NAVIGATING UNCHARTED WATERS

GENERATIVE AI GUIDANCE FOR ORGANIZATIONS

SEPTEMBER 2023

U.S. Cybersecurity Group

ASPEN DIGITAL
aspen institute

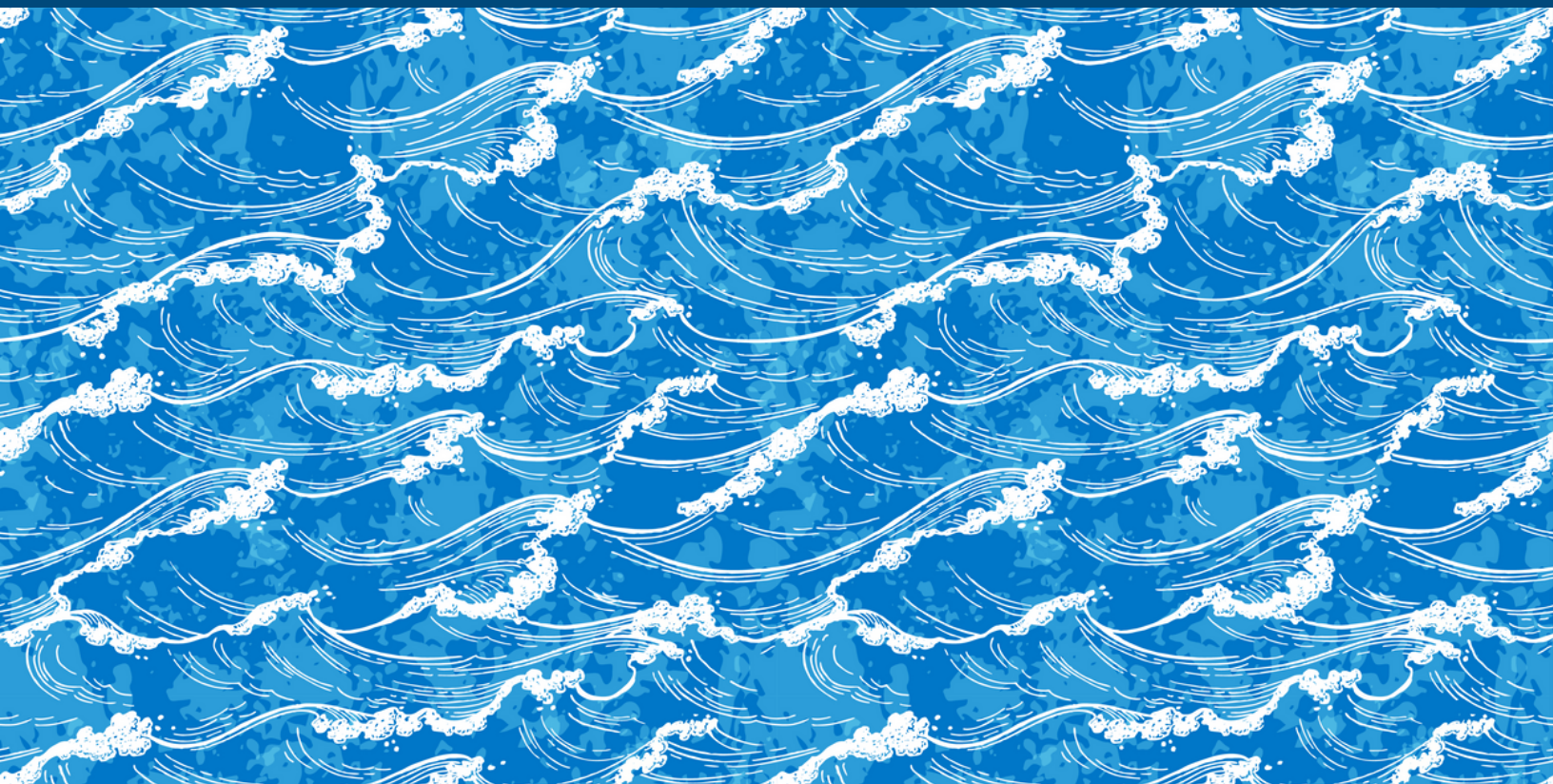


TABLE OF CONTENTS

INTRODUCTION

In recent months, the growing use of Generative Artificial Intelligence (GenAI) technology – including general-purpose and publicly available foundational models like GPT-4, LLaMA, and DALL-E – has captured the public's attention and dominated news headlines.It has also generated widespread concern about the potential consequences that could emerge from the rapid, unrestricted use of GenAI-based tools. While AI has been around for years, the public availability of large, general-purpose foundational models combined with a significantly lower barrier to access the power of those models is new. This has introduced risks and questions that public and private sector leaders are only beginning to consider.

