

Register for the EDUCAUSE Annual Conference today. Extra 25% off in-person registration, through July 8. →

AI Ethical Guidelines

An EDUCAUSE Working Group Paper

As artificial intelligence becomes integrated into all corners of higher education, addressing ethical concerns is crucial to responsible implementation.



Credit: Canva Magic Studio AI

Introduction

The adoption of artificial intelligence (AI) in higher education is accelerating to support a broad range of functions, from personalized learning and grading assistance to campus operations and research. While AI presents considerable benefits such as enhancing student personalized learning experiences, streamlining administrative tasks, and advancing scientific discoveries, its increasing use also raises significant ethical concerns and challenges. As AI systems grow more sophisticated, it is imperative for institutions of higher education to establish and uphold comprehensive ethical frameworks. These frameworks will be crucial for ensuring the responsible implementation of AI technologies while (1) upholding core academic values of fairness, privacy, transparency, and accountability, and (2) mitigating risks such as bias, privacy violations, and misuse. Thoughtful governance frameworks informed by diverse stakeholder voices are essential to harness AI's potential while safeguarding against harmful consequences. Drawing upon widely endorsed ethical guidelines, this report adopts a pragmatic lens to inform institutional decision-making. This report is designed to provide a structured foundation for critical discourse and actionable strategies concerning the ethical integration and responsible deployment of AI technologies in higher education.

Published in 1979, ***The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*** laid out foundational ethical principles for conducting research involving human subjects. Its core tenets of respect for persons, beneficence, and justice have been widely adopted across disciplines and sectors as the bedrock of ethical research practices. The *Belmont Report* is a pioneering and influential model for developing comprehensive ethical guidelines to govern emerging technologies and practices. For decision-makers at higher education institutions grappling with the increasing adoption of artificial intelligence, the framework outlined in the *Belmont Report* can serve as a valuable starting point. Its emphasis on protecting the autonomy of individuals, maximizing benefits while minimizing harm, and ensuring fair and equitable treatment provides a strong basis for ethical AI principles in academia. Ultimately, these principles must ensure the protection of the rights of students, faculty, and staff while also advancing the significant potential benefits that AI can contribute to education and research.

The ethical principles outlined in these guidelines address the multifaceted considerations that higher education institutions must navigate when implementing AI technologies. These principles include the following:

- **Beneficence:** Ensuring that AI is used for the good of all students and faculty.
- **Justice:** Promoting fairness in AI applications across all user groups.

- **Respect for Autonomy:** Upholding the rights of individuals to make informed decisions regarding AI interactions.
- **Transparency and Explainability:** Providing clear, understandable information about how AI systems operate.
- **Accountability and Responsibility:** Holding institutions and developers accountable for the AI systems they deploy.
- **Privacy and Data Protection:** Safeguarding personal information against unauthorized access and breaches.
- **Nondiscrimination and Fairness:** Preventing biases in AI algorithms that could lead to discriminatory outcomes.
- **Assessment of Risks and Benefits:** Weighing the potential impacts of AI technologies to balance benefits against risks.

These principles target specific ethical dimensions pertinent to AI use in higher education, emphasizing the importance of fair and equitable use, protection of individual privacy, transparency in decision-making processes, and a balanced assessment of risks and benefits.

Importantly, these principles are profoundly interconnected and should be considered holistically rather than in isolation. For instance, the principles of Transparency and Explainability are closely tied to Accountability and Responsibility, given that being transparent about how AI systems make decisions is crucial for holding colleges and universities accountable for their use. Similarly, the principles of Nondiscrimination and Fairness often intersect with Justice because ensuring that AI is used in an unbiased manner is essential for promoting just and equitable outcomes for all stakeholders in higher education. By recognizing the interdependence of these principles, institutions can develop comprehensive ethical frameworks that address the complex challenges posed by AI in a cohesive and integrated manner.

Implementing AI in higher education engages various stakeholders, each with unique perspectives and priorities. For instance, students might focus primarily on how AI influences their learning experiences, privacy, and potential employment opportunities. Faculty researchers are likely to focus on how AI can revolutionize their research methods and enhance their findings while also considering the ethical implications of incorporating AI into their work. And a chief academic officer may view AI through the lens of institutional strategy, resource allocation, and the impact on overall educational outcomes. Given these diverse perspectives, establishing a shared ethical framework

using these guidelines is crucial. This will facilitate a common understanding of AI's role and promote effective collaboration across the institution. By adopting these principles—either entirely or partially—or by adapting them to their individual contexts, institutional leaders affirm their dedication to the responsible and ethical use of AI for the benefit of all stakeholders. This commitment may manifest in various forms across different institutions, reflecting their distinct missions, values, and community needs. Some may integrate these principles thoroughly into their policies and practices, whereas others may use them as foundational elements for developing customized frameworks. Regardless of the method employed, this commitment is essential for maintaining trust and integrity and for ensuring the well-being and success of students, faculty, staff, and community partners amid rapid technological shifts.

Principles

Beneficence

Definition: What Is It?

Beneficence, defined as "the act of doing good," was highlighted alongside Respect for Persons and Justice in the 1979 *Belmont Report* as a cornerstone of the ethical framework for research in colleges and universities in the United States. It ensures that "persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by making efforts to secure their well-being." In the context of AI in higher education settings, Beneficence entails the ethical responsibility of institutions to develop and implement AI technologies that actively promote the well-being of students, faculty, and the academic community. This principle requires educators to integrate AI systems that enhance learning, support research, improve administrative functions, and encourage community engagement while carefully assessing and mitigating risks such as privacy concerns, biases, and falsehoods. It also prioritizes AI applications that align with institutional values over those driven by commercial interests.

Relevance to AI in Higher Education: Why Is It Important?

Beneficence is particularly important given AI's transformative potential within academia and society. It guides all aspects of AI integration—from shaping research priorities and developing curriculum to supporting the student experience. In research, this principle encourages focusing on projects that prioritize societal benefits and rigorously evaluate potential risks. In teaching, it advocates educating students not only in AI techniques but also in the ethical implications of these technologies, fostering responsibility among

future AI developers and users. For institutional leaders, Beneficence calls for thoughtful consideration and implementation of AI tools that enhance learning, institutional experience, and operational efficiency without compromising the well-being of any community member. Additionally, it requires institutions to use AI ethically in community partnerships to address broader societal challenges, ensuring that technological advancements serve the mission of education to improve the human condition.

Application in Key Areas: How Should It Be Applied?

