Generative AI -- Best Practices and Resources

Last Modified on 03/13/2024 11:14 am EDT

This article shares what Wharton Computing has learned about key concepts and emerging best practices in the rapidly evolving field of AI. It provides resources and suggests some useful paradigms that may help you think about the use of AI in our academic setting.

Faculty
Please see AI Best Practices -- Teaching and Research (for faculty and staff -- login required) to explore additional resources that support teaching and research activities.

Not sure how to log in? See Logging into the Wharton Computing Knowledge Base.

> **Check Back Regularly for Updates**
> This article will be regularly updated because the knowledge about generative AI is quickly evolving.

Key Concepts Everyone Should Know

Protect Information (yours and others')

- Be cautious of sharing confidential information with publicly available models, such as ChatGPT, Bard, Midjourney, and others.
- When in doubt, please collaborate with Wharton Computing and the Wharton Information Security Office.

> **Submission of University Confidential Data is prohibited**
> See the Security and Privacy Requirements section below for examples of confidential data.

If you are **considering any generative AI services** in your research for official use within Wharton (e.g. interconnections between Wharton Systems or Applications, or use of University Confidential Information), **please collaborate with Wharton Computing, the Research and Analytics team, and the Wharton Information Security Office**. These groups will partner with you to evaluate your project, provide the appropriate security risk evaluations, and help answer many of the IT support questions you may have.

Table of Contents

General Information

These general guidelines apply to everyone who uses AI in a Higher Education setting:

- **Stay Up to Date:** Continue to check news and recent legal opinions, especially regarding copyright enforcement and the security of individual products, and be prepared to adjust accordingly. For suggestions on staying up to date, see Additional Resources below.

- **Confirm Image Sources**: If you're using an image generator, research the dataset that the model was trained on. If there is information available from the company about their use of images, check whether they used open-source content. There may be a risk of copyright infringement if non-open-source content is used. If you do decide to use content generated by one of these tools, include a credit to the source (e.g. *Image generated by <insert tool name here>*).

- **Consider Authorship**: Since legislation trails technology, be aware that any content created by AI tools may not be legally protected or your work may be called into question. Authors who use AI tools are expected to disclose usage, preferably by identifying the one that was used. For example, "Foreground model created by Midjourney and sketched by the author."

  If seeking to copyright, remember the goal is to show how the tool was used to support the author's traditional tasks of authorship. For more information on copyright issues, see the US Copyright Office's guidance.

- **Know what the AI is doing with your data:** Be sure to understand what the AI tool is doing with any data you provide while using it. Many tools offer a selection of Service Level Agreements (SLAs) and Data Use Agreements (DUA). We recommend, when possible, opting out of tracking history and training usage.

Chat GPT's Opt Out option -- deselect "Chat history & training"
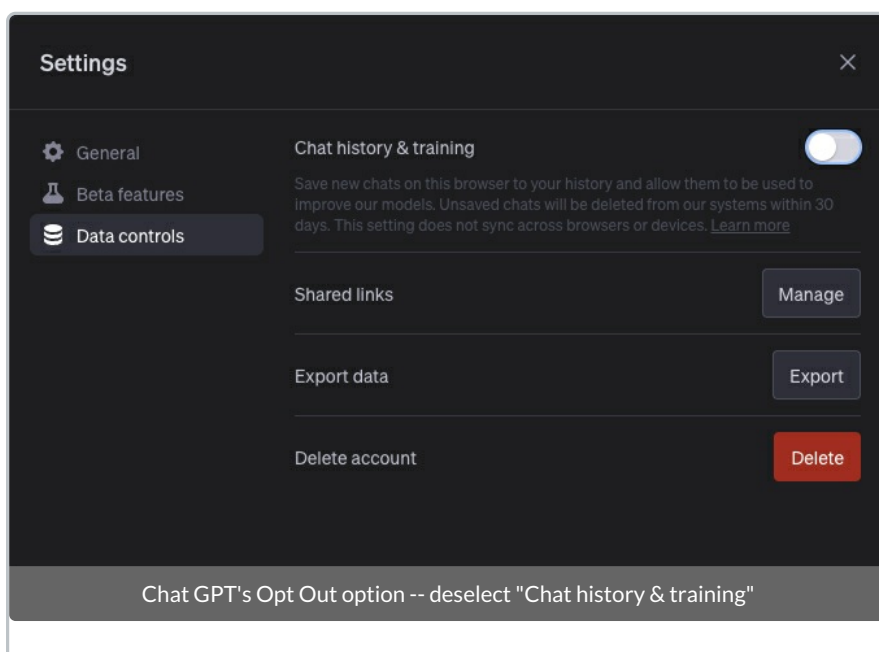
Security and Privacy Requirements

Submission of University Confidential Data to a generative AI tool is prohibited. Please read this section

carefully.

- **Examples of University Confidential Data Include (but are not limited to):**

Private Personal Information	Financial / University Policy Information	IP /Authorship/Research
○ Personal Identifiable Information (PII)	○ Payment Card Industry (PCI)	○ Unpublished research data (at data owner's discretion), subject to any IRB restrictions
○ Personal Health Information (e/PHI)	○ Financial Information	○ PENN-owned Intellectual Property (including code)
○ Biometric Data	○ Non-public Penn policies and documentation (e.g. system designs, budgets, plans, etc.)	○ Federal government research
○ Student, Faculty, and Staff Information or Records, including directory information.	○ System Credentials	
○ Any data covered by the University's FERPA policy	○ Data prohibited from sharing under terms of University-signed contract.	

(For a complete description of how the University treats confidential data, see University Data Classification.)

- **Interconnection to any University-managed system or application -- including the use of an API -- is prohibited.** Interconnection is the direct integration or connection between the AI tool and a system managed by the University. Managed system examples include (but are not limited to):

  - Endpoints (Wharton Computing-managed computers, printers and other devices).

  - Servers

  - Applications (on-premise and cloud-based)

  - Data Sources

  - IoT Devices

> **Remember**: Any information uploaded by faculty or staff may be both unrestricted and available to the AI tool (such as ChatGPT) and the AI tool's end users.

- **Terms of Use.** Review the Terms of Use and Privacy Policy of the application or service periodically. These documents may change frequently, as the technology, services, and other elements of AI evolve.

- **Penn Email Considerations**. A Penn/Wharton e-mail address for registration and use of AI tools is not prohibited. However, if a Penn/Wharton e-mail is selected, please ensure that the password is unique and different from any other Penn/Wharton applications, solutions or systems currently used. Additionally, e-mail may often be used to identify an individual, the type of information submitted, and other possibly private information.

  If you are using a Penn/Wharton email to register an AI account, we recommend using a team account rather than a personal account. This will preserve business continuity while accommodating situations such as vacations or departures from the University.

> For detailed information on the University's definitions of data sensitivity, see University of Pennsylvania Data Classification Standard

We encourage you to collaborate with Wharton Computing and the Wharton Information Security Office if you are considering any generative AI services for official use within Wharton (e.g. interconnections with Wharton Systems, Applications, or use of University Confidential Information). These groups will partner with you to provide the appropriate security risk evaluations, and help answer many of the IT support questions you may have.

Authorship

There is an ongoing debate regarding authorship and the differences between original work and machine-generated work. In addition, there are questions being raised about how machine-generated work should be used. The US Copyright Office provides helpful guidance regarding the distinction of authorship and ability to copyright works when generative AI tools are used. A few key points include:

- Machines cannot be authors

- All works developed by generative AI tools are not copyrightable.

- Authors may use generative AI tools to create original works, but certain conditions must be met:

  - generally, humans have to employ the "traditional methods of authorship" during creation of a work, which means making choices that are dependent upon the medium.

  - "How" the work is created is more important in determining originality than determining "what" created the work. This distinction can be difficult and we encourage you to seek consultation on this issue.

- Authors who use AI tools are expected to disclose that they have used an AI tool, preferably including which tool was used.

Faculty

While the information on this page is valuable for everyone, faculty may also want to review topics specific to teaching and research. We explore these in our faculty-focused article, AI Best Practices -- Teaching and Research (for Faculty). Please be sure you have logged in above (top right of screen) before launching this webpage. If you are not sure how to log in, see Logging into the Wharton Computing Knowledge Base.)

Staff

Depending on your role at the University, Wharton staff are likely to have widely varying needs for engaging with AI. To protect the University, staff should contact Wharton Computing and the Wharton Information Security Office prior to using any AI tool. Reasons for this level of concern include:

Read More →

Students

The General Information section above provides guidance around many questions regarding the use of AI in the course of your Wharton experience. Additional points to keep in mind include:

 Read More →

Additional Resources

- Penn ISC's AI Guidance
- Penn's Almanac: Guidance on Using Generative AI
- University's Policy on the Confidentiality of Student Records.
- Sources for Up-to-date information on generative AI in Education:
    - **Copyright topics**: the nonprofit Copyright Clearance Center's Velocity of Content is well-recognized (and not dominated by any one publisher or rightsholder.)
- University of Pennsylvania Data Classification Standard
- Chicago Manual of Style Online recommends various formal and informal approaches for citing AI contributions.

Questions?

Please contact your Wharton Computing representative with any questions regarding technology and the use of AI. For best results, please include answers to as many of these questions as possible when requesting help or guidance around AI:

- Do you need help with a new or existing project?
- Which generative AI tool(s) are you using?
- Who will be using the tool?
- What type(s) of data will be submitted and used?
- Who will be consuming the output?
- Will you need services such as prompt engineering, fine tuning, etc.?