Generative AI Policy Recommendations Report

Date: August 31, 2023

To: Interim CTO Jim Loter

From: CTO Policy Advisory Team

Subject: CTO Policy Advisory Team Report on Generative AI

Contents

Purpose

The purpose of this document is to present the findings and recommendations of the 2023 Generative Artificial Intelligence (AI) Policy Advisory Team. The Advisory Team, comprised of local thought leaders, researchers, subject matter experts, and City stakeholders, was tasked with delivering a set of recommendations for how the Chief Technology Officer (CTO) should set Citywide ITD Policy and/or adopt procedures and guidelines to responsibly manage the acquisition and use of Generative AI systems by City employees. The primary focus of the group was to develop the CTO's policy recommendations, however, a-broader set of City-wide recommendations around responsible AI principles and governance of AI tooling emerged.

Definitions

Artificial Intelligence (AI)

Artificial Intelligence (AI) commonly refers to a combination of: machine learning techniques used for searching and analyzing large volumes of data; robotics dealing with the conception, design, manufacture and operation of programmable machines; and algorithms and automated decision-making systems (ADMS) able to predict human and machine behavior and to make autonomous decisions (EU Parliament).

Generative AI

Generative AI is a class of artificial intelligence systems, including algorithms, deep-learning, and machine learning models, capable of generating new content, including but not limited to text, images, video, and audio, based on the inputs of training datasets. These also include systems capable of ingesting input and translating that input into another form, such as text-to-code systems.

Background

Open AI's ChatGPT launched in November of 2022 and had over 100 million active users by February of 2023 (Reuters). To contextualize the rapid scale of adoption accurately--it took TikTok approximately 9 months and Instagram approximately 2.5 years to reach the same number of active users (Reuters).

With the proliferation of Generative AI, there has been a corresponding global rise in concern around potential for risks and associated harms in the development and use of these systems, as well as a notable gap in comprehensive regulation to address these issues. The identified risks range from algorithmic bias, to legal implications surrounding intellectual property, and manipulation of Generative AI for malicious cyber security attacks.

In response to this rapid increase in Generative AI adoption, associated risk, and an increase in department requests for Generative AI technology in early 2023, the City's acting Chief Technology Officer (CTO) identified a need to provide City employees with timely guidance on use of Generative AI tools for official City business and drafted an Interim Policy on Generative AI use at the City. The interim policy was intended to provide early guidance for technology use while the Information Technology Department (ITD), in collaboration with a group of thought leaders in this space, gained a deeper understanding of its use for City business.

To ensure due diligence on long-term policy development, in May 2023, the CTO established a Policy Advisory Team on Generative AI systems, charged with researching and identifying potential implications for City government use of Generative AI systems and services.

Issues: Municipal Context for AI Use

AI has proven useful across multiple domains, from detection and diagnosis in healthcare, to improved resource management through robust econometrics. Similarly, Generative AI can be used to create efficiencies that enable users to shift their focus on more complex work tasks. Within the City, to date, we have explored or considered several potential use cases for AI and Generative AI including:

AI

- Supporting traffic management and planning goals including Vision Zero initiatives
- Using historical calls to evaluate opportunities for improved dispatch of appropriate first responder services

Generative AI

- Creating efficiencies in developing text-based communications
- Simplifying complex or technical language in outreach materials and communications
- Summarizing meetings and emails, and preparing meeting agendas and supporting materials

While these use cases have the potential to create efficiency in City employees' work, improve and innovate City services, and support the City's efforts to remove barriers to justice and opportunity for Seattle residents, without identifying and comprehensively addressing risk associated with these powerful technologies, there is potential for harmful outcomes for Seattle residents and the City at large.

NIST describes three areas of potential harms that arise from unmitigated risks associated with AI systems and their use: harm to individuals, organizational harm, and harm to the ecosystem. These harms apply across multiple topical areas of concern including equity, bias and discrimination, cyber security, data privacy, law, and more.

