Digital Intelligence International Development Education Alliance

北京大学 PEKING UNIVERSITY

Guide to the Boundaries of

ARTIFICIAL INTELLIGENCE

Applications in Higher Education 2.0

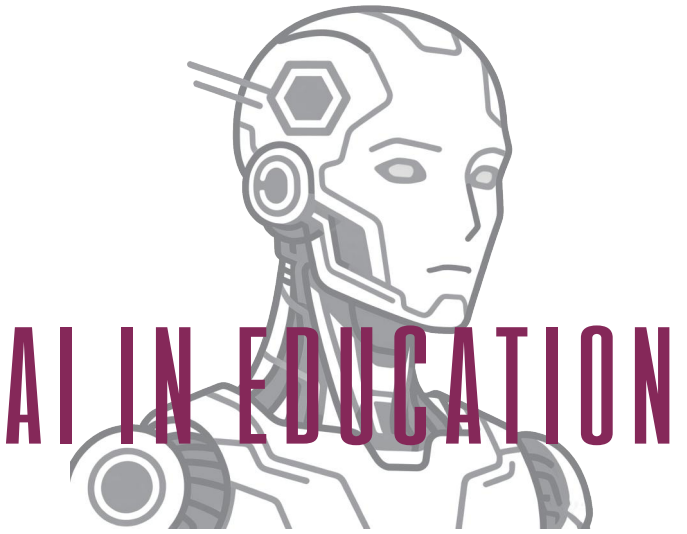
The Essential Edition

2025.11

Digital Intelligence International Development Education Alliance

DI-IDEA Guideline Research Group

Preface



Artificial Intelligence (AI) is rapidly transforming global industries and social structures. As a representative branch built upon large language models technologies, Generative Artificial Intelligence, with its powerful capabilities in content creation and information synthesis, has opened new avenues for innovation in higher education, fostering human–AI collaborative classrooms, personalized learning, and emerging models such as "AI for Science." However, the lack of clearly defined application boundaries has also given rise to concerns including academic misconduct, the erosion of students' critical thinking, and data security risks in teaching and research. How to rationally steering technology while preserving the essence of education has become a pressing global challenge for higher education.

UNESCO's Guidance for Generative AI in Education and Research (2023) establishes the fundamental position for generative AI applications, emphasizing a human-centered principle: technology should serve human development, assist rather than replace teachers, learners, and advance fairness and inclusiveness in education. It also defines key norms such as data privacy protection, age-appropriate use, ethical accountability, and educational adaptability. The United States and the European Union have likewise issued strategic plans and reports that specify AI usage regulations and transparency requirements. China's Plan for Building an Education Power (2024–2035) promotes the integration of AI across the entire educational process, while universities such as Oxford, Stanford, Peking University, and Tsinghua University have also issued institutional policies. Although these global practices have begun to take shape, higher education still faces unresolved challenges: blurred application boundaries, inconsistent standards, and weak risk -control mechanisms constrain the realization of AI and pose challenges to the fundamental mission of fostering virtue through education.

In response, the DI-IDEA developed the Guide to the boundaries of AI Applications in Higer Education 2.0. At its core, the Guideline aims to redefine AI from an auxiliary "tool" to a collaborative "partner" in teaching and learning, establishing a triadic relationship among teachers, learners, and AI systems. Guided by the vision of "co-creating a new human–AI ecosystem in higher education," it sets out six guiding principles to promote the responsible use of AI, identifies eight categories of potential risks — including technological limitations, content bias, privacy breaches, and academic integrity concerns — and analyzes eight representative scenarios across key stages such as pre-class preparation, classroom engagement, post-class learning, and research exploration. Finally, the Guideline advances the concept of an education-oriented and adaptively bounded use of AI, proposing concrete recommendations for teachers, students, and administrators: teachers should model ethical guidance and learner-centered pedagogy, students should practice self-regulation and literacy development, and administrators should foster an open, inclusive, and safeguarded environment. Together, these actions aim to enhance the quality of education and innovation in research within the intelligent era.

In the future, the Guideline will be continuously refined and updated in alignment with technological progress and educational practice, fostering the coordinated development of intelligent technology and the essential mission of cultivating people.

FUNDAMENTAL PRINCIPLES

PRINCIPLE 1

Promoting Fairness and Inclusive Access

Support the reasonable use of AI tools by teachers and students in teaching and research, avoid both outright prohibition and unrestricted openness. To achieve a balance between technological inclusiveness and effective boundaries, differentiated regulations should be developed based on disciplinary characteristics and course objectives, thereby promoting the joint enhancement of educational quality and research effectiveness.

PRINCIPLE 2

Education-Oriented and Human-Centered

Education must remain rooted in the mission of nurturing learners. The central roles of teachers and students in the educational process should be fully affirmed. AI should be positioned as an assistive participant that supports teaching and research through human–AI collaboration, fostering pedagogical innovation and improving research quality while ensuring that technology ultimately serves the fundamental essence of education.

PRINCIPLE 3

Literacy Development and Risk Awareness

A systematic mechanism for cultivating AI literacy should be established to help teachers and students understand the principles and limitations of technology and to strengthen their capacity for critical evaluation. A regularized risk identification system should be implemented to raise awareness of potential risks such as technological dependency and data leakage.

PRINCIPLE 4

Security Assurance and Privacy Protection

Establish qualified standards for trusted AI platforms and clear regulations for data usage to define the safety baseline of human–AI collaboration, providing a reliable safeguard for teaching and research activities.

PRINCIPLE 5

Clear Responsibility and Ethical Commitment

A framework of clearly defined responsibilities, transparent processes, and traceable outcomes should be established. Teachers and students should actively define the boundaries of AI use based on specific teaching and research contexts, take responsibility for the quality and compliance of AI-generated content, and uphold academic integrity and ethical standards throughout the process.

PRINCIPLE 6

Collaborative Innovation and Value Amplification

Emphasize the establishment of a new human–AI partnership and an innovative ecosystem. By leveraging a continuous feedback loop of design, practice, and reflection, it aims to achieve the sustained improvement of teaching and research quality, promoting the co-evolution of educational systems and AI capabilities.

POTENTIAL RISKS

01

Risk of Technological Limitations

Systemic risks may arise from algorithmic design flaws, limited data-processing capacity, or inadequate adaptation to complex educational contexts, leading to decreased accuracy and reliability in AI-based teaching applications.

02

Risk of Content Bias

Biased training data, algorithmic preferences, or flaws in generative mechanisms may lead to outputs that diverge from teaching objectives, academic standards, or factual accuracy, posing systemic risks.

03

Risk of Information Leakage

Inadequate security measures, weak management systems, or improper operations may expose sensitive personal information and critical data of teachers and students to unauthorized access, posing an information leakage risk.

04

Risk of Cognitive Inertia

Excessive reliance on AI assistance may discourage active thinking, deep inquiry, and critical analysis among teachers and students, leading to reduced cognitive engagement and the development of habitual dependence on AI systems.

05

Risk of Academic Misconduct

Improper use of AI may obscure the originality of academic outputs and undermine scholarly norms, creating systemic risks for individuals, institutions, and the academic community.

06

Risk of Teaching Ineffectiveness

Errors, functional limits, or misuse of AI may reduce teaching efficiency, compromise instructional quality, and hinder the achievement of educational goals, posing systemic risks.

07

Risk of Competence Erosion

When students rely on AI to perform essential learning processes—such as independent thinking, experiential trial and error, and critical construction—the development pathway of their professional competence may be disrupted, weakening the foundation of disciplinary literacy and practical capability.

08

Risk of Educational Inequality

Variations in AI access, literacy, or algorithmic bias may generate new inequities in learning resources and opportunities, undermining fairness and inclusiveness in education.