Applying Beneficence to AI in higher education requires a comprehensive and proactive approach:

- **Oversight Committee Formation:** Institutions should form committees to conduct thorough risk-benefit analyses and develop clear ethical guidelines that require adherence to the principle of doing good.
- **Training on AI Applications:** Offering training on AI applications before implementation maximizes benefits for students, educators, and experts.
- **Enhancement of Student Experiences:** AI systems should support the institution's educational mission and improve student support services, such as personalized learning recommendations, early warning systems for at-risk students, and efficient resource allocation.
- **Routine AI System Assessments:** Evaluations should be conducted on a regular basis to ensure that AI tools meet educational needs effectively while balancing efficiency and personalization, with considerations of student privacy and autonomy.
- **Equity Assurance in AI implementation:** Oversight processes should examine whether AI tools are working equitably for all student groups to ensure compliance with the principle of Beneficence.

Scenario: AI Feedback and Assessment in a Large-Scale Writing Course

In a foundational undergraduate writing course with over 250 students per semester, a faculty member oversees instruction with the support of 10 teaching assistants (TAs). In the past, students have raised concerns about inconsistent grading practices across multiple sections led by different TAs, while TAs have reported that the time it takes to grade the numerous essays in the course is prohibitive. The institution just licensed a tool that purportedly provides real-time feedback to students as they write their essays, potentially

resulting in final submissions that are stronger. Although most students enjoy the feedback they receive and find it generally useful to help them improve their writing, some students feel their unique voice and writing styles are unduly penalized. The TAs regain significant time, and they feel that the tool provides more consistent practices across all sections of the course. The instructor (1) is concerned that this tool is causing overreliance on AI by both students and TAs and (2) questions if students are truly mastering writing skills, as opposed to simply learning how to get favorable feedback/prompts from the tool.

Considerations:

- What risks arise from using AI-assisted grading tools in terms of reducing instructor–student engagement?
- How does the institution ensure that AI grading complements rather than replaces human evaluation of student work?
- Can the AI tool take into account diverse writing styles and cultural expressions, and evaluate them fairly?
- What impact does utilizing the tool have on teaching assistants, particularly when it provides feedback that may or may not align with individual faculty perspectives?
- Does this tool promote "normative/monoculture" language?
- What are the implications for student privacy and intellectual property rights?
- How much are students able to customize or influence the AI responses?
- What assumptions about audience and author are built into the tool?

Other ethical principles that apply to this scenario:

- Beneficence
- Transparency

Justice

Definition: What Is It?

Justice is the fair distribution of benefits and burdens among research subjects. According to the *Belmont Report*, "injustice arises from social, racial, sexual, and cultural biases institutionalized in society." In the context of AI, Justice refers to the integration of all voices into the design of AI systems. It emphasizes the inclusion of marginalized voices, images, and stories that have traditionally been omitted, which has led to disproportionate information from the majority.

Relevance to AI in Higher Education: Why Is It Important?

Justice is vital in AI to ensure that all voices have equal input in the conversation. This principle is crucial in higher education because AI systems can perpetuate existing social inequalities if not carefully designed and implemented. By integrating Justice, institutions promote diversity, equity, and inclusion, ensuring that AI technologies do not reinforce biases but instead contribute to a more equitable academic environment.

Application in Key Areas: How Should It Be Applied?

Applying Justice to AI in higher education requires the following:

- **Ethical Sourcing of Information:** AI design and implementation must ethically source information from various cultures, ethnicities, genders, and underrepresented groups.
- **Vetting AI Sourcing:** Leaders and stakeholders must vet AI sourcing to assess how designers address equality in AI training.
- **Ensure Equal Access:** Leaders and educators must ensure that students from all socioeconomic backgrounds have equal access to AI-driven resources, preventing disparities based on wealth, location, or access to technology.
- **Diverse Representation:** Diverse perspectives must be incorporated in AI datasets to avoid bias and promote fairness.
- **Policy Development:** Policies must be established that mandate inclusivity and fairness in AI applications across the institution.

Scenario: Institutional Challenges in AI Access

The institution cannot provide funding for all students to have access to the premium version of ChatGPT. As a result, an undergraduate student on a limited budget expresses concern about being at a disadvantage when completing an assignment that requires—or strongly encourages—the use of generative AI. Without premium access, the student feels they are unable to engage with the tool as fully as peers who can afford the subscription.

Considerations:

- How can institutions ensure equitable access to AI tools for all students?
- How might unconscious bias on the part of faculty affect assessments involving AI use?
- What are the academic implications of unequal access to generative AI tools across the student body?

- Does access to premium AI tools offer a significant advantage that impacts student performance and outcomes?
- How can instructors fairly assess work when students may be using different versions or capabilities of AI tools? Should faculty be informed about which students are using premium versus free AI models? If so, how?
- What accommodations or shared-access solutions (e.g., campus labs, library subscriptions) can institutions provide for students without premium access?
- How should institutions respond to GenAI companies seeking a role in education while offering tools at unaffordable prices?
- Why are academic-use AI models not currently supported or subsidized more broadly?
- How can students be directed to effective and accessible free AI tools?
- What would a fair and sustainable roadmap for educational pricing of GenAI tools look like?

Other ethical principles that apply to this scenario:

- Nondiscrimination and Fairness

Respect for Autonomy

Definition: What Is It?

Respect for Autonomy is a fundamental ethical principle that recognizes and upholds an individual's right to make informed decisions about their life, actions, and participation in various situations and processes that impact them. In the context of AI in higher education, it means acknowledging and preserving the ability of students, faculty, administrators, and staff to make independent choices regarding their engagement with AI systems without undue influence or coercion.

Relevance to AI in Higher Education: Why Is It Important?

Respect for Autonomy is crucial because it safeguards all stakeholders' academic freedom, personal agency, and privacy rights. It ensures that AI systems are understood to enhance rather than diminish human decision-making capabilities. As AI becomes more prevalent, it can influence learning paths, research directions, and administrative processes. By respecting autonomy, institutions maintain trust, promote ethical AI use, and prevent the erosion of individual rights within learning environments designed to support academic integrity.

Application in Key Areas: How Should It Be Applied?

Applying Respect for Autonomy in the use of AI in higher education requires the following:

- **Consent Mechanisms:** Institutions provide clear consent mechanisms and opt-out options for AI-driven educational tools.
- **Decision-Making Support:** AI systems are designed to support rather than replace human decision-making, fostering critical thinking and independent learning.
- **Preference Accommodation:** Students who resist including AI in their education have access to non-AI-influenced pathways.
- **Faculty Autonomy:** Faculty retain autonomy in choosing whether and how to incorporate AI tools in teaching and research and are involved in related decision-making processes.
- **Inclusive Discussions:** Administrative personnel are included in discussions about AI implementation, recognizing their expertise and preferences.
- **Data Privacy Transparency:** Institutions are transparent about data collection and use, and stakeholders have options to limit how their personal data is used.
- **Educational Resources:** Resources are made available to help stakeholders understand AI-related data collection and protection mechanisms.

