International Journal of Asian Social Science
ISSN(e): 2224-4441
ISSN(p): 2226-5139
DOI: 10.55493/5007.v15i10.5638
Vol. 15, No. 10, 362-375

URL: www.aessweb.com

Navigating ai in China: Access, censorship, and trust among preservice teachers

Ziyue Zhang[1]
Sharon Friesen[2]
Jian Liu[3+]

[1]*Beijing Language and Culture University, China.*
Email: zhangziyue.blcu@gmail.com
[2]*University of Calgary, Canada.*
Email: sfriesen@ucalgary.ca
[3]*Hebei Normal University, China.*
Email: liujian@hebtu.edu.cn

(+ Corresponding author)

Article History
Received: 4 July 2025
Revised: 8 September 2025
Accepted: 29 September 2025
Published: 6 October 2025

Keywords
AI literacy
Artificial intelligence
Censorship
Chinese higher education
Preservice teachers
Trust in AI.

ABSTRACT

This study aims to explore how preservice teachers at a Chinese Normal University perceive and engage with artificial intelligence (AI) tools, with particular attention to issues of access, censorship, and trust. Drawing on questionnaire responses from 50 sophomores and follow-up unstructured interviews with five participants, the research investigates students' awareness and usage of both international platforms (e.g., ChatGPT) and Chinese-developed systems (e.g., Wenxin, Deepseek). The study further examines students' perceptions of content differences across platforms, especially in politically sensitive contexts, and considers how these perceptions influence levels of trust in AI-generated information. Findings reveal that participants demonstrate a nuanced awareness of censorship and its implications, noting divergences in information quality and availability depending on platform origin. While many students acknowledge the educational potential of AI tools, they also express skepticism toward politically restricted outputs, underscoring how sociopolitical conditions shape digital trust. The results highlight that AI literacy plays a critical role not only in shaping preservice teachers' trust in emerging technologies but also in guiding their future pedagogical choices. This research contributes to broader discussions of AI ethics and cross-cultural digital engagement, offering insights into how higher education in China can better prepare future educators to critically evaluate and responsibly integrate AI into their teaching practices.

Contribution/ Originality: This study is one of the few that have investigated Chinese preservice teachers' perceptions of both domestic and international AI tools, highlighting how censorship, language ability, and sociopolitical context shape AI trust. It contributes novel insights into developing critical AI literacy in cross-cultural higher education settings.

1. INTRODUCTION

The pervasive integration of artificial intelligence (AI) has profoundly reshaped the global educational landscape, fundamentally altering how knowledge is accessed, processed, and disseminated (Holmes, Bialik, & Fadel, 2022; Luckin, Holmes, Griffiths, & Forcier, 2016). From sophisticated personalized tutoring systems that adapt to individual learning paces to advanced automated writing assistants that refine prose, AI-powered tools are increasingly becoming indispensable components of both formal pedagogical frameworks and informal learning environments. Within this burgeoning technological ecosystem, large language models (LLMs) like OpenAI's ChatGPT have emerged as particularly transformative, garnering immense attention for their remarkable capacity

to generate human-like text, facilitate content creation, and provide comprehensive support across diverse academic disciplines (Kasneci et al., 2023; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019). However, the effective utilization, nuanced perception, and fundamental trust in these powerful AI tools are not uniform; they are deeply contingent upon a complex interplay of local technological infrastructures, prevailing regulatory environments, and unique sociopolitical contexts.

In parallel with this global AI ascendancy, China has witnessed a rapid and robust development of its own domestic AI tools. Platforms such as Baidu's Wenxin Yiyan (ERNIE Bot) and Alibaba's Tongyi Qianwen have emerged as significant counterparts to international offerings like ChatGPT. These Chinese-developed platforms are meticulously tailored to local linguistic nuances, cultural sensibilities, and stringent regulatory frameworks, frequently operating under more pronounced government oversight and rigorous content filtering mechanisms (Chen, Zhang, & Ma, 2021; Liu & Graham, 2023). Consequently, Chinese users frequently encounter palpable differences between domestic and international AI tools, extending beyond mere variations in language and performance to encompass critical distinctions in how sensitive information, particularly politically sensitive content, is handled. These divergences raise profound questions regarding user trust, equitable access, and the overarching role of state-imposed censorship in shaping AI-mediated knowledge acquisition and understanding within the Chinese digital sphere.

The concept of trust in AI-generated information holds paramount importance within educational settings, where both students and educators are increasingly reliant on these tools for study support, writing assistance, and comprehensive information retrieval. Yet, the formation of trust is a multifaceted construct, not solely determined by the accuracy, fluency, or technical proficiency of AI responses. Instead, it is profoundly influenced by perceived transparency in algorithms, consistency in output, and, critically, ideological neutrality (Floridi & Cowls, 2022; Shin, 2021). In politically sensitive environments, such as the People's Republic of China, users often exhibit a heightened awareness of potential limitations or deliberate omissions in AI-generated content, particularly concerning topics deemed politically sensitive. This awareness can inevitably lead to discerning skepticism, prompting users to critically question the reliability, completeness, and ultimate credibility of the information provided by AI systems (Tang & Wu, 2023). The implications for critical thinking and information literacy are significant, as users must navigate an information landscape potentially shaped by regulatory directives.

This study specifically focuses on preservice teachers, a demographic of exceptional importance within this unfolding discourse. As the architects of future educational paradigms, their current experiences, evolving perceptions, and developing attitudes toward AI tools will undeniably dictate the pedagogical approaches they adopt and the extent to which they integrate such technologies into their future classrooms. Their foundational ability to critically evaluate and contextualize AI-generated information is not merely central to their own digital literacy development but is also absolutely vital for cultivating sophisticated critical thinking skills and fostering responsible AI use among the generations of students they will eventually teach (Zhang & Luo, 2024; Zhao, 2023). Understanding their perspectives offers invaluable insights into the future trajectory of AI adoption in Chinese education.