CASE DECODING

CASE 1

AI-Assisted Lesson Preparation — Integrating Frontier Knowledge

Background

In traditional lesson preparation, teachers must manually search for literature, filter materials, and organize content on their own—a process that is time-consuming, labor-intensive, and prone to information overload, which may cause key findings to be overlooked. Moreover, integrating interdisciplinary knowledge often proves challenging. By contrast, AI, with its powerful information processing and pattern recognition capabilities, can rapidly select relevant academic papers and extract key content, efficiently assisting teachers in acquiring and integrating the necessary instructional resources.

Case Description

Mr. Liu, a lecturer of Advanced Mathematics, aimed to make his course more engaging by introducing frontier examples on the "intersection of AI and mathematics." Using an AI teaching assistant platform, he retrieved recent papers, classroom cases, and practice exercises, greatly saving preparation time. However, during class, students pointed out that key data in the AI-generated example were unreliable and inconsistent with current scholarly consensus. Mr. Liu realized that AI tools face clear limitations in academic source tracing and accuracy of frontier content transmission, and that AI-generated materials must be cross-verified against authoritative databases to ensure scholarly rigor.

Risks Identified

- **Risk of Content Reliability**

AI-generated summaries of frontier literature may contain discrepancies between core data and original conclusions. Owing to its probabilistic generation logic, AI often lacks reliable academic traceability and may overlook authoritative research sources.

- **Risk of Academic Misconduct**

When teachers cite AI-generated content without verification, they may unintentionally commit academic misconduct by reproducing biased or inaccurate information.

- **Risk of Knowledge Obsolescence**

AI models are trained on data with inherent update cycles. Newly published research may not be promptly included, leading to retrieval results that fail to capture the latest academic developments.

- **Risk of Technological Dependence**

The efficiency and convenience of AI tools may foster psychological dependence among teachers, diminishing their motivation to actively track scholarly progress and engage in in-depth reading of original literature.

Recommendations

- **Ensure Human Verification**

All key data, examples, and references generated by AI should be verified through authoritative academic platforms or expert review before use to ensure accuracy and reliability.

- **Standardize Source Attribution**

In course materials and presentations, clearly indicate which sections were assisted or generated by AI. This practice not only aligns with academic integrity standards but also models responsible technology use for students.

- **Focus on Teaching Quality**

Teachers can use AI to save time on routine tasks such as literature organization, allowing more effort to be devoted to instructional design and pedagogical innovation, thereby improving teaching clarity and learning effectiveness.

- **Clarify Primary Responsibility**

AI may serve as a collaborative partner in lesson preparation; however, teachers must apply their own professional judgment to review AI-generated content and bear final responsibility for the accuracy and quality of all instructional materials.

CASE DECODING

CASE 2

AI-Assisted Pre-Class Preparation — Reviewing Learning Materials

Background

During pre-class preparation, students often face extensive and complex reading materials, and traditional reading methods can be time-consuming and labor-intensive. AI tools can rapidly generate literature summaries, outlines, and term explanations, helping students efficiently construct a preliminary knowledge framework.

Case Description

Wang, an undergraduate student, used an AI tool to process several English articles assigned by his instructor in order to save time during pre-class preparation. He asked the AI to extract key ideas and technical terms, and the tool quickly produced summaries and definitions. Satisfied with the results, Wang believed he had fully understood the readings. However, during class discussion, he was unable to follow when the instructor explained a central concept. Wang realized that excessive reliance on AI-generated summaries without critical engagement had created an illusion of mastery, which affected his comprehension and subsequent learning.

Risks Identified

- **Risk of Hallucinated or Inaccurate Generation**
  When processing learning materials, AI may produce inaccurate or fabricated information due to "hallucination" effects, leading to students' misconceptions and distorted understanding.
- **Risk of Logical Distortion**
  AI tools may fail to accurately interpret the complex logical relationships within academic texts, preventing students from constructing a coherent and accurate conceptual framework.
- **Risk of Learning Dependence**
  Students who rely entirely on AI for pre-class preparation without deep reading or critical reflection may weaken their capacity for autonomous learning and critical thinking.
- **Risk of Learning Ineffectiveness**
  Errors in AI-generated information may require additional time and effort to correct, reducing overall learning efficiency and comprehension quality.

Recommendations

- **Ensure Human Verification**
  After obtaining AI-generated results, students should engage in close reading and comparison with the original sources to verify the accuracy of core concepts and key formulas. They must correct any potential errors from AI outputs to ensure the reliability of acquired knowledge and lay a solid foundation for deep learning.
- **Clarify Individual Responsibility**
  Students should strengthen their awareness of AI literacy and take active responsibility for their own learning outcomes. They must not attribute misunderstandings or learning gaps to AI limitations but instead reflect on whether excessive reliance or misuse has hindered their independent learning.
- **Focus on Competence Development**
  Students should identify their learning priorities and explore how to apply AI appropriately to enhance understanding and problem-solving, ensuring that technology serves as a means to strengthen — not replace — their cognitive and analytical abilities.
- **Set Autonomous Boundaries**
  Students should define clear boundaries for AI use — determining which stages of learning can be supported by AI tools and which must be completed independently — to maintain ownership and authenticity in the learning process.

2

CASE DECODING

CASE 3

AI-Assisted Classroom Teaching — Enhancing Interactive Learning

Background

In large-scale university classrooms, teachers often find it difficult to monitor every student's level of understanding in real time, while students struggle to receive personalized feedback. AI-powered teaching assistants can enhance classroom interaction by providing instant feedback and support, helping to establish a triadic "teacher–student–AI" interaction model that improves teaching efficiency and student engagement.

Case Description

In a Product Design course, Professor Chen asked students to submit conceptual sketches and written explanations of their product ideas. After collecting the assignments, she uploaded all submissions to an AI teaching assistant to save evaluation time. The AI quickly categorized the works by theme and generated comments. After a brief review, Professor Chen accepted the AI's results, planning to showcase examples from each category and use the AI's feedback to guide class discussion. However, during the subsequent class presentation, she discovered several obvious errors in the AI's analysis. For instance, one "environmental product design" was mistakenly classified as "smart home appliances," and the AI even fabricated a "modular design concept" that the student had never proposed. These inaccuracies caused confusion in the classroom discussion and disrupted the normal teaching flow.

Risks Identified

- **Risk of Hallucinated Generation**

When processing visual and textual assignments, AI systems may experience hallucination errors, producing incorrect categorizations or fabricated comments that directly compromise teachers' accurate assessment of students' academic performance.

- **Risk of Teaching Efficiency and Quality Loss**

If teachers adopt AI-generated evaluations without human verification, inaccurate feedback may be presented in class or used to guide instructional adjustments, preventing students from receiving meaningful guidance and thereby reducing teaching efficiency and quality.

- **Risk of Undermined Teacher Authority**

When students identify errors in AI-generated analyses, they may begin to question the teacher's professional judgment and subject expertise, which can disrupt classroom order and erode the trust essential to the teacher–student relationship.

Recommendations

- **Appropriately Define AI's Pedagogical Role**

Teachers should recognize AI as an assistive tool rather than a decision-making authority. Core instructional judgments must be made by teachers themselves, grounded in professional expertise, disciplinary knowledge, and understanding of students' learning needs.

- **Strengthen Human Verification**

Before applying AI-generated results, teachers should conduct sampling checks and targeted reviews to identify and correct errors, ensuring that AI outputs contain no misclassifications or fabricated content.

- **Optimize Interaction Design**

When evaluating complex visual or written assignments, teachers may adopt a "AI pre-screening plus human evaluation" model to improve the accuracy and validity of assessments.