Individual harms resulting from bias outputs of AI systems have been well documented. For example, a Bloomberg study evaluating a Generative AI image generator found that images produced related to prompts for various high paying and low paying occupations resulted in lighter skin tone in the people depicted in the outputs for high-paying occupations, and darker skin tones associated with prompts for low-paying occupations. Similar results were also found when using the occupational prompt based on gender (images depicting men were produced for high paying occupations, and women for low paying occupations). In the City context, deploying systems that don't have oversight into mechanisms of bias evaluation and risk reduction not only means the City may run afoul of its Race and Social Justice and equity commitments, but also severely damage public trust and harm the residents we serve.

An example of end user risk can be seen in the instance of Samsung employees entering sensitive proprietary information into a Generative AI system[1]. The entry resulted in compromised confidential

[1] Samsung workers made a major error by using ChatGPT | TechRadar

information through end user prompts of sensitive code. At the City, an incident like this may result in compromised critical infrastructure that impacts basic needs of Seattle residents, such as access to potable water or electricity.

In the City context, such risks have potential to affect decision-making that could negatively impact things like resource allocation, equitable service provision, and law enforcement actions. Responsible use, governance, and policy around AI systems can help address such risks, increase public trust, and position the City of Seattle as municipal leaders in the space.

Scope

This report outlines recommendations on the City's use of Generative AI tools for the CTO to consider in development of a Citywide ITD policy on Generative AI, as well as additional Executive considerations to drive the development and adoption of Responsible AI Principles & associated Citywide program.

Contributors

External contributors to the work of the Generative AI Policy Advisory Team are as follows:

Name	Organization
Emily M. Bender Professor of Linguistics, University of Washington Adjunct Professor, Computer Science and Engineering, University of Washington Director, Professional Master's Program in Computational Linguistics Associate, **UW Tech Policy Lab**	University of Washington; UW Tech Policy Lab
Jan Whittington Associate Professor, University of Washington Urban Design and Planning Department Director, **Urban Infrastructure Lab** Associate, **UW Tech Policy Lab**	University of Washington; Urban Infrastructure Lab; UW Tech Policy Lab
Nicole DeCario Director, AI & Society, AI2	Allen Institute for AI (AI2)
Jacob Morrison Predoctoral Young Investigator, AI2 Public Policy Lead, AI2	Allen Institute for AI (AI2)
Isabel J. Rodriguez, CIPM City of Seattle CTAB Member	City of Seattle Community Technology Advisory Board (CTAB)

Co-chair CTAB Privacy & Cybersecurity subcommittee	
Omari Stringer, CIPT City of Seattle CTAB Member at large	City of Seattle Community Technology Advisory Board (CTAB)

City of Seattle contributors include:

Advisory Team Members	Stakeholders/Ex Officio Advisors
Jim Loter (Executive Sponsor): Interim Chief Technology Officer, *ITD*	**Aaron Valla:** Assistant City Attorney – Government Affairs & Records and Transparency Supervisor – Civil Division, *CAO*
Ginger Armbruster (Executive Sponsor): Director Data Privacy, Accountability, and Compliance, *ITD*	**Alexandra Nica:** Assistant City Attorney – Constitutional and Complex Litigation Section, *CAO*
Sarah Carrier (Team Co-Lead): Privacy Program Manager, *ITD*	**Joe Levan:** Assistant City Attorney, *CAO*
Eleonor Bounds (Team Co-lead): Data Privacy & Accountability Strategist, *ITD*	**Jennifer Dawson-Miller:** Race and Social Justice Program Advisor, *ITD*
Vinh Tang: IT Governance Advisor & Technology and Performance Advisor to Mayor's Office, *MO & ITD*	**Luv Sharma:** HR & Finance Applications Manager, *ITD*
Aurilee Gamboa: Architecture and Strategy Manager, *ITD*	**Destiny Cram:** Security Operations Manager, *ITD*
Monica Smitz: Applications Strategy Manager, *ITD*	**Dylan Morris:** Cybersecurity Risk Program Manager, *ITD*
Ed Odom: RSJI Program Lead, *ITD*	**Julie Kipp:** Public Disclosure Program Manager, *ITD*
Ana LaNasa-Selvidge: Change Management Lead, *ITD*	**Don Beherend:** Sr. Vendor Manager, *ITD*
Maria Hall: IT Service Desk Manager, *ITD*	**Jennifer Winkler:** City Records Manager, *LEG*
Ben Dalgetty: Digital Strategy Lead, *MO*	**Andrea Bettger:** Records Management Analyst, *LEG*
Mark James: Workplace Productivity Manager, *ITD*	

