242. THE HUMAN FACTOR IN HUMAN – AI INTERACTION: HYPER – RELIANCE AS A SYSTEMIC COGNITIVE AND ORGANIZATIONAL RISK

Senkevich P.V., Ermakova A.D.
Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus

Shevaldysheva E.Z. – Associate Professor

The paper analyzes hyper-reliance on AI as a cognitive and organizational phenomenon in which users consistently treat algorithmic outputs as more trustworthy than their own expert judgments, effectively lowering the level of critical self-assessment. Drawing on international studies, statistical data, and documented incidents, the paper shows that this tendency is becoming systemic and generating new professional and economic risks. Unlike works focused on the technical improvement of AI, the paper highlights the human factor as the most vulnerable element of the system.

The rapid expansion of artificial intelligence in professional environments does not simply accelerate information processing; it gradually reshapes the way individuals approach analytical tasks and make decisions. A noticeable shift has taken place: tools that were once perceived as supportive instruments are increasingly treated as autonomous and more competent decision-makers. This shift leads many users to manifest automation bias, lowering their own level of critical evaluation when interacting with algorithmic outputs.

The relevance of this issue is supported by empirical evidence. A survey conducted by MTS Link and Jinn indicates that 97 % of HR specialists already use or intend to adopt AI tools in the near future, while only 38 % express concern about potential errors in personnel decisions caused by algorithmic inaccuracies [1]. At the same time, only about one third of organizations introduce AI in a structured, centralized manner. Most employees begin using such tools independently, without formal training or supervision, which further increases the risks associated with uncritical over-reliance.

The purpose of this study is to examine the psychological mechanisms that contribute to the formation of AI over-reliance and to outline practical approaches that help preserve critical thinking when interacting with algorithmic systems. To achieve this aim, the study identifies the factors that reduce critical evaluation of AI-generated outputs, presents statistical indicators illustrating the extent to which AI competence is overestimated, analyzes documented cases of financial losses caused by overtrust, and proposes recommendations aimed at reducing cognitive dependence on AI.

A study published in October 2025 by researchers from Aalto University revealed a paradoxical pattern: although AI improves task performance, it simultaneously reduces the accuracy of users' self-assessment [8]. Participants who relied on ChatGPT demonstrated higher performance metrics, yet consistently overestimated their personal contribution to the results. This effect resembles an "inversion of the Dunning–Kruger phenomenon": the more confident individuals felt about their AI literacy, the greater the gap between their perceived and actual performance. These findings suggest that understanding how technologies operate does not automatically guarantee an adequate calibration of trust in their outputs.

Additional insight is offered by the concept of the "competence paradox", discussed in AI & Society [3]. According to this research, specialists often confuse the ease of interacting with an interface with a genuine understanding of the underlying technology. This illusion of competence emerges from cognitive biases that transfer a sense of control from the familiar interface to the system itself, as well as from the social pressure to appear knowledgeable and technologically proficient. The issue of AI overtrust extends beyond individual cognitive tendencies and is deeply embedded in contemporary corporate culture. In many organizations, AI is implicitly presented as an infallible solution to human error. When management frames algorithmic tools as authoritative, employees begin to perceive them as digital experts whose conclusions should not be questioned [4]. In such an environment, expressing doubt may be subconsciously interpreted as resistance to innovation or insufficient digital adaptability. As a result, specialists may hesitate to challenge AI-generated outputs, even when they notice inconsistencies. A significant barrier to thorough verification is the corporate "economy of time". Under constant pressure from deadlines and performance indicators, employees often view the verification of each AI-generated response as an unnecessary step for which they lack resources. When organizational priorities emphasize speed, individuals are encouraged – directly or indirectly – to delegate complex cognitive tasks to machines, which leads to an uncritical acceptance of algorithmic decisions.

A deeper analysis shows that over-reliance in corporate settings is not merely a consequence of cognitive inertia. In many cases, it becomes a forced adaptation strategy. Employees consciously reduce their own analytical involvement to meet efficiency requirements and avoid appearing slower than algorithmic tools. This tendency is reinforced by the diffusion of responsibility: individuals begin to assume that any potential error will be attributed to the algorithm rather than to them. As a result, caution is unintentionally discouraged, and AI gradually shifts from a supportive instrument to a source of systemic vulnerability.

