CSCanada
Higher Education of Social Science
Volume 29, Number 1, 2025, pp. 1-5
DOI:10.3968/13880

ISSN 1927-0232 [Print]
ISSN 1927-0240 [Online]
www.cscanada.net
www.cscanada.org

The Predicament and Path Forward for University Teachers' Digital Literacy from the Perspective of Large Model Security

CHEN Jinying[a],*

[a] Guangdong Polytechnic of Industry and Commerce, Guangzhou, China.
*Corresponding author.

Received 2 August 2025; accepted 15 September 2025
Published online 26 September 2025

Abstract
Artificial intelligence technologies, represented by Large Language Models (LLMs), are profoundly reshaping the educational ecosystem, presenting new era-specific demands for teacher digital literacy. However, LLMs are not purely empowering technological tools; their inherent security risks, such as "hallucinations," privacy leakage, and prompt injection, pose systemic challenges to the authenticity, security, and ethical boundaries of educational activities. This paper, from the perspective of large model security, analyzes the dual effects of LLMs in the educational field. On this basis, it delves into the four-fold predicament that currently confronts the development of teacher digital literacy: cognitive, competency-related, practical, and systemic. The study posits that addressing these challenges requires the construction of a four-dimensional development strategy centered on "Prudent Integration": at the cognitive level, fostering a critical view of technology; at the competency level, constructing an integrated skill chain; at the practical level, deepening contextualized pedagogical innovation; and at the ecological level, cultivating a multi-layered, collaborative support system. This paper aims to provide theoretical references and practical pathways for the professional development of teachers in the age of artificial intelligence, emphasizing that the core role of the teacher will shift from that of a knowledge transmitter to a prudent constructor of knowledge and a guide of wisdom within a human-AI collaborative environment.

Key words: Large language models; Model security; Teacher digital literacy; AI hallucination; TPACK

Chen, J. Y. (2025). The Predicament and Path Forward for University Teachers' Digital Literacy from the Perspective of Large Model Security. *Higher Education of Social Science*, *29*(1), 1-5. Available from: URL: http://www.cscanada.net/index.php/hess/article/view/13880
DOI: http://dx.doi.org/10.3968/13880

INTRODUCTION
Generative artificial intelligence, represented by Large Language Models (LLMs), is permeating the field of education with unprecedented depth and breadth. From planning personalized learning paths to the instantaneous generation of teaching content, it demonstrates immense transformative potential, yet "education also faces unprecedented challenges." Some scholars even argue that guiding teachers to "embrace the digital transformation of education is the next critical proposition for the cause of teacher education in our country." In this context, whether teachers can effectively command and utilize emerging technologies has become a key factor in determining the success of the digital transformation in education. It has been pointed out that university teachers must rapidly enhance their digital literacy. Teacher digital literacy has evolved from early-stage IT application skills to a comprehensive literacy encompassing the "consciousness, ability, and responsibility to use digital technology to optimize, innovate, and reform educational and teaching activities." It pertains not only to technical operation but also to instructional design, assessment innovation, professional development, and even ethical responsibilities in a digital context.

However, the integration of LLMs into educational practice is not a straightforward path. Unlike traditional educational technology tools, LLMs possess "black-box" characteristics and inherent uncertainty. A series of severe security issues exist within large language models, including but not limited to "hallucination" at the content

level, and privacy leakage and "Prompt Injection" attacks at the interaction level. Artificial intelligence also brings forth issues such as "digital discrimination, deepfakes, and privacy and data protection." These security concerns are not mere technical flaws but systemic risks that directly impact the core values of education, posing a potential threat to the reliability of knowledge and the security of the teaching process.

Therefore, this study re-examines the development of teacher digital literacy from the perspective of "large model security," aiming to answer the following core questions: Under the inherent security risks of LLMs, what unprecedented predicaments does the development of teacher digital literacy face? How should the education system construct a systematic response strategy to help teachers move beyond mere technological application to achieve a "Prudent Integration" of LLMs, thereby transforming risks into opportunities for development?

1. THE DUAL EFFECTS OF LARGE LANGUAGE MODELS: EDUCATIONAL EMPOWERMENT AND SECURITY RISKS

The impact of large language models on education presents a typical duality. They are powerful cognitive aids while also introducing unprecedented security challenges, profoundly affecting the professional practice of teachers.

1.1 Educational Empowerment: A Cognitive Partner for Teachers

LLMs can significantly empower teachers, with artificial intelligence "deeply intervening in the entire process of teachers' design, implementation, and evaluation of student learning." Firstly, as a **generator of teaching** resources, LLMs can quickly create lesson plans and exercises based on specific teaching objectives, alleviating the administrative burden on teachers. Secondly, as a **personalized tutoring assistant**, they can provide students with immediate answers and differentiated exercises, assisting teachers in achieving individualized instruction. Thirdly, as a **"sparring partner" for professional development**, LLMs can offer diverse perspectives on instructional design and classroom reflection, stimulating teachers' pedagogical innovation.

1.2 Security Risks: Systemic Threats to Educational Practice

Accompanying their empowering potential are security risks that cannot be ignored. These risks permeate educational practice at the content, interaction, and ethical levels.

Content Security Risk: The Erosion of Factual Truth by "AI Hallucination." Research has systematically expounded on the issue of "hallucination," where the model generates content that appears plausible but is factually incorrect. Its causes are complex, involving multiple layers such as training data, model architecture, and decoding strategies. In an educational context, this means that if teachers directly accept content generated by an LLM—be it a description of a historical event, an explanation of a scientific concept, or a news summary—they risk transmitting false information to students, directly shaking the knowledge foundation of education.

Interactional Security Risk: The Threats of Privacy Leakage and Prompt Injection. Studies emphasize that LLMs pose significant privacy security risks during interaction. In the course of their use, teachers might input sensitive information involving students' personal situations, class management, or teaching reflections. Once such data is improperly collected or leaked, the consequences could be severe. Furthermore, "prompt injection" attacks allow malicious users to construct specific instructions to bypass the model's safety guardrails, inducing it to generate harmful content or execute unauthorized operations, posing a direct threat to the security of the classroom environment.

Ethical Security Risk: The Challenges of Algorithmic Bias and Value Neutrality. The knowledge of LLMs is derived from their massive training data, which inevitably carries the existing biases and stereotypes of human society. It has been noted that "large models may confidently output erroneous or non-existent answers," and "when parts of the facts are replaced by incorrect but similar information, humans often find it difficult to identify." When the model unconsciously reproduces or even amplifies these biases while generating content, it can have a subtle negative impact on the formation of students' values. This also presents a serious test for teachers striving to uphold the principles of educational equity and value neutrality.

2. THE FOUR-FOLD PREDICAMENT OF TEACHER DIGITAL LITERACY FROM THE PERSPECTIVE OF LARGE MODEL SECURITY

The security risks of large models are not isolated technical problems; they directly translate into real-world predicaments for the development of teacher digital literacy, manifesting at the cognitive, competency, practical, and systemic levels.

2.1 Cognitive Predicament: Wavering Between "Techno-Myth" and "Techno-Panic"

Faced with powerful yet uncertain LLMs, the teaching

community is prone to falling into two cognitive extremes: one is blind worship under a "techno-myth," treating the LLM as an omniscient expert, which leads to "cognitive offloading" and the abdication of professional judgment. The other is wholesale avoidance driven by "techno-panic," where the fear of risks leads to a refusal to use the technology, thereby missing developmental opportunities. The root cause lies in a lack of profound understanding of the LLM's nature as a probabilistic tool.

2.2 Competency Predicament: A Structural Deficit in Discernment and Collaboration Skills

The core literacies in the era of LLMs have shifted towards higher-order skills of discernment and collaboration, whereas traditional training has often focused on operational skills. Teachers currently exhibit a common competency gap. First is a deficiency in **discerning evaluation ability**, lacking a systematic methodology to assess the validity, accuracy, and reliability of AI-generated content, such as through multi-source cross-verification and logical chain analysis. Second is an insufficiency in **precise guidance ability**, namely, the so-called "Prompt Engineering" skill, making it difficult to guide the model to produce high-quality, low-risk content through precise instructions. Third is a weakness in **human-AI collaborative ability**, failing to establish a mature "Human-in-the-loop" working model, which risks the teacher's agency being usurped by technological logic, a concern echoing the emphasis on teacher subjectivity.

2.3 Practical Predicament: The Persistent Challenges of Pedagogical Integration and Ethical Dilemmas

How to safely and effectively integrate a tool fraught with "hallucination" risk into the rigor of daily teaching is a significant practical challenge for teachers. Concurrently, ethical dilemmas are ever-present: How to define the reasonable boundaries for student use of AI to maintain academic integrity? How to protect student privacy and avoid algorithmic discrimination when using AI for assessment? The lack of clear guidelines for these issues often places teachers in a difficult position.

2.4 Systemic Predicament: A System-Wide Absence of Professional Development Support

The efforts of individual teachers require systemic support, yet the current ecosystem development lags significantly. Firstly, the **professional training system is slow to update**, with few specialized courses on the security and ethics of LLMs. Secondly, there is a **dearth of institutional policies at the school level**, with most schools yet to issue clear guidelines on AI use. Lastly, the **function of professional communities has not been fully leveraged**, lacking the "mutual learning mechanisms" needed to share "failure cases" and "success stories" in managing risks and forming "collective wisdom."

