CITY USE OF ARTIFICIAL INTELLIGENCE (AI)

SCOPE

This regulation applies to all city departments, agencies, and entities involved in the development, procurement, deployment, or use of AI technologies. It also covers any third-party contractors or vendors that work with the City of Boise on AI-related projects.

REGULATION OWNER

Business Owner: Director, Innovation and Performance

Technical Owner: Director, Information Technology

1. REGULATION PURPOSE

   The purpose of this Regulation is to establish guidelines and principles for the responsible and ethical use of Artificial Intelligence (AI) technologies within the City of Boise. This document aims to harness the potential of AI to enhance public services, improve efficiency, and drive innovation, while ensuring that its deployment upholds the values of transparency, accountability, fairness, privacy, and inclusivity.

   Of particular interest is Generative AI, which is a set of relatively new technologies that leverages very large volumes of data along with some machine learning techniques to produce very sophisticated 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), and often cannot be distinguished from human-generated content. These tools are evolving rapidly and are still the subject of active research to improve the technical community's understanding of how they actually work, and identify the potential impacts, both good and bad, of their use in society. These tools are not actual intelligence in the human sense. Rather, they are very sophisticated analytical models that respond to an input request by predicting what the language, text, or video will best satisfy it.

   Because of Generative AI's impact and potential usefulness, as well as its risks, the City of Boise is adopting the guidelines presented in this Regulation to serve as an

interim resource for City employees. But this regulation applies equally to all forms of AI.

Some of the content of this document was in fact generated by ChatGPT v3.5, and subsequently revised and extended by the authors.

2. REGULATION STATEMENT

Usage of Artificial Intelligence technologies within the City of Boise shall ensure that confidential City information is not compromised, that generated content is always validated by a person before publishing, and that responsible City staff know, and are comfortable with, the extent to which data we provide to an AI tool may be shared with non-City audiences.

AI is a tool, much like a Google search but more sophisticated. Nevertheless, ultimately the people using the tools are responsible for the outcomes. City users of AI tools must remain mindful of this reality. Technology enables our work; it does not replace our judgment nor our accountability.

3. DEFINITIONS

Generative AI: Generative AI refers to AI systems capable of generating new content, such as images, text, audio, or videos, that imitates or is indistinguishable from human-created content. It includes technologies such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and other deep learning-based models. Generative AI models are trained on large datasets and learn the underlying patterns and structures of the data. They can then generate new content that is similar to the examples they were trained on. The generated content can take various forms, such as text, images, music, or even videos.

Narrow or Weak AI: Narrow AI refers to AI systems designed to perform a specific task or a set of specific tasks. These systems excel at the task they are programmed for but lack the ability to generalize beyond their specific domain. Examples include voice assistants, image recognition systems, and recommendation algorithms.

Machine Learning: Machine Learning (ML) is a subset of AI that focuses on the development of algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. ML algorithms learn patterns and relationships from training data and use that knowledge to make predictions or take actions.

Natural Language Processing: Natural Language Processing (NLP) focuses on enabling computers to understand, interpret, and generate human language. NLP involves tasks such as text classification, sentiment analysis, machine translation, and question answering systems. NLP techniques are used in various applications like chatbots, virtual assistants, and language translation tools.

Training Dataset: The "corpus" of information used to train a generative AI tool in the questions it may expect to be posed and how to formulate its response. For City-

specific generative AI tools, the software vendor may provide a baseline corpus of data, to be extended & refined by City employees to tailor the responses to ensure appropriateness and relevance. For publicly available generative AI tools, the training data may include arbitrary sources from across the Internet and will not be within City control. For example, ChatGPT is trained on a large corpus of text data from the internet. It learns patterns, relationships, and statistical properties of language by processing billions of sentences and uses that knowledge base to formulate its responses. In this case, note that there is no guarantee that its training dataset does not include false or misleading data.

Prompt: Prompts are the inputs or queries that a user or a program gives to a Generative AI tool in order to elicit a specific response. Prompts can normally be expressed as natural language questions and can be successively refined to tailor the response provided. The prompts and responses may be used by the AI tool to expand its knowledge base, so care must be taken not to expose any sensitive data in the prompt input.

