

Civic AI Agents for Boston and the Identity–Authorization Challenge

Executive summary

The discussion produced a clear architectural direction for building **AI agents for Boston residents**: Boston does not necessarily need to create, own, or control every resident's personal AI agent. Instead, the city should build a **trusted civic access layer** through which citizen-selected, business-operated, nonprofit-operated, and government-operated agents can interact safely with Boston services.

This reframes the initiative from “build a Boston chatbot” to:

Create the identity, authorization, delegation, discovery, and accountability infrastructure that allows any trustworthy AI agent to serve a Boston resident.

The group identified identity and authorization as the central challenge. A civic transaction may involve at least four distinct entities:

1. The resident or organization receiving the service.
2. The AI agent acting on their behalf.
3. The company or institution operating that agent.
4. The city service agent or system being accessed.

Boston must be able to determine not only **who the resident is**, but also **what the agent is, who operates it, what authority it has been delegated, what civic functions it is certified to perform, and whether the requested action is appropriate in the current context**.

The discussion therefore supports a model built around three city responsibilities:

- **Identity and verification**: Know the citizen, business, representative, and agent.
- **Authorization and civic charters**: Define what each type of agent may access or do.
- **Observability and accountability**: Record actions, investigate misuse, revoke access, and assign responsibility.

These elements form the foundation for safe, population-scale civic agents.

1. The core problem the group identified

The group began with several foundational questions:

- What should a civic agent identity contain?
- Should an AI agent have its own identity, or merely use a resident's credentials?
- How should Boston's identity system connect with MyMassGov and other state systems?
- Which attributes should move with residents across agent providers—for example language, accessibility settings, permissions, service history, and preferences?

The most important conclusion was that it is unlikely that Boston—or any single government entity—will issue the one AI agent every resident uses. Residents will obtain agents from many sources: major technology companies, employers, health systems, banks, nonprofits, community organizations, neighboring municipalities, or open-source providers.

A person might live in Cambridge, work in Boston, receive healthcare from MGH, use a commercial personal agent, and interact with Massachusetts state services. No single institution will own the entire relationship.

Boston's problem is therefore not primarily:

How does the city build the best personal AI agent?

It is:

How does the city safely interact with AI agents it did not create, does not operate, and cannot fully inspect?

This is a much more consequential infrastructure problem. It requires Boston to establish a civic trust boundary: the rules, credentials, protocols, and controls that apply whenever an external agent attempts to access Boston services.

The city would not need to regulate what a resident's agent does when checking sports scores, making a restaurant reservation, or shopping. Boston's authority would apply specifically when that agent crosses into civic systems—for example requesting a permit, accessing a resident's case, filing a service request, or establishing eligibility for a benefit.

2. Why this strongly supports “AI agents for every Boston resident”

The discussion validates the broader vision because civic services are fragmented by design. A resident may have to interact with:

- Several city departments.

- Massachusetts state agencies.
- Schools and libraries.
- Housing providers.
- Utilities.
- Nonprofits and social-service organizations.
- Healthcare institutions.
- Private businesses whose activities trigger government requirements.

Residents are currently responsible for understanding which institution handles which service, finding the correct website, repeatedly entering the same information, determining their eligibility, and following up when something goes wrong.

A personal civic agent could absorb much of this complexity. Instead of requiring residents to understand government organization, the resident could express an outcome:

- “I am moving next week. Handle everything I need.”
- “My trash was not collected.”
- “Find the housing programs for which my family qualifies.”
- “I bought a home. Make sure I receive the exemptions available to me.”
- “My mother needs an accessible recreation program near her home.”
- “I need to reserve a moving space for a furniture delivery.”

The agent could then identify the appropriate jurisdiction, locate relevant services, assemble the required evidence, ask for narrowly defined permission, complete allowable actions, and report the result.

This is particularly valuable for residents who have difficulty navigating government because of language, disability, limited time, legal uncertainty, digital literacy, or unfamiliarity with government structures. The agent becomes a translation and coordination layer between the resident’s intent and government’s institutional structure.

The objective is therefore not merely to add conversational interfaces to existing websites. It is to let agents coordinate services across organizational boundaries while preserving the resident’s control.

