https://doi.org/10.1038/s41562-025-02124-6

The global landscape of academic guidelines for generative AI and LLMs

Junfeng Jiao, Saleh Afroogh, Kevin Chen, David Atkinson & Amit Dhurandhar

The integration of generative artificial intelligence (AI) and large language models (LLMs) in academia brings benefits for access and collaboration as well as challenges that include misinformation and threats to academic integrity. We examine 80 academic guidelines and recommend balanced approaches for the responsible integration of generative AI and LLMs in education.

The integration of LLMs and generative AI technologies in academia offers automated essay grading, content creation and language translation, but also raises important pedagogical, ethical and legal concerns. Universities worldwide have responded in diverse ways; some Australian universities, for instance, have reverted to pen-and-paper examinations to maintain academic integrity, and others have revised plagiarism policies. A survey by BestColleges found that 31% of participants faced restrictions in AI tool use from instructors or school policies[1].

Yet, less than 10% of schools and universities have formal policies or guidance for generative AI[2]. This gap underscores the need for comprehensive frameworks that balance innovation with ethical considerations.

Global and cross-national directives

Global and national discourse on the use of LLMs and generative AI in academia reflects a spectrum of perspectives, which we have collated in a non-peer-reviewed dataset[3]. Some directives highlight the positive effects and opportunities of these technologies, whereas others emphasize caution regarding fairness, privacy and equitable access.

Recent studies on the use of LLMs, such as ChatGPT, in education reveals mixed sentiments. One study that analysed Twitter data found that 40% of relevant tweets discussed the opportunities, limitations and consequences of ChatGPT: sentiment was divided into 40% positive, 30% negative and 30% neutral. The primary concern was the potential for cheating, and the most positive aspect was the reduction in educational costs[4]. Another study, which interviewed college students, faculty members and experts, echoed these mixed views, and highlighted the ease of use and potential benefits of the technology alongside concerns about overreliance and threats to academic integrity[5].

Mixed reactions to generative AI in academia. In academia, one common current approach is to acknowledge potential issues with generative AI while also providing guidance to maximize its benefits and minimize harms. Some directives focus on practical uses such as effective prompt writing, deleting chat history and testing systems before classroom use, without explicitly acknowledging potential drawbacks[6]. Other directives expressly note positive uses of generative AI, such as personalized learning, adaptive tutoring, assisting educators with administrative tasks, increasing access to education, aiding professional development, and providing individualized feedback[7,8].

On the basis of our exploration, no country has a national directive that bans the use of generative AI in academia except for those (such as North Korea) that generally lack access or ban it. However, many directives urge educators to consider potential harms, such as bias, privacy violations and unequal access. Some less frequently cited concerns include reduced social interaction, overdependence on technology, and threats to educators' autonomy and preparedness[7,8]. These issues highlight the need for careful considerations of the effect of generative AI on education that balance its benefits with potential risks.

Academic guidelines for generative AI

Universities have a crucial role in integrating generative AI and LLMs into academic settings by developing guidelines that balance their transformative potential with risk mitigation[3].

Our analysis of generative AI and LLM use draws on data from 80 universities, which we selected to represent a diverse range of institutions globally and which includes top-ranked universities known for academic excellence from six continents[9]. We included various types of universities — such as those focused on humanities and on technology, as well as both public and private institutions — to ensure a broad spectrum of perspectives.

Nine codes of generative AI and LLM regulation. We consider nine key themes of AI regulation and ethical practice in higher education institutions (Table 1). Universities emphasize the critical importance of responsibility and safety in integrating generative AI and LLMs into academic settings. They prioritize maintaining academic integrity, transparency and privacy, and aim to ensure responsible AI use to prevent plagiarism and safeguard confidential data. Addressing the ethical complexities posed by AI involves ensuring fairness, privacy and equal accessibility. Institutions work to mitigate biases and misinformation through critical assessment of AI-generated content and diversified training data, and by maintaining human oversight[3].

