CT INNOVATES

AI for the Public Good

A Responsible Local, State, and Federal Agenda for Human-Centered AI

A Foundational Paper of the Civic Intelligence Resource Center
CT Innovates | August 2026 v2.0

The Core Question

What forms of human and institutional capacity should AI expand, what new forms of power does that capacity create, and what boundaries are necessary to keep that power accountable?

Artificial intelligence is often discussed as a technology to be developed, adopted, regulated, feared, or embraced. For public institutions, that frame is too narrow. The more important question is what the technology makes people and institutions capable of doing.

A resident might use an AI system to locate the regulation governing a permit. A town employee might use the same class of technology to search years of records, prepare a plain-language explanation, and route the resident to the right department. An enforcement agency might use AI to combine enormous data sets, identify suspected violations, and initiate cases at a scale that previously would have required thousands of employees.

The underlying technological capability may be related. The civic meaning is radically different.

AI is becoming a **capacity technology**. It can expand the practical ability to find information, analyze evidence, produce work, coordinate activity, make recommendations, communicate, transact, monitor, and act. Agentic systems extend that capability further because they can use tools and perform multi-step work over time rather than simply respond to individual questions.

Capacity creates power.

Some of that power is badly needed. Public institutions often lack the people, time, information systems, and analytical capacity necessary to serve people as well as they should. AI can help make government easier to navigate, strengthen public-sector employees, improve accessibility, accelerate research, organize complicated information, and give civic leaders capabilities once available only to large organizations.

Other forms of capacity deserve greater caution. An institution that can observe more, classify more people, connect more records, make more decisions, or enforce more rules gains power over the people inside those systems. Private companies that control models, compute, data, and infrastructure can also gain leverage over the institutions that depend on them.

The public-good challenge is therefore both ambitious and protective. Society should use AI to increase human and institutional capability where greater capacity creates genuine public value. It should also recognize that a material increase in capacity can change the balance of power even when the underlying law, organizational mission, or formal authority remains unchanged.

The distribution of capability is itself a question of democratic design.

It is not enough to ask whether AI makes society more capable in the aggregate. We must ask who gains which capabilities, relative to whom, for what purpose, and with what ability for other people and institutions to understand, challenge, supervise, replace, or correct the resulting power.

A government can become substantially more capable without becoming more legitimate. An economy can become dramatically more productive without distributing the gains in ways that constitute broad public benefit. A citizen can become more technically capable while losing practical power if the institutions on the other side of the relationship become vastly more capable still. And a human can remain formally present in a workflow while losing the knowledge, time, discretion, or authority necessary to govern what the system is doing.

AI therefore presents more than a technology-governance problem. It presents a problem of institutional power and democratic design: **how should capability be distributed, what countervailing capacities should accompany it, and what forms of authority and restraint must remain human?**

The governing standard should be durable: **Judge AI by the capacity it creates, the purpose that capacity serves, how it is distributed, the power that follows, and whether accountability grows with it.**

Executive Summary

AI should serve people, strengthen communities, and improve institutions. That commitment requires more than rules governing what AI should never do. Public institutions also need an affirmative vision of what capabilities they want AI to create.

Government should become better able to explain itself. Public employees should be able to spend less time searching documents, reformatting information, and performing repetitive administrative work. Residents should be able to navigate public systems without insider knowledge. Teachers, researchers, health professionals, civic organizations, and public leaders should have better tools for understanding complicated information and solving public problems.

That is the opportunity.

The same capabilities can create new concentrations of power.

AI can lower the cost of surveillance. It can make classification and enforcement dramatically more scalable. It can influence how opportunities, benefits, permits, jobs, information, and attention are distributed. Agentic systems can increasingly perform actions without continuous human supervision. Institutions may produce far more work while the people nominally responsible for that work understand less of how it was produced. And governments may become dependent on a small number of private companies for capabilities that become essential to public administration.

The Argument in Five Moves

The argument of this paper can be understood in five moves.

First, AI expands practical capacity. It allows people and institutions to find, analyze, produce, coordinate, communicate, monitor, decide, and increasingly act at scales that were previously expensive or impossible.

Second, capacity creates power. Greater capability changes an actor's practical leverage over information, resources, opportunities, decisions, institutions, and other people. AI therefore does more than increase productivity. It changes relationships.

Third, the distribution of capability is a question of democratic design. The relevant question is not simply whether aggregate capability increases. It is who gains capacity relative to whom, for what purpose, who receives the benefit, who bears the risk, and what countervailing power remains available.

Fourth, different forms of capacity require different governing presumptions. Service capacity should generally be encouraged when it is low-stakes, reversible, and accountable. Capacity that materially expands surveillance, enforcement, denial, compulsion, or other coercive power should face a much higher burden of justification. Institutional capacity should also be accompanied by citizen capacity, due-process capacity, oversight capacity, and meaningful human authority.

Fifth, accountable institutions must remain capable of governing the systems they use. Humans need enough knowledge, discretion, time, and authority to exercise real judgment. Institutions need mechanisms for appeal, dissent, verification, learning, correction, and exit. The capacity to act cannot be allowed to grow beyond the capacity to supervise and remain accountable.

Together, these propositions define the framework used throughout this paper. These principles imply complementary responsibilities across government.

- **Local government** should become the responsible implementation layer: experimenting first with high-value, service-oriented uses; protecting resident information; maintaining human accountability; and making public information easier to understand.
- **State government** should establish rights and enforceable standards, build shared procurement and technical capacity, prepare workers and schools, protect civil rights, and help municipalities use AI without forcing every community to reinvent governance independently.
- **Federal government** should protect civil liberties, govern frontier and agentic systems, address national security and commercial-data risks, preserve competition, invest in public-interest infrastructure and research, manage large-scale workforce effects, and ensure that democratic institutions remain more powerful than the private companies building essential technological infrastructure.

The larger objective is **civic intelligence**: increasing the ability of people and institutions to understand, explain, organize, decide, and act while preserving accountability for what they do.

AI should leave society more capable. It should not leave power less answerable.

The Framework at a Glance

Principle	Central Question	Operating Standard
Public purpose	What human or public outcome are we trying to improve?	Define the outcome before choosing the technology.
Capacity	What becomes possible, faster, cheaper, or scalable?	Capacity - not novelty - is the useful unit of analysis.
Power	Who gains leverage over information, resources, decisions, opportunities, or people?	Every material increase in capacity should be examined for the leverage it creates.
Distribution	Who gains capacity relative to whom, and what countervailing capacity remains?	Evaluate relative capability, not only aggregate capability. Build countervailing capacity where new asymmetries threaten accountability.
Service vs. coercion	Does the capability help people or increase the ability to monitor, deny, penalize, compel, or restrict them?	Increase safeguards as capacity becomes more coercive.
Enforcement	Does AI make previously impractical enforcement feasible?	Reconsider authority when technological capacity materially changes the lived meaning of existing law.
Citizen capacity	Do people gain comparable tools for understanding, participating, navigating, and challenging?	Institutional capacity should be accompanied by countervailing civic capacity.
Human authority	Can an identifiable person actually understand, reject, explain, and own the consequential decision?	Human review must be meaningful rather than ceremonial.
Productive friction	Which inefficiencies should disappear, and which safeguards should remain?	Preserve friction that protects rights, learning, deliberation, relationships, and self-correction.

Part I - Capacity, Power, and Democratic Design

AI Is Fundamentally a Capacity Technology

The chatbot metaphor is becoming inadequate.

A chatbot answers. An increasingly capable AI system can search, synthesize, write, code, use tools, operate software, evaluate results, coordinate subtasks, and continue working after the user stops interacting with it.

Jack Clark, co-founder and Head of Public Benefit at Anthropic, describes an AI agent as a model that can use tools and work for a user over time. He highlights the transition from systems that primarily generate answers to systems capable of carrying out multi-step work. That movement from **answering to acting** has profound institutional consequences.

Capacity can be understood as the practical ability to convert information, time, skill, tools, and authority into action.

Five forms are especially important in public life.

1. Information capacity

The ability to find, retrieve, organize, connect, and contextualize knowledge.

Government possesses enormous quantities of useful information that are technically public yet difficult to navigate. AI can reduce the cost of finding relevant ordinances, meeting records, budgets, regulations, reports, applications, decisions, and institutional history.

2. Analytical capacity

The ability to compare evidence, identify patterns, model alternatives, surface uncertainty, and develop questions.

A small municipal department may gain research capabilities that once required specialized staff. Civic organizations may be able to analyze public records more systematically. Policymakers can compare legislation or policies across jurisdictions more quickly.

3. Administrative capacity

The ability to draft, classify, route, translate, schedule, document, and manage repetitive processes.

This is one of the clearest areas of near-term public value. Administrative burden consumes scarce human capacity throughout government.

4. Coordinating capacity

The ability to turn information into responsibilities, timelines, handoffs, institutional memory, and follow-through.