2. LITERATURE REVIEW

The advent of artificial intelligence (AI) has ushered in a new era for education globally, fundamentally transforming pedagogical approaches, learning processes, and administrative efficiencies (Holmes et al., 2022; Zawacki-Richter et al., 2019). The landscape of AI in education (AIED) is vast, encompassing a spectrum of applications from sophisticated adaptive learning systems that personalize content delivery to intelligent tutoring systems providing real-time feedback, and from AI-powered writing assistance tools to automated assessment platforms (Crompton & Burke, 2022; Miao, Holmes, & Lee, 2021). A particularly disruptive development has been the emergence of Large Language Models (LLMs), exemplified by tools such as OpenAI's ChatGPT. These models,

trained on colossal datasets, demonstrate an unprecedented capacity to generate coherent, contextually relevant, and even creative text, revolutionizing how students and educators engage with information. Their integration into educational contexts for writing support, translation, ideation, and even complex problem-solving is rapidly expanding (Kasneci et al., 2023; Trust, Whalen, & Mouza, 2023). The rapid adoption of these technologies necessitates a thorough understanding of their multifaceted impacts on teaching and learning worldwide.

2.1. AI in Education: Global Trends and Local Developments

Globally, the proliferation of AI in education is characterized by innovation and diverse applications. AI is being leveraged to offer personalized learning pathways, identify individual learning gaps, and provide targeted interventions, thereby enhancing learning effectiveness and efficiency (Demartini, Mariani, & Benussi, 2024; Ng et al., 2025). Beyond direct student interaction, AI also streamlines administrative tasks for educators, facilitates data-driven decision-making, and supports the development of more objective evaluation systems (Crompton & Burke, 2022; Kannan & Zapata-Rivera, 2022). However, the implementation and development of AIED are not uniform across national borders; they are significantly shaped by distinct geopolitical, economic, and socio-cultural factors.

In China, the trajectory of AI development in education mirrors, yet diverges from, global trends. While international platforms like ChatGPT have captivated worldwide attention, China has concurrently invested heavily in developing its own robust domestic AI ecosystem. This has led to the release of prominent Chinese AI tools such as Baidu's Wenxin Yiyan (ERNIE Bot), Alibaba's Tongyi Qianwen, and iFLYTEK's Xinghuo. These platforms are designed to offer comparable functionalities to their international counterparts but are meticulously tailored to align with local linguistic nuances, cultural values, and, crucially, stringent regulatory frameworks (Chen, Suen, & Wang, 2021; Liu & Graham, 2023). While these domestic tools offer obvious advantages in native language support and greater cultural alignment, their operation is notably characterized by tighter state control and oversight. This inherent difference raises critical questions regarding their transparency, the scope of information they provide, and their adherence to national ideologies (Rowsell, 2025; The Diplomat, 2025). The Chinese government's "Made in China 2025" initiative and subsequent AI development plans explicitly recognize the strategic importance of AI, not only for economic growth but also for societal governance, including content regulation.

2.2. Access to AI Tools and the Emergence of a "Digital Divide"

Access to AI technologies is far from equitable and is profoundly influenced by a complex interplay of national infrastructure, governmental regulatory policies, and prevailing geopolitical boundaries. In the context of China, direct and unfettered access to international platforms like ChatGPT is typically restricted without the use of virtual private networks (VPNs) or other technical workarounds. Conversely, domestic AI tools are more widely and readily available, but their functionality for certain politically sensitive topics is demonstrably constrained by content filtering and censorship mechanisms (Tang & Wu, 2023).

This creates a nuanced "AI digital divide" in China, which extends beyond the mere availability of technology. It describes a disparity where not only the presence of AI tools but also the ideological scope and breadth of information accessible through these tools differ significantly depending on the platform used (Lau & Wu, 2023). This bifurcation can have profound implications for educational equity. Research on AI access in educational settings consistently highlights that disparities in tool availability and functionality can directly impact student engagement, hinder the development of critical thinking skills, and impede overall digital literacy (Holmes et al., 2022; Luckin et al., 2016). Students primarily exposed to censored or constrained AI systems may consequently develop a limited or skewed understanding of AI capabilities, its inherent biases, and its real-world boundaries. This can, in turn, affect their expectations of the technology, their ability to critically evaluate its outputs, and ultimately, their capacity to become digitally literate citizens capable of discerning reliable information in a complex media environment (Eaton

et al., 2025; Sand Technologies, 2025). The implications of such a divide extend to the depth and breadth of knowledge students can acquire, particularly on topics that might be sensitive within the national context.

2.3. Trust in AI-Generated Information: The Role of Competence, Transparency, and Integrity

Trust stands as an indispensable factor in the successful adoption and effective integration of AI, particularly within educational settings where learners and educators fundamentally depend on the accuracy, impartiality, and overall reliability of information (KPMG, 2025; Shin, 2021). The concept of trust in AI tools is multifaceted, encompassing several critical dimensions:

1. Perceived Competence: This refers to the user's belief in the AI tool's accuracy, efficiency, and ability to perform its designated tasks effectively (Floridi & Cowls, 2022). In an educational context, this means an AI tutor providing correct answers or an AI writing assistant offering grammatically sound and relevant suggestions.
2. Transparency: This dimension relates to how understandable the AI tool's reasoning, processes, and potential limitations are to the user (Floridi & Cowls, 2022). A transparent AI system would ideally make its sources clear, acknowledge potential biases, or explain how it arrived at a particular answer.
3. Integrity: This crucial dimension assesses how free the AI tool is from systemic biases, manipulation, or alignment with specific agendas (Floridi & Cowls, 2022). Users often evaluate AI based on whether it appears to be objective and serves their best interests.

