	Georgia Technology Authority	
Title:	**Generative AI Guidelines for Responsible Use**	
PSG Number:	**GS-25-001**	
Effective Date:	7/1/2025	**Review Date:**
Synopsis:	Guidelines for state entities when using Generative AI	

PURPOSE

AI technologies offer significant benefits for state operations, from improving data analysis to automating routine tasks and supporting informed decision-making, to streamlining processes and helping employees across departments. However, it is important to follow established AI guidelines to maximize these benefits responsibly.

Proper use helps prevent the leakage of sensitive information to external models, reducing the risk of data breaches, and helps maintain the accuracy and reliability of outputs, keeping generated content free of errors and bias. By following these guidelines, state employees can use AI responsibly, ensuring high-quality work and maintaining public trust in state-led initiatives.

These guidelines apply to all forms of generative AI, including but not limited to text, image, video, and audio generation.

When engaging with generative AI, all state employees should abide by the *5 Guiding Principles* to ensure the safety and welfare of all stakeholders with a vested interest in the data used and created from generative AI tools.

SCOPE and AUTHORITY

O.C.G.A 50-25-4(a)(10) – *State Government, Georgia Technology, General Powers*
O.C.G.A 50-25-4(a)(21) - *State Government, Georgia Technology, General Powers*
PM-04-001 – *Information Technology Policies, Standards and Guidelines*
PS-08-005 – *Enterprise Information Security Policy*

GUIDELINES

Use AI Tools Safely and Properly

Use only pre-vetted tools: State employees should only use AI which GTA has vetted, with pre-approved vendors. However, one version of an AI model being approved does not necessarily imply that other versions of that same model are

also approved. Additionally, an approved AI model can always have its approval revoked. Employees are encouraged to consult with GTA regularly for updates and to verify they are using compliant versions of AI tools.

Record prompts: Employees should create a record of queries and responses outside of the generative AI software or platform for future reference. This practice allows for accurate tracking and auditing, aiding in tracing decision-making processes, and can also help address issues that may arise from AI tool use.

Review AI-generated content: To manage the risks associated with AI-generated content, employees should apply careful review methods. This includes cross-checking AI outputs with trusted sources, verifying the accuracy of information, and assessing for possible biases or inaccuracies. Refer to the strategies below for more details.

Keep a Human-In-The-Loop (HITL) approach: AI tools, including generative AI and other automated decision-making systems, should not be the only agents involved in any decision-making. For instance, while an HR department may consult with AI platforms in making hiring decisions, it may not use AI to scan and reject resumes automatically. By keeping a "Human-In-The-Loop", AI systems can be properly kept in check.

Cite properly: Anything produced with the assistance of AI should be cited correctly. See section *How to Properly Cite AI-Generated Content.*

Double-check AI-generated facts: One of AI's greatest weaknesses is hallucinations, meaning that it sometimes generates inaccurate facts and presents them as truth. For any fact provided by an AI, find a reliable source that corroborates it.

Do not enter personally-identifying or confidential information: Generative AI models use user-inputted information to further train their models. For this reason, do not provide them with confidential or private state information, or personal records. If employees are uncertain whether specific data should be entered into a generative AI tool, they should first consult with their department's data privacy officer or designated compliance authority to clarify. When in doubt, it is better to err on the side of caution and avoid inputting potentially sensitive information without approval. Additionally, avoid using personal accounts, such as personal email addresses, within any software to ensure that all content generated for

government purposes remains secure. For applications requiring users' data, ensure that such data is only used with user consent.

Create a culture of transparency: All state employees should be open and honest about their AI use. Cite and acknowledge all AI-generated, brainstormed, or edited content so that others know where your information is from. Teach and learn from other employees. Employees should help one another and make sure that all AI information is cited honestly. Employees should rely on the honor system, openly disclosing AI usage without fear of severe repercussions. This transparency enables the early detection of improper AI use, preventing potential impacts on sensitive information.

Assess risk level: AI has the potential to massively improve productivity, but this is not a guarantee. Employees should use common sense and exercise caution in deciding whether AI is right for a given task. See section Mitigating the Risks of Generative AI.

Be Vigilant in Virtual Meetings

The use of AI note-taking tools in Microsoft Teams meetings or any other conferencing platforms is strongly discouraged for meetings held within the State. Meeting hosts may opt to use the recording and transcription services built into Teams.

Meeting hosts should be diligent when admitting participants to virtual meetings to ensure no AI bot note-takers join. You can look for the ".ai" suffix to confirm.

While AI capabilities for recording and transcribing virtual meetings offer convenience, they should be used thoughtfully. Agencies should ensure that the use of AI tools in virtual meetings complies with all relevant laws and regulations, including those related to data protection, intellectual property, and labor laws. AI should not be used to manipulate or misrepresent the contributions of participants. Be mindful of the potential for AI to misinterpret or misrepresent human communication. AI outputs should be used only to supplement, not replace, human judgment.