3. THE PATH FORWARD: A FOUR-DIMENSIONAL STRATEGY FOR DEVELOPING TEACHER DIGITAL LITERACY

In response to the aforementioned predicaments, the development of teacher digital literacy requires a profound paradigm shift. We propose a four-dimensional strategic framework with "Prudent Integration" as its core concept, aimed at helping teachers achieve professional growth while navigating risks.

3.1 Cognitive Reshaping: Cultivating a Critical Technological View of "Prudent Optimism"

The cornerstone of this strategy is the internal cognitive reshaping of teachers. It is essential to abandon binary thinking and establish a form of "Prudent Optimism." This means actively embracing the opportunities brought by LLMs on one hand, while maintaining high vigilance towards their limitations and risks on the other. The key is to redefine the human-AI relationship, viewing the LLM as a knowledgeable but occasionally fallible "cognitive partner" or "co-pilot," rather than an "absolute authority" that replaces the teacher's thinking. Under this cognition, the teacher's role evolves from a traditional "knowledge transmitter" to a "wisdom guide" in a new educational ecosystem—an architect of the human-AI collaborative environment, a discerner of complex information, and a facilitator of students' higher-order thinking.

3.2 Competency Construction: Forging an Integrated "Discern-Verify-Integrate" Skill Chain

Building upon cognitive reshaping, it is necessary to cultivate teachers' core skills in a targeted manner, forming a closed loop for managing AI content risks.

The first step is to **develop discernment skills**, which serve as the first line of defense against AI security risks. Teachers need to understand the technical origins of issues like "hallucination" to maintain a healthy "professional skepticism" towards all AI-generated content, internalizing a sense of critical awareness. The second is to **master verification skills**, translating discernment into concrete action. This includes: (1) learning high-quality **Prompt Engineering** to constrain the model's output through precise questioning; (2) mastering **cross-verification methods** and becoming proficient in checking against authoritative sources; and (3) cultivating **source-tracing and logical scrutiny abilities** to question information sources and examine argumentation processes. The third is to **enhance integration wisdom**, the highest level of competency, which is demonstrated by strategically integrating LLMs into teaching. For example, using AI for "brainstorming" to inspire instructional design, but the final plan must be based on one's own pedagogical knowledge (PK) and judgment of the student's situation. A highly innovative practice is to **transform the limitations**

of AI into a teaching resource, designing inquiry-based activities such as "AI fact-checking" or "challenging the AI" to guide students in developing critical thinking through interaction.

3.3 Deepening Practice: Promoting Contextualized Ethical Norms and Pedagogical Innovation

The development of teacher digital literacy must ultimately be grounded in authentic teaching practices. To this end, abstract ethical principles and security norms need to be made concrete and contextualized. Schools and educational research institutions should collaborate to develop case libraries for the safe application of LLMs in different subjects, providing operational suggestions and ethical considerations for specific scenarios. At the same time, teachers should be encouraged to act as "action researchers," exploring LLM integration models suitable for their own school and class contexts. Furthermore, AI ethics and security should be incorporated as core components of digital literacy education, integrated into relevant curricula to help students establish a responsible and critical view of AI use from an early age.

3.4 Ecological Support: Fostering a Multi-Layered, Collaborative Professional Development System

The growth of individual teachers depends on a strong external support ecosystem.

At the **macro level**, education authorities should research and issue guiding policies on the use of generative AI in basic education, providing an authoritative ethical framework, security standards, and codes of conduct for schools and teachers. At the **meso level (school/regional)**, high-quality, systematic special training programs should be organized, focusing on deeper issues such as AI risk identification and critical thinking cultivation. Concurrently, drawing on research about "mutual learning mechanisms," the establishment of teacher Professional Learning Communities (PLCs) should be vigorously promoted to provide teachers with a safe and open space for regularly sharing confusions, lessons from failures, and successful experiences in using AI. At the **micro level (individual teacher)**, within a supportive ecosystem, teachers should be encouraged to achieve dynamic, autonomous, and lifelong development of their digital literacy through continuous self-reflection, participation in professional communities, and reading cutting-edge research.

4. RESEARCH OUTLOOK AND PRACTICAL PATHWAYS

This study is primarily based on theoretical speculation and strategy construction derived from existing literature. Its main limitation is the lack of large-scale empirical data support, and the effectiveness of the proposed strategies awaits subsequent verification.

Looking ahead, as LLM technology iterates, its security features may change, and related countermeasures will need to be dynamically adjusted. A valuable research direction is to shift the perspective from mere risk prevention to creatively utilizing AI's "imperfections" (such as hallucinations) to design innovative teaching activities, transforming security issues from "challenges" into "opportunities." To implement the above strategies, the following project-based paths can be explored:

School-Based Research Projects: Develop teaching modules for the prudent integration of LLMs around specific subjects (e.g., history, science), organizing teachers for collective lesson preparation, classroom practice, and reflective seminars.

