



City of Boston Interim Guidelines for Using Generative AI

Version 1.1

Prepared by Santiago Garces, Chief Information Officer, City of Boston

Published: 5/18/2023

Applies to: *all City agencies and departments with the exception of Boston Public Schools*

Purpose

Generative AI is a set of relatively new technologies that leverages large (very large) volumes of data along with some machine learning (ML) techniques to produce content based on inputs from the users known as prompts. The new content can be written (e.g. ChatGPT or Bard), or visual (e.g. Dall-E). These tools are evolving rapidly, and are still the subject of active research: improving our understanding of how they actually work, and the impacts of their use in society. These tools are not actual intelligence in the human sense, rather, they are very sophisticated models that predict what the language, text, or video that satisfies the prompt should be. **Because of their impact and potential usefulness, as well as the risks and dangers, these guidelines serve as an interim resource for employees of the City of Boston.**

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 nor our accountability.

These guidelines should be replaced in the future with policies and standards. But we want to encourage responsible experimentation and we encourage you to try these tools for yourselves to understand their potential. The Department of Innovation and Technology will support events and workshops that can support people and teams interested in learning more about these technologies. For the time being we encourage you to watch this video from Innovate.US about how to get started with generative AI in government:

<https://bit.ly/InnovateUS-AI>

You can also share your experiences, thoughts, and concerns via this online form:

<https://forms.gle/BptUcVhRdnTwHdxJ7>



Sample Use Cases

These are some of the types of uses that could be beneficial. Additional good practices and examples can be found at the end of this document.

1. Writing a memo. In government we often have to write short documents that present an argument why a policy should be adopted or a decision should be made. For instance try the prompt in ChatGPT, Bard, and other generative text tools:

Unset

```
Write a memo to the Chief Innovation Officer about the potential benefits of the use of generative AI in city government.
```

2. Writing a job description. Generative AI can produce job descriptions that aggregate and average parts of similar job descriptions, giving you a very good generalized version of a job description. For instance try the prompt in ChatGPT, Bard, and other generative text tools:

Unset

```
Write the job description for a Chief Information Officer of a large city
```

Principles

Empowerment

- The use of AI should support the work of our workforce to deliver better, safer, more efficient and equitable services and products to our residents.
- We rely and trust in our public sector professionals to do the right thing given the right tools and guidance. You will need to exercise your judgment to make sure we



get the benefits from the tools while avoiding the negative impacts for the City and its constituents.

Inclusion and Respect

- The use and development of AI should support the development of work that repairs damage done to racial and ethnic minorities, people of all genders and sexual orientations, people of all ages, people with disabilities, and others. Our work should uplift these communities and connect them more effectively with the resources they need to thrive.
- Everything we do, regardless of the tools, are a reflection of the City and ourselves. We are stewards of the public, and we will use tools respectfully and responsibly.

Transparency and accountability

- We embrace the possibilities of technology and community. We acknowledge that we do not have all the answers nor can we foresee all consequences. But when we act transparently, we build trust and we gain the ability to learn collectively.
- We also acknowledge that experimentation might have costs and impacts in of itself including the usage of power, greenhouse gas emissions. Being purposeful and accountable to these impacts is important.

Innovation and Risk Management

- We understand that there is value to be had in the use of technology, particularly new generative AI, but there are also risks, some of which will not be apparent or fully understood upfront.
- We embrace a culture of responsible experimentation, where we maintain control and understanding of the use of new tools while we develop new uses that drive efficiency, delight, civic dialogue or other outcomes in service of our residents.

Privacy and Security

- Every technology tool that we use has an impact on the security of our overall environment, and the privacy and digital rights of our constituents.

Public Purpose

- The best known of these new tools are developed for commercial purposes. While they can be adapted for mission-driven work by public professionals, it is important to maintain service to the public at the center of our work.



Guidelines

Fact Check and review all content generated by AI, especially if it will be used in public communication or decision making.

- **Why:** While Generative AI can rapidly produce clear prose, the information and content might be inaccurate, outdated, or simply made up. It is your responsibility to verify that the information is accurate by independently researching claims made by the AI.
- **What to look for:**
 - Inaccurate information including links and references to events or facts.
 - Bias in the positions or information. We want to make sure that vulnerable populations are not harmed by these technologies. Think about how racial and ethnic minorities, women, non-binary, people with disabilities or others could be portrayed or impacted by the content.