Empirical studies unequivocally demonstrate that when users perceive or suspect that an AI tool is deliberately censoring information, manipulating responses, or overtly aligning its outputs with political or commercial agendas, their trust in the tool significantly erodes (Tang & Wu, 2023; Zhou & Lee, 2022). This erosion of trust occurs even if the tool maintains a high level of technical accuracy in other, non-sensitive areas. In the specific context of China, the pervasive role of the state in regulating and overseeing AI-generated content can lead users to inherently perceive domestic tools as less objective or neutral. This suspicion is often amplified when users encounter inconsistencies, evasive responses, or complete omissions on politically sensitive topics, such as historical events or contemporary political debates (The Diplomat, 2025). Such experiences can foster a critical, albeit potentially cynical, approach to AI-generated information, compelling users to seek alternative sources or to apply heightened scrutiny to AI outputs. These dynamics underscore the complex relationship between technological utility and perceived ideological control.

2.4. Preservice Teachers and AI Literacy: Preparing Future Educators

Preservice teachers constitute an exceptionally vital population in the ongoing discourse surrounding AI in education. As the future custodians of learning environments, their foundational experiences, evolving attitudes, and proficiency with AI tools will directly shape how they integrate these technologies into their pedagogical practices and, critically, how they prepare the next generation of learners for an AI-augmented world. Consequently, AI literacy defined as the comprehensive ability to understand the principles of AI, effectively use AI tools, and critically evaluate their outputs and societal implications is increasingly recognized as an indispensable component of contemporary teacher preparation programs (Chiu, Wang, & Lo, 2024; Zhang & Luo, 2024; Zhao, 2023).

However, extant literature reveals a concerning gap: many preservice teachers currently lack a sufficient and nuanced understanding of how AI functions, its inherent limitations, the potential for algorithmic biases, and how political, commercial, or even ethical interests can profoundly shape its outputs (Popenici & Kerr, 2017; Zhang, He, & Zhao, 2024). This deficit in AI literacy can compromise their ability to effectively leverage AI in their teaching and to guide students responsibly.

In the Chinese higher education context, research specifically addressing AI literacy among preservice teachers is a nascent but rapidly expanding field. Recent studies indicate that while Chinese education majors often express considerable enthusiasm for AI's transformative potential in education, their levels of critical awareness, particularly

concerning issues of bias, censorship, and ethical use, vary significantly (Zhao, 2023). Some studies point to a correlation between AI literacy and general attitudes towards AI, though this relationship may not always be strong (Özden, Örge Yaşar, & Meydan, 2025). Furthermore, research highlights the necessity of tailored support and hands-on training to develop teachers' AI competencies (Kohnke, Zou, & Zhang, 2023). This underscores a critical need to better understand precisely how preservice teachers in China perceive trustworthiness in AI-generated content, especially given their unique exposure to both domestically controlled and internationally restricted AI platforms. Such an understanding is pivotal for developing targeted AI literacy curricula that prepare future educators to navigate complex digital environments, foster critical thinking in their students, and responsibly harness the power of AI in their classrooms.

3. METHODOLOGY

3.1. Research Design

This study employed a mixed-methods design to explore Chinese preservice teachers' access to and trust in AI-generated information. The combination of quantitative survey data and qualitative insights from unstructured interviews enabled a broader and deeper understanding of the participants' experiences and perceptions (Creswell & Plano Clark, 2018). Special attention was given to how cultural exposure, language proficiency, and access to both domestic and international AI tools influenced students' views, particularly in the context of politically sensitive topics.

3.2. Participants

Fifty sophomore students from a Normal University in China participated in the study. All participating students were Chinese nationals and preservice teachers expected to work in the K–12 education system after graduation. While the survey included students from multiple education-related majors, five follow-up interview participants were all English majors. This distinction is important, as English majors in China are typically more exposed to Western culture through coursework, media consumption, and engagement with English-language resources.

Their higher English proficiency not only enabled them to interact effectively with international AI tools like ChatGPT but also gave them the ability to recognize nuanced differences in language, tone, and content between responses generated in English and those in Chinese. Moreover, due to their cultural immersion and access to foreign media sometimes facilitated through tools like VPNs these students were better positioned to ask politically sensitive questions in both languages and evaluate the degree to which AI responses differed by platform and language.

3.3. Data Collection

A structured questionnaire was administered to all 50 participants. It consisted of five closed and open-ended questions aimed at understanding students' awareness, usage, and perceptions of both domestic and international AI tools. The questions were:

1. Are you aware of international AI tools such as ChatGPT?
2. Have you used Chinese-developed AI tools (e.g., Baidu Wenxin, Alibaba Tongyi)?
3. Do you think there are noticeable differences in responses between Chinese and international AI models?
4. When dealing with politically sensitive topics, do you believe AI-generated responses in China are influenced by censorship?
5. How do these variations in AI responses impact your trust in AI-generated information?

The responses were collected anonymously, and the results provided a general overview of how Chinese preservice teachers perceive the trustworthiness and accessibility of AI-generated content.

In the second phase, unstructured face-to-face interviews were conducted with five students, all of whom were English majors. The unstructured format was deliberately chosen to avoid restricting participants' thoughts or leading them toward predefined answers. This approach allowed participants to guide the conversation and surface issues that might not have been anticipated by the researcher (Minichiello, Aroni, Timewell, & Alexander, 2008).