- **Ensure Data Security Compliance**

Both teachers and students should use platforms with compliant data protection agreements and avoid uploading assignments containing personal or sensitive information to ensure data security and privacy.

CASE DECODING

AI

CASE 4

AI-Assisted Experimental Instruction — Supporting Virtual Simulation

Background

In experimental teaching, certain chemistry and biology experiments are difficult to conduct due to expensive reagents, high operational risks, or strict environmental requirements. AI-driven virtual simulation systems, with their highly realistic environments and real-time interactivity, can effectively replicate actual experimental processes and provide immediate feedback on procedural accuracy—reducing costs and safety risks while improving teaching efficiency and laboratory safety.

Case Description

Li, a student enrolled in an advanced organic synthesis course, needed to complete a complex experiment involving costly reagents and toxic solvents. For safety reasons, he first conducted a preliminary practice session on an AI-based virtual simulation platform. The AI assistant guided him through each step, providing real-time feedback and indicating reagent names and operations. Believing the AI feedback sufficient for error prevention, Li completed the simulation successfully and received a "Operation Successful" prompt. However, upon reviewing the operation log, the instructor discovered that Li had used an incorrect reagent concentration. This caused significant deviations in conversion rates and product purity within the simulated reaction, yet the AI system failed to detect the error. In a real experiment, such a deviation could have led to failed synthesis or even serious safety hazards.

Risks Identified

- **Perception and Recognition Error Risk**

The AI system lacked sensitivity to critical parameters such as reagent concentration, relying only on surface-level name matching and thus failing to provide accurate data validation.

- **System Robustness and Reliability Risk**

The system did not intercept operations that appeared correct but were in fact erroneous, nor did it recognize subtle deviations, reflecting insufficient reliability and a potential breach of laboratory safety standards.

- **Loss of Teaching Efficiency and Quality**

Incorrect AI feedback forced students to repeat experiments, wasting both time and laboratory resources.

- **Cognitive Process Outsourcing Risk**

Real-time algorithmic judgment replaced students' own anticipation and evaluation of experimental risks, weakening their development of essential safety awareness and experimental control skills.

Recommendations

- **Strengthen Risk Awareness**

Students should develop a mindset of "virtual assistance plus real-world verification." They should proactively check parameters after simulations, while teachers review logs and present real accident cases to enhance students' risk perception.

- **Reinforce Educational Anchoring**

Teachers should position virtual simulation as a conceptual learning tool. Supplementing simulations with recorded or live demonstrations can bridge sensory experience gaps and help students build a coherent "virtual–real–application" cognitive chain.

- **Uphold Responsibility and Ethics**

Students must take initiative in verifying experimental details and reflecting on procedural logic. Teachers should pay particular attention to reviewing blind spots in AI assessment to prevent incorrect feedback from undermining teaching integrity.

- **Promote Human–AI Synergy**

Teachers and students should employ a "virtual experiment plus real operation" collaborative model to achieve unified goals of risk reduction, efficiency improvement, and competency development.

CASE DECODING

CASE 5

AI-Assisted Environmental Monitoring — Inspecting Experimental Materials

Background

In experimental courses such as chemistry and biology, ensuring the accuracy and quality of experimental materials is fundamental to maintaining safety and the reliability of results. Traditional manual inspection is prone to oversight due to visual errors or limited experience and is difficult to conduct in real time. By leveraging image recognition and Internet of Things (IoT) technologies, AI systems can scan reagent labels and equipment parameters, compare them with preset protocols, and issue real-time alerts. These capabilities enable teachers to monitor laboratory activities remotely and safeguard both experimental safety and instructional quality.

Case Description

During an organic synthesis experiment in a smart laboratory, Lin, a student, accidentally picked up a reagent bottle that looked similar but contained the wrong chemical. The AI system, using its camera-based recognition module, immediately identified the error, issued a warning, and prompted Lin to select the correct reagent. However, in a later stage of the experiment, Lin used another reagent that had degraded over time, resulting in reduced purity. Because the AI system lacked the capacity to detect micro-level quality changes, it failed to recognize the issue or issue any alert. The experiment repeatedly produced substandard results until the instructor identified the cause through manual inspection. Lin ultimately realized that while AI can perform basic visual and compositional checks, it cannot replace human expertise in evaluating microscopic quality variations or latent material issues.

Risks Identified

- **Perception and Recognition Error Risk**

The AI system lacks the capability to perceive hidden quality indicators such as reagent purity or sample freshness, making it unable to accurately assess intrinsic material properties.

- **System Robustness and Reliability Risk**

When facing non-explicit material quality issues, the AI's verification mechanism fails, preventing effective detection or early warning.

- **Loss of Teaching Efficiency and Quality**

Students misled by the AI's undetected errors may repeatedly fail experiments, wasting both time and resources and significantly reducing instructional efficiency.

- **Cognitive Process Outsourcing Risk**

Overreliance on AI's basic judgments can inhibit students' development of critical skills in risk anticipation, material evaluation, and experimental reasoning.

Recommendations

- **Strengthen Risk Awareness**

Teachers and students should recognize the limitations of AI in material monitoring and adopt a mindset of "AI-assisted screening plus manual verification" to avoid sole dependence on AI outcomes.

- **Clarify Core Responsibility**

AI should be treated as a preliminary monitoring tool. Teachers and students remain ultimately responsible for the quality control and risk management of experimental materials and must not delegate these responsibilities to AI systems.

- **Maintain an Education-Centered Focus**

AI applications should serve the goal of cultivating scientific rigor. Teachers can use case analysis and guided practice to help students build a complete cognitive chain from material inspection to experimental reasoning.

- **Promote Collaborative Enhancement**

Students can reasonably rely on AI for preliminary material identification, while focusing their own effort on advanced tasks such as evaluating material quality and understanding experimental principles, thereby improving the overall effectiveness of laboratory learning.

CASE DECODING

CASE 6

AI-Assisted Post-Class Assessment — Evaluating Programming Code

Background

In computer science courses, the evaluation of programming assignments is a key process for testing students' coding proficiency and logical thinking. AI-assisted assessment can perform syntax checks, basic functionality validation, and code style analysis through static analysis and automated testing, thereby reducing teachers' workload.

Case Description

To improve the efficiency of code evaluation, Professor Zhang uploaded a student's final project code to an AI assessment platform. The AI-generated report indicated that the code contained no syntax errors and passed basic tests but exhibited logical redundancy and poor readability. However, upon manual inspection, Professor Zhang discovered that the AI had failed to identify deeper logical flaws and algorithmic inefficiencies in handling edge cases. More importantly, the code's unusually consistent formatting and uniform structure lacked the typical traces of individual problem-solving and reasoning found in student-written code. This raised his suspicion that the project might have been generated by AI. After discussing the matter with the student, Xiao Ming, Professor Zhang found that he could not explain the core logic of his own program. Xiao Ming eventually admitted that the project code had indeed been generated by an AI system rather than written independently.

Risks Identified

- **Risk of Logical and Knowledge Comprehension Limitations**
  AI systems struggle to detect deep logical flaws or inefficiencies in algorithm design, leading to inaccurate or incomplete evaluations.
- **Risk of Teaching Efficiency and Quality Loss**
  Current AI assessment tools cannot reliably determine whether code is AI-generated, thereby diminishing the validity and educational value of the evaluation process.
- **Risk of Assessment Reliability and Outcome Bias**
  Overreliance on AI-generated evaluations may prevent teachers from accurately gauging students' genuine mastery and practical abilities.
- **Risk of Academic Misconduct**
  Students may use AI to generate code that passes automated evaluation undetected, potentially encouraging dishonest practices.

