Asia Pacific Science Press

Journal of Educational Theory and Practice
ISSN(O):3005-9313
ISSN(P):3005-9305
Vol. 2 No.3 (2025)

From Clarity to Conviction: Instrumental Limits and Integration Pathways for Generative Artificial Intelligence in University Ideological and Political Education

Zhihao Wei[1]*, Zhen Liu[1], Tao Wang[2]

1.Chongqing College of Humanities, Science & Technology, Chongqing, 401524, China

2. Guangzhou Xinhua University, Guangzhou, 523133, China

Corresponding author: Zhihao Wei, 20191150019@mail.gdufs.edu.cn

Abstract: This qualitative study examines how generative artificial intelligence is being integrated into university ideological and political education (IPE) in China and delineates the conditions under which its instrumental rationality reaches its practical limits. We conducted semi-structured interviews with 17 instructors from five universities in Chongqing (45–120 minutes, in Chinese), audio-recorded, transcribed verbatim, and analyzed using reflexive thematic analysis (RTA). Sampling and stopping were guided by information power; we judged data adequacy when the developing patterns were sufficiently rich and useful for the research questions. NVivo 12 supported data management. We identified three themes: attenuation of affective and faith dimensions; content complexity and the limits of AI understanding; and insufficiency of high-quality, compliant training data. Building on these findings, we propose an integration framework that aligns classroom practice with platform support and institutional governance, and we formulate actionable recommendations for policymakers, universities, and instructors.

Keywords: Generative Artificial Intelligence; Ideological and Political Education (IPE); Chinese Higher Education

Published: Sept 9, 2025

DOI:https://doi.org/10.62177/jetp.v2i3.552

Introduction

The deep integration of artificial intelligence (AI) into university-level ideological and political education (IPE) is both an imperative of the era and an intrinsic requirement of educational governance. At the policy level, the 2019 Opinions on Deepening the Reform and Innovation of Ideological and Political Theory Courses in the New Era explicitly call for applying AI and other modern information technologies in IPE, thereby providing direction and impetus for an updated instructional paradigm. In line with this mandate, a transition from traditional models to an AI-driven paradigm has been framed as both necessary and consequential, with the stated goal of enhancing the ideological and theoretical depth, affinity, and relevance of courses so that reasoning becomes more compelling, confusion is dispelled, and learning is both engaging and convincing to students (General Office of the CPC Central Committee & General Office of the State Council, 2019). In practice, the integration of AI into IPE has emerged as a salient direction for teaching reform, underpinned by institutional efforts to build smart platforms, pilot innovative teaching, and redesign assessment, trends that are broadly consistent with international syntheses on AI's role in higher education (Selwyn, 2019; Zawacki-Richter et al., 2019). At a more macro level, national strategies such as the New Generation Artificial Intelligence Development Plan and China's Education Modernization 2035 signal a combined era-driven and policy-driven momentum, emphasizing the use of intelligent systems to reshape educational provision and governance. Collectively, these initiatives offer dual institutional and technological safeguards for improving the quality and efficiency of IPE in the context of digital transformation (Central Committee of the Communist Party of China & State Council, 2019; State Council of the People's Republic of China, 2017).

From the standpoint of feasibility and implementation, artificial intelligence (AI) can provide systemic support for IPE through three complementary mechanisms: embedding across instructional processes, scenario-based applications, and robust technical infrastructure. Drawing on large scale educational data and algorithmic modeling, institutions can develop digital learner profiles that link tasks with content, offer personalized learning pathways for students, and generate data-informed recommendations for instructors (Ferguson, 2012; Luckin, Holmes, Griffiths, & Forcier, 2016; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019). These capabilities can facilitate whole process and whole cycle enhancement of teaching and assessment, while strengthening teachers' instructional agency and broadening students' autonomy with respect to time, place, and access to resources (Selwyn, 2019; UNESCO, 2021). At the operational level, principles for intelligent pedagogy, learning resources, and evaluation are increasingly well specified in the research literature and in policy guidance (UNESCO, 2021; Zawacki-Richter et al., 2019). Learning analytics enable continuous diagnostic tracking and personalization (Ferguson, 2012). Curriculum aligned digital repositories and analytics platforms support evidence based design and iterative improvement (Zawacki-Richter et al., 2019). Large scale models can be deployed for formative and process-oriented evaluation with near real time feedback, provided their use remains aligned with curricular standards and pedagogical goals (Selwyn, 2019; UNESCO, 2021).

However, embedded applications also entail structural risks and context specific challenges. Algorithmic distribution in online environments may reduce the reach of mainstream discourse, intensify risks associated with ideological identification, and weaken value guidance, as research on platform curation and political exposure has shown (Bakshy, Messing, & Adamic, 2015; Bail et al., 2018). Related problems include the dilution of teachers' authority and discursive leadership, the erosion of learner agency, and the emergence of data privacy and algorithmic bias concerns in educational settings (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016; Selwyn, 2019; Slade & Prinsloo, 2013). At the classroom level, clogged channels of information sharing, the influence of online subcultures that attenuate the authority of mainstream ideology, and the fragmentation between online and offline learning contexts further amplify mismatches between instructional supply and learner demand, leading to imbalances in interaction and participation (Garrison & Kanuka, 2004; Garrison, Anderson, & Archer, 2000; Selwyn, 2019). In addition, current initiatives reveal application gaps, including smart systems that do not translate into genuinely intelligent pedagogy, interactive features that do not yield authentic dialogue, and incomplete data coverage. These gaps coincide with compounded ethical and governance challenges related to privacy protection, algorithmic bias, and technological dependence, indicating that the integration of AI remains in a critical transition from pilot to scale and requires stronger guidance and safeguards (Mittelstadt et al., 2016; OECD, 2021; UNESCO, 2021).

Grounded in this context, the present study examines the challenges and responses associated with integrating artificial intelligence into IPE and seeks to address a critical gap through an empirical lens. On the one hand, bibliometric and review evidence indicates that research on AI in education is expanding rapidly yet remains fragmented, with limited consolidation of core author groups and institutional collaborations and with a persistent imbalance between technical development and pedagogical inquiry (Zawacki-Richter et al., 2019). On the other hand, existing scholarship frequently privileges descriptive or conceptual discussion while providing insufficient classroom-based evidence and process data, which underscores the need for mixed method and cross disciplinary approaches that can connect design, implementation, and evaluation in authentic settings (Roll & Wylie, 2016; UNESCO, 2021). The literature has also called for rigorous examination of the actual effects of AI in teaching and learning in order to assess benefits and risks in a transparent and comparable manner and to narrow the gap between theory and practice (UNESCO, 2021; Zawacki-Richter et al., 2019). Responding to these needs, this article investigates the experiences of 17 university IPE instructors and employs RTA to identify salient problems, elucidate underlying mechanisms, and propose actionable strategies, thereby generating evidence that is both verifiable and transferable for the prudent integration of AI and for the high quality development of IPE (Braun & Clarke, 2006; Braun & Clarke, 2019).

Methods

This qualitative study was conducted in Chongqing, China. It examined university instructors of ideological and political theory courses, focusing on their lived experiences of integrating generative artificial intelligence into teaching and assessment, and on their views of AI's roles and limits in value guidance, instructional organization and assessment redesign. Prior to data collection, ethical approval was obtained from the authors' institution, and all procedures adhered to recognized research-ethics and data-protection requirements.

Recruitment information was disseminated via teaching and research offices responsible for ideological and political theory courses and through departmental academic-affairs channels at five universities. Instructors who expressed interest were referred by office contacts to the research team, which then scheduled interviews at mutually convenient times and locations. Written informed consent was obtained before each interview; participants were informed about audio recording, de-identification and intended uses of the data. A combination of purposive and snowball sampling was employed, yielding a final sample of 17 eligible instructors (eight lecturers, six at the rank of associate professor or above, and three teaching assistants). All had, within the previous year, either engaged with or explicitly refrained from using generative AI in lesson preparation, classroom instruction or course assessment, and collectively they covered the principal course modules, including Ideological and Moral Cultivation and the Rule of Law, Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics, Outline of Modern Chinese History and Situation and Policy. Inclusion criteria were: (1) current responsibility for teaching ideological and political theory courses at a university; (2) documented engagement with, instructional use of, or explicit refusal to use generative AI within the previous year; and (3) provision of written informed consent, including agreement to audio recording. Sampling and stopping were guided by information power; recruitment ceased when data adequacy was reached—i.e., when developing patterns were sufficiently rich and useful for the research questions. After the fifteenth interview, the marginal emergence of new codes declined, and by the seventeenth interview the thematic structure had stabilised.

All interviews were conducted by the first author using a semi-structured guide, in Chinese, either in on-campus meeting rooms or via an approved encrypted videoconferencing platform; each session lasted 45–120 minutes. The interview guide was informed by a prior literature review and policy documents. It began with basic demographic and professional information and proceeded to open-ended prompts covering: overall views of AI and the most recent episode of use or deliberate non-use; understandings and classroom enactment of boundaries regarding what is permitted, what requires disclosure and what is prohibited; task design following the introduction of AI and requirements for process evidence (e.g., prompts, conversational excerpts, version comparisons, reflective journals); strategies for guarding against factual hallucinations and bias, and for handling sensitive topics; conditions for platform access and resources (e.g., institutional accounts, concurrency and bandwidth, logging and student profiling); and organizational supports together with professional-development needs. With permission, interviews were audio-recorded in full and transcribed verbatim; transcripts were de-identified prior to analysis to protect confidentiality.

Data were analyzed using RTA following the six phases (Braun & Clarke, 2006, 2019): (1) familiarization with verbatim transcripts alongside analytic memoing; (2) generating initial codes across the semantic–latent continuum, applied line by line by the first author; (3) collating conceptually related codes into candidate subthemes and early theme clusters, accompanied by a preliminary thematic map; (4) reviewing extracts, codes and themes for internal coherence and boundary clarity, consolidating overlaps and retiring weakly evidenced candidates while returning to the corpus to check representativeness; (5) defining and naming stabilized themes with attention to mechanisms and potential transferability, supported by anonymized illustrative quotations; and (6) producing the analytic narrative in alignment with the research questions. NVivo 12 supported data management and queries.

Rigour was supported through reflexive memoing; an audit trail (interview-guide versions, coding records, thematic maps, decision memos); and peer debriefing with a second researcher to achieve negotiated coherence on coding boundaries and thematic logic. We also invited limited participant consultation by sending one-page theme summaries to a subset of participants to consider the resonance of interpretations, without positioning participants as arbiters of analytic correctness.

Researcher positionality and reflexivity. The first author is a university instructor-researcher with professional engagement in ideological and political education and an interest in how generative AI mediates curricular aims. This proximity afforded contextual and theoretical sensitivity but also the risk of interpretive over-familiarity. To manage this, reflexive memos documented assumptions, moments of surprise and shifts in interpretation across design, analysis and write-up; alternative readings were explored during peer debriefing; and disconfirming evidence was actively sought in the corpus. These procedures emphasized transparency and interpretive responsibility consistent with the RTA paradigm.

Iterative analysis yielded three themes: (1) attenuation of affective and faith dimensions; (2) content complexity and the limits of AI understanding; and (3) insufficiency of high-quality, compliant training data. Their scope, evidentiary bases and interrelations are detailed in the Findings.

Findings

Demographic characteristics

Among the 17 university instructors of ideological and political education included in this study (see Table 1), 13 were female and 4 were male. Eight held the rank of lecturer, six were at associate professor level or above, and three were teaching assistants. Participants ranged in age from 27 to 57 years. With respect to course responsibilities, two taught Ideological–Moral Cultivation and Legal Foundations, four taught Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics, three taught Outline of Modern Chinese History, five taught Current Affairs and Policies, and three taught Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era. Years engaged in ideological and political education ranged from 3 to 21.

Table 1: Social demography of participants (n = 17)

ID	Gender	Age(years)	Academic Rank	Primary Course Taught	Years in IPE Teaching
T01	Female	57	Associate professor or above	Current Affairs and Policies	4
T02	Female	41	Associate professor or above	Outline of Modern Chinese History	3
T03	Female	46	Lecturer	Ideological–Moral Cultivation and Legal Foundations	10
T04	Female	42	Lecturer	Current Affairs and Policies	21
T05	Male	33	Teaching assistant	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	12
T06	Male	38	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	7
T07	Male	51	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	3
T08	Female	30	Teaching assistant	Current Affairs and Policies	4
T09	Female	53	Associate professor or above	Outline of Modern Chinese History	10
T10	Female	31	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	4
T11	Female	27	Teaching assistant	Ideological–Moral Cultivation and Legal Foundations	4
T12	Female	41	Associate professor or above	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	9
T13	Male	43	Associate professor or above	Current Affairs and Policies	21

ID	Gender	Age(years)	Academic Rank	Primary Course Taught	Years in IPE Teaching
T14	Female	55	Lecturer	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	12
T15	Female	34	Associate professor or above	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	6
T16	Female	40	Lecturer	Outline of Modern Chinese History	7
T17	Female	50	Lecturer	Current Affairs and Policies	9

Thematic findings

This study identified three themes from the interview data: (1) the attenuation of affective and faith dimensions; (2) the complexity of substantive content and the limits of AI understanding; and (3) the scarcity of high-quality and compliant training data.