During the interviews, the participants reflected on their experiences using AI tools in both English and Chinese, comparing responses and noting perceived censorship or ideological filtering. Their proficiency in English, familiarity with Western discourse, and, in some cases, access to foreign news and social media via VPNs, allowed them to explore sensitive topics more freely and critically. These capabilities influenced the questions they chose to ask AI tools and informed their assessments of the differences in tone, depth, and transparency between domestic and international platforms. The interviews were conducted in Chinese, audio-recorded with participants' consent, and transcribed for analysis.

3.4. Data Analysis

The quantitative questionnaire data were analyzed using descriptive statistics to identify common patterns in AI tool awareness, usage, and perceived reliability. These results were used to frame the broader landscape of AI engagement among preservice teachers.

For the qualitative interview data, a thematic analysis was conducted following (Braun & Clarke, 2006) six-step framework. Given the unstructured format of the interviews, coding was done inductively, allowing key themes to emerge organically from the participants' narratives. Particular attention was paid to themes related to critical awareness of censorship, linguistic comparison of AI responses, and trust in AI tools based on perceived openness and cultural bias.

3.5. Confidentiality

Participation in both the questionnaire and interviews was voluntary, and informed consent was obtained from all participants. Interviewees were reminded that they could withdraw at any time. To protect participants' identities, pseudonyms were used, and all data were stored securely. The study followed institutional ethical guidelines for research involving human subjects.

4. FINDINGS

This study thoroughly examined the landscape of AI tool awareness, usage, and perceptions among 50 Chinese preservice teachers, providing robust insights into their interactions with both domestic and international AI models. The findings reveal a nuanced and often critical engagement with AI, shaped by factors ranging from accessibility to ideological content, ultimately impacting trust in AI-generated information.

The data clearly illustrate a significant bifurcation in the awareness and practical application of AI tools among the participants. As shown in Table 1, a majority of preservice teachers (58.49%) were aware of international AI tools like ChatGPT but had not yet used them, indicating widespread recognition without commensurate direct engagement. In contrast, 41.51% had actively used these international platforms, suggesting a segment of the preservice teacher population is already navigating global AI technologies. This widespread recognition, even in the absence of direct interaction for a large proportion, underscores the pervasive global influence and reputation of these AI models.

Table 1. Question 1: Are you aware of international AI tools like ChatGPT? (Percentage).

Yes, I have used them.	41.51%
Yes, but I have not used them.	58.49%
No, I have never heard of them.	0%

A distinct and compelling pattern emerged with Chinese-developed AI tools (Table 2), which demonstrated significantly higher rates of active use. Over half of the participants (50.94%) reported frequent use, with an additional 41.51% using them occasionally. This stark disparity strongly suggests that while international AI tools are recognized, their practical utility among Chinese preservice teachers is likely constrained by various factors. These constraints could include access limitations, prevalent VPN restrictions that hinder seamless connectivity, or a general unfamiliarity with English-language interfaces common among international platforms. Consequently, domestically developed tools appear to be the more accessible, convenient, and culturally aligned option for daily use, leading to their predominant adoption.

Table 2. Question 2: Have you used Chinese-developed AI tools (e.g., Baidu Wenxin, Alibaba Tongyi)? (Percentage).

Yes, frequently	50.94%
Occasionally	41.51%
Rarely	5.66%
Never	1.89%

Further qualitative insights from five unstructured interviews with English-major students provided crucial context. These students, due to their advanced English proficiency, reported stronger access to and more frequent usage of international AI tools. Their linguistic capabilities enabled them to navigate platforms like ChatGPT with greater ease, and several openly admitted to employing VPNs to bypass geographical restrictions. This proactive approach highlights their heightened digital literacy and determination to access a broader range of AI resources. Their exposure to Western media and academic content often acted as a powerful catalyst, fueling their curiosity and prompting them to actively compare how international and domestic AI systems respond to sensitive topics, thereby showcasing a proactive and critical engagement with AI technology across different cultural and political contexts.

A striking and widely held belief among preservice teachers, as reflected in the survey, was the presence of noticeable differences between domestic and international AI models. As detailed in Table 3, a clear majority (67.92%) perceived "some differences," with a substantial 18.87% believing these differences to be "significant." This widespread conviction was consistently echoed by interviewees, who attributed the variations not merely to superficial linguistic quality but to deeper, more fundamental aspects such as ideological content and the depth and scope of explanations.

Table 3. Question 3: Do you think there are noticeable differences in responses between Chinese and international AI models? (Percentage).

Yes, significant differences	18.87%
Some differences	67.92%
No, responses are similar	3.77%
Not sure	9.43%

Several interviewees consistently noted that responses from international AI models, particularly those in English, often provided more critical, nuanced, or multi-faceted perspectives on complex social, political, or ethical issues. This suggests a greater willingness or programming within international models to explore controversial topics from various angles. Conversely, domestic tools tended to offer more conservative, generalized, or vague replies, often avoiding direct engagement with sensitive subjects. This observation was vividly illustrated by one English major: “I asked the same question about democracy to both ChatGPT and a Chinese AI. The Chinese one gave me a general answer, but ChatGPT discussed the pros and cons with examples from different countries. It felt more informative and balanced.” This comparative analysis underscores that English majors, by virtue of their linguistic capabilities and critical engagement, possess a keen awareness of underlying content filtering mechanisms. They are capable of assessing the scope, neutrality, and reliability of AI responses across different platforms,

suggesting a discerning approach to AI-generated information that recognizes the potential for bias or limited perspectives depending on the model's origin and inherent programming. The survey findings revealed a profound and pervasive belief in the censorship of AI-generated responses within China, particularly concerning politically sensitive topics. As shown in Table 4, a significant 41.51% of participants believed AI content is "definitely" influenced by censorship, with an additional 47.17% indicating "somewhat." This widespread sentiment, totaling nearly 90%, points to a deep-seated skepticism about the ideological neutrality of AI content within the Chinese context, suggesting users are keenly aware of potential state-imposed controls on information dissemination, even through AI systems.

