



ARTIFICIAL INTELLIGENCE (AI) PROCUREMENT HANDBOOK

DC Government: **Office of Contracting and Procurement**

Last Updated: February 2025



Table of Contents

1. Purpose of this Handbook	4
2. Basic Capabilities of Broad Categories of AI Tools	4
2.1. What is Artificial Intelligence (AI)	4
2.2. Potential Use Cases.....	5
2.3. How to Identify AI.....	5
2.4. Categories of AI Tools.....	5
3. How to Structure and Scope AI Tool Procurements	6
3.1. Pre-Solicitation Phase.....	7
3.1.1. Scope of Work (SOW).....	7
3.1.2. Independent Government Estimate (IGE)	11
3.1.3. Request for Information (RFI).....	12
3.2. Solicitation Phase.....	12
3.2.1. Selecting the Right Procurement Method	13
3.2.2. Selecting the Right Contract Type	14
3.2.3. AI Factsheet for Third Party Systems	15
3.2.4. Technical Proposal Instructions and Evaluation Factors	15
3.2.5. Standard Provisions and Requirements for AI Systems	15
4. Post Award Phase: How to Effectively Monitor Performance of Procured AI Tools	16
4.1. Performance Evaluation Requirements.....	17
4.2. Tips for Monitoring AI Tool Performance	17
4.3. Changes to AI Technology Post Award.....	18
APPENDIX A – Definitions	19
APPENDIX B – AI Factsheet Template	21
APPENDIX C – Organization and Content of Proposal.....	27
C.1. Technical Proposal Instructions	27
C.2. Technical Capability	27
C.3. System Management and Oversight	28
C.4. Experience and Past Performance	29
APPENDIX D – Evaluation Factors	30
D.1. Evaluation for Award	30
D.2. Technical Rating	30

D.3.	Evaluation Criteria	31
D.3.1.	Technical Criteria (75 Points Maximum)	31
D.3.2.	Price Criterion (25 Points Maximum)	31
APPENDIX E – Addendum: Requirements for AI Systems		33
E.1.	Definition of AI System	33
E.2.	Guiding Principles	33
E.3.	Risk Mitigation	33
E.4.	Requirements for Contractors when Operating AI System(s).....	34
APPENDIX F – References and Helpful Resources		36

1. Purpose of this Handbook

Pursuant to [Mayor's Order 2024-028](#), the Office of Contracting and Procurement (OCP) was directed to work with the District's Artificial Intelligence (AI) Taskforce to develop an AI procurement handbook.

The purpose of the AI procurement handbook is to provide guidance to agencies on:

1. The basic capabilities of broad categories of AI tools,
2. How to structure and scope AI tool procurements, and
3. How to effectively monitor the performance of procured AI tools.

With [Mayor's Order 2024-028](#), Mayor Muriel Bowser firmly committed District agencies to explore and deploy AI tools in careful alignment with [DC's AI Values](#).

It is important that when considering the purchase of AI tools, District agencies should consult with OCTO and carefully evaluate their costs, capabilities, levels of data security, and access controls against the specific needs of their organization. This handbook aims to walk government buyers through many of these considerations during the procurement process.¹

2. Basic Capabilities of Broad Categories of AI Tools

Like the private sector, government agencies are increasingly interested in purchasing artificial intelligence (AI) to help their employees work more efficiently. AI tools have evolved considerably over the past few years with commercial vendors now offering enterprise-level AI systems specifically tailored for use in enterprise business environments.

2.1. What is Artificial Intelligence (AI)

As described in [Mayor's Order 2024-028](#):

AI refers to the broad class of technologies developed or marketed to be capable of performing tasks otherwise requiring an intelligent human agent. Relevant tasks include, without limitation, natural language processing including text and speech generation, image analysis and generation, and a wide variety of probabilistically determined classifications, predictions, scorings, and assessments.

Simply put, AI is technology that performs tasks that otherwise would require human intelligence.²

Please see **Appendix A** for a comprehensive list of commonly used AI terms and definitions.

¹ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

² Reference: [McKinsey & Company](#)

2.2. Potential Use Cases

AI tools can be applied in a variety of different ways. Below are some common applications:

- ❖ Document creation, editing, and summarization
- ❖ Information retrieval, especially for new staff* (Example: Finding information within a SharePoint site or other shared document space)
- ❖ Meeting summary and follow-up* (Example: A live meeting)
- ❖ Email management*
- ❖ Data analysis and visualization
- ❖ Task automation* (from plain text to automation without any additional steps)
- ❖ Software code production and quality checking
- ❖ Personalized tutoring in a subject
- ❖ Drafting a statement of work (SOW)
- ❖ Market research
- ❖ Reverse market research (summarizing requirements for suppliers)
- ❖ Solicitation or contract translation for suppliers
- ❖ Data analysis to identify fraud

**Note: The use case typically requires the AI tool to have access to an agency's internal environment or to be embedded into an agency's document space (e.g., Copilot in Microsoft 365 or Gemini in Google Workspace).³*

2.3. How to Identify AI

A technology may be considered an AI system if it elicits positive answers to any of the following questions:

1. Does the technology use data to provide predictions, recommendations, insights, or decisions?
2. Does the technology augment human decision-making?
3. Does the company use words such as “personalized”, “tailored”, and “adaptive” in its marketing?⁴

2.4. Categories of AI Tools

The following list identifies broad categories of AI functionalities. Many common AI tools offer multiple functionalities, simultaneously:

- **Language Models**
Language models process textual information. This broad category includes large language models (“LLM”), visual language Models (“VLM”), text generation, text-

³ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

⁴ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

based virtual agents and web agents, machine translation tools, chatbots, sentiment analysis tools, and every modern software tool with a natural language interface.

- **Computer Vision Systems**

Computer vision systems process images as pixel information. This broad category includes facial recognition, object detection, object localization, object segmentation, image generation, the portion of robots or self-driving systems that take in and process visual information, and any other system that processes an input as a function of pixels.

- **Virtual Agents**

Virtual agents model decision-making processes over time. This broad category includes text-based virtual agents and web agents, driver assist and self-driving systems, robotics, game-playing bots, and any other system processing discounted rewards to aid in, or autonomously engage in, multi-step decision making.

- **Content Creation**

The use of AI to create, improve, and optimize content, such as text, images, videos, and more.

- **Anomaly Detection**

A technique that uses machine learning and AI to identify unusual patterns in data, finding outliers that deviate from a normal baseline.

- **Process Automation**

The transition of all or part of a given workflow to reduce the need for active, human intervention.

