

# THE CAPACITY DIVIDEND

How Small Local Governments  
Can Use Artificial Intelligence  
Without Surrendering  
Public Authority.



813 STRATEGIES  
AUGUST 2026



## Executive Summary

Artificial intelligence matters more to small local government than it does to organizations with abundant staff. Many local governments are constrained less by a lack of policy ideas than by the number of people available to research them, write them, explain them, process applications, answer residents, maintain records and move routine administrative work through the institution.

Those constraints are measurable. Among state and local governments that recruited for particular occupations in the 2025 MissionSquare workforce survey, 71 percent reported difficulty filling engineering positions, 51 percent reported difficulty filling permitting and inspection positions, and 40 percent reported difficulty filling both planning and information-technology positions. The United States contains more than 90,000 local governments, making it unrealistic to assume that every jurisdiction can respond to artificial intelligence by creating a specialized technology bureaucracy.

AI offers a genuine capacity dividend because useful administrative applications no longer require a city or county to build its own software system. Writing, summarization, document comparison, records search, routine research, translation, meeting support and resident-information retrieval can increasingly be performed through general-purpose systems or through AI already embedded in software governments use.

The evidence now supports measured adoption. Pennsylvania's initial enterprise pilot involved 175 employees across 14 agencies, and participants reported saving an average of 95 minutes on days they used the system; controlled public-sector research has separately shown that AI can materially improve some document tasks while degrading performance on others. The relevant question is therefore not whether a model is impressive, but whether it improves a specific government task after correction time and error risk are counted.

Adoption is also outrunning formal governance. A 2026 review of California's counties and 25 largest cities found published AI policies in only 12 counties and six cities, while the government-led GovAI Coalition now includes more than 900 public agencies. The institutional question has moved beyond whether public employees will encounter AI. It is whether government determines the terms on which they use it.

The correct model for a small jurisdiction is broad assistance and narrow authority. Employees should have relatively easy access to approved tools for low-consequence work they already understand, while controls become stricter when AI receives nonpublic information, connects to internal systems, communicates directly with residents, influences consequential decisions or acquires the ability to act on government's behalf.

That operating model must also account for records law, labor rules and recurring cost. Prompts, chat histories and agent logs can become government records; workforce changes can trigger bargaining or classification issues in jurisdictions with collective-bargaining or civil-service systems; and recurring software or usage charges have to survive an annual budget after the pilot ends.

The objective is not to automate local government. It is to increase the amount of competent work a limited public workforce can perform while keeping governmental judgment, legal accountability and control of the public record where democratic government requires them to remain.

## I. The Real AI Case for Small Government Is Capacity

The debate over government AI has often started at the wrong scale. Large states and major cities can create AI offices, hire data scientists, negotiate enterprise agreements and devote teams to model evaluation; most American local governments cannot.

The workforce data make the constraint difficult to dismiss as temporary. In the 2025 MissionSquare survey, permitting and inspection positions were difficult to fill for 51 percent of governments recruiting for them, planning positions for 40 percent and information-technology positions for 40 percent. Engineering reached 71 percent, while dispatch, skilled trades, policing and several health occupations also remained persistent recruitment problems.

Generative AI is relevant because much of local-government work is document-intensive, repetitive and reviewable. Agendas, reports, ordinances, manuals, applications, meeting transcripts, correspondence and public questions are exactly the kinds of unstructured material that current systems can process without requiring a jurisdiction to build a bespoke application for every task.

The capacity dividend comes from reducing the amount of skilled staff time consumed by intermediate work while preserving professional review at the point where judgment matters. A government still needs someone capable of determining whether the ordinance says what a draft claims, whether an exception applies, whether a permit packet is complete and whether a final communication should be issued.

Lebanon, New Hampshire provides a useful small-city example. The city reports that an AI-assisted emergency medical services reporting system reduced documentation time from approximately 40 minutes to 10 minutes, a 75 percent reduction in a repetitive administrative burden that still leaves medical personnel responsible for the underlying report.

