Carnegie Mellon University
The Block Center
FOR TECHNOLOGY AND SOCIETY

University of Pittsburgh
Institute for
Cyber Law, Policy, and Security

PROCURING PUBLIC-SECTOR AI

Guidance for Local Governments

April 2025

Elise Silva, PhD: Director of Policy Research at the University of Pittsburgh Institute for Cyber Law, Policy, and Security
Nari Johnson: PhD Student at Carnegie Mellon University
Ravit Dotan, PhD: AI Ethics Advisor at TechBetter
Motahhare Eslami, PhD: Assistant Professor at Carnegie Mellon University
Hoda Heidari, PhD: Assistant Professor at Carnegie Mellon University
Beth Schwanke, JD: Executive Director at the University of Pittsburgh Institute for Cyber Law, Policy, and Security

Acknowledgements

The authors thank Leila Doty, Privacy and AI Analyst with the city of San José, for providing valuable feedback during the drafting process.

The authors thank Carnegie Mellon's Block Center for Technology and Society for grant funds that supported this project.

Table of Contents

Introduction: Context, Scope, and Audience

The past two years have marked an explosion of interest in public sector adoption of artificial intelligence (AI) technologies. Governments nationwide are increasingly procuring enterprise AI solutions, such as decision support tools, resident-facing chatbots, and smart cities technologies, to automate bureaucratic processes.

However, the lines clearly delineating between procurements involving AI have become blurred. Cost is no longer a barrier for government employees due to the increased availability of free, low, or no-cost AI tools like ChatGPT. Software procurements are beginning to roll-out new AI features, like Microsoft Copilot. Finally, vendors are increasingly relying on AI to fulfill their scope of work.

AI has the potential to support public sector goals of increasing access to government services, increased efficiency, or offering novel insights that improve bureaucratic decision-making. However, **AI also poses unique risks relative to traditional software procurements due to its opacity even to experts, dependency on its training data, and propensity for unpredictable outputs.** These risks are compounded by the lack of standardized guidelines for evaluating AI technologies, the potential for bias embedded in training data to perpetuate systemic inequalities, and challenges in ensuring accountability when AI systems fail or produce harmful outcomes.

AI also poses unique risks relative to traditional software procurements due to its opacity even to experts, dependency on its training data, and propensity for unpredictable outputs.

For these reasons, we studied current city government procurement practices across the United States to understand how governments are procuring AI systems and what challenges and barriers they face in responsible procurement of AI.[1] Our methods included interviews of 19 city employees from small, medium, and large cities across the country. Through these conversations, we uncovered several key challenges that governments face, and their needs for support. Our research surfaced a clear need for accessible educational resources that discuss how AI differs from other procured technologies and provide actionable steps that cities can take to revise their purchasing practices for AI.

In this white paper, we discuss the unique governance challenges posed by procured AI systems and provide actionable guidance on first steps that governments can take today to manage these emerging risks. **Our audience for this paper is U.S. local government employees** involved in procurement, IT, innovation, and related departments. We have designed our guide to be accessible as a "first step" for readers who have not yet considered adopting their governments' existing procurement processes for AI.

We designed our resource to complement existing resources and AI governance initiatives, such as the GovAI Coalition's deliverables. In particular, we contribute to the existing literature on AI and procurement by providing a comprehensive and accessible overview of emerging AI risks that is contextualized to the public sector. We share concrete examples of recent actions that local governments have taken to address these novel AI risks.

Understanding AI in Government Contexts

It is important to have a shared understanding of what we mean when we say "AI." For our purposes, we adopt a wide definition of AI as any machine-based system that can make predictions, recommendations, or decisions—a wide and well-known definition adapted from the Organization for Economic Co-operation and Development (OECD).[2] When defining AI, we are interested in any algorithmic system trained on data with human consequences.

There are multiple examples of AI systems used in city government contexts from decision support tools, to resident-facing chatbots, to smart city technologies. Select documented use cases include:

RESIDENT COMMUNICATION/ACCESS

- Translation services [3] [4]
- Meeting summaries and document digitization [5]

LAW ENFORCEMENT

- Predictive policing [6]
- License plate readers and gunshot detection [7]
- Facial recognition [8]
- Automated police reports [9]

SOCIAL SERVICES [10]

- Resource allocation for social work [11]

WORKPLACE PRODUCTIVITY [12]

- Image or voice generation for communication
- Chatbots to automate writing tasks
- AI for code generation

SMART CITIES/URBAN PLANNING

- Public transport/traffic management
- Waste collection
- Noise control
- School bus routing [13]

Why AI is Different from Other Technologies

Because AI systems involve software, many cities have applied their existing review processes for software technologies to assess the risks posed by AI. While many of these existing review processes involve important considerations like cybersecurity and privacy that are still applicable for AI technologies, **AI systems also pose unique risks that differ from traditional software procurements**, due to how AI models are designed and developed.

In the past, software systems were comprised of rules that were hand-coded by developers, who often understood more precisely how particular system inputs mapped to system outputs. While software systems could have "bugs" that resulted in unexpected behaviors, in many situations, software systems could (in theory) be programmed to be "accurate" by design.

In contrast, many modern AI systems are developed using machine-learning algorithms that map inputs to outputs using a set of rules that are not hand-coded but are "learned" using optimization processes to maximize performance on a dataset during model "training". The rules that AI systems use to map inputs to outputs often involve thousands of mathematical operations that are inscrutable even to their developers. As a result, **AI systems often have behaviors that are not fully predictable or controllable by those that make them.**

Given their potential use of high-risk systems that could affect the rights of citizens, governments should be especially aware of the unique challenges they will face in acquiring and deploying AI-driven technologies.

