Capital University

AI Guidebook:

Ethical Considerations of Generative AI

Ethical Considerations of Generative AI

This *AI Guidebook* section is written for staff, faculty, and students.

Ethics in AI involves the design and ongoing use of AI systems in ways that are transparent, fair, and accountable. Some of the major ethical challenges of AI are bias, privacy, and environmental impact. In education, ethical AI use means ensuring that AI tools enhance learning while maintaining academic integrity, privacy, and intellectual property rights.

Frameworks such as Microsoft's Six Responsible AI Principles emphasize fairness, reliability, privacy, inclusiveness, transparency, and accountability. These principles serve as a foundation for evaluating AI tools and their use in academic settings. As AI technologies evolve, regular reviews, updates to policies, and training sessions for faculty, staff, and students are critical for sustained ethical practices. For more information, check out trainings from Coursera or EdX.

Security Risks

Ensuring FERPA Compliance in AI Tools: The Family Educational Rights and Privacy Act (FERPA) protects the privacy of student education records. When using AI tools, ensuring FERPA compliance is non-negotiable.

Select tools that have strict data handling policies to ensure that student information is used solely for educational purposes and not for unauthorized commercial use. Contracts with AI vendors must clearly specify the responsibilities regarding data protection, including limitations on data sharing and data retention policies. The AI vendor must explicitly state FERPA compliance in their privacy policy, terms of service, or data processing agreement.

Ongoing audits of AI systems ensure that they remain compliant with FERPA regulations, especially as system updates and new features are implemented. When reviewing AI tools for FERPA compliance, look for (1) a description of how student data are collected, stored, used, and shared and (2) a statement prohibiting the sharing or selling of student data.

Duty to Protect Confidential and Personal Information: In higher education, we have a legal obligation and ethical responsibility to ensure that private personal information and confidential data are secure. AI systems often process large amounts of sensitive data. Minimization use of private personal information and confidential data reduces risk in the event of a data breach. Data minimization can include restrictions on the types of data collected, how the data are used, and the duration of data storage. Only *essential* private personal information and confidential data should be collected and processed by AI tools. Data breaches and unauthorized access can occur without proper encryption and access controls. Regular audits and testing help identify and address vulnerabilities in AI systems. This proactive approach minimizes the risk of external attacks and internal misuse.

Duty to Respect Intellectual Property Rights: Proper use of AI respects intellectual property rights. Both the output of AI and the data used for training these systems can involve proprietary content. AI tools should be designed to avoid infringing upon copyrighted materials. Faculty, staff, and students should understand the limitations and proper use of content generated by or fed into AI systems. Establish clear guidelines regarding data ownership when AI systems are used for content creation or research. Informed consent from content creators is essential when their work might be used as training data.

As AI tools become more sophisticated, it is important to ensure that any content produced respects academic integrity by properly citing sources and acknowledging original creators. Transparency in the use of AI is essential.

Global Community Impact

Cultural Sensitivity: AI tools should be developed and deployed with sensitivity to diverse cultural contexts. This includes language support, culturally relevant content, and inclusivity in AI training data.

International Collaboration: Educational institutions are encouraged to collaborate globally to share best practices, research, and ethical guidelines related to AI. These partnerships can foster innovation and address shared challenges.

Equitable Access: Global efforts should focus on bridging the digital divide. This means ensuring that students from diverse regions have equitable access to AI tools, which requires targeted investment and infrastructure support.

Worker Protections: There are serious concerns about protections for people who are hired to flag or tag inappropriate material in AI tools. They work for low wages and may be exposed to abusive material that takes a psychological toll.[1]

Environmental Impact

The rapid development and deployment of AI technologies have environmental implications, particularly regarding energy consumption and the associated carbon footprint. As faculty, staff, and students, we have a responsibility to minimize environmental impact related to our use of AI technologies.

Generative AI requires significant computational power, often involving the use of massive data centers with high electricity use. Most of the energy used occurs during the training phase, where complex algorithms are processed over weeks or months. The energy used in training these models often comes from non-renewable sources, contributing to significant carbon emissions.

Lifecycle analyses consider energy consumption during training, model deployment, and inference, as well as the eventual disposal or repurposing of hardware used in these processes. Careful lifecycle analyses can contribute to more sustainable, ethical use of AI by identifying areas for reducing environmental impact over the course of the AI tool's life. Sustainability metrics can guide the evaluation of AI tools. Some considerations for sustainability metrics are:

Environmental Sustainability Metrics

- **Energy Consumption**
  - Total electricity used during training and inference.
  - Energy used per task (e.g., per prediction, per query).
- **Carbon Footprint**
  - CO2 emissions during development and deployment.
  - Use of renewable vs non-renewable energy sources.
- **Hardware Lifecycle**
  - Longevity and recyclability of hardware used.
  - E-waste generation and management.
- **Data Center Efficiency**
  - Power usage effectiveness (PUE) of hosting environment.
  - Cloud provider's sustainability rating.

Social Sustainability Metrics

- **Workforce Impacts**
  - Net job creation vs. displacement.
  - Employee reskilling/upskilling rates.
  - Changes in job satisfaction and workload equity.
- **Accessibility and Inclusion**
  - Is the AI tool accessible to users with disabilities?
  - Is it inclusive across languages, cultures, and/or literacy levels?
- **Bias and Fairness**
  - Algorithmic fairness audits (e.g., disparate impact analysis).
  - Reduction in biased outcomes or disparities over time.
- **User Trust and Engagement**
  - Transparency of AI decisions (explainability).
  - Stakeholder acceptance and ease of use.

