Association for Computational Linguistics* **2024**, *1*, 14274–14284. https://doi.org/10.18653/v1/2024.acl-long.770
14. Salma Fatih Tofiq; Darya Abdulwahab Latif. Educating for the Algorithmic Age: Mapping AI Literacy. *World Journal of Advanced Research and Reviews* **2025**, *27*(2), 2170–2180. https://doi.org/10.30574/wjarr.2025.27.2.3045
15. Haugsbaken, H.; Hagelia, M. A New AI Literacy for The Algorithmic Age: Prompt Engineering Or Eductional Promptization? In *2024 4th International Conference on Applied Artificial Intelligence, ICAPAI 2024*; Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/ICAPAI61893.2024.10541229
16. Chee, H.; Ahn, S.; Lee, J. A Competency Framework for AI Literacy: Variations by Different Learner Groups and an Implied Learning Pathway. *British Journal of Educational Technology* **2025**. https://doi.org/10.1111/bjet.13556
17. Rapanta, C.; Bhatt, I.; Bozkurt, A.; Chubb, L. A.; Erb, C.; Forsler, I.; Jandrić, P.; et al. Critical GenAI Literacy: Postdigital Configurations. *Postdigital Science and Education* **2025**, *7*(4), 1296–1333. https://doi.org/10.1007/s42438-025-00573-w
18. Walters, S. E. Transmitting Understanding: Imitation, Transformation and Delayed Epistemic Vigilance. *Erkenntnis* **2025**. https://doi.org/10.1007/s10670-025-00990-0
19. Reisinger, D.; Kogler, M. L.; Jäger, G. On the interplay of gullibility, plausibility, and criticism: A computational model of epistemic vigilance. *JASSS* **2023**, *26*(3). https://doi.org/10.18564/jasss.5136

20. Austin, J. L. How to do things with words: Lecture II. In *The Performance Studies Reader: Fourth Edition*; Taylor and Francis, 2025; pp. 177–182. https://doi.org/10.4324/9781003282969-27
21. Searle, J. R. A classification of illocutionary acts. *Language in Society* 1976, *5*(1), 1–23. https://doi.org/10.1017/S0047404500006837
22. Goldberg, S. C. *Relying on Others: An Essay in Epistemology*; Oxford University Press, 2010. https://doi.org/10.1093/acprof:oso/9780199593248.001.0001
23. Lackey, J. *Learning from Words: Testimony as a Source of Knowledge*; Oxford University Press, 2008. https://doi.org/10.1093/acprof:oso/9780199219162.001.0001
24. Lavrakas, P. *Encyclopedia of Survey Research Methods*; Sage Publications, Inc., 2012. https://doi.org/10.4135/9781412963947
25. Parasuraman, R.; Manzey, D. H. Complacency and bias in human use of automation: An attentional integration. *Human Factors* 2010, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055
26. Goddard, K.; Roudsari, A.; Wyatt, J. C. Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association* 2012. https://doi.org/10.1136/amiajnl-2011-000089
27. Wickens, C. D.; Clegg, B. A.; Vieane, A. Z.; Sebok, A. L. Complacency and automation bias in the use of imperfect automation. *Human Factors* 2015, *57*(5), 728–739. https://doi.org/10.1177/0018720815581940
28. Liu, P.; Yuan, W.; Fu, J.; Jiang, Z.; Hayashi, H.; Neubig, G. Pre-train, Promotion, and Prediction: A Systematic Survey of Promotion Methods in Natural Language Processing. *ACM Computing Surveys* 2023, *55*(9). https://doi.org/10.1145/3560815
29. Shumailov, I.; Shumaylov, Z.; Zhao, Y.; Papernot, N.; Anderson, R.; Gal, Y. AI models collapse when trained on recursively generated data. *Nature* 2024, *631*(8022), 755–759. https://doi.org/10.1038/s41586-024-07566-y
30. Garry, M.; Chan, W. M.; Foster, J.; Henkel, L. A. Large language models (LLMs) and the institutionalization of misinformation. *Trends in Cognitive Sciences* 2024. https://doi.org/10.1016/j.tics.2024.08.007
31. Knar, E. Integrating Scientific Readiness and Added Scientific Value: A Multiplicative Conceptual Model for Research Assessment. *Publications* **2025**, *13*(4), 57. https://doi.org/10.3390/publications13040057
32. Zhumashov, Y.; Azambayev, S.; Bisenbaev, A.; Yessenzharov, A.; Bulatbayeva, K. Economic efficiency in grant funding evaluations: streamlining knowledge-intensive applications in Kazakhstan. *Economic Annals-XXI* **2023**, *(201)*, 1–2.