Artificial Intelligence Guidelines

AI

Understanding and Using AI for Work at Rhodes College

Rhodes College

Table of Contents

Artificial Intelligence Guidelines

Understanding and Using AI for Work at Rhodes College

Overview

Artificial Intelligence (AI) will allow the Rhodes College community to work in new and exciting ways. Generative AI is used in industry and academia to streamline administrative processes and increase operational efficiencies. While AI offers undiscovered possibilities, it also presents new ethical and security challenges. Our goal as a college community is to find new ways to enhance our work while using AI responsibly and ethically.

This guide is designed to provide the Rhodes College community with the base knowledge and tools needed to use AI responsibly and effectively. It outlines practical applications and best practices to ensure that our use of AI aligns with the ideals of our college community. Following these guidelines will allow us to leverage AI's potential while maintaining our values of transparency, integrity, and respect for privacy. This document is meant to empower our community to make informed decisions that allow us to integrate AI into our everyday work in beneficial ways. As a best practice, always consult your supervisor before using AI to perform work tasks.

Artificial Intelligence Policy Framework

The college's new AI policy, which can be reviewed in the Appendix of this document, was drafted to help our community use AI ethically, transparently, and securely. Crafted by the AI Working Group and approved by the Senior Leadership Team, the policy will be included in the next iteration of the employee handbook.

At its core, the AI policy emphasizes the respectful acknowledgment of intellectual property, the protection of data privacy, and the importance of adhering to ethical standards to prevent bias and misuse. Employees are advised to avoid entering sensitive information into AI tools to safeguard privacy and security. Sensitive data types include personally identifiable information, confidential business information, health records, and security credentials to protect privacy and security.

The policy also highlights the need to document AI tools' purpose, scope, decision-making processes, and potential risks. It also outlines plans for regular audits to ensure compliance with privacy laws and ethical standards, affirming our community's dedication to responsible and secure AI use.

What is Artificial Intelligence?

Artificial Intelligence refers to computer systems capable of performing tasks that typically require human intelligence. These systems are designed to mimic human thought processes such as learning, problem-solving, and decision-making. AI uses algorithms and data to identify patterns, adapt from experiences, and generate outputs across various domains and of varying types.

Generative AI is a subset of artificial intelligence designed to create new content based on patterns learned from existing data. Generative AI is creative by nature. This new content could be text, images, music, or anything a human could create.

Today's Generative AI Tools

AI has historically been used for specialized tasks like automated data entry, predictive analytics, and basic decision-support systems. These early applications paved the way for today's more sophisticated generative AI tools, which can handle more complex tasks like drafting reports, summarizing data, or generating creative content.

Generative AI tools are built using large language models (LLMs) which are trained on vast datasets, enabling them to produce human-like outputs across various disciplines. Several new generative AI tools exist and are used to perform various tasks. Today, well-known tools like OpenAI's ChatGPT, Google's Gemini, and Anthropic's Claude are paving the way and allowing users to create more efficient workflows, helping to create content, and allowing users to perform complex data analysis more quickly.

The AI tools listed below are all generative AI tools that can be used to accomplish a wide variety of tasks. The tools below all have free and paid versions. Microsoft Copilot is the only enterprise tool listed below.

- **Microsoft Copilot** (https://copilot.microsoft.com): Can be Integrated with Microsoft Office products or used as a standalone tool. Microsoft Copilot is the recommended AI tool for the campus. Copilot can be used to draft documents, analyze data, create presentations, or for other generative AI tasks. *Please be aware that our version of Copilot is the standalone version. The standalone version does not integrate directly into the Microsoft Office 365 suite of products (Microsoft Word, Excel, PowerPoint, Outlook, etc.)*.
- **Claude**: A tool designed for natural language understanding and human-like conversation, often employed for both academic and administrative purposes.
- **Gemini**: Known for its versatility and long-context understanding, Gemini is used to produce in-depth analyses and creative content.
- **Perplexity**: An AI-powered tool that offers advanced information retrieval and question-answering capabilities.
- **ChatGPT**: Developed by OpenAI, this tool is widely used for generating text, answering questions, and aiding in content creation.
- **DALL-E**: Designed by OpenAI this tool is used to generate high-quality images and visual content from text descriptions.
- **Adobe Firefly**: Adobe's AI tool is designed for creative projects, allowing users to generate high-quality images and visual content.

