Generative AI Usage Toolkit

Introduction

Welcome to the NIH Library's GenAI Usage Toolkit, a comprehensive guide designed to help you effectively and ethically integrate generative AI tools into your workflows. This toolkit provides a structured approach to evaluating, implementing, and verifying the use of AI in various tasks, ensuring that you can harness the power of AI while maintaining high standards of accuracy, transparency, and accountability.

This toolkit is designed to be a practical resource for researchers, professionals, and teams seeking to leverage AI tools responsibly. By following the guidelines and using the provided worksheets and checklists, you can make informed decisions about AI use, enhance the quality of your work, and uphold ethical standards in your AI applications.

We hope you find this toolkit valuable and empowering as you explore the potential of generative AI in your work.

This is the fourth edition of the NIH Library Generative AI Usage Toolkit, updated in November 2025 to reflect new NIH-supported platforms and expanded verification guidance.

Table of Contents

Executive Summary **2**

AI Policies at NIH and IC-Specific Guidance **3**

Guide to LLMs and Related Tools for NIH Staff **4**

Task Appropriateness Checklist for AI Use **5**

AI Application Assessment Matrix *(Example)* **6**

AI Application Assessment Matrix *(For Your Personal Use)* **7**

AI Output Verification Methods Worksheet **8**

Generative AI Usage Documentation Form **9**

Attribution Requirement Checklist **10**

Drafting an Attribution Statement **11**

Disclaimer **12**

Executive Summary

The NIH Library Generative AI Usage Toolkit is a practical resource designed to help NIH staff and affiliates make thoughtful, responsible use of generative AI tools. This updated edition reflects the evolving AI landscape and incorporates clearer guidance on secure NIH-supported platforms, data privacy, and verification practices.

Key components include:

- **NIH and IC-Specific Guidance** – Overview of policies and links to Institute and Center resources.
- **Guide to LLMs and Related Tools** – Quick-reference to secure NIH platforms (e.g., CHIRP, ChatGPT Enterprise, GENESIS, Copilot Chat) and appropriate use of public-facing tools.
- **Task Appropriateness Checklist** – Framework for identifying when AI is suitable in NIH contexts.
- **AI Application Assessment Matrix** – Guidance for evaluating applications by risk, boundaries, transparency, and monitoring.
- **AI Output Verification** – A tiered framework for scaling human oversight and a menu of verification methods.
- **Documentation and Attribution** – Tools for recording AI use and drafting attribution statements when needed.

The toolkit is intended to:

- Support NIH staff in identifying appropriate and ethical AI use cases.
- Encourage consistent human oversight and verification of AI outputs.
- Promote transparency and accuracy in research, administrative, and communication workflows.
- Provide adaptable worksheets and checklists to simplify decision-making.

This is a **living resource**. It does not replace IC-specific policies or HHS guidance but is meant to be used alongside them. By following these guidelines, NIH staff can take advantage of AI's potential while safeguarding scientific integrity and protecting sensitive information.

AI Policies at NIH and IC-Specific Guidance

NIH Artificial Intelligence (AI) Cybersecurity Guidance

- This document is intended to provide provisional guidance for the usage of external (public) and HHS/NIH approved generative AI services for NIH information. Moreover, this guidance is aimed to promote the responsible use of the technology and ensure its risks are fully understood while enabling NIH to benefit from these advanced capabilities.

AI Guidance - Foundational Information, Generative AI, and Risks

- Provides foundational information on AI, including definitions, key concepts, and the scope of AI technologies.
- Addresses specific risks associated with generative AI, such as data privacy concerns, potential biases, and ethical considerations.
- Offers recommendations for mitigating these risks.

The Use of Generative Artificial Intelligence Technologies is Prohibited for the NIH Peer Review Process

- Explicitly prohibits the use of generative AI technologies in the NIH peer review process.
- Ensures the integrity and objectivity of peer review by preventing potential biases or manipulations from AI-generated content.

Artificial Intelligence in Research: Policy Considerations and Guidance

- A new resource issued by the Office of Science Policy (OSP) to help the research community understand how NIH policies guide AI-related research.
- Illustrates the applicability of existing policies and guidance to research involving AI technologies.