Balancing innovation with academic integrity is another key focus: institutions have implemented strategies to restrict unauthorized AI use in generating academic work to promote originality and encourage critical thinking. They have also explored alternative assessment methods and restrictions on AI use in ethically sensitive research topics, to safeguard academic integrity. Universities emphasize clear expectations and responsible AI use to foster academic honesty. They also address the risks of AI-generated misinformation

Table 1 | Major and minor codes included in academic general guidelines

Themes of generative AI and LLM regulation (major codes)	Keywords and key concepts (minor codes)
Responsibility and safety	Nurturing critical thinking, transparency, confidential information protection, privacy features, security features, independent critical thinking, AI tool usage disclosure, unauthorized use condemnation, plagiarism prevention, cheating prevention
Ethical complexities and human-centric use	Ethical complexities, human-centric usage, fairness concerns, privacy concerns, accessibility concerns, societal biases, discriminatory outputs, misinformation, critical evaluation, diversifying training data, human oversight, human-in-the-loop approach, unequal access, digital divide
Balancing innovation and integrity	Ethical AI and LLM integration, limitations, restrictions, responsible integration, academic integrity, proper disclosure, critical evaluation, reflective usage, alternative assessment methods, oral examinations, project-based tasks, research restrictions, risk management
Truth assurance and misinformation risks	Generative AI and LLM outputs, reviewing and verifying content, accuracy and reliability, knowledge formation, plausible-sounding outputs, factual accuracy, fact-checking, scepticism fostering, transparent AI models, accountable AI models, reasoning transparency, bias detection, bias mitigation, diversifying training data
Pedagogical innovation, and generative AI and LLM literacy	Pedagogical innovation, generative AI and LLM literacy, educational settings, learning experiences, personalized instruction, knowledge creation, exploration tools, critical thinking skills, brainstorming, creative content generation, feedback provision, engaging learning experiences, independent learning skills, discerning accurate information, human oversight
Collaborative creativity and codesigning	Coauthoring processes, content creation, editor role, research assistant role, content refinement, teaching assistant role, self-study assistant role, personalized learning paths, coding assistance, summarizing, content development, skill democratization, scalable tutoring systems, intelligent tutoring systems, continuous learning
Empowering educators, and staff and faculty members	Evaluation methodology, grading practices, fair and efficient assessment, educational management, administrative processes, predictive analytics, syllabus creation, research proposal creation, teaching tool, intelligent tutoring systems, virtual assistants, AI-enabled educational technologies, equitable education, efficient education
Empowering students to study and research	Responsible use, academic honesty, research methodologies, critical thinking, human oversight, AI-assisted research, brainstorming, drafting, revising work, deep engagement, content understanding, personal growth, feedback loops, self-reflection, curiosity, AI-integrated education, digital learning landscape
Tailored guidance and specific guidelines	Tailored guidance, safety guidelines, educational guidelines, ethical guidelines, discipline-specific guidelines, student guidelines, educator and faculty guidelines, staff guidelines, digital learning environment, personalized learning, teaching methodologies, industry-specific skills, harmonious balance, stakeholder guidance

by recommending critical review and verification against reliable sources. To empower pedagogical innovation, institutions recognize the potential of generative AI and LLMs to enhance learning opportunities and personalize teaching, and emphasize the need for AI literacy for students and educators. Comprehensive guidelines for staff, faculty members and students promote responsible and innovative AI use to ensure equitable access and continuous learning across academic environments[3].

Quantitative and qualitative patterns

We conducted a text-mining-based analysis to analyse academic guidelines. This analysis quantified the prevalence of specific keywords within the text corpus, which we classified into the nine themes listed in Table 1.

We performed a literature review for these themes, and then used a qualitative approach to identify the five most crucial concepts for each of the nine themes. We used two criteria for this: the frequency with which they are mentioned in guidelines (which indicates consensus on importance) and their unique presence in discourse (which indicates individual significance). Examples of frequently mentioned concepts include 'integrity' and 'fairness'; less frequently mentioned but important concepts include 'skepticism' and 'democratization'. Figure 1 illustrates a frequency analysis of these qualitative findings.