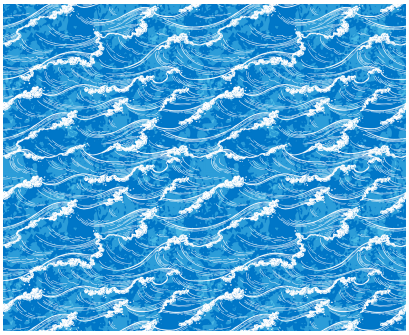
To help add clarity to this rapidly unfolding conversation, the Aspen US Cybersecurity Group convened experts to draft high-level guidance on how companies can inform employee use of openly available GenAI technology. What follows is a template guidance document that the Group developed for use by a broad range of organizations. Organizations can revise to fit their specific needs and share with employees and business units that use, rely on, or are considering how employees can or should use openly available GenAI-based solutions (as distinct from use of company-approved GenAI enterprise products). The guidance is targeted towards general employee populations and is designed to serve as a baseline document for companies to adapt to fit their specific organizational needs. No organization should adopt it without first reconciling it against its own unique policies, procedures, and legal and regulatory guidelines.

Importantly, the guidance provided below is relevant as of this report's publication in September 2023. The GenAI playing field is changing quickly, and before using this document, organizations should assess whether it is still relevant to their specific organizational needs and to the current state of technology, law, and regulation. **We encourage organizations to use any portion of this guidance that is relevant to their needs and to modify it as they see fit. No attribution is necessary.**

BACKGROUND AND DOCUMENT PURPOSE

Generative AI (GenAI) is the latest development in the field of AI and is an advanced form of machine learning that creates an opportunity to produce new content using large language models (LLMs) that have been trained on large amounts of data, including audio, text, images, code, simulations, and videos. GenAI, as with all other forms of AI, has the potential to transform our business, our industry, and our competition.

We encourage all Staff to explore and innovate with GenAI, but to do so responsibly.[1] This means understanding GenAI's risks and limitations as well as abiding by the general guiding principles outlined below. We require that you use GenAI-based tools consistent with this Guidance. If you are in doubt, please consult with your supervisors and leaders. As we continue to define further use cases for our businesses and operations, we will provide additional guidance. **This guidance is current as of September 2023.**



GenAI has the potential to transform our business, our industry, and our competition.

[1] This guidance addresses how you can use openly available instances of GenAI, such as ChatGPT, Bard, or Bing AI, and not company developed or approved enterprise products. For those tools, please follow guidance specific to them.

RISKS AND LIMITATIONS

Before beginning to use GenAI, it is important to understand its limitations and associated risks. Our ability to control and monitor the following aspects of model implementation is key to our ability to exercise responsible AI practices.

Foundational GenAI models are built on a very large pool of training data, but these sources are still finite. As a result, that data, its uses, and outputs may:

- **Not be accurate. Outputs may be inaccurate, unverifiable, and possibly out of date. This type of erroneous output includes overly confident responses not merited by training data, sometimes referred to as AI "hallucinations."**
- **Infringe third party intellectual property rights. Due to the way GenAI creates outputs based on the data available to it, results could lead to intellectual property rights infringements.**
- **Breach privacy rights. Outputs may violate privacy and data protection laws.**
- **Include data which is biased or discriminatory. This may lead to outputs which in turn amplify that bias.**
- **Disclose confidential information. Data provided in prompts is used to train the model and can be incorporated into the model itself. This could result in the disclosure of confidential information.**
  - **As a result, it is important to understand the full scope of our data classification policies, and what distinguishes "confidential" from "public" information.**

The data issues that can occur with openly available models generally do not exist with our internal GenAI models, which are typically built for a specific purpose and which we can manage. Therefore, while this technology brings great potential, when using openly available models you must consider the risks to our business and reputation in every engagement.