When purchasing an AI tool, it is important to understand the functionality and intended purpose of the technology. Using the wrong AI tool may generate inaccurate results, although it may first appear to be generating useful information. For example, large language models are a type of machine learning model that is designed for tasks such as language generation. An AI tool using a large language model would not be the right technology to statistical modeling, data mining, or to predict future outcomes or behaviors. Instead, large language models are best used for text summarization, question and answering systems, search engines, language translation, and similar tasks. It is therefore imperative to procure the right category of AI tool for the intended use.

3. How to Structure and Scope AI Tool Procurements

The Office of Contracting and Procurement's (OCP) partners with District agencies to purchase quality goods and services in a timely manner and at a reasonable cost while ensuring that all purchasing actions are conducted fairly and impartially.

The procurement process is typically completed in three primary phases:

1. Pre-Solicitation Phase

The client agency develops requirements and submits a request to purchase goods or services.

2. Solicitation Phase

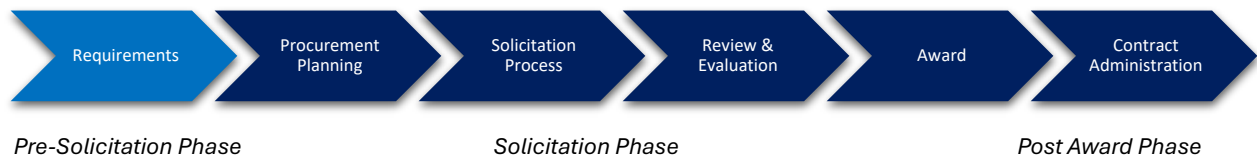
The procurement agency negotiates and enters a contract to acquire the goods or services following specific guidelines prescribed by District laws and regulations.

3. Post Award Phase

After the contract has been awarded, the procurement agency and the client agency work together to ensure the goods and services are received in accordance with the terms of the contract.

The following guidance is organized by procurement phase to help client agencies and procurement professionals successfully navigate the District's procurement process for the purchase of AI technology.

3.1. Pre-Solicitation Phase



The procurement process begins when a District agency identifies and assesses the need for specific goods or services—the agency's *requirement*. The timely and accurate identification of a requirement is necessary for an agency to meet its mission and functions.

For AI tool procurements, **District agencies** are required to include the following documentation:

- ❖ Scope of Work (SOW)
- ❖ Independent Government Estimate (IGE)

The following section provides guidance on how to develop and submit the mandatory requirements for AI tool procurements.

3.1.1. Scope of Work (SOW)

District agencies are responsible for drafting a scope of work (SOW) as part of the requirements phase of the procurement process. A well-written SOW provides the overall outline, goals, and requirements of the procurement and is used to:

1. Communicate what goods or services are to be delivered to the District,
2. Help vendors determine if they are capable of providing the goods or services,
3. Allow the vendor to submit an accurate bid or proposal, and
4. Allow the District to monitor the delivery of goods or services after contract award.

For AI tool procurements, the SOW should document or request the necessary information to complete an agency's [AI Values Alignment Report](#). Pursuant to [Mayor's Order 2024-028](#), District agencies are required to develop an AI Values Alignment Report for each AI tool they hope to deploy in support of their agency's mission. Agencies should refer to the [Handbook For AI Values Alignment](#) and the [AI Solution Intake Form](#) for more information about how to develop and submit their AI Values Alignment Report to OCTO.

The following guidance is provided to assist District agencies align their SOW with their AI Values Alignment Report. Although the AI Values Alignment Report does not have to be completed prior to the solicitation of an AI tool, aligning the SOW with the AI Values Alignment Report will help ensure agencies have the information it needs to make an informed decision regarding which AI tools meet the District's requirements, are the best fit for the agency, and comply with [Mayor's Order 2024-028](#).

❖ Clear Benefit to Residents

District agencies should clearly explain the purpose of the AI tool and the target population in the SOW. This explanation should also describe how the AI tool will improve service delivery by:

- Offering more services,
- Delivering services faster (reduced wait times),
- Delivering services with greater accuracy (fewer errors), or
- Improving service quality (better user experience).

❖ Safety and Equity

District agencies should clearly explain the safety and equity requirements in the SOW. Safety and equity requirements include but are not limited to the identification and mitigation of risks. It is important to document how risks will be carefully governed, mapped, measured, managed, and documented over time. As outlined in the [Handbook for AI Values Alignment](#), these risks should be categorized into the following areas:

- Direct and Indirect Physical Harm,
- Deprivation of Rights, and
- Exacerbating Inequity.

❖ Accountability

District agencies should clearly explain in the SOW how they will ensure responsibility for all government action flows clearly to an appropriate DC government official and how they will measure performance throughout the AI tool's lifecycle. Accountability requirements should include but not be limited to:

- Meaningful Accountability – requiring that District actions involving the AI tool remain traceable to human decision-makers,
- Preserving Human Control – requiring AI tools are designed to guarantee human control over AI outputs,

- Testing and Validation – explaining how the agency will want to test the AI tool’s performance under realistic conditions before deployment, and
- Ongoing Monitoring – explaining how the agency will require continuous testing and validation of the AI tool’s performance.

❖ Transparency

District agencies should include requirements in the SOW regarding how the usage of the AI tools will be clear and understandable to the public. Some specific requirements should include:

- AI Disclosure – SOWs should have requirements to explain how and when residents will be informed when they are interacting with an AI tool, like a chatbot, instead of a human. This could involve visual cues, audio messages, or upfront disclaimers.
- Human Review of AI Content – SOWs should have requirements that explain their process for human review of AI generated content, such as reports, and approvals before release of the content.

❖ Sustainability

The District should always be proactive to avoid situations where an AI tool creates new issues or concerns (such as environmental impacts or job loss) or becomes too expensive to maintain due to a vendor’s unilateral price and cost increases. The SOW should therefore ask the vendor to explain how the AI tool will be sustainable in the long term, considering the factors outlined in the [Handbook for AI Values Alignment](#), such as:

- Environmental Costs – require the vendor to disclose all the costs to maintain the AI tool, such as energy consumption for computing power,
- Impacts on the Workforce – require training for District staff,
- Financial Sustainability – ask for a clear description of the billing model to avoid cost overruns or unforeseen price increases, and
- Vendor Stability – require a transition plan so the agency understands how it can avoid “vendor lock-in” and switch providers without significant disruption, if needed.

❖ Privacy

District agencies should include in their SOWs privacy and legal requirements as outlined in the [Handbook for AI Values Alignment](#). These requirements should include, but may not be limited to:

- Legal Compliance
- Privacy Risks and Implementations
- Data Management & Security
- Potential Security Risks
- Notification Requirements

❖ Cybersecurity

As outlined in the [Handbook for AI Values Alignment](#), AI tools must be deployed in a way that promotes the confidentiality, integrity, and availability of the District's information technology assets. The SOW should require the vendor to explain:

- How the AI tool will be configured,
- What data will be collected, stored, or used,
- Where the data will be stored or used,
- Whether the system will be public facing,
- How the tool will be actively supported, and
- The vendor's risk mitigation strategies.

