Asia Pacific Science Press

Critical Humanistic Social Theory
ISSN(O):3005-9291
ISSN(P):3005-9283
Vol. 2 No.3 (2025)

Ethical Risks and Point-of-Use Governance of AI in IPE: Insights from Interviews with 17 Instructors

Zhihao Wei[1]*, Zhen Liu[1], Tao Wang[2], Lacong Yongzhen[3]

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

2.Guangzhou Xinhua University, Guangzhou, 523133, China

3.Lijiang Culture and Tourism College, Lijiang, 674199, China

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

Copyright: 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This study examines ethical risks and workable governance for artificial intelligence in university ideological and political education in China. Semi-structured interviews with 17 instructors from five universities in Chongqing conducted from March to June 2025 were analyzed using reflexive thematic analysis. Six themes characterize current practice: privacy and consent remain fragile in attendance, proctoring, and analytics; the teacher role shifts from authority to curator and ethical gatekeeper; recommendation and moderation shape visibility and the continuity of deliberation; assessment integrity benefits from process-based evidence and explicit disclosure; metric-driven activity targets can crowd out value reasoning; and governance and accountability depend on institutional capacity and consistent rules. The findings indicate that responsible integration requires governance at the point of use across classroom, platform, and institution, including course-level disclosure and granular consent, explainable moderation with instructor overrides and traceability, process-based assessment with AI-use disclosure, curated corpora linked to retrieval-augmented generation with citation binding, routine audits, and faculty training.

Keywords: Generative Artificial Intelligence (AI); Ideological and Political Education (IPE); Chinese Higher Education; Ethical Risks and Governance; Thematic Analysis

Published: Sept 24, 2025

DOI: https://doi.org/10.62177/chst.v2i3.599

1.Introduction

Artificial intelligence (AI) integration in education is rapidly expanding under policy impetus and institutional experimentation, driven by data-intensive learning ecosystems, widening access goals, and expectations for responsive governance. In China, national policy has positioned AI as a strategic driver of educational modernization and of IPE (Fuxiang & Shuangli, 2023). The New Generation Artificial Intelligence Development Plan (2017) set out a system-level blueprint for AI's role in public services and governance, including education (China's State Council, 2017). The General Offices of the CPC Central Committee and the State Council later issued the Opinions on Deepening the Reform and Innovation of Ideological and Political Theory Courses in the New Era (2019), calling for modern information technologies to be embedded in IPE to enhance depth, affinity, and effectiveness (General Office of the CPC Central Committee & General Office of the State Council, 2019). Recent directives continue this push: the Ministry of Education reports steady progress on the National

Education Digitalization Strategy Action, and a 2025 multi-agency opinion specifies accelerating education digitalization with intelligent technologies (Ministry of Education, 2024; Ministry of Education et al., 2025). Within this policy context, Chinese scholarship describes a shift in IPE from technology insertion to technology integration, supported by data, algorithms, and immersive scenarios that promise personalized pathways and expanded learning spaces. Rather than asserting structural necessity, this article offers context-bound empirical evidence from instructor interviews on how these trends translate into point-of-use practices and governance.

Despite these opportunities, embedded AI also introduces risks that are magnified by the epistemic and value characteristics of IPE. Studies warn that recommendation logics can narrow horizons, displace value-oriented dialogue, and erode teachers' discursive leadership, while datafication may crowd out humanistic work in the classroom (Slade & Prinsloo, 2013; Mittelstadt et al., 2016). IPE-specific analyses further document tensions between tool rationality and value rationality, concerns about privacy and algorithmic bias, and shifts in teacher–student relations as AI becomes a quasi-subject in interaction (Baker & Hawn, 2022; Holmes et al., 2022; Guilherme, 2019). Empirical work in "intelligent IPE" notes risks of information cocoons, fairness challenges tied to data provenance, and value drift when platforms prioritize engagement metrics over deliberative learning (Baker & Hawn, 2022; Gebru et al., 2021; Bender & Friedman, 2018). At the model level, large language models exhibit well-known tendencies toward hallucination and compressed reasoning chains, which are misaligned with IPE's demand for conceptual precision, historical sequencing, and cross-text consistency (Bender et al., 2021; Ji et al., 2023). These technical limits reinforce calls to anchor AI-supported IPE in authoritative, compliant corpora and to combine automation with teacher-led value guidance.

A clear research gap remains. Policy and conceptual essays have proliferated, yet systematic evidence from frontline IPE instructors on how ethical risks materialize and how safeguards can be made workable across classroom, platform, and institutional layers is comparatively limited. Existing reviews identify the promise of whole-process enhancement and the need for governance, but they seldom specify conditions under which AI helps or harms value formation in actual IPE settings (Zawacki-Richter et al., 2019; UNESCO, 2023). Field studies in China have begun to surface issues such as narrowed exposure, weakened affective engagement, and corpus heterogeneity, but they call for richer qualitative accounts to explain context, mechanisms, and boundary conditions. Recent qualitative work specifically on IPE underscores that credibility and authority are fragile without curated knowledge bases, retrieval-augmented citation binding, and auditability—yet these provisions are uneven across institutions and platforms.

The present study responds to this gap with an empirical investigation using thematic analysis of interviews with 17 university IPE instructors. First, the study delineates how teachers perceive and delimit ethical risks and value boundaries in AI-supported IPE, clarifying where instrumental gains meet pedagogical limits. Second, it explains mechanisms that generate these risks in context, including the interaction of data practices, recommendation logics, corpus governance, and classroom discourse. Third, it contributes an integration framework that centers teacher leadership, theoretical anchoring, and compliance safeguards, and it specifies operational elements such as retrieval-augmented generation with verifiable citation, process-based assessment integrity, and continuous audits for privacy and bias. This contribution directly addresses weaknesses identified in the literature by linking classroom practice to platform design and institutional policy within China's active policy environment.

2.Methods

This study adopted an interpretivist stance to understand how university IPE teachers make sense of AI in teaching and assessment. An interpretivist approach was appropriate because the research questions concern meanings, judgments, and practical reasoning in context. We used a qualitative exploratory design and applied reflexive thematic analysis to identify patterned meanings across participants' accounts. We followed the six-phase procedure described by Braun and Clarke: familiarization with the data, generation of initial codes, construction of candidate themes, review of themes, definition and naming of themes, and production of the report (Braun & Clarke, 2006; Braun & Clarke, 2021).

2.1 Participants and sampling

Participants were 17 IPE teachers from five universities. Teaching experience ranged from 3 to 21 years. We used purposive sampling to recruit instructors who had recent exposure to AI-supported teaching or management tools in IPE. We then used snowball sampling to expand the pool. Each participant was assigned an anonymous identifier from T01 to T17. We sought variation in institutional tier, course types, and prior AI tool experience so that the sample could capture a range of practices and views.

2.2 Data collection

Data were collected from March to June 2025 through semi-structured interviews with IPE instructors. Each participant completed one interview of 60 to 90 minutes. The interview guide covered AI use scenarios in IPE, perceived benefits and risks, data governance and privacy practices, assessment and academic integrity, teacher–student interaction, and institutional support. With written informed consent, interviews were audio-recorded and transcribed verbatim. Direct identifiers were removed during transcription, and role or context details were generalized when needed to reduce re-identification risk.

2.3 Data analysis technique

We employed reflexive thematic analysis to interpret participants' accounts (Braun & Clarke, 2006). Transcripts were read repeatedly to build immersion, and initial coding proceeded inductively at both semantic and latent levels. NVivo 12 was used to organize excerpts, memos, and iterative code sets. Codes were collated into candidate themes by comparing meanings within and across cases, and themes were refined by checking internal coherence and clear distinctions against the whole corpus. Theme salience reflected qualitative judgment rather than frequency counts; we considered three criteria: coverage across cases and contexts, explanatory reach for the research questions, and clarity of the mechanisms described by the data. We judged the analysis to be interpretively sufficient when later coding cycles added nuance without altering core patterns. We did not calculate inter-coder reliability because reflexive thematic analysis treats coding as an interpretive act; instead, we held regular analytic meetings, kept reflexive journals, and maintained an audit trail to support credibility and transparency (Nowell et al., 2017). The team had prior experience in IPE teaching and educational technology; we documented assumptions about AI benefits and risks before coding and revisited them during analysis, and reflexive memos recorded how our positionality shaped attention and the resolution of interpretive tensions.

