Legacy Procurement Practices Shape How U.S. Cities Govern AI: Understanding Government Employees' Practices, Challenges, and Needs

NARI JOHNSON, Carnegie Mellon University, USA

ELISE SILVA, University of Pittsburgh Institute for Cyber Law, Policy, and Security, USA

HARRISON LEON, Carnegie Mellon University, USA

MOTAHHARE ESLAMI, Carnegie Mellon University, USA

BETH SCHWANKE, University of Pittsburgh Institute for Cyber Law, Policy, and Security, USA

RAVIT DOTAN, TechBetter, USA

HODA HEIDARI, Carnegie Mellon University, USA

Most AI tools adopted by governments are not developed internally, but instead are acquired from third-party vendors in a process called public procurement. In this paper, we conduct the first empirical study of how United States cities' procurement practices shape critical decisions surrounding public sector AI. We conduct semi-structured interviews with 19 city employees who oversee AI procurement across 7 U.S. cities. We found that cities' legacy procurement practices, which are shaped by decades-old laws and norms, establish infrastructure that determines which AI is purchased, and which actors hold decision-making power over procured AI. We characterize the emerging actions cities have taken to adapt their purchasing practices to address algorithmic harms. From employees' reflections on real-world AI procurements, we identify three key challenges that motivate but are not fully addressed by existing AI procurement reform initiatives. Based on these findings, we discuss implications and opportunities for the FAccT community to support cities in foreseeing and preventing AI harms throughout the public procurement processes.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; • **Social and professional topics** → **Computing / technology policy**.

Additional Key Words and Phrases: public sector AI, public procurement, AI governance

1 Introduction

Motivated by promises of increased efficiency, governments are increasingly adopting AI to automate bureaucratic workflows and assist decision-making processes that impact residents [27, 42, 62, 66, 70, 94, 116]. Often, such public sector AI applications are not developed in house, but are purchased from external vendors through a process called *public procurement* [72, 87, 107]. In a 2023 statement, U.S. Senator Gary Peters shared that over half of the AI tools used by federal agencies were purchased from commercial vendors [86]. Experts estimate that this number is even higher at lower levels of government, such as state and local governments which are less likely to have internal expertise to develop AI [101]. Thus, most public sector AI systems in use today are developed by and acquired from *private vendors*.

A growing number of academic and advocacy efforts have pointed out how procured AI systems have predominantly targeted narrowly defined notions of efficiency or performance, resulting in adverse effects that disproportionately impact marginalized communities [42, 53, 99, 109]. Facial recognition technologies, for example, have faced sustained criticism due to concerns about civil liberties, racial biases, and privacy violations [2, 54, 117]. Despite these concerns,

Authors' Contact Information: Nari Johnson, narij@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, PA, USA; Elise Silva, University of Pittsburgh Institute for Cyber Law, Policy, and Security, Pittsburgh, PA, USA, elise.silva@pitt.edu; Harrison Leon, hleon@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, PA, USA; Motahhare Eslami, meslami@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, PA, USA; Beth Schwanke, beth.schwanke@pitt.edu, University of Pittsburgh Institute for Cyber Law, Policy, and Security, Pittsburgh, PA, USA; Ravit Dotan, ravit@techbetter.ai, TechBetter, Pittsburgh, PA, USA; Hoda Heidari, hheidari@cmu.edu, Carnegie Mellon University, Pittsburgh, PA, USA.

in 2022, a U.S. government watchdog revealed over 20 federal contracts with private companies that either specialized in facial recognition or had awards related to its use, with a total budget exceeding $7 million [97].

While such incidents have exposed flaws in individual systems, they highlight deeper issues in how AI is acquired, used, and governed in the public sector. The AI procurement process encompasses decisions of which AI tools to ask for, adopt or reject, and the manner in which they are developed and deployed: decisions of critical importance for communities susceptible to AI harms. Scholars, civil society, and community advocates have pointed out how these procurement decisions not only influence AI performance and risks, but also play a significant role in shaping broader governance practices and ethical standards by which AI operates in the public sector [55, 79, 95, 101, 107].

The past few years have marked an explosion of attention directed towards imagining how AI procurement might be *reformed* to prepare governments to anticipate AI harms [8, 30, 55, 79, 96, 101]. However, there is a lack of empirical research to understand how AI procurement actually *occurs* in practice. Past scholarship has established that AI must be acquired using *legacy procurement practices* that apply widely to all goods and services, including pencils, school buses, and AI [79, 94, 101, 107]. Yet, little is known about what exactly these practices entail, how they vary across localities, and whether governments have made any changes to their practices when assessing AI specifically. Furthermore, important information about governments' AI procurement practices is not always publicly available [52, 96].

Our Contributions. In this paper, we conduct a formative empirical study to understand U.S. cities' current procurement practices for AI. Through semi-structured interviews with 19 local government employees in roles that involve technology governance or procurement, we pull back the curtain to show how cities' existing purchasing practices apply to AI. We draw from employees' learnings from past technology procurements to identify key challenges government employees face in anticipating and addressing harms caused by procured AI systems. We organize our findings and their implications for the FAccT community in two sections:

- In Section 4, we offer an empirical account of how AI procurement occurs on-the-ground. We illustrate how legacy purchasing practices that differ across cities shape how AI is governed by establishing important infrastructure (like solicitations and contracts), determining who within a city has power over different types of AI acquisitions, and influencing the socio-technical values by which AI is assessed. By detailing the organizational factors that influence how procurement occurs, we show how cities' purchasing practices both enable and preclude cities' ability to address AI harms: for instance, historic norms such as cost thresholds [38] allow employees to acquire low- or no-cost AI without going through the accountability measures typically associated with government purchasing. We also highlight emerging steps some cities have taken to incorporate AI-specific criteria and ethical considerations into their AI procurement processes.

- In Section 5, we outline three key challenges that still remain to be addressed by cities' existing (and emerging) practices. We argue that more work is urgently needed to (1) address information asymmetries between governments and vendors, (2) support cities in asking more of AI vendors, and (3) support cities in sharing and assuming ongoing responsibilities of AI governance. We discuss implications and opportunities for the FAccT community to support public sector workers in anticipating and mitigating harms caused by procured AI.

2 Related Work

2.1 AI in the Public Sector

Governments' increased adoption of data-driven technologies has transformed work in the public sector [13, 41, 66, 70, 75, 110]. While many technologies marketed as "artificial intelligence" today bear resemblance to historical data mining

technologies of the 1990s (*e.g.,* to detect fraud [41, 70]), recent years have marked a rapid increase in the number and popularity of AI tools that employ machine learning [41, 70, 79, 101]. Governments' use cases for AI, as taxonomized by [41], range from law enforcement and adjudication, such as supporting criminal sentencing [46, 89], to applications that support government service delivery, such as chatbots to facilitate communication with residents [69, 91]. Like private sector workers, government employees are increasingly adopting general-purpose productivity tools like ChatGPT [26, 35, 80] to do the "street-level" work of governance [13], *e.g.,* to adjudicate child welfare cases [84, 112], thus further blurring the lines between AI's articulated scope and ultimate application.

Scholars and activists have documented the harms resulting from public sector AI systems, with the recognition that historically oppressed communities are disproportionately subject to AI-assisted punishment [42, 71, 90, 109]. Public sector algorithms have faced criticism for algorithmic bias (*e.g.,* in the case of predictive policing [16, 57]), validity (*e.g.,* in the case of child welfare risk assessments [32, 48]), and functionality (*e.g.,* in the case of facial identification mismatches [56]). To this end, a growing body of empirical scholarship has studied government employees' practices, desires, and concerns surrounding governments' use of AI systems [23, 25, 64, 68, 102, 109, 114, 121]. Much research has focused on the perspectives of *frontline workers* (*e.g.,* social workers [64]) who consume AI outputs. Studies have characterized how AI systems' "misalignment" with policy objectives and workers' ethical values can result in harms to marginalized communities [64, 100, 109], highlighting implications for technology design. Other past work illuminates the organizational factors that facilitate critical decisions about AI adoption and governance [62, 64, 65, 89, 114], realizing that governments "should not be treated as monolithic entities" [62], and that workers at varying levels of institutional power often hold varying perspectives about the appropriateness and ethical implications of AI tools.

Despite this growth in research, governments' *procurement processes* remain empirically under-explored as a stage that can impact how AI is adopted and governed in the public sector. In the following section, we describe how scholars, civil society, and activists envision public procurement as both a site of opportunity and concern for AI governance.

2.2 Governing AI through Public Procurement

The term "public procurement" generally refers to the processes governments use to bring in goods and services that are developed externally [72, 87, 107], often involving paid transactions with third-party organizations.[1] In this work, our core focus is on studying public procurement practices in the United States, particularly for *local* (city) governments. Public administration scholars have characterized how procurement laws, organizational structures, and activities vary between levels of government and localities [29, 38, 72].

A growing coalition of policymakers, scholars, and civil society groups have called on governments to implement AI-specific procurement guidelines [1, 30, 55, 96, 101]. Governments must be prepared, scholars argue [36, 79, 101], to interrogate the value-laden trade-offs embedded by design decisions such as the vendor's choice of training dataset, optimization function, and predictive target. In response, several groups have explored how to embed these and other ethical considerations within cities' existing procurement processes [8, 96, 101, 103]. These efforts have produced practical guidance and readily adoptable resources for government employees, such as guidelines, tools, vendor repositories, and templates to guide AI procurement practices (*e.g.,* [3, 22, 30, 37, 88, 96, 101]). However, there is very little work that examines whether and how these tools help government employees in practice.

Many AI procurement reform initiatives draw from a long history of governments adapting their purchasing practices to enact social change [49, 76, 101], *e.g.,* by creating processes that prioritize minority-owned businesses

---

[1]We note that as described by past work [87], there is no single precise agreed-upon definition for what it meant by the term "public procurement" – rather, the definition is "muddled" and varies across contexts. See Appendix A for a more detailed discussion of definitions.

[76], enable public oversight over government surveillance [33, 119], or incentivize other ethical behaviors, such as sustainability [73, 113]. These initiatives call attention to the *purchasing power* governments hold "not as regulators, but as participants" in a free market [76]. Drawing from this tradition, scholars point to the "gatekeeping role" of public procurement in deciding which AI systems are (and are not) purchased [30, 55, 83, 107]. The ultimate hope expressed is that governments' interactions with vendors can "change markets" [3]: that governments can exercise their purchasing power to advocate for AI vendors to prioritize social values such as fairness, accountability, and transparency [103]. Given this potential role of procurement in AI governance, there have been a small number of research efforts that have uncovered challenges faced by AI procurement practitioners in Kenya and Uganda [83] and the US federal government [15], such as knowledge gaps in public sector workers' understanding of AI risks or an absence of "clearly defined standards" for AI governance. However, what remains less known is the current *practices* that governments follow when acquiring AI, *e.g.,* the assessment criteria governments apply, how purchasing decisions are made, and who holds the power to make them.