Table 4. Question 4: When dealing with politically sensitive topics, do you believe AI-generated responses in China are influenced by censorship? (Percentage).

Yes, definitely	41.51%
Somewhat	47.17%
No, AI responses are objective	3.77%
Not sure	7.55%

Interview participants further elaborated on these perceptions, describing firsthand experiences where Chinese AI tools either explicitly avoided sensitive questions entirely or offered conspicuously limited and generic responses. These observations were not anecdotal; they were part of a consistent pattern that led students to actively engage in experimental queries. Some students proactively experimented with phrasing questions in both Chinese and English to directly compare outcomes, a testament to their critical literacy and active efforts to test the boundaries of AI responses. One student shared a particularly compelling observation: “If I ask about historical events in English, sometimes I get more detailed responses. But in Chinese, it’s like the AI avoids the question.” This dual-language approach employed by the students reflects not only their critical literacy but also their acute awareness of language as a mediating tool for accessing different types of knowledge through AI. Their deliberate manipulation of language to test AI boundaries highlights a sophisticated understanding of how information can be controlled or restricted through technological means, and their active strategies to circumvent such limitations.

The responses to the fifth survey question, while diverse, illuminated critical patterns regarding the preservice teachers' trust in AI-generated information. A notable number of students expressed "skepticism or conditional trust" in AI, particularly when encountering inconsistent, overly vague, or ideologically framed responses. For instance, one participant succinctly stated: “If AI responses are inconsistent, it can lead to doubt about the reliability of the information provided.” This highlights that consistency and clarity are crucial factors in building user trust. Many participants also emphasized the importance of "cross-checking," indicating a proactive and responsible approach to verifying AI output by comparing it with other AI tools, traditional search engines, or authoritative human-generated sources. Conversely, a smaller but significant group viewed AI positively, citing its ability to offer "diverse perspectives" and generate multiple interpretations as a valuable strength, suggesting that for some, the very variability can be seen as an intellectual asset. Interview data strongly reinforced this complex interplay of trust and skepticism, with English majors frequently emphasizing the "need for critical thinking" and thorough verification. One interviewee noted, “I don’t fully trust AI, especially in topics like politics or history. I always double-check with books or foreign media.” This statement underscores a clear understanding of AI's limitations and a commitment to independent verification. Another expressed palpable concern over "unexplained variations" in AI output: “Sometimes the same question gives different answers. That’s confusing, and it makes me question whether AI can really be objective.” This frustration with inconsistency points to a fundamental user expectation of reliability and predictability from AI systems, especially when they are intended to provide factual information.

Despite these significant concerns and observed inconsistencies, some interviewees also viewed response variation as "intellectually stimulating." They felt it encouraged them to explore different viewpoints, critically

evaluate information, and engage in deeper inquiry, effectively transforming a potential drawback into an opportunity for enhanced learning and intellectual growth. This fascinating tension between skepticism and intellectual curiosity defines a "conditional trust framework." Within this framework, AI is neither fully relied upon as an infallible source nor entirely dismissed as unreliable. Instead, it is approached with a discerning and critical eye, its outputs subject to human evaluation, verification, and interpretation. This nuanced perspective underscores that trust in AI is not a binary concept but rather a dynamic and evolving relationship influenced by direct experience, perceived reliability, and sophisticated critical engagement with the technology itself. It suggests that future educators will need to cultivate this conditional trust in their students, fostering critical thinking skills to navigate the complexities of AI-generated information.

5. DISCUSSION

This study offers a comprehensive exploration into Chinese preservice teachers' evolving engagement with AI, revealing a dynamic interplay of awareness, accessibility, and trust, particularly when comparing domestic and international AI tools. The findings underscore the critical role of digital literacies among Chinese university students, especially English majors, who consistently bridge the divide between domestic and Western knowledge systems. These insights hold significant implications for future educational practices and the cultivation of an informed digital citizenry. The observed discrepancy between awareness (over 58%) and actual usage (41.72%) of international AI tools like ChatGPT among participants highlights a significant access barrier for many Chinese users. These obstacles are multifaceted, extending beyond mere technological restrictions such as internet firewalls and VPN limitations. As Feng and Wang (2023) and Qiang (2023) suggest, linguistic and sociopolitical factors also play a crucial role. The predominantly English interface of many international AI platforms presents a significant hurdle for non-English proficient users, while the broader sociopolitical climate can influence the perceived safety or utility of accessing foreign digital resources. Crucially, the study reveals that English majors are more adept at circumventing these restrictions, frequently employing VPNs to gain access. This proficiency underscores the notion that language functions as a form of digital capital (Van Dijk, 2020). Their command of English not only facilitates navigation of these platforms but also provides a deeper engagement with the content, often rooted in Western epistemological frameworks. This finding reinforces the idea that English education in China extends beyond mere linguistic acquisition, offering a gateway to broader ideological and epistemological perspectives that might otherwise be inaccessible. This access to a wider range of information and perspectives through international AI tools can significantly shape their understanding of global discourse and critical thinking.