813 Strategies assesses that small local governments should treat AI primarily as a capacity technology. The initial target should be work that is repetitive, document-intensive and independently checkable, because that is where additional administrative capacity can be created without transferring the underlying governmental responsibility.

## II. Adoption Is Moving Faster Than Formal Government

Local government does not control when a widely available technology becomes part of its workforce. Employees encountered ChatGPT, Claude, Gemini, Copilot and AI features inside ordinary software before most councils or county boards had held a formal discussion about them.

The evidence shows public-sector use moving from novelty toward ordinary work. In Boston's 2025 workforce survey, roughly six in ten respondents said they had tried generative AI and one in four said they were using it weekly at work; the National Conference of State Legislatures reported in 2026 that 56 percent of responding legislative staff were already using generative AI for legislative work. These are not a census of local government, but they are consistent with a broader shift toward employee-level use before institution-wide policy catches up.

Formal governance has moved more slowly. ICMA's 2024 national survey of 635 local-government practitioners found that only 9 percent had an organization-wide AI policy and only 10 percent had assigned personnel to oversee AI efforts. The more recent California assessment found policy adoption still uneven in 2026 even among comparatively large counties and cities.

The result is an institutional gap in which employees can have access to more computational capability than their employer has formally decided how to govern. A blanket prohibition can reduce visible use without eliminating access through personal devices, browsers, search engines or software updates, pushing experimentation outside managed government systems where administrators have less visibility.

The better response is controlled abundance. Give employees a sanctioned place to perform ordinary, low-consequence work, then make the boundaries around information, authority and external action much harder to cross.

Training should reflect that objective. Employees need to know what system they may use, what information may be entered, when the answer must be checked against an original source and which functions remain outside the system's authority; they do not need a technical seminar on neural-network architecture before they can use an approved assistant.

### **III. Adopt by Task, Not by Model**

The strongest empirical evidence for government AI points in the same direction: performance varies substantially with the task. That should end the practice of conducting an "AI pilot" as though a model itself were the unit being tested.

Pennsylvania's initial experiment remains important because of its scale and duration. Over a year, 175 employees from 14 agencies used ChatGPT Enterprise for activities including writing, research and summarization; more than 85 percent reported a positive experience and users reported an average of 95 minutes saved on days they used the system. By April 2026, Pennsylvania said more than 3,000 employees in 35 agencies were using generative AI tools.

The 95-minute figure should not be converted into a statewide labor-savings estimate because it comes from employee reporting rather than a controlled time study. It nevertheless establishes that enough users found recurring value for the Commonwealth to move from a 175-person experiment to thousands of employees.

Controlled research provides a sharper lesson. In a pre-registered public-sector experiment, generative AI produced 17 percent higher-quality answers and 34 percent faster completion on a document-understanding task; on a data-analysis task, AI users performed 12 percent worse and received no statistically significant completion-time benefit. A separate 2026 quasi-experimental study in a South Korean government agency found that a domain-specific system reduced average drafting time for citizen-inquiry responses by roughly four minutes, although it did not measure substantive quality.

The common element is not the model. It is a bounded job with an observable output, which means the jurisdiction can compare AI-assisted performance against the prior process rather than rely on user enthusiasm or vendor demonstrations.

#### **A Five-Question Task Screen**

1. **Consequence.** Does the task make, recommend or execute a decision affecting rights, employment, benefits, enforcement, taxes or fees, licensing, final permitting, public safety or access to services? If yes, it should not enter the ordinary low-risk pilot track and must preserve independent human decision-making.
2. **Information.** Does the task require confidential, privileged, protected or otherwise nonpublic information, or broad access to an internal repository? If yes, the use requires an approved managed environment, least-privilege access and a records/security review appropriate to the information involved.
3. **Source.** Can the system be grounded in an authoritative and reasonably bounded source set, or can every material factual claim be checked against a controlling record? If no, the task is a poor candidate for factual government output unless another verification method exists.
4. **Verification.** Can an existing employee detect a material error before the output is used, without reconstructing the entire task from scratch? If no, the apparent time savings are not reliable and the use should be narrowed or rejected.
5. **Reversibility.** Can an error be corrected before it changes a payment, official record, resident status, legal deadline or government action? If no, the task requires a higher level of approval, testing and monitoring than ordinary drafting or research support.

A task that uses public or ordinary internal information, produces a verifiable draft and remains reversible is a strong candidate for broad employee access. A task that introduces sensitive data, public-facing output or system connections belongs in a controlled pilot, while a task that makes consequential decisions or performs difficult-to-reverse actions should remain human-determined unless the jurisdiction deliberately builds a separate legal and technical governance structure around it.

## **IV. Public AI Is an Information Architecture Problem**

Resident-facing AI appears to be one of the most ambitious uses of the technology because the public sees it directly. In practice, it can be one of the more controllable uses if government limits the system to information the government has already decided to publish.

Boston's current AI search illustrates the model. The city grounds answers in Boston.gov, displays the underlying information and tells users that generated answers can contain errors; as of December 2025, Boston reported positive feedback on 34.3 percent of AI-search interactions compared with 10.9 percent for traditional search.

Amarillo built its Emma assistant around a similarly restricted purpose. The city says the system is trained on its website and city-service material, does not answer outside that scope and warns users that responses can be incomplete or inaccurate; during roughly its first three months, residents submitted more than 16,400 questions.

The important comparison comes from a system that did not earn continued deployment. Long Beach tested its Ask Elby assistant with more than 4,000 users, but city officials later reported stale information, incorrect answers and missing links and concluded that it did not materially improve the user experience over ordinary search. The city declined to proceed with the rollout and shifted attention back to the underlying website and search infrastructure.

That failure is more analytically useful than another successful chatbot announcement. It demonstrates that a language model cannot repair an information system whose source material is contradictory, obsolete or poorly owned.

Bellevue's permitting experiment takes the model further while preserving the same basic logic. Its system uses city code and municipal information to guide applicants and assist intake; Bellevue reports that three times as many applications in the pilot arrive complete on first submission, that required documents are correctly identified 96 percent of the time and that the project generated an estimated 152 hours of staff savings in one month. The city also requires AI responses to link to underlying data or code citations and states that human reviewers retain final authority.

The lesson is not that every jurisdiction needs a chatbot. Public-facing AI is strongest when it makes existing government information easier to find, improves the quality of an application before staff review and preserves a clear exit to a person or controlling document when the question becomes consequential.

## **V. Access Changes Risk Faster Than Intelligence Does**

The first generation of generative AI was easy to conceptualize because the employee typed a prompt and received text. The next stage is different: systems can search government repositories, read email, access calendars and databases, invoke software tools and perform actions.

That distinction matters more than the difference between one frontier model and another. A mediocre model with access to an entire document repository can create a larger institutional problem than a much more capable model confined to public information.

The National Institute of Standards and Technology made agent security a dedicated priority in 2026. NIST describes agents as systems capable of operating autonomously for extended periods and interacting with external resources, and its review of public submissions found broad agreement that agents introduce security threats requiring adaptations to conventional cybersecurity practices.

For local government, the immediate risk is excessive permission. An employee may legitimately have access to hundreds of folders accumulated over a career, shared mailboxes inherited from a predecessor or files whose access controls were never revisited after an organizational change; allowing AI to retrieve across those sources at machine speed can turn years of poor information hygiene into an active exposure problem.

The same issue appears when AI can act. A system that drafts a proposed email is an assistant, while a system authorized to send the email has acquired operational agency; a system that identifies a potentially incomplete application is assisting staff, while a system that automatically changes the application's status in the official record is participating in government action.

Those transitions should require explicit approval rather than appear automatically when a software vendor releases a new feature. The relevant questions become what systems the AI can reach, what credentials it uses, what actions it can take, what gets logged and what happens when an action is wrong.