In this section, we provide an overview of several unique risks that may be relevant to consider for procurements involving AI, relative to a standard technology procurement. When possible, we contextualize our discussion using examples from the public sector. Below each risk category, we also provide links to further reading.

Unique Risks of AI

ACCURACY AND FAIRNESS

AI systems are only as good as the data they are trained on. As such, they can be wrong and/or biased due to imperfect or limited training data. Much data is collected by, cleaned by, and structured by humans, leading to different values being represented (or not represented) in the sets.

This affects the quality and slant of algorithmic outputs. Biases exist not only in the data, but also in the context in which AI systems are deployed, making it even more important to consider how fair and accurate they really are.

SOME BIASES CAN BE MITIGATED AND UNDERSTANDING HOW THEY ARE EMBEDDED IN AI SYSTEMS IS AN IMPORTANT PART OF THAT PROCESS.

REFERENCES AND FURTHER READING

- The Fallacy of AI Functionality (2022), by Inioluwa Deborah Raji et al. *ACM Conference on Fairness, Accountability, and Transparency*.
- Towards a Standard for Identifying and Managing Bias in Artificial Intelligence (2022), published by *NIST*.
- Combatting Bias in AI: Helping Civil Servants Mitigate Bias for Equitable Use of AI (2022), published by *The United States General Services Administration*.[14]

TRANSPARENCY

Machine learning AI systems operate as “black boxes." This makes it difficult for humans (even experts) to understand how the systems create their outputs and to relay such logic to anyone impacted by such an output. This opacity creates unique challenges in transparently explaining to the public how decisions are made.

Another dimension has to do with how to tell residents when AI is being used and to what effect. Disclosing AI use in a timely, accessible, and meaningful way can be challenging as policies must balance making AI use visible and transparent while at the same time not overwhelming the public.

REFERENCES AND FURTHER READING

- Federal Procurement of Artificial Intelligence: Perils and Possibilities. (2020), by David S. Rubenstein. *The Great Democracy Initiative*.
- Procurement as Policy: Administrative Process for Machine Learning (2019), by Deidre K. Mulligan & Kenneth A. Bamberger. *Berkley Tech*.

Unique Risks of AI

ACCOUNTABILITY

A central tenet of governance is ensuring those in positions of power are accountable for their actions, decisions, and performance. However, with the opacity of AI systems, coupled with the multilayered nature of their development and deployment, new challenges related to responsibility emerge.

AI systems involve an extensive ecosystem including engineers, vendors, buyers, assessors, and those who deploy and maintain them. This begs the question: who is accountable for what, especially if something goes "wrong" In essence, AI is creating new forms of legal and social liability that have yet to be widely applied.

REFERENCES AND FURTHER READING

- When AI Gets it Wrong, Will It Be Held Accountable? (2024), published by *RAND*.
- Dislocated Accountabilities in the "AI Supply Chain": Modularity and Developers' Notions of Responsibility. (2023), by David Gray Widder & Dawn Nafus. *Big Data & Society*.

OVER-RELIANCE ON ALGORITHMIC SYSTEMS

Even when AI decisions are not used to fully replace human discretion, cognitive science research has shown that humans tend to "over-rely" on AI outputs, even when they are nonsensical or wrong. While humans often develop their own "mental models" of an AI system's behavior, these mental models often have inaccuracies due to the complex and often surprising decision-making logics encoded by AI.

More generally, due to popular misconceptions about AI's capabilities, workers often attribute an unwarranted degree of authority to AI that requires education to be corrected.

REFERENCES AND FURTHER READING

- Overreliance on AI: Literature Review. (2022), by Samir Passi & Mihaela Vorvoreanu. *Microsoft AI Ethics and Effects in Engineering and Research*.
- Is AI in the Eye of the Beholder? (2023) by Adam Zewe, *MIT News*.
- "Because AI is 100% Right and Safe": User Attitudes and Sources of AI Authority in India (2022), by Shivani Kapania et al. *CHI Conference on Human Factors in Computing Systems.*
- An Overview of Catastrophic AI Risks. (2023), by Dan Hendrycks, Mantas Mazeika & Thomas Woodside. *Center for AI Safety*.

Unique Risks of AI

PRIVACY

AI systems may leak private or proprietary information that was present in their training data; or can be mis-used to create content that infringes on individuals' privacy. For example, large language models have been shown to "memorize" phrases present in their training datasets. In one prominent attack, researchers were able to extract personal information about individuals from the ChatGPT model.

This may pose risks if proprietary government data (for example, about government business or data about residents) is entered into AI models, and as a result winds up in the model's training dataset. Further, data is valuable, and vendors may want to retain and train their models on government employees' or residents' data, leading to worries about data ownership and surveillance.

REFERENCES AND FURTHER READING

- Deepfakes, Phrenology, Surveillance and More! A Taxonomy of AI Privacy Risks. (2024), by Hao-Ping (Hank) et al. *Proceedings of the CHI Conference on Human Factors in Computing Systems*.
- What Does it Mean for a Language Model to Preserve Privacy? (2022), by Hannah Brown et al. *ACM Conference on Fairness, Accountability, and Transparency (FaaCT)*.
- Scalable Extraction of Training Data from (Production) Language Models. (2023), by Milad Nasr et al. *ArXiv Preprint*.

USE MISALIGNMENT

While much software is designed for specific use cases, general-purpose AI models can be applied in contexts beyond their initially procured use parameters. This can introduce risks as oftentimes vendors do not restrict how foundation models are applied. If they are used in ways models were not designed, tested, or risk-assessed for, the results can be misaligned with organizational values.

REFERENCES AND FURTHER READING

- Lessons Learned from Ten Generative AI Misuse Cases. (2024), by Sameer Hinduja. *Cyberbullying Research Center*.

Unique Risks of AI

CYBERSECURITY

AI models are susceptible to a novel range of attacks from adversaries, who can manipulate AI systems to produce their desired behaviors. For example, adversaries can "game" AI systems by altering inputs to avoid adverse determinations, such as avoiding detection of fraudulent behavior.

Similarly, adversaries can conduct "data poisoning" attacks where they strategically alter the AI's training data to change the modified system's behavior. Large language models are particularly prone to "jailbreak" attacks that elicit undesired behaviors that these models were specifically trained to avoid, such as answering requests for harmful information, aiding crime, or leaking organization data.

REFERENCES AND FURTHER READING

- Cybersecurity Training: AI Risks Video. (2024), by the *City of San José* and the *GovAI Coali*tion.
- The Path to Defense: A Roadmap to Characterizing Data Poisoning Attacks on Victim Models. (2024), by Tarek Chaalan et al. *ACM Computing Surveys*.
- Jailbroken: How Does LLM Safety Training Fail? (2023), by Alexander Wei et al. *ArXiv Preprint.*

ENVIRONMENTAL IMPACT

AI has a unique risk of environmental impact due to the extraction of materials needed to manufacture the material compute, the massive computational power required to train large models, and the energy needed when users interface with the technology, all of which leads to high energy consumption and carbon emissions.

Data centers are particularly known for their high energy usage due to the need for constant cooling and uninterrupted power supply.

AS OF NOW, AI'S ENERGY USAGE IS OUTPACING EXISTING POWER INFRASTRUCTURE.

REFERENCES AND FURTHER READING

- The Environmental Impacts of AI – A Primer. (2024), by Sasha Luccioni et al. *Hugging Face Community Article*.
- Carbon Emissions in the Tailpipe of Generative AI. (2024), by Tamara Kneese and Meg Young. *Harvard Data Science Review*.
- Net 0++: Big Dirty Data Centres. (2024), hosted by Alix Dunnl. *Says Maybe Podcast*.

Unique Risks of AI

JOB DISPLACEMENT

While robots are unlikely to replace all public sector jobs anytime soon, AI-driven automation is already reshaping human workflows. Some organizations have explicitly replaced human workers with AI, such as using chatbots instead of hotline staff or automating translations services. AI's broadest impact may be less about outright job elimination, and more about fundamentally shifting the nature of the work people do. This reallocation of labor creates unique challenges like the need for reskilling and upskilling.

However, the rapid pace of technological advancement makes it difficult to predict which skills will be most valuable. Different sectors will likely experience the impact of AI-driven automation unevenly, potentially creating economic disparities. The psychological impact on workers who feel threatened by automation is also a concern in a time of rapid transition.

REFERENCES AND FURTHER READING

- AI Eliminated Nearly 4,000 Jobs in May, Report Says. (2023), by Elizabeth Napolitano. *CBS News*.
- 4 in 10 Translators are Losing Work to AI. They Want Remuneration from Devs. (2023), by Chloe Xiang. *Vice*.
- Machines Will Do More Taks Than Humans by 2025, but Robot Revolution Will Still Create 58 Million Net New Jobs in Next Five Years. (2018), by Oliver Cann. *World Economic Forum*.
- The Future of Work in the Age of AI: Displacement or Risk-Shifting? (2020), by Pegah Moradi & Karen Levy. *Oxford Handbook of Ethics of AI*.
- Reskilling in the Age of AI. (2023), by Jorge Tamayo et al. *Harvard Business Review*.
- A.I. is Going to Disrupt the Labor Market. It Doesn't Have to Destroy It. (2023), by Rebecca Stropoli. *Chicago Booth Review*.

Four Considerations When Procuring AI Systems

AI systems pose a unique set of risks to public sector users, residents, and society at large. How can governments adapt their practices to prepare for these risks, and protect the public interest?

From our discussions with government employees, we learned that local governments across the U.S. face shared challenges in revising their procurement practices for AI, such as a lack of clear accountability structures for acquisitions under cost thresholds, or a lack of expertise about the unique risks posed by AI. We also learned about several promising steps that governments have taken to address these challenges by setting up new AI governance strategies, oversight mechanisms, and risk assessments.

In this section, we discuss four key considerations for governments as they revise their procurement practices for AI. We believe that taking action to address these four points can establish a strong foundation to enable governments to foresee and prevent AI harms.

We recognize that all governments are different, and that it is difficult to formulate generalizable recommendations on how governments ought to structure and resource their AI governance efforts. As such, we point towards concrete examples of ongoing AI governance initiatives from local governments across the United States to inform different approaches. Using these examples, we share important take-aways and learnings for cities hoping to revise their AI practices.

1. DOES YOUR GOVERNMENT HAVE AN AI GOVERNANCE STRATEGY?
2. ARE YOU AWARE OF, AND REVIEWING, ALL AI ACQUISITIONS?
3. HOW ARE YOU CONDUCTING AI RISK ASSESSMENTS?
4. ARE YOU ENGAGING IMPACTED STAKEHOLDERS THROUGHOUT THE PROCUREMENT PROCESS?

Four Considerations When Procuring AI Systems

1 Does your government have an AI governance strategy?

The term “AI governance” describes the policies, processes, procedures, and practices across your government so that employees are empowered, responsible, and trained to manage AI impacts and risks[15]. It entails preparing your workforce to understand these novel risks of AI solutions and establishing a consistent culture across departments that considers and communicates AI risks.

In our conversations with government employees, we found that organizations with a strong culture of risk management with clearly articulated accountability structures were much more likely to foresee and take action to mitigate potential risks posed by AI. In many such cities, expert employees who were trained on how to assess proposed AI solutions played a critical role in supporting their colleagues across departments throughout their AI procurements.

AI governance: the policies, processes, procedures, and practices across your government that guide and empower employees to manage AI impacts and risks.

LEARNING FROM OUR PEERS

ESTABLISH AN ORGANIZATIONAL AI POLICY

Many cities' AI governance efforts began by establishing a government-wide AI policy. AI policies established consistent standards for all AI systems used by the organization. They often articulated a common definition for "AI", guiding ethical principles, and requirements for the organization's AI use.

Some policies also included explicit implementation details, for example, by naming clear roles and responsibilities, or allocating funding towards AI governance efforts.

Note that "AI policies" are different (and oftentimes more encompassing) than generative AI *usage* policies (such as the City of Boston's), which specify the appropriate conditions in which specific types of AI tools, like chatbots, can be used.

EXAMPLE: The GovAI Coalition developed a customizable AI Policy template. It contains a common definition of AI, outlines guiding "responsible AI" values, designates roles and responsibilities for those who are responsible for coordinating AI reviews and governance efforts, and states requirements for AI vendors under contract with the city.

The coalition recommends that cities modify the template to designate who specifically within their organization (e.g. the CIO) will be responsible for overseeing the policy's implementation.

EXAMPLE: The City of Tempe's Ethical AI Policy states the city's commitment to "designing, developing, and deploying AI technologies in a responsible and ethical manner".

To do this, the policy establishes clear roles, responsibilities, and accountability structures. For instance, all city departments are required to collaborate with IT employees who will complete a mandatory AI Review process for every acquisition.

This policy, one of the first of its kind, was adopted by City Council in June 2023.

LEARNING FROM OUR PEERS

ESTABLISH REVIEW FOR ALL AI ACQUISITIONS

Assessing the risks of employees' or vendors' AI usage requires proactive and continuous review. Someone should be in the room to ask important questions about novel AI risks like inaccuracies, privacy and security risks, and others, both before acquisition, and throughout the AI's deployment. If your city has not had these conversations before, this may also entail doing ongoing reviews of the AI that your employees are already using today.

Different cities have taken different approaches to how they delegate responsibility for these reviews. For example, in some cities this might happen through formal AI policy, or in others, through another authoritative mandate like a city council directive.

In many cities, employees across departments contact AI governance champions (whom are often IT employees) to conduct reviews. These AI governance champions are then able to maintain a centralized, holistic view of what AI is being used across the city. We further discuss different ways that such AI reviews can be conducted and organized in the following two sections.

EXAMPLE: The City of San José's policy establishes mandatory review for systems that are algorithm-based. Departments wishing to procure an AI system submit a procurement proposal that is reviewed by IT divisions, including privacy and AI, cybersecurity, architecture, etc. The Chief Privacy Officer then facilitates the risk analysis and AI review for the proposed system.

LEARNING FROM OUR PEERS

INVEST IN TRAINING AI GOVERNANCE CHAMPIONS

Anticipating the potential societal impacts of AI systems, interpreting information reported by AI vendors, and conducting continued system monitoring all requires specialized training.

Many cities began this learning process by designating specific personnel as “AI governance champions”. These knowledgeable employees could then be consulted or brought in to advise various AI procurements.

EXAMPLE: AI governance champions can form a community of practice within your city to share learnings and participate in professional development events together.

One great way to learn about AI is to join peer networks such as the GovAI Coalition, or the MetroLab Network. Budget can be allocated for employees to attend relevant conferences and professional development events, such as the GovAI Coalition summit.

EXAMPLE: In addition to specialized training for some, several cities have created generalist AI trainings to socialize their new AI policies, with everyone from administrators, to engineers, to public safety officials. For example, IT employees in one medium-sized city we interviewed used one of their city’s monthly all-hands meetings to speak briefly about their new AI policy and show a demonstration of generative AI tools.

AI governance champion: a designated individual or office who is adequately trained and responsible for ensuring that AI acquisition and deployment align with ethical, legal, and policy standards.

Four Considerations When Procuring AI Systems

2 Are you aware of, and reviewing, all AI acquisitions?

Understanding the scope of AI use in your organization may be trickier than you think—especially if low or no-cost AI systems are being purchased and used without having gone through a formal acquisition process. For instance, cities across the U.S. have obtained predictive policing technologies for free from philanthropic donations, academic collaborations, and other acquisition gateways that did not go through the typical oversight processes of a formal procurement.

Shadow IT—technology or software used within an organization without explicit approval or oversight—can make it difficult to track AI adoption or compliance. For example, several AI chatbots or media generation tools like ChatGPT or Canva Image Generator have free versions that users can access online, without purchasing any software. Employees might be using AI-powered tools to aid critical decision-making processes, draft reports, generate images, or analyze data without realizing the potential risks in doing so. Additionally, technologies that have already been acquired may be "upgraded" without your knowledge, adding new AI features that were not part of the original product. Finally, even when AI is not explicitly provided as a service, vendors may use AI to fulfill their scope of work.

Because AI has evolved so quickly, there may be gaps in understanding the full scope of AI use within your organization. Ensuring the safe and effective use of AI starts with understanding where AI is already in use and establishing clear protocols for future acquisitions, regardless of cost.

Shadow IT: technology or software used within an organization without explicit approval or oversight.

LEARNING FROM OUR PEERS

ORGANIZE PRE-ACQUISITION AI REVIEWS AROUND EXISTING PROCUREMENT INFRASTRUCTURE, LIKE E-PROCUREMENT SOFTWARE OR TICKETING PROCESSES

IT procurement is organized very differently across U.S. cities. Some cities have established procurement infrastructure and workflows (like e-procurement software) that enables centralized oversight. In these cities, integrating AI review into the existing e-procurement systems might be as easy as asking one or two extra questions to flag new AI systems being acquired. By utilizing familiar structures already in place to identify AI tools, cities can: 1) ensure proper vetting of the systems, and 2) be able to catalog all AI being used.

For cities without centralized IT procurement oversight, we suggest considering establishing a centralized process for AI reviews. This will require an understanding of how procurement operates across different parts of your organization and promoting the importance of AI review within these existing workflows. One path forward is to consider how different departments within your cities currently procure goods and services and identify patterns or overlap you might leverage as familiar intervention points for centralized oversight.

EXAMPLE: One interviewed city incorporated their AI review into their existing e-procurement software. When employees stepped through their existing workflow to make a request to make a purchase, they checked a box to note if their purchase involved AI. These requests were then directed to trained reviewers through the city's existing software.

LEARNING FROM OUR PEERS

INCLUDE LANGUAGE IN CONTRACTS TO REQUIRE VENDORS TO NOTIFY YOUR ORGANIZATION OF ANY USE OF AI TO FULFILL THEIR SCOPE-OF-WORK

As AI evolves, vendors are quickly integrating its capabilities into their products, sometimes after you have already acquired them. Further, sometimes systems are complex, and vendors are not clear about what systems might have AI features integrated into them or how they work.

Contracting can be a place where you require that vendors disclose any AI components used in their tools. They should also be transparent about new AI features they might roll out in tools already acquired and provide "opt out" options for cities that do not wish to use such features. Many cities are still working on writing contracting language for AI, and as such it is a ripe area for development and leadership.

EXAMPLE: One example of language to include in solicitations comes from the State of California's Generative AI Toolkit (see page 23 of the Toolkit for Procurement, Use, and Training).

LEARNING FROM OUR PEERS

COLLECT INVENTORIES BY EXAMINING CONTRACTS AND PURCHASES OF IN-USE AI TOOLS

Creating an inventory of the AI systems in use by city employees —even if retroactively—can help identify AI acquisitions that were not subject to the usual procurement process such as those acquired through piggybacking contracts or bought under a cost threshold using a purchasing card, or even ones subscribed to by individual employees themselves.

As a first step, you might establish processes to identify all incoming AI by creating documentation and tracking standards. Establishing inventory architectures to identify AI in ongoing and future acquisitions is a first step to getting ahead of unchecked acquisitions.

Once governance has been established, then your city might consider retroactively mapping the AI that has already been acquired. By systematically reviewing past contracts and purchases, and surveying employees about how they are using AI in their workflows, you will gain a broader picture of how and where AI is being implemented in your city.

EXAMPLE: One interviewed city opened a form to anonymously survey employees about what tools they had already implemented into their workflows for the purpose of understanding the scope of AI usage. This exploratory approach avoided a sense of employee surveillance or punitive tone that could pit leaders against staff. After understanding the scope of AI use, the city planned to use that knowledge to influence and enforce usage policies.

LEARNING FROM OUR PEERS

SOCIALIZE THE EXPECTATION THAT FREE OR LOW-COST TOOLS ARE NOT EXEMPT FROM REVIEW.

All AI acquisitions, regardless of cost, should be expected to go through a review process. While p-cards can often be used for purchases under a certain cost threshold, or piggybacking contracts are used to acquire similar services using already established contracts from another government agency, all AI tools should be thoroughly vetted given your city's unique context and needs.

Cultivating a culture of consistent review of AI tools, even acquisitions that might seem innocuous, will help with ongoing oversight and scrutiny of these systems.

EXAMPLE: In one interviewed city, employees are required to submit a ticket that is reviewed by IT employees before they can use any AI system to complete their work. Reviewing employees do research into the systems and employees' use cases to proactively identify and mitigate risks. The city also maintains a list of already-approved tools that employees can adopt without further review.

3 How are you conducting AI risk assessments?

AI risk assessments describe structured processes that your government can follow to identify both the potential benefits and harms posed by procured AI. Earlier in this white paper we mapped how AI can pose a wide set of novel risks relative to traditional software procurements. Robust AI risk assessment processes can help your organization get specific in understanding the kinds of risks that are the most likely to occur for a specific AI system, given the specific technology that is adopted, and context of its use.

For example, a government interested in procuring a computer vision AI tool to detect potholes in streets using street cameras may consider the privacy risks to residents of collecting such video, and the equity implications if the tool can accurately detect potholes in some types of neighborhoods (e.g., those with newly paved roads) but not others. When done early (i.e., before acquisition and adoption), risk assessments can help inform purchasing decisions and socio-technical mitigations to request of AI vendors, such as monitoring model performance across neighborhoods, or arranging camera feeds so they do not capture pedestrian traffic. Governments should consider risk assessment and mitigation to be an ongoing process that continues even after the technology is deployed, as it is often difficult to fully anticipate a system's impacts before deployment.

In this section, we share practical implementation guidance for AI risk assessments: how to organize them and who should be involved.

AI risk assessments: structured processes that your government can follow to identify both the potential benefits and harms by procured AI.

LEARNING FROM OUR PEERS

STRUCTURE AND DOCUMENT YOUR RISK ASSESSMENT PROCESSES USING A CONSISTENT TEMPLATE

Mapping the possible impacts—both positive and negative—of an AI system is often difficult to do without a starting place or additional guidance. To provide consistent structure to their discussions about AI risk, several cities created their own AI risk assessment instruments. These instruments often consisted of lists of questions that guided employees in imagining specific types of impacts (e.g., to individual users of the technology, impacted communities, and society). Completed AI risk assessments that contain written documentation of potential risks and mitigation strategies could facilitate communication between city employees.[16]

Below, we provide several examples of risk assessment instruments that your government can use as a starting place for creating your own risk assessment process. Note that several risk assessments can be initiated during the planning phases of procurement, e.g., once a potential use case for AI is identified, even before your city has identified a specific system, vendor offering, or AI model.

EXAMPLE: One sample AI impact assessment template that has been developed for public sector agencies is from the city of San José.

Cities can create their own risk or impact assessments by modifying the template linked here.

EXAMPLE: The Canadian federal government has developed an algorithmic impact assessment tool that contains a list of 51 risk and 34 mitigation questions. The tool was created to help government departments and agencies address policy, ethical, and administrative law implications for proposed AI systems. The tool also assesses the mitigation measures in place to manage the risks identified.

LEARNING FROM OUR PEERS

TRIAGE SYSTEMS INTO "LOW" OR "HIGH" RISK CATEGORIES

Several risk assessment instruments sort proposed AI solutions into categories of "low" versus "high" risk, based on the nature, likelihood, and severity of their potential to cause harm.

For instance, a department's use of an algorithm used to determine residents' access to critical government services may be considered "high risk", while an employees' use of an AI writing tool to make grammar suggestions may be considered "low risk".

Triaging AI systems can increase process efficiency by enabling expert reviewers to focus their energy on managing the riskiest acquisitions. Risk categories also allow organizations to institute additional governance requirements (such as public engagement) for high-risk systems.

EXAMPLE: The GovAI Coalition's AI Governance Handbook proposes beginning the risk assessment by triaging to determine if the system into should qualify for a "full-fledged AI review". Systems deemed "minimal risk" can be approved without further review.

The handbook puts forward a "AI Risk Threshold Analysis" method to triage systems into low vs. high risk using two axes: (1) the impacted individual's (in)ability to opt-out of AI use, and (2) the severity of potential harm.

The handbook provides several examples of systems that either qualified or did not qualify for needing an AI review.

LEARNING FROM OUR PEERS

TRIAGE SYSTEMS INTO “LOW” OR “HIGH” RISK CATEGORIES

EXAMPLE: The Biden administration’s Office of Management and Budget provided procurement guidance to federal agencies in their (since-rescinded) AI M-Memo.

The memo specifies that AI systems that qualify as “rights- and safety-impacting" must comply with a higher standard of review. The memo defines which systems qualify as “rights-impacting” and “safety-impacting”, categories that consider the extent to which AI outputs have a legal, material, binding, or significant effect on individuals’ rights, life and wellbeing, and ability to access critical government services.

LEARNING FROM OUR PEERS

INVOLVE A WIDE VARIETY OF INTERNAL STAKEHOLDERS IN BOTH MAPPING RISKS AND IMAGINING MITIGATIONS

AI systems and their impacts are fundamentally socio-technical in nature. They concern not just questions of technical engineering but also understanding the social systems that the AI is deployed within.

For example, to understand the risks posed by introducing a 311 chatbot requires a deep understanding of the communication services being automated: for instance, which residents have already historically had difficulties communicating with government employees (e.g., due to language barriers) and accessing government services. We encourage your government to reflect on who within your organization holds expertise about the domain in which AI is being applied, and the communities who may stand to be most impacted.

EXAMPLE: When conducting a risk assessment for an AI technology to be procured by their city's public safety department, the City of Tempe invited their Chief Diversity Officer to participate.

The CDO's role was constructed to understand and advocate on behalf of minoritized groups. Together, IT reviewers and the CDO discussed how the risks posed computer vision technologies used by local law enforcement. The CDO also participated in imagining mitigation steps and usage restrictions to reduce the possibility of misuse.

4 Are you engaging impacted stakeholders throughout the procurement process?

Multiple groups are potentially impacted by AI procurement, including employee and end users and the boarder community. Consulting those who will use the technology, those whose workflows will be shaped by these systems, is a good first step in gathering feedback. Beyond these consultations, those impacted should also have opportunities to participate in some procurement processes. While not all AI procurements require public involvement, high-risk procurements that could affect rights or have direct consequences for residents warrant public engagement.[17]

Many cities we interviewed expressed concerns about involving the public in procurement decisions citing legal risks, staffing limitations, and public education challenges. Some cities may have laws or regulations that restrict the ability to share certain types of information publicly throughout the procurement process. To determine where public input is feasible, your city can start by consulting procurement officials, legal advisors, and AI experts to identify points in the process where engagement aligns with existing policies and operational constraints. Engagement ideas might include community advisory boards, public comment periods, or pilot testing with end users.

We recognize the fear of public scrutiny is real. However, meaningful public engagement can build trust, reduce backlash that might otherwise lead to the abandonment of truly useful systems (especially if the public learns about them only after implementation), and enhance accountability and transparency for high-risk AI tools. Involving end users—such as city employees—in testing and acquisition ensures the technology improves their ability to perform their jobs effectively and fosters a sense of buy-in.

Stakeholder engagement ideas: community advisory boards, public comment periods, or pilot testing with end users.

LEARNING FROM OUR PEERS

CORRELATE THE LEVEL OF POTENTIAL RIGHTS, LIBERTY, AND SAFETY-IMPACTING RISK THE SYSTEM POSES TO THE LEVEL OF PUBLIC ENGAGEMENT NEEDED

The higher risk the system, the more public involvement should be sought, ranging from inviting feedback to actively soliciting it.

EXAMPLE: From our discussions with cities, some utilized a high/medium/low-risk scale to help them determine which acquisitions needed more public input.

A high-risk system might be something like a pre-trial algorithmic risk assessment which could influence if someone would go to jail; a medium risk procurement might be something like an automated license plate reader used to collect and document vehicular data; and a low-risk system might be something like an urban planning traffic light algorithm used to influence traffic flows.

LEARNING FROM OUR PEERS

ENGAGE FEEDBACK THROUGHOUT THE LIFECYCLE OF THE PRODUCT

Be sure to engage feedback throughout the lifecycle of the product from acquisition to deployment, and through any substantive changes the system undergoes.

EXAMPLE: One city hired a community engagement consultant to help liaise with the public.

Another described using surveys to engage with residents about high-risk tools before planning in-person public engagement activities.

EXAMPLE: The City of Long Beach runs the LB Co-Lab Program, a community-centered civic-engagement model that brings together residents who learn about technology and deploy a technology project in their neighborhood. They are engaged in the project from start to finish—identifying a community need, designing a project proposal, assessing different technology solutions, and evaluating vendors.

LEARNING FROM OUR PEERS

SEEK INPUT TO ENSURE USEFULNESS

Seek input to make sure the system will be useful to impacted communities and to the workers using it.

> **EXAMPLE**: In one city in our study, a translation system was procured to provide better access for a sub-community within that city, but it turned out that the system did not include the right dialect for the intended population, so it wasn't useful.
>
> In other cities, end users mentioned that they were not consulted in the procurement process and as such the AI tool they were expected to use was not tailored enough for their needs to be helpful in their actual work.

PUBLISH INFORMATION ABOUT AI SYSTEMS PUBLICLY

AI system information should be publicly available to ensure transparency.

> **EXAMPLE**: Interviews mentioned several ideas to boost AI transparency such as cities proactively posting impact assessments, publishing fact sheets or registries outlining AI systems in use, or using marketing departments to spread word about the systems being acquired and deployed by the city.

Find suggestions and examples of AI registries at the University of Pittsburgh's Institute for Law, Policy, and Security **AI & Algorithmic Registries** webpage.

Conclusion and Further Resources

The adoption of AI technologies by city governments presents significant opportunities to spread resources further and optimize workflows as well as poses unique challenges in terms of new risks. City governments must work to establish robust AI governance frameworks, conduct thorough risk assessments, and engage the public throughout the AI procurement and deployment process.

One issue that arose in our interviews was that many cities felt they had lack of leverage when dealing with AI vendors. They reported vendors unwilling to provide detailed information about their AI systems due to “trade secrets” or “intellectual property” concerns. This made it hard for cities to make informed decisions about what systems to buy in a safe and responsible manner. To combat this issue, cities can band together to increase their collective bargaining power in the way that the GovAI Coalition has done. This is one way to demand greater transparency and accountability from vendors and learn from one another.

The shared experiences and strategies from various city governments highlighted here provide some models for how cities are seeking to procure AI systems that serve the public interest. For more resources and guidance, please see a curated list of public procurement-related tools and frameworks in the Appendix: Practical AI-Related Procurement Resources on page 36.

Endnotes

1 See our research paper Legacy Procurement Practices Shape How U.S. Cities Govern AI: Understanding Government Employees' Practices, Challenges and Needs, by Johnson et al. accepted into the 2025 *ACM Conference on Fairness, Accountability, and Transparency*: https://arxiv.org/pdf/2411.04994

2 Artificial Intelligence & Responsible Business Conduct (2019), published by the *OECD*: https://mneguidelines.oecd.org/RBC-and-artificial-intelligence.pdf

3 New Multilingual Chatbot Expands Access to City Services (2023), by Thad Rueter. *Industry Insider*: https://insider.govtech.com/california/news/new-multilingual-chatbot-expands-access-to-city-services

4 Cities Bring Live, Automated Translation to Public Meetings (2023), by Jule Pattison-Gordon. *Government Technology*: https://www.govtech.com/civic/cities-bring-live-automated-translation-to-public-meetings

5 Cities Using AI for Transparency, Resident Engagement (2025), by Skip Descant. *Government Technology*: https://www.govtech.com/artificial-intelligence/cities-using-ai-for-transparency-resident-engagement#:~:text=Cities%20are%20using%20artificial%20intelligence,language%20that's%20relevant%20for%20residents

6 Note predictive policing technologies are especially controversial; Predictive Policing Explained (2020), by Tim Lau. *Brennan Center for Justice*: https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained

Endnotes

7 How Cities Can Pair Audio Detection with AI-Powered License Plate Recognition and Video Technology to Revolutionize Addressing Gun Violence (2024), by Josh Thomas. *National League of Cities*: https://www.nlc.org/article/2024/07/08/how-cities-can-pair-audio-detection-with-ai-powered-license-plate-recognition-and-video-technology-to-revolutionize-addressing-gun-violence/

8 Facial Recognition Guide for Cities (2021), published by the *National League of Cities*: https://www.nlc.org/wp-content/uploads/2021/04/FacialRecognitionSummary_NLC.pdf

9 Police Officers Are Starting to Use AI Chatbots to Write Crime Reports. Will They Hold Up in Court? (2024), by Sean Murphy & Matt O'Brien. *The Associated Press*: https://apnews.com/article/ai-writes-police-reports-axon-body-cameras-chatgpt-a24d1502b53faae4be0dac069243f418

10 Generative Urban AI Is Here. Are Cities Ready? (2024), by Timothy Papandreou. *Forbes*: *https://www.forbes.com/sites/timothypapandreou/2024/02/18/generative-urban-ai-is-here-are-cities-ready/*

11 Social Workers in England Begin Using AI system to Assist Their Work (2024), by Robert Booth. *The Guardian*: https://www.theguardian.com/society/2024/sep/28/social-workers-england-ai-system-magic-notes

12 State and Local Governments Are Using AI for Work. But Should They? (2023), by Libby Denkmann & Alec Cowan. *KUOW*: https://thankyou.kuow.org/stories/more-governments-are-using-ai-for-work-but-should-they

Endnotes

13 How One School District is Turning to AI to Solve its Bus Driver Shortage (2024), by Meg Oliver & Analisa Novak. CBS News: https://www.cbsnews.com/news/how-one-school-district-is-turning-to-ai-to-solve-bus-driver-shortage/

14 Note the toolkit was taken offline in 2025, but is still available on the Internet Archive.

15 See the "Govern" section of the NIST AI Risk Management Framework Playbook in the Trustworthy & Responsible AI Resource Center, published by the *National Institute of Standards and Technology*: https://airc.nist.gov/airmf-resources/playbook/govern/

16 Raji et al. (2020) discusses the importance of documenting and writing down potential risks that might be caused by an AI system (e.g., "creating document trails") as an integral part of system governance in Closing the AI Accountability Gap; Deng et al. (2024) similarly discuss the importance of documenting AI impacts to facilitate multi-stakeholder communication in Supporting Industry Computing Researchers in Assessing, Articulating, and Addressing the Potential Negative Societal Impact of Their Work.

17 Report of the Pittsburgh Task Force on Public Algorithms (2022), published by *Pitt Cyber*: https://www.cyber.pitt.edu/algorithms

Appendix: Practical AI Procurement-Related Resources

AI PROCUREMENT LAB REPOSITORY

(announced 2025), published by AIPL

A set of resources for anyone looking for best practices while procuring high-risk AI solutions. The resources include risk mitigation frameworks, indexes, databases, articles, practical procurement tools, educational tutorials, policy and standards, use cases, and more.

AI PROCUREMENT IN A BOX: AI GOVERNMENT PROCUREMENT GUIDELINES

(2020), published by the World Economic Forum

An AI governance toolkit and accompanying workbook with guidelines that can be incorporated into an AI policy, including key questions to consider when assessing the risks of procured AI. These resources contain templates to help government employees structure AI risk assessments, questions to ask of vendors, and detailed case studies.

BEST PRACTICES FOR GOVERNMENT PROCUREMENT OF DATA-DRIVEN TECHNOLOGIES: A SHORT GUIDANCE FOR KEY STAGES OF GOVERNMENT TECHNOLOGY PROCUREMENT

(2021), by Rashida Richardson

A guide that contains lists of questions for employees to consider at different phases of the procurement process (e.g., pre-solicitation; evaluation; contract negotiation, etc.). The guide also contains an appendix that summarizes other resources relevant to AI procurement.

GENAI FOR CALIFORNIA TOOLKIT

Published by the State of California

A "choose your own journey" toolkit that provides pathways through procuring various kinds of AI systems. Each pathway shows entities how to move forward and procurement steps to consider.

GOVAI COALITION TEMPLATES AND RESOURCES

Published by the City of San José

Various templates and resources for procuring and using AI in government. The resources are created specifically for government agencies. They include policy and use case resources including incident response plans, fact sheets, and vendor agreements.

A GUIDING FRAMEWORK TO VETTING PUBLIC SECTOR TECHNOLOGY VENDORS

(2024), published by the Ford Foundation

The framework includes a list of questions to ask of vendors that support evaluation of new technology-based proposals. It outlines a list of red flags to consider when deciding which proposals to fund.

RESPONSIBLE AI QUESTION BANK: A COMPREHENSIVE TOOL FOR AI RISK ASSESSMENT

(2024), by Sung Une Lee et al. ArXiv Preprint

An AI risk assessment toolkit that incorporates a question bank and framework designed to support many kinds of AI initiatives. This toolkit considers emerging regulations and incorporates best practices other well-known frameworks.

SITUATE AI GUIDEBOOK

(2024), by Anna Kawakami et al. Carnegie Mellon University

This guidebook is an in-development resource by academic researchers. It is being developed to help public sector agencies decide in early stages if they should move forward with developing or implementing a new AI tool.