Scenario: Ethical Considerations in AI-Assisted Grading

Some faculty members are beginning to use generative AI tools to assist with grading student assignments, citing benefits such as improved efficiency, consistency, and turnaround time. However, this raises concerns about student intellectual property, especially if AI platforms store, reuse, or train on submitted work.

Students may wish to opt out of AI-assisted grading due to privacy or fairness concerns. This introduces challenges for instructors seeking to maintain consistent grading standards across a course. As questions arise about who has the right to decide—faculty, students, or the institution—it becomes clear that ethical principles such as transparency, explainability, data protection, and fairness are at stake.

Considerations:

- How is student intellectual property protected when assignments are processed by AI tools?

- Are students clearly informed if AI will be used in the grading of their work?
- Should students have the right to opt out of AI-assisted grading, and under what conditions or frameworks?
- If students opt out, how can consistency and fairness in grading be preserved across the class?
- Who ultimately has decision-making authority over the use of AI in grading—faculty, students, or the institution?
- What standards are in place to ensure transparency and explainability in how AI-generated evaluations are used to inform grades?
- How are AI grading systems monitored for bias or discriminatory patterns across diverse student populations?

Other ethical principles that apply to this scenario:

- Transparency and Explainability
- Privacy and Data Protection
- Nondiscrimination and Fairness

Transparency and Explainability

Definition: What Is It?

Transparency and Explainability requires a commitment to clarity and openness about how AI systems are implemented and used within the institution. This commitment to openness includes documenting the system's data sources, processes regarding decision-making and ethical assessments, and potential impacts on students, faculty, and institutional policies. Transparency extends beyond surface-level disclosures to cover the ethical foundations and objectives driving AI use, which allows the campus community to engage critically with these tools. Without transparency, decision-makers lack the context needed to evaluate application of AI across the variety of emerging situations presented by rapid technological changes.

Relevance to AI in Higher Education: Why Is It Important?

These principles are crucial because they foster informed engagement and trust among stakeholders. Understanding how AI operates ensures alignment with institutional values, ethical standards, and privacy protections. Without transparency, there is a risk of significant unexpected consequences, especially with open-source AI systems that may

not be as "open" as they are perceived to be. Transparency allows institutions to hold AI vendors accountable and prevent inadvertent harm.

Application in Key Areas: How Should It Be Applied?

Applying Transparency and Explainability in the use of AI in higher education requires the following:

- **Vendor Collaboration:** Institutions collaborate with AI vendors to obtain detailed documentation and data-handling disclosures, including explanations of algorithms, data sources, and inherent biases.
- **Assessment of Alignment:** Both proprietary and open-source AI systems are evaluated for alignment with institutional values, ethical standards, and privacy protections.
- **Development of Resources:** Guidelines and workshops are developed to help educators and students interpret AI outputs and assess their benefits and limitations.
- **Promotion of Understanding:** Tools and resources are provided to communicate AI's advantages, challenges, and privacy implications.
- **Stakeholder Engagement:** The campus community is actively involved in discussions about AI implementation to foster shared understanding.
- **Ethical Foundations:** AI use is guided by clear ethical objectives, and these objectives are communicated transparently.

Scenario: AI Advising and Planning System

Facing substantial budget cuts and staff reductions in the Registrar's office, the institution implements an AI-powered course-planning system to streamline advising and improve scheduling efficiency. The system analyzes historical course data—including grades awarded, declared majors, course load patterns, and section availability—to generate individualized course recommendations for students. Each recommendation includes a "probability of B or higher" metric to help students make informed enrollment choices based on predicted academic outcomes. Students receive these personalized course paths upon logging into the registration system, which provides a new level of autonomy and decision-making. There are some concerns about what data is being leveraged by the recommender, and there are some questions about the usefulness of the recommendations. Some students report that the recommendations do not align with their academic interests or long-term career goals, instead reflecting what they perceive as emphasis on institutional priorities. Additionally, advisors across the institution have

observed informally that the system appears to perform better for students in certain majors, raising concerns that its recommendations may prioritize efficiency over individualized academic exploration and growth.

Considerations:

- How are AI-generated course recommendations explained to students and faculty in ways that are understandable and nontechnical?
- How is trust established and maintained across stakeholder roles—including students, advisors, and administrators?
- How is the purpose of the system framed? Is it designed to support meaningful academic journeys, maximize GPA outcomes, or optimize institutional efficiency?
- Is there transparency about how the AI system prioritizes different factors/inputs when making recommendations?
- How is human interpretation ensured in the registration process?
- How does the institution evaluate and ensure that the recommendation system is free from bias and avoids unequal outcomes for different student groups?

Other ethical principles that apply to this scenario:

- Respect for Autonomy
- Nondiscrimination and Fairness
- Assessment of Risks and Benefit

Accountability and Responsibility

Definition: What Is It?

Accountability and Responsibility emphasize transparent decision-making processes, clear ownership of actions taken by AI systems, and a commitment to addressing unintended consequences. Recognizing that AI provides acceleration but not direction, institutions must understand that AI-produced results are not neutral. Humans making decisions based on AI outputs remain responsible for any harms created.

Relevance to AI in Higher Education: Why Is It Important?

These principles are critical because deploying AI without safeguards can lead to harms ranging from data breaches to social impacts on students, staff, and faculty. For example, an AI-powered grading system might perpetuate biases, affecting academic

records and opportunities. Ensuring accountability prevents unintended consequences—such as stifling creativity or reinforcing biases—and upholds ethical standards.

Application in Key Areas: How Should It Be Applied?

Applying Accountability and Responsibility in the use of AI in higher education requires the following:

- **Assignment of Responsibility:** A designated human is made responsible for the AI system's output, ensuring a human remains in the loop.
- **Critical Evaluation:** Ensure that leadership asks critical questions—about such topics as vendor transparency, bias mitigation, data ownership, and supervision responsibilities—when evaluating AI products.
- **Avoidance of Black Boxes:** Institutions remain cautious with cloud-based solutions that are opaque, recognizing that responsibility without control can lead to unaddressed harms.
- **Policy Integration:** AI guidelines are incorporated into student honor codes and syllabus agreements, with expectations and consequences clearly communicated.
- **Ongoing Supervision:** Processes are established for continuous monitoring and intervention when AI systems fail to achieve desired outcomes.
- **Stakeholder Inclusion:** Relevant people or groups are included in evaluating potential harms prior to AI adoption.

Scenario: AI, Academic Integrity, and Student Accountability

A professor assigns a project in a course governed by the school's honor code, requiring students to submit work that is truly their own. Some students use generative AI tools by entering prompts and receiving detailed responses, while others rely on software with AI features embedded more subtly—offering suggestions, rephrasing, or auto-completion—without realizing they are using AI. As AI tools become more seamlessly integrated and harder to detect, questions of authorship and responsibility grow more complex. While institutions expect students to be accountable for their submissions, students may not always understand the extent of AI involvement, especially when features are embedded in everyday productivity tools and software they don't associate with "using AI." This raises several key questions: How should AI use be defined and detected in coursework? What's the difference between prompted and embedded AI? And is it fair to hold students responsible in all cases, especially when autonomy, fairness, and risk are at stake?

Considerations:

- How does anyone know how AI is being used?
- What are the ethical and practical differences between embedded AI features and intentionally and actively prompted AI?
- To what extent can students be held personally responsible for work influenced by AI tools?
- How can course design clarify acceptable and unacceptable uses of AI in assignments?
- Should academic integrity policies distinguish between types of AI use, and if so, how?
- What guidance or training should institutions provide to help students navigate AI-integrated tools responsibly?

Other ethical principles that apply to this scenario:

- Respect for Autonomy
- Justice
- Assessment of Risk and Benefits

Privacy and Data Protection

Definition: What Is It?

Privacy and Data Protection emphasize safeguarding personal information against unauthorized access, breaches, or exploitation. This includes student data, staff and faculty records, and sensitive institutional information. The goal is to ensure that AI systems handle data with integrity, transparency, and accountability, respecting individuals' rights to privacy and confidentiality.

Relevance to AI in Higher Education: Why Is It Important?

Compromised privacy and data protection can lead to these significant harms:

- **Personal Exposure:** Leaks of students' personal documents or prompts can expose academic struggles, financial situations, or health information.
- **Research Integrity:** Breaches of faculty research data can compromise projects and damage reputations.
- **Economic Risks:** Leakage of valuable information and intellectual property can have severe consequences for institutions and individuals.

Ensuring privacy and data protection maintains trust, complies with legal obligations, and protects against such risks.

Application in Key Areas: How Should It Be Applied?

Applying Privacy and Data Protection to the use of AI in higher education requires the following:

- **Evaluation of Vendor Practices:** Institutions take into account vendors' measures for secure data transmission and storage, transparency in data handling, and guidelines on personal information use.
- **Opt-Out Options:** Opt-outs are made available at both individual and organizational levels.
- **Data Use for Training:** Institutions determine whether user data is used for ongoing model training, which could lead to the risk of de-anonymized data appearing in future responses.
- **Legal Compliance:** Vendors are verified for compliance with all applicable privacy and data protection laws.
- **Risk Minimization:** AI systems that rely on on-campus or edge computing resources are preferred to reduce the risk of exposing sensitive data to cloud providers.
- **Institutional Policies:** Institutions develop and enforce strict data privacy policies, including the use of encryption technologies and transparent data practices.
- **Continuous Monitoring:** AI systems are regularly assessed for compliance with privacy standards, and any identified vulnerabilities are addressed promptly.

Scenario: AI-Personalized Student Support Systems

An Advising Center is preparing to launch an AI-powered system designed to enhance student support. The system collects and analyzes a combination of student data—including academic history, course performance, advising center visits, self-reported preferences, and motivational indicators—to generate personalized outreach recommendations for advisors. Students are excited about more personal and actionable support; however, they have concerns about how their data might be accessed, viewed, and shared among other institutional staff. Questions have also emerged about whether third-party vendors may use student data to refine their tools or for marketing purposes, as well as whether students will have control over how their information is used.

Considerations:

- What safeguards are in place to prevent unauthorized access to student data?
- How are students informed about how their data is collected, stored, used, and potentially shared?
- Can students opt out of data collection or AI-generated advising outreach?
- Is the vendor contractually restricted from using student data for tool training, marketing, or other non-advising purposes?
- How is the data for preferences and motivation being gathered and checked for quality/validity?
- What governance is in place to ensure that the AI system's outputs align with institutional values of trust, privacy, and equity?

Other ethical principles that apply to this scenario:

- Transparency and Explainability
- Respect for Autonomy

Nondiscrimination and Fairness

Definition: What Is It?

Nondiscrimination and Fairness are ethical principles that emphasize treating all individuals equally and justly, without bias or prejudice based on characteristics such as race, gender, age, socioeconomic status, or any other protected attributes. These principles mandate that AI systems be designed and implemented to avoid perpetuating existing biases or creating new forms of discrimination. This involves ensuring that AI algorithms and datasets do not favor or disadvantage any particular group and that outcomes are equitable for all users.

Relevance to AI in Higher Education: Why Is It Important?

Nondiscrimination and Fairness are critical because AI systems are increasingly used in admissions, personalized learning, grading, and administrative processes. Biased systems can perpetuate inequalities, hinder diversity, and compromise the integrity of educational institutions. For example, an AI-powered admissions tool might inadvertently favor applicants from certain backgrounds if trained on biased data. Ensuring fairness helps institutions uphold their commitment to equal opportunity, foster an inclusive

learning environment, and maintain public trust. It also prepares students to participate in a society that values justice and equality.

Application in Key Areas: How Should It Be Applied?

Applying Nondiscrimination and Fairness to the use of AI in higher education requires the following:

- **Bias Audits:** Regularly evaluate AI algorithms and datasets for potential biases by testing systems with diverse data to identify and mitigate discriminatory outcomes.
- **Inclusive Data Practices:** Use diverse and representative datasets when training AI models to reflect a wide range of experiences and perspectives.
- **Transparent Algorithms:** Promote transparency by making AI algorithms understandable to stakeholders, and explain decision-making processes to allow for scrutiny and accountability.
- **Policy Development:** Establish clear policies that prohibit discrimination and require fairness in all AI applications, and outline procedures for addressing identified biases.
- **Stakeholder Engagement:** Involve a diverse group of students, faculty, and staff in developing and implementing AI systems to ensure that multiple perspectives are considered.
- **Education and Training:** Provide training for developers, administrators, and users on recognizing and preventing bias in AI systems.
- **Legal Compliance:** Ensure that AI applications comply with all relevant anti-discrimination laws and regulations, such as the Americans with Disabilities Act (ADA) and the Civil Rights Act.

By implementing these strategies, institutions can leverage AI technologies while upholding Nondiscrimination and Fairness, promoting an equitable academic environment for all.

Scenario: Transparency in AI-Driven Admissions

A college launches a pilot program using an AI-driven system to assist in evaluating student applications. The system generates a set of "probability of success" scores intended to support admissions officers in their decision-making. After one year of implementation, an internal review reveals troubling patterns: the tool disproportionately favors applicants with

certain background characteristics. The results indicate that the model is producing a fundamentally unfair outcome for reasons that are not understood, raising questions about model transparency, input data quality, and institutional control.

Considerations:

- Does the college have insight and control over how the AI model is trained, and can other datasets be added for consideration? Can the overall weights of different elements be adjusted to promote greater equity in the evaluation process?
- Has the college reviewed the application data to determine that it is as free from bias as possible, thereby ensuring input data is clear from systemic bias?
- What processes are in place for human oversight of the AI-driven admission recommendations?
- What mechanisms exist to continuously review, validate, and refine the model as new data sources become available?

Other ethical principles that apply to this scenario:

- Beneficence
- Accountability and Responsibility
- Justice

Assessment of Risks and Benefits

Definition: What Is It?

Assessment of Risks and Benefits refers to a systematic process in which researchers carefully weigh the potential harms that participants might face against the study's possible outcomes. This ensures that the research is ethically justifiable by exposing participants to reasonable risks related to the anticipated benefits. The assessment should thoroughly examine all aspects of the study, consider alternative methods, and communicate the risks and benefits to potential participants.

Relevance to AI in Higher Education: Why Is It Important?

Assessing the risks and benefits of AI in higher education is essential for integrating AI responsibly and equitably. This assessment is critical in deciding whether introducing an AI system will truly benefit students and educators or if the risks are too significant without adjustments or safeguards. Consistently assessing AI integration ensures students' rights, promotes fairness, and maximizes technology benefits while minimizing

potential harm. It helps institutions make informed decisions that uphold ethical standards and foster an inclusive educational environment.

Application in Key Areas: How Should It Be Applied?

Applying Assessment of Risks and Benefits to the use of AI in higher education requires the following:

- **Privacy Safeguards:** Institutions enforce strict data privacy policies, use encryption technologies, and maintain transparent data practices to protect personal information.
- **Evaluation of Benefits against Risks:** Institutions assess whether the potential benefits of AI tools, such as personalized learning, outweigh the risks associated with data collection and ensure that proper safeguards are in place.
- **Bias Audits:** AI systems are regularly audited for bias, with transparency maintained in all decision-making processes.
- **Diverse Datasets:** Diverse datasets are implemented in AI systems, and outcomes are continuously monitored to promote equitable access and fairness.
- **Impact Monitoring:** Institutions ensure that no group of students is disproportionately impacted by carefully weighing all options and considering possible harms to inclusion and equality.
- **Stakeholders:** Risks and benefits are clearly communicated to all stakeholders—including students, faculty, and staff—to uphold transparency.

Scenario: Faculty Use of AI-Generated Course Content

While attending a conference, a professor discovers a new AI-powered application that generates custom course content. Drawing from a range of sources, it produces material (written and video) for lectures, discussion sections, and assessments. In addition to creating suggestions for activities and assignments, it can also suggest activity ideas, generate assignment prompts, and design a sequence map aimed at optimizing student learning. The developers of the application also hope to integrate open educational resources (OERs) and existing MOOCs to reduce costs for students. However, the platform is not institutionally supported, raising critical questions about instructional quality, data privacy, academic transparency, and content ownership.

Considerations:

- What are the institutional policies on the use of an unsupported tool? How are student privacy and data security protected?
- What are the concerns about the quality of instruction and the breakdown of trust between faculty and students?
- What practices are in place to ensure transparency about the co-creation of material via AI?
- How are intellectual property rights determined for AI-generated teaching materials?
- How are the accuracy and credibility of the AI-generated content validated?
- What safeguards are in place to prevent AI from reinforcing the prioritization of dominant narratives?
- How do faculty and students provide feedback about and/or report issues with the AI-generated material? How responsive are the owners of the application in addressing issues?

Other ethical principles that apply to this scenario:

- Accountability and Responsibility
- Transparency and Explainability
- Privacy and Data Protection
- Nondiscrimination and Fairness

Why We Developed the Principles

The development of these ethical principles was driven by the recognition of the transformative impact that AI technologies can have on education. As AI applications become increasingly pervasive across various sectors—including higher education—the imperative for a comprehensive ethical framework to guide decision-making and ensure responsible implementation becomes evident. Education is a field in which great trust is placed in the faculty and staff members who make decisions that impact students' lives, and the details of AI adoption—both in concept and practice—can greatly influence students' futures in and out of the classroom.

The group that developed these principles engaged in detailed discussions and reviews of existing practices, and its members represented a wide spectrum of stakeholders within the academic community. This inclusive strategy was designed to capture a

diverse array of perspectives, ensuring that the framework comprehensively addressed the multifaceted nature of AI integration across different educational elements.

Several key considerations shaped the development of these principles. The potential impact of AI on privacy and data security emerged as a primary concern, given the sensitive nature of educational data. As a result, the principles emphasize robust data protection and the ethical handling of student and staff information. Furthermore, the group considered the implications of AI on educational equity. This led to guidelines that promote fairness, prevent bias, and ensure equal access to AI-enhanced educational resources. Transparency in AI applications was also identified as critical because it fosters open communication about the use and impact of AI tools within both the academic community and industry.

To underscore the need for an ethical framework for decision-making and the potential implications of adopting AI without thorough consideration, parallels may be drawn to the *Belmont Report* as a document central to higher education and research policy. The *Belmont Report's* ethical principles of respect for persons, beneficence, and justice serve as a foundational reference that points to the entangled relationship between people and the increasingly complex systems influencing their lives in both visible and invisible ways. This approach ensured that the principles uphold fundamental ethical considerations universally recognized across scholarly and practical domains.

Ultimately, the development of these principles aimed to create an environment in which AI technologies enhance educational outcomes without compromising ethical standards or the well-being of the academic community. This framework serves not only as a guide for current practices but also as a flexible, evolving tool capable of adapting to future challenges and innovations in AI technology.

Why Use These Principles? Why Do They Matter?

The ethical principles outlined in this framework provide essential guidance for harnessing the transformative potential of artificial intelligence in higher education. By adhering to these principles, institutions can enhance student experiences and learning outcomes while mitigating potential risks. For example, an AI-powered adaptive learning system developed and guided by these ethical principles can deliver personalized educational content that specifically caters to the needs of individual learners while simultaneously ensuring their privacy. Additionally, it maintains transparency about

efforts to promote fairness and mitigate potential biases. Such an approach not only enhances individual student performance but also facilitates equitable access to academic resources, thereby aligning with the overarching goals of the ethical framework. Moreover, these principles help institutions navigate complex decisions about AI implementation, aligning technological advancements with core educational values and objectives. Universities can use these principles to create a robust foundation for responsible AI adoption that benefits all academic community members.

Adopting these ethical AI principles is not just a matter of compliance but a strategic move to future-proof academic excellence and institutional effectiveness in an increasingly AI-driven world. As AI technologies evolve rapidly, these principles provide a flexible yet robust framework for navigating emerging challenges and opportunities. For instance, adherence to these principles in research settings can guide the responsible development and use of AI tools, ensuring that research integrity and ethical standards are maintained as AI capabilities grow more sophisticated. This approach not only safeguards the reputation of institutions but also positions them at the forefront of ethical AI innovation. In the realm of teaching, these principles can inform the ongoing integration of AI technologies into pedagogical practices, guaranteeing that even as AI-enhanced learning tools evolve, they remain aligned with educational goals and values.

In addition, an ethical framework for AI equips students with critical skills and knowledge essential for their career readiness in an AI-infused job market. Students trained in environments that prioritize ethical AI use are better prepared to navigate the complexities of modern workplaces, where AI ethics increasingly influence professional practices and decisions.

On the administrative front, these principles guide ethical AI practices that support institutions in adapting to changing operational needs. By emphasizing transparency, fairness, and accountability in AI implementations, universities can manage their resources more efficiently while fostering an environment of trust among faculty, staff, and students—and enhancing their capacity to attract top talent and resources in an increasingly competitive higher education landscape.

Furthermore, the integration of multiple AI platforms from different technology providers presents unique challenges for institutions. By adhering to a unified set of ethical principles, universities can ensure a cohesive approach to AI integration despite the diversity of systems and tools involved. This unified approach is crucial for maintaining a seamless operational environment that upholds ethical standards across all technological interactions.

The implementation of these principles can also a powerful tool for fostering a sense of belonging within higher education. For instance, an AI-powered tutoring system that is thoughtfully designed to accommodate the diverse learning styles, needs, and backgrounds of students can play a crucial role in ensuring that all learners receive personalized support and resources. By considering language proficiency, cultural context, and prior knowledge, such a system can promote a more equitable learning environment—one in which every student feels supported and empowered to succeed.

Integrating AI in art, media, and design offers significant benefits but also raises concerns such as bias and diminished creative authenticity. AI can overshadow human creativity, potentially leading to homogenized outputs, especially if it relies on nondiverse datasets that perpetuate stereotypes. In media, biases in AI algorithms can compromise journalistic integrity and public trust. These challenges highlight the need for a balanced approach in AI education, emphasizing the importance of diverse datasets and critical scrutiny of AI outputs. This ensures that students develop as technologically proficient and ethically aware creators.

In STEM fields, AI integration can accelerate research and innovation, enabling students to efficiently process large datasets and discover patterns that enhance areas such as personalized medicine and sustainable engineering. For example, biology students might use AI to analyze genetic sequences rapidly, whereas engineering students could apply AI in designing and testing new models and materials more efficiently. Additionally, AI in labs and virtual simulations offers real-time feedback, allowing students to learn from errors safely. However, the reliance on AI raises concerns about bias—particularly if the data or algorithms reflect existing prejudices—which could skew research outcomes and solutions. By providing students with these advanced tools, institutions must also emphasize the importance of scrutinizing AI outputs to prevent the perpetuation of biases, ensuring that future scientists and engineers are not only technologically proficient but also ethically informed.

Moreover, the application of inclusive AI practices can have implications beyond individual student experiences, enhancing broader institutional efforts to promote student and faculty success. By leveraging AI to identify and address disparities in student academic outcomes or faculty and staff retention rates, colleges and universities can demonstrate their commitment to creating a welcoming and supportive environment for all. This commitment to ethical and inclusive AI practices not only benefits individual students and faculty but strengthens the institution by fostering a shared sense of belonging and purpose.

Furthermore, the process of discussing these principles offers valuable opportunities for community-wide dialogue about the role of AI in reinforcing institutional values. These conversations can engage diverse stakeholders—including students, faculty, staff, and administrators—in a joint effort to align AI practices with the institution's mission and values. This collaborative and inclusive approach to integrating AI underscores the potential of technology to support and enhance the foundational values of the academic community, promoting a more connected and responsive educational ecosystem.

Implementing AI in higher education engages diverse stakeholders with different perspectives and priorities. For example, university administrators might leverage AI to improve operational efficiency, reduce costs, and support data-driven decision-making. However, they must also consider the potential impact of AI on student learning outcomes, faculty workload, and institutional reputation. A shared ethical framework is pivotal, promoting a unified understanding of both the opportunities and challenges posed by AI adoption and fostering collaboration among these diverse groups. Such collaboration is essential for the advancement of research and learning in higher education, given that it enables institutions to harness the full potential of AI while mitigating risks and unintended consequences.

By working together, administrators, faculty, staff, and students can create a more innovative, effective, and inclusive educational environment through the development and implementation of ethical AI practices. Why should these principles be adopted? It demonstrates an institution's commitment to using AI responsibly and ethically for the benefit of all stakeholders. This commitment is crucial for fostering a sense of belonging within the institution and building public trust and credibility. As AI becomes increasingly prevalent, institutions prioritizing ethical considerations will be well positioned to shape the future of higher education and society at large.

How to Apply the Principles

Applying the principles developed by the working group involves a deliberate and systematic integration of ethical considerations into the deployment and use of AI technologies in higher education settings. These principles are designed to serve as both a compass and a benchmark for institutions as they navigate the complexities introduced by AI, ensuring that technological advancements enhance educational outcomes without compromising ethical standards.

Applying Ethical Principles

To implement these principles effectively, institutions should consider the following strategic actions:

1. **Assessment and Alignment:** Existing AI systems and proposed projects are regularly evaluated against the ethical framework to ensure alignment with core values such as privacy, fairness, and transparency.
2. **Training and Development:** Training sessions are conducted for faculty, administrators, and IT professionals to deepen their understanding of ethical principles and their practical implications.
3. **Policy Integration:** These principles are embedded in institutional policies and governance structures to formalize the commitment and guide consistent decision-making.
4. **Continuous Monitoring:** Mechanisms are established for ongoing monitoring and evaluation of AI initiatives to ensure adherence to ethical guidelines and to address emerging issues.
5. **Tools Usage and Selection:** AI tools are selected and used based on their alignment with institutional values and ethical standards, with preference given to systems that promote transparency, accountability, and equitable access.

Stakeholder Considerations and Scenario Planning

The principles were formulated to address specific ethical challenges identified in the deployment of AI within educational settings. By prompting stakeholders to analyze the implications of AI technologies from multiple ethical perspectives the group aimed to provoke critical thinking and cultivate a heightened awareness of the potential risks and benefits.

Different stakeholders—such as students, faculty, researchers, and administrators—interact with AI technology in distinct ways. The team established four stakeholder groups:

- Students
- Faculty and Researchers
- Administrators
- Staff (Information Technology and Instructional Designers)

Utilizing scenarios addressing the needs of these groups is a key strategy in helping stakeholders understand the practical implications of AI integration. These scenarios provide a tangible context for applying the ethical guidelines in decision-making processes. For example, the deployment of an AI-driven advising system could serve as a scenario to explore strategies for respecting student data privacy while simultaneously enhancing the effectiveness of academic advising. This approach ensures that stakeholders are able to visualize and comprehend the real-world impact of AI, facilitating informed and ethical decision-making across all levels of the institution.

Institutional AI Ethical Review Board (AIERB)

To oversee the consistent application of these ethical principles, institutions should consider establishing an AI Ethical Review Board (AIERB). The AIERB is a group of key stakeholders that reviews and monitors the usage of AI by members of the institution. It is responsible for ensuring that the principles of the ethical framework are being applied in a manner that is consistent and that protects faculty, staff, students, researchers, and administrators.

Consisting of faculty from various disciplines, instructional technologists, computer security experts, student representatives, and administrators, the AIERB will offer diverse perspectives that protect student and faculty rights, administrative goals, IT infrastructure, and pedagogical innovation.

The AIERB's responsibilities will include the following:

- Critically assess the validity of AI-generated content generated by tools developed and/or used by the institution
- Consider ethical uses of AI-generated content in alignment with institutional or instructional policies regarding the use of AI in coursework, research, and dissemination
- Review strategies to protect data by anonymizing when a public or private large language model (LLM) is used for analysis
- Establish acceptable standards and benchmarks for academic and operational platform adoption and implementation

By incorporating these diverse perspectives and creating structured frameworks such as the AIERB, institutions will ensure that the principles are not only comprehensive and robust but also practical and applicable across various contexts within higher education. This proactive approach will foster responsible innovation and strengthen the ethical

culture within academic institutions, positioning them to adeptly navigate the complexities of AI in higher education in the future.

Conclusion

The use of AI in higher education is growing rapidly, and its integration within higher education poses both opportunities and challenges, necessitating a careful and thoughtful approach that is guided by well-defined ethical principles. To address these concerns, colleges and universities need clear guidelines to ensure that AI is used responsibly.

This report proposes a framework for developing such ethical guidelines. It draws inspiration from the *Belmont Report*, a set of core principles for ethical research practices. The articulation of these principles—which include Beneficence, Justice, Respect for Autonomy, Transparency and Explainability, Accountability and Responsibility, Privacy and Data Protection, Nondiscrimination and Fairness, and Assessment of Risks and Benefits—reflects a comprehensive response to the multifaceted ethical dilemmas presented by AI technologies. It is to be expected that different stakeholders in higher education, such as students and faculty, will have different priorities when it comes to AI ethics. However, a shared ethical framework is crucial for fostering common understanding and collaboration across the institution. Universities can adapt these principles to their specific needs, thereby demonstrating their commitment to ethical AI use. This commitment is essential for maintaining trust and ensuring the well-being of everyone involved in higher education. Institutions need to have commitment to a future with AI if they are going to adopt it at all, and by working with their constituencies, they can establish appropriate norms. What is appropriate for one institution may not be appropriate for another.

By encouraging open discussions and collaborative decision-making processes that involve multiple stakeholders, institutions can better navigate the complexities of AI implementation. This collective approach guarantees that AI tools are used effectively, aligning with the institutional values and safeguarding the interests of all parties involved.

Ultimately, by adopting and adhering to these ethical guidelines, higher education institutions can harness the potential of AI to enhance educational outcomes and operational efficiency while upholding a commitment to ethical standards. This strategic approach not only fosters innovation and inclusivity but also strengthens the trust and credibility of educational institutions in an increasingly AI-driven world. Through these

efforts, institutions are well positioned to lead by example in the ethical application of AI, thereby shaping the future of higher education and influencing broader societal norms concerning the responsible use of technology.

Appendix: Ethical Principles

Principle/Category	Questions and Considerations / Additional Detail	Notable Differences by Stakeholder	Framing (What does this apply to? What does this not apply to?)
Beneficence AI should be used to maximize benefits and minimize harm.	How do we measure and monitor the benefits and risks of AI in an educational setting? What safeguards are in place to prevent harm—such as algorithmic bias—from AI? What are the goals of AI at this institution? What are the institution's goals for AI adoption? Who benefits from AI, who is excluded, and who may be harmed by its implementation?	Students: AI as a tool to enhance personalized learning and support. Faculty: AI as a tool for research enhancement and reduced administrative burden. Staff: AI for improving efficiency and job satisfaction.	Teaching and Learning • Instruction • Faculty Expertise • Academic Integrity
Justice AI should be used in ways that ensure fair and equitable	How does AI impact different groups, and how	Students: Equitable access to AI resources	• Admissions and Enrollment

treatment for all individuals.	can we ensure fair treatment? Are there protocols to address disparities amplified or created by AI? What knowledge, skills, or competencies are being measured through AI, and are they aligned with course or institutional learning outcomes? Which features of AI will enable these to be assessed, and how will assessments be conducted in practice? In what ways is AI being used to enhance and demonstrate the value of formative approaches to assessment,	and educational opportunities. Faculty: Fair use of AI in evaluation and assessment. Staff: Just application of AI in performance review and management.	• Curriculum and Instruction • Student Support Services • Assessment and Grading • Institutional Research and Decision-Making
---------------------------------------	---	---	--

	studying learning processes as well as outcomes, and supporting social and emotional development and learner well-being?		
Respect for Autonomy Recognize and respect the autonomy of individuals in decisions influenced by AI.	How do AI systems accommodate individual preferences and choices? What measures are in place to protect those with diminished autonomy from potential AI biases? What risks are associated with AI-generated predictions in this context? What safeguards exist to prevent harm, and how can positive interventions be prioritized?	Students: Consent and opt-out options for AI-driven educational tools. Faculty: Choice in adopting AI tools for teaching and research. Staff: Involvement in decisions about AI implementation in administrative tasks.	Student Agency: AI should enable student decision-making autonomy. Consent and Privacy: The use of AI with student data should have an opt-out option for students. Academic Freedom: Faculty should have the option to opt out. Workplace Autonomy: Staff members should be included in decision-making processes pertaining to AI.

