Artificial Intelligence (AI) Policy 1.7.12

PURPOSE

This policy establishes a governance structure that allows the City of San José (hereafter referred to as "City") to utilize Artificial Intelligence (AI) and AI systems (systems) while providing the necessary safeguards for purposeful and responsible use.

The key objectives of the AI Policy are to:

- Provide guidance that is clear, easy to follow, and supports decision-making for the staff, interns, consultants, contractors, partners, and volunteers who may be purchasing, configuring, developing, using, maintaining, or leveraging AI to provide services to the City;
- Ensure that the use of AI systems adheres to the Guiding Principles with regard to how systems are purchased, configured, developed, operated, or maintained;
- Define roles and responsibilities related to the usage of AI;
- Establish and maintain processes to assess and manage risks presented by AI;
- Align the governance of AI with existing data governance, security, and privacy measures in accordance with the City's Information and Systems Security Policy and City Council's Digital Privacy Policy;
- Define prohibited uses of AI systems;
- Establish "sunset" procedures to safely retire systems that no longer meet the needs of the City; and
- Define how AI may be used for legitimate purposes in accordance with applicable local, state, and federal laws, and existing agency policies.

AI systems and the data contained therein will be purchased, configured, developed, operated, and maintained using the City's AI Handbook, which will be managed by the Chief Information Officer (CIO).

SCOPE

This policy applies to:

1. All systems deployed by the City; and
2. Staff, interns, consultants, contractors, partners, and volunteers who may be purchasing, configuring, developing, using, or maintaining the AI or who may be leveraging systems to provide services to the City (collectively referred to as "users").

GUIDING PRINCIPLES FOR RESPONSIBLE AI SYSTEMS

These principles describe the City's values with regard to how AI systems are purchased, configured, developed, used, or maintained.

1. **Effectiveness:** Systems are reliable, meet their objectives, and deliver precise and dependable outcomes for the utility and contexts in which they are deployed;
2. **Transparency:** The purpose and use of systems is proactively communicated and disclosed to the public. A system, its data sources, operational model, and policies that govern its use are understandable and documented;
3. **Equity:** Systems deliberately support equitable outcomes for everyone. Bias in systems is

effectively managed with the intention of reducing harm for anyone impacted by the system's use;

4. **Accountability:** Roles and responsibilities govern the deployment and maintenance of systems, and human oversight ensures adherence to relevant laws and regulations;
5. **Human-Centered Design:** Systems are developed and deployed with a human-centered approach that evaluates AI powered services for their impact on the public;
6. **Privacy:** Privacy is preserved in all AI systems by safeguarding personally identifiable information (PII) and sensitive data from unauthorized access, disclosure, and manipulation in accordance with the City Council's Digital Privacy Policy;
7. **Security & Safety:** Systems maintain confidentiality, integrity, and availability through safeguards in accordance with the City's Information and Systems Security Policy. The integrity of information into and out of the City is maintained in light of fake AI-generated content. Implementation of systems is reliable and safe, minimizing risks to individuals, society, and the environment; and
8. **Workforce Empowerment:** Staff are empowered to use AI in their roles through education, training, and collaborations that promote participation and opportunity.

RESPONSIBILITIES

Several roles are responsible for enforcing this Policy, outlined below.

- The Information Technology Department Director / Chief Information Officer (CIO) is responsible for directing technology resources, policies, projects, services, and coordinating the same with other departments. The CIO shall designate the City Information Security Officer (CISO) and City Digital Privacy Officer (CDPO) to actively ensure the security, resilience, privacy, and policy compliance of the systems used by the City.
- The CISO and CDPO are responsible for recommending updates to this policy and the AI Handbook.

