

# CITY OF DETROIT

## MUNICIPAL ARTIFICIAL INTELLIGENCE POLICY

**MEMORANDUM TO:** The Honorable Mayor and City Council, City of Detroit, MI

**FROM:** Chief Artificial Intelligence Officer (CAIO), City of Detroit

**DATE:** March 3, 2026

**SUBJECT:** Draft Municipal Artificial Intelligence (AI) Policy for the City of Detroit

---

Honorable Mayor and City Councilmembers,

As the City of Detroit continues to modernize its municipal services, artificial intelligence (AI) offers unprecedented opportunities to improve efficiency, expand staff capacity, and deliver better outcomes for our residents. However, the unregulated adoption of AI has the potential to erode privacy, amplify existing societal biases, and widen equity gaps. To harness these benefits while mitigating the risks, I present the following draft Municipal AI Policy for your review and approval.

This document has been modeled to reflect a comprehensive and formal regulatory framework drawing on best practices from leading jurisdictions including the European Union, the State of California, the State of Mississippi, the City of Seattle, the District of Columbia, and the City of Boston.

---

# CITY OF DETROIT MUNICIPAL ARTIFICIAL INTELLIGENCE POLICY

## **1.0 PURPOSE AND OBJECTIVES**

The purpose of this policy is to establish the standards that City of Detroit departments, staff, and authorized representatives must observe when acquiring, developing, and deploying Artificial Intelligence (AI) solutions in municipal service delivery and decision-making.

The primary objectives of this policy are to:

- Define clear processes that enable the city to solve municipal challenges and enhance service delivery through responsible AI use.
- Provide actionable guidance for city staff, vendors, and partners who purchase, configure, operate, or maintain AI systems.
- Detect, assess, and manage risks presented by AI tools, including the prevention of strictly prohibited uses.
- Ensure that the adoption of AI reflects Detroit's commitment to equity, transparency, and the protection of resident data.

## **2.0 SCOPE**

*This policy applies to all City of Detroit departments, full-time and part-time staff, interns, consultants, vendors, contractors, partners, and volunteers who provide City services or otherwise act on behalf of the City. This includes the procurement and use of both free-to-use software and enterprise software-as-a-service tools.*

### 3.0 DEFINITIONS

For the purposes of this policy, the following definitions apply:

- **Artificial Intelligence (AI):** An engineered or machine-based system that, for explicit or implicit objectives, infers from the input it receives how to generate outputs (such as predictions, content, recommendations, or decisions) that can influence physical or virtual environments.
- **Generative AI:** A class of computer software and systems that use large language models, algorithms, deep-learning, and machine learning to generate new content, including text, images, video, audio, and software code, based on the patterns and structures of input data.
- **High-Risk AI:** AI systems deployed in sectors deemed critical to safety or fundamental rights, including critical infrastructure, law enforcement, employment/worker management, and the administration of justice.
- **Automated Decision System:** A computational process derived from machine learning or AI that issues simplified outputs (example: a score or classification) used to assist or replace human discretionary decision-making that materially impacts natural persons.
- **Algorithm:** A series of logical steps through which an agent (typically a computer or software program) turns specific inputs into specific outputs.

### 4.0 CORE PRINCIPLES AND RULES

The City of Detroit will evaluate and govern all AI systems based on the following core principles, translated into mandatory operational rules:

#### 4.1 Accountability and Human Oversight

- **Principle:** AI is a tool to enhance, not replace, human judgment. Mechanisms must be in place to track, audit, and address adverse effects.

- **Rule (Human-in-the-Loop):** Outputs of Generative AI systems must be reviewed and validated by a human prior to use in any official City capacity. Higher-risk decisions require explicit supervisor approval. AI tools must be designed to guarantee human control and override capabilities.

#### 4.2 Transparency

- **Principle:** The processes, decisions, and outcomes of AI tools should be understandable and available to the public to build trust.
- **Rule (Attribution and Disclosure):** All text, images, and videos created substantively by Generative AI must be clearly attributed to the specific AI system used. The city must inform residents when they are interacting with an AI tool (a chatbot) instead of a human employee.

#### 4.3 Privacy and Data Protection

- **Principle:** Individuals' personal data must be protected from unauthorized use, and AI implementation must comply with existing data privacy policies.
- **Rule (Data Ingestion Limitations):** City employees shall not submit data classified as Confidential or sensitive personal information to public Generative AI systems. City data cannot be used to train or parameter-tune AI models outside the City's secure control.

#### 4.4 Fairness, Equity, and Bias Reduction

- **Principle:** AI applications must not create or exacerbate inequities. Measures must be taken to mitigate discrimination and biased outcomes.
- **Rule (Bias Auditing):** Departments must conduct regular algorithmic risk assessments, leveraging racial equity toolkits, to ensure that AI-generated outputs and data inputs are accurate, representative, and free of discrimination against protected classes.

#### 4.5 Safety and Security

- **Principle:** AI tools must be reliable, rigorously tested, and protected from cyber-threats.
- **Rule (Incident Response):** All AI solutions are subject to cybersecurity incident monitoring. Serious faults that breach safety or fundamental rights must be reported immediately, and the City reserves the right to forcibly withdraw the AI system from use.

