SDCRAA HUMAN RESOURCES STANDARDS AND PROCEDURES

Section:	**Workplace Practices**
Standard:	**Artificial Intelligence Tools Usage Standard**
Section #:	D-21
Effective:	5/1/2024

See Also: Workplace Privacy; Personal Use of Authority Property; Misconduct; Formal Discipline; Use of Authority-Provided Computer Equipment Off-site; Use of Authority-Provided Commercial Cell Phone Devices

General Standard

The Authority's Artificial Intelligence (AI) Tools usage standard sets forth the requirements all Authority employees will observe when acquiring and using software that meets the definition of generative artificial intelligence "Generative AI."

The rapid growth and development of AI tools is transforming the way we work. These tools have the potential to automate tasks, improve decision-making, and provide valuable information back to us. However, as with most things that can be highly beneficial, they can also be used in nefarious ways. This presents challenges in terms of information security and data protection. This standard provides the rules that an employee must follow in order to ensure the safe and secure use of AI tools especially when it involves the sharing of potentially sensitive, secure, or customer information.

The purpose of this standard is to ensure that all employees use AI tools in a secure, responsible, and confidential manner. This standard outlines the requirements that employees must follow when using AI tools, including the evaluation of security, risks, and protection of confidential data.

Definitions

Generative Artificial Intelligence (Generative AI) is a class of computer software and systems, or functionality within systems, that use large language models, algorithms, deep-learning, and machine learning models, and are capable of generating new content, including but not limited to text, images, video, and audio, based on patterns and structures of input data. These also include systems capable of ingesting input and translating that input into another form, such as text-to-code systems.

Approved Practices

All employees shall adhere to the following security practices when using AI tools for work purposes.

1. Evaluation of AI Tools: AI and LLM (Large Language Model) tools that are purpose-built or significantly modified by the Authority, such as standalone AI/LLM solutions akin to ChatGPT, must be reviewed and approved by the I&TS Cybersecurity team before deployment. This is distinguished from AI functionalities that are integrated into vendor-supplied SaaS solutions which are not subject to this policy. AI and LLM tools must be reviewed for security features, terms of use, and privacy policies. A thorough evaluation of the company, developer, current known vulnerabilities, and plugins should also be performed.
2. Protection of Data: Employees must not upload, share, or use for querying any data that is confidential, marked as sensitive, proprietary, or protected by regulation. This includes data related to contracts, employees, customers, intellectual property, credentials, proprietary code and security sensitive information. Proprietary code in this context refers to any code that is not common or generic in nature and is characterized by its unique, non-obvious nature, including any novel methodologies, business logic or algorithms it may contain.
3. Access Control and Compliance: Employees must apply the same cyber security best practices we use for all company, sensitive, and user data. This includes the use of complex strong password management, keeping software up to date, and following all data retention and removal polices. Access must not be provided to third parties outside the company, including the sharing of login credentials or other sensitive information.
4. Data Privacy: Use of Generative AI tools shall be consistent with the Authority's Privacy Policy. Employees must exercise discretion when sharing information publicly or using data for search criteria in online AI tools. Generative AI tools like ChatGPT are public and all information used for generating output can be used to train their AI models. When asking questions of conversational AI, use generic examples and do not reference real people or projects. When using AI tools, if there is an option like "Chat history & training" to turn off, employees are encouraged to utilize this feature. Turning off such features ensures that chat histories are not utilized for the further training of the AI model, thus safeguarding the confidentiality of discussions.
5. AI Solutions: All AI applications must be run on approved, organization-controlled, and monitored platforms to ensure the utmost security and compliance with our cybersecurity policies. I&TS may revoke authorization for a technology that adds AI capabilities, or may restrict the use of those AI capabilities, if, in its judgment, those AI capabilities present risks that cannot be effectively mitigated to comply with this standard. Currently, the installation and operation of *downloadable* AI tools and large language models on the Authority's local devices is prohibited.

Approved AI Solutions	AI Solutions Pending Review
GPT-4 (OpenAI) GPT-3.5 (OpenAI) Gemini (Google)	LlaMA (Meta) BARD Falcon

Bing (GPT-4) CoPilot (Microsoft)	Cohere Claude v1 (Anthropic) PaLM 2 (Bison-001) Guanaco-65B Vicuna 33B MPT-30B 30B-Lazarus WizardLM, GPT4All BERT RoBERTa T5 MidJourney Sora StableDiffusion *All locally downloaded models*

6. Avoiding Bias: Generative AI tools can reflect the bias of their inputs and while many use controls to avoid content with certain biases, their outputs should always be carefully reviewed for accuracy and relevance. Conversational AI usually provides correct answers, but it can also create very plausibly sounding answers that are factually incorrect, so it is very important to validate your results. Available models are not trained on current events and their answers will be based on knowledge up until their last programming date.
7. Public Records: All records generated, used or stored by Generative AI vendors or solutions may be considered public records and must be disclosed upon request. Authority employees who use Generative AI tools are required to maintain, or be able to retrieve, upon request, records of inputs, prompts, and outputs in a manner consistent with the Authority's records management and public disclosure policies and practices.
8. Disclosure: Information you enter into Generative AI systems could be subject to a Public Records Act Request (PRA), may be viewable and usable by the company, and may be leaked unencrypted in a data breach. Do not submit information to a Generative AI platform that should not be available to the general public (such a confidential and personally identifiable information).
9. Fact Check: Review, revise, and fact check via multiple sources any output from a Generative AI. Users are responsible for any material created with AI support. Many systems, like ChatGPT, only use information up to a certain date (e.g., 2021 for ChatGPT 3.5).
10. Accountability: All images and videos created by Generative AI systems must be attributed to the appropriate Generative AI system.

Refer to this document quarterly, as guidance will change with the technology, laws, and industry best practices.