3. The city should enable agents, not necessarily issue the agent

A recurring theme was that Boston should play a **governance and infrastructure role**, rather than becoming the exclusive provider of personal AI.

The city would provide:

- Authoritative civic service agents or APIs.
- A method for discovering those services.
- Authentication and authorization requirements.
- Agent certification or registration.
- Rules governing allowable uses.
- Secure delegation mechanisms.
- Audit records and enforcement.
- Resident-facing controls and recourse.

Third parties could continue to build the resident-facing agents. This gives residents choice while preventing a single commercial provider from becoming the mandatory gateway to Boston government.

It also avoids requiring Boston to maintain every possible user interface, language model, device surface, accessibility mode, or personal-assistant application. Boston could instead maintain a stable, public civic interface that many approved agents can call.

This is analogous to a city maintaining roads and traffic rules without manufacturing every vehicle that uses them.

4. Identity is a multi-party problem, not simply resident login

Conventional digital government usually asks one identity question:

Is this person the legitimate account holder?

Agentic government adds several more:

- Is this a real and verified resident?
- Is the resident eligible for this particular service?
- Is the requesting software actually an AI agent?
- Who operates the agent?
- Has the agent been altered since it was approved?
- Has the resident authorized this particular agent?
- Has the resident authorized this particular action?
- Is the agent allowed by city policy to perform that category of action?
- Is another agent allowed to delegate part of the task to it?
- Which actor is accountable if something goes wrong?

The group described this as an extension of existing “Know Your...” frameworks:

- **KYC:** Know Your Customer or Citizen.

- **KYB:** Know Your Business.
- **KYA:** Know Your Agent.

All three may be required in one civic interaction.

For example, a furniture company's agent might coordinate with a resident's personal agent to request a temporary parking permit. Boston would need to know:

- That the resident is associated with the address.
- That a real delivery or move is taking place.
- That the personal agent is authorized by the resident.
- That the company agent represents the stated business.
- That each agent has permission to perform its part of the workflow.
- That neither agent receives unnecessary information about the resident.

A resident login alone cannot answer all of these questions.

5. Agents should have identities separate from human identities

One of the most important design choices concerned whether an agent should simply use the resident's OAuth credentials or be registered as a separate actor.

Allowing an agent to act entirely through the resident's identity is initially simpler. But it creates serious problems:

- Human actions and agent actions become indistinguishable.
- The agent may inherit every permission held by the resident.
- The city cannot revoke the agent without also affecting the resident.
- Audit records cannot reliably explain who or what initiated an action.
- Different agents operated by the same resident cannot be managed separately.
- Compromising one agent could expose the resident's entire civic account.

A stronger model gives the agent its **own cryptographic and administrative identity**, while linking it to the resident through a delegation relationship.

A civic transaction would therefore involve:

- **Resident credential:** Who the human or organization is.
- **Agent credential:** What software agent is requesting access.
- **Operator credential:** Who is legally or operationally responsible for the agent.
- **Delegation credential:** What the resident has authorized the agent to do.
- **City authorization:** What Boston permits this class of agent to do.

This resembles the difference between giving someone your password and issuing them a limited power of attorney. The latter records the representative's identity and limits their authority.

The group also noted that city identity systems already confront related forms of agency:

- A parent acting for a child.
- A guardian acting for another person.
- A company representative acting for a business.
- An employee permitted to change some corporate records but not others.

AI agents extend this existing problem rather than creating it from nothing. Boston can build on established concepts of representation, guardianship, organizational roles, and delegated authority.

6. A three-part trust model emerged

The discussion identified three related responsibilities for the city.

A. Identity: Who are the parties?

Boston needs mechanisms to identify or verify:

- Residents.
- Businesses and institutions.
- Human representatives.
- AI agents.
- Agent operators.
- Civic services and city-operated agents.

The assurance required should depend on the transaction. Looking up a public trash schedule should not require the same identity proof as changing a property record.

B. Charter and authorization: What may each agent do?

Identity does not by itself create permission.

An MGH-operated agent might legitimately receive authorization to interact with certain health-related services. A retail or transportation agent should not be able to retrieve health information merely because the resident clicked “allow.”

This is a crucial principle:

User consent should not override every institutional or public-interest restriction.

