

Guidelines for the Use of AI Tools at Chapin Hall

Updated: May 2024



Introduction

These guidelines are meant to ensure that AI (Artificial Intelligence) tools used at Chapin Hall are implemented responsibly and ethically, with an overarching commitment to protecting the sensitive data we work with, maintaining the privacy and rights of individuals, and contributing positively to the decision-making processes impacting children, youth, and families. We expect the field to advance rapidly and that these guidelines will evolve and become more detailed with greater specificity as that evolution occurs.

Data Privacy and Security

AI tools should only be used in strict accordance with all applicable laws and regulations related to data protection and privacy, such as the Family Educational Rights and Privacy Act (FERPA), Health Insurance Portability and Accountability Act (HIPAA), and the Privacy Act of 1974.

- Users should be aware that any data shared with a third-party AI should be considered public.
- Under no circumstances should sensitive data (e.g. PII / PHI) be shared with third parties, such as online AI chat bots like ChatGPT, unless explicitly authorized by all relevant parties, in compliance with legal obligations.
- Users should be aware that generative AI tools will send data to an external server even when accessed through desktop applications

Transparency and Accountability

- Users of AI tools for analysis or subject interaction must be transparent about their use, sharing the purpose, limitations, and potential outcomes. This applies both internally and with external partners or affected communities. For example, if AI is part of research methodology this must be mentioned in reports or articles.
- Authorship: APA recommends that text generated from AI be formatted as "Personal Communication." As such, it receives an in-text citation but not an entry on the References list. Examples: (OpenAI, personal communication, January 16, 2023).
- There should be clear accountability for decision-making processes involving AI, addressing any ethical or practical issues that arise. For example, all sources cited, or content generated by AI should be confirmed as accurate, applicable, appropriate, and not plagiarized. Plagiarism can be checked by using an internet search for any text generated or using a plagiarism checker like Grammarly. This should include administrative decisions as in HR and hiring.
- All mathematical results should be double checked for accuracy.

Bias and Fairness

- Users of AI tools should be aware that many of these tools have been shown to have significant problems with bias and fairness. This must be accounted for and monitored.
- AI tools should be examined for potential biases in their algorithms and data, particularly those that may perpetuate or amplify social inequalities. Mitigation strategies should be employed to minimize such biases. This may include the reviewer's judgment regarding any output generated by an AI tool.
- The output of AI systems should be continuously monitored and evaluated to ensure

Appropriate Use

- AI tools should only be used when they can add value, increase efficiency, or improve the quality of our work. They should not replace human decision-making where nuance, personal judgment, or direct human contact is vital.
- AI should be used cautiously with administrative decision making. In hiring processes, for example, AI can quickly scan application packets for specific qualifications or characteristics. However, depending on AI exclusively or using it to rank candidates can undermine our broader understanding of expertise and can miss nuanced cultural fit characteristics.
- Prompts when using AI tools should be thorough and carefully crafted. Since large language models will insert false information and plagiarized content, all AI-generated content must be carefully edited and reviewed. Custom GPTs are created by uploading selected reports and other information that then serve as the exclusive sources to be used in response to queries. These closed GPT systems eliminate the risk of inaccuracies and hallucinations, but still should be reviewed carefully by those familiar with the source content. (Chapin Hall is in the process of developing its own GPT, which will rely exclusively on Chapin Hall research and policy reports.)
- Extreme care should be used when using predictive analytics and machine learning in ways that could affect individuals or communities e.g., ignoring impact of systemic racism and equitable access to resources and opportunities.
- Specific details should never be provided to third party hosted AI tools unless a specific data sharing agreement is in place. Some Examples:
 1. PII / PHI or any non-public data
 2. Server names / paths
 3. File Names
 4. Data set and field names from non-public data sets.
 5. Code that includes any of the following: Server info, field names, dataset names,
 6. Anything else you don't want shared outside of Chapin Hall.
 7. Assume any information shared with a generative AI like ChatGPT is public.

User Consent and Rights

- If AI tools involve direct interaction with individuals, such as chatbots or survey tools, users must be informed about the use of AI and their consent should be obtained wherever possible.
- AI tool use must respect the rights of individuals, including the right to information, the right to opt out, and the right to explanation about decisions made by AI that affect them.

Ongoing Evaluation and Improvement

- AI tools should undergo regular evaluation to assess their effectiveness, ethical considerations, and social impact.
- Lessons learned from evaluations should be used to improve future use of AI tools and inform development of AI policies and guidelines, including these.
- Users of AI tools should be aware that third-party tools can change or be updated without warning.

Training and Awareness

- These guidelines will be disseminated on an annual basis or sooner if significantly changed or updated.
- Chapin Hall should cultivate an environment of ongoing learning and awareness about AI, including potential future uses and developments in the field.
- AI, particularly generative AI, is a rapidly evolving technology and users have a responsibility to pay attention to the state and evolution of the technology.

AI Service Licenses

- You may use free AI services like ChatGPT, Bard, Co-Pilot, and Claude.ai
- If you'd like to use a paid AI service, please contact Research Technology for licenses.

By following these guidelines, Chapin Hall will be able to use AI tools in a responsible and ethical manner, ensuring that the benefits of AI are realized while minimizing potential harm and maintaining our commitment to protecting children, youth, and families