Disclose that you have used AI to generate the content. You should also include the version and type of model you used (e.g, [Open AI's GPT 3.5](#) vs [Google's Bard](#)). You should include a reference as a footer to the fact that you used generative AI:

- **Why:** even when you use AI minimally, disclosure builds trust through transparency and it might help others catch errors.
- **Suggestions:** document how you used the model, the prompts you used etc. it could be helpful to you and your colleagues to better understand how you can use these technologies better and more safely.
- **Sample credit line:** “This description was generated by ChatGPT 3.5 and edited by Santiago Garces”
- **Sample credit line:** “This text was summarized using Google Bard”

Do not share sensitive or private information in the prompts

- **Why:** data including prompts used in generative AI might be used by the companies that power these systems. Any information that includes personally identifying information about our residents, other public servants, etc. could inadvertently be shared with others. Basically if you wouldn't share with other people or want to put the prompt in a public place, avoid sharing the information in the prompt. *If you have an application that requires sensitive information to be used with a generative*



AI, contact DoIT so we can help you provision access to enterprise secure resources to do so.

More Examples, Do's and Don'ts:

These are some suggestions on the kinds of uses that seem to be particularly useful for City uses. By encouraging responsible experimentation we are hoping to expand the potential uses, while minimizing risks.

Drafting documents or letters:

Generative AI provides a great opportunity to get started on a memo, letters, job descriptions. Note that when creating a prompt for ChatGPT for this context, it can consider including any specific format preferences such as essay, bullet points, outline or dialogue. Additionally, you can request the use of specific keywords or phrases, or technical terms to be included or avoided in the response. This will help ChatGPT provide you with a more tailored and efficient response to your request.

- Example: generate guidelines for the use of ChatGPT at the City of Boston
- Example: write a letter requesting support for funding digital equity initiatives in the next budget session.
- Example: you can ask Chat GPT to generate letters that express points of view specified in the prompt. This might allow you to understand an issue from different perspectives.
- Example: **You can ask Generative AI to help you write a more effective version of a prompt. You can say “ help me write a better prompt to [insert the initial objective of the prompt].**

Do's:

1. Try to be specific in the prompt. If you give more context, the answer becomes more relevant.
2. Edit and review the content. Regardless of how the content was authored, you and the City will bear responsibility over its use in the public.

Don'ts:

1. Do not include confidential information in the prompt.
2. Do not rely on generative AI to provide accurate answers.

B

3. Do not use generative AI to create communication regarding sensitive topics. For instance, a renowned institution was criticized for using generative AI to write a press release regarding a shooting.

Drafting Content In Plain Language

Generative AI can help you write clearer and simpler language. You can use the prompt to indicate the reading level or audience for a text.

- Example: use ChatGPT or Bard to write a version of the Declaration of Independence of the United States for a person in elementary school.
- Example: use tools such as AISEO, Wordtune or others to modify a sentence. These tools are similar to a thesaurus but for sentences and often allow you to optimize for the length of the sentence, or the audience.

Do's:

1. Specify in the prompt if you have a specific audience in mind.
2. Try different prompts, or request different versions of the same sentence until you find what works best.
3. You can pass the output of the text by a readability app that can identify challenging sentences, as well as the reading level for the text.

Don'ts:

1. Do not include confidential information in the prompt.
2. Review the text to ensure that the language is inclusive and respectful. The models might use language or patterns that appear regularly, but that might exclude some people. For instance, a model might suggest: "Dear Sir/Ma'am" does not include non-binary people, and could be replaced with "Dear Colleague" or "Dear neighbor".

Drafting Content In Other Languages

AI can help you draft communications in another language. It is not well documented the extent to which ChatGPT and other models can use other languages, but users report over 50 languages being available for ChatGPT, including some native american languages.

- Example: use ChatGPT to translate these guidelines into Spanish and French, just ask "translate [your text] into Spanish and French."
- Example: you can ask generative AI in what language some text is written in, just ask "what language is [original language] written in?"

B

Do's:

1. Try different languages. ChatGPT, Bard and other models were trained using text from many languages. ChatGPT told me it didn't speak Quechua in Quechua!
2. You can also ask generative AI to perform similar tasks as the ones in this document in other languages, such as summarizing text, etc.

Don'ts:

1. Do not include confidential information in the prompt.
2. Do not use content generated in a language you do not understand before consulting someone with proficiency in the language. You still need to check for accuracy, bias, etc.
3. Language generated in other languages might be confusing to people who speak different regional dialects. Do not assume that some text will be easily understood by all speakers. Use the prompt to get regional diction.

Summarizing Text

Generative AI does a great job of summarizing longer pieces of text into summaries. If you have a few pages that you want to condense into a few bullet points, or you have been struggling with converting a long set of notes into a paragraph, these tools could be very helpful.

- Example: copy notes taken from a meeting to generate a short summary of the meeting.
- Example: summarize citizen comments in response to an engagement
- Example: write a paragraph summary of a 5 page report.
- Example: use Fathom, Wudpecker, or the transcript tools in Google Hangouts to transcribe audio into text. You can then summarize the text further using generative AI. This summarization is included in some of these tools.