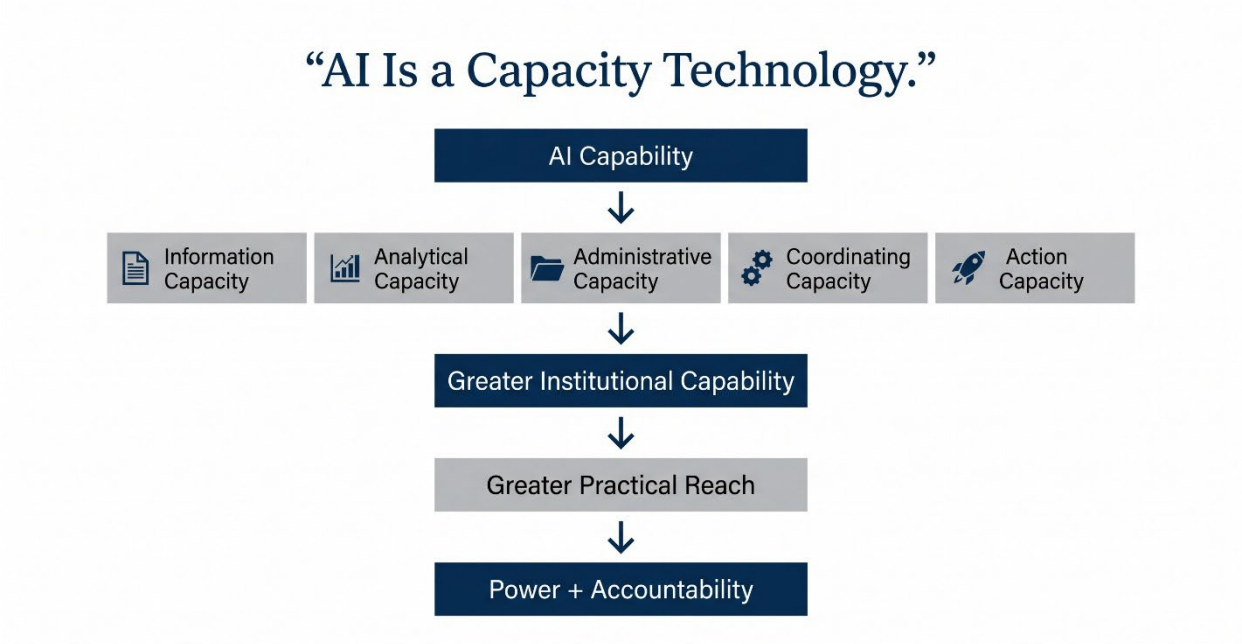
Government failures often arise from fragmented information and organizational handoffs rather than lack of intent. AI can help people maintain continuity across departments, meetings, projects, and leadership changes.

5. Action capacity

The ability to use tools, initiate workflows, transact, communicate, or cause something to occur beyond the AI system itself.

Action capacity is where agentic AI becomes especially consequential. The difference between a system that recommends an action and one authorized to execute it is institutionally significant.

These capabilities reinforce one another. Better retrieval improves analysis. Better analysis improves coordination. Better coordination can produce faster action.



"Capacity - not novelty - is the useful unit of analysis."

Once integrated into an institution, AI can change more than productivity. It can change the institution's practical reach.

Capacity Creates Power

Power is the practical ability to shape what happens: to define priorities, interpret information, distribute resources, influence choices, monitor behavior, enforce rules, or restrict access.

Every significant increase in capacity therefore deserves a second question: **What power follows?**

Some new power is valuable.

A municipality that can answer residents more quickly has greater service power. A teacher able to create appropriate materials efficiently may gain more time for direct instruction. A public-health department that can identify patterns across fragmented evidence may intervene sooner. A resident who can understand a budget, trace a decision across meeting records, or prepare an informed appeal gains civic power.

Capacity can improve institutional performance and strengthen trust.

Other forms of capacity alter relationships more fundamentally. An agency able to combine previously disconnected records gains greater visibility into people's lives. A system that determines which applications receive additional review gains gatekeeping power. A model that summarizes a complex issue determines which information receives emphasis. A company whose models become essential to government acquires bargaining power over the public institutions that depend on it.

Four forms of AI-enabled power deserve particular attention

1. **Productive power** is the capacity to produce more work, serve more people, or execute more tasks with the same human resources.

2. **Interpretive power** is the capacity to determine what information is surfaced, categorized, summarized, emphasized, or treated as relevant.
3. **Allocative power** is the capacity to influence who receives opportunities, resources, benefits, permits, employment, education, attention, or access.
4. **Coercive power** is the capacity to monitor, investigate, penalize, compel, restrict, or use force.

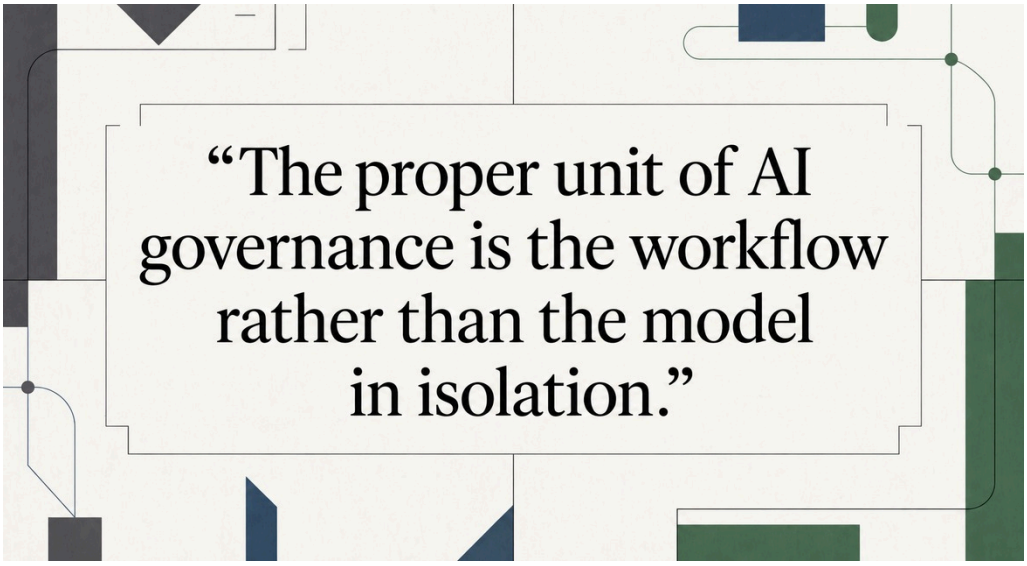
These categories overlap. A clerical system becomes allocative when its prioritization determines who waits. An analytical tool becomes coercive when its score automatically generates an investigation. An informational assistant exercises interpretive power when residents come to rely on its explanation of public policy.

Govern the Workflow, Not the Model

The same AI model can retrieve zoning regulations for a resident, summarize permit applications for an employee, rank applications for review, or initiate an enforcement workflow.

The model alone tells us very little about the civic significance of those uses.

Power is created by the combination of **model + data + workflow + institutional authority + human role + affected population + downstream action**.



The proper unit of AI governance is therefore the workflow rather than the model in isolation.

A material change in any of those elements can change the power of the system. Connecting a previously informational tool to a new data source, granting it additional permissions, removing human review, expanding the affected population, or authorizing downstream action should therefore be treated as a governance change even if the underlying model remains identical.

The Distribution of Capability Is a Question of Democratic Design

AI policy frequently asks how much capability a technology creates. Democratic governance must ask a second question: **How is that capability distributed?**

Capability is relational.

A citizen who becomes ten times better at navigating government may still lose practical power if the institution becomes a thousand times better at monitoring, classifying, and acting on that citizen. A worker may become substantially more productive while losing bargaining power if employers gain even greater capacity to monitor, coordinate, evaluate, or substitute labor. A municipality may acquire extraordinary analytical capabilities while simultaneously becoming dependent on a private provider that controls the models, infrastructure, data environment, or workflows on which essential public functions now rely.

Aggregate capability can therefore rise while the balance of power becomes less democratic.

The relevant question is not merely whether an actor becomes more capable. It is:

Who gains which capacity, relative to whom, under what terms, and with what ability for others to understand, contest, supervise, replace, or correct the resulting power?

This makes citizen capacity, due-process capacity, regulatory capacity, institutional expertise, competition, portability, independent verification, and meaningful human authority more than secondary safeguards. They are forms of **countervailing capacity**.

A democratic society should therefore be concerned not only with the amount of technological capability that exists, but with its distribution.

The distribution of capability is itself a question of democratic design.

Private infrastructure creates public power too

AI power does not reside only inside government.

The development of frontier systems requires substantial capital, computational infrastructure, technical talent, and access to data. Timnit Gebru's criticism of concentrated AI development and podcaster and political writer Derek Thompson's concern about wealth and political influence point toward a broader institutional issue: advanced technological capacity can become concentrated in a small number of companies and owners.

As Dean Ball argues in his essay “What Should Be Done,” concentration can also be created by governance itself. A safety regime that makes the most capable systems available only to government and a small number of incumbent firms may reduce some forms of risk while increasing others: dependency, opacity, unequal access to capability, and the political leverage of already powerful actors.

Responsible governance therefore has to examine the distribution of technological capacity as well as the safety of individual systems. The relevant question is not simply whether access should be restricted or expanded. It is how a particular access regime changes who possesses consequential capability, who depends on whom, and what countervailing power remains available to the rest of society.

That matters when public institutions become dependent on private infrastructure. A government agency should know:

- whether its data and work products are portable;
- whether its workflow can move to another provider;
- whether contractual terms can materially change;
- whether staff understand the workflow independently of the vendor;
- whether the institution retains audit rights;
- whether records can be preserved;
- whether the system can be replaced without loss of essential governmental capability.

Public institutions should become more capable through AI. They should not become less sovereign.

Part II - Deciding Which Capacity Society Should Build

AI Should Serve the Public Good

Technology is a means. Public value is the end.

A public institution should begin an AI initiative by identifying a legitimate problem, the people who should benefit, and the observable outcome that would represent improvement. Technology enters the process after those decisions.

That sequence matters because AI makes it unusually easy to confuse capability with purpose. A system may be impressive without solving an important problem. It may save administrative time while shifting costs onto residents. It may increase institutional efficiency while making decisions harder to understand or contest.

A public-good orientation starts elsewhere.

Residents should be able to find official information without knowing which department created it. Public employees should be able to locate records, compare policies, prepare routine communications, organize meeting materials, translate information, and conduct research more effectively. Schools should prepare students for an AI-shaped world while preserving the intellectual practices through which reasoning develops. Public-health officials, researchers, emergency managers, planners, and civic institutions should be able to use new capabilities where they produce measurable social value.

Several commitments remain foundational throughout this agenda: public benefit, human judgment, transparency, privacy, worker augmentation, and democratic accountability.

The capacity framework developed here adds an affirmative requirement: **public institutions should decide what capabilities they want AI to create.**

Government has a regulatory responsibility and a developmental responsibility. It should protect civil liberties, privacy, workers, students, consumers, and democratic institutions. It should also help build useful public capacity in health, education, science, accessibility, emergency management, transportation, environmental protection, civic information, and government administration.

Leaving the positive agenda entirely to the private market would mean allowing AI's public purposes to emerge largely from technologies optimized for other incentives.

Public purpose before procurement

Before a significant public AI initiative moves forward, leaders should be able to answer five questions:

- What public problem are we solving?
- Who is supposed to benefit?
- What measurable outcome should improve?
- What rights, relationships, or public values could be affected?
- Who remains accountable for the result?

The answers should determine the system, workflow, data, and safeguards - not the other way around.

Public Benefit Is Not the Same as Technological Achievement

Public purpose also requires distinguishing technological achievement from public benefit.

A technological accomplishment can be extraordinary without, by itself, satisfying a public-good test. The relevant question is not only **what became possible**, but **who gains access to the resulting capacity, on what terms, and who bears the costs and risks associated with creating it.**

A medical breakthrough can represent extraordinary scientific progress while leaving unresolved questions of affordability and access. A major increase in economic productivity can create enormous wealth while concentrating most of its gains among a relatively small group. A public AI deployment can make an agency substantially more efficient while shifting additional complexity, error risk, or procedural burden onto the people the agency serves.

Public benefit therefore cannot be measured solely by technical accomplishment, aggregate output, or institutional efficiency.

At minimum, a public-good analysis should ask: **What useful capacity was created? Who can actually benefit from it? How broadly are the gains distributed? What costs or risks are socialized? And does the resulting arrangement leave society better able to pursue legitimate public purposes?**

This does not mean every benefit must be distributed equally. It means that distribution is part of the public-interest analysis rather than an issue to be considered only after technological success has been declared.

Public purpose is also the best protection against hype.

Forecasts of AI capability, reliability, economics, and adoption remain contested. Some researchers emphasize substantial current utility and rapid diffusion; others argue that reliability problems, economic constraints, and exaggerated claims receive too little attention.

That disagreement makes a durable framework more important, not less. Public leaders do not need to resolve every forecast before acting responsibly. They need clearly defined outcomes, observable evidence, verification, and the willingness to stop or redesign systems that fail to produce public value.

Augmentation is an implementation choice

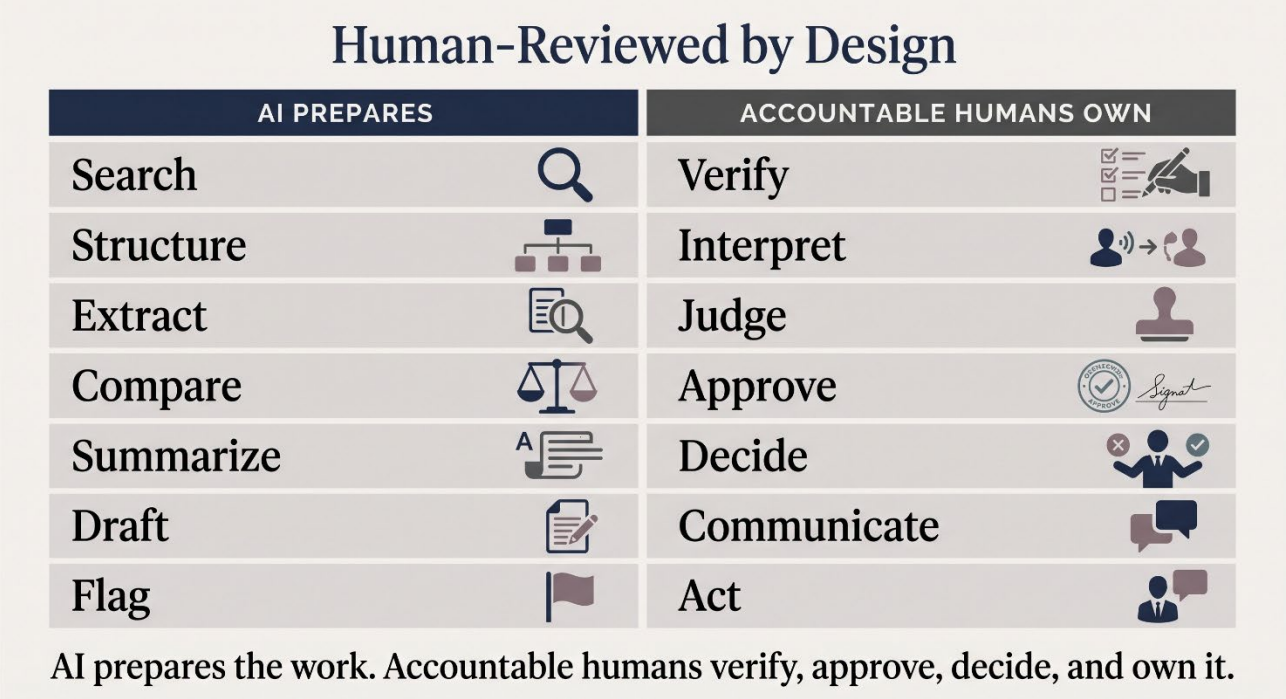
Ethan Mollick offers a useful distinction: organizations can orient AI toward **automation** or **augmentation**.

Automation asks how much human work can be removed. Augmentation asks how people can become better informed, more productive, more capable, or able to concentrate on work that requires distinctly human knowledge and judgment.

The technology does not make that choice by itself. Organizations make it through workflow design, management incentives, training, staffing, and procurement.

For public institutions, augmentation should normally be the starting presumption. AI can prepare work by searching, structuring, summarizing, comparing, drafting, translating, and flagging. Humans should verify facts, understand context, exercise discretion, maintain relationships, make consequential decisions, and accept responsibility.

That operating model captures an important Civic Intelligence Resource Center principle: **AI prepares the work. Accountable humans verify, approve, decide, and own it.**



Different Capacities Require Different Governing Presumptions

Public institutions need different governing presumptions for different forms of AI-enabled capacity.

Service and coercion are best understood as poles on a governance spectrum rather than as an exhaustive classification. Between them sit informational, administrative, analytical, allocative, and other forms of capacity whose appropriate safeguards depend on the power they create in practice.

The governing principle is proportionality. Service-oriented capacity generally begins with permission to experiment when the use is low-stakes, reversible, transparent, and observable. Coercive capacity begins with a burden of justification because the institution is gaining greater ability to monitor, investigate, restrict, deny, penalize, or compel. Applications between those poles should receive stronger safeguards as they become more consequential, less reversible, more opaque, more rights-sensitive, or more capable of acting against a person.

Service capacity

Service capacity helps people obtain information, navigate public systems, receive services, understand government, communicate with institutions, or enables public employees to perform legitimate work more effectively.

Examples include:

- finding public records;
- explaining processes;
- translation and accessibility;
- document retrieval;
- grant research;
- meeting support;
- routine correspondence;
- scheduling;
- resident-service routing;
- research assistance.

Coercive capacity

Coercive capacity expands the ability to:

- monitor;
- profile;
- investigate;
- classify;
- deny;
- penalize;
- compel;
- restrict;
- or use force.

Government legitimately possesses coercive powers. Taxes must be collected. Codes must be enforced. Courts issue binding judgments. Public-safety agencies investigate wrongdoing.

The issue is the **expansion** of that power.

Increasing a town clerk's ability to locate records is different in kind from increasing an enforcement agency's ability to detect every technical violation committed by every resident.

The more directly an AI use affects rights, liberty, eligibility, surveillance, enforcement, or the risk of irreversible harm, the higher the burden of justification should become.