The strong consensus among participants (nearly 87%) regarding perceived differences between domestic and international AI models is a critical finding, echoing global discussions on AI alignment, localization, and ideological bias (Hao, 2021; Zeng, Lu, & Huangfu, 2022). The consistent observation that domestic tools provide vague or limited responses, particularly on politically sensitive topics, while international models like ChatGPT offer more detailed, nuanced, and pluralistic answers, points to the profound impact of algorithmic design and content moderation on user experience and trust (Zhao & Lin, 2023). This is not merely a technical difference but a reflection of the embedded values and regulatory environments in which these AI systems are developed. Domestic AI, often operating within a more controlled information environment, is likely programmed to align with specific national narratives and censorship policies, leading to self-censorship or avoidance of certain topics. In contrast, international AI, developed in different regulatory landscapes, may prioritize a broader range of perspectives and comprehensive information, even on contentious issues. This divergence highlights a global tension in AI development: whether AI should be a neutral information provider or a tool for upholding specific societal or political values.

The proactive approach adopted by some students, posing the same questions to both domestic and international AI tools often in both Chinese and English demonstrates a sophisticated form of critical digital inquiry. This practice

transcends simple information retrieval; it reflects a conscious effort to test the consistency, depth, and ideological orientation of AI outputs across linguistic and cultural contexts. Such behavior resonates with Jiang and Whittaker (2022)'s conclusion that Chinese youth are increasingly perceiving AI not merely as a technological tool but as a discursive actor embedded within wider sociopolitical narratives. This active comparison suggests a burgeoning awareness among users that AI responses are not monolithic or inherently objective, but rather products of their training data, programming, and the socio-political environments in which they are developed. This critical engagement empowers users to identify potential biases and limitations, fostering a more discerning approach to AI-generated content. It also aligns with the concept of "algorithmic literacy," where users understand the underlying mechanisms and potential influences shaping algorithmic outputs (Seaver, 2017).

The study highlights that trust in AI is a complex and context-dependent phenomenon. While some students expressed skepticism due to inconsistent responses across platforms, others viewed such variability as a sign of adaptability and the evolving nature of AI systems. This ambivalence mirrors broader public attitudes toward AI, where trust is dynamically shaped by transparency, reliability, and user experience (Shin, 2021). The findings strongly reinforce the imperative of explainability and consistency in fostering user trust in AI systems (Eiband, Schneider, Bilandzic, Fazekas-Con, & Butz, 2021; Raji, Binns, Veale, Van Kleek, & Shadbolt, 2020). Users are more likely to trust AI when they understand how it arrives at its conclusions and when its outputs remain consistent for similar queries. Unexplained variations erode confidence and raise questions about the AI's underlying logic or potential biases. Conversely, when AI systems demonstrate a clear and consistent logic, even if their responses are limited due to external constraints, users are more likely to maintain a degree of conditional trust. This suggests that future AI development must prioritize not just accuracy, but also the clarity and predictability of its output to build stronger user relationships. The interviews further underscored that students are actively cultivating critical AI literacy. Rather than passively accepting AI-generated content, they engage in deliberate comparison, verification against alternative sources, and critical reflection on the content's implications. Such behaviors are essential for developing a resilient, informed digital citizenry capable of navigating increasingly complex information environments (Kimmons, Rosenberg, & Allman, 2023; Livingstone & Helsper, 2007). This also suggests that trust in AI is not a fixed binary (either fully trusted or fully distrusted), but rather a spectrum mediated by users' capacity to interpret and evaluate algorithmic content. This active, evaluative stance represents a crucial step towards fostering responsible AI use.

The participants' disciplinary background as English majors significantly influenced their responses and perceptions. Their sustained exposure to Western discourses encompassing concepts like critical thinking, human rights, and democratic values through both their curriculum and extensive language engagement likely heightened their sensitivity to ideological bias and content censorship within AI outputs. This aligns with previous research demonstrating that academic discipline significantly affects how students interact with and evaluate AI-generated information (Tang, Wang, & Luo, 2023). English majors, by virtue of their academic training, are arguably better equipped to discern subtle ideological leanings and critically analyze information presented through different cultural lenses. Moreover, the study reveals a compelling intersection of language and ideology. Many participants perceived English-language AI content as more objective or balanced, while viewing Chinese-language responses as more guarded or state-aligned. These perceptions reflect the broader sociopolitical context of digital information in China, where state narratives coexist with and sometimes conflict with transnational knowledge flows (Liu, 2022). This highlights how language itself can be perceived as a vector for different ideological frameworks. The choice of language for AI interaction, therefore, becomes not just a matter of convenience but a conscious decision that influences the type and perceived neutrality of the information received. This perception can lead to a fundamental difference in how users approach and trust information from different linguistic and cultural AI sources.

Taken together, these findings highlight an urgent and critical need to integrate AI literacy into teacher education programs, particularly in contexts where access to and trust in information are mediated by complex

sociotechnical constraints. As future educators, preservice teachers will play a pivotal role in shaping how students understand technology, knowledge, and truth in an increasingly AI-driven world. Fostering critical thinking, ethical awareness, and robust source verification in relation to AI tools must be a core component of their professional preparation (Blikstein, 2018; Zawacki-Richter, Marín, Bond, & Gouverneur, 2022). This includes equipping them with the skills to identify algorithmic biases, understand data privacy implications, and guide their future students in navigating diverse AI-generated content responsibly.

Future research could further explore the long-term impact of perceived censorship on trust in AI among Chinese users and investigate how educational interventions can effectively cultivate critical AI literacy across different academic disciplines. Additionally, comparative studies involving preservice teachers from other cultural contexts could offer richer insights into the global dynamics of AI awareness and adoption in education.

