CITY OF LONGBEACH

Generative AI Guidance

V 1.1

Generative Artificial Intelligence (Gen AI) systems have become increasingly popular and prevalent in all areas of society today. The City recognizes the opportunity for a guided and responsible approach to using Gen AI that acknowledges the benefits to our efficiency while minimizing the risks around AI bias, privacy, and cybersecurity.

We believe that Gen AI is a valid and useful exploratory "personal assistant" tool that can enhance our productivity and support essential City tasks; however, we encourage employees not to rely heavily on this technology or use it to support critical processes until the City develops further policies.

Currently, the City does not recommend any particular Gen AI system and does not have an Enterprise Gen AI system in place.

Integration of any non-contracted Gen AI systems without an executed City license agreement with the City's Microsoft environment (which includes SharePoint and Teams) is **not permitted** due to privacy and security risks that these systems pose to City data. These free and/or non-Enterprise AI systems may access and expose sensitive information, share sensitive data with external entities without your consent, and introduce vulnerabilities to your devices. This applies to common systems like Otter.ai and Read AI, as well as other AI meeting assistants such as Atlas, Avoma, Colibri, Compas, Dubber, Fathom, Fireflies.ai, Gong, Grain, Meet Record, MeetGeek, and Sembly.

How to Use this Document:

This Guidance is our first step in developing a policy intended to minimize risks that may arise from the use of Gen AI technology while we conduct additional research and analysis. Refer to this document quarterly, as guidance will change with the technology, laws, and industry best practices. Key sections include:

- **Guidelines:** Practices intended to encourage responsible and innovative use of Generative AI but minimize risks to the City and to the public.
- **Sample use cases:** Several Gen AI prompts that you might use for your work, along with best practices for each.

- **Principles:** Core codes of conduct that reflect our Long Beach values and responsibilities as public servants.

This Guidance is intended to govern Gen AI only – AI systems that specifically generate content, not traditional AI systems at large. Refer to the Definitions section below for details.

Guidelines:

1. **<u>Do not</u> submit sensitive or private information to a Generative AI platform you wouldn't want available to the public.** Everything you enter into a Generative AI system is subject to the California Public Records Act (PRA). The information you provide may be viewable and usable by the Gen AI service provider and may be compromised if they experience a data breach. Know your data! Remember that the [City Code of Conduct](#) requires that we safeguard all information and data entrusted to us. Sensitive and private information includes but is not limited to the following:
    a. Personally identifiable information (PII), Protected Health Information (PHI), Tax IDs, passport numbers, driver's license numbers, addresses, contact details, financial information, credit card numbers, and more.
    b. Passwords, passcodes, and other confidential information.
    c. Sensitive business data:
        i. Any information that if exposed publicly would tarnish or harm the reputation of the City.
        ii. Any information that if exposed publicly would open the City to regulatory or legal action.

2. **You should create a Generative AI system account just for City use** to ensure that public records are kept separate from personal records.

3. **Carefully review, verify, and fact-check via multiple sources the content generated by Generative AI** to eliminate instances of bias, offensive, inaccurate, or harmful material. Acknowledge that Generative AI systems can reflect cultural, social, and economic biases of source material. You are responsible for verifying that the information is accurate to mitigate biases and misinformation.

4. **Please cite or reference when you use Generative AI within your documents and communications.** Audiences should know when Gen AI was involved with producing content, and we encourage you to lead by example and disclose this information whenever possible. Our goal is to build a positive culture around using Gen AI systems, and being transparent with its usage is the first step in challenging any stigmas that may exist. Gen AI is not a shortcut; it is a tool used to enhance our creativity, optimize our efficiency, and strengthen our effectiveness.
   a. Examples of citations are below in the "How to Cite" section.
   b. Your citations should include which Gen AI system was used and confirmation that you have fact-checked everything.

5. **Please complete the form below to disclose to TID when you use Generative AI**. The intent of disclosing Gen AI usage is so TID may continue learning when Gen AI best fits in our work and potentially implement further systems and policies dedicated to addressing these emergent needs.
   a. This form is <u>anonymous</u> and will take 1 minute to complete. Your responses will help us explore how we can support the responsible and safe use of Generative AI across the City.
   b. Please note that you only need to complete the disclosure form for each <u>unique</u> use of Gen AI.
      i. For example, if one day you use Chat GPT to generate a program logo, and the following day you use Chat GPT to generate a meeting recap, you'd fill out the disclosure form twice.
      ii. However, if you use Chat GPT two days in a row to generate a meeting recap, you'd only fill out the form once, even if the meeting topics are different or you use a slightly different prompt. As our goal is to learn about the use cases/purposes of Gen AI, we are not interested in recording each individual time it is used.
   c. Form: https://forms.office.com/g/Lx3BthE34H

6. **Please opt out of data collection and sharing whenever possible**. Gen AI retains the data you input and the content it generates as part of its training and self-learning processes. Sometimes, this data is shared with third parties. This means every interaction with Gen AI may be saved and utilized to improve its functionality over time. Many Gen AI systems allow you to opt out of data collection (or delete your chat history), meaning they will not keep any of the data you provide, and those data will not be used in the systems' models.

      a. Instructions on how to opt out can be found in the "How to Opt-Out of Data Collection" section.

7. **City Departments may provide additional rules for their staff regarding Generative AI.** If your Department, Bureau, or Division has additional regulations for Generative AI, please disclose this with TID by completing the form below.
         a. Form: https://forms.office.com/g/Lx3BthE34H

8. **We encourage you to participate in the City's Gen AI Governance Workgroup** to help advance AI usage, guidance, and best practices in the City. If you are interested in joining, please complete this form:
         a. Form: https://forms.office.com/g/Lx3BthE34H

9. **If you are interested in using a Gen AI tool but unsure of how to proceed, contact TID.** There are many Gen AI tools out there, and we are a resource to help you evaluate the benefits and risks.
         a. To schedule a consult, email smartcity@longbeach.gov

Sample Use Cases:

We have compiled several use cases for Generative AI in a Use Case Registry. City staff identified these as ways in which they utilize Gen AI. Please keep in mind that these are just example use cases; they do not represent the only ways you can use Gen AI systems. We are still in the experimental phases, and there are many use cases we don't know about yet.

Principles:

These principles describe general codes of conduct that represent the City's values and align with our responsibilities to the residents we serve. They guide City employees in their use of Generative AI technology. **City employees should adhere to these principles and should champion accountability and compliance among their peers.**

The following principles reflect what we heard from City staff in TID's Generative AI survey. They build upon the City's Code of Conduct and Ethics, Smart City Initiative Guiding Principles, and Data Privacy Guidelines.

Safety & Security
- Every technology we employ impacts the safety and security of our overall environment, both physically and digitally. We must be mindful of safe data privacy practices for our constituents, and we will consider implementing mechanisms to minimize security risks presented by Generative AI and protect personal data to the greatest extent possible.
- We must be aware that Generative AI applications and models can be compromised with malicious links and prompts if they are not kept up to date and combined with sound security practices.

Empowerment
- The use of Gen AI should support our work to deliver better, safer, more efficient, and equitable services and products to our residents.
- We trust our staff to do the right thing given the right tools and guidance. Staff will need to exercise judgement to ensure we get the benefits from Generative AI tools while avoiding negative impacts on the City and its constituents.

Exploratory Mindset
- We embrace the possibilities of technology and innovation and commit to responsibly exploring and evaluating creative AI technologies that can advance positive outcomes.
- We acknowledge we are still learning about the full scope of what Generative AI can provide us with, so let's learn together.

Transparency & Accountability
- We acknowledge that we do not have all the answers, nor can we foresee all consequences of adopting this technology. We recognize further research into this technology may reveal issues that require more restrictions on its use.
- We will be transparent when updates to this Guidance are implemented and will hold ourselves accountable to the possible impacts this technology may have.
- The City will practice data integrity and use content produced by Gen AI for stated purposes and purposes that serve the public good.
- Generative AI is a tool. We are responsible for the outcomes of our tools. For example, if autocorrect unintentionally changes a word, changing the meaning of something we wrote, we are still responsible for the text. Technology enables our work; it does not excuse our judgment or accountability.