2.4 Ethics

Before data collection, all participants received information about the study purpose, procedures, potential risks, and their rights, including voluntary participation and the option to withdraw at any time without penalty. Written informed consent was obtained for participation and audio recording. Personal identifiers were removed during transcription. Audio files and transcripts were stored on an encrypted drive with access limited to the research team. Data were used only for academic research and teaching improvement. The study protocol received approval from the authors' institutional research ethics committee.

3.Findings

3.1 Demographic characteristics

Among the 17 university instructors of ideological and political education in this study (see Table 1), 13 were female and 4 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 and had 2–21 years of IPE experience. Course responsibilities were distributed as follows: 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. Regarding prior use of AI/AI-enabled tools, 12 instructors reported some use and 5 reported none; reported tools centered on LMS analytics dashboards (n = 4) and platform recommendation/moderation (including override/appeal features; n = 4), with additional mentions of a generative assistant for classroom prompting (n = 1), an automated quiz engine (n = 1), workflow automation for pacing/scheduling (n = 1), and biometric access control for platform log-ins (n = 1).

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

ID	Gender	Age	Rank/Role	Main Course	Prior AI Use (Yes/No; tool)	Years in IPE
T01	Female	57	Associate professor or above	Current Affairs and Policies	No; None reported (addressed AI-assisted answers via process-based assessment)	4
T02	Female	41	Associate professor or above	Outline of Modern Chinese History	Yes; LMS analytics dashboards (plus permitted generative brainstorming with disclosure)	3
T03	Female	46	Lecturer	Ideological–Moral Cultivation and Legal Foundations	Yes; Generative assistant for classroom prompting (source-tracing prompts)	10
T04	Female	42	Lecturer	Current Affairs and Policies	Yes; Platform moderation & bias-check settings (trained via targeted clinics)	21
T05	Male	33	Teaching assistant	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	Yes; Platform biometric login (facial/voice) for access control	2
T06	Male	38	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	No; None reported (used version history + brief viva; non-AI feature)	7
T07	Male	51	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	Yes; LMS analytics dashboards (weekly targets)	3
T08	Female	30	Teaching assistant	Current Affairs and Policies	Yes; Platform recommender/curation features	4
T09	Female	53	Associate professor or above	Outline of Modern Chinese History	No; None reported (issued privacy/consent notice)	10
T10	Female	31	Lecturer	Introduction to Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era	Yes; Automated quiz engine	4
T11	Female	27	Teaching assistant	Ideological–Moral Cultivation and Legal Foundations	Yes; Platform moderation & appeal workflow	4
T12	Female	41	Associate professor or above	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	Yes; LMS analytics dashboards (shared screenshots)	9
T13	Male	43	Associate professor or above	Current Affairs and Policies	No; None reported (affected by provider terms change)	21
T14	Female	55	Lecturer	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	Yes; Moderation override with rule/keyword display	12
T15	Female	34	Associate professor or above	Introduction to Mao Zedong Thought and the Theoretical System of Socialism with Chinese Characteristics	Yes; Automation for pacing/task flow (auto-quizzes/scheduling)	6
T16	Female	40	Lecturer	Outline of Modern Chinese History	Yes; LMS analytics dashboards (rebalanced via reflection tasks)	7
T17	Female	50	Lecturer	Current Affairs and Policies	No; None reported	9

Note: “Prior AI Use” reflects tools directly used in teaching/learning as evidenced in instructor interviews. LMS analytics/targets and platform recommendation/moderation are counted as AI-enabled features. Version history and ordinary document tools are not counted as AI.

3.2 Thematic findings

Analysis of interviews with 17 IPE teachers generated six connected themes that capture how AI is used, evaluated, and governed in practice. Themes are ordered by salience in the corpus. Illustrative interview excerpts are formatted with anonymized identifiers (T01–T17).