GUIDING PRINCIPLES AND DOS & DON'TS

You are responsible and accountable for your use of GenAI. To help guide your use of GenAI-based tools, we provide below basic principles that should inform your thinking as you prepare to use this technology. We also provide several "dos and don'ts" that we believe to be leading practices. However, it is important to remember that the GenAI landscape is still rapidly developing, and the capabilities, norms, and rules around use of the technology are not yet solidified. As a result, whether and how these principles and best-practices apply to your specific facts might change with time. If in doubt, ask.

You are responsible and accountable for your use of GenAI.

GUIDING PRINCIPLES

As you begin to use GenAI you will need to:

- Identify existing policies, processes, tools and frameworks that apply to the use of GenAI, including in the area of corporate security, technology, privacy, data handling, copyright and third-party risk management.

- Develop a standardized risk assessment framework that incorporates risks due to limitations of the technology as well as the risks of applying the technology in your specific business process, as you would with any third-party software or tool.[2]

- Upskill teams so they are able to detect issues unique to GenAI, such as concept drift and hallucinations.

- Implement governance measures that address key areas of risk such as cybersecurity, privacy, legal, compliance, and regulatory risks; and consider establishing a Governance Council to facilitate multi-stakeholder engagement and provide senior oversight over the use of GenAI.

- Protect proprietary data and intellectual property via technical and contracting mechanisms.

- Monitor the use of the technology for compliance with evolving laws as well as policies and procedures we have implemented.

[2] You can consider models such as the National Institute of Standards and Technology's AI Risk Management Framework, https://www.nist.gov/itl/ai-risk-management-framework, or the International Organization for Standardization's guidance on managing AI risks, ISO23894, https://www.iso.org/standard/77304.html.

GUIDING PRINCIPLES (CONTINUED)

- Learn how GenAI works. Several publicly available sources explain the concepts behind GenAI and how it operates, and having a basic understanding of these principles will help you ask the right questions when working with these technologies.[3]

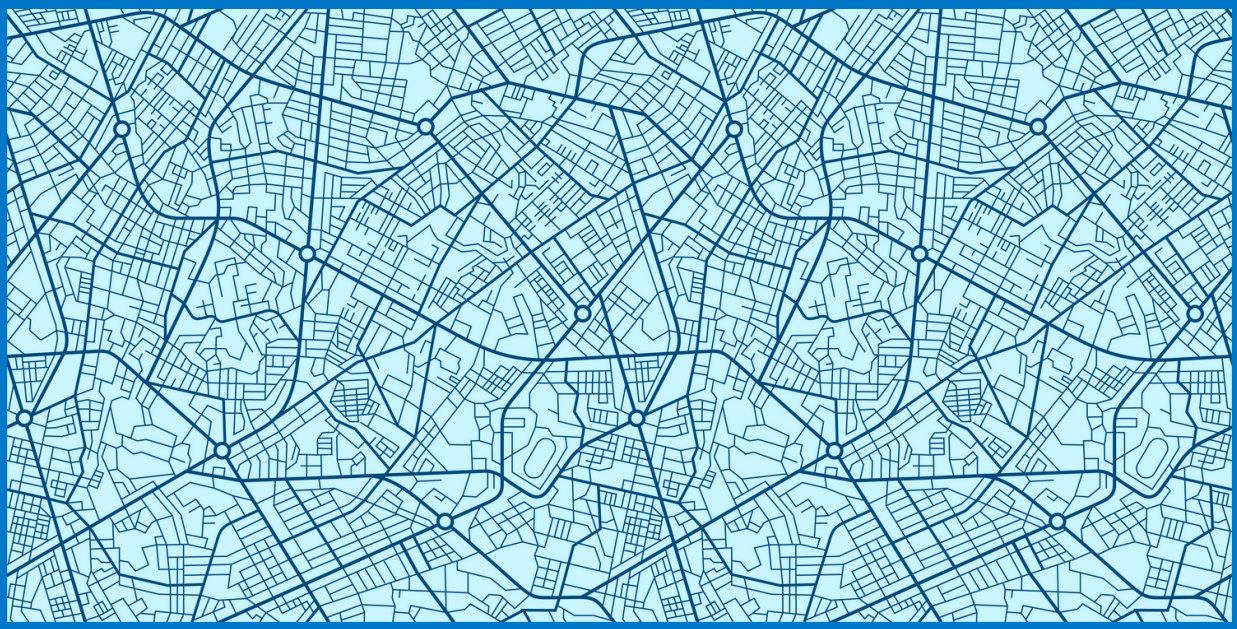
- Realize that GenAI may not necessarily be the right tool for every task or problem. It's okay to decide that GenAI is not appropriate in certain situations.