Recommendations

- **Enhance Human Review**
  Teachers should recognize the limitations of AI and adopt a mindset of "AI assistance plus human verification." Particular attention should be paid to detecting logical flaws, assessing algorithmic efficiency, and verifying code originality through manual inspection.
- **Clarify Teachers' Core Responsibility**
  Teachers bear ultimate responsibility for the accuracy of evaluation results and for judging students' true competence. They must uphold academic integrity and safeguard fairness in assessment.
- **Return to the Essence of Education**
  Teachers should adjust assessment strategies, shifting the focus from merely determining whether "the code runs" to evaluating students' comprehensive abilities in problem analysis, logical design, and algorithm optimization.
- **Leverage Collaborative Value**
  Teachers can rely on AI to efficiently handle basic evaluation tasks while devoting more time and effort to deeper instructional work — such as algorithm refinement and logic analysis — achieving a synergy of AI-enabled efficiency and human-driven depth.

CASE DECODING

CASE 7

AI-Assisted Review and Reinforcement — Consolidating Key Knowledge Points

Background

In post-class review, students often engage in self-testing, practice correction, and other consolidation activities that can be time-consuming and labor-intensive. AI tools can rapidly generate diverse self-assessment questions based on uploaded learning materials, provide instant grading and error analysis, and recommend targeted exercises according to performance. This helps students efficiently strengthen weak areas and enhance retention.

Case Description

When reviewing Ancient Chinese History, a student named Li uploaded his lecture notes to an AI system to generate self-test questions in order to improve study efficiency. The AI quickly produced a test paper with corresponding standard answers. Seeing that the questions covered the lecture content and the answers appeared well-organized, Li was satisfied with the AI-generated review material and used it directly for memorization without manual verification.

During the final exam, he encountered several questions that resembled those from the AI-generated quiz and answered them based on the memorized AI responses. However, after the exam, he learned from classmates that his answers significantly deviated from the correct information. Upon checking the textbook, Li discovered that the AI-generated questions contained factual errors regarding the timing of historical events and the relationships between key figures. Because he had not verified the material earlier, these mistakes ultimately led to point deductions in his exam.

Risks Identified

- **Risk of Hallucinated Content Generation**

Due to "hallucination" effects, AI may fabricate or distort facts, producing inaccurate information related to key knowledge points. Students who rely on such outputs risk constructing flawed knowledge systems.

- **Risk of Superficial Learning**

Blind trust in AI-generated content may discourage active verification, giving students a false sense of mastery and impairing the depth of learning.

- **Risk of Cognitive Outsourcing**

When students delegate essential processes — such as comprehension, logical reasoning, and fact-checking — to AI, their capacity for independent thinking and self-directed learning diminishes.

Recommendations

- **Reject Knowledge Outsourcing**

Students must take personal responsibility for core learning processes such as understanding key concepts and organizing logical structures. Through active thinking and practice, they ensure the authenticity and depth of knowledge internalization.

- **Maintain Critical Thinking**

Students should not accept AI-generated content uncritically. They should consult textbooks and authoritative academic sources for cross-verification, comparing multiple perspectives to distinguish between accurate and misleading information.

- **Prevent Cognitive Inertia**

Teachers should guide students to treat AI as a "learning companion." By incorporating diverse evaluation methods — such as classroom questioning, practical tasks, and open-ended assignments — teachers can encourage students to actively engage in knowledge construction, preventing overreliance on AI that may lead to cognitive passivity and reduced learning outcomes.

CASE DECODING

CASE 8

AI-Assisted Scientific Research — Reading Academic Literature

Background

Researchers often struggle to extract key information and synthesize insights efficiently from vast amounts of academic literature. AI-powered literature review tools can generate summaries, keywords, key arguments, and research methods, as well as categorize viewpoints to help scholars quickly understand their field, identify research gaps, and accelerate the research process.

Case Description

During a study on "smart classrooms," a graduate student named Li uploaded a large collection of academic papers to an AI platform for extraction and summarization. The AI swiftly produced a detailed summary report. However, upon deeper reading, Li discovered that the AI's extracted core ideas were biased—it had omitted crucial experimental data and even fabricated comparative conclusions between two papers. Realizing that AI systems are prone to "hallucination" and logical inconsistency when processing complex scholarly content, Li ultimately returned to the original papers for close reading and verification.

Risks Identified

- **Risk of Hallucinated Generation**
  AI's limited understanding of academic logic may cause it to overlook essential premises, misinterpret findings, or even generate false conclusions.
- **Risk of Logical Distortion**
  When integrating multiple papers, AI may lack coherence and academic reasoning, resulting in disorganized aggregation or omission of key arguments.
- **Risk of Reduced Research Efficiency and Quality**
  Although AI aims to save time, researchers may end up spending more effort cross-checking and correcting its errors, undermining overall efficiency.
- **Risk of Cognitive Outsourcing**
  Overreliance on AI-generated summaries weakens researchers' independent reading, synthesis, and critical analysis skills, leading to the erosion of core academic competence.

Recommendations

- **Strengthen Risk Mitigation**
  Researchers should recognize the technical limitations of AI and actively cross-check AI-generated findings with original sources, verifying key arguments and critical data to prevent errors at their root.
- **Clarify Final Responsibility**
  Researchers remain the ultimate bearers of responsibility for academic conclusions and must not delegate core scholarly judgments or interpretive authority to AI systems.
- **Reject Capability Substitution**
  AI can assist with repetitive tasks such as literature categorization and basic information extraction, but essential processes—such as validating core arguments and conducting critical academic analysis—must be carried out independently by the researcher.
- **Balance Efficiency and Quality**
  Researchers may leverage AI to handle time-consuming preliminary tasks, allowing them to focus their efforts on higher-order activities such as deep analysis and scholarly innovation.

RECOMMENDATIONS FOR IMPLEMENTATION

This Guideline emphasizes an education-oriented and adaptively bounded approach to AI application, offering practical recommendations for teachers, students, and administrators to promote responsible and effective human–AI collaboration in higher education.

For Teachers: From Knowledge Transmitters to AI-Integrated Learning Designers and Student Mentors

Teachers are encouraged to move beyond the traditional role of knowledge transmitters and become facilitators of AI-integrated instructional design and guardians of student development.

1. Demonstrative Guidance

Teachers should lead by example, clearly identifying AI contributions and delineating reasonable human–AI divisions of labor. Through designing comparative tasks, verifying AI outputs, providing in-depth explanations, and emphasizing formative feedback, teachers can guide students to critically evaluate and integrate AI-generated content, transforming AI into an effective partner for mutual growth and professional development.

2. Education-Centered Practice

AI integration into teaching must remain grounded in the fundamental mission of "moral and intellectual cultivation." Teachers should employ differentiated assessments across disciplines, include non-AI evaluation components, and strengthen the manual review of subjective assignments. The focus should be on assessing students' logical reasoning, creative thinking, and problem-solving abilities to ensure the solid mastery of professional knowledge and the attainment of high-quality learning outcomes.

For Students: From Passive Learners to Autonomous, Critical Co-Creators

Students should shift from passive knowledge receivers to active, responsible participants capable of setting boundaries, engaging in self-directed learning, and exercising rational critique.

1. Autonomous Boundary Setting

Students should always use AI with the goal of enhancing learning outcomes, consciously defining its appropriate scope of use. While drawing on AI for inspiration and assistance, they must maintain agency in the construction of knowledge and avoid reliance that weakens deep thinking and understanding, ensuring that AI serves learning objectives and cognitive development.

2. AI Literacy and Reflective Competence

Students should actively learn about AI's underlying mechanisms, recognize its limitations and risks, and develop verification habits suited to different learning tasks. They should seek teacher guidance, periodically engage in non-AI training sessions to reinforce core abilities, and record and reflect on their AI use to continuously optimize learning strategies — transitioning from passive adoption to critical, informed engagement.

For Administrators: From Regulators to Ecosystem Builders, Resource Enablers, and Risk Mitigators

Administrators should evolve from top-down regulators into enablers of open, sustainable environments that balance innovation, equity, and accountability.

1. Inclusive and Open Governance

Within the framework of academic integrity, institutions should encourage innovative AI practices suited to disciplinary contexts. This includes developing tiered and categorized regulatory systems, establishing pilot evaluation and dissemination mechanisms, and supporting the deep integration of AI into teaching and research processes — fostering a sustainable and collaborative ecosystem for responsible AI use among faculty and students.

2. Secure and Equitable Implementation

Institutions should build or adopt secure, compliant AI platforms to ensure equitable access to resources. They should systematically design AI literacy curricula to strengthen rational understanding and risk awareness among faculty and students. Robust governance mechanisms must be established to conduct pre-assessment and continuous supervision of critical AI applications that affect data security or academic integrity.

These recommendations integrate core principles with actionable practices, providing an operational framework for all stakeholders in higher education. Together, they aim to foster the healthy development of a human–AI co-creative ecosystem, achieving synergy between technological innovation and the enduring mission of education.

The Guide to the Boundaries of AI Applications in Higher Education 2.0 developed by the project team, is grounded in the core vision of "co-creating a new human–AI ecosystem for higher education." The Guideline systematically establishes a comprehensive framework encompassing principles, risks, scenarios, and implementation recommendations, providing valuable references for institutional policy development and pedagogical practice in universities.

We sincerely invite education authorities, industry experts, and members of the academic community to share their insights, feedback, and exemplary cases of AI application in higher education. The project team will carefully consider and integrate contributions from all stakeholders to further refine and enhance the Guideline, promoting both the ethical governance and innovative application of AI in higher education.

Guideline Research Group

Email: gdsjglyyyhy@pku.edu.cn

DI-IDEA

Email: diidea@pku.edu.cn

Please scan the QR code above to access the full Guideline.

Digital Intelligence International Development Education Alliance

北京大学 PEKING UNIVERSITY

高等教育人工智能应用边界指南2.0

Guide to the Boundaries of Artificial Intelligence Applications in Higher Education 2.0

精华版

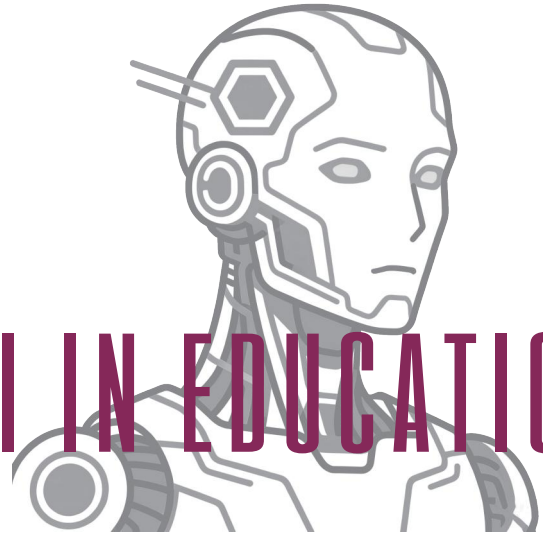
2025.11

数智教育发展国际大学联盟（DI-IDEA）

指南研制课题组　编研

前言

PREFACE



AI IN EDUCATION

人工智能（Artificial Intelligence, AI）技术的飞速演进正深刻重塑全球产业格局与社会形态。基于大语言模型的生成式人工智能（Generative Artificial Intelligence, GAI）凭借强大的内容生成与信息整合能力，为高等教育教学创新与科研活动变革开辟空间，推动人机协同课堂、个性化培养、“AI for Science”等新形态落地。但同时，因AI应用边界不明，也引发了学术不端、学生批判性思维弱化、教学科研数据泄露等问题。如何“理性掌控技术、坚守教育本质”成为全球高等教育的难题。

联合国教科文组织在2023年《Guidance for Generative AI in Education and Research》中，确立“以人为本”核心原则，明确数据隐私保护、使用年龄限制、伦理与教育适配性等关键应用规范；美国、欧盟分别通过战略计划、报告明确AI使用规范与透明度要求；中国以《教育强国建设规划纲要（2024—2035年）》等政策推动AI融入教育全流程；英国牛津大学、美国斯坦福大学及国内北京大学、清华大学等高校也相继出台相关规范。尽管全球实践初见成效，但高等教育AI应用仍存边界界定不明、标准规范不一、风险防控薄弱等问题，制约了技术价值释放，也对“立德树人”的根本任务构成挑战。

基于此，数智教育发展国际大学联盟牵头研制《高等教育人工智能应用边界指南2.0》。核心在于将AI从辅助“工具”升级为师生协作“伙伴”，构建“教师—学习者—人工智能”三元协同关系，并围绕“构建人机共创的高等教育新生态”这一核心理念，提出六大应用原则，为高等教育中人工智能的合理使用提供行动遵循；系统梳理技术局限、内容偏差、隐私泄露、学术诚信等八类潜在风险，为安全可控的人机协同提供风险防范依据；围绕“课前准备—课堂实施—课后学习—科研探索”等关键教育环节，遴选八类典型应用场景与案例，为原则落地提供具体参考；最后，强调“育人为本的动态适配AI应用边界”，从教师、学生、管理者多主体出发提出具体举措：教师应注重“示范引导、育人为本”，学生须践行“自主立界、素养建构”，管理者要营造“开放兼容、保障使用”，共同服务于高等教育质量提升与科研创新的核心目标。

未来，我们将结合技术演进与教育实践，在坚守育人本质的前提下对指南内容持续进行迭代优化，推动智能技术与教育使命的协同发展。

基本原则

PRINCIPLES

原则 1

公平普惠与多元包容

支持师生在教学科研中合理运用AI，避免简单禁止或完全开放。为实现技术包容与有效边界的平衡，需依据学科特性与课程目标制定差异化规范，推动教育质量与科研效能同步提升。

原则 2

教育为本与人本导向

坚持以育人为根本，明确师生在教育过程中的主体地位。将AI定位为协作性教育参与者，通过人机协同实现教学方法创新与科研质量提升，确保技术应用始终服务于教育本质。

原则 3

素养形成与风险意识

构建系统化AI素养培养机制，引导师生理解技术原理与应用局限，形成批判性评估能力。建立常态化风险识别体系，强化技术依赖、数据泄露等潜在风险的防范意识。

原则 4

安全稳健与隐私保护

确立可信AI平台的准用标准、明确数据使用规范来划定人机协作的安全底线，为教学科研提供可靠保障。

原则 5

责任明晰与伦理坚守

构建“权责明确、过程透明、结果可溯”的责任框架。师生应根据具体教学场景与目标，主动界定AI使用边界，对其生成内容的质量与合规性负责，恪守学术诚信规范与科研伦理准则。

原则 6

协同共创与价值增效

协同共创强调建立新型的人机伙伴关系和创新生态系统。依托设计、实践和反思的闭环机制，实现教学科研质量的持续提升，推动教育体系与AI能力的协同进化。

潜在风险

RISK

01

技术局限风险

因算法设计缺陷、数据处理能力或复杂场景适配不足，导致在教学应用中出现准确性下降、可靠性降低等问题的系统性风险。

02

内容偏差风险

因训练数据片面、算法设计倾向或生成机制缺陷，导致输出内容偏离教学目标、学术规范或客观事实的系统性风险。

03

信息泄漏风险

因技术防护不足、管理制度缺失或操作不规范等原因，导致师生个人敏感信息或重要数据被未授权获取、使用或披露的潜在威胁。

04

认知惰性风险

因过度依赖AI协助而减少主动思考、深度探究与批判性分析等认知活动，逐渐形成思维依赖的潜在威胁。

05

学术诚信风险

因不当使用AI技术，导致学术成果原创性难以界定、学术规范被突破，进而对个体、机构及学术共同体造成系统性危害的风险。

06

教学成效风险

因AI的内容错误、功能局限或使用不当，导致教学效率下降、教学质量受损且难以实现预设教育目标的系统性风险。

07

能力养成风险

因将本应独立亲历的自主思考、实践试错与批判性建构等关键学习环节借助AI完成，导致学生专业能力的养成路径中断、专业素养根基不牢的风险。

08

教育公平风险

因AI技术获取差异、使用能力差距或算法设计倾向等因素，导致不同群体在教育资源获取和教育发展机会上出现新的不平等，从而破坏教育公平性的系统性威胁。

CASE DECODING

案例解码

AI

1

案例一

AI辅助课前备课——整合前沿内容

背景 在传统的备课过程中，教师需手动检索文献、筛选资料并自行整理，过程耗时耗力，且易因信息过载而遗漏关键成果，跨学科知识的整合难度高；而AI则凭借其信息处理与模式识别能力，可快速完成相关学术论文的选取与内容提炼，高效协助教师获取与整合所需的教学参考内容。

案例 刘老师在为高等数学课程备课时，为增强教学趣味性，计划引入“AI与数学交叉应用”的前沿案例，于是借助AI助教进行智能检索。AI助教迅速生成一份涵盖最新论文摘要、典型案例及配套习题的参考资料，极大地节省了备课时间。然而，在课堂讲授过程中，当他引用AI所生成案例中的数据时，被学生指出这些数据问题且与当前学术内容存在偏差。这让刘老师意识到，AI助教在学术文献溯源、观点精准传递上有明显局限，其生成的前沿内容需经权威数据库交叉验证，才能有效确保内容的准确性与学术严谨性。

风险

内容可靠性风险： AI生成的前沿文献摘要存在核心数据与结论脱离原始语境等问题，因基于概率生成逻辑缺乏可靠学术溯源，且易遗漏权威研究。

学术不端风险： 教师未经核实直接引用存在偏差的AI生成内容，实质上构成了非主观意愿的学术不端。

知识滞后风险： AI的训练数据存在更新周期，对于刚发表的最新前沿成果可能存在收录延迟，导致检索结果无法反映学科最前沿动态。

技术依赖风险： AI的高效易用可能导致教师形成依赖心理，降低主动追踪学术动态、深入研读原始文献的动力。

建议

坚持人工核查： 对AI提供的关键数据、案例和参考文献，需通过权威学术平台或人工专业性进行核实，确认无误后再使用。

规范使用标注： 在课件和教学材料中，建议明确标注AI辅助生成的内容范围。这既符合学术规范，也为学生树立正确使用技术的榜样。

聚焦教学质量： 教师可借助AI节省处理文献整理等基础工作的时间，将精力投入到内容讲授模式设计等工作，让学生能更易懂易学，切实提升教学质量。

明确主体责任： 教师可将AI定位为备课协作伙伴，但需基于自身专业能力把关AI生成内容与质量，并对教学内容的准确性负最终责任。

本案例及其它课前备课场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

AI

案 例 二

AI辅助课前预习——浏览学习资料

背景 在课程预习环节，学生常面临海量且复杂的阅读材料，需要快速梳理知识框架、提炼关键信息，传统阅读方式既耗时又费力；而AI可快速生成文献摘要、梳理知识大纲并解释专业名词，助力学生高效搭建初步知识框架。

案例 王同学在预习《数据挖掘》课程时，需阅读教师布置的10篇英文定量文献。为节省时间，他使用AI辅助阅读，要求其提炼核心思想、梳理内容关联并解释专业术语。AI快速生成摘要、关联逻辑与名词解释，王同学阅读后十分满意，认为已掌握学习资料的核心内容，因此未对照原始文献核查细节，也未进行深度阅读梳理知识关联。但课堂上教师深入讲解某篇文献的核心算法时，他完全无法理解相关逻辑，对核心概念了解也不够准确。王同学这才察觉过度依赖AI预习且无批判意识，会产生“虚假掌握”错觉，进而影响后续学习。

风险

虚假生成风险： AI处理资料时，易因“幻觉”生成不准确信息，若学生未加核查直接采信，易引发认知偏差甚至错误。

逻辑失真风险： AI难以精准拆解文献等学习资料中的复杂逻辑关系，输出的关联梳理可能存在偏差，阻碍学生构建正确知识框架。

学习依赖性风险： 学生若完全依赖AI完成预习，省略深度阅读与批判性思考环节，长期下来会逐渐弱化自主学习能力。

学习效果受损风险： AI提供的信息若存在偏差或错误，学生需额外耗费时间核查修正，反而增加时间成本，降低学习效果。

建议

坚持人工核查： 学生获AI结果后，需通过深度阅读、对照原文，核查核心概念与关键公式准确性，修正AI可能的偏差，确保接收的知识真实可靠，为深度学习打基础。

明确主体责任： 学生需强化AI应用素养意识，对自身学习成效主动担责，不可将知识理解不足归咎于AI缺陷，主动反思是否过度依赖或使用不当行为，牢牢掌控学习主动权。

聚焦能力养成： 学生需结合自身，明确学习重点，进而思考并践行AI的合理利用方式。

自主设立边界： 学生需自主设立 AI 的使用边界，明确哪些学习环节可借助 AI 辅助、哪些必须独立完成。

本案例及其它课前预习场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

AI

案例三

AI辅助课堂教学——辅助课堂互动

背景 在大学课堂的规模化教学过程中，教师难以及时掌握所有学生的理解情况，学生也难以获得个性化反馈；而AI助教可通过即时反馈等方式辅助课堂互动，构建“师—生—机”三元互动结构，以提升教学效率与学生参与感。

案例 《产品设计》课程的陈老师在布置课堂作业时，要求学生提交产品概念草图及设计说明。陈老师回收作业后，为节省评阅时间，将全班同学的作品上传至AI助教。AI随即完成主题分类与点评，陈老师初步浏览后便采纳了AI生成的结果，计划先展示不同分类下的作业案例，再结合AI点评引导学生讨论。然而，在后续课堂讲解时，陈老师才发觉AI的分析结果出现明显错误，例如将一份“环保类产品设计”错误归类到“智能家电类”，并虚构了学生并未采用的“模块化设计理念”进行点评。这些错误引发了课堂讨论的混乱，打断了正常的教学节奏。

风险 **虚假生成风险：**AI在处理图文作业时可能出现幻觉，生成错误的分类与点评，会直接干扰教师对学生学业表现的准确判断。

教学效率与质量损失风险：若教师未对AI输出的错误评估结果进行人工审核，直接将其用于课堂反馈或教学调整，学生将无法获得针对自身问题的真实有效指导，会降低教学效率与质量。

教师专业权威受挑战风险：AI的错误分析一旦被学生察觉，可能引发学生对教师判断力、专业能力的质疑，进而影响课堂秩序稳定，直接冲击课堂中的教育信任关系。

建议 **合理定位AI角色：**教师需明确AI为辅助角色而非决策主导，核心的教学判断仍需由教师基于专业经验、学科认知及对学生的了解来完成。

强化人工核查环节：教师在使用AI结果前应通过抽样比对、重点复核的方式排查偏差，发现问题及时修正，确保AI输出无错误归类、虚构内容等问题。

优化互动环节设计：教师在审核图文复杂任务时，可采用“AI初筛+人工校评”模式以提升准确性。

严守数据安全底线：师生应选择具备合规数据保护协议的平台，避免上传包含个人信息的作业内容。

本案例及其它课堂教学场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

AI

案 例 四

AI辅助实验指导——辅助虚拟仿真

背景 在实验教学中，部分化学、生物等实验因试剂昂贵、操作风险高或环境要求严苛而难以开展；而AI虚拟系统凭借高仿真环境与实时交互能力，能模拟真实实验流程并即时反馈操作正误，既降低成本与风险，又提升教学安全与效率。