Theme 1: Attenuation of Affective and Faith Dimensions

Based on an in-depth analysis of interviews with 17 university instructors of ideological and political education, the application of generative artificial intelligence in classroom practice appears to produce a marked decoupling among cognition, emotion, and values. Most participants acknowledged the efficiency of artificial intelligence for presenting structured knowledge. At the same time, they described an educational gap that technology cannot close in the key moments of affective resonance, value guidance, and the formation of convictions. This finding aligns with the pedagogical expectation that ideological and political education should educate through persons and move through emotion. Digital tools may support making the logic clear at the cognitive level, but they are unlikely to accomplish reaching the heart at the level of value formation. The literature likewise cautions that integrating artificial intelligence into education should not reduce complex moral and civic formation to computable procedures, which risks marginalizing affective and value related work.

First, participants consistently reported that artificial intelligence cannot substitute for the teacher's embodied presence and context sensitive ethical judgment during interaction. Sensitivity to students' emotional states, careful handling of sensitive topics, and role modelling through words and deeds were described as necessary conditions for the internalisation of values. Such context bound commitments cannot be replaced by models that rely on statistical association.

AI can help me clarify the structure, but the emotional turning points and value hesitations still need to be caught by my gaze, pauses, and follow up questions in class. (T03)

When value conflicts arise, I have to judge the appropriate tone and boundaries on the spot. This is not a responsibility a model can assume in real time. (T14)

Second, the affective bridge from emotional resonance to rational endorsement is difficult to construct through artificial intelligence alone. Instructors widely regarded storytelling and personal reflection as pivotal narrative practices that connect affect and reason. Although artificial intelligence can supply background materials, it shows a distinctly nonhuman character in narrative vitality, emotional immersion, and real time responsiveness to the classroom climate.

AI speaks smoothly, but it lacks the human touch. The same case, when narrated by a person, carries a persuasive force that can light up students' eyes. (T07)

Immersive experiences supported by AI can spark emotions, but if they are not guided toward sound value judgments and historically grounded positions, the excitement becomes emotional movement without effect. (T11)

In addition, the surrounding algorithmic environment may erode the ecological basis for affective resonance even before teaching begins. Some instructors worried that prolonged exposure to platform recommendations creates information cocoons that narrow the breadth of students' intake and consolidate preexisting beliefs. This not only weakens the reach of mainstream values but also reduces the common vocabulary available for classroom discussion, thereby making it harder to build emotional connection and value consensus.

If students live in a recommendation driven comfort zone for too long, the common language in the classroom becomes scarcer, and emotional connection is more difficult to establish. (T05)

A further concern is that platform-oriented quantification can crowd out essential emotional labour. A number of participants observed that an emphasis on process data and visible interaction counts tends to shift effort toward what is easily measured, while undervaluing forms of labor that are crucial for value formation yet resistant to quantification, such as one to one conversations and context specific reflection.

The platform values logs and the number of interactions, so I tend to devote time to activities that leave traces in the system. Heart to heart conversations are vital, but they are difficult to quantify. (T02)

Theme 2: Complexity of Content and Limits of AI Understanding

Another core finding is that participants frequently identified a tension between the intrinsic complexity of political theory and the limited semantic capacities of artificial intelligence. Most instructors emphasized that the Marxist and Chinese socialist theoretical frameworks taught in ideological and political education constitute an integrated system that combines dialectical reasoning, historical development, and high sensitivity to context rather than a static set of definitions. Although AI can rapidly generate text that appears correct, it often exhibits excessive simplification, temporal and contextual misalignment, and arbitrary recombination when interpreting key concepts, tracing historical evolution, and capturing subtle discursive nuances. This finding aligns with the course principles of teaching in depth, with clarity and vitality, and with the unity of ideological orientation and scholarly rigor, which together require that knowledge construction be grounded in coherent theoretical reasoning rather than surface level aggregation of information.

In the interviews, instructors repeatedly noted that the historical development and contextual variation of core categories are frequent sources of distortion in AI generated content. Especially, AI models tended to splice discourse drawn from different historical periods or match terms to the wrong context:

Emphasis for the same term varies across historical junctures. Sometimes AI blends the discourse of the 1950s with contemporary policy language. (T06)

For instance, it tends to reduce "new quality productive forces" to technical productivity, which obscures the accompanying reconfiguration of production relations and the mechanisms of factor allocation. (T04)

In addition, instructors reported that chains of reasoning associated with dialectical logic are often reduced to binary, low tension treatments by AI. Several noted that when handling complex arguments such as the unity of opposites or the negation of the negation, models tended to compress dynamic, stage specific contradictions into static yes or no judgments, creating a line of reasoning that appears smooth but is in fact distorted:

It turns contradictions directly into choices, without the necessary historical and practical scaffolding. Students may feel that "the reasoning sounds right" but do not understand why it should be articulated in this way here and now. (T15)

Theme 3: Shortage of high-quality and compliant training data

Moreover, the integration of artificial intelligence into teaching ideological and political education faces a threefold bottleneck involving data, corpora, and compliance. The core tension is as follows. On the one hand, the theoretical system carried by these courses is highly disciplinary, holistic, and time sensitive, which imposes stringent requirements on the authority, accuracy, and traceability of teaching content. On the other hand, the Chinese language training corpora currently available for large AI models are marked by uneven quality, heterogeneous sources, and limited transparency in review processes. This state of corpus provision stands in structural tension with the normativity, unity, and ideological attributes of ideological and political education. Participants therefore generally held that without a high-quality, compliant, and authoritative Chinese corpus to anchor generation and inference, AI cannot meet the pedagogical requirement to teach in depth, with clarity, and with vitality.

This fundamental tension translates directly into a crisis of credibility and authority in classroom practice. Several instructors pointed out that the heterogeneity of corpora across platforms produces inconsistency in knowledge outputs and makes value orientation difficult to control. For example,

Different platforms provide two even opposite answers to the same theoretical question. It is difficult for me to convince students that AI is accurate. (T08)

What I fear most is content with no citation or outdated wording appearing in students' work. It looks polished, but once I ask for the source it falls apart. (T09)

For AI to support ideological and political education, rules and content supply must be aligned with the principle that ideological orientation and scholarly rigor remain unified. At the same time, the use of data must be strictly regulated within the legal and institutional framework to prevent information leakage and misuse. As participant noted:

There is no university-wide corpus for AI at present. Everyone uses different platforms with different data sources, so the risk points are not the same. (T04)

What we fear most is inconsistency between AI formulations and textbook descriptions. (T06)

Participants widely agreed that any model genuinely suited to political education in Chinese universities must be supported by a large scale, authoritative, rigorously reviewed, and continuously updated corpus. Building such a domain specific resource is essential, yet it currently faces substantial structural challenges of cost and complexity. As one instructor explained:

The human and financial resources needed to build such a model are beyond what a single university, or even several universities together, can address at this stage. (T17)

Discussion

Drawing on RTA of the interview materials, this study identifies three interlocking tensions that delineate the performance boundary of artificial intelligence as a form of "instrumental rationality" in university ideological and political education. First, there is an educative gap in the affective and conviction dimension. Although artificial intelligence can markedly increase the efficiency of knowledge transmission, the core process of value internalization, which rests on a chain of interaction grounded in trust, emotional resonance, and value commitment, continues to depend on the teacher's embodied presence and personal guidance. Excessive proceduralization and reliance on metrics risk reducing classroom practice to a merely computable process and marginalizing humanistic work such as care and the cultivation of conviction. This finding converges with scholarship that emphasizes educating through human relationships and affect, and underscores the need to position artificial intelligence as an assistant rather than a substitute. Second, there is a mismatch between theoretical complexity and generative mechanisms. The dialectical, historical, and context sensitive character of political theory sits in structural tension with generation based on statistical association, which makes artificial intelligence prone to conceptual simplification, temporal misalignment, and compression of reasoning chains. This observation echoes the pedagogical requirement to teach in depth, with clarity and vitality, and reaffirms that the unity of ideological orientation and scholarly rigor must be grounded in coherent theoretical support and an integrated logic. Third, there is a foundational constraint related to high quality and compliant corpora. Whether artificial intelligence can deliver accurate and stable outputs in this high stakes domain depends on the quality and compliance of its training data. The current heterogeneity of platforms and weak review pipelines directly introduce instructional risks and may erode classroom credibility and value guidance. Taken together, these results suggest that artificial intelligence in this domain should be understood as an amplifier of instrumental tasks that optimizes organization and presentation, but cannot by itself generate the value laden and humanistic outcomes that education seeks. Effective integration therefore requires a triadic anchor of teacher leadership, theoretical grounding, and compliance safeguards, in order to avoid the datafication, fragmentation, and decontextualization of a complex theoretical system.

The findings not only corroborate prevailing calls to move from isolated applications toward comprehensive enablement in educational uses of artificial intelligence, but also clarify the internal logic that such a shift requires. A pathway of holistic enablement must proceed in a coordinated manner across classroom practice, scholarly construction, and governance. At the classroom level, the priority is to preserve the irreplaceable work of affect and conviction; immersive and interactive experiences supported by artificial intelligence yield durable value only when teacher guidance and value clarification are present, which is consistent with evidence that education depends on human relationships and purpose rather than technical delivery alone (Biesta, 2009; Selwyn, 2019). At the level of scholarly construction, it is necessary to move beyond statistical association and to re-anchor instruction in the dialectical and historical specificity of theory. This requires designs that use controlled vocabularies, temporal mapping, and cross textual comparison to counter models' tendencies toward conceptual simplification and compressed chains of reasoning, a pattern well documented in research on hallucination and brittle reasoning in large language models (Bender et al., 2021; Ji et al., 2023). At the level of governance, capacity rests on building authoritative and compliant corpora together with mechanisms that enable retrieval augmented generation and verifiable citation, supported by transparency instruments such as datasheets for datasets and model cards, as well as privacy protection and auditability in educational contexts (Lewis et al., 2020; Gebru et al., 2021; Mitchell et al., 2019; Slade & Prinsloo, 2013; UNESCO, 2021). Without a curated whitelist knowledge base, retrieval augmentation, and citation binding, platform applications struggle to ensure content quality and remain vulnerable to drift in canonical formulations that can dilute the effectiveness of mainstream discourse (UNESCO, 2021; OECD, 2021).

In sum, to advance responsible integration of artificial intelligence in university ideological and political education, coordinated action is required across government, institutions, and classrooms. Government should promulgate domain-specific standards for data governance, privacy, and algorithmic accountability; fund an authoritative, continuously updated Chinese corpus aligned with curricular standards and vetted for ideological and scholarly integrity; and institute certification and periodic auditing of educational models and platforms. Universities should translate national guidance into enforceable rules by defining permitted, disclosure-required, and prohibited uses; establishing a curated, pre-approved knowledge base connected to retrieval and citation verification services; and implementing access control, logging, and incident reporting. Faculty development should prioritize algorithmic literacy, prompt design, supervision of student use, and assessment integrity, supported by toolkits such as controlled vocabularies, historical timelines, and cross-text comparison templates. Instructors should retain leadership in value guidance and theoretical interpretation while using artificial intelligence for organization, presentation, and feedback. Recommended practices include requiring process evidence, such as prompts, dialogue excerpts, version histories, and reflective notes; aligning generated materials with course glossaries and historical sequences; using cross-text triangulation to prevent conceptual simplification or temporal misalignment; explicitly teaching about recommendation mechanisms and bias; and converting immersive resources into value-oriented dialogue through guided questioning and brief oral defenses.

Naturally, this study has limitations. First, the sample was drawn from five universities in Chongqing, China, which may limit the geographic and institutional representativeness of the findings. Second, the study relied primarily on instructor interviews and did not include systematic classroom observations or process data on student learning. Future research should proceed in three directions: conducting comparative case studies across multiple regions and institutional types to test the generalizability of the conclusions; integrating classroom process data with evidence chains from student work and employing quasi experimental or quantitative designs to evaluate the effects of the proposed pedagogy on higher order thinking and academic integrity; and implementing instructional interventions that use a curated, institutionally approved knowledge base together with retrieval augmented generation to estimate causal effects on the quality and stability of AI generated content. These efforts would help validate and refine the proposed integration framework centered on teacher leadership, theoretical grounding, and compliance safeguards, and would support a shift from isolated pilots to genuine system level enablement.

Conclusion

Drawing on RTA of interviews with 17 ideological and political education instructors from five universities in Chongqing, this study identifies three structural tensions that delineate the limits of integrating generative artificial intelligence into this domain. First, although artificial intelligence can markedly improve the integration and presentation of knowledge, it cannot assume the core educative work grounded in trust, empathy, and value commitment, and proceduralized, metric driven practices risk marginalizing affective and conviction oriented dimensions of teaching. Second, the dialectical logic, historical framing, and contextual complexity of political theory are not well matched to statistical text generation, which fosters conceptual simplification, temporal misalignment, and compressed chains of reasoning, thereby weakening the theoretical foundations required to teach in depth, with clarity and with vitality. Third, the absence of large scale, authoritative, audited, and traceable Chinese corpora and corresponding compliance mechanisms generates instructional risks such as inconsistent answers to the same question and unattributed patching, which can erode classroom credibility. In response, the study proposes a system integration framework that positions artificial intelligence appropriately through teacher leadership, theoretical grounding, and compliance safeguards. On the institutional side, the framework calls for a curated whitelist knowledge base, retrieval augmented generation with citation binding, and mechanisms for logging, audit, and algorithmic transparency. On the pedagogical side, it recommends a human–AI collaboration chain that spans problem definition, prompt design, collaborative content development, fact verification, and reflective consolidation, together with a formative assessment system that draws on version comparison, oral explication, and reflective journals. On the capability side, it emphasizes controlled vocabularies, course outlines, and historical timelines to align concepts precisely and locate them in context, supported by communities of practice and model courses that disseminate reusable templates. In sum, the proper role of artificial intelligence is to amplify the instrumental tasks of exposition and organization rather than to replace the human core of education; only through the coordination of classroom practice, platform support, and institutional governance can technical potential be translated into high quality learning outcomes and credible value formation.

Funding

This work was supported by the Center for Theory and Practice of Chinese style Modernization, School of Marxism, Chongqing College of Humanities, Science & Technology, under grant 2024ZGSXDH11. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Central Committee of the Communist Party of China, & State Council. (2019). China's Education Modernization 2035. https://www.gov.cn/zhengce/2019-02/23/content_5367987.htm

[2] General Office of the CPC Central Committee, & General Office of the State Council. (2019). Opinions on deepening the reform and innovation of ideological and political theory courses in the new era. https://www.qstheory.cn/yaowen/2019-08/14/c_1124876471.htm

[3] Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Learning, Media and Technology, 44(2), 77–91. https://doi.org/10.1080/17439884.2019.1583671

[4] State Council of the People's Republic of China. (2017). New Generation Artificial Intelligence Development Plan. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

[5] Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4(5–6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816

[6] Luckin, H., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson. https://discovery.ucl.ac.uk/id/eprint/1475756/

[7] UNESCO. (2021). AI and education: Guidance for policymakers. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000376709

[8] 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? Computers & Education, 137, 103–121. https://doi.org/10.1016/j.compedu.2019.103615

[9] Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. Proceedings of the National Academy of Sciences, 115(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

[10] Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science, 348(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

[11] Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text based environment: Computer conferencing in higher education. The Internet and Higher Education, 2(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

[12] Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education.

The Internet and Higher Education, 7(2), 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001

[13] Mittelstadt, D. A., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1–21. https://doi.org/10.1177/2053951716679679

[14] OECD. (2021). OECD digital education outlook 2021: Pushing the frontiers with AI, blockchain and robots. OECD Publishing. https://doi.org/10.1787/589b283f-en

[15] Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510–1529. https://doi.org/10.1177/0002764213479366

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

[17] Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[18] Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

[19] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

[20] Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. Educational Assessment, Evaluation and Accountability, 21(1), 33–46. https://doi.org/10.1007/s11092-008-9064-9

[21] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

[22] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., & Madotto, A. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), Article 248. https://doi.org/10.1145/3571730

[23] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuang, L., Lewis, M., Yih, W.-T., Rocktäschel, T., & Riedel, S. (2020). Retrieval-augmented generation for knowledge-intensive NLP. arXiv. https://arxiv.org/abs/2005.11401

[24] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–229. https://doi.org/10.1145/3287560.3287596