

Rules, Risks, and Ethics of AI

This page outlines why ethical and governance questions matter in artificial intelligence, what risks and harms have been identified, and how these issues connect to academic and professional settings. It is descriptive rather than instructional. The goal is to help students understand how AI is discussed in terms of responsibility, fairness, and accountability.

Why Governance and Ethics Matter

Artificial intelligence can improve efficiency and support decision-making across many fields. At the same time, international organizations and researchers have raised concerns about how AI systems can affect people and institutions. These concerns include privacy, surveillance, bias, transparency, and the difficulty of explaining how some AI systems reach their outputs. In many settings, AI systems rely on large datasets that reflect existing social and institutional patterns, which can reproduce or intensify inequalities.

Ethical questions also extend beyond technical performance. Scholars have identified concerns about the moral and legal status of AI systems, and about how autonomous systems interact with human dignity, rights, and accountability. This is why governance frameworks exist: they are meant to establish guardrails for how AI is built, deployed, and evaluated.

Common Risks and Harms

Bias and unequal impact. Data-intensive AI systems can make recommendations, predictions, or classifications that disadvantage historically marginalized groups. This is not because machines are neutral, but because the data and infrastructures that train AI systems can reflect human and institutional bias.

High-stakes decisions and explainability. AI systems are often used in decisions that materially affect people's lives, such as admissions, lending, hiring, or assessment. Many ethical frameworks cite explainability as a basic criterion for acceptability. However, the "black box" problem means that some AI outputs are not easily understood or justified in human terms, even by those who deploy the systems.

Academic integrity and learning. In higher education, surveys have reported concerns about generative AI enabling plagiarism and reducing opportunities for learning. Some educators distinguish between dishonesty and failure to learn, noting that both matter for academic outcomes. This distinction has shaped how institutions talk about AI use and academic integrity.

Policy, Standards, and Institutional Frameworks

Global and national bodies have developed frameworks to guide AI use. UNESCO's Recommendation on the Ethics of AI establishes broad ethical principles

and governance priorities for member states. In the United States, federal agencies have developed standards and guidance that emphasize trustworthiness, security, and accountability in AI systems.

In higher education, institutional policies are uneven. Some universities have developed clear guidelines that address privacy, accessibility, and acceptable uses of AI tools. Others are still developing their approaches. Research on institutional responses suggests that formal policy often lags behind rapid changes in AI capabilities, which creates ambiguity for students and faculty.

What This Means for Students

Students may encounter AI tools in coursework, research, internships, or campus services. Institutional guidance often focuses on what kinds of data can be entered into AI systems, especially when those systems are public or third-party services. This matters for privacy and confidentiality.

Academic integrity policies also vary. Some institutions treat AI assistance as similar to receiving help from another person, which means disclosure and transparency may be expected. In other settings, certain uses may be restricted or prohibited. Because policies differ by course, department, or institution, expectations can be inconsistent.

Accessibility and equity concerns also matter. AI tools are not equally accessible to all users, and reliance on these tools can disadvantage students who have limited access to technology or who require assistive tools. Some institutional guidance emphasizes usability, accessibility, and the impact of technology on learning outcomes, while also noting that AI systems can introduce harm if they are deployed without careful review.

Featured Foundational Sources

UNESCO — Recommendation on the Ethics of Artificial Intelligence

Global ethical framework outlining principles, governance priorities, and safeguards for AI systems.

NIST — AI Standards: Federal Engagement

Overview of U.S. federal standards activity and guidance focused on trustworthy AI.

AAUP — Artificial Intelligence and Academic

Higher education-focused analysis of risks, academic integrity, and institutional responsibilities.

Harvard DCE — Ethics in AI: Why It Matters

Plain-language explanation of why ethics and governance matter in AI adoption.

Internet Encyclopedia of Philosophy — Ethics of Artificial Intelligence

Academic overview of core ethical debates and foundational concepts in AI ethics.

Foundational Sources (for page build)

- UNESCO — Recommendation on the Ethics of Artificial Intelligence
 - <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- NIST — AI Standards: Federal Engagement
 - <https://www.nist.gov/artificial-intelligence/ai-standards-federal-engagement>
- AAUP — Artificial Intelligence and Academic
 - <https://www.aaup.org/reports-publications/aaup-policies-reports/topical-reports/artificial-intelligence-and-academic>
- Harvard DCE — Ethics in AI: Why It Matters
 - <https://professional.dce.harvard.edu/blog/ethics-in-ai-why-it-matters/>
- Internet Encyclopedia of Philosophy — Ethics of Artificial Intelligence
 - <https://iep.utm.edu/ethics-of-artificial-intelligence/>