< Back to We, the People

NARA Notice 2024-015: Responsible Use of Artificial Intelligence

Published on October 25, 2023
NARA News - NARA Notice - Information Services - Cyber Security & Information Assurance

Notice

To: All Employees

Attention supervisors: If you have employees who do not have access to a computer, please ensure that those employees receive a copy of this notice. This includes employees on LWOP or paid leave.

On May 4, 2023, the **Biden Administration announced new actions to promote responsible use of Artificial Intelligence (AI)**. These actions build on the administration's work to promote responsible AI innovation, including:

- the **Blueprint for an AI Bill of Rights**,
- the **AI Risk Management Framework**, and
- a roadmap for setting up a **National AI Research Resource**.

Additionally, in February, President Biden signed an **Executive Order** that directs federal agencies to root out bias in their design and use of new technologies, including AI, and to protect the public from algorithmic discrimination.

ChatGPT, Google's Bard, and other AI-enabled Large Language Models (LLMs) tools can help you with a variety of tasks, but it's important to use them safely and responsibly. LLMs are a type of artificial intelligence models that have been trained on vast quantities of text data to create human-like responses to dialogue or other natural language inputs. Chatbots like ChatGPT are trained to make predictions about "what comes next" after a written prompt. They use vocabulary and information while also understanding words in context. This helps these tools to mimic speech patterns while displaying knowledge that the models have learned.

LLMs are highly capable and are expected to rapidly improve in capabilities over the coming months and years. However, these tools are new and under development, and should always be used with caution.

NARA information should **never** be used with chatbots or other online AI applications. ChatGPT and other similar tools may produce results with biases, including racial and gender biases, in their responses. They may generate false or misleading responses that may negatively impact the quality of work performed or assisted by AI tools. Importantly, AI-enabled tools incorporate the inputs and responses of previous user queries into their responses to queries from other users. This means that information provided by NARA users will be incorporated into other responses, which may provide unrelated users with biased a otherwise inaccurate responses about NARA or using NARA information.

Here are some additional things to keep in mind when using LLMs:

- **Don't share personal or sensitive information with an AI chatbot**. These tools are publicly accessible, so anyone can potentially be presented with the information you're sharing.
- **Be aware of the biases that may be present in LLM responses**. These tools are trained on large amounts of data, which may contain biases that include racial, gender, and other social practices that may cause harm if not recognized. For example, if a chatbot is trained on a dataset of text that is mostly written by men, it may be more likely to generate responses that are biased towards men.
- **Do not rely on LLMs for factual information**. These tools are not always accurate and may generate false or misleading information. It's always best to fact-check information before you use it.
- **Be aware of LLM limitations**. These tools are still under development and don't have the same knowledge and understanding as a human being. For example, ChatGPT may not understand context or complex concepts, or provide nuanced answers, especially about recent events.
    - ChatGPT in particular, is trained on a large amount of internet data that likely includes copyrighted material. Therefore, its outputs have the potential to violate copyright or other intellectual property protections.
    - Applications that use LLM models, including ChatGPT, are also susceptible to “prompt injection,” a hacking technique in which malicious adversarial prompts are used to trick the model into performing tasks that it wasn’t intended for, such as writing malware code or developing phishing sites that resemble well-known sites.

If you have any questions about the security aspects of AI, contact the Cyber Security & Information Assurance Division at **information.assurance@nara.gov**.

SHEENA BURRELL

Executive for Information Services, Chief Information Officer

---

If you have questions about this notice, contact:

Keith Day, Cyber Security and Information Assurance Division Information Services

keith.day@nara.gov

Room:

National Archives at College Park

Phone: