Uncovering a Culture of AI Grassroots Experimentation in Boston City: Safety Risks and Mitigation

Audrey Chang [*1] **Jude Ha** [*1]

Abstract

Through a series of interviews with Boston City employees, we investigate the phenomenon of "grassroots experimentation," where municipal employees independently experiment with AI tools outside of formal procurement channels. This practice of informal, "under the radar" tech adoption is motivated by the inability of procurement guidance to keep pace with the recent proliferation in accessible, low cost AI tools. Our first case study reveals how this self-directed exploration influences AI integration in the highly sensitive application of municipal public services. In three subsequent case studies, we highlight the ethical and security concerns of experimental AI usage at the municipal level. Our final case studies identify strong team leadership and a supportive tech culture - one that recognizes and respects the grassroots experimentation phenomena - as crucial factors in mitigating the risks posed by informal adoption while harnessing the benefits of AI experimentation to best support public servants. These insights may be useful for future policy innovation that empowers employees to adopt AI tools to improve municipal service provision in a safe and appropriate manner.

1. Introduction

Municipal governments manage a wealth of sensitive data — including tax and property reports, criminal history records, and marriage certificates — that informs numerous high-risk, high-impact decisions. While municipal automation through AI may streamline this decision-making, there is a high possibility for unsafe use that puts resident data or well-being at risk. As a result, effective municipal technological procurement is crucial to ensuring AI systems enhance decision-making without compromising human expertise and judgment.

The City of Boston is a prime example of recent municipal tech innovation and was the first major US city to release its Interim GAI Guidelines (Garces, 2023) for employees. However, these guidelines - and similar ones released by the cities of San Jose, Seattle, and others - may lag behind the rapid growth of general-purpose AI and other computational tools. Driven by advances in Generative AI (GAI), computational tools have become increasingly accessible to use, with over 20% of all employees reporting regular use of GAI. As a result, it is increasingly relevant for any institution looking to safely harness the power of in-house AI to understand how on-the-ground adoption of AI tools by employees occurs.

We present six case studies from Boston City that identify a novel phenomenon we call "grassroots experimentation" that has arisen among individuals/micro-networks of employees as city guidelines struggle to keep pace with the astounding growth of general-purpose AI tools. Previous work on safe AI use in public-use settings typically takes a "top-down" perspective that focuses on AI-specific procurement policy and administrative considerations. Our work goes further by taking a "bottom-up" approach to understanding broader tech adoption from a city employee's perspective, characterizing grassroots experimentation, and the challenges to safety and effectiveness that it poses. We investigate computational tools, which frequently include, but are not limited to, low-cost AI tools in an effort to understand how potential AI governance will build on existing tech adoption structures, ultimately leading to tech use that is trustworthy and empowering for all public servants.

2. Related Works

The integration of AI in municipal governance is a nascent field with significant potential to enhance government capabilities by increasing public servants' efficiency and improving services (Autio et al., 2023; Intel, n.d.). However, there are risks, such as data breaches, amplified biases, and

[*]Equal contribution [1]Department of Computer Science, Harvard University. Correspondence to: Audrey Chang <audreychang@college.harvard.edu>, Jude Ha <jhchang@college.harvard.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

privacy infringements (Carlini et al., 2022; Hall et al., 2022; Lee et al., 2024). Municipal governments handle sensitive data, like personal tax information and health records, necessitating thoughtful AI adoption to avoid biases, mitigate security risks, and minimize harm to people, organizations, and communities (3M, n.d.).

Research on municipal tech procurement can be categorized into two perspectives: smart city literature and AI governance. Smart cities are built on embedding pervasive and ubiquitous digital devices into urban environments that are used to monitor and manage city processes in real-time (Kitchin, 2014). Procurement-related smart city research focuses on the infrastructural needs of data-driven cities, emphasizing efficient data systems as foundations for effective AI tools. This includes studies on implementation challenges of canceled Automated Decision Systems (Our Board & Our Associates, 2022), real-time city data infrastructure (Kitchin, 2014), and characterizing fragmented data governance in urban data ecosystems (Kitchin & Moore-Cherry, 2021). Our study complements existing literature by characterizing tech adoption from the employee perspective, examining how staff interact with AI tools and data. This approach informs policies on workplace practices rather than infrastructural recommendations, but emphasizes that both management and workplace culture must advance for thoughtful technological progress.

AI governance research addresses risks through governance frameworks, ethical considerations, and accountability measures for responsible AI use. Procurement is seen as an avenue for tech governance, with studies on privacy and accountability challenges of procurement (Noya et al., 2021), procurement as AI "soft law" governance (Ben Dor & Coglianese, 2021), and risks in federal AI acquisition (Autio et al., 2023). Our work provides a sociotechnical analysis of AI governance, focusing on pre-procurement dynamics often overlooked by formal policy development. Understanding of this ground-level perspective is crucial for effective policy intervention. Previous studies on interplay between governance and practical implementation have explored municipal AI capabilities (Patrick Mikalef a, Kristina Lemmer b, Cindy Schaefer b, Maija Ylinen c, Siw Olsen Fjørtoft a, Hans Yngvar Torvatn a, Manjul Gupta d, Bjoern Niehaves b), algorithmic audits (Radiya-Dixit & Neff, 2023), and policy adaptations to new technologies (ACM FAccT Conference, 2023).

This work uniquely contributes to smart city literature and AI governance, examining how current workplace dynamics evolve with low-cost, widely available AI tools. This holistic approach informs both municipal infrastructure and AI governance.

3. Background

The City of Boston has a history of technological progressivism, in terms of tech innovation, transparency, and regulation. The Mayor's Office of New Urban Mechanics (MONUM), established under Mayor Menino, exemplifies this with initiatives like a pay-by-phone parking program (Graham, 2014) and COVID-19 chatbots. Many departments maintain a "product-manager"-like internal structure, where those interacting with tech stakeholders and implementing tech can influence the policy itself (Ryan-Mosley, 2023). Boston's status as a tech leader makes it a prime candidate for studying the impacts of tech adoption within a large city government. Boston also faces other issues endemic to city government, such as inconsistent data governance (Kitchin & Moore-Cherry, 2021), lapsed or overshadowed programs, and employee mistrust or unfamiliarity with technological solutions (Clauss, 2016). The fragmented nature of technological experimentation may be a culprit; most of the initiatives mentioned above are project or department specific and do not constitute an institution-wide shift in culture. As the MIT Technology Review notes, a "central challenge" for the Wu administration is attempting "transformation within an organization that is built to move slowly." (Bliss, 2023) These problems are compounded by the fact that AI and newer computational tools are more likely than other services to not be covered by the city's existing procurement processes, which are meant to handle big ticket items, not low-budget software services. In order to meet the current AI wave, these challenges will need to be first characterized, then addressed, to create a more comprehensive, robust, and responsible framework towards technical adoption. The city's recently released Interim Guidelines for Using Generative AI (GAI), which encourages cautious employee use of GAI, is an early indicator of movement in this direction.

Note that computational tools are currently subject to the same procurement process as non-computational items and services. This process imposes limits on what can be procured, primarily based on dollar value (i.e. any contract over $50k must put out an open advertisement). Formal guidelines state that for procurements under $10k, departments should consult the Directory to identify possible vendors and periodically solicit price lists, but in practice, the process for procurements under $10k is significantly more free-flowing. In fact, interviewees only mentioned the $50k mark (and not the $10k mark). Furthermore, the nature of many computational tools is that they tend to be cheaper, easier to access, and more available on a trial basis than traditionally procured goods and services. This means that many technical solutions can be quickly and easily adopted without going through any formal procurement process.

4. Methods

We conducted a series of semi-structured interviews with 17 Boston City employees over the course of two months. Interviewees were sourced by recommendation from partners in the Department of Innovation and Technology (DoIT) and MONUM, and added until saturation was achieved:

1. Interviewees had worked at the city anywhere from a few months to 15+ years. The average tenure at the city was 4.5 years, and the median tenure was 2 years.

2. Interviewees came from 14 departments (of the city's 94). These departments were resident-facing (Arts and Culture, Community Engagement, Registrar, Returning Citizens, Women's Cabinet, Mayor's Policy Team, Early Childhood, Transportation, Environment), administrative (Operations, People's Operations), and technical (Cybersecurity, DoIT, IT).

3. Interviewees held multiple roles, ranging from intern to project manager, policy associate, data analyst, department head, and marketing coordinator.

Interviews lasted one hour each and were based off of 11 standardized questions (Appendix A.1). All interviews were transcribed by hand without the aid of AI-powered transcribers. Thematic content analysis was performed by authors independently to identify and verify core interview themes. All interviewees are referred to by pseudonyms or numbers. Tools referenced in case studies are also referred to by pseudonyms (names of popular games).

5. Results

The following characterizes a broad pattern of grassroots experimentation and its implications. We are unable to provide direct quotes from individual employees per a data privacy agreement with Boston City employees, but present our summarized case studies, of which five (of six) were corroborated from multiple interviewee accounts.

Tech advocates provide individualized support for tech experimentation. *(Case Study 1)* Gary, a member of the Mayor's Office, is an enthusiastic adopter of Tetris, a popular database service that allows free trials for teams of up to six users. As an avid Tetris user, Gary hosted a formal training on the tool with MONUM. Attendees who were interested in using Tetris for their own work then reached out to Gary, who holds informal "office hours" and helps them troubleshoot in an unofficial capacity. While we did not interview Gary, we heard about their work through two separate interviewees who did not know each other and worked in separate departments. One employee, Interviewee 8, explained that when they ran into issues with Tetris, their first course of action was to talk to Gary, describing the office hours as informal conversations.

Interviewee 6, who also attended Gary's MONUM training, took the initiative to begin a free trial for Tetris within their office. At the time of the interview, they were planning to present the results from their free trial to convince their department to procure a paid version of Tetris. Interviewee 6 described the critical importance of the hands-on, individualized support that they and Gary provide for colleagues, emphasizing that individualized support was far more effective in assisting their peers in adopting Tetris use.

Gary's advocacy for Tetris is emblematic of a larger pattern of grassroots experimentation and associated tech adoption observed among city employees. Experimentation relies on highly connected "tech advocates" such as Gary, who encourage their peers to adopt a tool through personal demonstration, onboarding, and support of tool use. Advocates tend to be newer employees, have previous technical experience, and typically focus on promoting one piece of technology at a time using a hands-on approach. Rather than establishing a formal program for procurement or training of a tool, they will often provide informal "office hours" or individualized IT-like support. They leverage informal relationships to push through Boston's usually cautious tech culture and are highly interconnected; we would hear about the same tech advocate from interviewees in different departments, affirming a "grassroots" network by which experimentation with new tools spreads.

Advocate-led AI adoption can lead to untrustworthy decision-making. *(Case Study 2)* AI tool Qwirkle collects phone data with user permission and tracks the user's movements by car or bike. Qwirkle was formally procured through an RFP (request for proposal) after a compelling and persistent sales pitch to a former team manager who wished to better understand how people move through the city. However, despite initial excitement and a months-long procurement process, the tool's performance had not been sufficiently vetted in a manner appropriate for a modern computational tool. After procurement, employees realized significant inaccuracies in the Qwirkle-collected data when compared with historical city data, leading to the tool falling into disuse. This experience with Qwirkle prompted reflection on the risks of advocate-led, hasty procurement contracts, again demonstrating the need for procurement guidance on low-cost tech tools. Without standardized governance and the infrastructure to transition from an informal free trial to a department-wide procurement, employee-led experimentation risks incompatible or near-sighted adoption, security breaches, and tools that end up detracting from, rather than empowering, employee decision-making.

Unregulated grassroots experimentation poses severe privacy risks. *(Case Study 3)* Uno is a voice transcription

autopilot that can attend virtual meetings and record meeting data. While Uno has not been formally procured by the city, one interviewee described how colleagues would personally install the service for use without consulting the rest of the department. Other employees did not realize they were being recorded while the tool was in use and were concerned about the autopilot recording constituent information or sensitive conversations when in use. This highlights a severe safety concern regarding the collection of data of both constituents whose information is discussed, and less technically-fluent employees who have not consented to being recorded.

The interviewee also told us that external partners who independently adopted Uno often send their autopilots to meetings when they cannot show up themselves. While unrelated to stratification of technical knowledge, this third-party adoption leads to confusion about how to collaborate with partners using tools, which is currently unaddressed under current guidelines. In the absence of official guidance for low-cost tools, less experienced employees are more conservative and wait for the "green light" from city leadership. This leads to uneven tech adoption and widens the gap between tech-fluent and less fluent employees on the same team.

Balancing safety with shifting employee needs remains a challenge. *(Case Study 4)* Atari is another automated transcription AI tool with a free trial version that employees were able to independently access through a website. The tool was reported to expedite tasks, and spread organically through word of mouth. Within a few weeks, enough employees had experimented with it that the administration called to put the tool on hold in order to conduct a more thorough vetting. A member of the Cybersecurity department expressed consternation that the department had not been allowed to vet the tool for safe use before its ad hoc adoption. Employees had mixed opinions about the tool: while some trusted the city's decision to pause Atari, others were frustrated by the city's hesitancy to adopt computational tools that clearly accelerated their work. Several employees mentioned being overwhelmed or burned out, seeking out tools like Atari to reduce their burden, even if at a cost to privacy. In the future, we anticipate that municipalities will see more tools that will gain traction among employees through word of mouth, with little prior knowledge of the tool's inner workings. Furthermore, while the Atari case study illustrates the need for a flexible approach that balances security concerns with the desire of early adopters for progress, we are skeptical that such a case-by-case review system will be sustainable in the future.

Team leadership may add structure, safeguarding to experimentation. *(Case Study 5)* Rummy is a licensed data visualization software that DoIT recently switched to from a competitor, Scrabble. The department underwent an experimental but rigorous process to choose a successor for Scrabble, creating demo accounts for several competitors, duplicating dashboards across the different tools, and regularly meetings to debrief their tool preferences. Once the tool was procured, the entire team underwent training run by staffers with previous experience with Rummy. Afterwards, the team ran trainings for staffers across the city in small cohorts. Rummy was a crucial tool for a number of departments, so departments sent a representative to familiarize themselves with the tool via training. This is a successful example of grassroots adoption supported by team leadership, where individual experimentation is followed by formalized presentation and discussion of results. This pattern of acquisition achieves the benefits of employee-led experimentation while preventing confusion and lapsed standards, and may provide a positive model for flexible, employee-centered tool adoption.

Strong citywide values compensate for lack of experimentation infrastructure. *(Case Study 6)* Interviewee 5 and several other interviewees recounted an incident with Catan, a no-code platform used to build custom databases and forms with added AI capabilities. Interviewee 5's team intended to send status updates to specific clients but due to a Catan glitch, the messages were sent to every client who had ever requested a project with the team. They resolved the incident with a mass email explaining the glitch and its impact. Despite the lack of existing protocols to address tool-specific incidents such as this one, Interviewee 5 emphasized that strong guiding values from their chief and deputy chief were sufficient to support them through the situation. Although the team was still sending out emails months after the incident, they believed the administration's priority on communication and transparency allowed them to resolve the situation as with speed and efficiency. Catan is an example of how strong, trustworthy leadership compensates where formal procedures do not exist. Guiding principles in action during the major failure of a widely adopted computational tool, highlight how trust in city leadership is crucial to supporting and balancing out responsible grassroots tech adoption and use.

6. Discussion

Our findings prompt a discussion on implications for policymaking, organizational culture, and the responsible integration of AI systems in high-risk domains like municipal governments.

We see that grassroots experimentation can present several disadvantages, including:

- **Near-Sighted Adoption:** Advocate-led adoption is often hasty, leading to the adoption of tools that are un-

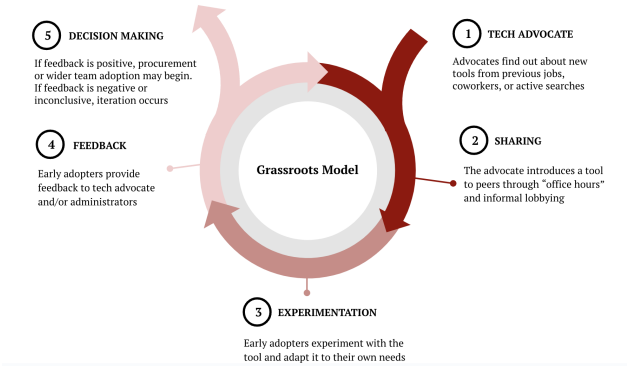

Figure 1. **Grassroots experimentation is characterized by tech advocates' sharing, experimentation, feedback-collection, and decision-making.** The cyclic process of grassroots adoption is driven by tech advocates' coordination and advocacy for a specific computational tool, most emblematic in Case Study 1.

trustworthy or unsafe, reducing the quality of decisions made using these tools.

- **Security Concerns:** Grassroots-acquired tools lack formalized vetting, potentially sacrificing security and privacy standards. Ad hoc experimentation may end up adversely affecting individuals (colleagues and constituents alike) who have not consented to the tool's use.

- **Uncoordinated Acquisition:** Tech-savvy staff push for experimentation, leaving less experienced employees struggling to catch up, slowing overall productivity and causing frustration. Simultaneously, exhaustive exploration of options that are both safe and useful for all employees is almost impossible.

Despite these disadvantages, we also identify several instances of healthy grassroots experimentation that is supported by citywide leadership and augments decision-making while avoiding potential pitfalls. Benefits of supported experimentation include:

1. **Rapid Adoption:** Informal tech support provided by advocates accelerates tool integration faster than official training, strengthening arguments for procurement when employees have tested and resolved issues.

2. **Immediate Feedback:** Small-scale experimentation allows quick adjustments and instant feedback on tool performance.

3. **A Healthy Tech Culture:** Grassroots experimentation fosters a supportive environment for discussing and testing new tools, promoting comfort with AI tools that benefits both city employees and residents.

These case studies highlight the need for policy frameworks that can accommodate the rapid adoption of AI tools while ensuring accountability, security, and effectiveness. Although these results apply to all computational tools, both the risks and potential benefits of grassroots adoption can be amplified by the growth in use of AI tools specifically. Traditional top-down procurement policies may not be agile enough to keep pace with the proliferation of low-cost AI tools, leading to instances of ad hoc adoption without sufficient oversight, requiring policy that balances fostering innovation and maintaining regulatory control.

Grassroots experimentation often relies on the enthusiasm and expertise of individual advocates who champion specific tools within their departments. While this bottom-up approach can drive innovation and agility, it also risks exacerbating disparities in technical knowledge across teams and departments. Recommendations presented to the City of Boston are included in the Appendix (A.2) highlight the role of organizational culture in shaping productive tech adoption and mitigating articulated risks.

Future studies could interview a larger sample of city employees to confirm findings and assess grassroots experimentation in other municipalities. Contrasting informal policy interventions in response to low-cost AI tools' proliferation among municipalities may also offer insights into effective methods for shifting workplace culture.

7. Conclusion

Our exploration into the intersection of top-down procurement policies and grassroots tech adoption within municipal governments sheds light on a crucial yet under-explored aspect of AI integration.

Our findings motivate comprehensive and feasible tech policy that is designed for technological ecosystems in their entirety, which include but are not limited to AI tools. For this reason, our conclusions apply broadly to tech adoption practices and computational tool use, which have meaningful implications for both AI and broader technological procurement.

Individual grassroots experimentation with AI tools, while offering advantages such as rapid adoption, immediate feedback, and a healthy tech culture, also presents significant safety risks, including security breaches and adoption of untrustworthy AI systems. Without standardized governance of individual experimentation and the infrastructure to transition an informal free trial into a department-wide procured tool, this form of experimentation observed in our case studies becomes haphazard and inefficient. Employees describe

a compromised ability to fulfill their roles and unnecessary privacy risks. However, amidst these challenges lie valuable insights and examples of how team leadership and strong city guidance can mitigate risks and foster responsible tech adoption.

To ensure that AI systems in sensitive application areas enhance decision-making, our findings emphasize that policymakers should not only develop robust procurement policies but also cultivate a supportive tech culture within municipal governments to facilitate intended implementation that prioritizes human judgment. Technical interrogation of systems themselves is insufficient — appropriate training, flexible team procedures, trusting attitudes towards values-based leadership, and an openness to experimentation are crucial to agile teams that can reap the benefits of AI integration in high-risk scenarios.

8. Acknowledgements

Thank you to our advisors Dr. Weiwei Pan, Dr. Dasha Pruss, and Hongjin Lin for their support and feedback, in addition to our City of Boston partners aleja jimenez jaramillo (DoIT) and Yo Deshpande (MONUM) for their insight and collaboration. This project was conducted through Harvard Responsible Computing Collective, ReCompute.

References

3M. Visual privacy: A critical area of data protection in state and local government. n.d. URL https://multimedia.3m.com/mws/media/9934580/a-critical-area-of-data-protection-in-state-and-local-government.pdf.

ACM FAccT Conference. Tutorial 10.2: The rubber meets the road: Implementing AI governance in a public agency in san josé, July 2023. URL https://www.youtube.com/watch?v=Bif3fwI_d20.

Autio, C., Commings, K., Elliot, B. S., Noveck, B. S., et al. A snapshot of artificial intelligence procurement challenges: Diagnosing perceived and actual risks impeding responsible ai acquisition in government. *a-snapshot-of-ai-procurement-challenges-june2023. pdf*, 2023.

Ben Dor, L. M. and Coglianese, C. Procurement as AI governance. *IEEE Transactions on Technology and Society*, 2(4):192–199, December 2021. ISSN 2637-6415. doi: 10.1109/TTS.2021.3111764. URL http://dx.doi.org/10.1109/TTS.2021.3111764.

Bliss, L. The startup CEO remaking city hall. *MIT Technology Review*, April 2023. ISSN 0040-1692. URL https://www.technologyreview.com/2023/04/25/1070381/the-startup-ceo-remaking-city-hall/.

Carlini, N., Ippolito, D., Jagielski, M., Lee, K., Tramer, F., and Zhang, C. Quantifying memorization across neural language models. February 2022. URL http://arxiv.org/abs/2202.07646.

Clauss, K. S. What ever happened to CityScore? https://www.bostonmagazine.com/news/2016/08/17/boston-cityscore-national/, August 2016. URL https://www.bostonmagazine.com/news/2016/08/17/boston-cityscore-national/. Accessed: 2024-5-30.

Garces, S. City of boston interim guidelines for using generative ai, May 2023. URL https://www.boston.gov/sites/default/files/file/2023/05/Guidelines-for-Using-Generative-AI-2023.pdf.

Graham, J. Boston taps zen of paying parking tickets. *Boston Herald*, September 2014. ISSN 0738-5854. URL https://www.bostonherald.com/2014/09/23/boston-taps-zen-of-paying-parking-tickets/.

Hall, M., van der Maaten, L., Gustafson, L., Jones, M., and Adcock, A. A systematic study of bias amplification. January 2022. URL https://github.com/facebookresearch/cv_bias_amplification.

Intel. The future of artificial intelligence (AI) in government. https://www.intel.com/content/www/us/en/government/artificial-intelligence.html, n.d. URL https://www.intel.com/content/www/us/en/government/artificial-intelligence.html. Accessed: 2024-5-30.

Kitchin, R. The real-time city? big data and smart urbanism. *GeoJournal*, 79(1):1–14, February 2014. ISSN 0343-2521, 1572-9893. doi: 10.1007/s10708-013-9516-8. URL https://doi.org/10.1007/s10708-013-9516-8.

Kitchin, R. and Moore-Cherry, N. Fragmented governance, the urban data ecosystem and smart city-regions: the case of metropolitan boston. *Reg. Stud.*, 55(12):1913–1923, December 2021. ISSN 0034-3404. doi: 10.1080/00343404.2020.1735627. URL https://doi.org/10.1080/00343404.2020.1735627.

Lee, H.-P. h., Yang, Y.-J., Von Davier, T. S., Forlizzi, J., and Das, S. Deepfakes, phrenology, surveillance, and more! a taxonomy of AI privacy risks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, number Article 775 in CHI '24, pp. 1–19, New York, NY,

USA, May 2024. Association for Computing Machinery. ISBN 9798400703300. doi: 10.1145/3613904.3642116. URL https://doi.org/10.1145/3613904.3642116.

Noya, M., Varela, M., and Ospina-Fadul, M. J. Innovation management and public procurement of AI. In *Artificial Intelligence and Innovation Management*, volume 38 of *Series on Technology Management*, pp. 259–286. WORLD SCIENTIFIC (EUROPE), August 2021. ISBN 9781800611320. doi: 10.1142/9781800611337_0013. URL https://doi.org/10.1142/9781800611337_0013.

Our Board and Our Associates. Automating public services: Learning from cancelled systems. https://carnegieuktrust.org.uk/publications/automating-public-services-learning-from-cancelled-systems/, September 2022. URL https://carnegieuktrust.org.uk/publications/automating-public-services-learning-from-cancelled-systems/. Accessed: 2024-5-30.

Patrick Mikalef a, Kristina Lemmer b, Cindy Schaefer b, Maija Ylinen c, Siw Olsen Fjørtoft a, Hans Yngvar Torvatn a, Manjul Gupta d, Bjoern Niehaves b. Enabling AI capabilities in government agencies: A study of determinants for european municipalities. *Elsevier*. URL https://www.sciencedirect.com/science/article/pii/S0740624X21000320.

Radiya-Dixit, E. and Neff, G. A sociotechnical audit: Assessing police use of facial recognition. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '23, pp. 1334–1346, New York, NY, USA, June 2023. Association for Computing Machinery. ISBN 9798400701924. doi: 10.1145/3593013.3594084. URL https://doi.org/10.1145/3593013.3594084.

Ryan-Mosley, T. How to make government technology better. *MIT Technology Review*, October 2023. ISSN 0040-1692. URL https://www.technologyreview.com/2023/10/23/1082123/how-to-make-government-technology-better/.

A. Appendix

A.1. Interview Questions

The following questions provided the basis for our semi-structured interviews.

Section 1: Contextualizing the interviewee's technical background

1. What are the main tasks and responsibilities that characterize your day-to-day work?

2. Can you tell us about the last time you discussed AI or computing systems with a coworker or supervisor?

3. What computational tools does your department use on a day to day basis? What do you use these tools for, and what is your experience working with these tools like?

Section 2: Assessing an individual's use of tools and their impact

1. When you use a tool, how do you know that it has performed as expected? Do these tools improve the quality or quantity of your work?

2. Are there any specific examples of the tool functioning or not functioning as intended? What do you do if this happens?

Section 3: Departmental adoption of new technology

1. Has your department recently adopted any new AI or computational systems? How do you hear about new tools that your department has adopted or may be considering?

2. If a new tool is adopted, is there a formal process for procuring these tools?

3. Once a tool is adopted, is there a formal training process for use of these tools? What risks are emphasized in training (E.g. protecting sensitive information, trustworthiness, equitable use)?

4. Are you in a position to make decisions regarding tech adoption? What factors do you weigh when considering adopting a new tool? Where do you turn to for information regarding a new tool?

Section 4: Siloization

1. Is use of AI tools primarily individual or collaborative? Is there interfacing with other departments when it comes to sharing data, sharing results, or making joint decisions using these tools?

2. Do you know officials in other departments who have adopted AI systems? Have instances of technological adoption or user guidelines in other departments influenced adoption and use within your own department?

A.2. Executive Summary Recommendations

In a longer report internal to the City of Boston, we presented a series of recommendations specific to the city, which aims to cultivate a technological culture of collaboration and continuous learning that maximizing the benefits of grassroots experimentation while mitigating its potential risks.

	1. Problem Definition	2. Market Research	3. Experimentation	4. Advocacy	5. Informal Acquisition	6. Procurement	7. Training
Challenge	Developing a clear and tool-agnostic understanding of individual or team needs *before* searching for solutions.	Identifying candidate solutions from a variety of sources in a fair and thorough manner. Understanding what is out there and what is allowed under city guidelines.	Allowing early adopters to experiment with candidate solutions on their own and gather feedback.	Articulating the benefits and drawbacks of a tool in a clear, accessible way and providing support to peers of all backgrounds.	Moving from an individual trial to a properly vetted tool that can be used by the whole team without requiring procurement. Providing a decision-making framework for tools that have gained enough traction to warrant proper vetting.	Formally procuring a tool for official use in a rigorous but straightforward process for all parties involved.	Onboarding the tool with training options that are both accessible and useful for all team members.
Recommendation	Teams should meet to predefine **success metrics** prior to exploring the current market for tools.	When beginning market research, teams should **meet with a DoIT liaison** to address questions such as: Does DoIT have a tool that already fills this need? If not, what tools are available?	Require individuals experimenting with tools to generate **one-pager briefs to document** experimentation. Based on previous iterations from the BTD and DoIT.	Provide support for tech advocates, such as **formal recognition** or official programming for their work providing individualized help to peers.	Ask advocates to **present results of experimentation** to designated decision-makers and colleagues (similar to Demo Days within DoIT). Teams or departments should meet to evaluate a new tool according to pre-adoption guidelines set during Problem Definition.	Consider creating a **separate tier** of adoption guidelines for low-cost tools. Designate **procurement experts** within teams to serve as a "single point of contact" between DoIT liaisons and employees	Ask advocates and early adopters to lead onboarding sessions in **small cohorts** and identify employees most in need of training (modeled after Case Study #5).