Microsoft Copilot

Rhodes College recommends Microsoft Copilot as the preferred AI tool for college business use. You can access Copilot by navigating to https://copilot.microsoft.com/. Our Microsoft 365 license gives us access to the standalone version of Microsoft Copilot. This version of Copilot works with OneLogin, allowing us to use our Rhodes account to log in. Copilot's built-in enterprise data protection features offer a significant advantage over free generative AI tools. The data protection provided by Microsoft is more robust and provides greater protection than what is offered in free versions of AI tools like ChatGPT, Google Gemini, or other non-enterprise AI tools. These protections include encrypting data during use and while stored, keeping our information separate from other users, and following strict security practices to protect our data physically and digitally. Copilot also meets privacy laws like GDPR to keep user data private. Additionally, Copilot follows our existing Microsoft security settings and permissions, supports audit logs for tracking usage, and can use data sensitivity labels.

Access to Copilot is secured through OneLogin, our multi-factor authentication (MFA) provider. Using MFA adds an extra layer of protection by helping to prevent unauthorized access. This combination of features makes Copilot a reliable and secure choice while maintaining the college's commitment to data security and the ethical use of AI.

Key Reasons to Use Microsoft Copilot

- Included with our Microsoft 365 license, making it easy and cost-effective.
- Built-in data protection measures, including encryption and data isolation.
- Compliance with privacy standards like GDPR and ISO/IEC 27018.
- Secure login through OneLogin MFA for added security.
- Integration with existing software and user permissions.

Microsoft Copilot Training Links

- **Get Started with Microsoft 365 Copilot**: A beginner-friendly course that outlines ways to integrate Copilot into daily workflows. https://learn.microsoft.com/en-us/training/paths/get-started-with-microsoft-365-copilot/
- **Copilot Foundations Training**: Introduction to the basics of Copilot and its capabilities. https://learn.microsoft.com/en-us/training/paths/copilot-foundations/
- **Optimize and Extend Microsoft 365 Copilot**: This module explores best practices to leverage Copilot's functionality. https://learn.microsoft.com/en-us/training/modules/optimize-and-extend-microsoft-365-copilot/
- **Microsoft Copilot Video Tutorials**: A collection of instructional videos demonstrating Copilot's functionalities across different applications. https://support.microsoft.com/en-us/topic/microsoft-copilot-video-tutorials-25a1b328-79be-4e8a-af96-1f894e52bcf6
- **Microsoft 365 Copilot Documentation**: Detailed information on Copilot's features and best practices. https://learn.microsoft.com/en-us/copilot/microsoft-365/

Please see the appendix to review instructions that detail how to access our Microsoft Copilot instance.

Tips and Best Practices for Effective AI Use

AI tools, particularly generative AI, offer significant potential to enhance productivity and creativity. However, AI's use, from the prompt entered to the content generated, requires careful consideration and oversight. Staff should always discuss the use of generative AI tools with their supervisors to ensure appropriateness for the work being completed. All AI-generated content must be thoroughly reviewed and aligned with institutional policies, especially when the task impacts decision-making. The list below is not exhaustive, and you are encouraged to explore additional AI applications that align with their work and institutional policies.

Recommended Uses

- **Creative Thought Partner:** Use AI for brainstorming, drafting, or enhancing creative content such as marketing materials, newsletters, and presentations. AI can also provide novel suggestions and alternative approaches that support innovative thinking.
- **Streamline or Automate Routine Tasks:** Automate processes like organizing meeting notes, generating email responses, and performing simple data analysis, which frees up time for more critical work.
- **Research and Information Gathering:** Use AI to quickly gather background information, explore recent trends, or identify relevant data on a given topic. AI can also compile summaries of large or complex documents to support informed decision-making.
- **Data Insights and Analysis:** Utilize AI to identify patterns, trends, or anomalies within datasets, providing predictive insights and enhancing reporting. AI tools can also help organize data for presentations, dashboards, or other reporting purposes.
- **Use AI for Idea Validation:** Before fully committing to a new project or concept, use AI to quickly generate potential pros, cons, and considerations.