Once created, meeting recordings and transcriptions become public records and are subject to retention rules. Further, there are storage and cost considerations associated with maintaining recordings. Meeting hosts are responsible for determining the need for recordings and automated transcriptions and keeping any

records created.

Do Not Put Private Data at Risk

The goal of these guidelines is to reduce the risk of employees using Shadow AI. Using AI without proper citation and disclosure can release sensitive state or personal information to generative AI models, communicate inaccurate information to the public, or log incorrect information into state records without accountability.

All state employees using AI should understand that transparency builds trust, supports accountability, and encourages collaborative efforts, leading to increased productivity, reduced repetitive tasks, and more efficient research. However, these benefits are only possible if everyone knows what work is AI-generated. All AI-generated content – including text, images, videos, and audio – must be clearly labeled as AI-generated and double-checked to ensure it is free of inaccurate information, AI hallucinations, or bias.

Strategies for Data Privacy and AI:

- Keep personal and work materials separate and create an account using your work email specifically for GenAI materials.
- Opt out of data collection on any tool you use, particularly those involving pictures.
- Protect all data used by AI systems from unauthorized access or breaches. This includes regularly changing passwords, minimizing data retention by regularly clearing chats, and conducting regular audits to ensure compliance.
- Do not enter personal and/or sensitive data into a GenAI model.
- The use of GenAI tools must be consistent with Georgia's privacy laws, such as the Georgia Computer Data Privacy Act (GCDPA).
- Beware of Bias

Generative AI models can inadvertently perpetuate biases based on race, color, ethnicity, sex (including pregnancy, childbirth, and related medical conditions, gender identity, intersex status, and sexual orientation), religion, age, national origin, disability, veteran status, genetic information, or any other classification protected by law. Bias can occur at any stage from data collection, labeling, model training, and deployment. The most common types of bias in AI include but are not limited to:

- **Algorithm Bias:** This bias results from users improperly asking a question to an AI or if the user's feedback provided after an AI's mistake is not specific

enough. When querying an AI, carefully word the question and be sure that the AI is responding to the question asked.

- **Cognitive Bias:** This bias results from implicit biases users possess. If a user's query to an AI model contains bias, the AI may reproduce this bias in its output. When querying an AI, think about whether the question being asked presupposes things that are not necessarily true.
- **Confirmation Bias:** This bias results from users being quick to accept output that matches their expectations. When querying an AI, carefully consider whether the answer is correct and consult a third party or non-AI source before proceeding, regardless of whether the answer is what was expected.

It is important to be aware of these types of biases in any work produced by AI and screen the work for bias before putting them into use.

Guiding Principles

To safeguard the welfare of and enhance the services provided to Georgians, GTA has established five guiding principles governing the design, implementation, and utilization of automated systems. Informed by industry research and experts, these principles are intended to guide state agencies as they integrate protective measures into their policies and operational procedures. These principles serve as a framework whenever automated systems have significant implications on the rights of Georgians or their access to essential services.

Implement Responsible Systems

- *User-centered Design and Development*
  State agencies should prioritize user research as an integral component in the procurement or development of automated systems. It's important to maintain the human element during the design of any service. Seek input and insights from user groups, diverse stakeholders, and domain experts to identify concerns, risks, and potential impacts associated with the system.
- *Comprehensive Testing*
  Automated systems must undergo pre-deployment user testing to identify potential risks and assess their intended functionality. Implement risk identification and mitigation strategies to ensure system safety and effectiveness, including addressing unintended consequences.
- *Ongoing Monitoring and Improvement*
  It's essential to confirm that the system continues to operate as intended; deviations should be addressed promptly. Adhere to domain-specific standards to ensure compliance and compatibility with industry best

practices. Regularly evaluate system performance, ethical adherence, and the impact on outcomes and take corrective actions as needed.