This paper contributes to these ongoing discussions by illuminating the organizational realities in which AI procurement occurs. We provide an empirical account of the emerging actions that U.S. local governments have taken to revise their AI procurement processes on the ground. Given this context, in Section 5 we revisit several critical challenges surfaced by past scholarship, such as knowledge gaps in government employees' preparedness to anticipate AI harms [41, 83, 120], governments' lack of leverage with AI vendors [55, 62], and uncertainty about how to share and delegate responsibilities of AI governance [11, 50]. We ground our analysis using examples of real-world procurements to reflect on both the progress cities have made and obstacles that remain to be overcome.

3   Methodology

To understand the organizational processes that shape how cities govern AI, we focus on the perspectives of municipal employees. We conducted semi-structured interviews over six months from December 2023 to June 2024. Our research coincided with a landmark year of government action on AI procurement, including the release and adoption of the U.S. "AI M-Memo" [4, 106], which instituted governance requirements for federal agencies' use of AI[2], and the formation of the Government AI Coalition [9, 40, 98], a grassroots coalition founded to "give local governments a voice in shaping the future of AI". As a result, our study took place at a moment when governments across the U.S. were just beginning to stand up and implement new governance processes for procured AI.

Recruitment. Our recruitment process aimed to capture a diverse set of employees' perspectives on the AI procurement process, across multiple cities. Many cities required employees to complete a formal approval process to participate in research, so we organized our recruitment efforts by city. Once we received approval, we used snowball sampling to ask our initial contacts within a city to introduce us to other employees whose present role was involved with technology procurement or governance. As shown in Table 1, most participants worked in technology-focused roles in their city's IT or Innovation departments. We also spoke with specialists in vendor relations and procurement, and one HR specialist who had conducted organizational training on AI. Participants included both leaders who made decisions on behalf of their department, and workers whose day-to-day responsibilities involved managing AI procurement.

We began the study by reaching out to contacts in our professional networks in four U.S. cities, who introduced us to contacts in three additional cities. From the contacts we were given, we intentionally invited cities that represented a wide range of regions and maturity surrounding AI (*e.g.,* whether or not they had adopted any public-facing AI policies).

---

[2]The Biden AI M-Memo has since been revoked by the new administration [81].

ID	Pseudonym	Department	City Size	Title
P1	Charles	Innovation	Medium	Chief Data Officer
P2	Mari	Information Technology	Large	Senior IT Manager
P3	Jennifer	Information Technology	Large	Privacy Program Manager
P4	Emma	Information Technology	Small	IT Business Relationship Manager
P5	Kai	Information Technology	Large	Privacy Specialist
P6	Michael	Information Technology	Small	Director of IT
P7	Eric	Information Technology	Large	Chief Privacy Officer
P8	Liz	Management & Budget	Medium	Sourcing Specialist
P9	Paul	Innovation	Medium	Innovation Specialist
P10	Hana	Information Technology	Large	IT Analyst
P11	Andrea	Information Technology	Medium	Chief Technology Officer
P12	Kathryn	Human Resources	Small	Talent & Culture Program Manager
P13	Rebecca	Information Technology	Medium	Chief Data & Analytics Officer
P14	Olivia	Management & Budget	Large	Director of Procurement
P15	Ryan	Information Technology	Large	Chief Technology Officer
P16	Adrian	Innovation	Large	IT Policy Director
P17	Steven	Information Technology	Large	Vendor Manager
P18	Hugo	Innovation	Large	Chief Information Officer
P19	Liam	Innovation	Medium	Senior IT Manager

Table 1. An anonymous description of participating municipal employees. Titles were modified to preserve anonymity. Small cities have under 200,000 residents, medium cities have 200,000 - 500,000 residents, and large cities have over 500,000 residents.

Although this approach has limitations, we learned that establishing trust through shared connections was important for government employees to feel comfortable speaking openly with academic researchers. For example, one participant shared that their department was hesitant to participate in research studies led by someone they were unfamiliar with. We invited eight total cities to participate, and seven agreed. Participating cities represented all four major regions (Northeast, West, Midwest, and South) defined by the U.S. Census Bureau [7].

Semi-structured interviews. Following Veale et al. [114], we adopted a semi-structured interview approach to allow for flexibility in discussions. This allowed participants to spend more time discussing the phases of the procurement process that were closest to their role and expertise. The interviews ranged from 60 to 90 minutes and our protocol included three sections. First, we asked participants about their current role and work responsibilities relating to AI. We defined "AI" for participants using the OECD definition of "any machine-based system that can make predictions, recommendations, or decisions" [44], and provided examples of qualifying systems. Second, we asked participants to walk through how an example procurement involving AI would occur in their city, paying particular attention to any differences between a standard technology procurement. Third, we asked participants to reflect more deeply on their perceived challenges, needs, and desires to improve the AI procurement process. We expressed to participants that we were interested in understanding potential risks or societal impacts of the AI systems they procured, but did not steer participants towards discussing particular types of impacts, as we aimed to understand how participants conceptualized risk and harm in their own words. The study was approved by a university IRB. We include our complete interview protocol in Appendix D and discuss ethics and participant safety considerations in Section 7.1.

Data analysis. We collected 23 hours of interview audio which were transcribed and coded by four team members. We adopted a bottom-up thematic analysis approach [21] to analyze interview transcripts. Each transcript was first open-coded by two authors, who met to discuss each transcript and resolve any differences in interpretation [77]. In total, we created 305 unique codes, such as "*proprietariness*" or "*indemnification from AI harms*". We then grouped these codes into themes corresponding to cities' purchasing practices, such as "*alternative procurement pathways*" and "*AI*

risk assessments", and challenges faced by employees such as "*lack of leverage with AI vendors*". We also complemented our interview findings by conducting a document analysis [61] of cities' written AI policies or procurement materials, such as solicitation or contracting templates that were referenced in interviews. When selecting our final set of themes, we aimed to emphasize findings that contributed new insights to the existing literature on AI and procurement.

Limitations. We acknowledge several methodological limitations of our study. Our reliance on snowball sampling skewed our sample of participating cities towards large cities, whose practices and needs likely do not represent the majority of local governments in the U.S. While our focus on U.S. cities allowed us to draw productive comparisons across jurisdictions, we believe that understanding how procurement differs across countries and levels of government is a critical direction for future work. Finally, our decision to focus on city employees neglects the perspectives of other important stakeholders such as AI vendors, impacted communities, and other members of the public. We hope that our research can lay empirical foundations that can inform this future work.

4 Legacy Purchasing Practices Shape How Cities Govern AI

In what follows, we shine a light on the processes that municipal employees currently use to procure AI, on-the-ground. We learned that all seven interviewed cities had already started to use procured AI technologies designed for a wide set of users and goals, for example, to facilitate resident communication, increase workplace productivity, aid law enforcement, and assist bureaucratic decision-making[3]. Notably, five out of seven cities we interviewed *did not* have the capacity to develop their own AI solutions internally, affirming our hypothesis that local governments rely heavily (and in many cases, entirely) on third-party AI systems.

Throughout interviews, employees repeatedly emphasized that AI must go through all existing legacy procurement processes, which are shaped by a complex web of hard policy (*e.g.,* laws and regulations) and soft policy (*e.g.,* guidelines, practices, organizational culture). We illustrate how these legacy practices shape how cities govern AI in three ways: (1) by structuring critical decision-making, (2) determining who within a city has power to shape procurement outcomes, and (3) shaping the socio-technical values by which AI is assessed. We conclude by summarizing the steps that some interviewed cities have taken to adapt their legacy purchasing practices for AI.

4.1 Public procurement structures critical decisions about AI

When asked what procurement entailed, employees often described the highly structured processes their city took to award and oversee contracts with vendors. Consistent with public administration scholarship [34, 45, 51], employees tended to follow a six-step process (Figure 1) for acquisitions that involved a competitive solicitation. However, not all steps occurred in every city. While some cities required employees to introduce legislation in front of City Council to approve every procurement contract, presenting an opportunity for public oversight [94], other cities did not need to introduce legislation for a purchase to be made.[4] For each step, cities applied highly standardized processes when purchasing AI. Standardization, employees explained, was important to "increase efficiency" of the procurement process and prevent government corruption [10, 85]. Analyzing cities' standardized processes reveals the implicit assumptions they carry, and how existing processes both facilitate and constrain cities' capacity to address AI harms.

Throughout interviews, we learned how cities' solicitation processes determined if, and how, the societal impacts of AI factored into vendor selection. One common type of solicitation is a Request for Proposal (RFP), a structured process where a government outlines their needs, expectations, and desired outcomes. Interested vendors then submit

---

[3]In Appendix B, we describe each of these use cases in further detail and provide example real-world AI solutions mentioned by participants.
[4]See Duchicela et al. [38] for a more detailed review of how City Council oversight processes vary across local governments.
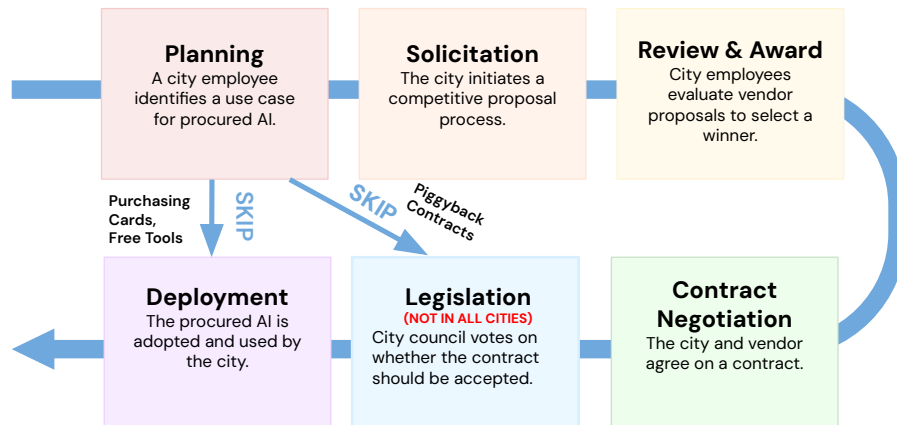
Fig. 1. Six common steps that occurred in cities' public procurement processes. While purchases that went through a full solicitation (*e.g.,* an RFP) proceeded linearly through these steps, as shown by the long blue arrow, AI acquired via alternative purchasing pathways such as cooperative purchasing agreements ("piggyback contracts") or under cost thresholds (*e.g.,* using purchasing cards) skipped past several steps, illustrated by the short blue arrows. In our paper, we discuss meaningful differences across cities' processes. For example, not all cities had a legislation step, and some cities still required all acquisitions to go through contract negotiation. As such, this figure serves as an illustrative example of how AI procurement occurs in many (but not all) U.S. cities.

detailed proposals that comprehensively address these requirements. After the solicitation has closed, cities review vendors' proposals to select a winner. Liz, a procurement specialist, walked through her city's RFP evaluation process by sharing her screen as she navigated her city's procurement management software for a past IT procurement: "So all the evaluators will show up here", Liz said, pointing to a list of ten city employees. "And when they go in, they score each vendor based on multiple criteria" such as "*Qualifications of the project team*", "*Project plan*", and others, assigning a score of 1 to 15. Once evaluators individually had assigned scores for every single proposal, they met as a committee to discuss the scores and select their "top three candidates to reach out and set up a live demonstration".

After a vendor is selected, the city and the vendor create a contract that specifies legally enforceable obligations for both parties. At minimum, contracts specify an agreed price for the city to pay over a specified time period, raising concerns of vendor "lock-in" [19] where cities are hesitant to break existing contracts. Steven, a procurement specialist, discussed his experiences negotiating "indemnification" clauses with AI vendors, which spelled out how parties would compensate each other for damages, *e.g.,* "Will the vendor cover [the city] if we get sued for bias?" Some contracts exactly spelled out the support AI vendors would provide cities: in one egregious example, a city that was struggling to reach a vendor learned that according to the contract, they were only allowed up to five phone calls of support.

Rather than write each contract from scratch, often both the city and the vendor relied on "standard agreement" templates written by lawyers. Vendor's standard agreement templates written for government customers included detailed *terms-of-service* [43] for the technology and the support services the company would provide. All interviewed cities had created their own standard agreement templates for IT procurements to "protect the interests of the city" by laying out clear expectations of vendors, *e.g.,* that the vendor must comply with minimum data security standards, or provide the city ownership over any data collected from the engagement.

Uncovering cities' standardized processes reveals how critical decisions about AI are currently structured, and how efforts to mitigate algorithmic harms might be operationalized within existing processes, *e.g.,* via solicitation questions, scorecards, and contract terms. However, a critical examination of these practices surfaces how they might limit cities'

ability to effectively address AI harms. For example, FAccT scholars have documented the limitations of relying on highly structured methods (*e.g.,* compliance "checklists" [74, 118] that are highly similar to RFP requirements) to assess the societal implications of AI. Similarly, FAccT research has highlighted the difficulty of anticipating algorithmic harms before deployment [20, 78, 104, 105], arguing that the socially situated nature of harms arising from lived experiences mandates their assessment "from the ground up" [78]. However, cities were required to negotiate and sign contracts with AI vendors before they could access (and deploy) AI systems behind vendor paywalls. Contracts were often not renegotiated, and indeed, several interviewed employees [P2, P4, P17] regretted how they negotiated a particular contract only after algorithmic harms surfaced post-deployment. Thus, procurement contracts codified both cities' continued financial obligations to vendors, and vendors' (lack of) obligations to redress algorithmic harms.

4.2 Public procurement determines who has visibility, oversight, and decision-making power over AI

Cities' legacy procurement practices often specified accompanying roles, responsibilities, and lines of communication for employees. Many cities conceived of procurement as a highly collaborative activity involving not only the purchasing customer, but also specialists from a wide variety of departments: procurement specialists to assist with solicitations, legal specialists to help with contracting, and IT specialists to provide expert review. Different employees who participated in an AI procurement had different vantage points and objectives depending on their roles. However, not all of these specialists were involved, or even aware of, every acquisition. Differences in cities' established purchasing processes shaped *who* within the city had visibility, oversight, and decision-making power over procured AI.

While all cities had designated IT staff whose responsibilities included providing support on technology procurements, cities' procurement processes differed in how they allocated decision-making power. Some cities established IT reviews as mandatory approval gateways, but other cities positioned IT reviews as optional consulting. For example, because Adrian's city historically granted individual departments (*e.g.,* the police, public schools) "a lot of autonomy in how they operate", city employees in these departments were not required to consult IT before making their own technology purchases. In contrast, Jennifer's city mandated that all city employees, regardless of department, must "open a ticket" to initiate an IT review before buying new software.

Another key factor that determined the actors involved in an AI procurement was the purchasing pathway used to make the acquisition. In Figure 1, we describe the six steps that occur for procurements that go through a full solicitation. However, our interviews indicate that AI procurement often does not take this classic route, sometimes skipping competitive solicitation or contract negotiation entirely. Instead, employees leveraged a variety of *alternative procurement pathways* to purchase AI. In many cities, acquisitions that occurred through these alternative procurement pathways "didn't have to go through a procurement [department]", and thus fell outside the scope of existing accountability measures for government purchasing. Due to broader shifts in the AI landscape, namely the availability of low- and no-cost AI tools, many AI acquisitions did not involve a competitive solicitation because they were under specified *cost thresholds* [38] that would require going to solicitation. Local policies specified that municipal employees could make purchases under a certain dollar amount (which varied city-to-city) at their own discretion, using a government-issued *purchasing card.* Cost threshold policies, procurement director Olivia explained, were established to "triage risk, by defining the [riskiest projects] based on dollar value". Types of AI tools that fell under cost thresholds included free online services like ChatGPT, services with paid subscription models, or AI donated through academic collaborations, foundations, or from for-profit companies[5].

---

[5]For instance, Palantir, a for-profit company, donated its predictive policing technology as a "philanthropic gift" to be used by the City of New Orleans, free of charge [108].

The governance processes applicable to procurements under cost thresholds varied significantly from city to city. Adrian, an IT leader, reflected on how technology acquisitions under cost thresholds were particularly difficult for their department to even be aware of until after they were already purchased: "[For purchases] below $x0,000, there's few oversight of regulatory mechanisms to control, or even have visibility of what departments do," they explained. "It's not routed through a centralized control mechanism." As a result, individual employees could purchase and deploy free and low-cost AI tools (*e.g.,* with prices under $x00 per month) at their own discretion, without notifying others in the city. In contrast, Mari's city still routed "anything involving a company or vendor" through a centralized procurement review: "Even if it's an in kind donation, it is still tied to a scope-of-work and a contract."

Cities' varying governance processes for different types of AI procurements had significant impacts on individual employees' ability to advocate for their desired outcomes: for instance, refusing to proceed with procuring an AI system due to concerns about its potential harms. In some cities, employees who were tasked with doing AI governance work often lacked basic visibility into the AI their colleagues were using due to these legacy purchasing norms.

4.3 Public procurement shapes the socio-technical values of AI assessment

In conversations, we found significant differences in the socio-technical values that shaped cities' approaches to assessing AI systems. These values, we found, were rooted in cities' existing public procurement norms and established technology review processes.

Public procurement norms, particularly cost-effectiveness, played an important role in how employees assessed AI proposals. "When we make decisions, we think about how we're stewards of residents' tax money," explained Emma, an IT procurement specialist. Indeed, cost was central factor that influenced employees' AI purchasing decisions. Employees shared many anecdotes where they ultimately decide not to procure an AI system because it was too expensive [P4, P6, P8, P14]. While some employees like Hugo expressed enthusiasm about the increased availability of low-cost consumer AI tools (such as a generative AI pilot that only cost Hugo's city $5), others expressed concern over the "extractive capabilities" and hidden costs of free AI tools. Adrian recalled an incident where an employee used a transcription AI tool that did not have "a consensual model for data collection". Although the transcription tool was free, its use could result in "divulging resident information" or "more secure" government business, data that the vendor could use to improve its models. Adrian's story highlights a tension between the long-standing norm of cost-effectiveness and the new reality of unique risks posed by low- and no-cost AI.

Beyond cost, employees also applied criteria from their established IT review processes when purchasing AI. Understanding these IT criteria is important because in many cities, IT review processes were the only applicable governance processes for procured AI. IT personnel were responsible for assessing review criteria that varied across cities, often operationalized as "technical" or "functional" requirements in RFPs. Common review criteria included the software's security (*e.g.,* to data breach attacks), usability, ease of set-up, and compatibility with the city's existing technology stack. In conversations, we found that these common criteria for software procurements were insufficient to anticipate AI harms. For instance, basic considerations of system functionality [93], validity [32], or train-test distribution shift [92] were not typically asked during a traditional software procurement, where systems are often assumed to be "programmed correctly" to achieve specified goals.

One significant difference between cities' IT review processes was the degree to which they prioritized values such as data privacy and protection from government surveillance. Although these considerations were never mentioned in most employees' descriptions of their review criteria, participants in a small number of cities called them out as organizational priorities. Understanding these priorities, we learned, was critical to understanding how cities procured

AI. For example, cities with established privacy programs often tasked privacy personnel to work on AI-focused initiatives, and adapted existing privacy governance infrastructure to include AI-specific considerations. Similarly, cities with surveillance ordinances had already established governance processes such as community oversight mechanisms for qualifying AI technologies [119]. Organizational values also shaped the way employees conceptualized AI's benefits and risks: Kai shared that it was the "privacy lens" adopted by her organization that led her to ask vendors questions about training dataset curation and consent. In contrast, employees in cities without surveillance or privacy policies spoke considerably more positively about AI-powered "smart city" [116] or policing technologies. Thus, cities' (lack of) existing privacy or surveillance policies played a critical role in determining what risks were considered, what mitigations were requested, and what technologies were ultimately purchased and deployed.

4.4 Emerging practices: New actions cities have taken to govern AI

Three interviewed cities had already implemented specific changes to their existing procurement practices for AI, and several other city leaders shared their future plans to change their practices. City employees felt that it was "early days" in revising their AI governance processes: for example, employees were in the midst of overseeing their first AI solicitation, conducting their first AI risk assessments, and revising their AI governance processes more broadly. With this rapidly evolving landscape in mind, we highlight important changes that cities have made so far.

Intervening within existing procurement processes. The most common action considered by cities was to modify their existing contract and solicitation templates to include additional items for AI-related procurements. Cities instituted additional reporting requirements for AI systems in their technology RFPs, adopting language from the GovAI Coalition's AI FactSheet [5] of questions for vendors. The factsheet asks vendors to report "essential technical details" such as on what data the AI was trained, under what conditions the system was tested, the values of relevant performance metrics, and steps that vendors have taken to promote values of fairness, robustness, and explainability. Cities also adapted their contracting practices by developing additional contract language specific to AI procurements. Many employees adopted language from GovAI's Vendor Agreement [6], a "plug-and-play legal addendum" of requirements for vendors, *e.g.,* to develop an AI incident response plan, remediate AI incidents, and provide governments with a means to monitor and audit AI performance. Most interviewed cities did not plan to adopt GovAI's resource templates as-is, but instead chose to "selectively adopt" the items that felt most important to their city.

Cities' legacy purchasing practices shaped how applicable these procurement interventions were to different AI systems. For instance, AI acquired under cost thresholds did not go through a competitive solicitation or contract negotiation, leading the cities to accept vendors' terms-of-use as is. Similarly, AI acquired by "piggybacking" by adopting an existing contract between the vendor and another government [67] often did not involve re-negotiation to include additional terms. As a result, these interventions did not apply to the many AI solutions acquired through these other pathways. In recognition of the limitations of relying exclusively on existing procurement infrastructure for solicitations and contracting, several employees decided to organize their AI governance efforts around a separate "AI review".

Establishing an AI review process. Three interviewed cities established an internal AI review process overseen by IT employees, that occurred outside of the formal review processes of procurement (*e.g.,* scorecards). When a city employee wanted to acquire an AI solution, they initiated an AI review by contacting these IT specialists directly, *e.g.,* by filling out a form or "opening a ticket". IT employees or committees trained to assess AI systems then reviewed the employees' request. AI reviews often involved conducting *risk assessments* to help city employees understand potential positive and negative impacts of the system. Employees used pre-acquisition risk assessments to shape subsequent conversations with AI vendors, for instance, by requesting specific risk mitigation steps to be implemented in the AI

system's design. Some cities used risk assessments to institute additional governance steps for "high-risk" systems, such as instituting internal usage protocols, public reporting requirements, and regular post-deployment monitoring. We discuss governments' emerging AI risk assessment practices in detail in Appendix C.

Separating AI reviews from cities' formal procurement processes had several benefits, such as improving process efficiency so that employees could contact expert AI reviewers directly. Cities could require AI reviews for acquisitions that would not typically be reviewed by their city's procurement departments, such as free or low-cost AI tools. However, employees' efforts to operationalize and socialize new AI reviews in practice often relied heavily on existing procurement infrastructure. For instance, Hana's city implemented their AI review within their city's online procurement system that had existed for years. This centralized procurement system, Hana explained, "allows us to be pretty confident that if there is some kind of AI system, we're going to catch it, because it's going through this procurement process". But this level of centralization was uncommon across other interviewed cities, where IT employees only became involved once a formal solicitation was initiated. Thus, the feasibility and implementation of cities' AI reviews was still heavily influenced by their legacy procurement practices.

5 Three Key Challenges at the Heart of AI Procurement Reform

Although some employees pointed us to success stories where their purchasing processes enabled them to walk away from red flags or negotiate for mitigations, we found that cities' present AI procurement practices often fell short of their ambition to protect the public from risks and harms of AI. In what follows, we reflect on three challenges that often motivate, but are not adequately addressed by existing AI procurement reform efforts. While considerable progress has been made, we believe that more work is urgently needed to (1) address information asymmetries between governments and vendors, (2) support cities in asking more of AI vendors, and (3) support cities in sharing and assuming ongoing responsibilities of AI governance. Under each challenge, we identify opportunities for how the FAccT community might contribute towards addressing local governments' needs for support.

5.1 Information asymmetry and (un)preparedness for AI

While some employees expressed confidence in their organization's preparedness to govern procured AI, many employees, particularly in smaller cities, felt behind. Employees of cities that had not yet introduced AI standards into their procurement processes were often uncertain of "what questions to ask" AI vendors. One procurement specialist reflected on how her unfamiliarity with AI made her feel: "I don't really know what the risks are to working with AI. If I can't protect us from those risks comfortably, then I'm not doing my job."

This information asymmetry, defined by [17] as a phenomena "where one business party possesses more information than the other party they are dealing with", resulted in procurements where important questions of system functionality, training data provenance, or other societal impacts were never raised. Sales representatives, who held significant sway in educating governments about AI use cases through pitch-like demos, shaped cities' narratives about procured AI technologies' potential benefits and harms. We learned from several anecdotes that AI vendors did *not* report basic information about AI systems unless asked. For instance, one participant shared that their team did not consider risks posed by hallucinations (providing incorrect information) in a recent procurement of an AI chatbot service, and that risks due to inaccuracy were "not part of the conversation" they had with the vendor.

Employees shared a variety of methods they used to increase their organization's AI literacy and preparedness, such as recruiting employees with AI expertise, resource sharing efforts (*e.g.,* through GovAI), and participating in intergovernmental AI task forces [39] or grassroots peer networks. Some employees began collaborations with

academic researchers, for example, to evaluate a pilot of a generative AI productivity tool. Finally, some cities hired paid consultants (*e.g.,* from Gartner) to provide strategic advising (*e.g.,* conduct market research) on AI. Employees in smaller cities were less likely to be aware of existing resources or peer networks to learn about AI, and felt overwhelmed by the broader information ecosystem of "hype cycles" surrounding emerging generative AI technologies.

While existing resources and standardization efforts helped cities identify questions to ask of vendors, employees repeatedly struggled to understand what to do with the information that vendors provided in response. Jennifer, a privacy specialist, shared how her department was uncertain about how to interpret the performance metrics reported by vendors: "We ask some sort of question: what is your R-squared value? How do we know if [what is reported] is good? Someone needs to be able to have the technical acumen to say what is acceptable here in terms of accuracy, error rates, thresholds." While employees like Jennifer, Eric, and Paul expressed a similar desire for clear guidance and thresholds, existing AI procurement reform initiatives intentionally left these decisions to the discretion of individual cities, who could determine their own organizational risk tolerance. However, even cities that knew what to ask still struggled to individually assess and make decisions based on information reported by vendors.

Recommendations for future work. Although much progress has been made in developing resources to support governments' AI procurement practices, our findings reveal that there is still an urgent need for government education and support. Misconceptions surrounding AI capabilities, often perpetuated by AI vendors, resulted in power imbalances where cities' existing review criteria for software proved insufficient to anticipate AI harms. We believe that the FAccT community has a critical role to play in doing this translation work. Our conversations revealed a number of avenues through which members of the public can influence AI policymaking, such as by collaborating on research [110], joining peer networks such as the GovAI Coalition [8], or even simply reaching out to one's own government to understand their AI governance practices and needs for support. Following interviews, our research team curated and shared educational resources that we hoped would address employees' open questions about AI. These conversations, we learned, were often a precipitating event that led interviewed cities to join AI-focused peer networks, revise their procurement practices, and even establish new ethical AI governance policies. The FAccT conference can explore models to build community and develop trust-based relationships across scholars, community advocates, and government employees, *e.g.,* following models such as the Public Tech Leadership Collaborative [12].

Future research should also explore tools and processes to help governments assess information provided about AI systems, to make decisions about *what to purchase*. Our conversations surfaced several obstacles to measuring pre-acquisition performance. A lack of standardized benchmark datasets for public sector use cases meant that vendors reported metrics using their own proprietary datasets, making comparison across vendors difficult. More broadly, participants expressed a desire for vendors' AI systems to be vetted *for them*, for instance, through a third-party certification program [28, 31]. Future efforts can aim to provide domain-specific guidance on what measures should be reported, develop standardized benchmarks, and explore the possibility of third-party certification, with a critical recognition of where the narrow framing of performance measurement might fail to capture the ethical implications of AI adoption [47, 59, 78]. We join past calls [55, 62, 78, 119] for more empirical research that examines how members of the public (*e.g.,* impacted communities and their advocates) can have a "meaningful opportunity to respond and, if necessary, dispute the use of a particular AI system" before acquisition occurs [94].

5.2   Cities lack leverage in relationships with AI vendors

Several employees from cities with new AI governance measures described positive experiences with "good vendors" who were transparent and responsive. For instance, an AI vendor's openness to disclose system limitations in their

autonomous drone technology helped a city refine its usage policies to avoid its deployment in low-accuracy environments. Despite these successes, we heard many more stories of uncooperative AI vendors that failed to provide basic information about their AI system, amend their contracting terms, or implement mitigations requested by the city. Interviewed employees repeatedly felt that they lacked leverage in advocating on behalf of their city.

AI vendors frequently withheld critical information about their AI systems from governments, hindering employees' ability to assess risks. Vendors often invoked intellectual property claims to deny basic requests, such as responding to AI RFP items or completing AI FactSheets. Consistent with past research [22, 62, 79, 107], participants were often denied access to vendors' AI models (*e.g.,* model weights) and training datasets. Additional denied requests included information about the presence of copyrighted content in training data [P3, P5], whether city data would be used to train vendor's models [P1, P17], and disaggregated accuracy measures, *e.g.,* across demographic groups [P2, P7, P10].

While bad vendor behavior could serve as an informative signal when making purchasing decisions, relational challenges still arose for procurements that were already under-contract. Past procurements were important to consider given the prevalence of vendors' "scope creep" [111] that introduced AI under existing contracts. Also, many cities' AI review processes were so new that they did not apply for most AI they already owned. Employees shared anecdotes where AI vendors did not provide adequate personnel training [P4, P17], refused to give cities access to data collected on city employees or residents [P2, P14, P15], and refused to let cities opt out of new AI features [P1, P17] or their data being used for AI training [P5, P13, P16]. In some cases, cities responded by "shutting off" purchased services, such as a city that stopped paying for Zoom because they were "uncomfortable with their AI notetaker". Yet, several employees discussed past procurements where they struggled to break ties with vendors. For instance, employees shared stories where they struggled to disable AI features that were deeply embedded in necessary technical infrastructure, such as new AI productivity features introduced into enterprise software [P1, P3, P5, P6, P11, P15].

More broadly, we observed that employees' optimism about their city's ability to hold vendors to account was shaped by the decision-making leverage they held within their own city. When vendors failed to meet employees' expectations or presented them with take-it-or-leave-it scenarios, employees who felt they had the authority to influence purchasing decisions felt comfortable walking away. For instance, Rebecca, an IT director, worked in a city that had passed a policy requiring AI reviews as a necessary approval workflow. When an AI vendor told Rebecca that basic information about their model's performance was "proprietary" (an "unacceptable answer" according to city policy), Rebecca made the decision on behalf of the city to not make the purchase. But in other cities, employees who oversaw cities' AI review processes lacked the decision-making power to influence final purchasing decisions. As a result, many AI procurements where vendors failed to meet cities' standards still ultimately moved forward.

Recommendations for future work. We found that many AI vendors refused to voluntarily comply with cities' requests to provide basic information or implement simple harm mitigation steps. This context of low compliance is critical to consider in this moment of unprecedented attention targeting governments' AI procurement practices. Several interviewed employees expressed their frustration in having to advocate for basic ethical behaviors (*e.g.,* transparency about how the vendors would use resident data) that they felt should be minimum requirements to even be able to participate in the market (*e.g.,* enforced as consumer protection rights or federal data privacy legislation). These stories call into question the hope implicit to initiatives that focus primarily on governments' power *as purchasers*: that when it comes to protecting the public interest, the market of public sector AI vendors will simply regulate itself.

When vendors failed to meet expectations, a variety of factors such as contract agreements or conflicting incentives across government employees meant that procurements still moved forward. Future research efforts can more deeply interrogate the role that legacy purchasing practices like contracts play in constraining cities' negotiation power, and

ways that cities might adopt their processes to address these issues head-on (*e.g.,* by explicitly giving AI reviewers the power to reject a proposed AI system due to its potential to cause harm).