Best Practices

- **Define Objectives**: Clearly define your objectives and expectations before using AI. Doing so will allow you to set realistic expectations and generate the output you are looking for.
- **Review and Verification**: Ensure that AI-generated content is up to date by verifying AI outputs, particularly when handling policy-sensitive or public-facing information.
- **Request Sources**: If you are using AI for research or data analysis, ask it to provide citations or references to verify the accuracy of the information.
- **Start with Simple Tasks:** Build your confidence and understanding of AI capabilities by using AI for simple tasks like summarizing short texts or generating basic ideas.
- **Adaptation and Continuous Learning**: Stay informed about updates in AI technology and adjust strategies to incorporate new tools responsibly and ethically.
- **Garbage In, Garbage Out**: The quality of AI outputs is directly related to the quality of the input data. Poor-quality data leads to poor-quality outputs. This principle underscores the importance of using clean, accurate, and relevant data.
- **Diverse Data Sources**: Use a variety of reputable data sources to give AI a comprehensive and well-rounded understanding of the topic. This approach helps reduce biases and ensures a more balanced perspective.

AI Prompting Recommendations

1. **Be Specific**: Provide clear, detailed prompts to get more accurate responses. For example, instead of asking, “How do I improve department efficiency?” try asking, “What strategies can increase departmental productivity based on graduation rates and student feedback?”
2. **Clarity and Context**: Provide background information and context when crafting prompts. Assign a specific role to the AI (e.g., "research assistant"), clearly state the task, specify the desired output format, and indicate the preferred response length. Following this approach will yield more accurate and relevant AI responses tailored to your needs.
3. **Step-by-Step Instructions (Chaining)**: Break your request down into smaller steps for more complex tasks. This helps the AI focus on the specific aspects of the problem and ensures it handles each element thoroughly.

Figure 1 – Image: AI Prompting Bot - Created using DALL-E from OpenAI.

4. **Limit the Scope**: Narrow the topic to ensure more precise responses. For example, instead of asking for general insights on “student engagement,” specify, “What strategies can improve student engagement in small liberal arts colleges?”
5. **Tone and Style**: If you need the AI to match a certain style, specify it upfront. For instance, ask it to “write this email in a professional tone” or “summarize this report in a casual style.”
6. **Define Output Expectations**: Clearly specify the desired format, such as tables, bullet points, or other formats. For example, you might request, “Present this data in a table”.
7. **Interview Me**: Ask AI to interview you, this will help it gain a greater understanding of the task (ex. *Interview me, ask me 3 questions if you need to gain a deeper understanding*.).
8. **Iterate and Experiment**: AI responses can vary, so iterative adjustments to prompts can help refine the results for more accuracy and relevancy.

AI Guidelines for Ethical and Responsible Use

Transparency and Documentation

All AI applications must be documented and openly communicated to relevant departmental stakeholders. This includes the purpose, scope, and the decision-making processes involved. The documentation should be clear and should outline the associated risks of the AI application.

Respect for Privacy and Data Security

Do not enter sensitive or confidential information into AI tools, particularly free or unapproved platforms. The following data types should never be entered into generative AI platforms:

- **Personally Identifiable Information (PII)**: Full names (your own or others), addresses, Rhodes IDs/R#s, Social Security numbers, financial information, etc.
- **Confidential Business Information**: Proprietary college data, internal business strategies, non-public financial data, etc.
- **Health Information**: Medical records, disability accommodations, mental health records, etc.
- **Legal Information**: Confidential legal documents, privileged communications, etc.
- **Sensitive Personal Information (SPI)**: Political opinions, religious beliefs, ethnic origins, etc.
- **Security Information**: Passwords, encryption keys, security codes, etc.

Risks and Limitations

AI models can produce convincing human-like outputs. As advanced as AI is, it is important to acknowledge its inherent risks and limitations. Artificial intelligence is still only an advanced pattern recognition system that does not possess a fundamental understanding of language or real-world knowledge. AI does not understand the context of the questions we ask it to answer.