Executive Overview

The Advisory team met over seven 60-minute long sessions to deliberate on a broad set of questions about Generative AI technology. A high-level overview of specific recommendations for the CTO to consider in development of ITD's Citywide Policy on Generative AI include content around the following topic areas:

- General recommendations
- Acquisition and Contracting for City Use

- Intellectual Property, Attribution, Accountability, and Transparency of Authorship
- Bias and Harm
- Privacy
- Public Records and City Records Management
- Cybersecurity
- Labor and Economic impact

In addition to the CTO's policy suggestions, the Advisory team identified a foundational need for establishing City-wide guiding principles around the City's values and commitments to responsible and transparent AI use. Without adopting an agreed upon Citywide principled approach that guides the implementation of governance frameworks, a stand-alone CTO policy on Generative AI may lack authority, accountability, and robust mechanisms for ensuring consistent governance of AI systems across City departments.

The Advisory team conducted extensive research on industry best practices and assessed the global regulatory landscape to identify a set of additional recommendations that are built upon a principles-based approach. These are also listed below and further elaborated on where appropriate throughout this document:

1. Adopt a City-wide set of principles for Responsible AI use. These will guide policy and inform future technology acquisition and use and should be enacted through a joint Mayoral/Council Resolution
2. Establish a City of Seattle Responsible AI Program (modeled after the approach to creating the City's Privacy Program). This holistic programmatic approach to AI governance should be in alignment with adopted principles and determine appropriate owners/drivers for operationalizing governance functions (Staffing/resourcing for a Responsible AI Program should be considered.)
3. Empower the Responsible AI Program to operationalize governance frameworks across all City departments
4. Develop a set of policies and standards for the CTO to enact for specific Citywide guidance about AI use

Additional details about reach of these recommendations is provided below.

Recommendations Detail

Generative AI Principles

The following Responsible AI principles described below are proposed in alignment with current industry best practices[2], which include applicability to Generative AI solutions.

---

[2]OECD: https://oecd.ai/en/ai-principles
EU AI Principles: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html
White House AI Bill of Rights: https://www.whitehouse.gov/ostp/ai-bill-of-rights/
IBM Ethical AI Principles: https://newsroom.ibm.com/Principles-and-Practices-for-Building-More-Trustworthy-AI
Institute for Ethical AI and Machine Learning Principles: https://ethical.institute/principles.html
NIST Trustworthy AI: https://www.nist.gov/trustworthy-and-responsible-ai

**Please note: these principles apply to all AI systems, components, and services leveraging AI (including but not limited to Generative AI).*

1. **Innovation and Sustainability:** The City values public service innovation to meet resident needs. We commit to responsibly explore and evaluate AI technologies which will improve our services and advance beneficial outcomes for both people and the environment.
2. **Transparency and Accountability:** The City values transparency and accountability and understands the importance of these values in our use of AI systems. The City will ensure development, use, and deployment of systems are evaluated for and compliant with all laws and regulations applicable to the City prior to use and will make documentation related to the use of AI systems available publicly.
3. **Validity and Reliability:** The City will work to ensure that AI systems perform reliably and consistently under the conditions of expected use, and that ongoing evaluation of system accuracy throughout the development and/or deployment lifecycle is managed, governed, and auditable, to the greatest extent possible.
4. **Bias and Harm Reduction and Fairness:** We acknowledge that AI systems have the potential to perpetuate inequity and bias resulting in unintended harms on Seattle residents. The City will evaluate AI systems through an equity lens, in alignment with our Race and Social Justice commitments, for potential impacts such as discrimination, unintended harms arising from data, human, or algorithmic bias to the extent possible.
5. **Privacy Enhancing:** The City values data privacy and understands the importance of protecting personal data. We work to ensure that policies and standard operating procedures that reduce privacy risk are in place, and applied to the AI system throughout development, testing, deployment, and use to the greatest extent possible.
6. **Explainability and Interpretability:** The City understands the importance of leveraging AI systems, models, and outputs that are easily interpreted and explained. We work to ensure all AI systems and their models are explainable to the extent possible, and that system outputs are interpretable and communicated in clear language, representative of the context for use and deployment.
7. **Security and Resiliency:** Securing our data, systems, and infrastructure are important to the City. We will ensure AI systems are evaluated to ensure they are resilient and can maintain confidentiality, integrity, and availability of data and critical City systems, through protection mechanisms to minimize security risks to the greatest extent possible, in alignment with governing policy and identified best practices.

Responsible AI Program

Regulatory trends and industry best practices clearly establish need for authority, ownership, and robust governance structures to realize responsible AI use at scale. Building on the principles-based approach described above, it is recommended that Executive direction outlay the City's commitment to Responsible AI (to encompass Generative AI) establishing a Citywide Responsible AI Program through Resolution. Modeling a similar approach as was taken to establish the City's Privacy Program (established in 2015), the Responsible AI Program will develop a governance framework and drive operational aspects thereof, ensuring compliance with the City's stated commitments and associated policies.

Additional details about the proposed scope of work for the Responsible AI Program can be found in Appendix A.

Citywide Governance Framework

The Policy Advisory Team identified the need for the development and implementation of a Citywide governance framework, building upon the City's Responsible AI Principles. The Advisory team did not identify a specific governance structure, but identified the *need* for governance, as many of the recommendations center around management and policy around people, process, and technology (including data).

The City's governance framework should be developed by the Responsible AI Program and City stakeholders. The City may also identify areas of overlap and opportunities to partner with Innovation and Performance's City Data Alliance effort on the governance work currently underway.

CTO Policy Considerations and Recommendations

In driving the City's AI direction through continued leadership in the Generative AI space, it is recommended that the CTO consider the following:

1) CTO should develop an AI strategy in alignment with the City's AI principles/values and over all technology strategy to enable growth in responsible use of AI tools to support innovative service delivery
2) Existing IT policies, guidelines, and procedures should be reviewed and updated to reflect alignment with the Generative AI Policy

General Policy Recommendations for the CTO

The following table reflects thematic recommendations that showed up in *multiple topical sections* discussed by the Generative AI Policy Advisory Team.

**Please note that the content in the tables below assume adoption of the Citywide approach recommendations (e.g. Principles adoption and creation of a City of Seattle Responsible AI Program).*

Category	Recommendation	Owner
ITD Policy	All vendor & professional services contracts must include City standard Generative AI contract terms and conditions (to be developed) to ensure the vendor, sub-vendors or other service provider's use of Generative AI systems on behalf of the City aligns with the City's needs, principles, and policies	ITD; FAS; Department
ITD Policy	Vendors of Generative AI solutions (or solutions with Generative AI functionality) must undergo a vendor evaluation process & agree to adhere to all applicable City, IT, and department policies	ITD & Department
Process	Vendor evaluation should be conducted to assess: - If vendors meet City Policy and Principles around Responsible AI	Responsible AI Program