4. FOUR CARDINAL RULES FOR AI USAGE

Information Technology develops, delivers, operates, and supports solutions that help City departments efficiently and effectively deliver equitable and responsive services to the public. We recognize that we are entrusted with responsibly stewarding the public's data and protecting our IT systems. We see the emergence of generative AI as providing both opportunities that can help us deliver our services, but it also has risks that can threaten our responsibilities.

Because the generative AI field is emergent and rapidly evolving, the potential policy impacts and risks to the City are not yet fully understood. The use of generative AI systems within the city can have unanticipated and unmitigated impacts.

A. City employees shall obtain IT Department approval before accessing or acquiring an AI product.

This is a standard operating practice in the city for all new or non-standard technology. The IT Department will maintain a list of approved AI products; anything not already on the approved list must be reviewed and approved by IT prior to introduction or use.

For introduction of new commercial software that includes an AI capability, requesters will provide answers to the following questions:

- What function(s) does the AI component provide or support?
- What data does it use for its training dataset?
- Is it public data, data shared with other vendor customers, or City-accessible only?
- Who creates the training dataset? Who maintains it?

- Does anyone outside the city have access to our data, including user submissions to prompts and the responses?
- Does the City have access to anyone else's data? If so, whose, and why?
- Can we disable the AI component if we so choose? What impact would this have on system functionality?

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

While generative AI can rapidly produce clear prose, the information and content might be inaccurate, outdated, offensive, or simply made up. So, it's essential to validate that the output of generative AI systems is accurate, properly attributed, free of someone else's intellectual property, and free of unintended or undesirable instances of bias and potentially offensive or harmful material. It is your responsibility to verify that the information is accurate and appropriate by independently researching claims made by the generative AI tool.

C. Reference AI usage when you use it for significant communications with the public or for other important purposes.

Like we did for this document – see Paragraph 1 above.

You may consider the sensitivity of your use case; for example, it's important to be transparent when creating a communication such as a PowerPoint presentation or a memo, but not nearly as important when drafting a thank-you email.

Even when you use AI minimally, disclosure builds trust through transparency, and it might help others catch errors. So be upfront about crediting your usage of a generative AI "assistant", and ideally include the version and type of model you used. For example, "This document was generated by ChatGPT 3.5 and edited (heavily | moderately | lightly) by John Doe".

D. Do not share sensitive or private information in the prompts!

Data provided in generative AI prompts, especially in the publicly accessible platforms such as ChatGPT and Bard, are used by the companies that power these systems to continuously grow their tool's knowledge base. Even generative AI components imbedded in other 3rd-party software may exhibit the same behavior. So be very careful about the information you provide in the prompt! Any information that includes personally identifying information about our employees (see Employee Handbook Personnel Files Regulation) and/or community members could inadvertently be shared with others.

Basically, if you wouldn't share information with the public, avoid sharing it in the prompt. If you have a use case that requires sensitive information to be used with a generative AI, contact IT so we can help you establish an appropriate solution.

5. REGULATION REVIEW AND UPDATES

    A. Review Period

    This regulation will be reviewed annually, and updates will be made as necessary based on technological advancements and emerging best practices.

    B. Continuous Learning and Improvement

    The City of Boise will promote continuous learning and improvement in AI governance and ensure that AI policies remain at the forefront of ethical and technological standards.

6. GOVERNANCE FRAMEWORK
    A. Roles and Responsibilities

    The Information Technology department will oversee all artificial intelligence initiatives, in concert with representatives from sponsoring departments.

    Each city department involved in AI projects will consult with Information Technology to coordinate and ensure compliance with this regulation.

    B. Oversight and Evaluation Mechanisms

    The IT department will establish an ongoing evaluation and monitoring process to assess the effectiveness, fairness, and safety of AI systems. Periodic audits will be conducted to identify and address any issues related to compliance and ethical considerations.

7. CONCLUSION

    The City of Boise AI Regulation reflects our commitment to harnessing AI technologies responsibly and ethically to enhance public services, foster innovation, and benefit our community. Through transparent, accountable, and inclusive AI practices, we aim to create a thriving and sustainable future for all citizens. This regulation sets the foundation for a human-centric approach to AI governance that prioritizes the well-being and values of our city's residents.

8. RELATED INFORMATION

    Employee Handbook – 4.45a Personnel Files Regulation. Defines requirements for protecting employee privacy.