"AI+" Mentorship Programs: Pair experienced veteran teachers with young teachers to provide personalized guidance, focusing on areas such as AI tool selection, instructional design integration, and ethical risk avoidance.

Regional Inter-School Alliances: Establish cross-school teacher professional learning communities to regularly hold online and offline salons for sharing high-quality cases, jointly researching practical problems, and co-building application resource libraries.

CONCLUSION

The security risks of large language models are a structural reality that education in the digital age cannot evade. They have profoundly reshaped the core meaning of teacher digital literacy—evolving from proficiently "using" a tool to prudently "navigating" a human-AI collaborative environment fraught with uncertainty. From a security perspective, the connotation of teacher digital literacy has been endowed with unprecedented critical, ethical, and strategic dimensions.

In the face of this challenge, the core of the "Prudent Integration" strategic framework constructed in this paper lies in the adherence to and reshaping of teacher subjectivity. In the age of artificial intelligence, the power of technology has not diminished the value of the teacher; on the contrary, it increasingly highlights the irreplaceability of the professional judgment, educational wisdom, ethical responsibility, and humanistic care that are unique to human teachers. To accurately assess and guide the enhancement of teacher digital literacy, the following quantitative assessment dimensions can be constructed based on the TPACK framework:

Table 1
TPACK-Based Assessment Dimensions for Teacher Digital Literacy in the Context of Large Model Security

TPACK Dimension	Core Literacy Indicators from the Perspective of Large Model Security	Assessment Suggestions
Technological Knowledge (TK)	Understanding the probabilistic nature and security risks of LLMs; mastering high-quality prompt engineering and multi-source cross-verification skills.	Scenario simulation tests, prompt design tasks.
Content Knowledge (CK)	Ability to use LLMs to expand subject knowledge boundaries and to prudently evaluate, verify, and correct the accuracy of AI-generated content.	Observation of lesson preparation process, subject knowledge verification.
Pedagogical Knowledge (PK)	Ability to design human-AI collaborative teaching activities; ability to transform LLM limitations into teaching opportunities for cultivating students' critical thinking.	Review of instructional design plans, analysis of classroom recordings.
Technological Content Knowledge (TCK) & Technological Pedagogical Knowledge (TPK)	Ability to select appropriate LLM application models based on subject and teaching objectives, and to foresee and avoid AI-related risks in teaching.	Cross-disciplinary case analysis, micro-teaching sessions.
Technological Pedagogical Content Knowledge (TPACK)	Comprehensively applying the above knowledge and skills to design, implement, and reflect on a complete teaching plan that can "prudently integrate" LLMs.	Portfolio assessment (including design, reflection, etc.).

The process of successfully navigating the security risks of large models is also a process for the teaching community to complete the reshaping of its professional identity and achieve a spiral ascent of digital literacy within the TPACK framework. The future educational landscape will not be one of simple competition or replacement between humans and machines, but one where wise teachers, who know how to collaborate prudently with machines, will lead students to navigate broader oceans of knowledge and richer spiritual worlds in this new era of coexisting opportunities and challenges.

REFERENCES

Cao, T. J., et al. (2024). Teaching practice of cybersecurity based on generative AI technology. *Journal of Science and Education*, (15).

Duan, D. P., et al. (2025). Implementation of faculty professional development programs in age of artificial intelligence: Characteristics, challenges and inspiration. *Heilongjiang Researches on Higher Education*, (6).

Liu, L. (2020). Dilemma and breakthrough in the transformation of teachers' role in the era of artificial intelligence. *Open Education Research*, *26*(3). https://doi.org/10.13966/j.cnki.kfjyyj.2020.03.009

Liu, Z. Y., et al. (2025). Survey on hallucinations in large language models. *Journal of Software*.

Lu, H., et al. (2025). Research on the components and mutual learning mechanisms of teachers' digital evaluation literacy from a limited evidence-based perspective. *e-Education Research*, (5), 15.

Mishra, P., & Koehler, M. J. (2008, March 24-28). *Introducing technological pedagogical content knowledge*[Paper presentation]. Annual Meeting of the American Educational Research Association, New York, NY, United States.

Tian, H. J. (2021). College and university teachers' development at the age of AI: Ideological breakthrough and roadmap unfolding. *Journal of South China Normal University (Social Science Edition)*, (4).

Wu, D., et al. (2023). Teacher digital literacy: Connotation, standards, and evaluation. *e-Education Research*, *44*(8), 15–23. https://doi.org/10.13811/j.cnki.eer.2023.08.015

Zhao, Y., et al. (2023). Security of large language models: Current status and challenges. *Computer Science*, *50*(8).

Zhu, X. D., et al. (2025). Holistic professional development of teachers in the artificial intelligence era. *Journal of Distance Education*, (4).