	- AI development processes (explainability mechanisms, data flows, data collection and training processes, privacy/security/compliance mechanisms, and approaches to evaluating and addressing bias)	
ITD Policy	Use and procurement of Generative AI solutions and associated use cases will follow ITD Technology Acquisition Policies, procedures, and associated assessment/risk evaluations (See Privacy Policy, IT Security Policy, technology exception policy, etc.)	ITD
Process	Develop and implement new Responsible AI Assessment (to ensure use meets Responsible AI commitments) and stakeholder review panel (including Privacy, Security teams, Business, RSJI Change Team, City Attorney's Office, CPRA, etc.) – to include tiered approach to AI system & use case risk (See *Appendix A)*	Responsible AI Program
ITD Policy	Outputs of Generative AI systems must be reviewed by humans prior to each use in an official City capacity (Human in the Loop). HITL review processes should be documented by owning departments	Department
ITD Policy	Users of Generative AI solutions must complete training on Generative AI solutions commensurate to their proposed use case and alignment with the City's AI Principles including but not limited to: Bias and Harm; Security & Resiliency; Validity and Reliability; Data Privacy; Public Records and Records Management; Explainability & Interpretability. (This may also include domain/job-role specific training)	Department
Process	Develop a robust training program for City employees using Generative AI tools & professionals looking to upskill in alignment with the CTO's AI strategy	Responsible AI Program; SDHR

Acquisition and Contracting for City Use

Please see *General Policy Recommendations* for other applicable considerations for this section.

Category	Recommendation	Owner
ITD Policy	RFPs involving Generative AI functionality or Generative AI solutions must include standard Generative AI RFP requirements (to be developed)	ITD, Responsible AI Program, Department
ITD Policy	Generative AI Policy should reference requirement to follow the City's established IT Acquisition processes	ITD & Department
Process	Produce and maintain living list of Generative AI solutions, use cases, and other transparency-related documentation	Responsible AI Program
Standard	Develop and leverage sandbox for pilot and solution evaluation	ITD & Department

Guideline	Free Generative AI tools should be avoided to ensure appropriate contract terms and compliance requirements are met	Department & ITD
Guideline	Consider WMBE vendors in procurement and contracting processes	Department & ITD

Intellectual Property, Attribution, Accountability, and Transparency of Authorship

Please see *General Policy Recommendations* for other applicable considerations for this section.

Category	Recommendation	Owner
ITD Policy	All **images and videos** created by Generative AI systems must be attributed to the appropriate Generative AI system. If a significant proportion of **text** generated by AI systems is used in a final product, attribution is required. Departments may establish a text-generation attribution threshold more rigorous than defined in ITD policy. At this time, departments may interpret "significance" thresholds. These thresholds should be documented in departmental policy on Generative AI. Future iterations of the Citywide Generative AI policy or associated guidelines may reflect attribution thresholds as the domain matures. Attributions should include the system name + a HITL assertion (which may include the department or group who reviewed/edited the content).	Department
Guideline	Appropriate methods of attribution should include but are not limited to: watermarks, footnotes, and headers	Department

Bias and harm

Please see *General Policy Recommendations* for other applicable considerations for this section.

Category	Recommendation	Owner
ITD Policy	Departments must leverage RSJI resources (e.g. the Racial Equity Toolkit) and/or work with departmental RSJI Change Team to conduct and apply the Racial Equity Toolkit prior to use of a Generative AI tool	Department; Responsible AI Program

Guideline & Process	Departments should develop equity metrics to assist in addressing the following: 1) Equitable deployment and use of the technology 2) Efficacy of the scoped solution These metrics should be evaluated and made publicly available annually	Department; Responsible AI Program

Privacy

Please see *General Policy Recommendations* for other applicable considerations for this section.