Equity & Inclusion

- We are stewards of the public and we will use Generative AI to support and uplift our communities.
- Everything we do, regardless of the tools, is a reflection of the City and ourselves. We acknowledge that AI systems have the potential to perpetuate inequities, discrimination, harm, and bias against Long Beach residents, and the City will commit to using a racial equity lens to examine the burdens, benefits, and unintended consequences of Gen AI.
- We acknowledge that Generative AI introduces concerns about job security and job displacement, and we recognize that we must be resilient as a workforce in the face of emerging technologies.

How to Cite:

Citing the use of Generative AI is more than a best practice, it is a powerful way to communicate to your audience how Gen AI is improving your work. **Citing your use of Gen AI is easy and shows that you are an early adopter!**

Generative AI can be cited as a statement within your document or as a footnote, endnote, or header. Citations for content created by Generative AI should include the following:

- GEN AI SYSTEM: Name of the Generative AI system (ChatGPT, Google Gemini, Copilot, Midjourney, etc.)
- DD/MM/YYYY: Date when content was generated.
- Confirmation that you fact-checked information.
- *Optional*: State whether the content is unmodified from the Generative AI system.

In-text citation example unmodified: *"This document was drafted with support from (GEN AI SYSTEM) on DD/MM/YYYY. The content was fact-checked by City staff. This information is unmodified."*

In-text citation example modified: *"This email was drafted with support from (GEN AI SYSTEM) on DD/MM/YYYY. The content was edited for brevity and fact-checked by City staff."*

Footnote/endnote/header example: *(GEN AI SYSTEM, DD/MM/YY. Fact checked by City staff.)*

How to Opt-Out of Data Collection:

Some Generative AI systems offer the option to opt-out of data collection, and we encourage you to do so whenever possible. These options are typically found within a system's privacy settings or user settings.

One Gen AI system that allows users to opt-out is ChatGPT. Once a user opts out of data collection, their information is not used to train the AI system for other users. Below is a step-by-step guide on opting out of ChatGPT's data collection.