Generative AI can assist administrators with tasks like drafting reports, composing emails, or creating event announcements. Rhodes will employ a Human-in-the-Loop model for working with AI because it works best when used to augment human work. The most advanced versions of AI cannot fully replicate human thought and are not capable of replacing skilled workers. Common shortcomings include:

- **Accuracy and Oversight:** AI tools, particularly generative AI, sometimes produce believable but inaccurate information, known as "hallucinations." Since these models lack true understanding, human oversight is critical to ensure the accuracy and reliability of AI outputs, especially in sensitive or complex contexts.
- **Empathy and Judgement:** Human involvement is essential for interpreting AI outputs because generative AI models are not replacements for human judgment and should not be relied upon for decisions that require empathy, context, or nuanced understanding.
- **Bias and Fairness Concerns:** AI systems may inadvertently perpetuate biases in the training data used to teach it. These biases may lead to unfair outcomes. Users must review AI outputs critically and take corrective actions if biases are identified.

- **Intellectual Property and Copyright:** Outputs from generative AI may inadvertently include material influenced by copyrighted works. Community members must ensure that all AI-generated content is appropriately credited and must comply with copyright laws.
- **Human Involvement and Ethical Decision-Making:** Artificial intelligence is meant to augment, not replace humans. Rhodes prioritizes a Human-in-the-Loop (HITL) model, where human participants actively engage with AI outputs to ensure reliability and ethical alignment. This kind of involvement helps minimize bias in our AI systems.

Feature	Human-in-the-Loop (HITL)	Human-on-the-Loop (HOTL)	Human-out-of-the-Loop (HOOTL)
Human role	Active participant	Oversight and intervention	Minimal to none
AI autonomy	Limited	Moderate	High
Decision-making	Collaborative	Machine-led with human oversight	Fully automated
Scalability	Limited	Moderate	High
Accuracy	High	Moderate to High	Variable
Bias mitigation	Strong	Moderate	Weak to Moderate

Figure 2 – Human/AI Involvement Chart

- **Limitations of Training Data and Context:** AI models depend on training data, which can lead to inaccuracies if the input data is flawed or incomplete. This "garbage in, garbage out" principle underscores the importance of vetting AI outputs against verified sources.

AI has significant potential to enhance productivity, but our community must remain aware of its limitations. Maintaining a Human-in-the-Loop model will allow us to leverage AI's capabilities while ensuring its output aligns with our institutional values and ethical standards.

Conclusion

Artificial Intelligence at Rhodes

These guidelines serve as both a roadmap and a commitment to harnessing artificial intelligence in ways that enhance our work while upholding our high standards of ethical and responsible use. By integrating generative AI tools like Microsoft Copilot within our established frameworks, we are empowered to streamline operations, foster creativity, and drive innovation—all while ensuring transparency, safeguarding data, and honoring the principles of intellectual integrity.

Moving forward, it is essential that every member of our community continues to engage with these tools thoughtfully, consult with supervisors on their application, and remain vigilant about the potential risks and limitations inherent in AI technologies. Through a steadfast commitment to a human-in-the-loop approach and ongoing dialogue about best practices, we will collectively navigate the evolving landscape of AI and secure a future that is both innovative and ethically sound.

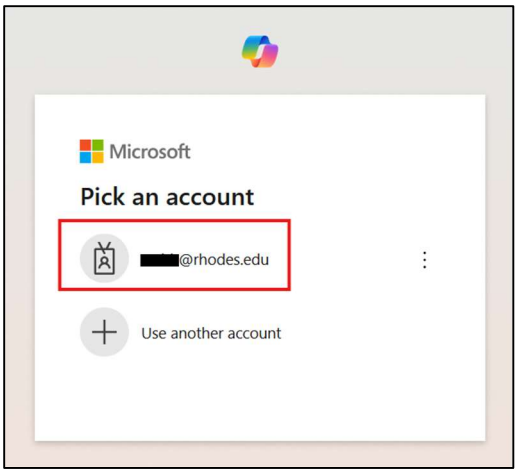
Appendix

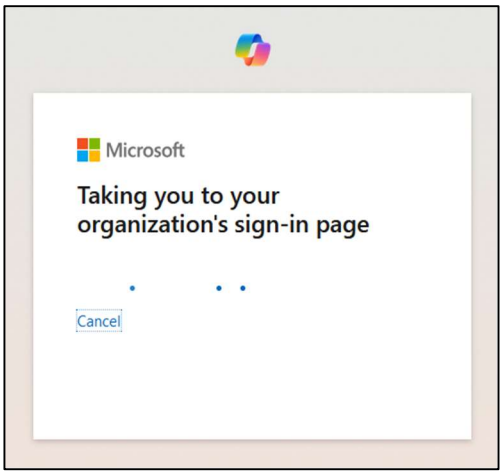
How to Access and Use Microsoft Copilot

Microsoft Copilot is the recommended AI tool for business use at Rhodes College. Follow these steps to get started:

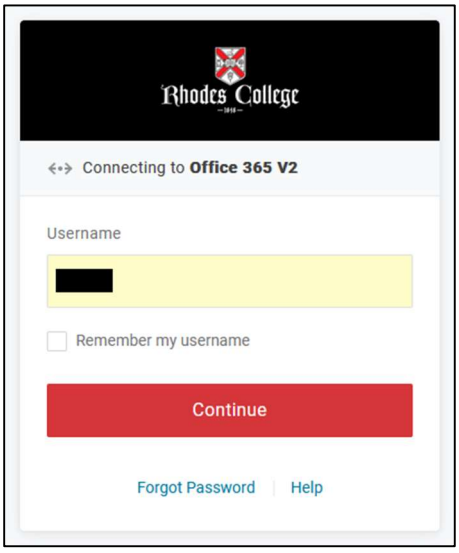
1. **Access Microsoft Copilot**
   a) Open a web browser and navigate to https://copilot.microsoft.com.

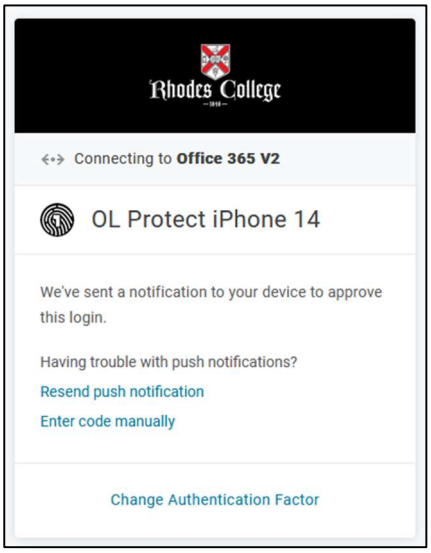
2. **Log into the Microsoft Copilot**
   a) When prompted select your Rhodes credentials "**Username@rhodes.edu**" to sign into your account.



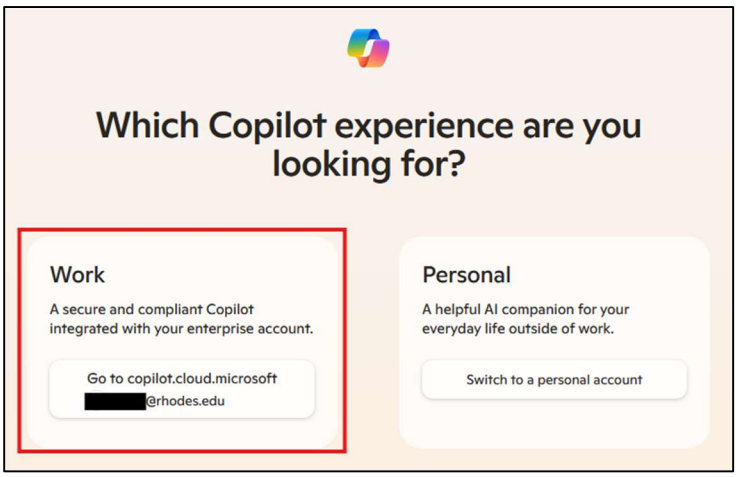


   b) Use your Rhodes College credentials to log in through **OneLogin** (multi-factor authentication).





c) Select "**Work**" when asked "Which Copilot experience are you looking for?"



3. Start Using Microsoft Copilot

a) Once the app launches, you'll see a prompt box at the bottom of the screen.

b) Enter your queries or commands to begin generating outputs with Copilot (Examples of tasks include drafting documents, analyzing data, and generating creative content).

Glossary of Terms

1. **Algorithms**: Step-by-step instructions that computers follow to solve problems or make predictions.
2. **Alignment (in AI):** Making sure an AI system works in a way that matches human values, goals, and expectations. For example, if an AI is designed to help with homework, alignment means ensuring it gives helpful and accurate answers, not misleading or harmful ones.
3. **Artificial Intelligence (AI)**: Technology that lets machines act like humans, helping with tasks like learning, problem-solving, or making decisions.
4. **Automation (in AI)**: Using AI to do task processes automatically, such as sending email replies to common inquiries of prospective students, etc.
5. **Bias (in AI):** When an AI system produces results that are influenced by imbalances or patterns in the data it was trained on, leading to inaccurate or unequal outcomes.
6. **Chatbots**: Computer programs that talk with you, often used for customer service or answering questions. For example: when the chat screen pops up in the bottom right corner of a website asking if you need help finding something.
7. **Computer Vision**: AI's ability to "see" and understand pictures or videos.
8. **Data Anonymization**: Removing personal information from data to keep it private.
9. **Generative AI**: A type of AI that creates text, images, or music. This includes models like GPT for text and other systems for visuals or audio.
10. **GPT (Generative Pre-trained Transformer)**: A specific type of generative AI model designed to understand and produce human-like text, such as Copilot, trained on large amounts of text data and can perform tasks like answering questions, writing stories, or summarizing information.
    - 10.1 **Custom GPT Development:** The process of creating a tailored version of a GPT model by providing it with specific data, fine-tuning and testing its responses, or integrating it into custom workflows. This allows the GPT to specialize in particular tasks or contexts, such as automating processes or providing targeted support to users.
11. **Guardrails (in AI):** Rules or safety measures built into an AI system to prevent it from making harmful, inappropriate, or incorrect outputs. For example, an AI chatbot might have guardrails to avoid giving medical advice or answering unsafe questions, ensuring it stays within its intended use.
12. **Hallucinations**: Mistakes where AI makes up answers that sound right but are not true.
13. **Human-in-the-Loop (HITL)**: A process where people check and guide the AI's work to make sure it is correct and ethical.
14. **Large Language Models (LLMs)**: Advanced AI systems trained on vast amounts of text to write and understand human language.
15. **Learning**. The process where AI improves its ability to perform tasks or make predictions by analyzing data. It is similar to how humans learn from experience, but for AI, it involves recognizing patterns in data to "know" more and get better over time. Machine Learning is one method by which AI achieves learning.

16. **Machine Learning (ML)**: A way for computers to learn and improve by looking at data, without being told exactly what to do. This is a specific implementation of learning in AI systems.
17. **Model**: A program or system in AI that has been trained to do specific tasks, like answering questions or writing text. For example, GPT is a language model that has learned how to understand and create human-like text.
18. **Neural Networks**: AI systems inspired by the way the human brain works, used to recognize patterns and make decisions.
19. **Pattern Recognition**: AI's skill in finding patterns in data, like grouping similar items or predicting what comes next.
20. **Prompting**: Writing clear instructions for AI to get the best results.
21. **Prompt Chaining:** The process of breaking a broad or complex prompt into smaller, more focused steps to help the AI provide better and clearer results. For example, instead of asking, "Tell me everything about climate change," you can break it into smaller prompts like, "What are the causes of climate change?" followed by "What are its effects on agriculture?" and "What are the solutions?" This makes it easier for the AI to give accurate and detailed answers.
22. **Training Data**: The information used to teach AI models how to perform specific tasks.

Sources

1. IBM: What is Generative AI? - https://www.ibm.com/topics/generative-ai
2. McKinsey & Company: What is Generative AI? - https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai
3. PCMag: How to Use Microsoft's Copilot AI, and 10 Things to Try Right Away. - https://www.pcmag.com/how-to/how-to-use-microsoft-copilot-ai-what-to-try
4. Harvard University IT: AI Prompting Guide. - https://huit.harvard.edu/news/ai-prompts
5. Microsoft Learn: Enterprise data protection in Microsoft 365 Copilot and Microsoft Copilot. - https://learn.microsoft.com/en-us/copilot/microsoft-365/enterprise-data-protection
6. Microsoft Learn: Overview of Microsoft Copilot. - https://learn.microsoft.com/en-us/copilot/overview
7. OpenAI Help: Best Practices for Prompt Engineering. - https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-the-openai-api
8. OpenAI Help: ChatGPT Capabilities Overview. - https://help.openai.com/en/articles/9260256-chatgpt-capabilities-overview
9. Label Your Data: Human-in-the-Loop in Machine Learning: Why Does It Matter? - https://labelyourdata.com/articles/human-in-the-loop-in-machine-learning
10. McKinsey Quarterly: Implementing Generative AI with Speed and Safety. - https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/implementing-generative-ai-with-speed-and-safety
11. The Chronicle of Higher Education: Is It Time to Regulate AI Use on Campus? - Is It Time to Regulate AI Use on Campus?

