

AI USE POLICY:

A Framework for Responsible, Educational, and Sustainable Technology Integration

1. INTRODUCTION AND PURPOSE

Our Vision

At Limestone Community School, we are committed to preparing students for a rapidly changing world while honoring our foundational values of academic excellence, environmental stewardship, and ethical citizenship. As artificial intelligence becomes increasingly integrated into daily life, we recognize both its potential to enhance learning and our responsibility to our planet and future generations.

Defining AI in Our Context

For the purposes of this policy, artificial intelligence (AI) refers to computer systems and applications that can perform tasks typically requiring human intelligence, including but not limited to: generative text tools (such as ChatGPT, Claude, and similar platforms), image generation and editing software, translation and language processing tools, data analysis and pattern recognition systems, educational adaptive learning platforms, and voice assistants and chatbots.

Our Educational Philosophy

We believe that technology should serve pedagogy, not replace it. AI tools, when used thoughtfully, can:

- Support differentiated instruction and personalized learning
- Provide immediate feedback during the learning process
- Help students explore complex problems and generate creative solutions
- Democratize access to information and expertise
- Enable students to tackle real-world challenges with powerful analytical tools

However, we also recognize that overreliance on AI can undermine the development of critical thinking, creativity, persistence, and genuine understanding. Our approach emphasizes AI as one tool among many in a rich educational ecosystem.

Connection to Our Sustainability Mission

As a school deeply committed to environmental protection and sustainable practices, we acknowledge that digital technologies, including AI, have significant environmental impacts through energy consumption, resource extraction, and electronic waste. This policy reflects our commitment to using AI intentionally and sustainably, teaching students to consider the environmental costs of their technological choices, and leveraging AI as a tool for environmental problem-solving and advocacy.

2. GUIDELINES FOR WHEN TO USE AI

The Intentionality Principle

Before using any AI tool, students and teachers should ask: *"Does this technology serve the learning goal better than available alternatives, and does the educational benefit justify its environmental cost?"*

Decision Framework

AI use could be considered when:

- Students need scaffolding or support to access challenging content
- Learners are exploring possibilities, brainstorming ideas, or generating creative options
- Immediate feedback would accelerate the learning cycle
- Students are analyzing large datasets or complex systems
- Diverse perspectives or information synthesis would enrich understanding
- The tool provides accessibility features that support students with learning differences
- Students are learning to use AI itself as a literacy skill for their future

AI use is discouraged or prohibited when:

- Students are developing foundational skills that require manual practice (e.g., early writing, basic math computation, initial research skills)
- The assessment measures independent mastery of content or skills
- The learning objective emphasizes process over product (e.g., developing persistence, creative struggle, revision skills)
- Human interaction, collaboration, or mentorship is the primary pedagogical goal
- The environmental cost (computational resources) outweighs the educational benefit
- Using AI would prevent students from engaging with primary sources, nature, or hands-on experiences

3. ACCEPTABLE USE GUIDELINES

For Students

General Principles:

- Always ask permission before using AI tools for schoolwork
- Be transparent about your AI use with teachers and peers
- Use AI to support your learning, not replace it
- Cite AI assistance appropriately in your work
- Consider the environmental impact of your technology choices
- Never share personal information with AI platforms
- Report concerning or inappropriate AI outputs to a teacher immediately

Documentation Requirements

When AI use is permitted, students must:

- Clearly indicate which portions of their work involved AI assistance
- Describe how AI was used (e.g., "Used Claude to brainstorm initial ideas" or "Used ChatGPT to get feedback on my thesis statement")
- Include relevant prompts or queries used
- Explain their own thought process and decisions
- Follow the citation format specified by their teacher

For Teachers

Instructional Integration:

- Clearly communicate AI policies for each assignment
- Design assessments that value the learning process, not just products
- Teach students how to use AI effectively and ethically

- Model critical evaluation of AI outputs
- Integrate lessons on AI's environmental impact
- Create opportunities for AI-free work that develops fundamental skills
- Use AI to provide personalized feedback and support when appropriate

Professional Use:

- Use AI to enhance, not replace, pedagogical expertise
- Protect student privacy when using AI tools
- Evaluate AI suggestions critically before implementing them
- Use AI for efficiency in administrative tasks (lesson planning, feedback drafting) while maintaining personal connection with students
- Stay informed about AI developments and sustainability implications
- Share effective AI integration strategies with colleagues

For Administration

Approved Applications:

- Student information system analytics to identify learning needs
- Scheduling optimization for efficient resource use
- Administrative task automation (report generation, communications)
- Data analysis/research to support institutional decision-making
- Brainstorming and organizational assistance in policy and grant writing

4. ENVIRONMENTAL IMPACT CONSIDERATIONS

The Hidden Cost of AI

While AI may seem "virtual" and therefore environmentally neutral, it requires significant physical infrastructure and energy. Training large AI models and processing queries consume substantial electricity, much of which comes from fossil fuels. Additionally, the servers, chips, and devices required for AI contribute to resource extraction, manufacturing emissions, and electronic waste.

Key Environmental Impacts

- *Energy Consumption:* a single AI query can use significantly more energy than a traditional search, large language model training requires massive computational resources, data centers housing AI systems require cooling and constant power, the carbon footprint varies significantly between AI providers and tools
- *Resource Extraction:* rare earth minerals and metals required for AI hardware, water consumption for cooling data centers, manufacturing impacts of devices used to access AI
- *Electronic Waste:* rapid obsolescence of AI-capable devices, disposal challenges for complex electronic components, limited recycling infrastructure for advanced technology

Guidelines for Minimizing Environmental Impact

- *Use AI purposefully, not casually:* Avoid unnecessary queries or experimentation when environmental cost isn't justified
- *Batch your work:* Prepare questions and prompts thoughtfully rather than having long, rambling conversations with AI

- *Choose appropriate tools:* Use simpler tools when they'll accomplish the task (e.g., calculator instead of AI for basic math)
- *Consider offline alternatives:* When possible, use books, observations, or human expertise instead of AI
- *Optimize your prompts:* Clear, specific prompts reduce the need for multiple attempts
- *Track your usage:* Be mindful of how often you turn to AI

Preference for Sustainable AI Platforms

When evaluating AI tools for school use, we will prioritize platforms that: use renewable energy for data centers, publish transparent environmental impact reports, commit to carbon neutrality or carbon negativity, design for energy efficiency, support device longevity rather than planned obsolescence, and demonstrate ethical supply chain practices.

Balancing Innovation and Responsibility

We recognize that AI can be a powerful tool for environmental education and problem-solving. Our goal is not to avoid AI entirely, but to use it intentionally and to teach students to make informed choices about technology use—a skill they will need throughout their lives.

5. FACT-CHECKING AND VERIFYING AI OUTPUTS

Critical Reality: AI Makes Mistakes

One of the most important things students, teachers, and parents must understand about AI is this: AI systems frequently generate false, misleading, or completely fabricated information. This phenomenon, often called "AI hallucination," occurs when AI confidently presents incorrect information as fact.

What Are AI Hallucinations?

AI hallucinations are instances where AI systems generate plausible-sounding but factually incorrect information. This can include:

- Inventing citations to non-existent books, articles, or research papers
- Creating false historical events or dates
- Fabricating statistics or data
- Generating incorrect scientific or mathematical information
- Attributing fake quotes to real people
- Inventing biographical details about real individuals
- Creating non-existent organizations, places, or events
- Mixing accurate information with fabricated details

Why This Happens

Understanding why AI hallucinates helps students use it more critically:

- *Pattern Recognition, Not Knowledge:* AI systems don't "know" facts the way humans do. They predict likely word sequences based on patterns in their training data. When uncertain, they may generate plausible-sounding but false information.
- *No Access to Real-Time Information:* Most AI systems have training data cutoffs and cannot access current information, news, or recent developments. Yet they may confidently provide outdated or invented "current" information.
- *No Built-In Fact-Checking:* AI systems generally don't verify their outputs against authoritative sources before responding. They generate text that sounds confident regardless of accuracy.

- *Training Data Limitations:* AI training data may include inaccuracies, biases, or contradictions. AI systems may reproduce or amplify these errors.
- *The Confidence Problem:* AI systems present all information with similar confidence, making it difficult to distinguish between well-supported facts and complete fabrications.

Mandatory Verification Requirements

When using AI for any academic purpose, students **MUST** verify information through authoritative sources before using it in their work. This is not optional—it is a fundamental requirement of academic integrity and responsible AI use.

Step-by-Step Verification Process:

1. Identify Claims: What specific factual claims is the AI making?
2. Assess Importance: How critical is accuracy for this information? Higher stakes = more rigorous verification
3. Search Authoritative Sources: Use library databases, academic search engines, institutional websites
4. Cross-Reference: Find the same information in multiple independent sources
5. Evaluate Source Quality: Consider authorship, publication standards, peer review, reputation
6. Document Verification: Keep records of where you verified information
7. When in Doubt, Ask: Consult teachers, librarians, or subject experts

Information Literacy Integration

Fact-checking AI is part of broader information literacy education:

- Evaluating source credibility and bias
- Understanding different types of sources and their appropriate uses
- Distinguishing primary and secondary sources
- Recognizing propaganda, misinformation, and disinformation
- Understanding peer review and editorial processes
- Developing healthy skepticism balanced with openness to new information

Developing strong fact-checking and verification skills serves students far beyond school. In an age of rapid information flow and sophisticated AI systems, the ability to distinguish accurate information from plausible-sounding fiction is a critical life skill.

6. ACCESSIBILITY AND EQUITY CONSIDERATIONS

Fundamental Commitment

At Limestone Community School, we believe every student deserves equitable access to educational tools and opportunities. AI should reduce barriers to learning, not create new ones. This section outlines our commitment to ensuring AI supports rather than undermines educational equity.

Ensuring Equal Access to AI Tools

All students will have access to approved AI tools through school devices during the school day when needed. The school will provide necessary technology rather than requiring personal devices. Internet connectivity will be available to all students during school hours. Technical support will be provided to ensure all students can use AI tools effectively.

AI as Assistive Technology

AI tools can provide crucial support for students with learning differences and disabilities. Reading support, writing support, executive function support, language support, and mathematical support are just a few of the avenues in which well-vetted and responsibly used AI tools can improve outcomes for students.

7. AI SAFETY AND WELL-BEING

Fundamental Commitment

While AI tools offer educational benefits, we prioritize student safety and well-being above technological innovation. This section outlines protections to ensure AI use supports healthy development.

Age-Appropriate Content Filtering Protection Systems

- Elementary Level (K-5)
 - All AI interactions must occur through school-approved, filtered platforms
 - Teacher supervision required for all AI use
 - Pre-screened AI tools designed specifically for children
 - Immediate adult intervention for any concerning outputs
 - No conversational AI without direct teacher involvement
- Middle School Level (6-8):
 - Filtered AI platforms with age-appropriate content restrictions
 - Monitored use with regular check-ins
 - Approved tool list vetted for safety features
 - Training on recognizing and reporting inappropriate content
 - Increased independence with continued oversight

Protocols for Handling Inappropriate AI Responses

AI systems can occasionally generate inappropriate, biased, inaccurate, or disturbing content. Students and teachers should be prepared to recognize and respond to problematic outputs. These might include: violent, graphic, or disturbing material, sexually explicit content, hate speech or discriminatory language, instructions for dangerous or illegal activities, promotion of self-harm or eating disorders, misinformation presented as fact, privacy violations, or manipulation or deceptive content.

Student Response Protocol:

1. Stop: Discontinue the conversation with the AI immediately
2. Don't share: Do not share concerning content with peers
3. Report: Tell a teacher, counselor, or trusted adult right away
4. Document: If possible, take a screenshot for reporting purposes (teacher will handle)

Teacher Response Protocol:

1. Immediate support: Ensure the student feels safe and supported
2. Documentation: Save evidence of the inappropriate content
3. Reporting: Report incident to administration and technology coordinator
4. Follow-up: Check in with student about their well-being
5. Review: Evaluate whether the AI tool should remain approved
6. Education: Use as learning opportunity about AI limitations and safety (when age-appropriate)

Administrative Response:

1. Review the incident and content
2. Contact AI platform provider if necessary
3. Assess whether tool should remain in use
4. Update filtering or monitoring as needed
5. Communicate with parents as appropriate
6. Document incident and response

8. AI EDUCATION

Fundamental Commitment

Effective AI integration requires ongoing education for all community members. This policy succeeds only when students, teachers, and parents understand both the opportunities and responsibilities that come with AI tools.