案例 李同学需要完成一项涉及昂贵试剂和有毒溶剂的复杂有机合成实验，为安全起见，他先利用AI虚拟仿真平台进行前期操作训练。李同学登录系统后，AI助手全程提示试剂名称与操作步骤。他认为AI的实时反馈能有效规避低级错误，快速完成实验并收到“操作正确，实验成功”的提示页面后，便未进一步核查细节。但指导教师复核详细的操作日志时发现，小李选用了错误浓度配比，导致后续虚拟反应的转化率、产物纯度等参数出现较大偏差，而AI系统未能识别此细节。若在真实实验中，这一偏差可能导致实验失败甚至严重的安全事故。

风险 **感知与识别错误风险：** AI系统对试剂浓度这类参数缺乏识别能力，仅能匹配名称等表面信息，无法为实验提供精准数据支撑。

系统鲁棒性与可靠性风险： AI系统未能拦截“表面正确但实质错误”的操作，也未能识别细微偏差，可靠性不足，易偏离实验安全规范。

教学效率与质量损失： AI系统给出错误反馈后，学生需重新开展实验，造成时间与实验资源的双重浪费。

认知过程外包风险： AI算法的实时判定替代了学生对实验的风险预判，可能导致学生实验风险把控等关键能力退化。

建议 **深化风险意识：** 学生需树立“虚拟辅助+现实验证”意识，在实验后主动核查参数；教师需复查实验日志、展示真实风险场景，并通过展示真实风险场景，强化学生风险认知。

锚定育人核心： 教师应将虚拟仿真定位为原理学习工具，通过实验录像、现场演示弥补感官体验，帮助学生建立“虚拟—现实—应用”的完整认知链。

坚守责任伦理： 学生应主动核对实验细节、分析操作逻辑；教师需重点复核AI系统评估存在的盲区，避免错误反馈影响教学。

深化协同增效： 师生应通过“虚拟实验+现实操作”协同模式，实现“降风险、提效率、育素养”的统一。

本案例及其它实验指导场景案例可扫描上方二维码查看详情

4

CASE DECODING

案例解码

案例五

AI辅助环境监测——监测实验材料

背景 在化学、生物等实验教学中，保障实验材料的准确性与质量是确保安全和结果可靠性的前提。传统人工核查易因视觉误差或经验不足出现疏漏，且难以实时监控；而AI可借助图像识别、物联网等技术，扫描试剂标签与设备参数，与预设方案比对并实时预警，辅助教师进行远程监控，从而保障实验安全与质量。

案例 林同学在智慧实验室进行有机合成实验时，AI系统先通过摄像头识别出他误拿了外观相似但成分错误的试剂瓶，及时发出警告并提示正确试剂；但在后续实验中，当小林使用另一瓶因存放过久导致纯度下降的试剂时，AI系统因无法检测微观质量变化，并未识别出试剂问题，也未发出任何提醒。最终实验结果始终不达标，直到人工凭经验排查才发现是试剂纯度问题。这让小林意识到，AI仅能完成基础的外观、成分等显性检测，无法替代人工对微观质量、隐性问题的专业判断。

风险 **感知与识别错误风险：**AI系统对剂纯度、样品新鲜度等隐性质量指标缺乏感知能力，无法精准判断材料内在属性。

系统鲁棒性与可靠性风险：AI系统在面对材料品质变化这类非显性问题时，其验证机制失效，无法提供有效预警或判断。

教学效率与质量损失风险：学生因AI系统未识别材料隐性问题而被误导，导致实验反复失败，既浪费时间与实验资源，也大幅降低教学效率。

认知过程外包风险：学生过度依赖AI系统的基础判断，可能抑制自身实验风险预判、材料质量鉴别等核心能力的形成。

建议 **树立风险意识：**师生需认识到AI在材料监测中的局限性，建立“AI辅助筛查+人工深度核验”的防控意识，避免单一依赖AI结果。

明确核心责任：师生应将AI定位为初级监测工具，需对实验材料的质量与风险把控负最终责任，不可转嫁判断和审查义务。

坚持教育为本：AI应服务于培养学生实验严谨性，教师需通过案例分析、实操指导等方式帮助学生建立完整认知链。

实现协同增效：学生可以合理利用AI完成基础材料识别，将精力聚焦于材料质量的深度评估和实验原理探究等高阶环节，提升实验教学实效。

本案例及其它环境监测场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

AI

案例六

AI辅助课后评估——评估程序代码

背景 在计算机课程中，代码作业评估是检验学生编程能力和逻辑思维的关键环节。传统人工评估需教师逐行审阅代码，耗时耗力且易因疲劳出现判断偏差；而AI辅助评估通过静态分析与自动化测试完成语法检查、基础功能验证和代码规范性评估，减轻教师负担。

案例 为提高代码评估效率，张老师将学生小明的期末项目代码上传至AI评估平台，AI报告显示语法无误、通过基础测试，但存在逻辑冗余、可读性不足等问题。然而，张老师人工审查后发现，AI未能识别出代码在处理极端情况时的深层逻辑漏洞和算法效率低下的问题。更关键的是，代码风格过于工整统一，缺乏学生独立编程时常见的思考痕迹，这让他怀疑代码可能由AI生成。随后张老师与小明线下沟通，过程中发现小明无法正确解释代码核心逻辑，最终小明承认该项目代码确实是通过AI生成，而非自主编写。

风险 **逻辑与知识理解局限风险：** AI难以发现代码的深层逻辑漏洞和算法低效问题，仅能完成基础检测，导致评估结果无法反映代码真实质量。

教学效率与质量损失风险： AI无法有效辨别代码是否由AI生成，难以检验学生真实能力，削弱了评估价值。

评估信度与结果偏差风险： 若过度依赖AI评估，教师会缺失对代码深层问题的判断，无法准确掌握学生在编程逻辑、算法思维上的真实掌握程度。

学术不端风险： 学生可能利用AI生成规范代码，而AI评估难以识别此类 “代写” 行为，可能间接纵容学术不端，破坏课程的学术诚信环境。

建议 **深化人工审查：** 教师需认识到AI的局限性，树立“AI辅助+人工审查”的意识，重点针对逻辑漏洞、算法效率、代码原创性进行人工核验。

明确教师核心责任： 教师需对评估结果的准确性和学生真实能力的判断负最终责任，坚守学术诚信与教育公平的底线。

回归教育本质： 教师需调整考核策略，从单纯关注“代码是否可运行”转向考察“学生问题分析、逻辑设计、算法优化”的综合能力。

发挥协同价值： 教师可利用AI高效处理基础评估任务，将节省的精力投入到算法优化、逻辑漏洞分析等深层教学环节，实现“AI提效+人工深化”的协同效果。

本案例及其它课后评估场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

AI

案例七

AI辅助课后复习——巩固知识要点

背景 在课后复习中，学生进行自我出题、批改练习等巩固知识的活动通常耗时耗力，而AI能够根据学生输入的学习资料，快速生成多样化的自测题，提供即时批改与错因分析，并根据学生的答题表现推荐针对性的练习，从而帮助学生高效复习薄弱环节。

