Security Policy for Generative Artificial Intelligence (AI) Large Language Models (LLMs)

Number: CIO IL-23-01
Status: Active
Signature Date: 06/09/2023
Expiration Date: 06/30/2024

Full Directive PDF

1. **Purpose**. This Instructional Letter (IL) provides an interim policy for controlled access to Generative Artificial Intelligence (AI) Large Language Models (LLMs) from the GSA network and Government Furnished Equipment (GFE).

2. **Background**. Generative AI is a type of AI system that uses LLMs for generating natural language text. These models are trained on vast amounts of textual data scraped broadly from the internet or from specific focused data sets. They are able to generate new text that is similar in style and content to the data it was trained on, generating summaries, translations, predictive text, and other content based on the patterns learned from its training. Well-known examples include OpenAI ChatGPT, Google BARD, and Salesforce Einstein. While this disruptive technology has many promising applications, it is not without risk. As LLMs train on public data sources and information inputted to them, it is possible for federal data to be leaked to a platform that is not authorized to house or transact it. Further, inputted data can train the publicly available LLM model, and reveal it later to others.

3. **Scope and Applicability**. This instructional letter applies to all GSA employees and contractors.

a. This IL also applies to the Office of Inspector General (OIG) to the extent that the OIG determines it is consistent with the OIG's independent authority under the Inspector General Act of 1978 (5 U.S.C. App. 3) and does not conflict with other OIG policies or the OIG mission.

b. This IL applies to the Civilian Board of Contract Appeals (CBCA) only to the extent that it is consistent with the CBCA's requisite independence as defined by the Contract Disputes Act (CDA) and its legislative history. 41 U.S.C. §§ 7101-7109 (2012) and S. Rep. No. 95-1118 (1978).

4. **Policy**. This IL provides policies for the responsible use of Generative AI LLMs.

a. Access to publicly available, third party Generative AI LLM endpoints and tools shall be blocked from the GSA network and GFE devices.

(1) Exceptions will be made for research (relevant to the role, position, or organization of the requestor) and non-sensitive uses involving data inputs already in the public domain or generalized queries.

(2) Exceptions require completing this request form detailing intended usage and acknowledgment of GSA's IT General Rules of Behavior to not expose internal federal data.

(3) The usage of personal devices for government work is prohibited; exceptions are made for Bring Your Own Devices (BYOD) that comport with GSA's IT General Rules of Behavior and Rules of Behavior for Mobile Devices.

b. Federal nonpublic information (including work products, emails, and conversations that are meant to be pre-decisional or internal to GSA), such as financial disclosure information, protected acquisition, controlled unclassified information (CUI),

personally identifiable information (PII), and Business Identifiable Information(BII), shall not be disclosed as inputs in LLM prompts as they are not authorized to process or store such data. Such information may be used to expand the training data set and could result in an unintentional breach.

c. Deployment or use of locally deployed LLMs, such as Alpaca or Open-LLaMA, on GFE shall abide by GSA Information Technology (IT) Standards Profile, CIO 2160.1F CHGE 2.

d. GSA-deployed and managed LLMs shall be assessed and authorized to operate by GSA and require specific authorization to handle PII; have privileged access to GSA systems; or transfer data to systems that are not authorized to operate by the GSA. GSA-authorized information systems are listed in GSA EA Analytics & Reporting (GEAR).

e. LLMs for the purpose of generating software code shall:

(1) Be limited to publicly available code only; code in GSA private repositories shall not be inputted into publicly available LLMs.

(2) Not have access to nonpublic, sensitive or proprietary GSA code, unless there is an explicit policy from the LLM's owner stating that code will never be retained for nor used to train LLMs.

f. The output from LLMs used to generate code or publishable material shall be manually reviewed by the approved user for accuracy, functional effectiveness and suitability, and intellectual property, including copyrights and trademarks, terms of service, or end-user license agreements, as LLMs may have been trained on data that AI providers may not have had full legal rights to use.

e. LLMs shall not be used to generate malicious, inappropriate, or illegal material.

f. GSA performs electronic monitoring of internet communications traffic including publicly available LLMs.

g. All inquiries about this policy may be directed to the Security Operations Division (ISO) at secops-fw@gsa.gov.

5. **Responsibilities**.

a. The Office of the Chief Information Security Officer (OCISO). OCISO in GSA IT shall implement appropriate network and device level blocks to facilitate controlled access to Generative AI LLM solutions; develop a review and approval framework allowing access to individuals and groups in agreement with the policy provisions in this IL; and update annual security and privacy awareness training material on the appropriate use of this new technology.

b. The Office of Digital Management (IDR). IDR in GSA IT shall develop communications and training material on the provisions of this IL and the appropriate use of this new technology.