Integrated Approach to Student Digital Literacy

AI literacy is not taught in isolation but integrated across the curriculum and through dedicated lessons.

K-2 (Early Elementary): Introduction

- *Key Concepts:*
 - What is AI? (Very simple explanations: "computers that can learn and help us")
 - AI is a tool, like other tools we use
 - Computers can make mistakes, just like people
 - We don't share our names or personal information with computers
 - Playing outside is more important than screen time
 - Technology uses energy and materials from the Earth
- *Learning Activities:*
 - Teacher demonstrations of AI tools (voice assistants, photo recognition)
 - Comparing what AI says to what we learn from books and teachers
 - "Can the computer be wrong?" discussions
 - Very brief AI-assisted nature identification during outdoor time, followed by close observation without devices
 - Simple conversations about where devices come from

3-5 (Upper Elementary): Foundations

- *Key Concepts:*
 - How AI works: computers learn from lots of examples
 - AI is helpful but not perfect it makes mistakes
 - Understanding that AI can be biased
 - Basic privacy and safety with AI tools
 - When to use AI and when not to use it
 - How our technology choices affect the environment
 - AI should help us learn, not do our work for us
- *Learning Activities:*
 - Exploring how AI-powered tools work (recommendations, search results, translations)
 - Comparing AI answers to books, expert knowledge, and multiple sources
 - Practicing safe technology habits with AI

- Nature walks using AI identification apps, then deeper observation and field guide verification
- Beginning fact-checking: testing if AI information is correct
- Discussions about technology and the environment
- Simple AI citation practice
- Exploring when AI helps vs. when it gets in the way of learning

6-8 (Middle School): Development

- *Key Concepts:*
 - How AI learns from data and training
 - Understanding AI biases, limitations, and hallucinations
 - Critical evaluation of AI outputs
 - Academic integrity with AI
 - Digital citizenship and AI ethics
 - Environmental impact of AI and technology
 - Effective AI prompting and usage
 - Comprehensive fact-checking and verification skills
 - AI's role in society and future implications
- *Learning Activities:*
 - Experimenting with AI tools and comparing results across platforms
 - Analyzing AI-generated content for accuracy, bias, and hallucinations
 - Practicing appropriate AI citations in various formats
 - Researching AI's environmental footprint and comparing alternatives
 - Using AI for environmental data analysis projects
 - Structured discussions and debates about AI's role in society
 - Systematic fact-checking exercises: identifying AI hallucinations and verifying information against authoritative sources
 - Designing verification protocols for different types of information
 - Exploring AI ethics through case studies
 - Creating media literacy projects about AI misinformation

9. PRIVACY AND DATA PROTECTION

Fundamental Commitment

Protecting student privacy is paramount. As we integrate AI tools, we must be vigilant about data collection, storage, and use to safeguard our students' personal information and rights.

Data Minimization Principle

Students and teachers should share the minimum amount of personal information necessary with AI platforms.

- **Prohibited Information.** Never input into AI systems:
 - Full student names (use first names or initials only when necessary)
 - Student ID numbers or other unique identifiers
 - Addresses, phone numbers, or email addresses
 - Birthdates or ages
 - Medical or health information
 - Family financial information

- Social security numbers or other government IDs
- Photographs of students (unless using approved, secure platforms)
- Sensitive personal or family information
- Allowed Information. With appropriate context:
 - First names or initials for feedback purposes
 - Academic work content (essays, problems, projects)
 - General grade level
 - Non-identifying learning needs (e.g., "a student struggling with fractions")

Vetting Process:

Before any AI tool is approved for use, we will evaluate:

- Privacy Criteria:
 - Clear, transparent privacy policy
 - Compliance with COPPA (Children's Online Privacy Protection Act) for students under 13
 - Compliance with FERPA (Family Educational Rights and Privacy Act)
 - No sale of student data to third parties
 - Limited data retention with clear deletion policies
 - Strong data security measures
 - Parent/student rights regarding data access and deletion
 - Clear information about how data is used
 - Opt-out options available
- Educational Criteria:
 - Clear educational value
 - Age-appropriate content and interface
 - Accessibility features
 - Alignment with curriculum standards
 - Teacher controls and monitoring capabilities
- Sustainability Criteria:
 - Transparency about environmental impact
 - Commitment to renewable energy or carbon offsets
 - Energy-efficient operations
 - Responsible hardware lifecycle practices

Current Approved Tools List:

There are currently no AI tools approved for student use at Limestone Community School as of December 2025.

10. DIGITAL CITIZENSHIP INTEGRATION

Fundamental Principle

AI literacy is one component of comprehensive digital citizenship. Students need to understand how to participate ethically, safely, and responsibly in an increasingly digital world—while maintaining connection to the natural world that sustains us.

Connection to Broader Digital Citizenship Curriculum

Our digital citizenship education encompasses:

- Digital Literacy
 - Finding, evaluating, and using information effectively
 - Understanding how digital technologies work
 - Recognizing misinformation and manipulation
 - AI literacy as critical component
 - Fact-checking and verification skills
- Digital Ethics
 - Academic integrity in digital environments
 - Respecting intellectual property and copyright
 - Ethical AI use
 - Privacy and data protection
 - Environmental responsibility in technology use
- Digital Safety
 - Protecting personal information
 - Cybersecurity basics
 - Recognizing online risks
 - Appropriate help-seeking
- Digital Well-being
 - Healthy technology habits
 - Balancing screen time with other activities
 - Maintaining real-world relationships
 - Connection to nature
- Digital Communication
 - Effective online communication
 - Respectful discourse and empathy
 - Understanding audience and context
 - Productive collaboration
- Digital Participation
 - Positive contribution to online communities
 - Digital activism and advocacy
 - Creating rather than just consuming
 - Using technology for social and environmental good

AI Within Digital Citizenship:

- Literacy: Understanding how AI processes information and verifying outputs
- Ethics: Using AI with integrity and considering societal impacts
- Safety: Protecting privacy when using AI tools
- Well-being: Avoiding over-reliance on AI, maintaining human skills
- Communication: Using AI to enhance rather than replace human interaction
- Participation: Leveraging AI for positive environmental and social impact

Developmental Progression:

- Early Elementary: Foundation of digital citizenship with supervised technology use
- Upper Elementary: Building independence and judgment in digital spaces
- Middle School: Development of more sophisticated digital citizenship

Integration Across Curriculum:

- Library/media specialist coordinates digital citizenship education
- Reinforced in Project-Based Learning as appropriate
- Classroom teachers integrate in relevant contexts

Environmental Literacy in Digital Context

- The Cloud is Physical:
 - Data centers require energy, water, and physical infrastructure
 - Internet cables, servers, and devices have material reality
 - Electronic waste and resource extraction
 - Carbon footprint of digital activities
- Design Choices Matter:
 - Energy-efficient vs. energy-intensive applications
 - Planned obsolescence vs. device longevity
 - Repair vs. replacement
 - Sustainable vs. extractive supply chains
- Digital Minimalism:
 - Purposeful vs. mindless technology use
 - Quality over quantity in digital interactions
 - Reducing unnecessary data and digital waste
 - Choosing appropriate tools for tasks
- Technology for Environmental Good:
 - Environmental monitoring and data collection
 - Climate modeling and research
 - Efficiency optimization
 - Environmental education and awareness
 - Advocacy and organizing

Curriculum Integration:

- Elementary:
 - Where does technology come from and where does it go?
 - Simple concepts of energy use
 - Outdoor time as essential, not competing with screens
 - Technology as one tool among many
- Middle School:
 - Carbon footprint of technology
 - Electronic waste and recycling
 - Sustainable technology choices
 - Comparing digital vs. analog environmental impacts
 - Critical evaluation of "green" technology claims
 - Systems thinking about technology and environment
 - Innovation in sustainable technology
 - Personal and institutional carbon footprints

CONCLUSION

Our Commitment

This AI Use Policy reflects Limestone Community School's commitment to preparing students for the future while remaining grounded in timeless educational values and our deep commitment to environmental stewardship.

Balancing Innovation and Wisdom

We embrace AI as a powerful educational tool while recognizing that:

- Human connection, creativity, and wisdom remain irreplaceable
- Direct experience with the natural world is essential
- Critical thinking must be developed, not outsourced
- Academic integrity is foundational to learning
- Environmental responsibility extends to our technology choices
- Education is about becoming, not just knowing
- AI outputs must be verified before use
- Distinguishing fact from fiction is a crucial life skill

Living Document

This policy will evolve as technology advances and we learn from experience. We commit to:

- Regular review and revision
- Responsiveness to community feedback
- Transparency about our decision-making
- Flexibility in the face of rapid change
- Consistency in our core values

Partnership

Success requires partnership among students, teachers, parents, and administrators. We are all learning together how to navigate this new landscape responsibly and sustainably.

Our Hope

We hope this policy helps our students become:

- Critical thinkers who evaluate AI outputs thoughtfully
- Skilled fact-checkers who verify information before using it
- Ethical technology users who maintain integrity
- Environmental stewards who make sustainable technology choices
- Effective communicators who enhance rather than replace human connection
- Innovative problem-solvers who leverage AI for good
- Confident learners who trust their own capabilities
- Informed citizens prepared for an AI-integrated world

Together, we will prepare our students not just to use AI, but to shape its role in creating a more just, sustainable, and flourishing world.

SOURCES

LibraryReady.AI <https://libraryready.ai/>

News Literacy Project <https://newslit.org/>

Common Sense Media AI Initiatives <https://www.common sense media.org/ai>

ACKNOWLEDGMENTS

This policy was developed collaboratively by students, staff, and administrators of Limestone Community School during 2025. Fifth and sixth grade students embarked on a year-long digital and media literacy curriculum in 2024-2025 during library instruction; the foundations of this document are the culmination of their collaborative work and learning.

This comprehensive policy provides a framework for responsible, educational, and sustainable AI integration at Limestone Community School. Questions, concerns, or suggestions should be directed to Bailey Pike, Librarian/Systems Management, library@limestoneschool.com.

This policy takes effect December 2025 and supersedes all previous guidelines regarding AI use.

Next Scheduled Review: December 2026

APPENDICES:

APPENDIX A: GLOSSARY OF AI TERMS

AI (Artificial Intelligence): Computer systems that can perform tasks typically requiring human intelligence, such as understanding language, recognizing patterns, and making decisions.

AI Hallucination: When an AI system generates plausible-sounding but factually incorrect information, often presented with confidence.

Constitutional AI: AI systems designed with built-in ethical guidelines and principles that govern their behavior to promote safety, honesty, and helpfulness.

Generative AI: AI systems that can create new content, such as text, images, or code, based on patterns learned from training data.

Large Language Model (LLM): A type of AI trained on vast amounts of text data that can understand and generate human-like text.

Machine Learning: A method of teaching computers to learn from data and improve their performance without being explicitly programmed for each task.

Prompt: The input or question given to an AI system to generate a response.

Training Data: The information used to teach an AI system, which influences what it can do and what biases it might have.

Verification: The process of checking AI-generated information against authoritative sources to confirm accuracy.

APPENDIX B: PREFERENCE FOR CONSTITUTIONAL AI SYSTEMS

Our Commitment to Ethical AI Design

At Limestone Community School, we recognize that not all AI systems are created equal. Beyond evaluating AI tools for educational value, privacy protection, and environmental impact, we also consider the underlying design principles and safety frameworks that guide AI behavior.

What Are Constitutional AI Systems?