- Ask questions of your peers, supervisors, and leaders and share your knowledge with them. Many companies, including ours, are creating new policies, procedures, and infrastructure to guide their use of GenAI in a manner that preserves data privacy, security, confidentiality, and intellectual property—and we all benefit from ongoing dialogue and questions.

[3] Aspen Digital's emerging technologies team has developed a new primer on this topic, Intro to Generative AI, https://www.aspendigital.org/report/intro-to-generative-ai/.

DOS AND DON'TS

The following are examples of acceptable and unacceptable uses of publicly available GenAI technologies. Nonetheless, these dos and don’ts are built on the following baseline assumptions: (a) you only provide non-personal, non-confidential, and non-proprietary and/or public information to the foundational models; (b) no output is used for a commercial purpose; and (c) any code or model generated is not put into production unless there is specific guidance and approvals allowing you to do so.



DOS

- Conduct queries and searches multiple times, and in multiple ways, to understand variance in outputs.
- Create training, tutorials, and images for internal purposes, based on textual instructions (e.g., for inspiration in brainstorming, team meetings, internal presentations).
- Debug code using test samples not sourced or derived from company code.
- Experiment with code creation capabilities to learn how GenAI can create code and to evaluate the potential utility of such code.
- Expand perspectives and new ideas for inspiration and brainstorming.
- Comply with existing approval processes as applicable. This guideline includes submitting to the governance council [if established] for review and for review and approval all GenAI use cases that leverage foundational models, such as ChatGPT. However, use cases leveraging single-purpose models (e.g., translation models), or models that are built internally, may follow standard processes and principles, such as AI Governance[4] and Privacy by Design.
- If you are working with a GenAI use case that has been approved through the governance process, you still must:
  - Exercise human oversight over all uses of GenAI.
  - Document how you used GenAI and any results you discarded, as well as why a specific output was used.
  - Communicate about the instances in which GenAI was used, including to the colleagues exposed to the output.
- Turn off chat history if you plan to make any use of the chat.

[4] AI governance is a set of principles, policies, and practices that govern the development, use, and ethical implications of artificial intelligence (AI).

DON'TS

When exploring GenAI with commercially available tools, don't:

- Provide or input any proprietary, confidential, or sensitive data, including any confidential, personal, and/or customer information into commercially available or public instances of GenAI, unless otherwise set forth in these guidelines, or if the software has already been reviewed and approved for your specific use case.

- Use any outputs (or otherwise utilize any GenAI tool) in any commercial product or solution (including proofs of concept or market tests and marketing material) or otherwise in a commercial context unless specifically approved.

- Provide any output from GenAI to external third parties.

- Retain or otherwise partner with any service provider that leverages GenAI without reviewing their use of the technology. While the vendor may have done due diligence, this new technology has launched numerous tools that may not have been vetted during that process.

- Use GenAI to write or debug any software code used in any production system or infrastructure unless there is additional, specific guidance and approvals allowing you to do so.

- Share any outputs externally. If you are sharing internally for non-commercial uses you still must have any outputs reviewed for privacy and legal compliance, accuracy, bias or offensive material, as well as label the material as being created by AI.

Individuals are encouraged to cite this report and its contents. In doing so, please include the following attribution:

"Navigating Uncharted Waters: Generative AI Guidance for Organizations." Aspen Digital, a program of the Aspen Institute, Sept. 2023. CC BY-NC. https:// creativecommons.org/licenses/ by-nc/4.0/. https://www.aspendigital.org/report/navigating-unchartered-waters/