Governance and Ethical Use Metrics

- **Data Privacy and Security**
  - Compliance with FERPA, HIPAA, IRBs.
  - Incidents of data breaches or misuse.
- **Accountability**
  - Clearly defined responsibility for AI decisions.
  - Human-in-the-loop implementation.
- **Policy Alignment**
  - Alignment with institutional sustainability goals.
  - Documentation of ethical review or risk assessment processes.
- **Lifecycle Cost Analysis**
  - Total Cost of Ownership (TCO) including environmental/social impact.
  - Long-term maintenance vs. short-term efficiency.

Users of AI have responsibility to monitor the energy usage of AI systems and advocate for greener, energy-efficient technologies through continued research, innovation, and policy change. When evaluating environmental impact, consider optimizing model architectures, using more efficient hardware, and transitioning to renewable energy sources to power data centers. Institutions

[1] Thank you to the faculty member who reviewed this AI Guidebook section and provided this information.

and policymakers are also beginning to incorporate environmental impact assessments as part of the standard evaluation for AI projects.

Labor Impact of AI Tools

The use of AI in education can transform labor dynamics by automating routine tasks while creating opportunities for new roles and responsibilities. However, this transformation should be managed to avoid undermining the professional roles of educators and support staff. Adopt AI in ways that support and enhance human work, rather than eliminate it. Human oversight of AI-facilitated task automation is critical for maintaining quality and ethical standards.

Assess How AI will Affect Roles & Workflows	•Which tasks will be automated? •Which jobs will be replaced, redefined, or augmented? •Who is most affected (e.g., administrative staff, lower wage workers, etc.)?
Upskill and Reskill the Workforce	•Offer training on AI literacy and tools. •Build soft skills: critical thinking, ethics, communication, adaptability. •Create pathways into jobs that AI cannot replace easily, such as jobs requiring emotional intelligence, judgment, and leadership.

Availability and Accessibility of AI Tools

Ensuring that AI tools are available and accessible to all students is critical for promoting equity in education. Accessibility extends beyond physical access to include usability by individuals with diverse learning needs and abilities. Efforts should be made to bridge the digital divide by providing the necessary hardware, software, and connectivity to students from all backgrounds. Similarly, when selecting an AI tool, consider universal design and accessibility, including features that support diverse learning styles, disabilities, and language proficiencies.

Bias Within AI Tools

Bias in AI is a significant issue, as algorithms can perpetuate social inequities by reflecting or amplifying biases present in their training data. Regular audits of AI algorithms are necessary to detect biases in decision-making processes. Diversified training sets and transparency in methodology, including data sources and design choices, helps students, faculty, and staff to critically assess the outputs and the limitations of the tools.

Tools such as IBM AI Fairness 360 and Google's What-If Tool can assess and help mitigate bias in AI tools.

Academic Integrity and Student Learning

The integration of AI in academic settings raises concerns about originality, accountability, and the overall impact on student learning. Maintaining academic integrity while leveraging AI benefits and preparing students with AI skills is essential.

Transparency: Students and faculty are encouraged to disclose when and how AI tools are used in research, writing, and other academic activities. Clear syllabus statements and assignment guidelines help prevent plagiarism and ensure that AI is used as a supplement, rather than a replacement for critical thinking.

Course/Curricular Integration: Inclusion of discussions on the ethical use of AI and its implications for academic work can clarify expectations, both fostering awareness and preparing students to critically engage with emerging technologies. Curricular revision may be necessary as the technology evolves to prepare students with the skills needed for AI integration in the workforce.

Assessment and Evaluation: New assessment strategies may be necessary to distinguish between student work and AI-assisted work. Alterations to assignments may be necessary as AI technology continues to evolve (e.g., video content creation, live debate, etc.). Verification of the originality of student work can occur through AI-detection software (e.g., Turnitin, GPTZero, Originality.ai, Scribbr, Quillbot, Grammarly, etc.).

Cognitive Impact of Outsourcing Writing, Research, and Organizing

Reliance on AI for cognitive tasks such as writing, research, and organizing information can affect students' learning and critical thinking skills. Balancing the benefits of AI with the need for cognitive development is essential.

Skill Development vs. Dependency: While AI can serve as a valuable aid in organizing ideas and generating preliminary content, overreliance may hinder the development of essential research and writing skills. Educators should stress the importance of independent critical thinking and analytical skills.

Critical Evaluation of AI Outputs: Teaching students to evaluate the reliability and accuracy of AI-generated content is crucial. This includes cross-checking facts, analyzing sources, and understanding the limitations of algorithmic processes.

Balanced Integration in Pedagogy: Curricula should be designed to integrate AI tools in ways that enhance rather than replace traditional cognitive skills. Educators might assign projects that require both AI support and independent analytical work to foster a balanced approach.

Sources

IBM. (2024). What is AI ethics? https://www.ibm.com/think/topics/ai-ethics

Microsoft. (n.d.) *The Microsoft Responsible AI Standard.* https://www.microsoft.com/en-us/ai/principles-and-approach

Northern Michigan University. (2025). *Faculty – Ethics in Academia, FERPA, and Inclusivity*. https://nmu.edu/ai-literacy-initiative/faculty-ethics-academia-and-inclusivity

OpenAI. (2025). *ChatGPT* [Large Language Model]. https://chat.openai.com/chat

Swaak, T. (2024). Adapting to AI: How to understand, prepare for, and innovate in a changing landscape. *The Chronicle of Higher Education.*

Document History

Initial Publication July 2025: Drafted as part of Capital University's AI Initiative and made available to the campus community for comment and input prior to publication.