Other Considerations and Resources

The agency should include sufficient background information in the SOW so the vendor can answer the questions listed above. For example, the agency should disclose how the AI tool is intended to be used and any systems the AI tool will interface.

It is also important to document how the performance of the AI tool will be measured. This information is needed for the contract so the District can hold the vendor accountable for specific performance outcomes. The best way to evaluate an AI tool is to compare it to a human doing the same task. Examples include, but are not limited to:

- ❖ Question and answer
Accuracy of statements in the response, relevance of the response to the question.
- ❖ Document summarization
Quality of summary, relevance of summary to the document as determined by the user.
- ❖ Meeting summarization
Accuracy of assigning next steps or action items, correctly recording decisions reached in the meeting.
- ❖ Document retrieval
Relevance and recency of the documents recommended (i.e., providing the documents that an experienced staff-member familiar with the document space would have provided).
- ❖ Memo or policy drafting
Time required to correct or edit the document, adherence to agency's writing standards.
- ❖ Translation
Cultural sensitivity of translation, reading level of translation, ROUGE score (the quality of document translation).
- ❖ Coding

Accuracy of code generated and ability to find and resolve bugs, speed, and efficiency.⁵

Refer to the Government AI Coalition's [Guide to Measuring AI Performance](#) for additional information on measuring AI tool performance.

All District employees are highly encouraged to complete SOW training provided by OCP. District personnel can register for the training through [Peoplesoft](#). District agencies may also refer to the Harvard Kennedy School Government Performance Lab's [Guidebook for Crafting a Results-Driven RFP](#) for additional SOW best practices.

3.1.2. Independent Government Estimate (IGE)

Agencies must conduct market research to develop an Independent Government Estimate (IGE) to accompany their scope of work. The IGE estimates the cost of the purchase and is required when submitting a requisition.

When preparing the IGE, it is important to note that pricing schemes will differ by vendor, but they are generally either per user or consumption based.

❖ Per User

A per user price model is a flat fee per user of the system, per month. This is common for systems that staff might use for internal work functions, such as Microsoft 365 Copilot and ChatGPT Teams.

Per user price models may simplify cost calculations but can lead to unused licenses costing an agency unnecessary funds. If an agency chooses an AI tool with this price model, the agency should be mindful of usage rates in the agency and consider requiring staff to undergo a training on how to use the technology before they are given a license.

❖ Consumption-Based

A consumption-based price model charges by usage of the system. This price model is also known as a token-based scheme. For example, an agency might be charged for each time a user prompts the AI tool with a question or charged by the number of tokens used in each prompt (for reference, 1 token is approximately 4 English characters, or 0.75 words). This is common for public-facing systems that anyone can use at any time, such as external facing chatbots.

Sometimes pricing schemes may be a mix of per user and token-based models. Many AI tools (e.g., chatbots) also have a token or message limit. For instance, even with per user pricing, an agency may be limited to the number of tokens that can be used per day or per hour.⁶

⁵ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

⁶ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

When conducting market research to prepare the IGE, it is important to understand the difference between these two pricing schemes to accurately estimate the cost of the procurement.

If developing a custom AI tool, District agencies are highly encouraged to consult OCTO who can advise on proposed architecture, security requirements, and help to develop a cost estimate for contractual resources.

An accurate IGE is also important because it will be used by the contracting officer during the course of the procurement process to determine if a vendor's proposed pricing is fair and reasonable.

Once completed, agencies should include the IGE as an attachment to the agency's AI tool requisition.

3.1.3. Request for Information (RFI)

To assist agencies with their market research and requirements, it may be appropriate to issue a Request for Information (RFI). Pursuant [Title 27, Chapter 1601 of the DC Municipal Regulations \(DCMR\)](#), RFI's are used:

When information necessary for planning purposes cannot be obtained from potential sources by more economical and less formal means....

An RFI is used to obtain information from potential vendors and prepare them for an upcoming solicitation. An RFI is not a proposal and is not used to make an award. A formal solicitation is still required following the RFI via a separate process.

The RFI is non-binding and is generally used to obtain feedback and comments, including on best practices, new technology, and industry standards. RFIs generally do not request cost information. The RFI may focus on what a vendor may provide the District but may also focus on what information the District should provide about its needs to receive more meaningful proposals.

OCP will determine if an RFI is appropriate, after consulting with the client agency and OCTO.

3.2. Solicitation Phase



After receiving the agency's requirements, contracting staff will start the solicitation phase. The solicitation phase consists of four primary steps:

1. Procurement Planning

The procurement agency will determine the best procurement method and contract type to acquire the goods or services.

2. Solicitation Process

The procurement agency will seek vendors to provide the goods or services.

3. Review and Evaluation

The procurement agency will work with the client agency to select a vendor to provide the goods or services.

4. Award

The procurement agency will negotiate and enter a contract with a vendor to provide the goods or services.

The following section provides guidance for procurement professionals during the solicitation phase of the procurement process.

3.2.1. Selecting the Right Procurement Method

When preparing an AI procurement, procurement professionals should:

- ❖ Use procurement processes that focus not on prescribing a specific solution but rather on outlining problems and opportunities and allow room for iteration,
- ❖ Implement a process for the continued engagement of the AI provider with the acquiring entity for knowledge transfer and long-term risk assessment, and
- ❖ Create the conditions for a level and fair playing field among AI solution providers.⁷

The following section provides guidance to assist procurement professionals when determining the right procurement method for an AI tool purchase. Typically, the following procurement methods are recommended for most AI tool purchases but any procurement method may be used pursuant to [D.C. Official Code § 2-354.01](#) and [Title 27 of the DC Municipal Regulations \(DCMR\)](#):

- ❖ Competitive Sealed Proposals
- ❖ Special Pilot Procurement
- ❖ Cooperative Purchasing Agreements

Procurement professional should consider the following factors when selecting the right procurement method.

Competitive Sealed Proposals

⁷ Reference: [AI Government Procurement Guidelines](#)

Pursuant to [D.C. Official Code § 2-354.03](#), competitive sealed proposals shall be solicited through a request for proposals (RFP). Generally, RFPs should be used when purchases require specialized expertise, innovative solutions, or custom services. RFPs are most appropriate when purchasing goods or services based on overall value, quality, or capability and not just the lowest price.