Empirical data confirms that over-reliance has already become widespread. Studies show that 71 % of users consider algorithmic decisions objective, while 20 % of managers allow AI to make final decisions without

human involvement [7]. At the same time, only 32 % of users have received formal training in AI tools, which contributes to an illusion of understanding. Interaction with AI is often limited to a single query, without follow-up verification, reinforcing the "quick answer" effect. These tendencies become even more concerning in light of expert forecasts: by 2027, more than 90 % of online content may be generated or modified by AI. In such an environment, the refusal to verify information becomes a systemic risk that affects the quality of managerial decisions and the overall resilience of organizations.

At the same time, it is important to recognize that the growing integration of AI into everyday professional routines creates a background environment in which automated recommendations gradually become perceived as a standard element of decision-making. As employees interact with such systems on a regular basis, the boundary between auxiliary support and authoritative guidance may become increasingly blurred. This gradual normalization of algorithmic involvement does not necessarily manifest through explicit organizational directives; rather, it develops through repeated exposure, convenience, and the implicit expectation that digital tools streamline routine tasks. Over time, these subtle dynamics can reinforce a general tendency to accept AI-generated outputs with reduced scrutiny, even in situations that would traditionally require careful human judgment.

Several real-world cases illustrate the consequences of uncritical over-reliance. One of the most widely discussed examples is the Air Canada incident [2]. A corporate chatbot incorrectly informed a passenger that they could apply for a bereavement fare retroactively. Trusting the official service, the client incurred financial losses, and the court rejected the company's attempt to treat the chatbot as an independent legal entity. The resulting compensation exceeded $50,000, demonstrating that failure to verify AI-generated information can lead to legal and reputational damage.

Another example involves an investment fund where an analyst relied entirely on a credit-risk forecasting model [6]. The algorithm failed to account for geopolitical factors, resulting in losses of approximately $2 million. Researchers describe this as another manifestation of the competence paradox: the specialist assumed the model was more competent than themselves and refrained from verifying its conclusions. A similar pattern is observed in personnel management. According to Resume Builder, 66 % of managers use algorithms to make layoff decisions [7]. The issue is that models often reinforce existing biases. The financial consequences of such errors can be substantial. In all these cases, the damage resulted not from technical failures but from the absence of human verification.

If overtrust is both a cognitive and institutional phenomenon, it requires a comprehensive response. Several approaches have already demonstrated effectiveness and may serve as the foundation for a new culture of AI interaction. One promising approach is the "forced reflection method", proposed by researchers at Aalto University [8]. The system encourages users to articulate their reasoning, consider alternative explanations, and justify agreement or disagreement with AI conclusions. This method disrupts the illusion of automatic correctness and restores the user's active analytical role. Another approach involves "technological time-outs", implemented at Rowan University [5]. Students temporarily set aside digital tools and solve problems independently. Comparing their own conclusions with AI recommendations helps them understand the limits of algorithmic reasoning and strengthens autonomous analytical skills. In corporate settings, mandatory reflection periods before strategic decisions could serve as an effective analogue.

References:

1. MTS Link: [website] / MTS. – Moscow, 2025. – URL: https://mts-link.ru/blog/38-rossiiskih-hr-professionalov-opasayutsya-nevernyh-kadrovyh-reshenii-iz-za-oshibok-ii/ (date of access: 25.03.2026).

2. BC Civil Resolution Tribunal: [website] / Civil Resolution Tribunal. – Victoria, 2018. – URL: https://civilresolutionbc.ca (date of access: 25.03.2026).

3. Springer Nature Link: [website] / Springer Nature. – Cham, 2026. – URL: https://link.springer.com (date of access: 25.03.2026).

4. ACM Digital Library: [bibliographic database] / Association for Computing Machinery. – New York, 2026. – URL: https://dl.acm.org (date of access: 25.03.2026).

5. Rowan Digital Works: [institutional repository] / Rowan University Libraries. – Glassboro, 2026. – URL: https://rdw.rowan.edu (date of access: 25.03.2026).

6. Financial Stability Board: [website] / FSB. – Basel, 2026. – URL: https://www.fsb.org (date of access: 25.03.2026).

7. ResumeBuilder.com: [website] / Bold LLC. – Salt Lake City, 2026. – URL: https://www.resumebuilder.com (date of access: 25.03.2026).

8. *Aaltodoc: [institutional repository] / Aalto University. – Espoo, 2026. – URL: https://aaltodoc.aalto.fi (date of access: 25.03.2026).*