5.3 Sharing (and assuming) ongoing responsibilities of AI governance

Procurement agreements create a unique type of accountability structure: local governments established to protect public values must rely on third-party vendors to develop AI solutions. But the work of AI governance does not end at development, or even deployment. Review and oversight of procured AI technologies requires labor and expertise to conduct evaluations, respond to incidents, train users, and enforce compliance. In this section, we share cities' experiences working with AI vendors to evaluate system performance to explore cities' varying perspectives on delegating AI governance work.

In our conversations, we found that vendors often did not provide performance evaluations after deployment. In response, some participants [P7, P10, P15, P18] evaluated models "on their own". For example, Hana's department sought to monitor the effectiveness of a procured gunshot detection AI system to ensure it "was continuing to be effective and meet the needs of the city". However, the vendor neither shared performance evaluations nor offered guidance on how to assess the system. As a result, Hana and her colleagues had to "develop their own metrics" and monitor performance themselves. Some employees embraced the opportunity to conduct independent evaluations. Hugo, for example, valued the city's ability to shape evaluation criteria both as a safeguard against vendor's selective reporting and as an opportunity to strengthen local AI expertise. However, not all employees conducted their own evaluations of procured AI systems, and many participants [P2, P3, P4, P5, P8, P13, P17] believed that conducting performance monitoring and audits should be the responsibility of the vendor.

Considering performance evaluation as a case study reveals several insights for understanding how cities and vendors can share responsibilities of AI governance at large. Employees pointed out how conflicting incentives (*e.g.,* companies' unwillingness to report unfavorable results) might actually dis-incentivize vendors from fully participating in activities of governance – in this case, evaluating system efficacy. While employees who took ownership of doing evaluation work often found it to be an rewarding opportunity to promote the best interests of their city, other employees expressed hesitancy in their ability and capacity to do AI governance work.

Recommendations for future work. To date, there is a lack of guidance on how the ongoing responsibilities of AI governance should be shared between cities versus vendors. Although several city employees expressed concerns about their capacity, we believe that most AI governance activities are *best performed* by government employees, who are better positioned to represent the needs of city employees and residents throughout this process (*e.g.,* in defining what is most important to measure [58]). We join past work in arguing that internal capacity-building is also critical for establishing robust systems of *internal* accountability and ownership over procured AI [41, 94]. Thus, future research can expand on existing tools and processes to support governments in various aspects of AI governance, *e.g.,* training users of AI systems [63] and evaluating system performance [68]. To be effective, interventions should be developed with both the unique context and constraints of procurement in mind, such as employees' non-technical expertise and capacity, as well as constraints employees face in accessing vendors' "proprietary" AI models [24].

6 Conclusion

In this paper, we explore how cities' varying legacy procurement practices shape how public sector AI technologies are selected, assessed, and governed on the ground. We show how cities' ability to implement AI governance initiatives that prioritize ethical considerations is heavily shaped by the existing procurement infrastructure and power structures that

facilitate (or prohibit) cross-organizational oversight over different types of acquisitions. Understanding these present organizational realities of how AI procurement unfolds is a critical first step towards re-imagining interventions that both acknowledge and address governments' longstanding purchasing practices head-on.

While government employees are still in early stages of adapting their AI procurement practices, our conversations reveal a deeply concerning trend: many AI vendors are not cooperating with employees' efforts to understand and mitigate AI harms. The implications of vendors' non-compliance for the rights and freedoms of vulnerable communities cannot be understated as local, state, and federal governments increasingly outsource the critical work of governance to third-party AI systems [18, 55, 71, 79]. Research to support cities in navigating these power differentials, *e.g.,* by conducting their own evaluations or renegotiating existing contracts, is urgently needed. Finally, we encourage the FAccT community to reflect on what intervening within the structures of public procurement affords, and its limitations as a remedy for algorithmic accountability. Our study prompts us to ask: what types of reforms to cities' legacy purchasing processes might the FAccT community advocate for? What happens when cities' negotiations with AI vendors fail to be successful? When might we look towards other approaches, such as direct regulation of AI vendors [60, 115], to protect the public interest?

7   Endmatter Statements

7.1   Ethical Considerations

To preserve anonymity of participating employees and cities, we assured interviewees that their participation was voluntary, they could decline to answer interviewer questions, and their responses would be kept anonymous. For sensitive or potentially identifying interview quotes, we exclude participant IDs to preserve anonymity. When appropriate, we use the "x" character to omit exact dollar amounts to preserve confidentiality. To mitigate the risk that participating cities are identified, we limit the amount of detail we provide about cities' practices.

References

[1]   2020. *Guidelines for AI procurement.*   Retrieved September 12, 2024 from https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement

[2]   Logic(s) Magazine 2020. *Safe or Just Surveilled?: Tawana Petty on the Fight Against Facial Recognition Surveillance.* Logic(s) Magazine. Retrieved January 16, 2024 from https://logicmag.io/security/safe-or-just-surveilled-tawana-petty-on-facial-recognition/ Interview.

[3]   2020. *TAI Procurement in a Box: Workbook.* Retrieved January 16, 2024 from https://www.weforum.org/publications/ai-procurement-in-a-box/

[4]   2024.   *Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence.*   Retrieved September 10, 2024 from https://web.archive.org/web/20250119032834/https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

[5]   2024.   *AI FactSheet.*   Retrieved January 19, 2025 from https://web.archive.org/web/20241211185508/https://www.sanjoseca.gov/home/showpublisheddocument/109730/638672090591070000

[6]   2024. *AI Vendor Agreement.* Retrieved January 19, 2025 from https://web.archive.org/web/20241213170059/https://www.sanjoseca.gov/home/showpublisheddocument/109754/638458754641770000

[7]   2024. *Census Regions and Divisions of the United States.* Retrieved September 10, 2024 from https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

[8]   2024. *GovAI Coalition's open letter to the public, policymakers, and industry.* Retrieved September 12, 2024 from https://www.sanjoseca.gov/home/showpublisheddocument/109774/638459250725100000

[9]   2024. *The Government AI Coalition.* Retrieved September 12, 2024 from https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition

[10] 2024. *The nature of corruption in public procurement.* Retrieved September 10, 2024 from https://www.u4.no/topics/procurement/basics

[11] 2024.   *Request for Information: Responsible Procurement of Artificial Intelligence in Government.*   Retrieved September 10, 2024 from https://www.regulations.gov/document/OMB-2024-0004-0001

[12] 2025. *Public Technology Leadership Collaborative.* Retrieved January 19, 2025 from https://web.archive.org/web/20250108220513/https://datasociety.net/public-technology-leadership-collaborative/

[13] Ali Alkhatib and Michael Bernstein. 2019. Street-Level Algorithms: A Theory at the Gaps Between Policy and Decisions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19).* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300760

[14] Aquisition.gov. 2024. Part 2 - Definitions of Words and Terms. Retrieved September 9, 2024 from https://www.acquisition.gov/far/part-2

[15] Chloe Autio, Kate Cummings, Brinson S. Elliott, and Beth Simone Noveck. 2023. *A snapshot of artificial intelligence procurement challenges.* Retrieved July 21, 2023 from https://files.thegovlab.org/a-snapshot-of-ai-procurement-challenges-june2023.pdf

[16] Johana Bhuiyan. 2021. *LAPD ended predictive policing programs amid public outcry. A new effort shares many of their flaws.* Retrieved January 16, 2024 from https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform

[17] Andrew Bloomenthal. 2024. *Asymmetric Information in Economics Explained.* https://www.investopedia.com/terms/a/asymmetricinformation.asp

[18] Robert Booth. 2025. *'Mainlined into UK's veins': Labour announces huge public rollout of AI.* https://www.theguardian.com/politics/2025/jan/12/mainlined-into-uks-veins-labour-announces-huge-public-rollout-of-ai?CMP=Share_iOSApp_Other

[19] Sean Boots. 2021. *Rule number one: Avoid vendor lock-in.* Retrieved January 16, 2024 from https://sboots.ca/2021/05/12/rule-number-one-avoid-vendor-lock-in/

[20] Margarita Boyarskaya, Alexandra Olteanu, and Kate Crawford. 2020. Overcoming Failures of Imagination in AI Infused System Development and Deployment. arXiv:2011.13416 [cs.CY] https://arxiv.org/abs/2011.13416

[21] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa

[22] Robert Brauneis and Ellen P Goodman. 2018. Algorithmic transparency for the smart city. *Yale JL & Tech.* 20 (2018), 103.

[23] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward Algorithmic Accountability in Public Services: A Qualitative Study of Affected Community Perspectives on Algorithmic Decision-making in Child Welfare Services. In

Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300271

[24] Stephen Casper, Carson Ezell, Charlotte Siegmann, Noam Kolt, Taylor Lynn Curtis, Benjamin Bucknall, Andreas Haupt, Kevin Wei, Jérémy Scheurer, Marius Hobbhahn, Lee Sharkey, Satyapriya Krishna, Marvin Von Hagen, Silas Alberti, Alan Chan, Qinyi Sun, Michael Gerovitch, David Bau, Max Tegmark, David Krueger, and Dylan Hadfield-Menell. 2024. Black-Box Access is Insufficient for Rigorous AI Audits. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) *(FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 2254–2272. https://doi.org/10.1145/3630106.3659037

[25] Hao-Fei Cheng, Logan Stapleton, Anna Kawakami, Venkatesh Sivaraman, Yanghuidi Cheng, Diana Qing, Adam Perer, Kenneth Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 162, 22 pages. https://doi.org/10.1145/3491102.3501831

[26] Trey Childress, Sameer Chowdhary, Drew Erdman, Gayatri Shenai, and Tim Ward. 2024. *Deploying generative AI in US state governments: Pilot, scale, adopt.* Retrieved January 18 2025 from https://www.mckinsey.com/industries/public-sector/our-insights/deploying-generative-ai-in-us-state-governments-pilot-scale-adopt

[27] Alexandra Chouldechova, Diana Benavides-Prado, Oleksandr Fialko, and Rhema Vaithianathan. 2018. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, 134–148. https://proceedings.mlr.press/v81/chouldechova18a.html

[28] Peter Cihon, Moritz J. Kleinaltenkamp, Jonas Schuett, and Seth D. Baum. 2021. AI Certification: Advancing Ethical Practice by Reducing Information Asymmetries. *IEEE Transactions on Technology and Society* 2, 4 (Dec. 2021), 200–209. https://doi.org/10.1109/tts.2021.3077595

[29] Jerrell Coggburn. 2003. Exploring differences in the american states' procurement practices. *Journal of Public Procurement* 3 (03 2003). https://doi.org/10.1108/JOPP-03-01-2003-B001

[30] Cynthia Conti-Cook. [n. d.]. A Guiding Framework to Vetting Public Sector Technology Vendors. https://www.fordfoundation.org/wp-content/uploads/2023/03/final_ford-foundation-guiding-framework-r3-full-document-final2.pdf

[31] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a field scan of the algorithmic auditing ecosystem. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1571–1583. https://doi.org/10.1145/3531146.3533213

[32] Amanda Coston, Anna Kawakami, Haiyi Zhu, Ken Holstein, and Hoda Heidari. 2023. A Validity Perspective on Evaluating the Justified Use of Data-driven Decision-making Algorithms. arXiv:2206.14983 [cs.LG] https://arxiv.org/abs/2206.14983

[33] Catherine Crump. 2016. Surveillance policy making by procurement. *Wash. L. Rev.* 91 (2016), 1595.

[34] William Sims Curry. 2019. *Contracting for services in state and local government agencies.* Routledge.

[35] Emilia David. 2024. *Pennsylvania government workers will start using ChatGPT in test program.* Retrieved January 16, 2024 from https://www.theverge.com/2024/1/9/24031904/openai-pennsylvania-chatgpt-pilot-program-ai

[36] Hannah Quay de la Vallee, Ridhi Shetty, and Elizabeth Laird. 2024. *Report – The Federal Government's Power of the Purse: Enacting Procurement Policies and Practices to Support Responsible AI Use.* Retrieved September 12, 2024 from https://cdt.org/insights/report-the-federal-governments-power-of-the-purse-enacting-procurement-policies-and-practices-to-support-responsible-ai-use/

[37] Ravit Dotan, Emmaline Rial, Ana Maria Dimand, and Virginia Dignum. 2023. How to manage AI procurement in public administration. Retrieved May 9, 2024 from https://www.weforum.org/agenda/2023/07/how-to-manage-ai-procurement-in-public-administration/

[38] Shylah Duchicela, Amanda Jaffe, Laura Merryfield, and Hope Patterson. 2023. *Are Procurement Thresholds Slowing You Down?* Retrieved January 16, 2025 from https://gplpen.hks.harvard.edu/wp-content/uploads/2023/06/Are-Procurement-Thresholds-Slowing-You-Down.pdf

[39] Maddy Dwyer. 2025. *State Government Use of AI: The Opportunities of Executive Action in 2025.* https://cdt.org/insights/state-government-use-of-ai-the-opportunities-of-executive-action-in-2025/

[40] Julia Edinger. 2023. Local Governments Band Together to Address Use of AI. Retrieved May 9, 2024 from https://www.govtech.com/artificial-intelligence/local-governments-band-together-to-address-use-of-ai

[41] David Freeman Engstrom and Daniel E. Ho. 2021. Chapter 3: Artificially intelligent government: A review and agenda. In *Research Handbook on Big Data Law.* Edward Elgar Publishing, Cheltenham, UK. https://doi.org/10.4337/9781788972826.00009

[42] Virginia Eubanks. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

[43] Casey Fiesler, Cliff Lampe, and Amy S. Bruckman. 2016. Reality and Perception of Copyright Terms of Service for Online Content Creation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) *(CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1450–1461. https://doi.org/10.1145/2818048.2819931

[44] The Organization for Economic Cooperation and Development (OECD). 2019. Recommendation of the Council on Artificial Intelligence. Retrieved September 9, 2024 from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[45] Financial Oversight & Management Board for Puerto Rico. 2020. Essay: Improving Public Procurement in PR. Retrieved September 9, 2024 from https://oversightboard.pr.gov/essay-improving-public-procurement/

[46] Katherine B. Forrest. 2021. *When Machines Can Be Judge, Jury, and Executioner: Justice in the Age of Artificial Intelligence.* World Scientific.

[47] R. Geiger, Udayan Tandon, Anoolia Gakhokidze, Lian Song, and Lilly Irani. 2023. Rethinking Artificial Intelligence: Algorithmic Bias and Ethical Issues| Making Algorithms Public: Reimagining Auditing From Matters of Fact to Matters of Concern. *International Journal of Communication* 18, 0 (2023). https://ijoc.org/index.php/ijoc/article/view/20811

[48] Marissa Gerchick, Tobi Jegede, Tarak Shah, Ana Gutierrez, Sophie Beiers, Noam Shemtov, Kath Xu, Anjana Samant, and Aaron Horowitz. 2023. The Devil is in the Details: Interrogating Values Embedded in the Allegheny Family Screening Tool. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1292–1310. https://doi.org/10.1145/3593013.3594081

[49] Jolien Grandia and Joanne Meehan. 2017. Public procurement as a policy tool: using procurement to reach desired outcomes in society. *International Journal of Public Sector Management* 30 (04 2017). https://doi.org/10.1108/IJPSM-03-2017-0066

[50] Lara Groves, Jacob Metcalf, Alayna Kennedy, Briana Vecchione, and Andrew Strait. 2024. Auditing Work: Exploring the New York City algorithmic bias audit regime. arXiv:2402.08101 [cs.CY] https://arxiv.org/abs/2402.08101

[51] L Elaine Halchin. 2006. Overview of the federal procurement process and resources. Congressional Research Service, Library of Congress.

[52] Kirsi-Maria Halonen. 2016. Disclosure rules in eu public procurement: Balancing between competition and transparency. *Journal of Public Procurement* 16 (04 2016), 528–553. https://doi.org/10.1108/JOPP-16-04-2016-B005

[53] Bernard E. Harcourt. 2006. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age.* The Chicago University Press.

[54] Caroline Haskins. 2021. *The NYPD Has Misled The Public About Its Use Of Facial Recognition Tool Clearview AI.* https://www.buzzfeednews.com/article/carolinehaskins1/nypd-has-misled-public-about-clearview-ai-use

[55] Merve Hickok. 2024. Public procurement of artificial intelligence systems: new risks and future proofing. *AI & society* 39, 3 (2024), 1213–1227.

[56] Kashmir Hill. 2023. *Eight Months Pregnant and Arrested After False Facial Recognition Match.* https://www.nytimes.com/2023/08/06/business/facial-recognition-false-arrest.html

[57] The Coalition Against Predictive Policing in Pittsburgh. 2020. *Predictive Policing in Pittsburgh: A Primer.* Retrieved January 16, 2024 from https://capp-pgh.com/files/Primer_v1.pdf

[58] Abigail Z. Jacobs and Deirdre K. Mulligan. 2022. *The Hidden Governance in AI.* https://www.theregreview.org/2022/07/07/jacobs-mulligan-the-hidden-governance-in-ai/,lastaccessed={January16,2024},

[59] Nari Johnson, Sanika Moharana, Christina Harrington, Nazanin Andalibi, Hoda Heidari, and Motahhare Eslami. 2024. The Fall of an Algorithm: Characterizing the Dynamics Toward Abandonment. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) *(FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 337–358. https://doi.org/10.1145/3630106.3658910

[60] Cecilia Kang. 2024. *States Take Up A.I. Regulation Amid Federal Standstill.* Retrieved January 16, 2025 from https://www.nytimes.com/2024/06/10/technology/california-ai-regulation.html

[61] Kari Karppinen and Hallvard Moe. 2012. *What We Talk about When We Talk about Document Analysis.* https://doi.org/10.2307/j.ctv36xvj36.12

[62] Anna Kawakami, Amanda Coston, Hoda Heidari, Kenneth Holstein, and Haiyi Zhu. 2024. Studying Up Public Sector AI: How Networks of Power Relations Shape Agency Decisions Around AI Design and Use. arXiv:2405.12458 [cs.HC] https://arxiv.org/abs/2405.12458

[63] Anna Kawakami, Luke Guerdan, Yanghuidi Cheng, Kate Glazko, Matthew Lee, Scott Carter, Nikos Arechiga, Haiyi Zhu, and Kenneth Holstein. 2023. Training Towards Critical Use: Learning to Situate AI Predictions Relative to Human Knowledge. In *Proceedings of The ACM Collective Intelligence Conference* (Delft, Netherlands) *(CI '23)*. Association for Computing Machinery, New York, NY, USA, 63–78. https://doi.org/10.1145/3582269.3615595

[64] Anna Kawakami, Venkatesh Sivaraman, Hao-Fei Cheng, Logan Stapleton, Yanghuidi Cheng, Diana Qing, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 52, 18 pages. https://doi.org/10.1145/3491102.3517439

[65] Anna Kawakami, Jordan Taylor, Sarah Fox, Haiyi Zhu, and Ken Holstein. 2024. AI Failure Loops in Feminized Labor: Understanding the Interplay of Workplace AI and Occupational Devaluation. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, 1 (Oct. 2024), 683–683. https://doi.org/10.1609/aies.v7i1.31670

[66] Seyun Kim, Bonnie Fan, Willa Yunqi Yang, Jessie Ramey, Sarah E Fox, Haiyi Zhu, John Zimmerman, and Motahhare Eslami. 2024. Public Technologies Transforming Work of the Public and the Public Sector. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work (CHIWORK 2024, Vol. 31)*. ACM, 1–12. https://doi.org/10.1145/3663384.3663407

[67] Josh Klika. 2022. *The ABCs of Piggybacking.* Retrieved January 16, 2024 from https://mrsc.org/stay-informed/mrsc-insight/march-2022/the-abcs-of-piggybacking

[68] Tzu-Sheng Kuo, Hong Shen, Jisoo Geum, Nev Jones, Jason I. Hong, Haiyi Zhu, and Kenneth Holstein. 2023. Understanding Frontline Workers' and Unhoused Individuals' Perspectives on AI Used in Homeless Services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 860, 17 pages. https://doi.org/10.1145/3544548.3580882

[69] Colin Lecher. 2024. *NYC's AI Chatbot Tells Businesses to Break the Law.* Retrieved January 16, 2025 from https://themarkup.org/news/2024/03/29/nycs-ai-chatbot-tells-businesses-to-break-the-law

[70] Karen Levy, Kyla E. Chasalow, and Sarah Riley. 2021. Algorithms and Decision-Making in the Public Sector. *Annual Review of Law and Social Science* 17, 1 (Oct. 2021), 309–334. https://doi.org/10.1146/annurev-lawsocsci-041221-023808

[71] Kevin De Liban. 2024. *Inescapable AI: The Ways AI Decides How Low-Income People Work, Live, Learn, and Survive.* https://www.techtonicjustice.org/reports/inescapable-ai

[72] Robert E Lloyd and Clifford P McCue. 2004. What is public procurement? Definitional problems and implications. In *International Public Procurement Conference Proceedings*, Vol. 3. 2–18.

[73] George Lăzăroiu, Luminița Ionescu, Cristian Uță, Iulian Hurloiu, Mihai Andronie, and Irina Dijmărescu. 2020. Environmentally Responsible Behavior and Sustainability Policy Adoption in Green Public Procurement. *Sustainability* 12, 5 (2020). https://doi.org/10.3390/su12052110

[74] Michael A. Madaio, Luke Stark, Jennifer Wortman Vaughan, and Hanna Wallach. 2020. Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376445

[75] Rohit Madan and Mona Ashok. 2023. AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly* 40, 1 (2023), 101774. https://doi.org/10.1016/j.giq.2022.101774

[76] Christopher McCrudden. 2007. *Buying Social Justice: Equality, Government Procurement, and Legal Change.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199232420.001.0001

[77] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. https://doi.org/10.1145/3359174

[78] Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, Madeleine Clare Elish, and Jacob Metcalf. 2021. *Assembling accountability: algorithmic impact assessment for the public interest.* Retrieved July 21, 2023 from https://datasociety.net/wp-content/uploads/2021/06/Assembling-Accountability.pdf

[79] Deirdre K. Mulligan and Kenneth A. Bamberger. 2019. Procurement As Policy: Administrative Process for Machine Learning. In *Berkeley Technology Law Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3464203

[80] MetroLab Network. 2024. GenAI for Local Governments: Cybersecurity and Privacy Concerns. Retrieved January 18, 2025 from https://metrolabnetwork.org/wp-content/uploads/2024/12/MetroLab_AI_Cybersecurity.pdf

[81] Matt O'Brien. 2025. *Trump rescinds Biden's executive order on AI safety in attempt to diverge from his predecessor.* Retrieved February 7, 2025 from https://apnews.com/article/trump-ai-repeal-biden-executive-order-artificial-intelligence-18cb6e4ffd1ca87151d48c3a0e1ad7c1

[82] NYC Mayor's Office of Contract Services. 2024. Glossary. Retrieved September 9, 2024 from https://www.nyc.gov/site/mocs/about/glossary.page

[83] Pross Oluka, Godfrey Mugurusi, Peter Obicci, and Emmanuel Awuor. 2022. Human-centered artificial intelligence for the public sector: The gate keeping role of the public procurement professional. *Procedia Computer Science* 200 (04 2022), 1084–1092. https://doi.org/10.1016/j.procs.2022.01.308

[84] Elena Parmiggiani, Polyxeni Vassilakopoulou, and Ilias Pappas. 2024. Chapter 5: Reconfiguring human-AI collaboration: integrating chatbots in welfare services. In *Research Handbook on Artificial Intelligence and Decision Making in Organizations*. Edward Elgar Publishing, Cheltenham, UK, 97 – 111. https://doi.org/10.4337/9781803926216.00013

[85] Andrea Stefano Patrucco, Tommaso Agasisti, and Andreas Glas. 2020. Structuring Public Procurement in Local Governments: The Effect of Centralization, Standardization and Digitalization on Performance. *Public Performance & Management Review* 44 (12 2020). https://doi.org/10.1080/15309576.2020.1851267

[86] Gary Peters. 2023. *Testimony of Ritchie Eppink: Governing AI Through Acquisition and Procurement, U.S. Committee of Homeland Security.* Retrieved January 16, 2024 from https://www.hsgac.senate.gov/hearings/governing-ai-through-acquisition-and-procurement-2/

[87] Eric Prier and Clifford P McCue. 2009. The implications of a muddled definition of public procurement. *Journal of Public Procurement* 9, 3/4 (2009), 326–370.

[88] EdTech Equity Project. 2024. School Procurement Guide. Retrieved September 9, 2024 from https://coda.io/d/School-Procurement-Guide_dYBoc7ujwQA/School-Procurement-Guide_su-8G9mx#_luEofXPB

[89] Dasha Pruss. 2023. Ghosting the Machine: Judicial Resistance to a Recidivism Risk Assessment Instrument. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 312–323. https://doi.org/10.1145/3593013.3593999

[90] Dasha Pruss, Hannah Pullen-Blasnik, Nikki Stevens, Shakeer Rahman, Clara Belitz, Logan Stapleton, Mallika Dharmaraj, Mizue Aizeki, Petra Molnar, Annika Pinch, Nathan Ryan, Thallita Lima, David Widder, Amiya Tiwari, Ly Brown, Jason Sexton, and Pablo Nunes. 2024. Prediction and Punishment: Critical Report on Carceral AI. https://doi.org/10.2139/ssrn.5017321

[91] Keely Quinlan. 2024. *Chatbot snapshot: How state, local government websites use AI assistants.* Retrieved January 16, 2025 from https://statescoop.com/government-ai-chatbots-state-local-websites-2024/

[92] Joaquin Quionero-Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D. Lawrence. 2009. *Dataset Shift in Machine Learning.* The MIT Press.

[93] Inioluwa Deborah Raji, I. Elizabeth Kumar, Aaron Horowitz, and Andrew Selbst. 2022. The Fallacy of AI Functionality. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. ACM. https://doi.org/10.1145/3531146.3533158

[94] Dillon Reisman, Jason Schultz, Kate Crawford, and Meredith Whittaker. 2018. *Algorithmic Impact Assessments: A Practical Framework for Public Agency.* Retrieved July 21, 2023 from https://www.nist.gov/system/files/documents/2021/10/04/aiareport2018.pdf

[95] Rashida Richardson. 2021. Best Practices for Government Procurement of Data-Driven Technologies. *Available at SSRN 3855637* (2021).

[96] Rashida Richardson et al. 2019. Confronting black boxes: A shadow report of the New York City automated decision system task force. *AI Now Institute* (2019).

[97]   Tonya Riley. 2022. *Feds' spending on facial recognition tech expands, despite privacy concerns*. Retrieved January 16, 2024 from https://cyberscoop.
        com/feds-spending-on-facial-recognition-tech-continues-unmitigated-despite-privacy-concerns/

[98]   Skyler Rispens. 2024. *San Jose, Calif., forms nationwide government AI coalition*. Retrieved September 12, 2024 from https://statescoop.com/govai-
        coalition-support-local-artificial-intelligence-governance/

[99]   Dorothy E. Roberts. 2022. *I Have Studied Child Protective Services for Decades. It Needs to Be Abolished.* Retrieved January 16, 2024 from
        https://www.motherjones.com/criminal-justice/2022/04/abolish-child-protective-services-torn-apart-dorothy-roberts-book-excerpt/

[100] Samantha Robertson, Tonya Nguyen, and Niloufar Salehi. 2021. Modeling Assumptions Clash with the Real World: Transparency, Equity, and
        Community Challenges for Student Assignment Algorithms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*
        (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 589, 14 pages. https://doi.org/10.1145/3411764.
        3445748

[101] David S Rubenstein. 2021. Acquiring ethical AI. *Fla. L. Rev.* 73 (2021), 747.

[102] Devansh Saxena and Shion Guha. 2020. Conducting Participatory Design to Improve Algorithms in Public Services: Lessons and Challenges. In
        Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing (Virtual Event, USA) *(CSCW '20
        Companion)*. Association for Computing Machinery, New York, NY, USA, 383–388. https://doi.org/10.1145/3406865.3418331

[103] Chinmayi Sharma and Sam Adler. 2024. *The Power of the Collective Purse: Open-Source AI Governance and the GovAI Coalition.* Retrieved January
        16, 2025 from https://www.techpolicy.press/the-power-of-the-collective-purse-opensource-ai-governance-and-the-govai-coalition/

[104] Renee Shelby, Shalaleh Rismani, Kathryn Henne, AJung Moon, Negar Rostamzadeh, Paul Nicholas, N'Mah Yilla, Jess Gallegos, Andrew Smart, Emilio
        Garcia, and Gurleen Virk. 2023. Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction. arXiv:2210.05791 [cs.HC]

[105] Hong Shen, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. Everyday Algorithm Auditing: Understanding the Power of Everyday
        Users in Surfacing Harmful Algorithmic Behaviors. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 433 (Oct. 2021), 29 pages. https:
        //doi.org/10.1145/3479577

[106] Deepa Shivaram. 2024. *The White House issued new rules on how government can use AI. Here's what they do.* Retrieved January 16, 2024 from
        https://www.npr.org/2024/03/29/1241281892/biden-government-ai

[107] Mona Sloane, Rumman Chowdhury, John C. Havens, Tomo Lazovich, and Luis Rincon Alba. 2021. AI and Procurement - A Primer. Retrieved May
        9, 2024 from https://archive.nyu.edu/handle/2451/62255

[108] Jay Stanley. 2018. *New Orleans Program Offers Lessons In Pitfalls Of Predictive Policing.* Retrieved January 16, 2025 from https://www.aclu.org/
        news/privacy-technology/new-orleans-program-offers-lessons-pitfalls-predictive-policing

[109] Logan Stapleton, Min Hun Lee, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022.
        Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. In *Proceedings of the 2022 ACM
        Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York,
        NY, USA, 1162–1177. https://doi.org/10.1145/3531146.3533177

[110] Logan Stapleton, Devansh Saxena, Anna Kawakami, Tonya Nguyen, Asbjørn Ammitzbøll Flügge, Motahhare Eslami, Naja Holten Møller, Min Kyung
        Lee, Shion Guha, Kenneth Holstein, and Haiyi Zhu. 2022. Who Has an Interest in "Public Interest Technology"?: Critical Questions for Working
        with Local Governments & Impacted Communities. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work
        and Social Computing* (Virtual Event, Taiwan) *(CSCW'22 Companion)*. Association for Computing Machinery, New York, NY, USA, 282–286.
        https://doi.org/10.1145/3500868.3560484

[111] Jarrod Stevens. 2022. *How to Prevent Scope Creep Through Better Procurement Planning.* Retrieved January 16, 2024 from https://vendorpanel.com/
        blog/scope-creep-procurement

[112] Josh Taylor. 2024. *AI ban ordered after child protection worker used ChatGPT in Victorian court case.* Retrieved January 16, 2025 from https:
        //www.theguardian.com/australia-news/2024/sep/26/victoria-child-protection-chat-gpt-ban-ovic-report-ntwnfb

[113] Eva Varga. 2021. How Public Procurement Can Spur the Social Economy. Retrieved September 12, 2024 from https://ssir.org/articles/entry/how_
        public_procurement_can_spur_the_social_economy

[114] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public
        Sector Decision-Making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*.
        Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174014

[115] Moira Warburton. 2024. *US lawmakers push for federal data privacy law; tech industry and critics are wary.* Retrieved January 16, 2025 from
        https://www.reuters.com/world/us/federal-data-privacy-laws-gain-support-us-congress-critics-remain-2024-06-26/

[116] Cedric Deslandes Whitney, Teresa Naval, Elizabeth Quepons, Simrandeep Singh, Steven R Rick, and Lilly Irani. 2021. HCI Tactics for Politics from
        Below: Meeting the Challenges of Smart Cities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama,
        Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 297, 15 pages. https://doi.org/10.1145/3411764.3445314

[117] Timothy Williams. 2015. *Facial Recognition Software Moves From Overseas Wars to Local Police.* Retrieved January 16, 2024 from https://www.
        nytimes.com/2015/08/13/us/facial-recognition-software-moves-from-overseas-wars-to-local-police.html?_r=1

[118] Richmond Y. Wong, Michael A. Madaio, and Nick Merrill. 2023. Seeing Like a Toolkit: How Toolkits Envision the Work of AI Ethics. *Proc. ACM
        Hum.-Comput. Interact.* 7, CSCW1, Article 145 (April 2023), 27 pages. https://doi.org/10.1145/3579621

[119] Margaret Young, Michael Katell, and P M Krafft. 2019. Municipal Surveillance Regulation and Algorithmic Accountability. *SSRN Electronic Journal*
        (01 2019). https://doi.org/10.2139/ssrn.3429872

[120] Tom Zick, Mason Kortz, David Eaves, and Finale Doshi-Velez. 2024. AI Procurement Checklists: Revisiting Implementation in the Age of AI Governance. arXiv:2404.14660 [cs.CY] https://arxiv.org/abs/2404.14660

[121] Marta Ziosi and Dasha Pruss. 2024. Evidence of What, for Whom? The Socially Contested Role of Algorithmic Bias in a Predictive Policing Tool. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) *(FAccT '24).* Association for Computing Machinery, New York, NY, USA, 1596–1608. https://doi.org/10.1145/3630106.3658991

A Defining "procurement"

All the definitions of "public procurement" that we reviewed encountered share fundamental similarities: they all concern the process of bringing in goods and services that are developed externally, to achieve the goals of a public sector entity. They differ, however, in specific components of the process. For example, while the United States' federal definition emphasizes a "competitive" purchasing process, denoting the exchange of money as part of procurement, some local governments, like New York City, have definitions that are broader, encompassing all functions related to obtaining goods and services whether or not money changes hands [14, 82].

In this paper, we do not adopt a single definition of public procurement, as methodologically we chose to leave such distinctions to our interviewees who were encouraged to discuss whatever processes and components they personally and professionally associated with public procurement. Given the broader diversity of the term, as would be expected, we observed differences across municipalities in what types of acquisitions and activities participants deemed to fall under the umbrella of "procurement". For example, procurement departments often did not oversee governments' acquisition and use of free technologies, which we discuss further in Section 4.2.

B   Cities' AI Use Cases

Table 2 groups examples of AI adopted by municipalities into five categories based on their intended usage. In our discussions, employees in each city shared at least one example that they were aware of belonging to one of these five categories.

Interestingly, not all of the employees that we interviewed were aware that other employees in their city had already procured or adopted AI technologies: for example, one city employee stated that to their knowledge, their city "has never purchased anything AI related", whereas their colleagues stated that the city in fact has.

Type of AI technology	Examples
Facilitating resident communication	Translation services, chatbots, 311 assistance, public meeting summaries
Law enforcement	License plate readers, gunshot detection, object detection
Smart cities/urban planning	Sensors to track service utilization, accident tracking, snow plow routing
Assisting bureaucratic decision-making	Funding allocation, service allocation, school bus routing
Workplace productivity tools	Chatbots, image generation, voice generation, coding assistants

Table 2. We grouped the AI systems that municipal employees discussed procuring or adopting in interviews, into 5 categories based on their intended usage. We provide anonymized examples of types of AI systems that were mentioned in each category. Employees in each city shared at least one example that they were aware of belonging to one of these five categories.

C   Cities' emerging AI governance practices (extended)

In this section, we present an extended description of interviewed governments' emerging AI governance practices. We do not provide exact descriptions to preserve the anonymity of participating cities, and instead aim to pull out broad trends across cities.

AI risk or impact assessments. Participants conducted additional risk or impact assessments to better understand the possible positive or negative impacts of procured AI systems. While some cities conducted such assessments in an informal or ad-hoc way, others had started to standardize assessment processes by creating assessment templates with lists of questions and considerations. Different cities also conducted risk or impact assessments at different phases of the procurement process: some assessment instruments could be completed based on a "purpose statement" for AI, before a specific vendor or AI system is identified. In contrast, other risk assessments can only be completed once a concrete system has been identified, *e.g.,* they require knowledge of the system's performance.

The role and purpose of these assessments varied across interviewed cities. In many cities, the risk assessment had no immediate outcome, but employees were encouraged to take action to manage and if possible, mitigate potential risks identified in the process. Beyond informing mitigation steps, some participants also used risk assessments to triage AI solutions into "high" or "low" risk categories, which then determined subsequent requirements for review and oversight. For example, one city required high-risk AI to have additional reporting requirements, further risk assessment, usage protocols, and regular post-deployment monitoring. Participants viewed risk triaging as a way to reduce reviewing burden and better allocate their limited technical expertise. One participant who conducted AI risk assessments explained:

> "[When triaging risk], we're just trying to get a sense of how thorough a review we need to do, because we're working with very limited capacity and resources. So we've got to decide: is this a low-risk system that we can just do a really quick look at? Or is this going to be something really sensitive and safety-impacting, rights-impacting, that we need to dedicate a lot of our time to?"

Participants also noted how risk triaging was also time-saving for their colleagues on the other end trying to purchase the AI, as put by one employee: "If it's low risk, I'll approve it, and you'll be on your way tomorrow!"

Practice vs. policy? In many cities, employees made changes to their procurement practices simply by adjusting their existing practices, *e.g.,* by electing to include vendor reporting requirements in an AI RFP. Some cities decided to make these changes in their practices more formal or mandatory for vendors or city employees, by adopting policies or passing laws that required them. For example, Rebecca, an IT department leader walked through how their city's formal AI policy spelled out mandatory steps, such as a risk assessment, that city employees must complete for any AI procurement. The participant viewed the policy, which was passed by their city council, as an "accountability trigger" to incentivise compliance for both colleagues and vendors:

> "Council adopted the policy. So you can't just say no. I'm going to have some leverage to say, we can't just say we're not going to do this. [...] [The policy] is really meant to be a way to say the city is going to be taking this on, these are our values."

Participants in another city shared that while ideally someday they would like to institutionalize their practices via a formal policy, at the time of interviewing, they did not yet have one:

"We very intentionally have not put out a [formal] AI policy yet, because we wanted more [community and government] input on it. And the space, especially in 2023, was very new for us. So we wanted to get a better understanding before asking our leadership to pass a policy."

This city has since adopted a formal AI policy following engagement with the community, experts, and agency staff.

D   Interview Protocol

We began the interview by reminding the participant of our informed consent protocol (approved by our institution's IRB board), and asking for their consent to record.

Introduction. The goal of this interview is to learn more about existing procurement practices specifically for artificial intelligence, or AI, technologies in your city. We adopt a wide definition of AI as "any machine-based system that can make predictions, recommendations, or decisions". This would include technologies such as facial recognition, gunshot recognition technology, resume screening technology, ChatGPT, etc.

Our goal is not to assess your practices, but rather to identify needs and opportunities for researchers as partners to support US cities.

Q1.1: Can you tell me a bit about your current role, and any past work experiences or responsibilities relating to artificial intelligence?

Q1.2: Have you ever been involved in a past procurement of an artificial intelligence technology?

- If YES: How were you involved?
- If NO: Has your [agency] ever considered or talked about procuring AI?

Walk-through. The goal is to understand how a "typical" AI procurement occurs in the city. Our goal is not to impose structure on the participant's description, but rather allow them to describe how they personally view/understand the procurement process.

If it doesn't come up naturally, we can prompt them to reflect on specific parts of procurement, such as (1) Planning, (2) RFP writing, (3) Evaluating Vendors, (4) Contracting, (5) Designing/Building/Evaluating the AI, and (6) Deployment, and (7) Post-deployment.

Q2.1., Walkthrough. Can you briefly walk us through how a typical procurement involving an artificial intelligence technology would occur in your city? We're specifically interested in understanding any difference between a standard technology procurement, vs. a procurement involving AI.

- If never procured AI: *e.g.,* imagine your city is considering procuring an enterprise-level generative AI product, like a chatbot to screen 311 questions.

Drill-down prompts on specific parts of the procurement process:
Planning (Problem Formulation):

(1) What does your city do to plan for the procurement before the RFP (request for proposal) writing stage?
(2) (if not covered) Pre-RFP, how does the agency identify that an AI tool might be a part of the solution (rather than a tool that does not use AI)?
(3) (if not covered) Do you have a process for evaluating the risks of a proposed AI technology before RFP writing?
    - If YES: What about potential mitigation processes for these risks?

RFPs:

(1) Is there anything different in the content of the RFP for AI procurements, compared to standard technology procurements that do not involve AI?
(2) (if not covered) In the RFP, do you ask vendors questions about potential risks and mitigation strategies?

Evaluating Proposals:

(1) How does your city evaluate proposed AI solutions? We are especially interested in differences between evaluating standard technology vs. AI proposals.

(2) (if not covered) What information do you ask vendors to report in their proposal? Do you ever encounter "trade secrecy" claims?

(3) (if not covered) What measures do you expect them to report? Do they validate that the technology works as claimed using data from your city?

Contracting:

(1) Are there any differences in the contracting process for AI vs. non-AI (standard technology) technologies?

(2) Are there specific terms and conditions that you include in AI contracts?

(3) Can you share a past contract for an AI technology with us?

AI Design, Development, and Evaluation:

(1) How are people from your city involved with the design, development and evaluation of AI technologies under contract?
  - If YES: How were you involved? What type of feedback did you give?

(2) How often do vendors make changes to their technologies (like updating or improving it using data from your city) before they are deployed?

(3) How do vendors evaluate the AI solution they have designed and developed to make sure it fits your use case?
  - Do they use data from your municipality for evaluation?
  - What kind of measures do they look at and report to you?

AI Deployment:

(1) How often do vendors (or the city) provide training or onboarding for people who will be using the AI?

(2) How are agency workers involved in deciding the way the AI is used in their everyday practice?

Post-deployment:

(1) How do you oversee and monitor deployed AI technologies?
  - What is the vendors' responsibility?
  - What if something goes wrong? (liability)

Q2.3 (if unclear) Can you remind me of who in your city is involved or oversees each phase of this procurement process?

Q2.4 (if unclear) Do you believe the process that we just went through together is representative of most AI procurements in your city (if relevant: beyond that specific example)?

Q2.5. Are there any existing policies in place that target the procurement of AI technologies specifically?

- If YES:
  – Can you share your city's policies/guidelines with us?
  – How long have these policies been in place?
- If NO:
  – Is this something you anticipate being developed in the near future, or something that has been discussed?

Q2.6. Can you direct us to your city's general procurement policies that may be applicable to AI technologies? e.g., such as data privacy policies?

Q2.7. Are there any AI technologies that come to be used through processes outside of the traditional procurement process? (e.g., research partnerships, foundations, donations, or free tools?)

- Do these technologies undergo a similar "vetting" process to procured technologies?
  - Do similar people evaluate these proposals?
  - Do similar people oversee or monitor their deployment?

Q2.8. Does your city consider opportunities to engage with residents who may be affected by an AI tool during the procurement process?

Challenges & Desires. The goal is to understand the participant's needs and desires to improve the procurement process. For the last part of our interview, we'd like to understand your opinions and wishes for improving AI procurement.

Q3.1. What do you believe are the main challenges or "pain points" for AI procurement in your city?

- Do you have any suggestions as to how cities could improve their procurement of AI?
- (if relevant) Do you have any examples where [this challenge] happened in the past?

Q3.2. Can you imagine any new resources that could help you address these challenges?

- What resource format would be most helpful? ex: Checklists? Templates? Trainings?