Constitutional AI refers to AI systems designed with built-in ethical guidelines and principles that govern their behavior. Like a constitution guides governance, these systems are trained to follow explicit values and rules that promote:

- Helpful and honest responses
- Harmlessness and safety
- Respect for human dignity and autonomy
- Transparency about capabilities and limitations
- Refusal to assist with harmful activities
- Fairness and reduction of bias

Why We Prefer Constitutional AI

Our preference for constitutional AI systems aligns with our school's educational mission and values:

- **Safety and Student Protection:** Constitutional AI systems are designed with stronger safeguards against generating harmful, inappropriate, or dangerous content. This reduces risk for our students and provides additional layers of protection beyond content filtering.
- **Educational Alignment:** These systems are built to promote genuine learning rather than enable shortcuts. They're more likely to encourage critical thinking, ask clarifying questions, and guide students toward understanding rather than simply providing answers.
- **Transparency and Honesty:** Constitutional AI systems are designed to acknowledge their limitations, express uncertainty when appropriate, and avoid presenting speculation as fact—modeling the intellectual honesty we want to cultivate in our students. This includes being transparent about the potential for AI hallucinations and the need for verification.
- **Values-Based Interactions:** Just as we teach students to consider ethical dimensions of their choices, constitutional AI systems are designed to consider ethical implications in their responses, providing a model of thoughtful, values-informed engagement.
- **Harm Prevention:** These systems include explicit training to refuse assistance with activities that could cause harm—including academic dishonesty, dangerous activities, or circumvention of safety measures—supporting rather than undermining our community standards.

Evaluation Criteria

When we evaluate AI tools for approval, we consider:

- Design Philosophy: Does the company demonstrate commitment to ethical AI development and constitutional principles?
- Transparency: Is the company clear about its AI training methods, safety measures, and limitations?
- Safety Track Record: What is the history of harmful outputs or safety incidents with this system?
- Alignment with Educational Values: Does the AI promote learning, critical thinking, and intellectual honesty?
- Continuous Improvement: Does the company actively work to improve safety and reduce biases based on feedback?
- Factual Accuracy Measures: Does the system acknowledge uncertainty and encourage verification of factual claims?

Not a Perfect Solution

We acknowledge that even constitutional AI systems have limitations:

- No AI system is completely free from errors or biases
- Constitutional frameworks themselves reflect specific value choices
- AI responses should still be critically evaluated
- Human judgment and oversight remain essential
- All AI systems can produce hallucinations that require verification

Our Approach

While we prefer constitutional AI systems, we:

- Recognize this is an evolving field with varying approaches to AI safety
- Remain open to new developments in ethical AI design
- Continue to emphasize critical evaluation of all AI outputs regardless of system design
- Maintain multiple layers of protection: tool selection, content filtering, supervision, and education
- Teach students that responsible AI use depends on both system design AND user choices
- Emphasize verification of all factual information regardless of source

Teaching Opportunity

Our preference for constitutional AI provides educational opportunities to discuss:

- How values and ethics can be embedded in technology design
- The relationship between system design and user responsibility
- Critical evaluation of different approaches to AI safety
- Broader questions about technology ethics and governance
- How our choices as consumers can influence AI development
- The importance of verification even with well-designed systems

Questions to Explore with Students

- What values should guide AI behavior?
- Who decides what principles AI systems should follow?
- How do we balance different ethical considerations?
- What are the limits of building ethics into technology?
- How does this relate to human ethical development and decision-making?
- Why must we verify information even from ethically-designed AI?

Ongoing Commitment

As AI technology rapidly evolves, we commit to:

- Staying informed about developments in AI safety and ethics
- Prioritizing constitutional and ethically-designed AI systems when available
- Engaging with AI developers about our educational values and safety needs
- Participating in broader conversations about responsible AI development
- Regularly reassessing our approved tools against evolving ethical standards
- Teaching students to be critical consumers of all AI systems

By preferring constitutional AI systems, we reinforce our commitment to creating a learning environment where technology serves our values, protects our students, and models the thoughtful, ethical approach to innovation we hope to inspire in the next generation.

For More Information:

Bai, Yuntao, et al. "Constitutional ai: Harmlessness from ai feedback." *arXiv preprint arXiv:2212.08073* (2022).

<https://doi.org/10.48550/arXiv.2212.08073>