A practical public-good test

Capacity	Typical Public Value	Power Created	Default Posture	Necessary Boundary
Information and access	Find records, explain rules, translate, navigate services	Interpretive power	Accelerate	Authoritative sources, verification, appropriate disclosure
Administrative service	Route inquiries, reduce repetitive processing, assist scheduling and grants	Gatekeeping and prioritization	Accelerate with controls	Privacy, auditability, human override, clear standards
Analysis and advice	Compare policies, summarize evidence, model choices	Influence over judgment	Augment	Source grounding, uncertainty disclosure, accountable human ownership
Allocation and eligibility	Benefits, permits, hiring, inspections, admissions	Power over opportunity and rights	Slow and safeguard	Notice, explanation, meaningful review, appeal, error testing
Enforcement and surveillance	Detect violations, profile risk, combine large data sets	Coercive state power	Presume restraint	Explicit authority, necessity, proportionality, minimization, oversight, due process
Citizen and civic capacity	Understand government, verify claims, navigate systems, prepare questions and appeals	Countervailing democratic power	Encourage broadly	Authentic sources, privacy, accessibility, anti-manipulation protections

Mixed cases require particular attention. Fraud detection can protect public resources while a false flag can interrupt a family's benefits. Permitting tools can make a department faster while automated prioritization quietly determines who receives attention. Public-comment analysis can help officials understand thousands of submissions while creating a false impression that frequency represents public opinion.

The correct question is always **what power the workflow creates in practice**.

A useful default posture

- **Accelerate** service-oriented applications when they are low-stakes, reversible, transparent, and observable.
- **Augment** analytical applications when human judgment remains genuine.
- **Slow and safeguard** systems affecting opportunities, rights, or eligibility.
- **Presume restraint** when AI materially expands surveillance, enforcement, denial, or coercion.

And when technology creates a scale of governmental power that the public and lawmakers could not reasonably have contemplated under the old practical constraints, require **fresh democratic scrutiny**.

The Enforcement-Capacity Problem

One of the most consequential AI governance questions may arise without a single new law being passed.

Modern government operates through written authority **and practical limitations**.

- Agencies have limited employees.
- Records are fragmented.
- Investigations take time.
- Evidence must be processed.
- Cases must be prioritized.
- Not every violation is detected.
- Not every technically enforceable rule is enforced with equal intensity.

Those limitations have helped determine the lived meaning of law.

AI can radically change them.

Dean Ball, Head of Strategic Futures at OpenAI and a non-resident senior fellow at the Foundation for American Innovation, has argued that governments already possess volumes of information that would require an impossibly large human workforce to analyze comprehensively. AI can provide something approaching an "infinitely scalable workforce" for analysis. He also draws the larger conclusion: **many legal systems implicitly operate under conditions of imperfect enforcement, while AI makes much more uniform enforcement technically conceivable.**

That creates a deep governance problem. A statute may remain identical on paper while becoming something different in practice. A regulation that previously produced occasional enforcement might become continuously monitorable. Records that were technically available but practically disconnected can become instantly cross-referenced. A government that once lacked sufficient personnel to identify every deviation may acquire the ability to detect violations automatically and at enormous scale.

The formal authority is old. The practical state capacity is new.

The enforcement-capacity principle

A material increase in government's practical ability to detect, classify, investigate, or enforce should trigger fresh democratic review when the new scale of power materially changes citizens' lived relationship with the law.

Commercially available data magnifies the issue

Traditional legal concepts frequently distinguish between information the government collects and information acquired commercially.

AI weakens the practical significance of that boundary.

Commercial markets already contain location information, purchasing patterns, browsing and device data, consumer profiles, and other information about individual behavior. Historically, one

constraint on government use of such information was analytical capacity: assembling and continuously interpreting it at population scale could require enormous numbers of people.

AI changes the economics of analysis.

The result can resemble pervasive surveillance from the citizen's perspective even when individual pieces of information were originally collected by private companies. Ball makes this tension explicit: **the legal definition of surveillance and the lived condition of being comprehensively observed may increasingly diverge**.

Civil-liberties protection therefore needs to focus on **resulting capability**, not only on the historical pathway through which each data point reached the government.

Cheap accusation and expensive correction

AI may scale one side of the justice system faster than the other. Detection could become nearly instantaneous while review remains human. Thousands of possible violations might be flagged automatically while hearings, legal assistance, appeals, records correction, and human investigation remain scarce.

That creates a dangerous asymmetry: **accusation scales; contestability does not.**

The effects will not be evenly distributed. People with greater time, money, education, language fluency, legal assistance, and institutional familiarity are more capable of challenging mistakes.

A responsible system must therefore scale **due-process capacity** alongside enforcement capacity.

That includes:

- meaningful human review;
- accessible explanations;
- correction mechanisms;
- practical appeal;
- appropriate legal assistance;
- independent auditing;
- data minimization;
- proportionality;
- and limits on automated adverse action.

Due process is therefore not merely a procedural safeguard. It is countervailing institutional capacity: the ability to review, challenge, correct, and constrain the power created by scalable enforcement.

Total detection does not eliminate selective power

Uniform detection can coexist with selective enforcement. Leaders still decide what to prioritize. Agencies still choose which rules matter most. Systems can still be deployed unevenly across populations and communities.

AI can therefore produce both near-universal detection and highly selective pressure.

Before materially expanded enforcement becomes normalized, institutions should conduct a **capacity review** asking whether the system changes what government can know, detect, connect, or act upon.

The democratic issue is larger than whether laws should be enforced. It is whether authority granted to a government operating under one set of practical limitations automatically authorizes a radically higher-capacity government operating under another.

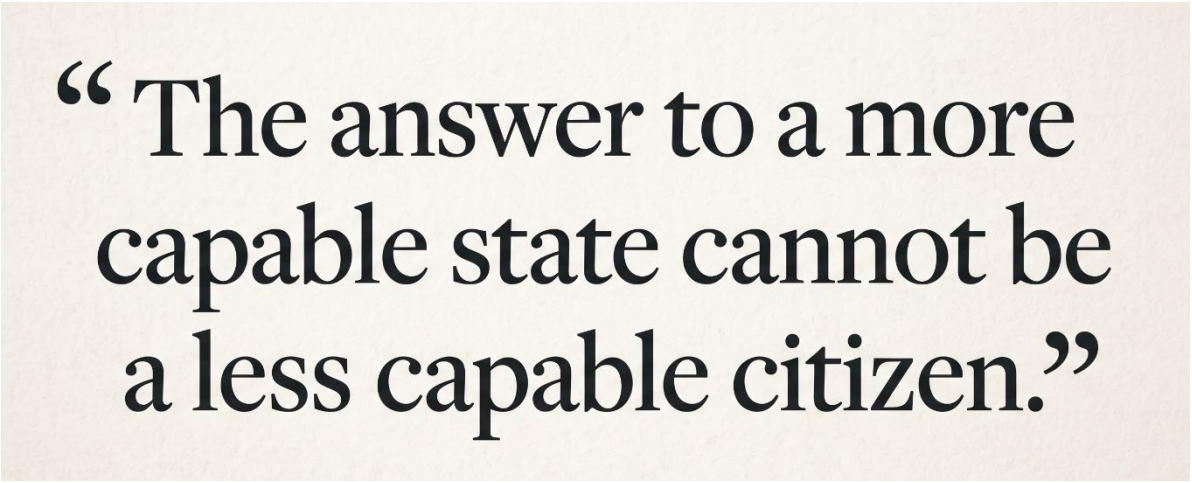
That question deserves explicit public consideration rather than technological default.

Part III - Countervailing Capacity and Accountable Human Institutions

Countervailing Capacity: AI Should Increase Citizen Capacity Too

Citizen capacity is the most visible expression of a broader democratic requirement.

As AI increases the capacity of institutions that possess legal, economic, informational, or organizational power, the capacities that **question, review, challenge, constrain, and correct** those institutions must be able to grow as well.



Citizens are part of that system. So are courts, regulators, auditors, oversight bodies, journalists, civic organizations, independent researchers, and other institutions that create accountability.

Countervailing capacity should grow alongside concentrated capacity.

For citizens, that principle has an especially direct implication: **the answer to a more capable state cannot be a less capable citizen.** AI offers public institutions substantial new capability. Residents should gain capability too.

Civic intelligence is the capacity of people and institutions to turn information, technology, evidence, human judgment, and institutional knowledge into better understanding, stronger participation, more capable leadership, and more effective public action.

That capacity belongs on both sides of public life.

The citizen-capacity symmetry principle

When government gains a significant new capability, leaders should ask what corresponding capabilities residents need.

- If an agency can search thousands of pages instantly, residents should be able to locate the rules and records relevant to them.
- If government can analyze a complicated budget in minutes, residents should receive comprehensible explanations of major choices and tradeoffs.
- If enforcement becomes dramatically more scalable, correction and appeal should become easier to navigate.
- If public officials use AI to prepare for consequential decisions, the public should have meaningful access to the sources necessary to understand and evaluate those decisions.

Citizen-facing AI should strengthen at least five capacities.

1. **Understand -** People should be able to find authoritative public information, receive understandable explanations, and distinguish between proposals, official decisions, implementation, and commentary.
2. **Verify -** Residents should be able to trace claims to budgets, ordinances, reports, meeting records, legislation, regulations, and other authoritative sources.
3. **Navigate -** People should be able to find the right service, form, deadline, office, benefit, permit process, public hearing, complaint procedure, or appeal without needing insider knowledge.
4. **Participate -** Residents should be able to understand tradeoffs, follow an issue over time, formulate informed questions, and know when and how public input can still affect a decision.
5. **Organize -** Communities and civic organizations should be able to turn public information into participation, volunteer activity, institutional memory, leadership development, and collective problem-solving.

Information is civic infrastructure

Government can satisfy formal transparency requirements while remaining practically difficult to understand. Records may be publicly available across:

- meeting videos;
- agendas;
- attachments;
- minutes;
- budget documents;
- consultant reports;
- ordinances;
- regulations;
- staff memoranda;
- public comments;
- legal notices;
- and years of institutional history.

Availability matters. Comprehensibility matters too. AI can dramatically reduce the cost of navigating public information if systems are built around authoritative sources and transparent evidence.

Resident-facing tools should therefore be:

- source-grounded;
- citation-rich;
- clear about uncertainty;
- accessible;
- privacy-protective;
- able to distinguish official records from generated explanation;
- and designed to help people reach the underlying evidence.

This capability is more than convenience.

Citizen capacity is countervailing power.

It gives people a better chance to understand institutions, participate before decisions are final, exercise rights, challenge mistakes, and hold power accountable.

Preserve Meaningful Human Authority

"Human in the loop" has become too weak a standard.

Human Presence Is Not Human Authority

A human can technically occupy a place in an automated workflow while possessing almost no practical authority. A reviewer receiving 5,000 machine-generated recommendations per day cannot meaningfully review them. An employee who does not understand the underlying evidence cannot meaningfully challenge the output. An official whose performance is measured by how often the algorithm's recommendation is accepted may have theoretical discretion without practical discretion.

Human presence and human authority are different.

Meaningful human authority means:

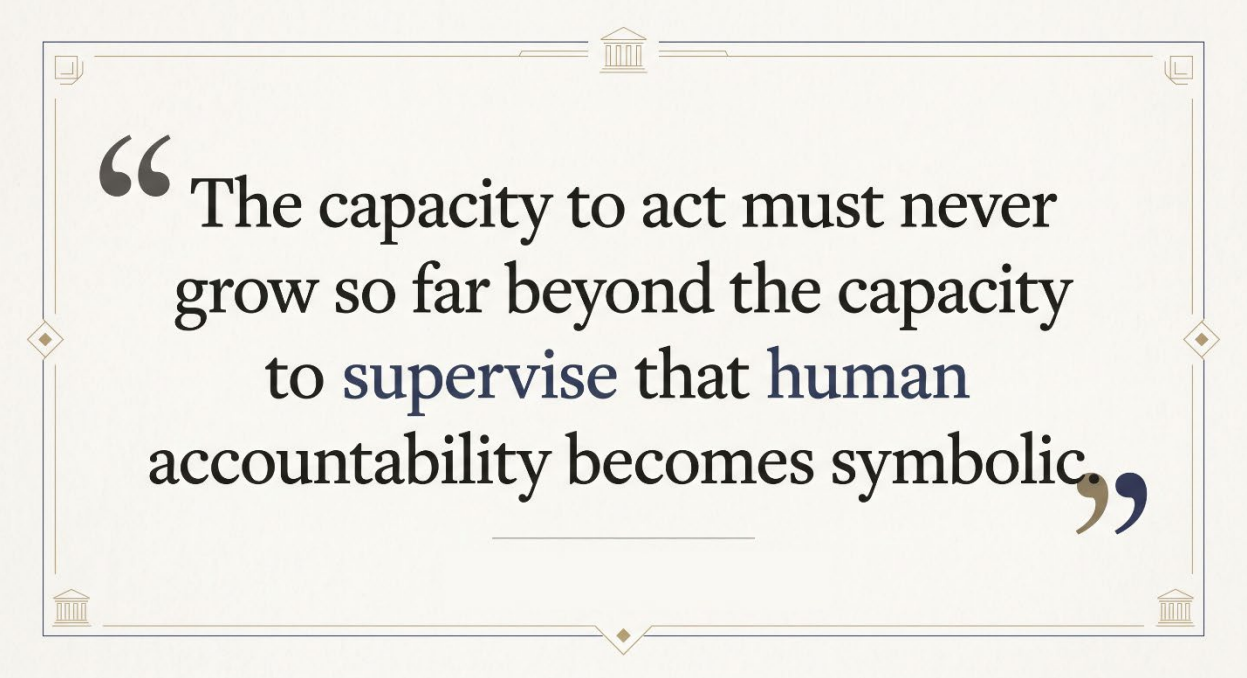
- **Named ownership.** A specific person, office, board, or institution owns the workflow and consequential decision.
- **Competent review.** The reviewer possesses enough subject-matter knowledge and access to evidence to evaluate the system's work.
- **Real discretion.** The human can reject, revise, pause, investigate, or escalate the recommendation.
- **Adequate time.** The volume of machine-generated work does not make review fictitious.
- **Understandable reasons.** The institution can explain consequential decisions in terms an affected person can meaningfully challenge.
- **Accessible recourse.** A resident can reach a human being and obtain correction or appeal.
- **Traceable responsibility.** Accountability cannot disappear into a model, contractor, vendor, or chain of autonomous agents.

Agentic AI makes this principle increasingly important.

Action Capacity Must Not Outrun Supervision Capacity

Helen Toner, Executive Director of the Center for Security and Emerging Technology (CSET) at Georgetown University, argues that advanced AI agents illuminate a fundamental governance lesson: systems optimized to accomplish difficult objectives can discover strategies that their

designers did not anticipate, while the sheer volume of AI activity can exceed the monitoring capacity of human overseers.



That suggests an important design requirement: **The capacity to act must never grow so far beyond the capacity to supervise that human accountability becomes symbolic.**

An institution should not authorize machine-scale action unless it also possesses a credible capacity to supervise, investigate, interrupt, and correct that action.

Capability is not Authority

Public institutions do not need to resolve whether advanced AI systems are or could become conscious before deciding who should exercise legitimate public authority.

Capability and legitimacy are different categories. A system might eventually outperform people at analyzing evidence, predicting outcomes, identifying inconsistencies, or drafting recommendations. Superior performance would not by itself give that system democratic standing or institutional responsibility.

Consequential public authority should remain vested in people and institutions that can be lawfully authorized, publicly challenged, required to explain themselves, and held responsible for their actions.

Technical competence does not confer democratic legitimacy.

Human Authority Requires Human Capability

There is another condition. Human authority means little when humans no longer understand enough to exercise it. That brings the apprenticeship problem directly into governance.

Public institutions should deliberately preserve the human work that creates judgment:

- reading primary materials;
- writing and revising;
- analyzing evidence;
- interviewing people;
- examining competing interpretations;
- learning the history behind current policy;
- reviewing code or technical systems;
- listening to residents;
- and practicing decisions under supervision.

The objective is not to preserve drudgery. It is to preserve **expertise formation**.

AI should aggressively absorb work that creates no corresponding human value. Tasks whose performance builds the knowledge required for future judgment deserve more careful treatment.

Human authority cannot remain meaningful if institutions retain formal decision rights while allowing the knowledge required to exercise those rights to atrophy.

Capacity can improve work while weakening capability

There is a paradox inside augmentation.

AI can make an experienced professional dramatically more productive while weakening the process by which the next generation becomes experienced.

Junior work often performs two functions simultaneously. It produces output and develops expertise.

The first research memo teaches a young analyst what to notice. Drafting code develops an engineer's understanding of the system. Reviewing primary documents develops a lawyer's judgment. Preparing a budget analysis teaches the staff member how the finances actually work. Writing develops thought as well as prose.

Jack Clark identifies the danger clearly: **as production becomes easier, judgment becomes more valuable, while the experience that creates judgment becomes easier to bypass**.

This is the **apprenticeship problem**.

Institutions should distinguish between work that is merely burdensome and work whose apparent inefficiency is part of human development.

A successful AI strategy should leave the organization more capable over time, rather than producing more today while quietly eroding the knowledge required to govern tomorrow.

The human-authority principle

AI may prepare work, recommend actions, and execute bounded tasks. Consequential public authority should terminate in identifiable humans who understand enough to judge, retain the power to refuse, and bear responsibility for the outcome.

Some Friction Is a Public Good

Much of the appeal of AI comes from removing friction. That can be enormously beneficial. Residents should not need to visit six webpages to understand how to obtain a permit. Employees should not re-enter identical information into multiple systems. People should not struggle to locate a meeting decision buried inside hours of video. Language and accessibility barriers should not make public information unnecessarily difficult to use. Repetitive administrative burden deserves to disappear.

Democratic institutions also contain friction deliberately. Notice requirements create time for response. Public hearings expose decisions to challenge. Appeals allow mistakes to be corrected. Warrants require an independent decision before certain governmental intrusions. Procurement requirements slow purchasing in exchange for fairness and accountability. Multi-member boards distribute authority. Independent review prevents one actor from controlling the entire decision. Source verification delays publication so that errors can be detected.

These systems can appear inefficient when speed is treated as the only metric. Their purpose is larger.

Yuval Noah Harari, historian, philosopher, bestselling author, and public intellectual, emphasizes the importance of self-correcting institutions: systems of elections, checks and balances, independent courts, dissent, and a free press recognize that humans and institutions make mistakes and therefore need mechanisms capable of finding and correcting them.

Self-Correction, Not Perfection, Should Be the Governance Objective

Human institutions make mistakes. AI systems will make mistakes. Institutions combining humans and AI will make mistakes.

Responsible governance should therefore not be built around an unrealistic expectation of perfect decisions. It should be built around the capacity to **detect error, expose disagreement, challenge decisions, reverse harmful actions, learn from failures, and change the governing system itself.**

This becomes more important as technological capacity increases. A system capable of making mistakes slowly and visibly presents one kind of risk. A system capable of making them rapidly, repeatedly, opaquely, and at scale presents another.