1. Select your name on the top right of the interface, then select "Settings"
2. Click "Data controls" in the settings menu and disable "Improve the model for everyone"
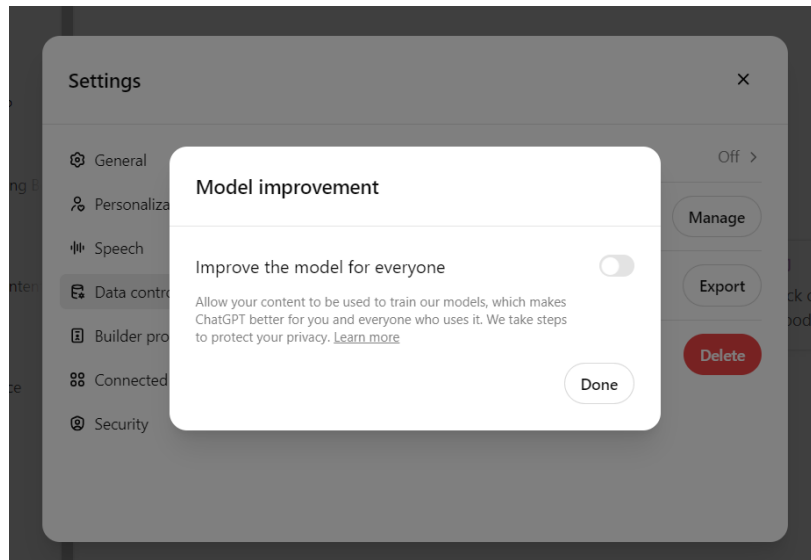
3. In addition, Chat GPT allows users to disable its "memory" so it does not remember details and preferences between your chats, so what you enter in one chat will not impact your next one. To do this, from the Settings menu, select "Personalization," then disable the "Memory" toggle. Chat GPT conversations will only be retained for 30 days in this mode.
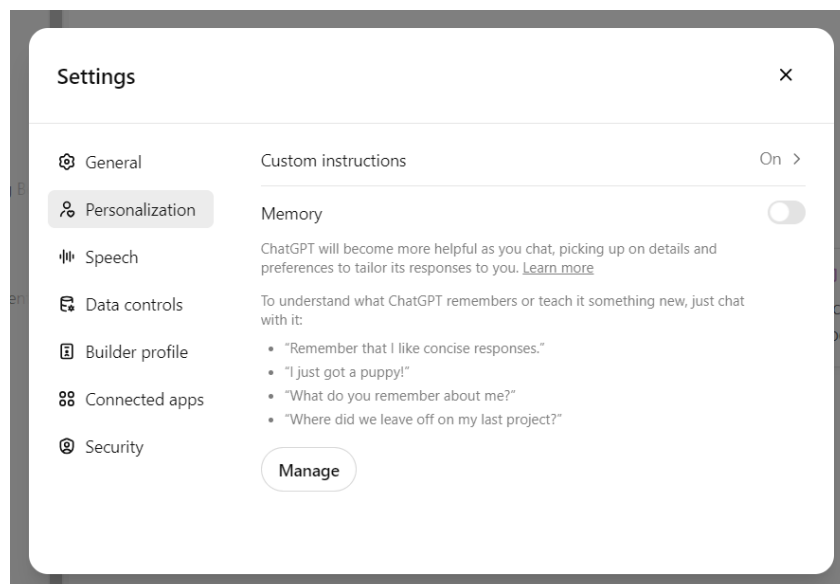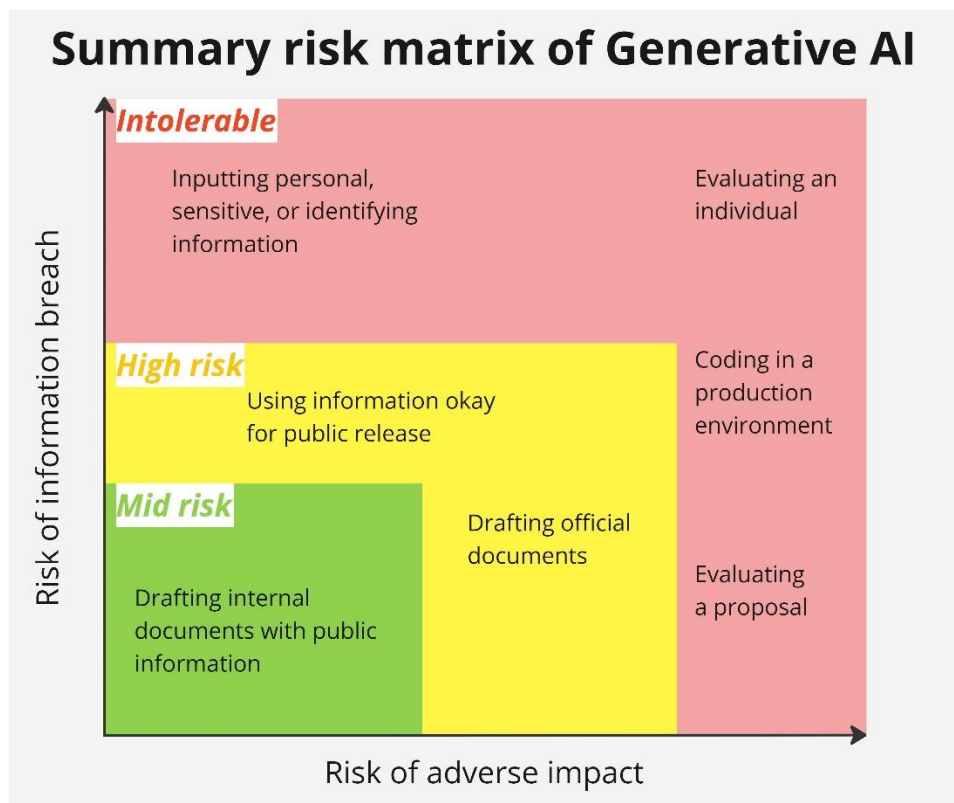
Assessing Risk in Generative AI Use Cases:

The risk presented by Generative AI tools varies by use case, with the risk spectrum ranging from mid-risk to high-risk to intolerable risk. San Jose's Data Privacy Office developed the following framework.

Two key factors determine Generative AI risk:

1. **Risk of information breach**: the potential harm if the information exchanged with a Generative AI is released to an unintended audience. This can include entering personally identifiable information, sensitive records, or confidential business information into Generative AI. Additionally, any information entered into Generative AI may be subject to the Public Records Act (PRA). If you wouldn't share the information in a public forum, don't share it with a Generative AI.
2. **Risk of adverse impact:** the potential harm of using the output for a decision, task, or service. This impact can be different for different populations and should be considered from an equity lens, such as adverse impacts on people of a certain race, age, gender identity, or disability status. Not only can AI be biased, but it can also provide false information. In general, if Generative AI is used in relation to City processes that can alter an individual or community's rights, freedoms, or access to services, multiple users should thoroughly review it before any document is finalized or action is taken.

