Generative AI Platform Usage and Data Security Policy

Last updated: 2023-XX-XX

Policy Purpose

This document outlines required practices and prohibitions that apply to the use of generative AI platforms such as ChatGPT, Google Bard, or other similar AI platforms using “large language models” (LLMs) for work tasks of any kind.

The purpose of the document and the practices and prohibitions outlined herein is to ensure that company private, personal, proprietary, customer, or other sensitive data is not intentionally or inadvertently transmitted to such platforms or related third parties in violation of data and compliance controls to which the company and its employees must adhere.

Policy Scope

All employees of the company share in the responsibility to ensure the continued health, safety, profitability, and growth of the company. Company data of all kinds is a key asset and enabler of these goals. Therefore all employees are expected to adhere to the company Generative AI Platform Usage and Data Security Policy, to encourage and support other employees in doing so, and to promptly report violations of the policy to the office of information technology (the office of the CIO).

Authorization

a. Employees must obtain explicit, documented authorization from their respective managers and/or authorized IT managers in the office of information technology before using any generative AI platforms for work-related tasks or for tasks of any kind, whether work-related or not work-related, while using or connected to work-related assets or systems.

b. This requirement expressly applies even in cases in which the employee in question does not believe that they have access to sensitive data or does not believe that the AI use in question presents a risk to the company.

c. Authorization must and will be granted on a narrowest-possible-purpose basis and with appropriate limitations and stipulations that follow from the nature of the task, the role of the requesting employee, and the sensitivity of the data to which the employee has regular access.

d. The employee may be required to complete training on company data classification standards, standards of data use and disclosure, and/or appropriate use of generative AI platforms before authorization is given.

e. Authorization may be withdrawn at any time by an employee’s manager or by IT managers in the office of information technology.

Data Classification

a. Employees using generative AI systems must clearly understand company data classification standards related to data type and sensitivity level (e.g., public, internal, confidential, customer, proprietary) before using generative AI platforms on work time or while using or connected to company devices and systems.

b. During AI platform use or prompting, the employee in question must adhere to guidance for appropriate applications and uses of data in relation to these standards.

c. Unless otherwise documented in writing by the employee's manager, the use of generative AI platforms is regarded by the company to be the equivalent of public disclosure of any prompt data, i.e. any data used, included, or implied in AI prompting will be regarded as having been disclosed to the public by the employee in question. Any outcomes, additional processes, or consequences related to public disclosure of data therefore apply in the case of generative AI use.

Prompt Selection

a. Employees must exercise due caution when creating prompts for generative AI platforms and ensure that prompts do not violate data classification and disclosure standards and practices.

b. Controlled or sensitive data may be replaced with anonymized data to enable an otherwise prohibited prompt only in cases in which the resulting prompt does not in any way suggest, imply, or implicitly convey either contents of the data in question or other company proprietary data, trade secrets, confidential information, customer information, etc.

Data Leakage Prevention

a. Employees must take ongoing affirmative steps to prevent the inadvertent release of proprietary or private information through generative AI platforms. These steps may include, but are not limited to, requesting regular peer or manager review of prospective prompts or of recent prompt history to ensure sound prompting practices.

b. Employees whose primary roles involve regular interaction with sensitive data may be required to complete additional steps or adhere to additional requirements in order to receive authorization to use generative AI.

c. Neither AI prompts nor AI-generated results may be shared beyond the employee's department without prior documented manager review and approval.

d. Employees authorized to use generative AI agree to periodic review of their complete AI prompt and result history, which may occur at any time, with or without the employee's knowledge.

User Authentication and Access Control

a. Employees authorized to use generative AI must do so in their own generative AI platform account associated with their company login or email address and may not use other platform accounts on company time or while using or connected to work-related assets or systems.

b. The sharing of generative AI accounts between two or more employees is expressly forbidden unless prior documented approval and an access plan have been received from the office of information technology.

c. Access to and maintenance of generative AI accounts must adhere to all standard user authentication, access control, and other information systems policies and procedures that apply to company employees.

Monitoring and Auditing

a. The employee understands that review of all generative AI access, prompting activity, and all returned results will regularly occur, and may occur at any time with or without the employee's prior knowledge.

b. The employee understands that all AI prompting activity and returned results remain property of the company to the extent allowed by controlling platform licensure. The employee has no expectation that any prompting activity or returned results, either in part or in whole, are the employee's property.

c. The employee agrees not to take measures to delete, clean, curate, or "sanitize" their prompt and result history for any purpose.

Management and Awareness

a. Employees with authorization to use generative AI will receive regular generative AI data safety training to ensure that awareness is maintained. This includes regular training beyond "initial" training prior to use.

b. All regular training must be company-provided or company-approved and completed promptly, with documentary evidence of completion provided to the employee's manager.

c. Managers must ensure that employees with authorization to use generative AI receive such trainings and that records of training dates and outcomes are strictly maintained.

d. Managers of employees with authorization to use generative AI understand that they are responsible for ensuring that a complete prompt and result history for every such employee is maintained. Such a history may be maintained by the platform in question or by other processes established in cooperation between manager and employee.

e. Managers of employees with authorization to use generative AI agree to conduct regular reviews of employee prompt and result history in order to maintain awareness of generative AI activity and employee understanding of, and adherence to, this policy.

f. Managers of employees with authorization to use generative AI agree to maintain auditable documentation of employee access, training, practices, and violations and understand that this documentation may be reviewed at any time with or without the manager's prior knowledge.

Policy Violations

a. Violation of the guidelines and requirements outlined in this document may result in loss of authorization to use generative AI for work purposes or further disciplinary action, up to and including immediate termination of employment.

b. Any known or suspected violation of the guidelines and requirements outlined in this document, even if inadvertent, must be reported immediately to the employee's manager and the office of information technology.

c. Any known or suspect instance of data "leakage," misuse, or conveyance to generative AI platforms in violation of company data classification standards, or company standards of data use and disclosure, must be reported immediately to the employee's manager and the office of information technology.

d. Failure to report a known violation or misuse in any of the above cases, or any affirmative steps taken to obscure such a violation or misuse, will subject the employee to immediate termination and possible further liability for any resulting consequences or outcomes.