Our semantic analysis of concepts and keywords in the guidelines reveals several crucial insights. First, 'privacy' is prominent (appearing in 177 instances), whereas 'disclosure' (which is crucial for responsible AI use) ranks low, with 14 mentions. However, emphasizing disclosure ensures that users are informed about AI capabilities and limitations, which promotes responsible decision-making and accountability.

Similarly, 'ethical' is frequently mentioned (216 instances), whereas 'human-centric' appears only twice; this highlights the need to incorporate user feedback, design accessible interfaces and consider ethical concerns such as privacy and consent. Guidelines should also include measures for data fairness, algorithmic transparency and fairness assessments to promote equitable AI systems.

Further, 'integrity' is common (211 mentions), whereas 'alternative' methods appear in only 47 instances. Exploring alternative methods for teaching, assessment and research can balance innovation and integrity, such as personalized learning pathways and AI-powered grading systems.

We also see a low frequency of 'misinformation' (9 instances) and 'skepticism' (2 instances), yet these are vital for truth assurance and risk mitigation. We believe that addressing the lack of attention to these two concepts is crucial, along with strategies for detecting misinformation and fostering critical thinking skills.

The term 'assistance' appears in 41 instances, whereas 'democratization' appears only once, which indicates a need to pay more attention to personalized innovation and democratization in generative AI and LLM usage. This involves customizing AI solutions to individual preferences and ensuring broad access to these technologies through educational materials and open-source projects.

Finally, although 'student' guidelines are frequently discussed, there is less focus on guidelines for teachers and educational management, and discipline-based guidelines appear the least, which suggests

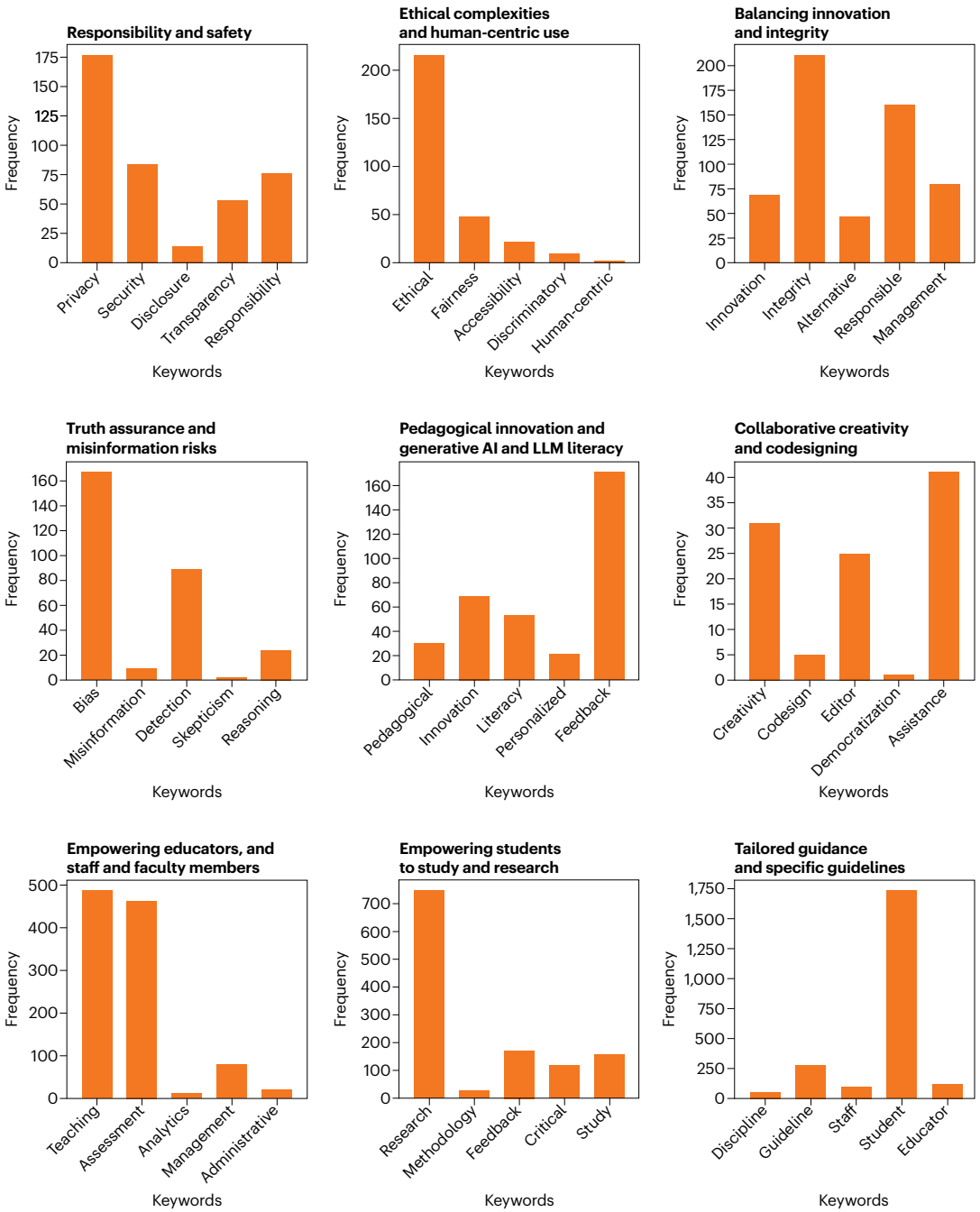

Fig. 1 | Key concepts. The frequency of key concepts in the nine major themes.

a substantial gap that needs addressing. Customizing guidelines for different fields will ensure responsible AI use tailored to each field's distinct needs and ethical considerations.

Synthesis and roadmap

These findings can help researchers, policymakers and institutions to identify crucial areas for development in future guidelines to effectively use generative AI and LLMs. We believe it will be possible to balance tradition with technology to ensure the ethical and effective integration of AI in academia. To achieve this, we call for a rethinking of pedagogy, decision-making processes and policies.

We must rethink pedagogy. Educators face a paradoxical choice between traditional teaching methods and integrating AI tools such as LLMs (Table 2). The debate centres on whether prohibiting LLMs fosters critical thinking and self-reliance or whether embracing them better prepares students for a future in which AI is ubiquitous in the workplace. Prohibiting LLMs might prevent dependency and intellectual

Table 2 | Benefits and drawbacks in the education preparation paradox

	Traditional approach	Generative AI-driven
Benefits	Emphasizes deeper thinking	Introduces students to a wider variety of ideas in a shorter amount of time
	Teaches how to conduct thorough research	Enables students to conduct research more quickly
	Emphasizes the inception, curation and refining of ideas	Allows students to draft and refine more rapidly, and potentially helps them to shift focus to the more human-necessary portions of a task
	Provides students with an opportunity to wrestle with various ideas, generating an understanding of how they do or do not relate	Showcases the need to be adaptable to innovative technologies in a constantly changing society
	Students are more likely to remember information gathered by this method	Prioritize efficiency and productivity
	Lessens overreliance on AI for critical thinking and advocates a more cautious approach	Embraces risk
	Follows empirically tested methods	Explores untested but potentially promising methods
	Relies on the educator to adapt methods to students in relatively small groups	Focuses on commodified tasks such as brainstorming ideas
Drawbacks	May underprepare students for 'real life' outside of academia	May become overdependent on AI, leading to an overemphasis on simple tasks such as brainstorming, and an underemphasis more useful skills such as collaborating and critical thinking
	Not as scalable; quality of the education will vary from instructor to instructor, so any benefit from the traditional approach is contingent in large part on the educator	The importance of students' fact-checking outputs may be ignored owing to automation bias, thus undermining fact-checking as a safeguard
	Overthinking ideas and theories might contradict efficiency and practicality	Constantly incorporating new technology means there is little opportunity to examine best practices and efficacy for learning

atrophy, but allowing their use could equip students with valuable skills and make them competitive and productive, without diminishing their intellectual capabilities.

The goal is to strike a balance that leverages AI for learning without compromising deep thinking skills. We recommend that educators:

- Integrate AI literacy into curricula and teach students how to use LLMs responsibly, by focusing on skills such as prompt engineering, accuracy evaluation and understanding limitations.
- Encourage critical thinking alongside AI use by assigning tasks that require students to analyse and critique AI outputs, which will promote both critical thinking and effective AI use.

Table 3 | Human-in-the-loop versus machine-in-the-loop approaches to educational decision-making

Dimensions	Description and comparison		
	Autonomous decision	Human in the loop	Machine in the loop
What it is	Total reliance on AI to make best decisions	AI develops recommendations	AI provides arguments for both sides
Benefits	Fastest and possibly simplest approach; requires the least time and effort from humans.	Ensures a human has an opportunity to review; may also require human sign-off before a decision is final	Requires humans to contemplate options, and that humans make the decisions
Drawbacks	AI could be biased, untrustworthy, hacked or any number of other problems	Possibility for automation bias (overly trusting the AI and simply agreeing with its recommendation each time)	Possibility for false equivalency (presenting both sides as if they are equally valid even when that is not true)
Efficacy	Not clearly better than either of the alternative approaches	Not clearly more useful than machine in the loop in most instances	Not clearly more useful than human in the loop in most instances

- Deploy a hybrid approach comprising traditional and generative AI-driven learning that combines traditional research methods with AI tools to create a balanced, hybrid learning experience.
- Evaluate student outcomes holistically by assessing both critical thinking and AI proficiency, to ensure that students excel in human and AI-driven approaches.

We must redefine decision-making in education. The concept of 'human in the loop' in educational decision-making is often seen as beneficial but lacks clarity in terms of its actual effects. Automation bias might cause humans to accept AI outputs uncritically, and make human involvement less effective in many cases. An alternative – 'evaluative AI' – involves AI providing evidence for and against decisions, and enables humans to maintain agency and accountability while benefiting from AI insights. The effectiveness of both approaches depends considerably on the timing of the AI intervention and how each method is implemented[10]. The key point to take away is that implementation details are crucial for success – possibly even more so than the choice of approach itself (Table 3). We recommend that department-level administrators:

- Train educators on evaluative AI by providing training on using it to analyse evidence for and against decisions, which will empower educators to retain agency.
- Implement hybrid decision models that combine human-in-the-loop and machine-in-the-loop approaches to balance AI insights with human judgment.
- Mitigate automation bias by educating staff and students to critically evaluate AI outputs rather than blindly accepting recommendations.
- Promote agency and accountability by encouraging both educators and students to take responsibility for AI-informed decisions, to ensure ethical and effective outcomes.

We must rethink policies. Institutions face challenges in encouraging deep student engagement with content in the era of AI tools such as LLMs, which can easily summarize articles or answer quiz questions. Deep discussions that probe student comprehension might be the best way to ensure engagement but are not scalable and may require more time than typical class durations allow. Hiring more educators to decrease student-to-educator ratios could help, but is often financially impractical. Although policies may differ across departments, it is worth examining whether uniform policies regarding AI use might be equally effective. Some fields, such as mathematics, may naturally limit AI use without needing distinct policies. We recommend that institutions:

- Invest in and support scalable discussions and smaller class sizes to promote engagement through small-group discussions and reduce student-to-educator ratios, while assessing the need for unified or department-specific AI policies.
- Standardize baseline AI policies to create foundational AI guidelines for all departments with flexibility for specific needs.
- Adapt AI policies by discipline to tailor AI use policies to the practical needs of each field.

Concluding remarks

The integration of generative AI and LLMs into academic education brings substantial opportunities and challenges.

We note that academic institutions generally approach AI policies with cautious optimism that recognizes potential harms (such as privacy issues, bias and fairness) while remaining open to the benefits of AI in education. Many institutions see the potential of AI to personalize learning and enhance research and brainstorming, and do not ban its use outright but instead advocate for wise application. However, some uses, such as AI-assisted grading, lack empirical support. Institutions might prudently limit AI grading to low-stakes assessments rather than important examinations to mitigate potential harms while still exploring the benefits of AI.

Overall, we recommend that academic institutions develop balanced strategies to leverage the benefits of these technologies while addressing ethical issues, ensuring fair access and enhancing educational outcomes. Emphasis should be placed on responsible innovation and ethical practices, and on prioritizing transparency, accountability and human-centric use — especially concerning privacy and fairness. Customized policies tailored to specific educational contexts and fields are necessary, which move away from a one-size-fits-all approach to using AI in teaching, assessment and research while maintaining academic integrity.

Junfeng Jiao[1], Saleh Afroogh[1] ✉, Kevin Chen[1], David Atkinson[2] & Amit Dhurandhar[3]

[1]Urban Information Lab, The University of Texas at Austin, Austin, TX, USA. [2]Allen Institute for AI (AI2), Seattle, WA, USA. [3]IBM Research, Yorktown Heights, NY, USA.

✉e-mail: saleh.afroogh@utexas.edu

Published online: 03 March 2025

References

1. Khedkar, S. Guidelines for generative AI use from universities worldwide. *Editage Insights* https://www.editage.com/insights/guidelines-for-generative-ai-use-from-universities-worldwide (2023).
2. UNESCO. AI: UNESCO mobilizes education ministers from around the world for a co-ordinated response to ChatGPT. *UNESCO* https://www.unesco.org/en/articles/ai-unesco-mobilizes-education-ministers-around-world-co-ordinated-response-chatgpt (2023).
3. Jiao, J., Afroogh, S., Chen, K., Atkinson, D. & Dhurandhar, A. AGGA: a dataset of academic guidelines for generative AIs. *Harvard Dataverse* https://doi.org/10.7910/DVN/XZZHA5 (2024).
4. Fütterer, T. et al. *Sci. Rep.* **13**, 15310 (2023).
5. Hasanein, A. M. & Sobaih, A. E. E. *Eur. J. Investig. Health Psychol. Educ.* **13**, 2599–2614 (2023).
6. African Observatory on Responsible Artificial Intelligence. Generative AI guidelines at South African universities. *africanobservatory.ai* https://www.africanobservatory.ai/social/generative-ai-guidelines-at-south-african-universities (2024).
7. European Commission. *Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators* (Publications Office of the EU, 2022).
8. Russell Group. Russell Group principles on the use of generative AI tools in education. *Russell Group* https://russellgroup.ac.uk/media/6137/rg_ai_principles-final.pdf (2023).
9. Jiao, J., Afroogh, S., Chen, K., Atkinson, D. & Dhurandhar, A. Preprint at https://doi.org/10.48550/arXiv.2406.18842 (2024).
10. Agudo, U., Liberal, K. G., Arrese, M. & Matute, H. *Cogn. Res.* **9**, 1 (2024).

Acknowledgements

This research is funded by the National Science Foundation (NSF) under grant number 2125858. The authors express their gratitude for the NSF's support, which made this study possible. Furthermore, in accordance with MLA (Modern Language Association) guidelines, we note the use of OpenAI's applications for assistance in editing and brainstorming.

Author contributions

J.J. and S.A. conceived and developed the main idea, contributed to the conceptualization and methodology, and conducted both qualitative and quantitative analyses. K.C., D.A. and A.D. contributed to the methodology and carried out qualitative and quantitative analyses.

Competing interests

The authors declare no competing interests.