Artificial Intelligence Guiding Principles
March 25, 2024

These guiding principles are intended to provide a framework that counties can use to develop policies that further establish county practices for the secure, ethical, effective, and responsible use of Artificial Intelligence.

Artificial Intelligence (AI), also known as machine intelligence, is the simulation of human intelligence processes. Counties are committed to using AI technology responsibly and ethically to improve processes, enhance services to residents, and support employees to do their best work. As counties increasingly integrate AI into operations, we realize the benefits of AI, while working diligently to address and mitigate the risks associated with its use.

AI tools are developing at an exponential rate. Counties should regularly review, update, and train users on their AI-related policy to keep it aligned with ethical and legal standards and technological advancements, as needed.

These Guiding Principles should apply to all county employees, contractors, and any other third-party individuals or entities who have access to AI technologies or are involved in using AI tools or platforms on behalf of Counties.

Accuracy:

1. The County strives to deploy algorithmic systems that have a high accuracy rate. The County will consider the impacts of AI on the authenticity of data supplied to them by outside sources. Users are responsible for ensuring the accuracy and validity of content that is ultimately used in County business and will verify content generated by AI.

   Examples of what the above bullets address:

   o False information: AI-generated content seems authoritative, but its "facts" may be inaccurate or even fabricated. Any statements presented as fact should cite a credible source rather than the AI system.

   o Outdated data: Be sure that any information presented as current in generated content is up to date. For example, Chat GPT-3.5's knowledge base was last updated in September 2021. AI-generated content does not always include its most recent Refresh date.

o Relevance: Generated content may not be applicable or relevant to the County's specific needs and users should ensure that all work products address the County's unique character and needs.

Transparency:

1. When conducting County business, users should disclose and attribute the use of AI-generated content for more substantive work (to be defined by the County).
2. The County discloses to the general public when AI is used in County operations, sharing its benefits and potential implications to foster transparency and understanding. Clearly communicate the role of AI in specific processes and services, assuring residents that their interests and rights are respected.
3. Inform the public when they are interacting with AI (e.g., conversational AI and chatbots), and provide an "opt out" alternative, whenever possible.
4. The use of AI systems should be explainable to those who use and are affected by their use. Commonly understood explanations of how AI features work will be maintained and available to convey a concrete sense of how or why a particular result has been arrived at based on inputs.

Examples of what the above bullets address:

o AI-generated content used in County business should be clearly labeled as such, and details of its review and editing process (how the material was reviewed, edited, and by whom) should be provided. This provides transparent authorship and responsible content evaluation.

o <u>Sample disclosure line</u>: *This document was drafted by [insert AI tool, e.g., Google Bard/ChatGPT] using the following prompt: "Does the inclusion of cybersecurity standards in contract Terms and Conditions result in better contractor security compliance?" The document was reviewed and edited by [insert name(s).]*

Equity:

1. The County only uses algorithmic tools in alignment with its commitment to equity and reducing racial and socioeconomic disparities. Documents or work produced using AI should be screened for biases and potentially offensive material. Users should take care to ensure that output (e.g., documents and communication produced using AI) used in County business is free of bias and toxicity and reflects the County's commitment towards equity and providing culturally responsive services.

2. Documents and communication should reflect respect for the County's workforce and community and should be screened to maintain the use of AI results in equitable services to all residents, regardless of their background, by considering diverse data and perspectives.
3. The County maintains human oversight over AI, especially in decision-making processes to prevent undue reliance on automated decisions and where bias could play a role.

Examples of what the above bullets address:

o Biases: Because AI knowledge bases contain so much unevaluated information, they may reflect political, racial, sexual, cultural, economic, or other biases.
o Potentially offensive material: Some AI systems may attempt to keep generated content within the bounds of language and ideas acceptable to most people, but their attempts don't always succeed. Other systems make no such attempts.

Accountability:

1. The County will strive to certify that data or content used by AI complies with copyright laws, and whenever possible, use data that is explicitly licensed for its intended use or that is in the public domain. The County will avoid using AI to generate content that directly copies copyrighted material.
2. The County will seek to procure AI products that have built-in protections to limit the use of copyrighted data.
3. AI vendors will be selected based on their reputation, adherence to data protection standards, and commitment to ethical AI principles. The County should verify that vendor contracts disclose the use of AI and explicitly define data usage, data ownership, security measures, and compliance with privacy regulations.
4. AI tools will not be used for any illegal, harmful, or malicious activities.
5. When creating accounts with AI providers for the use of AI tools, users should use their official County email address and a unique password for better security and accountability.
6. Each system will have responsible parties (to be identified by the County) that are accountable for the use and outcomes of AI systems.

Examples of what the above bullets address:

- It is not always easy to determine whether AI-generated content contains copyright-protected material. If you are sure the material is protected, obtain rights to it, or do not use it. If you are not sure, do not use the material.

Utility:

1. The County adopts algorithmic tools that have demonstrated a positive impact to residents. AI can be used to maximize the County's effective and efficient use of evolving technology to support its operations and to make services better, more just, and more efficient.
2. AI should not be used as a definitive source of medical, legal, or strategic advice; instead, it should be regarded as a tool that can be responsibly used to facilitate the County's efforts to provide collaborative and affordable public services that are beneficial and responsive to our wider community. Users should be conscientious about how AI tools are used to inform decision-making processes.

Confidentiality:

1. The County ensures a privacy-centric approach in all algorithmic systems. Never enter personal, confidential, or sensitive information into an AI system. Always assume any information you enter into an AI system may be subject to a California Public Records Act (CPRA) request.

   Examples of what the above bullets address:

   - It is documented that AI tools may store user data for use with future models.