Responsible use of Generative Artificial Intelligence in Research: Key Principles to Guide Institutional and Investigator Behavior

Purpose

Use of generative artificial intelligence (GenAI) can expand opportunity for research and lead to new developments. As we strive at this forefront, it is important to embark upon research with confidence in the safety and effectiveness of the approach. The purpose of this document is to share principles that are foundational to the integrity of research processes, and to guide research conduct and reporting when applying Generative AI. This guidance is relevant to creative works, and it also points to relevant policies in the sections below.

Research is the creation of new knowledge and information through exploration, investigation, and processing. Research involves a process of systematic inquiry including development, testing, analysis, interpretation or evaluation designed to contribute to generalizable knowledge or creative works when conducted according to prevailing methodological standards of the relevant academic disciplines. The life cycle of research includes stages such as: proposing and performing research (conception and execution), reporting (dissemination), and reviewing (consuming, translating, and funding including sponsor agreement compliance, etc.). Research can also provide a basis for creative works composed or performed.

Artificial intelligence includes systems or algorithms to predict, recommend, or advance defined objectives or tasks that have historically required a human.[1] Generative AI (GenAI) is a type of AI that can generate new content in response to prompts entered by the user. GenAI tools can generate text, images, audio, video, computer code, and other data types. Use of GenAI in research is a technical and a social act. GenAI tools and services[2] produce new capabilities and risks for research. University researchers are both legally and ethically obligated to limit risks to

[1] Source: National Artificial Intelligence Research Resource Task Force, 2023

[2] Examples currently include, but are not limited to, ChatGPT, Google Gemini, Microsoft CoPilot, ClaudeAI, Stable Diffusion, Midjourney, etc. This list is descriptive not prescriptive as it is a dynamically evolving market. For instance, GenAI tools are now embedded into other software packages routinely used in content creation, etc. In addition, many machine learning tools (e.g., image processing and recognition algorithms, data processing algorithms) are also used in the research process but have not attracted the same scrutiny as GenAI. The distinction between GenAI and more traditional machine learning algorithms is difficult to draw. We believe that many of the principles articulated above apply equally to other recent research practices using advanced computational or statistical models, and whether the tool is local on desktop CPUs, GPUs, or on external hardware systems for instance that are web-based.

the validity and integrity of their processes as they pursue valid, verifiable, reproducible, and original research, as expected by disciplinary standards.

This document was developed by a multidisciplinary team of faculty and research administrators as initial recommendations to encourage responsible use of GenAI to accelerate new research while protecting its integrity. Primary goals of research are to foster original thinking to build a body of knowledge while distributing costs/burdens effectively including protecting human subjects and enabling future inquiry. In concordance with principles developed from other research institutions,[3] the following guidelines are provided for use of GenAI in research by faculty, staff, and students or trainees.[4] Researchers are to consider the following principles as they explore the use of GenAI in developing research processes and products, while striving to maintain - and even improve upon - research integrity and standards relevant to their disciplines. This guidance also applies to research that is part of generating visual, compositional or performance-based creative works.

Principles

Three key duties of *data and information protection*, *verification*, and *transparency* frame appropriate use of GenAI in research.[5]

1) Protect Data and Information: Many popular GenAI tools collect and store information from users as part of their learning process. While some GenAI services offer a "no training use" policy as part of enterprise agreements, by default, researchers should first assume that any data or information entered into web-based GenAI could become part of its training data and thus be recoverable by third parties from future iterations of the model. Terms for the tool's use of information entered by the user should first be verified by the researcher in advance: do not assume that data entered is private, that output is already in public domain, nor that one can share information that is private, export-controlled, sensitive, or proprietary.

---

[3] https://research-and-innovation.cornell.edu/generative-ai-in-academic-research/, and https://it.wisc.edu/generative-ai-uw-madison-use-policies/

[4] For AI developers – there are additional considerations to incorporate in the design, development, use and evaluation of such products including environmental impact. This is beyond the scope of this document focused on the application of AI tools in the research context.