Generative AI Policy

Rhodes College Policy for Using Generative AI Tools

Introduction

This policy aims to provide Rhodes College employees with a framework for the responsible and effective use of ChatGPT and other generative AI tools for administrative purposes. The goal of this policy is to ensure that AI is used ethically, transparently, and safely to enhance operational efficiency while safeguarding data privacy and security.

General Principles

1. **Respect for Intellectual Property**: Acknowledge the use of AI tools where applicable and respect copyright laws and licensing agreements related to AI-generated content.
2. **Privacy and Confidentiality**: Do not share sensitive or confidential information with AI tools where data privacy is not guaranteed. Exercise caution even when data confidentiality is guaranteed.
3. **Ethical Use**: Utilize AI tools in a manner that aligns with the college's commitment to ethical standards, avoiding uses that may perpetuate bias, discrimination, or harm.

Guidelines for Administrative Use

1. **Ethical Framework and Transparency**
    a. **Ethical Guidelines:** AI technologies must be developed and used in accordance with ethical principles that ensure fairness, accountability, and transparency. The outputs from AI algorithms must be actively reviewed for bias and if found steps must be taken to eliminate the bias.
    b. **Transparency in AI Usage:** All AI applications must be documented and if required, openly communicated to relevant departmental stakeholders. This includes the purpose, scope, decision-making processes involved, and risks associated with the AI application.
2. **Data Privacy and Security**
    a. **Data Protection:** Rhodes College is committed to protecting the privacy and integrity of all data processed by AI systems. All data must be handled in compliance with applicable privacy laws and regulations, such as FERPA. Measures should be in place to prevent unauthorized access, data breaches, and misuse of personal information.
        I. **Prohibited Data Types**: The following data types should not be entered into generative AI platforms:
            - Personally Identifiable Information (PII): Full names, addresses, phone numbers, Social Security numbers, financial information, student ID numbers, academic records, financial aid details, disciplinary records, etc.

- Confidential Business Information: Proprietary college information, internal business strategies, non-public financial data, contractual agreements, etc.
- Health Information: Medical records, health insurance information, mental health records, disability accommodations, etc.
- Legal Information: Confidential legal documents, client-attorney privileged communications, pending litigation details, etc.
- Sensitive Personal Information: Political opinions, religious beliefs, sexual orientation, ethnic or racial origins, etc.
- Security Information: Passwords, security codes, encryption keys, access credentials, etc.

b. **Regular Audits:** Audits of AI systems will be conducted as needed to ensure compliance with this policy and relevant data protection laws. These audits will evaluate the effectiveness of security measures and the integrity of data processing activities.

3. **Operational Efficiency and Support**

a. **Administrative AI Applications:** AI tools may be used to enhance administrative tasks, including but not limited to, student inquiries, financial modeling, and workflow automation. These tools should be designed to augment human capabilities, not replace human oversight.

b. **Regular Evaluation:** Departments will regularly evaluate the impact of AI usage on business processes, and adjustments will be made as needed to optimize performance.

4. **Innovation and Continuous Improvement**

a. **Encouraging Innovation:** Rhodes College encourages the innovative use of AI to improve administrative processes. However, all innovative efforts must align with ethical guidelines and operational standards to ensure they contribute positively to the institution's goals.

5. **Monitoring and Review**

a. **Policy Updates**: These guidelines will be reviewed and updated as needed to adapt to technological advancements and emerging ethical considerations.

Conclusion

The responsible use of generative AI tools at Rhodes College may contribute to administrative efficiency. Adhering to this policy will ensure that these technologies are used in a manner that aligns with the college's values. Please contact the office of Information Services (901-843-3745) if you have questions or if you need further clarification about this policy.