Category	Recommendation	Owner
ITD Policy	No City data or records, including inputs or prompts, are to be used for training or parameter tuning for Generative AI models outside the City's control (e.g. boundary-less systems; systems not leveraging retrieval augmented generation (RAG), etc.)	ITD; Department
ITD Policy	Data classified as Confidential or Confidential Requiring Special Handling based on the City's data classification standard, should not be used in prompts for Generative AI solutions	Department
ITD Policy	Departments leveraging Generative AI tools must develop and maintain departmental policy reflective of approved use cases, controls, text attribution thresholds, and compliance with any applicable regulatory requirements	Department
ITD Policy	Generative AI Policy should reference requirement to follow the City's established Privacy Policy	Department

Public Records & City Records Management

Please see *General Policy Recommendations* for other applicable considerations for this section.

Category	Recommendation	Owner
ITD Policy	All records generated, used, or stored by Generative AI vendors or solutions must comply with records retention and public disclosure laws	Department
ITD Policy	All Generative AI solutions and/or vendors are required to support retrieval/export of all prompts and outputs (either via functionality or vendor contract assurances)	Department & ITD
Guideline	Departments should maintain a living registry of active users of each Generative AI tool to be provided to their departmental Public Disclosure Officer upon request	Department

Guideline	Departments should document data discovery and retrieval processes for each Generative AI solution	Department

Cybersecurity

Recommendations in this section include security considerations from both a cyber risk and security operations perspective.

Future state recommendations will ensure technical controls are in place to prevent data leakage, secure City data assets, and reduce risk associated with unapproved use of Generative AI solutions. These recommendations include:

- Staff a Security Enforcement Team (3 FTE Purview admins) to roll out and support Purview initiatives, including scaled Citywide implementation of Data Loss Prevention and Information Protection services
- Define and establish ownership, support structure and associated processes for CASB tooling and implementation

Category	Recommendation	Owner
ITD Policy	Adversarial Testing : applications should have a penetration test performed to test the product over a wide range of inputs and user behaviors. This cost should be included in any project budget. (This may include City-side testing, or additional contract terms with vendors)	ITD
Process	Establish process for platform risk evaluation (either in house or 3rd party) to include conducting a pen test and/or human evaluation of code to ensure code is secure	ITD
ITD Policy	Ensure incident response policies are in place (including for vendor incidents)	ITD & Department
ITD Policy	Generative AI Policy should reference requirement to follow the City's established security policy, standards, and guidelines	Department
*Consideration	Consider need for development of supplemental Generative AI Security Policy addressing considerations of acceptable use, supply chain risks, code review and validation, etc.	ITD

Labor and Economic Impact

Category	Recommendation	Owner
ITD Policy	Ensure accountability and enforcement terms around prohibited use or violations of acceptable use of	ITD; SDHR; Labor Partners

	Generative AI systems are represented in a long term Generative AI Policy	
Process	Partner with HR to develop or update associated policies and/or personnel rules with respect to potential Generative AI use and associated labor and workforce impacts	ITD; Responsible AI Program; SDHR
Process	Develop upskilling plan for IT workforce in alignment with AI strategy	ITD

Appendix A: Responsibilities of the Responsible AI Program

The work of the Responsible AI Program will include, but is not limited to:

- Develop framework for use-case based risk categorization in alignment with City's proposed values and principles around Responsible AI
- Create and roll-out an evaluation process for AI solutions and proposed use cases in collaboration with other stakeholders
    - Assessment generation and operationalization
    - Identify/develop and implement risk and controls framework for AI tools and use cases ***Developing this approach to categorization and risk management is dependent upon Citywide adoption of a principled framework that authorizes and enables formalized governance structures to be put in place and operationalized programmatically*
    - Identify process integration points
    - Establish stakeholder review and other collaboration opportunities
- Lead development of City standard contracting language specific to AI solutions; develop AI-specific RFP requirements
- Develop vendor evaluation criteria, and operationalize vendor evaluation and assessment process
- Partner with IT technical teams to establish sandboxed environments to enable evaluation and validation of outputs (efficacy, utility, bias, etc.) in piloting AI solutions
    - Establish appropriate technical and policy controls for sandbox use
- Maintain supporting and associated IT Policy on AI and Generative AI technology
- Partner with HR and City labor partners to address policy and personnel rule guidance and/or updates to ensure training requirements and accountability measures can be met
- Develop and/or identify and provide City-wide training on AI and Generative AI in partnership with stakeholders and SDHR
- Maintain public transparency efforts
- Provide resources and support for City departments in meeting City's stated commitments on Responsible AI and other applicable compliance requirements associated with City use of AI systems or components
- Establish standards and guidelines for responsible use of AI systems within the City
- Develop processes for ongoing continual feedback and evaluation / audit of AI solutions