POLICY

When purchasing, configuring, developing, using, or maintaining AI systems, users will:

1. Uphold the Guiding Principles for AI systems outlined above**;**
2. Conduct an AI Review to assess the potential risk of the AI system. The CDPO or designee is responsible for coordinating review of AI systems used by the City as detailed in the AI Handbook;
3. Obtain technical documentation about AI systems. The Finance Department, or other department overseeing the purchase of an AI system, is responsible for requiring vendors to disclose AI usage and to provide technical documentation (e.g., via the AI FactSheet as defined in the Terms and Definitions section, below) at the request of the CDPO; and
4. In the event of an incident involving the use of the AI system, follow the City's AI Incident Response Plan in accordance with the Information and Systems Security Policy. The CISO is responsible for overseeing the security practices of AI systems used by or on behalf the City.

Additionally, Finance is required to ask vendors to disclose the use of AI in procurement solicitations and to comply with the Requirements for AI Systems upon the request of the CDPO or designee.

Prohibited Uses

The use of certain AI systems is prohibited due to the sensitive nature of the information processed and severe potential risk. This includes, but is not limited to, the following prohibited purposes:

- Real-time and covert biometric identification;
- Emotion analysis, or the use of computer vision techniques to classify human facial and body movements into certain emotions or sentiment (e.g., positive, negative, neutral, happy, angry, nervous);
- Fully automated decisions that substantially impact the rights or safety of individuals with no meaningful human oversight;
- Social scoring, or the use of AI systems to track and classify individuals based on their behaviors, socioeconomic status, or personal characteristics; and
- Cognitive behavioral manipulation of people or specific vulnerable groups.

If staff become aware of an instance where an AI has caused harm, staff must report the instance to their supervisor, the CDPO, and the Office of Employee Relations no later than 24 hours after discovery.

Sunset Procedures

If an AI system operated by the City or on its behalf ceases to provide a positive outcome to the City as determined by the staff or CDPO, then the City must halt the use of that system unless express exception is provided by the CIO. If the abrupt cessation of the use of that AI system would significantly disrupt the delivery of services, a gradual phased out approach must be approved by the CIO before sunsetting. All measures to minimize the impact and recovery must be considered in the termination or phase out protocol, including but not limited to:

- Ownership and future access of data;
- Portability of the AI model, algorithm, and/or data; and
- Impact to services, users, and residents.

Public Records

The City must consider applicable public records laws before implementing an AI system and must comply with the City's Open Government and Ethics Provisions and the California Public Records Act. More information can be found in City's Administration Policy Manual 6.1.4 Open Government Policy.

Policy Enforcement

All employees and agents of the City, whether permanent or temporary, interns, volunteers, contractors, consultants, vendors, and other third parties operating AI systems on behalf of the City are required to abide by this Policy and the associated AI Handbook.

VIOLATIONS OF THE AI POLICY

Violations of any section of the AI Policy, including failure to comply with the AI Handbook, may be

subject to disciplinary action, up to and including termination. Violations made by a third party while operating an AI system on behalf of the City may result in a breach of contract and/or pursuit of damages. Infractions that violate local, state, federal or international law may be remanded to the proper authorities.

TERMS AND DEFINITIONS

Artificial Intelligence: "Artificial intelligence" or "AI" is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action.

Algorithm: A series of logical steps through which an agent (typically a computer or software program) turns particular inputs into particular outputs.

System: Any software, sensor, or process that uses AI to automatically generate outputs including, but not limited to, predictions, recommendations, or decisions that augment or replace human decision-making. This extends to software, hardware, algorithms, and data generated by these systems used to automate large-scale processes or analyze large data sets.

AI Fact Sheet: A template that captures the "nutrition facts," or essential technical details, of an AI system. Vendors are expected to complete the AI Fact Sheet during the procurement process. The AI Fact Sheet is a critical document that provides technical information needed to adequately understand, evaluate, and use AI systems. Maintained by the Information Technology Department.

Approved:

/s/	6/28/2024
Khaled Tawfik Information Technology Department Director / Chief Information Officer	Date

Approved for posting:

/s/	6/28/2024
Jennifer A. Maguire City Manager	Date