- *Consideration for Non-deployment*
  State agencies should be prepared to halt the deployment of an automated system or remove it from use if it fails to meet safety or effectiveness standards.
- *Data Protection*
  Ensure that the design, development, and deployment of automated systems protect against inappropriate or irrelevant data use. Mitigate the risks associated with the reuse of data, preventing compounded harm.
- *Independent Evaluation*
  GTA reserves the right to conduct an independent evaluation and report to confirm the safety and effectiveness of automated systems, including mitigation of potential harm. GTA will make evaluation results publicly available whenever appropriate, promoting transparency and accountability.
- *Ensure Ethical and Fair Use of Automated Decisions*
  - Fairness, Transparency, Accountability, and Privacy
    State agencies should adopt a set of ethical AI principles that prioritize fairness, transparency, accountability, and privacy in the design and deployment of AI systems for state services. Develop a set of ethical guidelines for AI system design and deployment. Users should be able to understand why a particular decision was made, building trust in the system. Assign specific individuals or teams to be responsible for monitoring AI systems for bias and corrective actions. Clearly defined accountability ensures that bias-related issues are addressed promptly.
  - Algorithmic Bias Awareness and Mitigation
    Provide training and educational programs for agency staff about the concept of algorithmic bias and its potential impacts on decision-making processes within state services. Awareness is the first step in addressing bias effectively. Develop and implement strategies to address and mitigate algorithmic bias whenever detected, such as refining algorithms, adjusting data inputs, or retraining models. Regularly assess how AI systems affect different user groups. Understand any disparities or unintended consequences that may arise and take action to rectify them.
  - Data Quality and Diversity
    Carefully curate and vet the data used to train AI algorithms. Make sure the data is diverse and representative of all relevant demographic groups. This helps prevent biased outcomes caused by skewed data.

- Regular Assessments
  Continuously monitor AI systems to detect and rectify biases where they emerge. Regular assessments are essential to maintaining fairness and effectiveness over time.

Maintain Data Quality and Privacy

- Data Governance Framework
  Establish clear guidelines for data governance to maintain integrity and privacy.
- Security and Data Handling
  Prioritize robust security measures and transparent data handling practices.
- Accuracy and Retention
  Ensure data accuracy, minimize storage, and dispose of obsolete data.
- Compliance and Accountability
  Maintain compliance with data protection laws, conduct regular audits, and involve the public in the decision-making process.

Keep AI Usage Transparent

- System Use
  Ensure that individuals are informed about the use of automated systems and understand how these systems contribute to outcomes that can affect them.
- Accessible Documentation
  Encourage designers, developers, and deployers of automated systems to provide plain language documentation that is easily accessible to the public. This documentation should include clear descriptions of system functionality and ownership, the role of automation, and explanations of outcomes.
- Up-to-date Notices
  Require that notice regarding the use of automated systems is kept current, and individuals impacted by the system should be notified of significant use cases or key functionality changes.
- Technically Valid and Accessible Explanations
  Ensure that individuals have access to information explaining how and why outcomes that affect them were determined by the automated systems, even when these systems are not the sole contributors to the outcome. Mandate that automated systems provide technically valid, meaningful, and useful explanations to affected individuals, as well as operators and stakeholders who need to understand the system. The level of detail in these

explanations should align with the level of risk involved.

- Public Reporting
  Promote the publication of summary information about automated systems in plain language. Assessments of the clarity and quality of notice and explanations should also be made public whenever possible to enhance transparency and public trust.

Keep Human Involvement at the Center

- Human Responsibility and Ownership
  State agencies should establish and adhere to policies that emphasize human responsibility and ownership of the outcomes produced by AI systems used in state services. AI systems should not operate in isolation. State agencies should ensure that humans retain control over the operation of AI systems and that human decision-makers remain responsible for the final decisions made with the support of AI.
- Ethical and Transparent Design and Use
  State agencies should prioritize transparency and accountability in the deployment of AI systems. Agencies should retain clear records of AI system use, their objectives, and the roles of individuals overseeing and interacting with these systems. Agencies should mandate that AI systems be designed and used in accordance with ethical principles that prioritize fairness, transparency, accountability, and privacy. Ethical considerations should be an integral part of AI system development and use.
- Clear Roles and Responsibilities
  Clearly define roles and responsibilities for individuals involved in AI system implementation. This includes specifying the duties of AI system operators, data stewards, and decision-makers.
- Human-AI Collaboration
  Encourage collaboration between humans and AI systems to enhance decision-making processes. AI should be viewed as a tool that complements human expertise rather than a replacement for human judgment.
- User Training and Education
  Promote user education to ensure that individuals interacting with AI systems understand the capabilities and limitations of these technologies. Users should be aware of how AI contributes to outcomes and that humans remain responsible for those outcomes. Agencies should invest in training and development programs to equip their staff with the skills and knowledge necessary to effectively use AI systems and make informed decisions.
- Ownership of Data and Models

State agencies should retain ownership and control over the data used to train AI models and the models themselves. This ownership ensures that AI systems serve the agency's mission and values.

How to Properly Cite AI-Generated Content

Citing AI-Generated Content

Generative AI can be used in many ways and take different forms. For example, it can help produce lists for brainstorming, act as a grammar checker, and even create rough drafts of text. However, in some circumstances, it can be misleading to take the content generated by AI and present it as your own. A helpful way to think about it is this: AI should be cited when its content is used as a product rather than a tool.

What does it mean for AI-generated content to be a product?