Transparency and Explainability AI systems should be made transparent, with decisions that are explainable to users.	How are AI decision-making processes communicated to stakeholders? Are students, faculty, and staff clearly informed about how AI is used in educational settings? If/How will student and faculty data be used to train AIs?	Faculty: Transparency in how AI tools impact teaching, grading, student evaluation, and research. Students: How AI is used in admissions, course placement, and academic support. Transparency about data collection, use, and privacy. Staff: Transparency about AI implementation in administrative processes and decision-making. Administrators: Transparency to all stakeholders about institutional AI policies and strategies. Highest-Level Leadership: High-level	Applies to: • Admissions • Grading • Advising • Instruction • Research • Administration • Strategy • Data Practices
---	--	---	--

		transparency on AI's alignment with institutional mission and values.	
Accountability and Responsibility AI should be implemented with clear accountability for decisions and outcomes, ensuring responsible oversight at every stage.	Who is responsible for the outcomes of AI decisions? What mechanisms are in place for feedback about and rectification of AI-related issues? How will implementors monitor and assess the extent to which the intended impacts and objectives are being achieved? If AI does not achieve its intended impact, what evaluation process is in place to adjust or improve its use?	Faculty: Accountable for the appropriate use and oversight of AI in teaching and research. Responsible for monitoring AI's impact on student learning outcomes and academic integrity. Students: Accountable for using AI tools ethically and responsibly for their learning. Responsible for reporting concerns or issues related to AI's impact on their educational experience. Staff: Accountable for the implementation and monitoring of AI in	Applies to: • Teaching • Research • Learning • Administration • Governance • Strategy • Values • Legal Compliance

administrative processes.
Responsible for ensuring that AI aligns with institutional policies and values.

Administrators:
Ultimately accountable for AI's impact on institutional operations and outcomes.
Responsible for establishing governance structures and policies for ethical AI use.

Highest-Level Leadership:
Accountable for the strategic direction and oversight of AI at the institutional level. Responsible for ensuring that AI aligns with the institution's mission, values, and legal obligations.

Privacy and Data Protection AI should safeguard personal and sensitive data in ways that are consistent with privacy laws, regulations, and ethical standards.	What data privacy measures are in place for AI systems handling student, faculty, and staff data? How is informed consent obtained, tracked, and managed for student, faculty, and staff data?	Faculty: Responsible for protecting student data privacy in AI-assisted teaching and assessment. Entitled to privacy protections for their own data used in AI systems for evaluation, research, and professional development. Students: The primary stakeholders whose personal and academic data is collected and used by AI systems. Consent and control over their data usage are critical privacy considerations. Staff: Responsible for handling and protecting sensitive data used by AI in administrative processes. Entitled to privacy protections for	Applies to: • Student Records • Faculty Records • Staff Records • Research Data • Administrative Data • Institutional Data • Data Governance • Legal Compliance
---	--	---	--

their own data used in AI systems for performance evaluation and management.

Administrators:
Accountable for establishing and enforcing institutional data privacy policies and practices.
Responsible for ensuring that AI systems comply with legal and ethical standards for data protection.

Highest-Level Leadership: Have oversight responsibility for ensuring that the institution's AI data practices align with its values and legal obligations.
Accountable for the reputational and legal risks associated with AI-related data breaches or privacy violations.

Nondiscrimination and Fairness AI should be used with a commitment to identifying and mitigating bias, striving to ensure nondiscrimination and fairness in all applications.	What steps are being taken to ensure that AI algorithms are free from biases? How is fairness maintained with respect to AI-driven decisions that affect stakeholders? What processes are in place to identify, mitigate, and disclose bias in AI training data and outputs, and how are those documented and communicated?	Faculty: Responsible for critically evaluating AI tools used in teaching, research, and assessment for potential biases. Students: AI should strive to minimize discrimination based on personal characteristics in admissions, grading, and support. Staff: AI used in hiring, evaluation, and management should be regularly audited for biases and disparate impacts. Administrators: Accountable for establishing policies and practices that prioritize nondiscrimination and fairness in AI adoption.	Applies to: • Admissions • Grading • Student Support • Hiring • Performance Evaluation • Promotion • Resource Allocation • Policy Development
--	--	---	---

		Highest-Level Leadership: Have oversight responsibility for ensuring that AI aligns with institutional commitments to nondiscrimination and that potential biases are addressed proactively.	
Assessment of Risks and Benefits AI should be evaluated continuously to weigh potential benefits against associated risks, ensuring that its use aligns with institutional values and ethical standards.	How are users informed about the risks, benefits, and limitations of the AI tool—including transparency around LLM training and data sources?	Faculty: Should be provided with clear information about the potential risks and benefits of AI tools in teaching and research, including limitations and uncertainties. Need transparency about the training data, algorithms, and performance metrics of LLMs used in educational applications to make informed decisions.	Applies to: • Educational Tools • Research Applications • Administrative Systems • Institutional Strategy

Students:
Deserve clear and accessible explanations of the risks and benefits of AI-powered educational tools, including potential biases, privacy implications, and limitations.

Staff: Need to be informed about the risks and benefits of AI tools used in their work, including potential impacts on job roles and decision-making.

Administrators:
Responsible for ensuring that all stakeholders are adequately informed about the risks and benefits of AI adoption, including the implications of LLM training and use.

Highest-Level Leadership: Need

		high-level understanding of the risks and benefits of AI adoption—including the implications of LLM training and use—to make informed strategic decisions.	
--	--	--	--

EDUCAUSE Working Groups provide an opportunity for community members to come together around a topic of shared experience, expertise, and passion. The members of Working Groups collaborate to envision and develop products they believe will help the higher education community understand and address the topic.

This resource was developed by the **AI Ethical Guidelines Working Group**. HP provided support for a 2023–24 AI Community of Practice, and this working group grew out of the rich discussions of that work. Statements of fact or opinion are the responsibility of the authors alone and do not imply an opinion on the part of the EDUCAUSE board of directors, staff, or members.

© 2025 EDUCAUSE. The content of this work is licensed under a **Creative Commons BY-NC-ND 4.0 International License**.



Connect with EDUCAUSE

Subscribe to our emails and hear about the latest trends and new resources.

Log in or create a profile

The views, information, or opinions expressed by speakers and authors are solely their own and do not necessarily reflect those of EDUCAUSE.

Topics

[Analytics](#)

[Artificial Intelligence](#)

[CIOs and Senior Technology Leaders](#)

[Cybersecurity](#)

[Digital Transformation](#)

[Federal Policy](#)

[Presidents and Senior Executives](#)

[Student Success](#)

[Teaching and Learning](#)

[Explore All Library Topics](#)

Insights

[Research](#)

[Core Data Service](#)

[EDUCAUSE Review](#)

Conferences & Learning

[Annual Conference](#)

[Cybersecurity and Privacy Professionals Conference](#)

[EDUCAUSE Institute](#)

[Professional Pathways](#)

[Face-to-Face](#)

[Online](#)

[Explore All Upcoming Events](#)

Community

[Community Groups](#)

[Member Directory](#)

[Mentoring](#)

[Member QuickTalks](#)

[Volunteering](#)

[EDUCAUSE Awards](#)

[Find or Post a Job](#)

Who We Are

[Discover Membership](#)

[Mission and Organization](#)

[Leadership and Staff](#)

[Member Orientation](#)

[Corporate Participation](#)

[Work at EDUCAUSE](#)

[Contact Us](#)

Copyright © 2026 EDUCAUSE

Privacy Notice

Terms of Use

.EDU Home

Web Accessibility Assistance