Special Pilot Procurement

Pursuant to [D.C. Official Code § 2-354.08](#), a special pilot procurement can be used when there is an unusual or unique situation, such as satisfying a new and unique District requirement or obtaining a new technology. Pilot procurements are typically recommended when the AI tool is going to be tested on a smaller scale before a full-scale deployment. Pilot procurements are beneficial when an agency wants to test an AI tool for technical issues, identify if there are integration challenges with other systems, identify and test risk mitigation strategies, or prove that the AI tool is worth a much larger investment.

Cooperative Purchasing Agreements

Pursuant to [D.C. Official Code § 2-354.11](#), District agencies are encouraged to use cooperative purchasing agreements. Essentially, a cooperative purchasing agreement is a procurement method that allows the District to purchase goods or services using a contract that has already been negotiated and fully executed between a vendor and another jurisdiction. For example, the U.S. General Services Administration (GSA) maintains a multiple award schedule called the GSA Schedule. The District can enter into contracts with vendors on the GSA Schedule using the cooperative purchasing agreement procurement method.

3.2.2. Selecting the Right Contract Type

Pursuant to [Title 27, Chapter 2401 of the DC Municipal Regulations \(DCMR\)](#), contracting officers are required to identify the type of contract, or combination of types, to be used prior to solicitation. Contracting officers are required to consider factors, such as:

- ❖ Nature of the work: complexity and type of goods or services.
- ❖ Market conditions: price competition and market stability.
- ❖ Cost and risk: estimating costs, administrative burden, and overall risk.
- ❖ Timelines: urgency and contract length.
- ❖ Other considerations: existing contracts, contractor capabilities, and any other relevant factors.

Essentially, the contract type should align with the specific characteristics of the procurement to maximize efficiency and minimize risk for both parties.

Generally, **cost reimbursement category** contracts are not recommended for AI tool procurements because of:

- ❖ High risk of cost overruns,

- ❖ Reduced incentive for cost control,
- ❖ Difficulty in performance measurement, and
- ❖ Increased oversight burden.

3.2.3. AI Factsheet for Third Party Systems

As part of the procurement process, vendors should be required to fill out an AI Factsheet. The AI Factsheet captures basic facts about the AI system and enables the District to better understand the technical details of the AI system and assess the risks and benefits it presents. The AI Factsheet is intended to capture information including, but not limited to:

- ❖ Training data
- ❖ Testing data
- ❖ Input and outputs
- ❖ Performance metrics
- ❖ Optimal conditions
- ❖ Poor conditions
- ❖ Bias

The AI Factsheet should accompany a vendor's proposal. It should not be submitted in place of the proposal.

Please see **Appendix B** for an AI Factsheet template.

3.2.4. Technical Proposal Instructions and Evaluation Factors

Please see **Appendix C** for instructions on organization and content of the proposal and **Appendix D** for evaluation factors.

Please use these appendices as a guide or rubric for best practices. It is the vendor's responsibility to address these questions, and it is the agency's responsibility to identify the information they require and ensure that the vendor provides meaningful responses in the technical proposal.

3.2.5. Standard Provisions and Requirements for AI Systems

When purchasing an AI tool, it is important to pay careful attention to the contractual terms of the agreement. Check for:

- ❖ Support offered to the District for integration,
- ❖ Security on data the District provides to the AI tool,
- ❖ Ownership of the data the District provides to the AI tool, and ownership of the AI's outputs, and
- ❖ Indemnification, including protection from copyright infringement.

It is common for AI providers to protect their users from copyright claims. However, it is also important to consider the limitation of liability, or the maximum dollar amount the vendor will cover per claim.⁸

AI tool solicitations and contracts should include the following:

- ❖ The District's [Standard Contract Provisions](#),
- ❖ Applicable solicitation templates,
- ❖ Standard Solicitation Clause for AI Tool Procurements, and
- ❖ **Appendix E** – Addendum: Requirements for AI Systems.

Standard Solicitation Clause for AI Tool Procurements

All solicitations for services and information technology, regardless of procurement method or contract type, must contain the following AI notification clause:⁹

AI Notification Clause

Offeror must notify the contracting officer in writing if their solution or service includes, or makes available, any AI including AI from third parties or subcontractors.

During the term of the contract, Contractor must notify the contracting officer in writing if their services or any work under this contract includes, or makes available, any previously unreported AI technology, including AI from third parties or subcontractors.

At the direction of the contracting officer, Contractor shall discontinue the use of any new or previously undisclosed AI technology that materially impacts functionality, risk or contract performance, until use of such AI technology has been approved by the District.

Failure to disclose AI use to the District may be considered a breach of the contract by the District at its sole discretion and the District may consider such failure to disclose AI as grounds for the termination of the contract. The District is entitled to seek any and all relief it may be entitled to as a result of such non-disclosure.

4. Post Award Phase: How to Effectively Monitor Performance of Procured AI Tools



⁸ Reference: [Buyer's Guide to Enterprise Generative AI Tools](#)

⁹ Reference: [California Generative AI Toolkit](#)

The final phase of the procurement process is called post-award, or contract administration. During contract administration, procurement professionals work with client agencies to ensure vendors provide goods and services in accordance with the terms of their contract.

The following section provides guidance on how to effectively monitor performance of procured AI tools, including what to do if a contractor adds or makes changes to AI technology after a contract has been awarded.

4.1. Performance Evaluation Requirements

Pursuant to [D.C. Official Code § 2-352.04\(7\)](#), the OCP is required to *prepare, establish, and implement a periodic review process for the evaluation of contractors who provide goods or services to the District*. Following this law, OCP requires contract administrators to complete evaluations for any contracts over \$100,000 using the District's [Contractor Performance Evaluation System](#) (CPES).

For AI tools in particular, contract administrators are highly encouraged to pay special attention to the following evaluation questions below, which will help OCP determine if the contractor is adhering to the terms and conditions of the contract, and whether modifications to option years or future contracts may be appropriate.

- ❖ How well did the product(s)/service(s) comply with contract requirements / specifications?
- ❖ How accurate and complete was the required reporting?
- ❖ How well did the contractor control the cost of the contract and its components?
- ❖ How was the contractor's performance in resolving issues for all involved stakeholders?
- ❖ How well did the contractor display reasonable and cooperative behavior?

In addition to rating each question on a 0 – 5 scale, contract administrators should provide detailed narrative explanations for each question.

4.2. Tips for Monitoring AI Tool Performance

It is important to evaluate an AI system's performance across a range of metrics. Since many AI tools are a bundle of multiple services, it may be necessary to utilize different metrics for different functions.

OCP contracting officers through the contract administrator should continuously monitor, assess, and validate AI contract deliverables and performance metrics for equitable outcomes, output inaccuracies, fabricated content, hallucinations, biases, and the need for human action for all decision-making processes, to ensure applicable District laws and policies are followed. Agencies should assign a subject matter expert to assist OCP contracting officers and contract administrators with assessing and validating contract deliverables and performance metrics.

4.3. Changes to AI Technology Post Award

If a vendor makes any changes or modifications to AI technology after contract award, the vendor is required to notify the contracting officer in accordance with the AI Notification Clause referenced in Section 3.2.5 of this Handbook. Once notified, the contracting officer must notify the client agency and OCTO. Agencies may be directed by OCTO to update their AI Values Assessment or Risk Analysis as part of OCTO's review of the changes and the vendor may be required to update their AI Factsheet. OCP will then decide if changes to the AI technology necessitate a contract modification.

APPENDIX A – Definitions

Algorithm – a clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result.

Artificial Intelligence (AI) – the broad class of technologies developed or marketed to be capable of performing tasks otherwise requiring an intelligent human agent. Relevant tasks include, without limitation, natural language processing including text and speech generation, image analysis and generation, and a wide variety of probabilistically determined classifications, predictions, scorings, and assessments.

AI Incident – a documentation event, like the filing of a judicial or administrative claim, a complaint, or an incident report, that alleges a harm or near harm event to people, property, reputation, technical integrity of the environment arising from the operation of an AI Tool.

AI System – any data system, software, hardware, application, tool, or utility that operates in whole or in substantial part using AI.

Anomaly Detection – a technique that uses machine learning and AI to identify unusual patterns in data, finding outliers that deviate from a normal baseline.

Automated Decision System – a computational process derived from machine learning, statistical modeling, data analytics, or artificial intelligence that issues simplified output, including a score, classification, or recommendation, that is used to assist or replace human discretionary decision making and materially impacts natural persons. An “automated decision system” does not include a spam email filter, firewall, antivirus software, identity and access management tools, calculator, database, dataset, or other compilation of data.

Chatbot – computer programs that simulate and process human conversation, either written or spoken, to allow humans to interact with digital devices as if they were communicating with a real person.

Content Creation – the use of AI to create, improve, and optimize content, such as text, images, videos, and more.

Generative AI (GenAI) – pretrained AI models that can generate images, videos, audio, text, and derived synthetic content. GenAI does this by analyzing the structure and characteristics of the input data to generate new, synthetic content similar to the original. (Also referred to as Strong AI or Creative AI).

Language Model – an AI tool capable of receiving natural language inputs and providing natural language outputs.

Machine Learning – techniques for automated performance improvement.

Predictive Modeling – a technique that uses artificial intelligence, particularly machine learning algorithms, to analyze historical data and identify patterns to predict future outcomes or behaviors, essentially forecasting what might happen based on past trends and current information.

Process Automation – the implementation of AI technologies, such as natural language processing (NLP), machine learning (ML), large language models, (LLMs) and data analysis, into an organization's process orchestration layer to enhance and optimize end-to-end business processes.

APPENDIX B – AI Factsheet Template¹⁰

Please provide details regarding your Artificial Intelligence (AI) product by filling out the AI Factsheet¹¹ template below. Vendors should submit the AI Factsheet along with their technical proposal. It should not replace the technical proposal.

AI Factsheet Template

Vendor Name	
System Name	
Overview	Brief summary of the AI system.
Purpose	What function does the AI system perform, and for what purpose? If the system performs multiple functions, list each discretely. For features that are configurable, please describe all configuration options and default settings.
Intended Domain	What domain is the AI system intended to be applied in?
Training Data	How was the AI system trained? What data was used? How often is data added to the training set? Was all training data legally obtained and its use fully licensed?
Test Data	What data was used to test system performance? Under what conditions has the system been tested?
Model Information	General description of the model(s) used (e.g., large language model, transformer, deep learning, supervised learning, built on an existing open source model, computer vision)
Update procedure	In general, how often are the models updated for users? Will the user have a choice in moving to the updated model or staying on the current model? What documentation is available for new versions of the model?
Inputs and Outputs	What are the inputs to the AI system? What are its outputs? What interfaces and integrations are supported?

¹⁰ Reference: [AI FactSheet PDF](#)

¹¹Reference: [IBM AI FactSheets 360](#)

Performance Metrics	What are the performance metrics? What is your current level of performance on these metrics? How can the user monitor performance in the deployment environment?
Bias	What biases does the tool exhibit and how does it handle that bias? This can include but is not limited to biases on human factors such as gender, race, socioeconomic status, disability, culture, age, or other protected classes, or biases on general factors such as a sampling bias, survivorship bias, detection bias, or observer bias.
Robustness	How does the AI system handle outliers? Do overwritten decisions feed back into the system to help calibrate it in the future?
Optimal Conditions	What conditions does the model perform best under? Are there minimum requirements for the quantity of records/observations?
Poor Conditions	What conditions does the model perform poorly under? What are the limitations of the AI system? What kinds of errors can it make (e.g., hallucinations) and what conditions make those errors more likely?
Explanation	How does the AI system explain its predictions? Are the outcomes of the AI system understandable by subject matter experts, users, impacted individuals, and others?
Jurisdiction-specific Considerations	Please describe any considerations relevant to local, state, industry, or other specific jurisdictional regulations.

Algorithmic Impact Assessment Questionnaire

How is the AI tool monitored to identify any problems in usage? Can outputs (recommendations, predictions, etc.) be overwritten by a human, and do overwritten outputs help calibrate the system in the future?	Problems in usage can include false negatives, false positives, bias, hallucinations, and human-reported quality issues (such as poor translations or poorly generated images).
How is bias managed effectively?	This can include ways to monitor bias, or abilities to toggle parameters to change observed bias in the model.

Have the vendors or an independent party conducted a study on the bias, accuracy, or disparate impact of the system? If yes, can the Agency review the study? Include methodology and results.	This can include bias impact reports, algorithmic impact reports, or others. ¹²
How can the Agency and its partners flag issues related to bias, discrimination, or poor performance of the AI system?	This can include ways to report inaccurate or concerning decisions/classifications made by the AI system, or ways to retroactively review past system actions.
How has the Human-Computer Interaction aspect of the AI tool been made accessible, such as to people with disabilities?	Has it been assessed against any usability standards, and if so what was the result?
Please share any relevant information, links, or resources regarding your organization's responsible AI strategy.	URL to any broad AI policy or strategy.

¹² See [Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms](#) for an example bias impact report template.

Example Factsheet

This is an example of the AI Factsheet¹³ above completed by a fictitious company.

Vendor Name	XYZ Technologies, Inc.	
System Name	Audio Classifier	
Overview	This document is a FactSheet accompanying the Audio Classifier model on IBM Developer Model Asset eXchange .	
Purpose	This model classifies an input audio clip.	
Intended Domain	This model is intended for use in the audio processing and classification domain.	
Training Data	The model is trained on the AudioSet dataset by Google. New data is added to the training set daily. The AudioSet database was legally obtained and its use is fully licensed.	
Test Data	The test set is also part of the AudioSet data. There was a 70:20:10% split of the data into train:val:test. The ratio of samples/class was maintained as much as possible in all the splits. The system has been tested in X,Y,Z conditions.	
Model Information	The audio classifier is a two-stage model: • The first model (MAX-Audio-Embedding-Generator) converts each second of input raw audio into vectors or embeddings of size 128 where each element of the vector is a float between 0 and 1. • Once the vectors are generated, there is a second deep neural network that performs classification.	
Update procedure	In general, the model is updated annually. If the user does not wish to move to the updated model, the user cannot continue to use the system. Documentation for all new versions of the model can be found on the website at this link.	
Inputs and Outputs	Input: a 10 second clip of audio in signed 16-bit PCM wavfile format. Output: a JSON with the top 5 predicted classes and probabilities.	
Performance Metrics	Metric	Value
	Mean Average Precision	0.357

¹³ Reference: [IBM Research AI Factsheets 360](#)

	<table><tr><td>Area Under the Curve</td><td>0.968</td></tr><tr><td>d-prime</td><td>2.621</td></tr></table> The user can regularly monitor these metrics [here].	Area Under the Curve	0.968	d-prime	2.621														
Area Under the Curve	0.968																		
d-prime	2.621																		
Bias	The majority of audio samples in the training data set represent voice and music content. Potential bias caused by this over-representation has not been evaluated. Careful attention should be paid if this model is to be incorporated in an application where bias in voice type or music genre is potentially sensitive or harmful.																		
Robustness	This audio classifier is not robust to the L-infinity and L2 norms for the HopSkipJump attack. <table><tr><td></td><td>L2</td><td>L-Infinity</td></tr><tr><td>5th Percentile</td><td>887.0 (200.9)</td><td>5.5 (4.9)</td></tr><tr><td>10th Percentile</td><td>1496.6 (720.6)</td><td>7.53 (5.73)</td></tr><tr><td>15th Percentile</td><td>3723.1 (4707.2)</td><td>52.8 (41.8)</td></tr><tr><td>25th Percentile</td><td>7187.9 (---)</td><td>187.6 (198.1)</td></tr><tr><td>50th Percentile</td><td>11538.6 (---)</td><td>502.8 (---)</td></tr></table> The susceptibility of the model to the two attacks. The parenthetical values in the table above represent the fitted curve evaluated at 11 iterations. (When we are unable to fit a curve, or the result is negative, we denote by ---.) Overwritten decisions are fed back into the system to help calibrate it in the future.		L2	L-Infinity	5 th Percentile	887.0 (200.9)	5.5 (4.9)	10 th Percentile	1496.6 (720.6)	7.53 (5.73)	15 th Percentile	3723.1 (4707.2)	52.8 (41.8)	25 th Percentile	7187.9 (---)	187.6 (198.1)	50 th Percentile	11538.6 (---)	502.8 (---)
	L2	L-Infinity																	
5 th Percentile	887.0 (200.9)	5.5 (4.9)																	
10 th Percentile	1496.6 (720.6)	7.53 (5.73)																	
15 th Percentile	3723.1 (4707.2)	52.8 (41.8)																	
25 th Percentile	7187.9 (---)	187.6 (198.1)																	
50 th Percentile	11538.6 (---)	502.8 (---)																	
Optimal Conditions	- When the input audio contains only one or two distinct audio classes. - When the audio quality is high with lesser noise.																		
Poor Conditions	The system can misclassify audio: - When the audio contains more than two distinct classes, and - When the audio quality is low with more noise.																		
Explanation	While the model architecture is well documented, the model is still a deep neural network, which largely remains a black box when it comes to explainability of results and predictions.																		
Jurisdiction-specific Considerations	N/A																		

Algorithmic Impact Assessment Questionnaire

How is the AI tool monitored to identify any problems in usage? Can outputs (recommendations, predictions, etc.) be overwritten by a human, and do overwritten outputs help calibrate the system in the future?	The system can be monitored in usage, and audio classification decisions can be retroactively overwritten by a human. The overwritten decisions can help calibrate the system in the future if desired.
How is bias managed effectively?	Users have access to performance metrics that can be used to understand if the bias in voice-type or music style is harmful.
Have the vendors or an independent party conducted a study on the bias, accuracy, or disparate impact of the system? If yes, can the Agency review the study? Include methodology and results.	Yes. Results from the third-party study can be provided upon request.
How can the Agency and its partners flag issues related to bias, discrimination or poor performance of the AI system?	The system provides a web portal to each customer to show the results of the system and its impact on transit performance in the form of reports and graphs.
How has the Human-Computer Interaction aspect of the AI tool been made accessible, such as to people with disabilities?	The system is embedded into a graphics user interface that is compliant with modern screen readers, and provides the option for auto-generated dictation of text on the screen.
Please share any relevant information, links, or resources regarding your organization's responsible AI strategy.	Information about our responsible AI strategy can be found on our website at this link.

APPENDIX C – Organization and Content of Proposal

C.1. Technical Proposal Instructions

Offerors must submit the following completed documentation with their technical proposal submission, along with the **Appendix B – AI Factsheet**. The recommended evaluation factors for the technical proposal are as follows:

1. Technical Capability
2. System Management and Oversight
3. Experience and Past Performance

C.2. Technical Capability¹⁴

Offerors should be evaluated on their ability to meet the technical system requirements. Specifically, Offerors should be required to provide:

- ❖ System Overview
 - A brief summary of the AI system, including a non-technical overview of how the AI tool operates and its key functionalities.
 - A copy of training materials and an implementation plan.
- ❖ Data Training and Model Description
 - A description of how the AI system learns information and what kind of data it has been trained on. For example, the vendor should explain how the AI system was trained, what data was used, and the conditions that were used to test the AI system.
 - A general description of the model(s) used.
- ❖ System Operations
 - A description of how often the models are updated.
 - An explanation if users have a choice in moving to an updated model or staying on the current model.
 - A summary of specific education or certifications that may be required for system operators.
 - If applicable, compatibility with the District's existing IT infrastructure.
 - A description of data security and privacy protocols.
- ❖ Interpretability and Explanation
 - An explanation of how the AI system explains its predictions.
 - Examples or scenarios illustrating how the AI system communicates its predictions in a way that is easy to understand to non-experts.