[5] These duties are derived from the Cornell Statement on Generative AI in Academic Research (https://research-and-innovation.cornell.edu/generative-ai-in-academic-research/). Beyond the short list above, researchers may also be obligated to follow the policies of external entities regarding GenAI, including funding agencies, publishers, suppliers, or other institutions where collaborating researchers reside. For instance, NIH and NSF prohibit the use of GenAI in peer review as a breach of confidentiality.

All members of the University community (faculty, staff, students and affiliates) wishing to enter institutional research data or other sources of information into GenAI must ensure that doing so will not breach any University policies, including: Code of conduct | HIPAA Policy | Student Records (FERPA) policy | HR Policy on Confidentiality | IP Policy | Data Stewardship Policy.[6] Attention should also be paid to the privacy components of the University Information Technology Policy. As a general guide, only data classified as low risk by the university Data Security Classification Security Policy should be considered suitable for entry to GenAI. The use of GenAI with medium or high-risk data (especially identifiable human data) should only be contemplated when the relevant tool has undergone appropriate internal review.

2) Verify: Researchers have a duty to verify the output of GenAI systems to ensure that they have not introduced inaccuracies or biases into the research process. GenAI can produce extremely convincing text, imagery, audio and video (and increasingly other types of data); but GenAI tools may not stay faithful to some ground truth or fairly represent the world.

Use of GenAI carries real risks regarding the reproducibility and reliability of research. This is particularly true in the context of research involving cultural practices or human diversity, where the risk of bias is high. The following are best practices for instances when the use of GenAI requires a higher degree of scrutiny:

a. *Inquiry as the basis of forming research questions:* verify the provenance of your resulting question(s) as the output of GenAI stems from others' content put in. To what degree have you invented the focus of the research question or hypotheses to test?
b. *Literature review*: verify to ensure that critical claims (i.e. effect size, statistical significance) from individual articles or sub-literatures are accurately reported. Do not rely on GenAI to represent the range of works available on a topic.
c. *Output of analysis and code*: in concordance with disciplinary standards, verify reproducibility of results, and ensure authorial control of your original content. Also, verify to ensure that generated code is human interpretable and performs the intended function.
d. *Interpretation* is an act of authorship: consider your discipline's standards when GenAI is used to develop interpretations, and be aware that you are responsible for the integrity of the interpretation (e.g., how theory connects to the data, limitations for discussion, etc.).
e. *Attribution*: verify to ensure that data sources and citations exist, and are correctly attributed. The text produced by GenAI tools may not include any attribution of authorship, and if attribution is included, it may attribute incorrectly. Content produced by GenAI tools is often based on copyrighted, web-scraped or publicly available material.

[6] A proposed Data Stewardship policy is currently in review for adoption by the UR Board of Trustees. Link to be updated as available.

In all cases, the author using the GenAI tool is responsible for correctly attributing prior work. Authors who fail to do this risk plagiarizing prior work, and/or infringing copyright.

f. *Research involving human meaning or difference*: verify to ensure that outputs which report on or represent human diversity (incl. biological, social and humanistic research) are accurate.
g. *Composing and editing or revising products for dissemination*: verify to ensure data summaries, or the information or knowledge generated (i.e., in creative works derived from the research inquiry) - including quotations, citations, author meaning or descriptions of methods - have not been altered.

3) Be Transparent: Consistent with disciplinary and publishing norms, the use of GenAI should be disclosed/identified at the time of proposing, reporting, publishing or when engaging in other dissemination activities. In reporting such use, the following existing policies should be adhered to: Research Misconduct Policy | Intellectual Property Policy. While norms around the reporting of GenAI use are evolving, we urge researchers to err on the side of acknowledging and reporting *all* use of GenAI tools in the research process, including use in idea development, assistance with composition, code assistance, or generation of presentations and figures. For instance, NSF currently recommends attribution in the context of proposing research: "Proposers are encouraged to indicate in the project description the extent to which, if any, generative AI technology was used and how it was used to develop their proposal."

If GenAI is used to conduct research that will result in a creative work with a registered copyright (e.g., exhibit, multimedia, music, script, book, dance, sculpture or other visual art form), then extra care should be taken. While the norms around authorship are evolving, if the *final work* contains more than *de minimis* AI-generated content as defined by your field, the nature and use of GenAI should be disclosed, and the parts of the work created by the author and by the GenAI should be distinguished and described.

Moreover, researchers should adequately acknowledge the use of GenAI in the research design. GenAI poses significant risks to the reproducibility of research processes, because of their inherently stochastic nature. The specific GenAI tool should be documented along with the parameters used (including prompts, context windows, and version numbers), for instance, when using GenAI to generate synthetic datasets. Use of GenAI systems where the model is freely available can aid reproducibility. Protecting provenance – that is, knowledge of the “ownership” history – or the contextual narrative of how information has been created through research – also maintains value in and authenticity of research.

Human-centered Research and Other Resources

As an individual or team-based user of GenAI, these principles guide us to reflect on such critical social concerns by asking broad questions. E.g.,:

- What is my role as a user of AI? What is my role as an AI developer? What is AIs role in my work?

- Do I understand how the tool works?
- Do I know the source of the data used to train the tool's model?
- Have I identified the relevant technical and human factors expertise needed when using GenAI tools and resources?
- Am I using the lowest risk tool from a socio-ethical-technical perspective? How is this being evaluated?

Researchers should also be aware of broader considerations that emerge from the social nature of research. We encourage the use of GenAI in alignment with our institution's broader goals—to foster creativity and expand human capacity. As these tools are deployed within research, we ask that users consider how GenAI can enhance learning and mentorship in AI literacy, support broad and meaningful access, and promote responsible development and use. As with any research capacity or tool – it is an aid; we must build our literacy to critically assess tools and their outputs. Stay up to date on academic as well as legal and policy developments including AI regulation, and choose trustworthy tools and services (e.g. the National Institute of Standards and Technology's (NIST's) trustworthy AI).

We should also consider the human and social impacts of GenAI use (e.g., human labor displacement, wellbeing, environmental protection, maintenance of content creators' rights and use in accordance with copyright policies). We thus encourage use of our institution's resources and risk management to evaluate use of such tools. For instance, UR researchers can consult legal services, chief security officers, etc. to review a tool supplier's terms of use, privacy policy, and other tool documentation before using any system to process sensitive or proprietary information in research (see URMC Office of Integrity and Compliance; URMC Privacy Statement; University IT Policies; Office of Human Subject Protection; Office of Counsel; Office of the Vice President for Research – Integrity and Compliance).

The practice of integrating GenAI will evolve dynamically given the rapid pace of discovery and of application development. We cannot yet know boundaries of use derived from either human or technical capacity. Since we will continue to innovate using GenAI, we must share knowledge about GenAI use in research. Such sharing will support adaptation across our varied research endeavors. Consequently, we welcome information about use cases from all disciplines to evolve this guidance.

To learn more about the related considerations on use of GenAI when engaged in complementary activities of education, IT system use, communication and marketing, or for more information about the AI Council and RPC, see the University of Rochester AI website.

Publication Date: December 13, 2024
AI in Research Committee (authors):
Ann Marie White
Jon Herington
Chris Kanan
Adam Purtee

Deb Rossen-Knill
Chris Dasilva
David Mathews
Andrew White
Ehsan Hoque
Erika Ramsdale
Yasemin Basdogan
Raffaella Borasi
Kristen Love
Chris Deeney
Ajay Anand
Brendan Mort
Rachel Remmel
Paul Brookes
Caitlin Dreisbach
Richard Libby
Yvonne Watson
Libby Reitz