IC-Specific Guidance

While this toolkit offers general guidance on the use of generative AI, each NIH Institute and Center (IC) may have its own specific policies and guidelines. These IC-specific policies may be more restrictive or include additional requirements beyond what is outlined in this general toolkit.

Guide to LLMs and Related Tools for NIH Staff

Data Privacy Notice

Follow all NIH policies on privacy and security. Do not enter PII, PHI, or sensitive data into public tools. Use only NIH-secure platforms when handling sensitive information.

Secure NIH Options: Use these for NIH-sensitive or internal work:

- CHIRP (NIH)
  Secure chatbot with GPT-4o, Claude, DALL·E
  • NIH VPN required • Limited seats • Safe for de-identified data
- ChatGPT Enterprise (NIH license)
  Enterprise-grade ChatGPT with gov email login
  • Research summaries • Structured analysis • Scientific writing
- Copilot Chat (Microsoft – coming soon)
  Dedicated secure chatbot interface (not Office Copilot)
  • For Q&A and general tasks using NIH data
- GENESIS (NIAID)
  Suite of GenAI tools (chat, summarizer, multi-doc analysis)
  • Designed for NIH research workflows

Major Public LLMs: Use only for non-sensitive, exploratory tasks (brainstorming, public-facing content):

- ChatGPT (OpenAI) – Summaries, patient education, lit reviews → chatgpt.com
- Gemini (Google) – Multimodal: text, images, code; diagnostic/training uses → gemini.google.com/app
- Claude (Anthropic) – Conversational support, admin automation, research ethics → claude.ai
- Mistral – Efficient LLMs for data analysis, code generation, text summarization → mistral.ai

Other Tools

- Elicit – Research paper summaries → elicit.com
- Perplexity – Answer engine with citations → perplexity.ai
- NIH STRIDES Initiative – Cloud-hosted AI/LLM resources → cloud.nih.gov.

Task Appropriateness Checklist for AI Use

Introduction: *This worksheet will help you assess which tasks in your workflow are suitable for the application of generative AI tools. By identifying appropriate use cases, you can ensure that AI is utilized effectively and ethically within your specific research or work context. Use this checklist to evaluate tasks and make informed decisions about integrating AI into your processes.*

	Circle one	
1. Is the task primarily one of the following? ○ Information retrieval or summarization ○ Data analysis or pattern recognition ○ Content generation or ideation ○ Language translation or paraphrasing	Yes	No

STOP! If you selected No, generative AI might not be suitable for this task.

2. Does the task require any of the following? ○ Deep subject matter expertise ○ Critical analysis of recent or specialized research ○ Ethical decision-making	Yes	No

If you selected yes, ensure human expertise leads the process.

3. Will the output be used for any of the following? ○ Public publication or sharing ○ Critical decision-making ○ Incorporation into official reports or papers	Yes	No

If you selected yes, implement a robust review process.

4. Can you clearly define the inputs and desired outputs for this task?	Yes	No
5. Do you have the means to verify the AI's output?	Yes	No
6. Is using AI for this task aligned with your team's or department's policies?	Yes	No

If you selected No on any of the above, address these issues before proceeding.

AI Application Assessment Matrix *(Example)*

Instructions: Use this worksheet to identify potential applications of generative AI within your workflows and establish guidelines for responsible implementation based on the considerations outlined in the columns.

Application *Specific task or process where generative AI could be applied.*	**Risk Level** *Potential risks based on data sensitivity, impact, and error potential.*	**Boundaries** *Clear limits and guidelines for AI use, including human oversight requirements.*	**Transparency & Attribution** *How AI use will be disclosed and if/how AI-generated content will be attributed.*	**Monitoring** *Processes for overseeing AI use, including quality control and performance evaluation.*
Drafting emails	Low	- Do not include confidential or personal information - Review content before sending	- No need to be transparent about AI usage - No attribution required	- Carefully review generated content before sending
Literature reviews	Medium	- Summarize key points, do not reproduce full text - Verify accuracy of AI-generated insights	- Disclose use of generative AI in the process - Attribute AI-generated summaries	- Review AI-assisted literature reviews for accuracy and completeness
Ideation and brainstorming	Low	- Use generative AI for initial idea generation only - Maintain human oversight and control	- No need for transparency about AI usage - No attribution required	- Ensure human review and refinement of AI-generated ideas
Report writing	High	- Avoid using AI to generate entire reports - Limit AI to specific sections (e.g., background, data summaries)	- Clearly indicate which sections were AI-generated - Attribute AI contributions accordingly	- Thoroughly review AI-assisted report sections for accuracy, completeness, and alignment with research goals
Data analysis	High	- Restrict AI to exploratory data analysis only - Ensure human validation of findings	- Disclose use of generative AI in the analysis process - Attribute any AI-generated insights or visualizations	- Implement rigorous checks and validations of AI-assisted data analysis