¹⁴ Reference: [AI Policy Manual](#)

C.3. System Management and Oversight¹⁵

Offerors should be evaluated on their ability to meet the management and oversight requirements. Specifically, Offerors should be required to provide:

❖ Performance Evaluation

- An explanation of how the accuracy and effectiveness of the system are measured. For example, what metrics are used, and why?
- A description of the range of accuracy of the AI system and how it may vary depending on the data.
- A description of what the system is optimizing for and under what constraints.

❖ Ethical Considerations

- A list of biases the tool exhibits and how the vendor handles that bias.
- A description of how the vendor reports bias or justifies why no bias would be present.
- A description of how the tool prevents or reduces harm to end users.

❖ System Reliability

- An explanation of how the AI system handles outliers.
- How the system is calibrated and if overwritten decisions feed back into the system to help to improve accuracy in the future.
- A description of the conditions the model performs best and a description of the conditions that the model performs poorly.
- A description of the limitations of the AI system.
- A summary of the expertise required for operation, debugging, modification, and troubleshooting.

❖ Monitoring and Correction

- A summary of how the AI tool is monitored to identify any problems in usage.
- An explanation of whether outputs (recommendations, predictions, etc.) can be overwritten by a human.
- An explanation of whether overwritten outputs help calibrate the system in the future.

❖ Studies and Transparency

- A copy of any studies on the fairness and accuracy of the system on topics such as bias, accuracy, or disparate impact.
- For each study, include a summary of the methodology and results, and who conducted the study (the vendor or an independent party).
- An explanation of whether the data used to train the system is representative of the communities it covers.

❖ User Interaction and Feedback

¹⁵ Reference: [AI Policy Manual](#)

- A description of how users provide feedback on any issues they encounter with the AI system, such as bias, discrimination, or performance.
- A summary of the measures that have been taken to ensure accessibility for all users.
- A summary of assessments against usability standards, and the results.
- A description of any other human factors, if any, that were considered for usability and accessibility of the system.

C.4. Experience and Past Performance

Offerors should be evaluated on their experience and past performance on projects of similar scope and complexity. Specifically, Offerors should be required to provide:

❖ Expertise with AI Technologies

- A description of the Offeror's expertise with AI technologies required under this solicitation, which may include but is not limited to: machine learning for predictive analytics, natural language processing for data extraction and analysis, computer vision for image and video recognition, and chatbots for citizen interaction.

❖ Previous Projects and Experience of Similar Size and Scope

- A summary of the Offeror's experience on three relevant projects of similar size and scope, focusing on what the Offeror considers being most relevant in demonstrating its qualifications.

❖ Experience and Qualifications of Key Personnel

- Identify key personnel and submit resumes with their qualifications.
- Define the roles and responsibilities of the key personnel identified.
- Describe the Offeror's organizational structure, including the position of the key personnel in the organizational structure.

APPENDIX D – Evaluation Factors

D.1. Evaluation for Award

The contract will be awarded to the responsible offeror(s) whose offer is most advantageous to the District, based upon the evaluation criteria specified below. Thus, while the points in the evaluation criteria indicate their relative importance, the total scores will not necessarily be determinative of the award. Rather, the total scores will guide the District in making an intelligent award decision based upon the evaluation criteria.

D.2. Technical Rating

The Technical Rating Scale for all AI tool procurements shall be as follows:

Numeric Rating	Adjective	Description
0	Unacceptable	Fails to meet minimum standards; e.g. no demonstrated capacity, major deficiencies, which are not correctable; offeror did not address the factor.
1	Poor	Marginally meets minimum requirements; major deficiencies which may be correctable.
2	Minimally Acceptable	Marginally meets minimum requirements; minor deficiencies which may be correctable.
3	Acceptable	Meets requirements; no deficiencies.
4	Good	Meets requirements and exceeds some requirements; no deficiencies.
5	Excellent	Exceeds most, if not all requirements; no deficiencies.

The technical rating is a weighting mechanism that will be applied to the point value for each evaluation factor to determine the offeror's score for each factor. The offeror's total technical score will be determined by adding the offeror's score in each evaluation factor. For example, if an evaluation factor has a point value range of zero (0) to forty (40) points, using the Technical Rating Scale above, if the District evaluates the offeror's response as "Good," then the score for that evaluation factor is 4/5 of 40 or 32.

If subfactors are applied, the offeror's total technical score will be determined by adding the offeror's score for each subfactor. For example, if an evaluation factor has a point value range of zero (0) to forty (40) points, with two subfactors of twenty (20) points each, using the Technical Rating Scale

above, if the District evaluates the offeror's response as "Good" for the first subfactor and "Poor" for the second subfactor, then the total score for that evaluation factor is 4/5 of 20 or 16 for the first subfactor plus 1/5 of 20 or 4 for the second subfactor, for a total of 20 for the entire factor.

D.3. Evaluation Criteria

The evaluation factors that will be considered in evaluating proposals must be tailored to each procurement. They shall include only those factors, and any subfactors, that guide the District to make a decision that best meets its needs.

D.3.1. Technical Criteria (75 Points Maximum)

The technical criteria shall include a maximum of 75 points. When developing the technical criteria for an AI tool procurement, agencies and procurement professionals should consider:

❖ **Factor 1 – Technical Capability**

The Offeror's proposal for this factor will be evaluated based on the following subfactors:

- System Overview
- Data Training and Model Description
- System Operations
- Interpretability and Explanation

❖ **Factor 2 – System Management and Oversight**

The Offeror's proposal for this factor will be evaluated based on the following subfactors:

- Performance Evaluation
- Ethical Considerations
- System Reliability
- Monitoring and Correction
- Studies and Transparency

❖ **Factor 3 – Experience and Past Performance**

The Offeror's proposal for this factor will be evaluated based on the following subfactors:

- Expertise with AI Technologies
- Previous Projects and Experience of Similar Size and Scope
- Experience and Qualifications of Key Personnel

D.3.2. Price Criterion (25 Points Maximum)

The price evaluation will be objective. The offeror with the lowest price will receive the maximum price points. All other proposals will receive a proportionately lower total score. The following formula will be used to determine each offeror's evaluated price score:

$$\frac{\text{Lowest Price Proposal}}{\text{Price of Proposal Being Evaluated}} \times \text{Weight} = \text{Price Score}$$

APPENDIX E – Addendum: Requirements for AI Systems

This Addendum defines special requirements agreed to by the District and Contractor regarding the AI system and / or subsystem provided as part of the Contract.

This Addendum governs over any contrary license terms and the District will not agree to any terms that conflict with the Addendum. Failure of the Contractor to comply with the terms of this Addendum shall constitute a material breach of the Contract.