Summary risk matrix of Generative AI

Intolerable

Inputting personal, sensitive, or identifying information

Evaluating an individual

High risk

Using information okay for public release

Coding in a production environment

Mid risk

Drafting official documents

Drafting internal documents with public information

Evaluating a proposal

Risk of information breach

Risk of adverse impact

History:

Gen AI is a new branch of AI technology that can generate helpful content—such as emails, research summaries, reports, and images — at a user's request. Popular Generative AI systems include OpenAI's ChatGPT, OpenAI's Dall-E, Google Gemini, Midjourney, Microsoft O365 integration with OpenAI Copilot, Microsoft GitHub Copilot, and Stable Diffusion. These systems have the potential to support many City business functions and services, however their use also raises important questions, particularly around the sourcing of training data, ensuring proper attribution of generated content, and the handling of sensitive or public data. Further research into this technology may uncover issues that require more restrictions on its use. Please remember that Generative AI technology is already incorporated into many existing services and products we use, such as Office 365.

Through a survey distributed to all City of Long Beach staff in October 2023, the Technology and Innovation Department (TID) found interest in using such systems to conduct City business but low levels of actual use. Over 70 percent of respondents described work-related benefits that Gen AI presents. However, just 14 percent of staff respondents said that they had used Gen AI in the past for work-related purposes.

Definitions:

- **Generative Artificial Intelligence**, commonly referred to as "Generative AI" or Gen AI, is an "automated system" used to generate "content." Generative AI uses massive datasets to develop content when given a prompt. Common Generative AI systems include ChatGPT, Google Gemini, Copilot, and others.

- **Training data** is data gathered from a wide range of sources available on the internet, such as books, websites, articles, and other texts. The goal is to cover a broad human knowledge and language use spectrum. The Gen AI system uses this data to understand user prompts and generate "content."

- An "**automated system**" is any artificial intelligence system, software, or process that uses computation as part of a system to generate outputs and outcomes, make or aid decisions, inform policy implementation, collect data or observations, or otherwise interact with individuals and communities.

- **Traditional AI** is an artificial intelligence system trained for a specific, narrowly defined task. Commonly used AI systems include Google's search algorithm, recommendation engines like Netflix or Amazon, and writing assistants like Grammarly.

- "**Content**" includes text, emails, presentations, images, video, audio, architectural documents, diagrams, and other forms of media.

- **Large Language Models (LLMs)** are a type of AI model trained on immense amounts of broad data. They are capable of understanding and generating natural language and other types of content to perform a wide range of tasks with minimal fine-tuning.

- "**Personally Identifiable Information,**" or "**PII**" for short, is any data used to identify someone.

- The **California Public Records Act (PRA)** gives residents the right to access the public records that the government maintains. https://www.ftb.ca.gov/your-rights/california-public-records-act.html

Resources and Trainings:

- The Emerging Use of Generative AI in City Government (City of Long Beach): https://www.longbeach.gov/smartcity/news/the-emerging-use-of-generative-ai-in-city-government/
- An Intro to Generative AI (City of Long Beach Data Learning Community): https://longbeach.sharepoint.com/:f:/r/sites/Data-Collaborative/Shared%20Documents/General/Trainings/AI?csf=1&web=1&e=jM3T5X
- Government AI Coalition: https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition
- About Generative AI (Harvard University Information Technology): https://huit.harvard.edu/ai/about
- What is Generative AI? An AI Explains (World Economic Forum): https://www.weforum.org/agenda/2023/02/generative-ai-explain-algorithms-work/

Change Log:

Version	Notes	Date
V1.0	Interim Guidance published	Jan 8, 2024
V1.1	Guidance updated to reflect the evolving AI system landscape, reinforce AI systems' privacy and security risks, and explain the City's posture on common AI systems. Updated glossary to distinguish between Gen AI and traditional AI. Added language about permitted systems. Added Guideline #9.	Aug 13, 2024

Change Log: