

Virginia Tech Responsible AI Principles



Virginia Tech Responsible and Ethical AI Principles

- The University AI Working Group drafted a set of principles to guide the use of AI.
- The draft principles were shared through a series of presentations around campus over several months in 2024-2025.
- Based on feedback from those conversations, a final revised set of principles was presented to sponsors, Provost Cyril Clarke and COO Amy Sebring, who approved them as a part of the Virginia Tech Responsible and Ethical AI Framework.

Virginia Tech Responsible AI Principles

1 Mission Alignment

We use AI only in ways that further Virginia Tech's teaching, research, and outreach mission and honor our Principles of Community.

2 Innovation for Good

We explore new AI tools boldly while weighing benefits against risks.

3 Human-Centered Benefit

AI should extend—not replace—human insight, creativity, and well-being.

4 Responsible & Ethical Use

We consider environmental impact, intellectual property rights, and social consequences before deploying AI.

5 Fairness & Transparency

We design, procure, and use AI systems that are explainable and strive to reduce bias. We disclose AI use to those affected by it.

6 Human Judgment & Accountability

People remain accountable for decisions influenced by AI. A human should always be in the loop for important decisions.

7 Data Security & Privacy

We safeguard personal and institutional data used in AI systems and interactions.

Mission Alignment

We use AI only in ways that further Virginia Tech's teaching, research, and outreach mission and honor our Principles of Community.

The use of AI should support Virginia Tech's core missions and align with our Principles of Community, enhancing learning, fostering discovery, promoting engagement, improving the human condition, advancing knowledge, and mitigating potential harms. Every AI implementation should be evaluated against these fundamental purposes.

Innovation for Good

We explore new AI tools boldly while weighing benefits against risks.

We will foster a culture of continuous learning and improvement and embrace working with innovative technologies to create new opportunities for our students and employees. We acknowledge that the inherent complexity and rapidly evolving nature of AI technologies will necessitate an approach in which we continually question, analyze, and evaluate AI solutions for robustness, suitability, and benefits that outweigh risks.

Human-Centered Benefit

AI should extend—not replace—human insight, creativity, and well-being.

AI serves us best when its purpose is clear and its tools are well-matched to complement or enhance outcomes—whether by informing decisions, strengthening results, or improving efficiency—while working in concert with human expertise. We commit to leveraging AI technologies in ways that amplify and enhance human intelligence, creativity, and decision-making capabilities. Further, we commit to considering the psychological and social impacts of AI integration across university functions, recognizing that "beneficence" means actively working for the common good while avoiding harm.

Responsible & Ethical Use

We consider environmental impact, intellectual property rights, and social consequences before deploying AI.

We will promote understanding and discourse regarding the responsible and ethical use of AI. We will closely monitor and consider the implications and impact of the intersection of AI and intellectual property. When choosing to use AI systems, we will consider quality, sustainability, and environmental impact. Further, we will consider the findability, accessibility, interoperability, and optimization of reuse of AI digital assets.

Fairness & Transparency

We design, procure, and use AI systems that are explainable and strive to reduce bias. We disclose AI use to those affected by it.

Transparency regarding AI usage is essential to maintain public trust. We should promote attribution and transparency regarding data collection practices and AI use in decision-making. We will strive for continuous improvement in inclusivity and equality of opportunity, treatment, and impact, considering fairness in all aspects of AI use. We acknowledge that there are several potential sources of bias in AI and that our responsibility to educate ourselves about and respond to bias extends to our interactions with AI. We will prioritize explainable AI solutions that clearly outline how specific results are produced and why, allowing users to understand the reasoning behind the outputs, anticipate expected impacts, and identify potential biases.

Human Judgment & Accountability

We design, procure, and use AI systems that are explainable and strive to reduce bias. We disclose AI use to those affected by it.

People remain accountable for decisions influenced by AI. A human should always be in the loop for important decisions.

Data Security & Privacy

We safeguard personal and institutional data used in AI systems and interactions.

We will prioritize the safety, security, privacy, and protection of our community, valuing the unique benefits that human interactions bring to university processes. We recognize that people are ultimately responsible for decisions, especially in situations that directly impact others. Further, we commit to safeguarding individual privacy rights as well as proprietary university data by understanding what data is being used by AI solutions and who can access it. We should protect data from disclosure and, where appropriate, obtain consent from individuals who interact with AI.