Contractor agrees to indemnify, defend, and hold harmless the District regarding any third party action rising out of or related to (1) any breach of any representation or warranty of Contractor contained in this Addendum; (2) any breach or violation of any covenant or other obligation or duty of Contractor under this Addendum or under applicable law; (3) any third party claims which arise out of, relate to or result from any act or omission of the Contractor related to the provision of an AI system; and (4) any violations or alleged violations of intellectual property rights; in each case whether or not caused in whole or in part by the negligence of the District, or any other indemnified party, and whether or not the relevant claim has merit.

E.1. Definition of AI System

Pursuant to [Mayor’s Order 2024-028](#), the District defines artificial intelligence (AI) as:

The broad class of technologies developed or marketed to be capable of performing tasks otherwise requiring an intelligent human agent. Relevant tasks include, without limitation, natural language processing including text and speech generation, image analysis and generation, and a wide variety of probabilistically determined classifications, predictions, scorings, and assessments.

The District defines an “AI system” to be any data system, software, hardware, application, tool, or utility that operates in whole or in part using AI.

E.2. Guiding Principles

The Contractor shall demonstrate that the AI system and its usage, deployment, and maintenance as it pertains to the services outlined in this agreement do not conflict with [Mayor’s Order 2024-028](#), the District’s [AI/ML Governance Policy](#), or the District’s [Handbook for AI Values Alignment](#).

E.3. Risk Mitigation

The Contractor represents that the AI system is suitable for its intended use by the District and has been developed and will perform in a manner that is in compliance with all applicable laws and regulations.

The Contractor shall work with the District to evaluate and minimize risks posed by the AI system.

AI Incident Response

In the event of an AI incident, at the request of the contracting officer, the Contractor shall thoroughly investigate their systems of any suspected AI incident and promptly report findings to the District. An “AI incident” is a documentation event, like the filing of a judicial or administrative claim, a complaint, or an incident report, that alleges a harm or near harm event to people, property, reputation, technical integrity of the environment arising from the operation of an AI Tool.

Remediation

At the contracting officer’s request, the Contractor will immediately discontinue the use of any AI system involved in providing services to the District. If the District, in its sole discretion, determines that the Contractor does not promptly resolve an AI incident, or that the system does not adequately support the District’s commitment to [Mayor’s Order 2024-028](#), the District’s [AI/ML Governance Policy](#), or the District’s [Handbook for AI Values Alignment](#), the District will provide the Contractor with notice that they have 10 calendar days to promptly assess and resolve the issue. Potential methods to address such issues include, without limitation, changing the behavior of the AI system or subsystem; supplementing the system or subsystem to achieve the necessary outcomes; replacing the system with a non-AI system that meets the District’s needs; or limiting the function of the AI system or subsystem. After 10 calendar days, the Contractor must provide evidence that the AI system is adequately fixed and ready for re-deployment, or that the AI system is not suitable for use.

E.4. Requirements for Contractors when Operating AI System(s)

To the extent permissible by law, the Contractor shall adhere to the following requirements while using any AI systems in the course of doing business with or for the District:

1. **Review:** Contractor attests that the previously completed AI Factsheet accurately represents the AI system. Contractor commits to update the AI Factsheet on an annual basis and within 30 calendar days of any substantive change to the AI system. Any substantive changes made to the AI Factsheet may be cause for termination by the District, if it is determined at the District’s sole discretion that the revised information renders the AI system unserviceable to the District.
2. **Performance:** Contractor will provide the District with the means to monitor the performance, including the accuracy, of the AI system it uses and report this accuracy to the District. This may include, but is not limited to, the false positive rate, the false negative rate, the true positive rate, the average percentage error, the mean-squared error, and human judgement scores.
3. **Algorithmic bias:** Contractor will provide the District with evidence that demonstrates that bias present in the AI system is effectively managed for the context in which it will be deployed. Contractor shall provide information describing in detail how bias is assessed.
4. **Human oversight:** Contractor will provide the District the means for a human to evaluate and override outputs of the AI system. The human evaluator must be able to override the outputs of the AI system and take precedence over all outputs.
5. **Explainability:** Contractor will provide the District with an explanation of how the AI system generates outputs, including what factors influence the system’s decisions, rule-based logic, training data sources, and probability-based decisions. The District holds the right to communicate its general usage of the AI system and explain its decision-making processes to the public.

6. **Notice:** If required by the District, provide written notice of the usage of the AI system to data subjects and/or end-users, preferably at the point of service.
7. **Process:** Contractor shall comply with existing local, state, and federal law for data access related to the use or operation of the system.
8. **Ongoing Monitoring:** Contractor shall regularly monitor the performance of the AI system to detect and rectify system behavior that violates any of the requirements in this Addendum. Contractor shall promptly communicate the discovery of system behavior that violates any of the requirements in this section to the District, including the potential impact to services.
9. **Training:** Contractor shall ensure that appropriate training is available to District staff who may operate the AI system, which may include how to:
 - Protect sensitive or personal information,
 - Mitigate harmful algorithmic bias,
 - Promote optimal performance,
 - Report system errors, and
 - Maintain service delivery if the AI system fails, to the extent possible.
10. **Auditing:** The District retains the right to observe or audit any relevant work processes, services, or documents in the course of doing business with the District to confirm that the Contractor (and any relevant subcontractors) is complying with this contract. Contractor shall provide access to information, documentation, and personnel required to complete this audit at no additional cost to the District.
11. **Data Security:** The Contractor shall implement appropriate technical and organizational measures to protect the District's from unauthorized access, disclosure, alteration, or destruction.
12. **Data Confidentiality:** The Contractor shall maintain the confidentiality of the District's data, if applicable, and shall not disclose or use the District's data for any purpose other than as expressly authorized in the Contract.
13. **Data Retention Policy:** The Contractor shall implement a data retention policy that specifies the duration for which the District's data will be stored.
14. **Deletion:** Upon termination of the Contract or upon the expiration of the retention period, the Contractor shall delete or return all of the District's data to the District, unless required by law to retain such data.

APPENDIX F – References and Helpful Resources

District of Columbia

- ❖ [AI/ML Governance Policy](#)
- ❖ [Handbook for AI Values Alignment](#)
- ❖ [Mayor's Order 2024-028](#)

External Resources

- ❖ [Brookings](#)
- ❖ [California Generative AI Toolkit](#)
- ❖ [Government AI Coalition](#)
- ❖ [Guidebook for Crafting a Results-Driven RFP](#)
- ❖ [IBM Research AI Factsheets 360](#)
- ❖ [National Association of State Procurement Officials](#)
- ❖ [World Economic Forum AI Procurement Guidelines](#)