Appendix B: Initial Thoughts on Risk Categorization in City Context

This section is early thoughts on City use case risk categorization but is included as a reflection of robust Advisory Group conversation and future consideration.

Details in this section follow leading global approaches to AI regulation (e.g. EU AI Act) by proposing development of risk categorization by use case in alignment with the City's proposed values and principles around trustworthy and Responsible AI.

Draft City Use Case Categorization

This section is based on the EU AI Act's risk categorization framework, and is used as a baseline for the City to adapt and build upon.

*Please note the EU AI Act Framework is for AI more broadly, but incorporates reference to Generative AI.

Categories for consideration:

Unacceptable Risk: AI systems considered to involve a level of risk to society, fundamental rights, and other values which is not acceptable. Use of these systems is prohibited.

High Risk: AI systems which are considered to involve a high level of risk to society, fundamental rights and other values. May include general purpose AI systems and/or foundational models.

Limited Risk: AI systems intended to interact with natural persons. Mainly subject to transparency obligations.

Minimal Risk: Not regulated under the EU AI Act.

Risk Tier	EU AI Act Example Use Cases	City Use Cases (deliberative)	City Requirements (deliberative)
Unacceptable Risk	Remote biometric identification in public spaces; social scoring; systems used to distort a person's behavior or manipulate vulnerable populations	Would recommend using EU + any additional considerations. We have less protections for people so may need more specific call out. Generative AI: Analysis and reporting used for resource allocation in public services or initiatives; Evaluating potential hires/resume review	EU: Prohibited from use
High Risk	**Generative AI;** Medical devices; financial services; transport; aviation; automotive; critical infrastructure; HR/Recruitment tools;	City: Job application review, potential for self-reinforcing bias, discrimination, legal risk.	EU: Subject to regulatory requirements and fines to include: Risk management system; data and

	credit-worthiness; law enforcement uses	**Generative AI** (inclusive of EU call out) -- plugins and integrations as high risk as well.	data governance; technical documentation; record Keeping; transparency and reporting requirements to users; human oversight; accuracy, robustness and cybersecurity; conformity assessment City: ?
Medium/ Limited Risk	AI systems generating deep fakes; AI Chatbots	Chatbot for public facing website searches (pointing to City-owned sites or resources)? This is conversational/general not the same as generative though (which would be classed as high risk **if** alignment is with EU currently (unsure though)	EU: users must be informed that they are interacting with an AI system (not a human) City: If chatbot, users must be informed that they are interacting with ai system and have option to opt out
Minimal Risk	Spam Filters; AI powered video games	City: Communications assistance (Jasper): text or graphics generated by AI as draft communications assets and reviewed by a human before shared with public.	

Questions to Consider through AI Assessment and Evaluation Process

The questions below could prove useful in operationalizing an evaluation process of Generative AI systems (such as through an AI Impact Assessment), once authority and ownership of such a responsibility is determined at a City level.

Input:

1) Is personal confidential information being input?

2) Is other privileged, deliberative, or confidential information being input?

3) Are inputs fed back into public data sets?

Output:

4) Are humans in the loop before output goes public (e.g. augmentation or replacement of human interaction)? (*those people responsible for this oversight should have sufficient time and expertise to verify the generated output and ensure it aligns with the City's stated values).

5) Are outputs being used to drive policy or budget decisions?

6) Are outputs being used to drive public safety decisions?

7) Are outputs fed back into public data sets?

8) Is the output being used to speak for the City?