6. SUGGESTIONS AND RECOMMENDATIONS

Based on our findings, we recommend a series of actionable steps to strengthen critical AI literacy among Chinese preservice teachers, with a particular focus on bridging access gaps, supporting ethical awareness, and enhancing curriculum development. One clear divide that emerged in our research is that students with stronger English proficiency are significantly more likely to access and critically engage with international AI tools, while others are limited to domestic platforms. To address this imbalance, institutions should offer targeted language support programs and provide vetted, institutionally sanctioned access to global AI tools like ChatGPT or Claude preferably through localized or translated interfaces. This not only democratizes access to diverse knowledge sources but also reduces reliance on VPNs and informal workarounds. Teacher education programs should also scaffold students' growing abilities to compare and interpret AI-generated content by incorporating comparative inquiry tasks, reflective writing assignments, and classroom discussions that explore the ideological and political underpinnings of different AI systems.

At the same time, fostering ethical awareness and transparency in AI use must become a foundational aspect of preservice teacher training. Many students in our study expressed skepticism toward ideologically constrained or inconsistent AI responses especially when engaging with sensitive topics. To address this, coursework should introduce key concepts such as AI transparency, accountability, and explainability, while also offering practical tools for identifying ethical dilemmas through case studies or simulations. Educators should also be encouraged to model open dialogue and critical engagement in the classroom, creating space for multiple perspectives. On a broader level, policymakers and AI developers should work collaboratively with educational institutions to design AI platforms and national standards that reflect the lived experiences of users and support both cultural relevance and academic freedom. Finally, interdisciplinary coursework and international exchange programs can expose preservice teachers to diverse epistemologies, helping them become globally literate educators who are equipped to guide students in critically navigating AI in an increasingly complex digital world.

7. CONCLUSION

This study thoroughly investigated Chinese preservice teachers' interactions with AI-generated information, focusing on their awareness, access, and trust in these tools, especially when comparing domestic and international AI models. Our findings, drawn from both quantitative questionnaire data and qualitative insights from unstructured interviews, reveal a nuanced picture of how these students navigate a complex digital environment profoundly shaped by technological access, linguistic ability, and sociopolitical context.

While a significant majority of participants demonstrated awareness of international AI platforms like ChatGPT, their actual usage was notably lower. This disparity often stemmed from practical limitations such as internet censorship and restricted access. However, a distinct trend emerged among English majors, who showed considerably

greater engagement with these international tools. This highlights how factors like language proficiency and exposure to Western discourse can significantly expand digital agency, enabling these students to bypass typical barriers. Their unique ability to directly compare responses from both domestic and international AI tools frequently by formulating queries in both Chinese and English demonstrates a developing and sophisticated form of critical AI literacy. This comparative approach allows them to discern underlying biases and limitations, fostering a more informed interaction with AI.

Crucially, trust in AI-generated content was not absolute but profoundly shaped by the perceived consistency, transparency, and ideological positioning of the information provided. Many students expressed healthy skepticism, particularly when encountering overt censorship or contradictory outputs from AI tools. This highlights a critical awareness of potential biases and manipulations inherent in AI systems. Conversely, a segment of participants appreciated the capacity of AI to offer a variety of perspectives, viewing this diversity as a valuable asset for learning and exploration. These varied reactions underscore that trust in AI is far from a binary concept; instead, it operates within a conditional framework, heavily influenced by both the specific content delivered and the broader sociopolitical context in which that information is accessed and interpreted.

In an era where AI is rapidly transforming educational and informational landscapes, this study underscores the urgent need to cultivate critical engagement with AI tools among future educators. Understanding how users perceive, evaluate, and ultimately trust AI-generated knowledge will be paramount. Equipping preservice teachers with these critical AI literacies is essential for fostering a generation of informed, reflective, and responsible digital citizens who can effectively navigate the opportunities and challenges presented by artificial intelligence. Moving forward, continued research into these dynamic interactions will be vital to ensure that educational practices evolve to meet the demands of an AI-driven world.

7.1. Use of Generative AI

This manuscript made use of ChatGPT to assist in paraphrasing and improving the clarity of the text. No AI tools were used for the generation of research hypotheses or the writing of conclusions. The final manuscript was authored and reviewed by the listed authors without additional AI input.

Funding: This study received no specific financial support.
Institutional Review Board Statement: The study involved minimal risk and followed ethical guidelines for social science fieldwork. Formal approval from an Institutional Review Board was not required under the policies of Hebei Normal University, China. Informed verbal consent was obtained from all participants, and all data were anonymized to protect participant confidentiality.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Blikstein, P. (2018). *Pre-college computer science education: A survey of the field.* Mountain View, CA: Google LLC. https://services.google.com/fh/files/misc/pre-college-computer-science-education-report.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Chen, L., Zhang, Y., & Ma, H. (2021). Development of artificial intelligence in China: A policy perspective. *International Journal of Emerging Technologies in Learning, 16*(14), 21–34.

Chen, Y., Suen, H. K., & Wang, D. (2021). Artificial intelligence in education: Applications and ethical concerns. *Educational Technology Research and Development, 69*(2), 897–900.

Chiu, T. K. F., Wang, X., & Lo, Y. (2024). AI literacy for preservice teachers: A framework and curriculum design principles. *Educational Technology & Society, 27*(1), 88–102.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). United States: SAGE Publications.

Crompton, H., & Burke, D. (2022). *Artificial intelligence in education: Emerging research and opportunities*. United States: IGI Global.

Demartini, C. G., Mariani, L., & Benussi, L. (2024). AI and personalized learning: The promise and limits of algorithmic education. *Journal of Educational Technology Systems, 52*(2), 123–145.

Eaton, S. E., Pethrick, H., Hayden, K. A., Brennan, R., Wiens, J., & McDermott, B. (2025). *Academic integrity and artificial intelligence in higher education contexts: A rapid scoping review*. Canada: Canadian Perspectives on Academic Integrity.