案例 李同学在复习《中国古代史》时，为提升备考效率，将讲义上传至AI生成自测题，AI迅速生成了试卷并给出了配套标准答案。小李见题目覆盖讲义内容、答案条理清晰，便对这份AI生成的复习材料十分满意，未再进行人工核查便直接用于记忆背诵。在期末考试中，他遇到多道与AI自测题知识点相似的题目，便直接按记忆中的AI答案作答；直到考试结束后与同学交流时，才知晓自己写的答案与正确知识点存在明显偏差。小李核对教材后发现，AI生成的自测题中，部分涉及历史事件时间、关键人物关联的信息存在错误，而自己因前期未及时核查，最终导致考试失分。

风险 **虚假生成风险：**AI因“幻觉”问题，可能捏造事实、篡改数据，生成与知识点相关的错误信息，学生若将这些错误内容作为学习依据，会逐步构建起偏离事实的错误知识体系。

学习浅表化风险：学生因过度信任AI输出结果，缺乏人工核查意识和深度思考环节，易产生“虚假掌握”的错觉，无法实现知识的真正理解与吸收。

认知过程外包风险：若学生将独立思考、事实核查等核心环节外包给AI，长期缺乏自主训练，会阻碍批判性思维、信息辨别等关键学习能力的提升。

建议 **拒绝知识外包：**学生必须自主完成知识点的理解、逻辑梳理等核心环节，通过主动思考与实践，保障知识内化的质量。

保持批判思维：学生不能盲目采信AI输出的内容，应主动查阅教科书、学术文献等权威资料进行交叉验证，通过多方比对辨别信息真伪，避免被错误内容误导。

防范认知惰性：教师应引导学生将AI视为“学习伙伴”，同时通过课堂提问、实践任务、开放性作业等多元化评估方式，要求学生主动参与知识内化过程，避免因过度依赖AI陷入认知懒惰，影响学习效果。

本案例及其它课后复习场景案例可扫描上方二维码查看详情

CASE DECODING

案例解码

案例八

AI辅助科学研究——阅读学术文献

背景 科研人员在面对海量文献时，难以快速提取核心信息并进行有效归纳；而AI文献阅读系统可生成摘要、关键词、核心观点和研究方法，并对观点分类归纳，帮助科研人员快速了解研究领域，确定研究切入点，加速科研进程。

案例 硕士研究生李同学在研究“智慧课堂”课题时，将收集到的大量文献批量上传至AI平台，让其进行提炼和总结。AI很快生成了一份详细的阅读报告，包括每篇文献的摘要、核心观点和分类结果等。然而，小李在后期深入阅读时发现：AI提取的多篇文献核心观点与原文表述存在偏差，关键实验数据被遗漏，甚至在对比两篇文献观点时，还给出了错误的结论。这让小李意识到AI在处理复杂学术内容时可能存在幻觉，无法保障学术内容的准确性和权威性，最终他不得不回归原始文献进行逐句精读和比对。

风险

虚假生成风险： AI对学术内容理解不足，在提炼信息时易遗漏关键前提，或得出错误结论，甚至凭空生成无文献依据的虚假观点，干扰学术判断。

逻辑失真风险： AI在整合多篇文献时缺乏学术逻辑，可能出现信息无序堆砌的情况，导致整合结果缺乏学术严谨性。

学习效果受损风险： 科研人员意在借助AI节省文献处理时间，提升研究效率，最终却因核对和修正AI的错误花费更多精力。

认知过程外包风险： 科研人员若过度依赖AI进行文献总结，会削弱独立阅读、归纳整合、批判分析等核心学术能力，长期下来可能导致自身科研思维与学术能力退化。

建议

强化风险应对： 科研人员应意识到AI的技术局限，在使用AI结果时，主动对照原始文献，对核心观点和关键数据进行交叉验证，从源头规避错误风险。

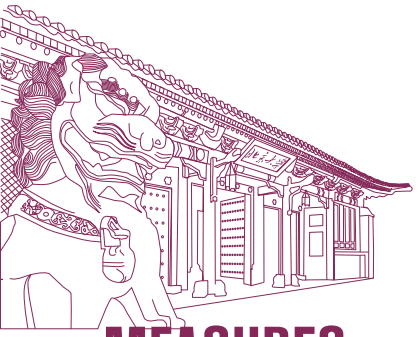
明确最终责任： 科研人员是学术结论的最终责任主体，不可将核心学术判断、结论生成等关键环节的责任转嫁于AI。

拒绝能力替代： AI可用于文献分类、基础信息提取等重复性工作，但核心观点校验、学术批判分析等任务，必须由研究者自主完成。

平衡效率与质量： 科研人员可以让AI承担基础低效的文献处理工作，将精力聚焦于学术创新、深度分析等高阶任务。

本案例及其它科学研究场景案例可扫描上方二维码查看详情

举措建议



MEASURES AND SUGGESTIONS

本指南强调“育人为本的动态适配AI应用边界”，面向教师、学生和管理者三大主体提出以下举措建议：

教师应突破传统知识传授者的定位，转型为融合人工智能的教学设计引领者与学生成长守护者。

1） 示范引导：教师须以身作则，明确标注AI贡献与合理的人机分工。通过组织对比任务、内容复核、深度讲解与过程反馈，引导学生对AI输出进行对等审视与批判性整合，最终将AI能力转化为促进师生智慧成长与专业发展的有效伙伴。

2） 育人为本：将AI深度融入教学，坚守“立德树人”根本使命。通过学科差异化评估、设置无AI环节、加强主观作业评阅等方式，重点考查学生的逻辑推演、创新思维与独立解决问题的能力，确保学生扎实掌握专业知识与能力，进而高质量地达成培养目标。

学生须从被动接收知识转向具备边界责任意识、自主学习能力与理性批判精神的积极建构者。

1） 自主立界：在使用AI过程中，应始终以提升学习效果为导向，主动划定其合理使用边界。在激发灵感、辅助学习的同时，注重保持自身在知识建构中的主体性，避免因便利性而弱化深度思考与理解过程，确保AI的使用真正服务于学习目标与认知发展。

2） 素养建构：主动了解AI的底层原理、辨识其固有局限及潜在风险，在不同任务中建立针对性验证习惯。积极寻求教师指导，定期开展无AI专项训练以巩固核心能力，并通过记录使用过程、复盘修正点，持续优化应用策略，实现从被动接受到理性驾驭的素养提升。

管理者要从自上而下的管控者转变为环境营造者、资源支持者与风险化解者。

1） 开放兼容：在恪守学术诚信的前提下，支持师生立足学科特色探索AI应用创新。通过构建分层分类的规范体系、建立试点评估与推广机制、支持AI深度融入教学科研环节等方式，共同培育师生共建、可持续的AI应用健康生态。

2） 保障使用：推动建设安全合规的AI平台，促进资源公平可及；系统规划AI素养课程体系，深化师生对AI的理性认知与风险防范意识；健全治理机制，对关键AI应用实行前置审查与持续监督。

本举措融合了核心理念与具体实践，为高等教育机构中的各方主体提供可操作性的建议，共同推动“人机共创高等教育新生态”的健康发展。

课题组研制的《高等教育人工智能应用边界指南2.0》，以“人机共创高等教育新生态”为核心理念，系统构建了涵盖原则、风险、案例与举措建议的完整框架，为高校制度建设与教学实践提供参考。我们诚挚欢迎各级教育管理部门、行业专家和广大师生提出宝贵意见，并通过邮件分享高等教育领域AI应用的典型案例。课题组将认真吸纳各方反馈，持续深化研究，完善指南内容，共同推动高等教育AI应用的规范发展与创新实践。

指南研制课题组
联系邮箱：gdsjglyyyhy@pku.edu.cn

数智教育发展国际大学联盟（DI-IDEA）
联系邮箱：diidea@pku.edu.cn

请扫描上方二维码
查看指南全文