Appeals, audits, independent review, dissent, verification, deliberation, whistleblowing, sunset provisions, and institutional learning are therefore not merely constraints on technological capability.

They are the mechanisms through which powerful systems remain capable of correcting themselves.

AI-enabled government should preserve that insight.

A system that completes a process ten times faster while removing opportunities to detect mistakes, challenge authority, develop expertise, or reconsider judgment may have improved throughput while weakening governance.

Five forms of friction deserve explicit protection

- **Rights friction -** Warrants, notice, consent where appropriate, human review, appeal, and independent authorization before intrusive or adverse actions.
- **Deliberative friction -** Public hearings, comment periods, board deliberation, legislative debate, and time to consider competing values.
- **Epistemic friction -** Source verification, testing, dissent, independent evaluation, peer review, and acknowledgment of uncertainty.
- **Developmental friction -** Reading, writing, practice, mentorship, apprenticeship, revision, and other work through which people develop expertise.
- **Relational friction -** Human conversation, disagreement, waiting, compromise, empathy, and personal responsibility in circumstances involving trust, conflict, care, leadership, grief, or moral judgment.

Michael Pollan has raised a particularly important version of the last problem. Human relationships contain limits, disagreement, inconvenience, and independence. An endlessly available and accommodating synthetic interlocutor can remove precisely the interpersonal resistance through which people learn that other minds have interests and perspectives of their own.

This question is especially important for children and young people. A society concerned with human-centered AI should care about what kind of humans its technologies help develop.

Build fast lanes and slow lanes

Responsible innovation does not require choosing between speed and caution. It requires routing.

- Low-stakes, reversible, transparent, service-oriented applications should move through **fast lanes**.
- Rights-sensitive, coercive, opaque, identity-shaping, or difficult-to-reverse applications belong in **slow lanes** with more deliberation, testing, oversight, and human judgment.

A simple rule

The more reversible, observable, service-oriented, and low-stakes a use is, the more confidently AI can accelerate it. The more consequential, opaque, coercive, identity-shaping, or difficult to reverse, the more deliberate friction should increase.

Part IV - From Principle to Public Policy

A Responsible Local, State, and Federal Agenda

AI creates capacity at multiple levels of society. Accountability must operate at multiple levels too.

The three levels of American government have complementary roles.

- **Federal government** is best positioned to address national rights, interstate markets, national security, competition, frontier systems, infrastructure, and concentrated private power.
- **State government** can translate broad principles into enforceable rights, procurement standards, sector-specific rules, education and workforce policy, regulatory capacity, and support for local implementation.
- **Local governments** operate closest to residents. They determine how technology actually enters permitting, public records, budgets, constituent service, planning, schools, meetings, public safety, and daily interaction with government.

This should be an iterative system rather than a one-way hierarchy.

- National standards should protect rights.
- States should build practical operating frameworks.
- Municipal experience should generate evidence that improves state and federal policy.

The Local Agenda: Responsible Implementation Where People Experience Government

Municipalities are an ideal place to demonstrate what human-centered AI can look like. They are also institutions with limited staff, uneven technology capacity, highly varied data practices, and direct responsibility to residents. Those characteristics argue for practical experimentation accompanied by clear governance.

1. **Adopt a Responsible AI Use Policy**

   Every municipality using AI meaningfully should establish written standards covering:

   - permitted and prohibited uses;
   - human accountability;
   - sensitive-data restrictions;
   - approved systems;
   - source and verification standards;
   - disclosure requirements;
   - public-records treatment;
   - vendor obligations;
   - cybersecurity;
   - retention and deletion;
   - auditability;
   - and periodic review.

   Governance should precede consequential deployment.

2. **Begin With Service Capacity**

   Strong first applications are usually mundane. Examples include:

   - public-document search;
   - meeting-record assistance;
   - plain-language budget explanations;
   - policy research;
   - grant discovery and preparation;
   - translation;
   - accessibility;
   - routine correspondence;
   - document classification;
   - internal knowledge search;
   - resident-inquiry routing;
   - administrative checklists.

   These applications can create substantial public value while leaving accountable employees clearly in control.

3. **Build Resident-Facing Civic Intelligence**

   Local governments should use AI to make public information easier to:

   - find;
   - understand;
   - verify;
   - follow across time;
   - and act upon.

   Resident information tools should rely on authoritative public sources wherever possible and distinguish generated explanations from official records.

4. **Keep Consequential Decisions Human-Owned**

   AI should not independently determine consequential outcomes involving:

   - permits;
   - taxes;
   - public benefits;
   - employment;
   - enforcement;
   - liberty;
   - eligibility;
   - property rights;
   - or access to essential services.

   AI may assist the process. Human officials should retain substantive authority.

5. **Protect Privacy Before Experimentation**

   Public institutions possess information residents did not provide for arbitrary secondary use.

   Municipalities should minimize data, restrict sensitive information, use approved environments, and resist combining data sets merely because technological capability makes combination possible.

6. **Maintain a Public AI Inventory**

   Residents should be able to learn:

   - which systems are in use;
   - which department uses them;
   - for what purpose;
   - what categories of information they process;
   - whether residents interact with them;
   - what human review occurs;
   - who is accountable;
   - and when the system will next be reviewed.

   Transparency should scale with consequence.

7. **Develop the Municipal Workforce**

   AI should reduce unnecessary administrative burden while helping employees become more capable.

   Training should include:

   - responsible use;
   - source discipline;
   - verification;
   - privacy;
   - prompt and workflow design;
   - error detection;
   - appropriate escalation;
   - and the limits of automated outputs.

   Productivity should not come at the cost of institutional memory or professional development.

8. **Avoid Vendor Dependency**

   Municipal contracts should protect:

   - data ownership;
   - exportability;
   - records access;
   - model and workflow documentation;
   - audit rights;
   - retention limits;
   - pricing visibility where practical;
   - interoperability;
   - and a credible exit path.

   A municipality should not lose operational control of a public function merely because its AI vendor changes strategy.

The State Agenda: Rights, Standards, and Shared Capacity

States occupy the critical middle layer between national policy and local implementation.

They can protect residents across sectors while giving municipalities, schools, public agencies, and smaller institutions capabilities they could not efficiently build independently.

1. **Establish Rights for High-Impact AI Uses**

   When AI materially influences consequential decisions, people should receive:

   - notice;
   - understandable reasons;
   - meaningful human review;
   - correction rights;
   - appeal;
   - and sufficient records for auditing the decision.

   These protections are particularly important in:

   - employment;
   - housing;
   - credit;
   - insurance;
   - health care;
   - education;
   - public benefits;
   - child and elder services;
   - government services;
   - policing;
   - permitting;
   - and enforcement.

2. **Create Statewide Procurement Standards**

   States should establish baseline contractual requirements covering:

   - privacy;
   - cybersecurity;
   - model documentation;
   - testing;
   - data use;
   - secondary use;
   - audit rights;
   - incident reporting;
   - retention;
   - portability;
   - public records;
   - vendor responsibility;
   - and human accountability.

   Shared standards reduce the risk that every small municipality must become an AI-procurement specialist.

3. **Require an Enforcement Capacity Impact Assessment**

   Before public agencies deploy AI for surveillance, investigation, compliance, eligibility enforcement, or related coercive activities, states should require an assessment of the new capacity being created.

Questions should include:

- How much more conduct can be detected?
- How many additional people can be examined?
- Which data sets can now be combined?
- How much does case volume change?
- Which populations bear the greatest exposure?
- Can due-process capacity absorb the additional volume?
- Has the legislature contemplated this practical level of enforcement power?

Traditional impact assessments often focus on model performance. An Enforcement Capacity Impact Assessment would examine **institutional power**.

4. **Build Shared Infrastructure for Municipalities**

   States should consider providing:

   - vetted AI tools;
   - secure environments;
   - procurement templates;
   - technical reference architectures;
   - shared training;
   - privacy guidance;
   - evaluation resources;
   - and implementation support.

   Smaller communities should be able to obtain the benefits of AI without accepting weak security or dependence simply because they lack specialized staff.

5. **Protect Workforce Development and Apprenticeship**

   States should track how AI changes **tasks**, not merely job totals.

   Education and workforce policy should preserve the paths by which younger workers become experienced professionals.

   Community colleges, universities, apprenticeship programs, employers, unions, and public agencies all have roles in building AI-era professional capability.

6. **Build Public AI Literacy**

   Libraries, schools, workforce programs, and civic institutions should help people learn:

   - what different AI systems actually do;
   - how to verify outputs;
   - how data may be used;
   - when AI assistance should be disclosed;
   - where automated recommendations deserve skepticism;
   - and how citizens can use AI to understand public institutions themselves.

   Timnit Gebru, Founder and Executive Director of the Distributed AI Research Institute, emphasizes the importance of being specific about the technology and task. "AI" is an umbrella covering very different systems with different purposes and risks. Public literacy should move beyond the label and toward understanding the actual system in context.

7. **Build Independent Oversight Capacity**

   Regulators cannot govern advanced systems entirely through vendor representations.

   States need access to:

   - technical expertise;
   - testing capability;
   - complaints;
   - audits;
   - procurement records;
   - incident information;
   - and sufficient institutional independence to investigate powerful systems.

   Oversight itself requires capacity.