The governing principle should be minimum useful access. If a task can be performed from a defined folder of current ordinances, there is no reason to give the system every document the employee can see.

## **VI. Public Records Follow the Work**

Artificial intelligence creates a records-management problem before it creates a novel legal category. Federal Freedom of Information Act rules generally do not govern cities and counties; state public-records, open-records and sunshine laws do, and those statutes differ substantially in exemptions, response procedures and retention requirements.

The safest operating assumption is that an AI interaction can become a government record when it is created, received, used or retained in connection with official business. Texas law, for example, defines public information to include electronic communications connected to official business and information maintained for a governmental body when the body owns it or has a right of access; the California Supreme Court has likewise held that communications about public business do not lose their public-record character merely because they reside in a personal account rather than a government system.

That logic matters directly to AI. Prompts, uploaded files, generated outputs, chat histories, agent execution logs, tool-call logs and intermediate drafts can all become responsive material depending on what they contain, how they were used and the law of the jurisdiction. A vendor-hosted conversation is not automatically outside the records law simply because the government does not store the database on its own server.

Record status, retention and disclosure are separate questions. A document can be a public record while an exemption protects it from disclosure, and a record can be subject to a short retention period even though it would have been disclosable while it existed; conversely, a final decision record, substantive analysis or communication that becomes part of an official transaction may need to be retained with the underlying record series.

Retention should follow function rather than interface. California records-management guidance makes the point using email: email is a format, not a record series, and retention should depend on the information and purpose of the message. The same approach is appropriate for AI interactions, because a chat used as transitory drafting scratch is different from a conversation that documents why a permit, expenditure or personnel action was taken.

The records officer and counsel should therefore decide, before deployment, which AI artifacts the government intends to preserve, where they will be stored and how normal retention schedules apply. Contracts and administrative settings should support search, export, legal holds and deletion according to government rules rather than allow a vendor's default history or auto-deletion settings to determine the public record by accident.

This is also a design reason to prefer managed government accounts over consumer tools. Central administration cannot eliminate disclosure obligations, but it gives the jurisdiction a realistic way to locate responsive material, apply retention rules and separate official work from personal use when a request arrives.

## **VII. Governmental Judgment Is the Boundary**

AI governance becomes unnecessarily complicated when every conceivable use is treated as a separate ethical category. The more durable distinction is whether the output is helping an employee perform work or becoming a principal basis for how government exercises authority over a person.

The federal government now uses a similar test. Office of Management and Budget guidance treats an AI use as high-impact when its output serves as a principal basis for decisions or actions with a legal, material, binding or significant effect on rights or safety, and expressly notes that a system can meet that definition even when a human remains nominally involved. The federal rule does not bind local governments, but the distinction is useful.

A human signature is not the same thing as human judgment. If software produces the recommendation, frames the relevant evidence, assigns the score and presents approval as the default, an employee who routinely clicks accept may remain legally responsible while exercising very little independent judgment.

Small governments should avoid that structure in the first generation of adoption. Benefits eligibility, hiring and discipline, policing and enforcement, tax assessments, final licensing decisions, final permit decisions and determinations affecting access to public services involve authority that government should not transfer merely because automation becomes technically feasible.

That does not make AI irrelevant to those departments. It means the system should work around the decision by locating provisions, comparing materials against a checklist, organizing evidence, summarizing correspondence or preparing a draft explanation, while the accountable employee remains able to reach and defend the conclusion from the underlying record.

The burden should rise with consequence. Government does not need the same review protocol for an internal meeting summary that it needs for a system influencing whether a family receives assistance, and proportionality is essential if AI governance is to preserve rather than consume administrative capacity.

## **VIII. AI Adoption Is a Workforce Change**

AI deployment in a municipal workforce is not only a technology decision. In jurisdictions with collective bargaining, civil-service systems or formal classification plans, a tool that changes how work is assigned, measured or performed can create labor obligations even when management retains authority to select the technology itself.