## 5.0 ACCEPTABLE AI USES

When applied responsibly, AI can greatly expand the capacity of local government. Acceptable use cases for City of Detroit personnel include, but are not limited to:

- **Improving Public Communication:** Utilizing AI-enabled translation services to make city information, websites, and public meetings accessible in multiple languages.
- **Enhancing Customer Service:** Deploying AI-powered web chat assistants (chatbots) into the 311 system to help residents navigate city services and submit requests 24/7.
- **Assisting Administrative Tasks:** Summarizing lengthy meeting transcripts, drafting internal memos, routing public inquiries, and assisting with grant writing, provided a human verifies the output.
- **Analytics and Infrastructure Management:** Using predictive AI to analyze patterns for infrastructure maintenance (example: predicting water pipe failures or detecting potholes), optimizing traffic signal timings to reduce congestion and emissions, and improving weather/flood preparation.

## 6.0 PROHIBITED AI ACTIVITIES

To protect the fundamental rights of Detroit residents, the following uses of AI are strictly prohibited within City government:

- **Social Scoring:** Utilizing AI to track, classify, or score individuals based on their behaviors, socioeconomic status, or personal characteristics in ways that could lead to discriminatory outcomes.
- **Subliminal Manipulation:** Using AI techniques to manipulate a person's behavior or exploit vulnerabilities (due to age, physical, or mental disability) in a manner that causes psychological or physical harm.
- **Consequential Unsupervised Decisions:** Using AI to make fully automated decisions without meaningful human oversight regarding hiring, termination, essential public services, or the allocation of financial grants/benefits.
- **Unregulated Biometric Identification:** Providing real-time remote biometric identification (for example: facial recognition) in publicly accessible spaces by law enforcement, except under specific, legally authorized scenarios accompanied by a warrant or established probable cause.
- **Deception and Deepfakes:** The creation or distribution of digitally generated or altered depictions of an individual without their consent, or generating misinformation with the intent to deceive the public.

## 7.0 GOVERNANCE AND OVERSIGHT STRUCTURES

To ensure the continuous safe deployment of AI, the City of Detroit will implement the following governance structures:

- **Chief AI Officer & AI Oversight Committee:** The CAIO will chair a cross-departmental AI Oversight Committee (like models in Tempe, AZ and Washington D.C.) responsible for tracking AI development, reviewing algorithmic systems, and managing policy compliance.
- **Algorithm Registry:** The City will maintain a public-facing "Algorithm Registry" to disclose the AI tools and applications actively used by municipal departments, providing transparency on how algorithmic tools support city decisions.

- **AI Review and Procurement Process:** All new software acquisitions, including free-to-use applications, must undergo an AI Review. Vendors supplying AI tools must submit an AI Factsheet detailing their model's training data, test data, performance metrics, and bias handling procedures.

## 8.0 DATA-HANDLING RULES AND GUIDELINES

Data is the foundation of artificial intelligence. The City of Detroit establishes the following data-handling rules:

- **Data Classification and Input Restrictions:** Employees must strictly follow City data classification guidelines. Confidential, Restricted, or Personally Identifiable Information (PII) must only be used in securely procured, closed AI systems approved by the IT Department, and never in open, public generative AI platforms.
- **Infrastructure Security:** AI tools must not process or store city data on infrastructure owned by foreign adversaries. All data storage must utilize encrypted, access-controlled environments.
- **Data Ownership and Vendor Contracts:** All procurement contracts involving AI must include standard provisions that guarantee the City retains ownership of its data and the outputs generated by AI. Vendors are prohibited from using City data to train public instances of their models without explicit authorization.

## 9.0 EMPLOYEE GUIDELINES AND TRAINING

- **AI Literacy and Upskilling:** The City recognizes that AI adoption requires an organizational culture shift. The Department of Information Technology will provide continuous AI literacy training programs, ensuring that employees understand how to protect sensitive information, mitigate algorithmic bias, and report system errors.
- **Prompt Libraries:** Departments are encouraged to develop standardized "Prompt Libraries" (as modeled by Chattanooga, TN) to help staff safely and consistently use generative AI to streamline daily tasks.

- **Records Retention:** All records, prompts, and output generated by AI tools used in an official capacity are subject to the city's records retention schedules and may be subject to public disclosure laws.
-

## References Utilized in the Formulation of this Policy:

1. *European Union Artificial Intelligence Act (AIA)*- Utilized for defining prohibited uses (social scoring, subliminal manipulation) and high-risk conformity frameworks.
2. *California Assembly Bill 2885 (Bauer-Kahan)*- Utilized for the foundational, technology-neutral definition of "Artificial Intelligence" derived from the OECD.
3. *City of Seattle Artificial Intelligence (AI) Policy & Generative AI Policy*- Utilized for policies surrounding Human-in-the-Loop (HITL), content attribution, bias mitigation via racial equity toolkits, and data privacy standards.
4. *State of Mississippi AI Acceptable Use Policy*- Utilized for guidelines on employee oversight, data handling, prohibition of deepfakes, and training rollout.
5. *Washington D.C. AI Procurement Handbook*- Utilized for procurement frameworks, the implementation of AI Factsheets, contract provisions, and continuous performance monitoring.
6. *City of Boston Interim Guidelines for Generative AI*- Utilized for baseline rules on fact-checking, disclosure, and limitations on sharing sensitive information in prompts.