AI Application Assessment Matrix *(For Your Personal Use)*

Instructions: Use this worksheet to identify potential applications of generative AI within your workflows and establish guidelines for responsible implementation based on the considerations outlined in the columns.

Application *Specific task or process where generative AI could be applied.*	**Risk Level** *Potential risks based on data sensitivity, impact, and error potential.*	**Boundaries** *Clear limits and guidelines for AI use, including human oversight requirements.*	**Transparency & Attribution** *How AI use will be disclosed and if/how AI-generated content will be attributed.*	**Monitoring** *Processes for overseeing AI use, including quality control and performance evaluation.*

AI Output Verification

Introduction: Verification is essential to ensure that generative AI outputs meet NIH's standards of accuracy, reliability, and ethical integrity. Human judgment should always play a role in reviewing AI-assisted work. The level of scrutiny can vary depending on the context: a quick summary may only need a light check, while materials for reports or publications benefit from more formal review. This section outlines a tiered framework for scaling verification, followed by a menu of methods to choose from.

Tiered Verification Framework

Low-Risk Tasks (drafts, brainstorming, internal notes)

- Include a human check for clarity, correctness, and usefulness.
- Light review is generally sufficient.

Medium-Risk Tasks (literature reviews, educational materials, meeting summaries)

- Involve a peer or subject matter expert where possible.
- Consider cross-referencing outputs with trusted sources (e.g., PubMed).
- A short note of how the output was reviewed can be helpful.

High-Risk Tasks (publications, official reports, policy/funding analyses, sensitive/clinical outputs)

- Strongly benefit from multiple forms of verification (SME review, cross-referencing, replication/testing).
- Keeping fuller notes on the verification process helps with transparency.
- If AI contributed substantially, consider including an attribution statement.

Verification Methods Menu

Depending on the level of risk, you may combine one or more of these approaches:

- Expert Review – Ask a peer, supervisor, or SME to evaluate accuracy and nuance.
- Cross-Referencing – Check against authoritative sources like PubMed or NIH databases.
- Comparative Analysis – Compare with human-generated work or established benchmarks.
- Empirical Testing – Test AI predictions or suggestions where feasible.
- Replication of Results – Run the AI multiple times to assess consistency.
- Consistency Check – Ensure alignment with established facts or prior research.
- Logical/Mathematical Proof – Validate reasoning or calculations formally (if applicable).
- Real-World Application/Observation – Pilot outputs in practice and monitor outcomes.
- Other Methods – Note any additional strategies used.

Generative AI Usage Documentation Form

1. Project/Task Name:_____
2. Date of AI Use: _____
3. AI Tool Used: _____
4. Version/Model of AI Tool (if known): _____
5. Interaction Environment:

   ☐ Public web interface
   ☐ Secure enterprise application
   ☐ Other (please specify):
   ☐ Institutional cloud instance
   ☐ Local installation

6. Purpose of AI Use:

   ☐ Information retrieval
   ☐ Content generation
   ☐ Other (please specify):
   ☐ Data analysis
   ☐ Language translation

7. Brief Description of Task: _____
8. Prompt(s) Used: _____
9. AI-Generated Output (summary or key points):_____
10. Human Modifications/Contributions:

    ☐ Prompt engineering/refinement
    ☐ Fact-checking and correction
    ☐ Content expansion or elaboration
    ☐ Reorganization of AI-generated content
    ☐ Stylistic edits for tone or voice
    ☐ Integration with human-generated content
    ☐ Critical analysis or interpretation of AI output
    ☐ Removal of irrelevant or incorrect information
    ☐ Addition of domain-specific knowledge
    ☐ Other:

11. Verification Method:

    ☐ Expert review
    ☐ Experimental validation
    ☐ Cross-reference with other sources
    ☐ Other (please specify):

12. Additional Notes: _____
13. Documented by: _____
14. Supervisor Approval (if required): _____

Attribution Requirement Checklist

Proper attribution of AI-generated content is important for maintaining transparency and ethical standards. This checklist will help you determine when and how to attribute the use of generative AI tools in your work. Use this worksheet to ensure that AI contributions are clearly acknowledged and that your work complies with relevant policies and guidelines.

If you select yes for any of these questions, attribution of AI use is likely required or strongly recommended.

		Circle one	
1.	Is the AI-generated content being used in a published work (e.g., research paper, report, article)?	Yes	No
2.	Does the AI contribution constitute a significant portion of the work?	Yes	No
3.	Has the AI-generated content been used with minimal or no modification?	Yes	No
	If you selected yes for any of the above, attribution is generally necessary for academic integrity.		
4.	Is the work being submitted to a journal or publisher that specifically requires AI attribution?	Yes	No
5.	Does your institution's policy require attribution of AI use?	Yes	No
	If you selected yes for any of the above, attribution is required to comply with specific policies.		
6.	Is the AI-generated content being used in a way that could impact scientific conclusions or interpretations?	Yes	No
7.	Are you using the AI-generated content in a teaching or presentation context where disclosing AI use would be relevant?	Yes	No
8.	Could failure to disclose AI use be considered a form of plagiarism or academic dishonesty in this context? Would concealing AI use undermine the trustworthiness or credibility of your research?	Yes	No
	If you selected yes to the above, attribution is advisable for transparency and ethical considerations.		

Drafting an Attribution Statement

Creating clear and accurate attribution statements is a key aspect of responsible AI usage. This worksheet offers a formula and examples for drafting effective attribution statements, ensuring that the use of generative AI tools is transparently communicated. Use this worksheet to craft your attribution statements, providing specific details about the AI tool, its purpose, and the extent of its contribution to your work.

Attribution Formula

[AI Tool Name] version [Version Number] was used by [the User] to [Purpose/Task] in this [Type of Work]. The output was [Level of Modification] by the authors. [Additional Details if necessary].

Using this formula, here is an example attribution statement:

"ChatGPT version 3.5, an AI language model developed by OpenAI, was used by the author to assist in brainstorming ideas and generating initial drafts for sections of this report. The AI-generated content was subsequently reviewed, edited, and expanded upon by the authors to ensure accuracy and relevance to the research context."

Here are a few more examples to show how this might vary:

1. For minimal use: "DALL-E 2 was used to generate the initial concept sketch for Figure 3. The image was then refined and finalized by the research team."
2. For more extensive use: "GPT-4 was employed to assist in the literature review process, helping to summarize key points from multiple sources. All AI-generated summaries were critically reviewed, fact-checked, and integrated into the final manuscript by the authors."
3. For use in data analysis: "A custom large language model, based on the GPT architecture, was used to analyze patterns in the textual data. The model's outputs were interpreted and validated by the research team, with all conclusions drawn by human experts."

When crafting an attribution statement, consider including:

1. The specific AI tool used
2. The version or model of the AI
3. The purpose or task for which the AI was used
4. The extent of AI contribution
5. The level of human oversight and modification
6. Any limitations or potential biases of the AI tool, if relevant



Disclaimer

This toolkit is intended as a general resource for NIH staff and affiliates. However, it does not supersede or replace any policies or guidelines set forth by individual NIH Institutes and Centers (ICs) or the Department of Health and Human Services (HHS). Users of this toolkit should always consult and adhere to the specific AI usage policies and guidelines of their respective IC and any applicable HHS policies.

Before implementing any AI tools or processes in your work, ensure that you are familiar with and following the most up-to-date guidance from your specific IC and HHS. For the latest information on AI policies and guidelines specific to your IC, please consult your IC's official communications channels or contact your IC's designated AI policy representative.

Last Updated: November 17, 2025