The federal National Labor Relations Act generally does not cover employees of states or political subdivisions, so local-government bargaining duties arise from state law, local law and the applicable collective-bargaining agreement rather than from the private-sector federal regime. That makes jurisdiction-specific labor review important before an AI pilot changes staffing assumptions, workloads, job duties, monitoring, performance standards or discipline.

California provides a useful technology analogy. In a Santa Clara County case involving a surveillance-technology ordinance, the Public Employment Relations Board concluded that the county did not have to bargain over the decision to adopt the ordinance itself, but did have to bargain over effects on employee workload and safety and over consequences for represented employees accused of misusing the technology. The rule is not universal, but the distinction between management's technology decision and bargainable effects is directly relevant to AI implementation.

Civil-service classifications create a related issue. A drafting assistant that reduces time spent formatting correspondence does not necessarily change an employee's classification, but a system that permanently transfers inspection, intake, case-screening or analytical duties across positions can require human-resources review under the jurisdiction's classification rules and may affect promotional or qualification standards.

The practical rollout strategy should therefore treat AI as employee support rather than an undeclared workforce-reduction program. Management should identify the task being changed, state who remains accountable for the work, avoid using unvalidated AI metrics as employee-performance measures, and engage bargaining representatives when state law or the agreement makes the effects negotiable.

That approach is not merely defensive. Employees are more likely to surface errors, weak workflows and useful applications when the institution makes clear that the pilot is testing whether a tool improves their work rather than testing whether the employee can be removed from it.

## **IX. Procurement and Budget Are Policy**

A local government can write an excellent AI policy and surrender most of its practical control in the contract. For the majority of jurisdictions that will buy rather than build AI, procurement determines what data the vendor receives, what changes without notice, what performance information government can obtain and how difficult it becomes to leave.

Federal acquisition policy provides several useful principles even where its requirements do not apply directly. Current guidance tells agencies to address government-data use, ongoing evaluation, documentation and vendor lock-in, including contractual protection against using nonpublic government inputs and outputs to train publicly or commercially available systems without explicit consent.

Small governments have more leverage here than they often assume because they no longer have to invent every term independently. The GovAI Coalition, created by San Jose and now encompassing more than 900 public agencies, has published adaptable policy, procurement and vendor-disclosure materials aligned with the NIST AI Risk Management Framework.

Shared templates reduce duplication, but a vendor questionnaire is not proof of system quality. Government still needs observable performance tied to the actual public problem: a permitting product should be evaluated on application completeness, correction rates or review time, while a resident-information system should be evaluated on answer accuracy, successful routing, escalation and user experience.

### **Budget structure matters after the pilot**

The two dominant cost models create different budget pressures. Per-seat products convert AI into a predictable recurring software expense, while application programming interface and agent services charge primarily for usage and can align cost more closely with transaction volume; the latter can be inexpensive at the model level but require technical integration, identity management, logging, monitoring and support that a seat license largely bundles into the product.

Current public pricing illustrates the difference without establishing a preferred vendor. As of August 2026, ChatGPT Business lists standard seats at \$20 per user per month when billed annually and \$25 when billed monthly, while ChatGPT Enterprise uses custom pricing; Microsoft 365 Copilot Business lists a regular annual price of \$21 per user per month for up to 300 users and requires an underlying qualifying Microsoft 365 plan. At those list rates, 25 seats represent roughly \$6,000 to \$6,300 per year before any additional licenses, premium usage, implementation or support.

Usage-based systems can produce a lower direct inference bill for narrow workflows, but they are less predictable and should not be budgeted from a demonstration month's token cost alone. A local government needs a monthly or annual ceiling, logging by department or use case, and enough margin for growth without allowing an agent or integration error to create an uncontrolled variable expense.

The budget test should be the same as the operational test: does the recurring cost buy enough measurable capacity to justify a permanent line item? Pilot money can prove a concept, but it should not hide the full-year cost of licenses, integration, records storage, security controls, staff time and contract administration that arrives after the experiment becomes routine.