Eiband, M., Schneider, H., Bilandzic, M., Fazekas-Con, C., & Butz, A. (2021). *Bringing transparency design into practice*. Paper presented at the Proceedings of the 2021 ACM CHI Conference on Human Factors in Computing Systems.

Feng, X., & Wang, Y. (2023). Digital inequalities in China: AI access, education, and the great firewall. *Journal of Asian Internet Studies, 7*(1), 34–49.

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Machine Learning and the city: Applications in Architecture and Urban Design*, 535-545.

Hao, K. (2021). *How China is policing AI*. United States: MIT Technology Review.

Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. United States: Center for Curriculum Redesign.

Jiang, J., & Whittaker, S. (2022). Chinese youth and digital media: Between national narratives and global flows. *Global Media and China, 7*(1), 3-19.

Kannan, A., & Zapata-Rivera, J. D. (2022). AI in education: Enhancing administrative efficiency and decision-making. *MDPI Sustainability, 14*(1), 1101.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kimmons, R., Rosenberg, J. M., & Allman, B. (2023). What should AI literacy look like for teachers? *Computers & Education: Artificial Intelligence, 4*, 100116.

Kohnke, L., Zou, D., & Zhang, R. (2023). Preparing teachers for AI-enhanced classrooms: A scoping review. *Education and Information Technologies, 28*, 4153–4175.

KPMG. (2025). *Trust in AI: Global education insights report*. Netherlands: KPMG International.

Lau, J., & Wu, Y. (2023). Understanding the AI digital divide in China: Technology access and ideological implications. *Journal of Technology and Society, 15*(2), 45–60.

Liu, Y. (2022). The digital transformation of Chinese media: State control, market forces, and global influences. *Journalism Studies, 23*(12), 1653–1670.

Liu, Y., & Graham, S. (2023). *Censorship practices of the People's Republic of China*. China: U.S.-China Economic and Security Review Commission.

Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New Media & Society, 9*(4), 671-696. https://doi.org/10.1177/1461444807080335

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. United Kingdom: Pearson Education.

Miao, F., Holmes, W., & Lee, S. (2021). *AI and education: Guidance for policymakers*. United Kingdom: UCL Institute of Education.

Minichiello, V., Aroni, R., Timewell, E., & Alexander, L. (2008). *In-depth interviewing: Principles, techniques, analysis* (3rd ed.). Australia: Pearson Education Australia.

Ng, A., Aboraya, W., Abdulghafor, R., Subramanian, K. P., Al Husaini, Y., & Al Hussaini, M. (2025). AI-powered learning pathways: Personalized learning and dynamic assessments. *International Journal of Advanced Computer Science and Applications, 16*(1), 454–461.

Özden, M., Örge Yaşar, F., & Meydan, E. (2025). The relationship between pre-service teachers' attitude towards artificial intelligence (AI) and their AI literacy. *Pegem Journal of Education and Instruction*, *15*(3), 121–134. https://doi.org/10.47750/pegegog.15.03.13

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

Qiang, X. (2023). The battle for the future of the internet: China's digital authoritarianism. *Journal of Democracy*, *34*(1), 52–66.

Raji, I. D., Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2020). *The fallacy of AI functionality*. Paper presented at the Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society.

Rowsell, J. (2025). *The dangers of political ideology in AI*. United States: The Diplomat.

Sand Technologies. (2025). *AI and the future of work: Insights from the World Economic Forum's Future of Jobs Report 2025*. United States: Sand Technologies.

Seaver, N. (2017). Algorithms as culture: Computation in the cultural sector. *Big Data & Society*, *4*(2), 1–12.

Shin, D. (2021). The impact of explainable artificial intelligence (XAI) on user trust: An empirical investigation. *Information Processing & Management*, *58*(1), 102717.

Tang, Y., Wang, L., & Luo, X. (2023). *Disciplinary influence on Chinese university students' attitudes toward AI-generated content*. United Kingdom: Springer.

Tang, Y., & Wu, Y. (2023). Political sensitivities and AI-generated content in China. *AI & Society*, *38*(1), 111–123.

The Diplomat. (2025). *Is China building the future of AI governance through open-source modeling?* United States: The Diplomat.

Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education*, *23*(1), 1–23.

Van Dijk, J. (2020). *The digital divide*. Cambridge, UK: Polity Press.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2022). A systematic review of artificial intelligence in higher education: Insights from 10,000 publications (1990-2019). *Journal of Educational Technology & Society*, *25*(1), 1-33.

Zeng, Y., Lu, E., & Huangfu, C. (2022). Ethical considerations of AI in China: Regulatory frameworks and future directions. *AI and Ethics*, *2*, 233–244.

Zhang, X., & Luo, L. (2024). AI literacy for preservice teachers: Preparing educators for intelligent digital futures. *Computers & Education*, *210*, 104716.

Zhang, Y., He, Y., & Zhao, J. (2024). Chinese education majors' perceptions of AI literacy: Challenges and opportunities. *Asia-Pacific Education Researcher*, *33*(1), 95–108.

Zhao, R. (2023). Teaching with AI: Chinese preservice teachers' digital literacy and ethical awareness. *Asia-Pacific Journal of Teacher Education*, *51*(4), 460–475.

Zhao, X., & Lin, Y. (2023). Algorithmic opacity and state narratives: Trust in AI in the Chinese context. *Chinese Journal of Communication*, *16*(1), 22–39.

Zhou, M., & Lee, F. L. (2022). Online trust and political information in authoritarian contexts: The case of China. *Information, Communication & Society*, *25*(6), 803–818.