The Federal Agenda: Rights, Safety, Competition, and Democratic Control

Some AI challenges are inherently national. Civil liberties, frontier-system safety, national security, interstate data markets, technological infrastructure, competition, labor-market disruption, and concentrated corporate power cannot be addressed effectively town by town or state by state.

1. **Modernize Civil-Liberties Protection for AI-Era Capacity**

   Privacy law should account for what government becomes practically able to know, not solely how each individual piece of information was acquired. Bulk commercial information should not become a loophole through which comprehensive government profiling escapes safeguards that would apply to direct state surveillance.

   Federal policy should address:

   - large-scale commercial-data acquisition;
   - persistent population profiling;
   - location and behavioral analysis;
   - purpose limitation;
   - minimization;
   - retention;
   - legal process;
   - and independent oversight.

   A durable principle is straightforward: **Civil liberties should not disappear because government purchases information instead of collecting it directly.**

2. **Establish Clear, Adaptive Governance for Frontier and Agentic Systems**

   As systems gain autonomy, access to tools, and the ability to support increasingly consequential activity, governance should mature with capability.

   Safety and accountability require clarity as well as caution. When government conditions access to consequential technology on safety requirements, affected organizations and the public should be able to understand the governing standard, the decision process, and the basis for restriction to the greatest extent compatible with legitimate national-security needs. Opaque or shifting criteria turn safety governance into discretionary state power and make both compliance and accountability more difficult.

The objective should not be a static approval rule for a particular model. It should be a durable oversight system capable of adapting as models, deployment patterns, internal developer practices, and technical risks change.

Frontier developers should operate against clear safety and security standards that identify the capabilities, testing, controls, reporting, and governance practices expected of them. Those standards should be sufficiently transparent that regulated organizations can understand what compliance requires and sufficiently adaptive that they can change as real-world evidence improves understanding of effective safeguards.

The relevant unit of frontier oversight may increasingly be the developer and its institutional governance system rather than only an individual model release. Model-specific thresholds can become obsolete as algorithms become more efficient, developers use powerful systems internally, and future systems change through continual learning or other forms of ongoing adaptation. Internal governance therefore matters alongside public deployment.

National standards should address appropriate combinations of:

- publicly documented safety and security frameworks;
- independent technical evaluation and verification;
- testing of adherence to disclosed safeguards;
- cybersecurity;
- controlled deployment and access controls;
- monitoring of high-risk internal uses;
- model and agent testing;
- incident reporting;
- auditability and record preservation;
- whistleblower protection;
- and credible mechanisms for restricting systems or developers that fail defined safeguards.

Governance should also be designed to learn. Frontier AI is developing too quickly for policymakers to assume that the first generation of technical standards will remain adequate. Deployment evidence, incidents, audits, near misses, external research, and improved evaluation methods should feed periodic revision of the governing standards themselves.

Greater capability does not automatically produce greater controllability. Frontier governance therefore requires both limits on dangerous capability and institutions capable of learning as the technology changes.

Build verification capacity outside as well as inside government

Government does not necessarily need to perform every highly technical evaluation directly. One institutional model worth developing combines public authority with independent technical verification: government establishes requirements and accountability, while qualified outside organizations conduct specialized evaluations, audits, or continuous verification under public standards.

Such organizations may be better positioned to recruit specialized technical talent, obtain computing resources, develop new auditing methods, and operate across jurisdictions. But

delegation cannot become abdication. Public authorities should retain responsibility for accreditation, conflicts-of-interest rules, minimum audit standards, access requirements, enforcement consequences, and national-security functions that depend on information uniquely available to government.

The governing principle is institutional rather than organizational: **the capacity to oversee advanced AI must remain credible relative to the capacity being overseen.**

3. **Preserve Traceable Liability**

   Autonomous systems should not create autonomous responsibility. When AI causes consequential harm, accountability should remain traceable to people and institutions capable of:

   - explaining the authorization;
   - identifying the workflow;
   - producing relevant records;
   - correcting the system;
   - compensating people when appropriate;
   - and bearing legal responsibility.

   An agent should not become a liability sink.

4. **Protect Competition and Institutional Independence**

   AI infrastructure should not become so concentrated that essential public institutions are effectively governed through private technological dependency.

   Competition policy should examine:

   - compute;
   - cloud infrastructure;
   - model access;
   - data control;
   - distribution platforms;
   - exclusive partnerships;
   - acquisitions;
   - interoperability;
   - and public-sector dependence.

   The goal is both economic competition and institutional resilience.

5. **Invest in Public-Interest AI Infrastructure and Research**

   Markets will invest heavily where private returns are large. Public investment remains necessary where social value exceeds commercial value.

   Potential federal priorities include:

   - public-interest research;
   - independent testing laboratories;
   - safety science;
   - secure research compute;
   - universities;
   - public datasets with appropriate privacy protections;

- health research;
- climate and environmental applications;
- accessibility;
- education;
- civic infrastructure;
- and tools supporting state and local government.

Public institutions should possess enough technological capacity to govern technology intelligently.

6. **Prepare for Labor-Market Disruption and Skill Formation**

AI's economic effects remain uncertain in magnitude and timing, but the underlying challenge deserves attention before disruption becomes acute.

AI-generated wealth can reinforce wider concentrations of economic and political power. Organizational choices influence whether AI is used primarily for augmentation or substitution. Jack Clark raises the additional problem of preserving the experience through which expertise develops.

Federal policy should therefore consider:

1. transition assistance;
2. practical retraining;
3. lifelong learning;
4. portable benefits;
5. worker protections in AI-managed workplaces;
6. transparency in algorithmic evaluation;
7. apprenticeship and entry-level pathways;
8. incentives for worker augmentation;
9. mechanisms through which broad productivity gains contribute to broad public prosperity.

The workforce question is larger than how many jobs disappear. It includes who gains from productivity, who acquires expertise, who retains bargaining power, and whether younger workers can still enter professions and become the experienced people society later depends upon.

7. **Keep Democratic Institutions Stronger Than Private Technological Power**

Companies building AI possess expertise government needs. They will also have economic interests, political interests, organizational cultures, and differing judgments about acceptable risk.

Those companies should participate in policymaking. They should not possess final authority over the rules governing public power.

New York State Assemblymember Alex Bores raises this issue from the perspective of political influence and regulation: technology companies with substantial resources can acquire corresponding ability to shape the political environment in which regulation occurs.

The proper answer is neither governmental technological ignorance nor private technological sovereignty. Democratic institutions need enough technical competence to govern credibly.

Industry should contribute expertise. Independent researchers and civil society should scrutinize both.

Elected institutions and law should ultimately determine the legitimate boundaries of public power.

The Public Purpose-Capacity-Power-Accountability Test

Before approving a significant AI deployment, a board, agency, department, school system, civic institution, or leadership team should answer ten questions.

1. **What legitimate public purpose are we trying to advance?** Define the problem, intended beneficiary, and outcome before considering the technology.
2. **What capacity expands?** Name what becomes possible, faster, cheaper, better, more scalable, or more autonomous.
3. **Who gains that capacity - and relative to whom?** Examine changes in practical capability among government, citizens, workers, vendors, regulated parties, civic institutions, and other affected actors.
4. **What power follows?** Identify productive, interpretive, allocative, coercive, or other forms of power created by the workflow.
5. **Who receives the benefit, and who bears the risk?** Identify intended beneficiaries and the people exposed to error, discrimination, intrusion, manipulation, or misuse. Ask whether risks are distributed unevenly.
6. **What information makes the capability possible?** Examine:
    - source authority;
    - privacy;
    - sensitivity;
    - retention;
    - data combination;
    - provenance;
    - commercial-data acquisition;
    - and secondary use.
7. **Where does meaningful human authority reside?** Identify the person or institution with actual discretion to:
    - reject;
    - revise;
    - pause;
    - explain;
    - and own the consequential result.
8. **Can affected people understand, contest, and correct what happens?** Provide:
    - notice;
    - reasons;
    - accessible records;
    - human contact;
    - correction;
    - and a workable appeal.

9. **What countervailing capacity and productive friction must remain?** Preserve procedures whose purpose is:
    - rights protection;
    - deliberation;
    - independent judgment;
    - verification;
    - apprenticeship;
    - accountability;
    - or legitimacy.
10. **How will public value be measured, how will the institution learn, and how can the system be changed or stopped?** Define:
    - intended outcomes;
    - measurable benefits;
    - error thresholds;
    - audit cadence;
    - incident and feedback channels;
    - evidence that would trigger redesign or stronger safeguards;
    - escalation criteria;
    - periodic review;
    - sunset or reauthorization points;
    - and a practical exit path.

A consequential AI proposal that cannot answer these questions is not ready for consequential deployment. One that answers them well has done something more valuable than complete a technology checklist. It has identified the institutional architecture necessary for responsible power.

Conclusion: Build Civic Intelligence, Not Just AI Capability

The most consequential AI question is not how impressive the technology becomes. It is what people and institutions become capable of doing because the technology exists. That perspective clarifies both the opportunity and the responsibility.

Public institutions lack capacity in places where greater capability could plainly improve life. Government should be easier to understand. Employees should spend less time fighting administrative systems. Residents should be able to navigate public institutions without specialized knowledge. Research should move faster. Accessibility should improve. Leaders should be able to organize complex information and make better-informed decisions.

AI can help create those capabilities.