Theme 1: Privacy and consent

Attendance, proctoring, and analytics tools collect biometric and behavioral data. Consent is often presented once at first login with few item-level opt-out options. Retention and secondary use are not always clear. Concerns increase when engagement screenshots or rankings circulate beyond the course. Some instructors use plain-language notices, itemized consent choices, and masked identifiers in exported reports.

Access to core functions required facial capture and voice activation. Consent appeared once at first login and most students clicked through to enter class. (T05)

Analytics screenshots were later shown in a college meeting. Students asked whether the initial consent covered this second use. (T12)

A course notice listed collected items, retention time, and an opt-out for facial capture. Questions became more specific and several students opted out while staying engaged. (T09)

We got an email saying our data would be shared with a “third-party partner” for “analytics purposes.” What partner? What purpose? It feels like we’re the product, and I don’t know who they’re selling us to. (T10)

The popup had a huge block of text and then “I Agree.” To get to the course materials, I had to agree. It mentioned eye-tracking, keyboard analysis, all of it. What if I was okay with the camera but not with them tracking my keystrokes? There was no way to say that. It was all or nothing, so I just clicked. (T11)

Theme 2: Teacher role

Teachers reported efficiency in preparation, examples, and feedback. Discussion sometimes shifted from dialogue to answer seeking. Many accounts described a move from sole authority to curator and ethical gatekeeper when students brought fluent but superficial outputs. The shift was smaller when AI use stayed in preparation and seminars emphasized argumentation.

Students arrived with compact answers from the assistant. Initial discussion focused on unpacking assumptions and rebuilding the chain of reasoning. (T01)

Automation supported pacing and task flow. Value clarification still depended on presence, tone, and carefully chosen stories. (T15)

Framing the model as a claim to test and using prompts such as justify and trace sources helped sustain teacher leadership. (T03)

For me, the role is the same. I use AI for prep materials, but the seminar is all about human debate. The AI isn’t part of that, so my role as facilitator hasn’t really shifted. (T08)

A student brought in a flawless summary, but when I asked about a key assumption, they couldn’t answer. The AI gave them the polish without the reasoning, and we had to reconstruct it together in class. (T11)

Theme 3: Recommendation and moderation

Recommendation and moderation shaped what students saw and what persisted. Repeated clicks on historical cases produced narrowing and theoretical readings with competing views surfaced less. Flags and removals without reasons interrupted threads and discouraged follow-up, especially near assessments. Brief explanations and instructor overrides with an audit trail were linked to fewer disruptions and higher trust.

Students who preferred historical cases kept receiving similar items. Theoretical texts with contested viewpoints were harder to surface. (T08)

A post about debate boundaries was marked sensitive without explanation. The appeal took a week and the thread lost

momentum. (T11)

When the platform showed the matched rule and keywords and allowed instructor restore with a reason, complaints fell and posts returned faster. (T14)

We had a great discussion thread going with study tips for the final. Two days before the exam, the whole thing was locked for a 'code of conduct violation.' No one knew why. Suddenly our main resource for last-minute questions was gone. The panic was real. (T17)

In my class seminar, the system is much better. It'll flag a post but it tells me and the student exactly why. I have a dashboard where I can review it and hit 'restore' in seconds. Because the process is transparent, students are more willing to tackle sensitive topics instead of shying away. (T06)

Theme 4: Assessment integrity

Automated scoring and rapid feedback increased pace and coverage. The same tools enabled AI-assisted paraphrasing and answer generation when grading focused only on final products. Many courses used process-based evidence such as drafts, version histories, planning memos, source trails, and short oral defenses to make authorship and learning visible. Limited and disclosed brainstorming followed by in-class defense was seen as consistent with integrity once expectations were clear.

Automated quizzes saved time. Regenerated short answers made ownership difficult to judge from the final script alone. Process evidence became necessary. (T10)

Combining version history with a brief viva revealed whether arguments were understood and owned. (T06)

Allowing AI for brainstorming with mandatory disclosure and in-class defense kept efficiency while aligning with integrity. (T02)

We used a mandatory 'Research Plan Memo' to capture the student's own thesis and source ideas before they drafted the paper. (T16)

In the oral defense, I just ask about their counterarguments and source choices. It quickly makes genuine ownership clear. (T09)

Theme 5: Metrics and values

Activity dashboards and weekly targets supported monitoring and pacing. They could displace value reasoning when treated as the main goal. Counts of clicks, posts, or minutes online did not show whether students could reason through value conflicts, engage opposing views, or justify a position. Reported countermeasures included reflective journals, short position statements with textual evidence, and rubrics that assessed stance clarity, engagement with counter-arguments, and quality of justification. Courses that displayed indicators of breadth and depth, such as diversity of sources, reported a more balanced focus.

High activity numbers did not show whether students could reason through value conflicts or hold a position with reasons. (T02)

Weekly targets produced visible activity but not necessarily conviction. Students learned to complete steps without engaging value questions. (T07)

Adding a reflection component and discussing excerpts publicly rebalanced the dashboard's influence. (T16)

Our rubric for the position statements was simple: ten points for a clear thesis, ten for using textual evidence, and ten for seriously addressing one counter-argument. It shifted their focus from post count to argument quality. (T05)

Theme 6: Governance and accountability

Institutional capacity and policy coherence shaped responsible use. Participants described uneven digital and ethical literacy, inconsistent guidance across departments, and reliance on vendor defaults. They also noted sudden changes to export formats or consent terms. Helpful supports included short clinics on privacy controls, bias checks, and appeal workflows. A standing cross-unit group aligned teaching affairs, IT security, legal counsel, and student affairs. An internal policy listed permitted, restricted, and prohibited coursework uses.

Procurement language, platform terms, and course rules pointed in different directions. When dilemmas arose, it was unclear which rule prevailed. (T07)

A change in provider terms affected data access and consent language. The institutional response lagged behind classroom needs. (T13)

Targeted clinics on privacy controls, bias checks, and moderation appeals made a practical difference. Instructors felt more confident in daily decisions. (T04)

Our policy has a 'Restricted Uses' section. For example, AI can generate feedback on drafts, but only if the final grade is determined by the instructor and students are notified of the tool's use in the syllabus. (T15)

The Teaching Center runs 30-minute clinics every month. The one on 'Checking for Algorithmic Bias' was really useful. It gave us a simple checklist to use when choosing new software. (T01)

4.Discussion

This study asked how university IPE teachers perceive and delimit ethical risk when using AI, what mechanisms generate these risks in teaching practice as reported by instructors, and what measures are workable at the course, platform, and institutional levels. Thematic analysis identified six findings. First, privacy and consent are fragile because access to core functions depends on bundled, one-time consent and because secondary use of data is not always transparent. Second, the teacher role shifts from sole authority to curator and ethical gatekeeper, especially when students arrive with fluent yet shallow outputs. Third, algorithmic recommendation and moderation shape what becomes visible and discussable, which narrows exposure and produces contestable takedowns. Fourth, assessment integrity is challenged by AI-assisted authorship, and can be supported by process evidence and bounded, disclosed use. Fifth, activity metrics can displace value formation unless courses adopt indicators that capture depth and justification. Sixth, responsible use depends on institutional capacity, including clear rules, training, cross-unit coordination, and timely responses to vendor changes.

These results are important because they link policy goals with classroom practice. National plans call for deep integration of intelligent technologies in education and IPE to expand access and improve quality (China's State Council, 2017; General Office of the CPC Central Committee & General Office of the State Council, 2019; Ministry of Education et al., 2025). The findings indicate that such integration is effective only when governance mechanisms are embedded at the point of use. Course-level data notices and itemized consent make privacy control substantive. Brief justifications for moderation decisions, together with instructor overrides, preserve space for reasoned disagreement. Process-based assessment and disclosure maintain efficiency while supporting credible authorship. Rubrics that emphasize clarity of stance, engagement with opposing views, and use of evidence align student activity with the mission of IPE to cultivate judgment and commitment rather than mere participation counts.

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). 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.