AI-generated content is considered a product when a substantial portion of the content is featured in the final product, even if the user extensively edits it. As an example, imagine a student having to write a paper for an English class. If that student asks an AI model for potential paper topics and then uses one of those suggestions to develop their own topic, the AI model has merely acted as a brainstorming tool. Therefore, it is acceptable to continue without citing the AI model. However, if the student directly uses one of the AI-generated topics in their paper (either verbatim or with changes to the phrasing), they would need to cite the AI model as a source. In this case, the AI model has gone beyond serving as a tool and has become part of the final work product. In short, if the AI model helps a user in generating their own content and ideas, citation *isn't necessary*, but if the AI model's content or ideas appear in the final product, a citation *is required.*

How do I cite an AI model?

This depends on the type of work product created by the user. If AI is used in the generation of an informal document like an email, then all that is necessary is to acknowledge somewhere in the text that AI was used in its development and which AI model was used (e.g. "This email was edited for style and content by GPT-4o"). However, on more official documents, it is necessary to use a more detailed citation. The in-text citation below serves as a good example. T

(GPT-4o, “Rewrite this document so that it would fit on one page.” *OpenAI*, June 1, 2024)

Be sure to include the name of the AI model, the name of the prompt used, the company to which the AI model belongs, and the date that the content was generated.

What about non-text mediums, like photos, video, and audio?

Photos and videos must have a clear watermark indicating that the following content was generated by AI. Any audio including AI-generated sound must include an audio disclaimer notifying listeners that some or all of the recording was generated with AI.

Mitigating the Risks of Generative AI: Rubric

- Human Oversight:

Implement robust human oversight in all AI processes; AI should never have autonomous decision-making capabilities. Ensure that a human is reviewing both the inputs to and outputs from the AI, across all stages of AI use.

- Accuracy of Information:

Carefully review AI-generated content for accuracy. Cross-check information with reliable sources to prevent the spread of false or misleading content.

- Processing Sensitive Information:

Always verify that the AI is not handling sensitive or private information, unless given explicit approval from the Georgia Technology Authority. In such cases, ensure that strict data protection protocols are being consistently enforced.

- Bias and Ethical Concerns:

Assess the potential for AI to introduce or perpetuate bias in its outputs. Regularly audit AI systems to proactively identify and address potential ethical issues.

- Transparency:

Clearly inform users when they are interacting with AI-generated content. Maintain transparency about how the AI is being used and what its limitations are.

- Compliance with Policies and Standards:

Ensure that all AI applications are fully compliant with relevant laws and regulations, both at the federal level, as well as those set forth by the State of Georgia.

Mitigating the Risks of Generative AI: Task-Specific

Information Processing

Permissible: Summarizing publicly available information (e.g., news and research journal articles).

Prohibited: Summarizing sensitive government documents and personal private information, such as biometrics and financial data.

Questions to Consider:

- Does the data include sensitive information?
- Is all the information in the summary accurate?
- How could this summary be biased?

Information Gathering

Permissible: Using an LLM to provide you with basic information on a topic.

Prohibited: Relying solely on an LLM and failing to verify your findings with other sources.

Questions to Consider:

- Does the data include sensitive information?
- Is all the information in the summary accurate?
- How could this summary be wrong?

Coding/Interpreting Software

Permissible:

- Generating documentation and/or comments for existing software code.
- Debugging existing software code.

Prohibited:

- Deploying AI-generated code without human validation/oversight.
- Generating code that violates Georgia's AI ethics standards.

Questions to Consider:

- Is GenAI being used to code entire scripts autonomously?
- Is the generated code thoroughly reviewed by a human?
- Does the software application deal with any sensitive information?

User Interaction

Permissible: Creating chatbots to streamline user support.

Prohibited: Storing information from user interactions with GenAI without explicit consent from the user.

Questions to Consider:

- Are users aware that they are interacting with AI-generated content?
- Does the AI application require sensitive or private user information?
- Is that information stored, or is it promptly deleted following the interaction?

Content Generation

Permissible:

- Brainstorming a list of ideas.
- Checking for grammatical errors and/or formatting a human-written document that does not include sensitive information.

Prohibited: Generating content that replicates or closely mimics copyrighted material.

Questions to Consider:

- Did the GenAI "hallucinate" or create other false information?
- Is the content plagiarized or involved in any form of copyright infringement?

RELATED ENTERPRISE POLICIES, STANDARDS AND GUIDELINES

Generative AI Responsible Use SS-25-001
Enterprise Artificial Responsible Use (PS-23-001)
Artificial Intelligence Responsible Use Guidelines (GS-23-001)
Artificial Intelligence Responsible Use (SS-23-002)
Data Security - Electronic Records (SS-08-003)
Reliance on Electronic Records **PS-08-007**