Procurement should also preserve exit. The jurisdiction should retain its source documents, logs and evaluation data, be able to export or delete government information, and avoid workflows that become so dependent on one provider's proprietary structure that switching vendors effectively means rebuilding the administrative process.

## **X. Small Government Needs Less AI Governance Than It Thinks**

The emerging public-sector governance architecture can look intimidating because the most visible examples come from large governments. Small jurisdictions should not copy the organizational chart of California, New York City or the federal government; they should copy the controls that matter.

The governing policy can remain short because its core principles change slowly. Employees remain accountable for government work; approved systems must be used for nonpublic information; consequential governmental decisions require independent human judgment; access to internal repositories and the ability to perform actions require elevated approval; public-facing systems must preserve access to authoritative sources and human service; records obligations apply to AI work just as they apply to other government work; and procurement must preserve government control over data, cost and exit.

The operational layer should change much faster. Product names, account types, approved functions, connectors, retention settings, training material and system configurations should be maintained administratively rather than embedded in legislation every time a model changes.

Transparency should attach to systems that matter institutionally rather than every employee interaction with an assistant. Technology that communicates directly with residents, uses resident information, screens applications, performs government actions or materially influences consequential decisions should have a named owner and enough public visibility for the jurisdiction to explain what the system does and who is responsible for it.

The California evidence shows how far even larger local governments remain from that basic visibility. Of the counties and 25 largest cities examined in the 2026 assessment, only four jurisdictions had publicly inventoried their AI uses.

An inventory does not make an AI system safe, but it forces the institution to answer useful questions: what are we using, who owns it, what information does it use, does it affect residents, what does it cost and is it still worth operating? Those questions are more valuable to a small jurisdiction than creating a ceremonial AI office with no authority over actual tools or contracts.

One employee should be responsible for coordinating the answers, but being responsible for "AI" need not mean a new position. In many jurisdictions the administrator, information-technology director, clerk or another senior official can coordinate the process and bring in counsel, records, procurement, finance, human resources or the relevant department when a specific use reaches their responsibilities.

The permanent governance structure should therefore be deliberately small. The organization needs enough central authority to stop uncontrolled adoption and elevate consequential uses, but not so much central process that an employee needs a committee meeting to summarize a public agenda.

## Conclusion

The first phase of public-sector AI produced a predictable choice between enthusiasm and fear. The evidence available in 2026 supports a more useful conclusion: artificial intelligence can increase the effective capacity of government, but the gain is highly dependent on the task and the institutional boundaries placed around the technology.

That matters disproportionately to small jurisdictions. Governments struggling to recruit planners, inspectors, engineers and information-technology employees cannot afford to devote skilled human time to every intermediate administrative step simply because that is how the work was performed before capable language models existed.

The technology has demonstrated value in the kinds of work that fill public offices: drafting, document understanding, information retrieval, application preparation and repetitive reporting. It has also demonstrated failure, including worse performance on some analytical work and resident-facing systems that produced stale or inaccurate information.

The proper response is discrimination between uses. Low-consequence assistance should be broadly available inside approved systems, internal-data access should be narrower, agents capable of acting should face another threshold and governmental authority should remain the hardest line to cross.

The operating model also has to survive contact with the institutions around the technology. Records law determines what government must preserve and disclose, labor and civil-service rules can constrain how work is reorganized, and recurring software or usage costs determine whether a promising pilot becomes a sustainable service.

Small local governments do not need to become technology organizations to do this well. They need a short policy, a known owner, controlled access, a task-level evaluation screen, sensible procurement, workable records rules and the institutional confidence to expand what works and kill what does not.

The jurisdictions that gain the most from AI will not necessarily be those that purchase the most sophisticated systems. They will be the ones that understand exactly which scarce human capacity they are trying to recover, measure whether the technology actually recovers it, and refuse to confuse administrative efficiency with governmental judgment.

That is the capacity dividend. It allows a small government to do more with the people it has without allowing the machine to become the government.