Don'ts:

1. Do not include confidential information in the prompt: make sure you have deleted confidential information from your notes or other inputs.
2. If you plan on making a decision based on the summary, you should read the entire document(s) to make sure you did not miss or miss characterized the original document.

B

3. Be aware that the resulting summary might have biases as it will tend to present language that is more frequent in the data used to train the model. You can use changes to the prompt to enhance the results by suggesting that the result incorporates perspectives from marginalized groups. Even better, you can engage with some individuals in these communities to better understand their perspectives on the text generated.

Summarizing Audio

Coding/Programming

Generative AI can be great at producing snippets or even help you build more complex components of code.

- Example: write code in Python that extracts tables in a PDF into a Pandas data frame.
- This can make it possible for less technical people, including interns and student workers, to get to work on technical projects.

Do's:

1. Explore new languages and libraries - but you should understand the code and read the documentation of the relevant components before using it.
2. You might have to adjust parameters, and your environment to make the suggestions from the AI model work. Generative AI can help you get started, but often you will have to edit before the code works.

Don'ts:

1. Do not include confidential information in the prompt. As in development best practices: do not include passwords, confidential keys, or other proprietary information in your code or in the prompts.
2. You should understand what the code is doing before using it in production.
3. You should understand the use of new libraries and dependencies, and become familiar with vulnerabilities and other security considerations of using a language or a library.

Images, Audio, and Videos

Generative AI can produce images, audio, and videos based on prompts. This can support the creation of appealing or insightful communication resources.

B

- Example: make an image in a medieval style of residents connecting to the wifi in order to create appealing collateral for a digital equity campaign.
- Example: create a training video that walks residents on how to schedule a bulky item pick up, by providing the script of the video.
- Example: write a jingle or song to remind them to switch to Boston's Community Choice Electricity to switch to 100% renewable energy.

Do's:

1. Visual, audio and video communication can be a powerful tool to communicate with others and get across a message. Generative AI can empower you to use these tools beyond your artistic skills.
2. Use generative AI as a tool to create drafts or mock ups that allow you to communicate more effectively with graphic designers, videographers, and other creative workers.
3. Contact your department or agency's public information officer about the image, audio, or video before publishing or using it. They have expertise on best practices in accessibility, branding, etc.
4. Engaging with members of the Equity Cabinet, or community organizations that represent groups that might be referenced or impacted by this content. Getting their perspective, in a respectful way, can help you identify when content might be hurtful, discriminatory, or misinterpreted.

Don'ts:

1. Do not include confidential information in the prompt: make sure you have deleted confidential information from your notes or other inputs. Some confidential information could include: people's faces, people's voices, their identifications, license plates, etc. Particularly, those who have not provided their consent.
2. Make sure the outputs of the generative AI will not be offensive or harmful towards people, particularly vulnerable residents that are susceptible to harm including ethnic and racial groups, diverse gender individuals, and others.
3. Make sure that any content adheres to the [City's Brand Guidelines](#)

Resources

You can contact the Department of Innovation and Technology [doit@boston.gov] to learn more about generative AI.

You can also contact the Mayor's Office of Arts and Culture [arts@boston.gov] or to the Mayor's Office of New Urban Mechanics [newurbanmechanics@boston.gov] if you want to

B

discuss important questions about the impact of generative AI on the arts and on our society.

The following resources include external links. We do not endorse any one of these resources.

Reddit, ChatGPT sub reddit: <https://www.reddit.com/r/ChatGPT/>

A great explanation on the mathematical principles behind generative language models:

Stephen Wolfram (2023), "What Is ChatGPT Doing ... and Why Does It Work?," Stephen Wolfram Writings.

writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work.

AI Principles from Microsoft:

<https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimar6>

AI Principles from Google:

<https://ai.google/principles/>

NIST AI Risk Framework:

<https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

A critical analysis of large language models (major paper that predicted much of the harms/risks we are experiencing now)

<https://dl.acm.org/doi/10.1145/3442188.3445922>

Acknowledgements

The development of these guidelines has benefited from the contributions of academic, community, and City of Boston team members.

Special thanks to Beth Noveck, Director of the Burnes Center for Social Change at Northeastern University; Saiph Savage, Director of the Civic AI Lab at Northeastern University; Catherine D'Ignazio, Director of the Data + Feminism Lab at MIT; Kimberly Lucas, Professor of the Practice at Northeastern University; Mitch Weiss, Professor at Harvard Business School; Alejandro Jimenez Jaramillo; Michael Evans from the Mayor's Office of New Urban Mechanics; Jerry Kelley, project manager at the Department of Innovation and Technology, Kerry Jordan, Chief of Staff at the Department of Innovation and Technology.