5.Limitations and future recommendations

This study draws on semi-structured interviews with 17 IPE instructors from five universities in Chongqing. The reliance on instructor self-reports and the single-region sample limit the transferability of the findings beyond similar institutional and policy contexts. No classroom observations or student-produced artifacts were collected, which constrains triangulation across data types and may leave some mechanisms inferred rather than witnessed in situ. The cross-sectional design also precludes claims about change over time or causal effects on learning outcomes and academic integrity behaviors.

Future work should broaden the sampling frame across regions and institutional types to test the scope conditions of the themes identified here. Mixed-methods designs that pair interview data with classroom observations and process evidence from student work would strengthen credibility and enable richer mechanism tracing. Where feasible, quasi-experimental or quantitative evaluations could estimate effects on higher-order thinking, integrity-related behaviors, and teacher–student interaction. Finally, instructional trials that integrate an approved institutional knowledge base with retrieval-augmented generation—and that report transparent prompts, outputs, and audit trails—can assess the quality and stability of AI-supported content. Together, these steps would help validate and refine a teacher-led, theory-grounded, compliance-aware integration framework and support movement from small pilots to coherent, system-level adoption.

Conclusion

In conclusion, the study sheds light on a critical intersection of educational modernization and value-oriented teaching. The findings underscore the practical challenges faced by instructors who are expected to integrate intelligent tools while protecting privacy, sustaining dialogue, and upholding academic integrity. These challenges call for a coordinated response that combines clear classroom protocols, transparent and controllable platform functions, and institution-level safeguards and training. With these supports in place, AI can contribute to access and efficiency while remaining aligned with the core aims of ideological and political education.

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.

Conflict of Interests

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

Reference

[1] Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. International Journal of Artificial Intelligence in Education, 32(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

[2] Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. Transactions of the Association for Computational Linguistics, 6, 587–604. https://doi.org/10.1162/tacl_a_00041

[3] Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (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.

[4] 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

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

[6] Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

[7] Fuxiang Dong & Shuangli Dong. (2023). Research on the optimization of ideological and political education in universities integrating artificial intelligence technology under the guidance of curriculum ideological and political thinking. ACM Transactions on Asian and Low-Resource Language Information Processing. Advance online publication. https://doi.org/10.1145/3611012

[8] Dong, Y. (2024). Ethical risks and countermeasures of ideological and political education integrated with artificial intelligence. Journal of Beijing University of Aeronautics and Astronautics (Social Sciences Edition), 37(1), 160–165. https://link.oversea.cnki.net/doi/10.13766/j.bhsk.1008-2204.2022.0955

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

[10] 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.gov.cn/zhengce/2019-08/14/content_5421252.htm

[11] Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. Field Methods, 18(1), 59–82. https://doi.org/10.1177/1525822X05279903

[12] Guilherme, A. (2019). AI and education: The importance of teacher and student relations. AI & Society, 34(1), 47–54. https://doi.org/10.1007/s00146-017-0693-8

[13] Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? Qualitative Health Research, 27(4), 591–608. https://doi.org/10.1177/1049732316665344

[14] Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

[15] 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

[16] Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in neural information processing systems, 33, 9459-9474. https://doi.org/10.48550/arXiv.2005.11401

[17] Ministry of Education et al. (2025). Opinions on accelerating education digitalization [In Chinese]. https://www.gov.cn/zhengce/zhengceku/202504/content_7019045.htm

[18] Ministry of Education of the People's Republic of China. (2024). National education digitalization strategy action

continues to advance [In Chinese]. https://www.moe.gov.cn/jyb_xwfb/s5147/202406/t20240612_1135098.html

[19] 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. In FAT '19: Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287596

[20] Mittelstadt, B. D., 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

[21] Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16, 1–13. https://doi.org/10.1177/1609406917733847

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

[23] Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1509–1528. https://doi.org/10.1177/0002764213479366

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

[25] UNESCO. (2021). AI and education: Guidance for policy-makers. https://doi.org/10.54675/pcsp7350

[26] UNESCO. (2021). Recommendation on the ethics of artificial intelligence. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligenc

[27] UNESCO. (2023). Global education monitoring report 2023: Technology in education—A tool on whose terms? https://doi.org/10.54676/BIUM3029

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