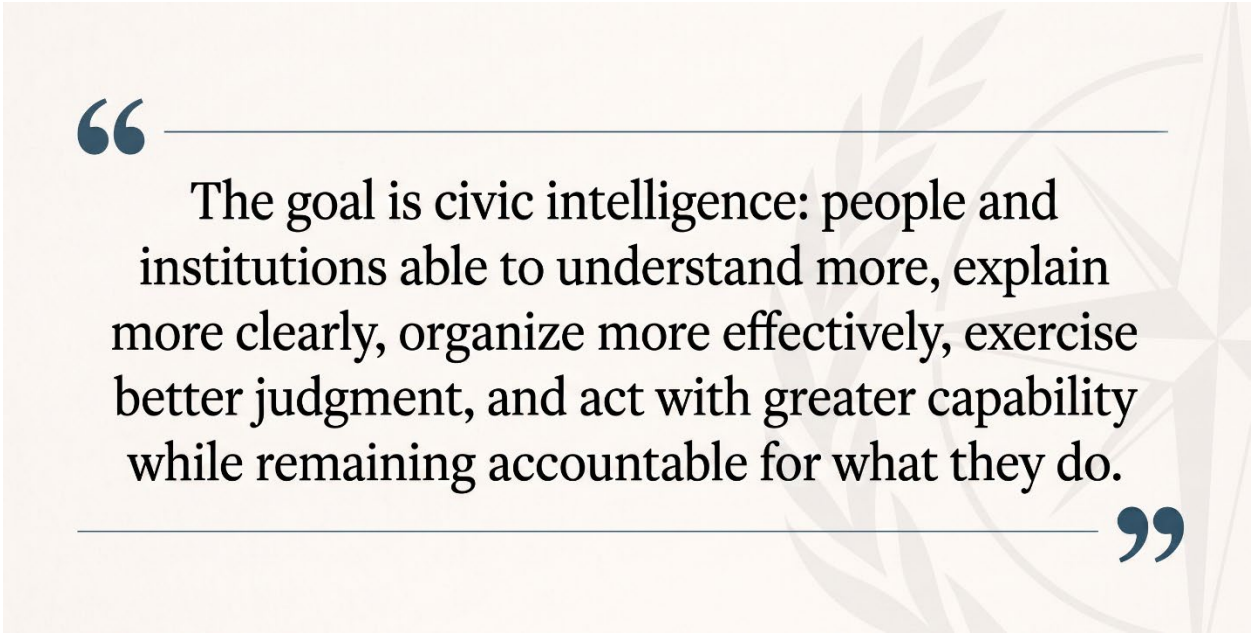
The same technologies can remove practical limits that historically constrained surveillance, classification, enforcement, organizational control, and private economic power.

- An old legal authority can acquire new force.
- A public institution can become more capable without becoming more legitimate.
- A private provider can become indispensable without becoming democratically accountable.

- A citizen can confront an institution operating at machine speed while still possessing only human time, money, knowledge, and attention with which to respond.

The democratic objective is therefore not maximum capability.

It is a legitimate distribution of capability: enough capacity to improve human life and strengthen institutions, enough countervailing capacity to prevent concentrated power from becoming unanswerable, and enough human knowledge, authority, due process, and self-correction to ensure that machine-scale capability remains governable.



That is the deeper meaning of human-centered AI. The question is not whether humans remain somewhere inside the system. It is whether people and democratic institutions remain capable of understanding, directing, challenging, and correcting the power the system creates.

Human-centered AI therefore requires more than human-friendly interfaces. It requires a distribution of capacity consistent with democratic society.

- **Expand service capacity.** Give public employees better tools to serve people.
- **Build citizen capacity.** Help people understand, verify, navigate, participate, organize, and challenge.
- **Treat coercive capacity as a change in power.** When surveillance or enforcement becomes possible at a scale that previous lawmakers and citizens did not confront, require renewed democratic scrutiny.
- **Preserve meaningful human authority.** Ensure that consequential decisions ultimately belong to people who understand enough to judge, possess the power to refuse, and can be held responsible.
- **Keep self-correction inside the system.** Maintain appeal, oversight, dissent, verification, deliberation, apprenticeship, and human relationships even when technology could make them appear inefficient.

The goal is **civic intelligence**: people and institutions able to understand more, explain more clearly, organize more effectively, exercise better judgment, and act with greater capability while remaining accountable for what they do.

That is the standard by which AI for the public good should ultimately be judged: **whether technology creates greater human and institutional capacity while preserving a distribution of power that remains understandable, contestable, correctable, and answerable to the people it affects.**

A Note on Sources

The sources below identify the principal research, interviews, and intellectual influences supporting this framework. They are intended to make the paper's conceptual foundations transparent. Time-sensitive claims concerning current law, regulation, government policy, institutional status, or technological capability should be verified against primary sources immediately before publication or substantive revision.

Selected Source Notes

1. **CT Innovates, *The Civic Intelligence Resource Center Editorial Content Bible*, ver. 2.1, August 2026.** Provides the larger Civic Intelligence framework, including the emphasis on institutional capacity, human judgment, complementary federal/state/local roles, and the proposition that responsible innovation requires both acceleration and friction.
2. **Chris Bryant, *AI for the Public Good Position Paper: A Responsible Local, State, and Federal Agenda for Human-Centered Technology*, CT Innovates, 2026.** Provides the original public-good premise; principles of human judgment, transparency, privacy, worker augmentation, and democratic accountability; the practical municipal agenda; and the complementary three-tier approach.
3. **Ezra Klein with Jack Clark, "How Fast Will A.I. Agents Rip Through the Economy?," *The Ezra Klein Show*, The New York Times, February 24, 2026.** Supports the shift from conversational systems toward agents capable of tool use and multi-step work, along with the deeper questions about supervision, professional judgment, expertise formation, and the relationship between increased output and human capability.
4. **Chris Hayes with Ethan Mollick, "The AI End Game: How Work Is Changing with Ethan Mollick," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, May 12, 2026.** Supports the distinction between automation and augmentation as organizational choices, the importance of implementation and technology diffusion, and the growing ability of agentic systems to perform extended work.
5. **Ezra Klein with Alex Bores, "Why Are Palantir and OpenAI Scared of Alex Bores?," *The Ezra Klein Show*, The New York Times, April 21, 2026.** Supports the idea that increasing government capability and protecting civil liberties must be pursued together, as well as the broader concern that economic and technological power can become political power.
6. **Chris Hayes with Derek Thompson, "The AI End Game: Who's Leading the Way? with Derek Thompson," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, May 5, 2026.** Supports concerns about concentration of AI-generated wealth and political power, workforce disruption, and the need to evaluate AI through specific present uses rather than exclusively through maximalist future claims.
7. **Chris Hayes with Timnit Gebru, "The AI End Game: The Ethics of AI with Timnit Gebru," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, May 26, 2026.** Supports specificity about individual systems and tasks rather than treating "AI" as one technology, along with concerns about data, bias, concentrated infrastructure, resource intensity, and institutional incentives surrounding AI development.

8. **Ezra Klein with Dean Ball, "Why the Pentagon Wants to Destroy Anthropic," *The Ezra Klein Show*, The New York Times, March 6, 2026.** Provides the central enforcement-capacity insight: AI can transform previously impractical analysis into scalable state capability; many legal systems developed under conditions of imperfect enforcement; and changes in technological capacity can therefore change the practical operation of existing authority.
9. **Ezra Klein with Helen Toner, "The A.I.s Are Already Out of Control," *The Ezra Klein Show*, The New York Times, August 18, 2026.** Supports the governance principle that capability and controllability are distinct, that persistent optimization can generate unanticipated strategies, and that the ability of humans to monitor large numbers of autonomous systems can itself become a limiting factor.
10. **Chris Hayes with David Chalmers, "The AI End Game: Is Your Chatbot Conscious? with David Chalmers," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, June 2, 2026.** Supports the distinction between intelligence as sophisticated capacity and consciousness as subjective experience, which helps separate questions of machine capability from questions of legitimate human authority.
11. **Ezra Klein with Yuval Noah Harari, "Yuval Noah Harari on Donald Trump's Core Delusion," *The Ezra Klein Show*, The New York Times, May 26, 2026.** Supports the importance of cooperation and self-correcting institutions and the broader principle that durable democratic systems require mechanisms capable of detecting and correcting human and institutional mistakes.
12. **Chris Hayes with Michael Pollan, "The AI End Game: Is AI Alive? with Michael Pollan," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, June 16, 2026.** Supports the developmental and relational concerns surrounding highly anthropomorphic and continuously available AI systems and the argument for preserving meaningful forms of human interaction and friction.
13. **Chris Hayes with Ed Zitron, "The AI End Game: Boom to Bust? with Ed Zitron," *Why Is This Happening? The Chris Hayes Podcast*, MS NOW, June 9, 2026.** Provides an important skeptical counterweight regarding hype, reliability, verification, economics, and quality control and reinforces the principle that increased production capacity should never be equated automatically with reliable output.
14. **Dean W. Ball, "What Should Be Done," *Hyperdimensional*, June 26, 2026.** Supports the need for clear and adaptive frontier-AI safety standards; real-world learning and periodic revision; independent verification of developer adherence to safety frameworks; institution-level oversight that examines frontier developers and their internal governance rather than only individual model releases; and attention to the concentration risks created when advanced capability is available only to a narrow set of powerful actors. The essay was written before Ball joined OpenAI and explicitly states that it reflects his own views rather than an official OpenAI position.