The city should define civic “charters” or capability policies for classes of agents. A charter would state:

- Which services the agent can access.
- Which data categories it may request.
- Which actions it may initiate.
- Whether it may perform the final submission.
- Whether it may delegate to another agent.
- Which resident confirmations are required.
- Which records it must retain.
- Which conduct is prohibited.

A resident delegation and a city charter must both permit the action. Neither is sufficient alone.

C. Observability and accountability: What happened?

Boston needs some ability to understand agent activity across the civic ecosystem. This does not require viewing every private conversation. It does require transaction-level records such as:

- Which agent contacted which service.
- Which resident or organization authorized it.
- What scope was requested.
- What evidence or credentials were presented.
- What action was taken.
- Which downstream agents were involved.
- Whether human confirmation occurred.
- What result was returned.
- Whether the transaction triggered an exception.

This supports auditing, dispute resolution, rate limiting, incident investigation, credential revocation, and policy improvement.

The meeting correctly recognized this third layer as both difficult and potentially a major area for Boston to innovate.

7. Authorization should be scoped delegation, not blanket consent

The discussion challenged the assumption that conventional digital consent is enough for AI systems.

Today, users frequently encounter choices such as:

- Allow once.
- Always allow.
- Accept all.
- Connect account.

These interfaces already encourage overbroad consent. Autonomous agents make the problem worse because users want the agent to complete work without interrupting them for every step.

A workable civic-agent model needs something stronger than a consent screen. It needs **structured delegation**.

A delegation should specify:

- **Agent:** Which agent receives authority.
- **Purpose:** Why the authority is being granted.
- **Service:** Which city system can be accessed.
- **Data:** Which attributes may be disclosed.
- **Action:** What the agent may read, submit, change, or execute.
- **Duration:** One transaction, one day, a recurring period, or until revoked.
- **Jurisdiction:** Boston, Massachusetts, or another relevant authority.
- **Delegation rights:** Whether the agent can call or authorize another agent.
- **Human checkpoint:** Which actions require confirmation.
- **Revocation:** How the resident or city can terminate authority.

For example:

“Permit my personal agent to prove that I reside at this address and apply for one moving permit for July 8, but do not disclose my full identity to the delivery company and do not authorize any other parking transaction.”

That is substantially safer than granting an agent general access to the resident’s Boston account.

8. Some actions may require human confirmation or remain non-delegable

The group distinguished among actions of different consequence.

Public information and read-only services are relatively easy. High-impact actions such as payments, legal declarations, demolition permits, changes in property records, or surrendering benefits may require stronger controls.

A useful risk ladder would be:

Level 0: Public read

Examples include general schedules, open datasets, public eligibility rules, and service descriptions. No resident authentication is required.

Level 1: Authenticated read

The agent can retrieve resident-specific information, such as case status or the correct trash schedule for an authenticated address.

Level 2: Low-risk or reversible write

The agent can report a missed collection, request a recycling cart, update a communication preference, or reserve a low-risk service.

Level 3: Consequential transaction

The agent can prepare and submit a permit, application, benefit request, or payment, but the resident may be required to approve the final action through step-up authentication.

Level 4: Restricted or non-delegable action

Certain legal declarations, property changes, waivers of rights, or unusually consequential actions may require direct human execution or in-person verification.

This graduated approach lets Boston deploy useful services without waiting to solve every high-risk case.

9. “Authenticated but anonymous” is a major opportunity

One of the discussion’s strongest privacy insights was that authentication and identification should not always be treated as the same thing.

Residents may want to submit 311 reports, request assistance, or access information without attaching their names. This can be especially important for undocumented residents, people

reporting unsafe conditions, or anyone concerned that seeking one government service could expose unrelated personal information.

An agent could present a short-lived credential proving only what the transaction requires, such as:

- The requester is a Boston resident.
- The requester is associated with a particular property.
- The requester is eligible for a service.
- The requester has not exceeded a permitted request limit.
- The request comes from a valid, non-revoked resident credential.

The city service would not necessarily receive the resident's full identity.

This is the difference between:

“This is Ramesh Raskar, with this complete account history.”

and:

“A verified resident associated with this address is authorized to make one request.”

Temporary tokens, selective disclosure, pseudonymous identifiers, and verifiable attributes could make civic interaction more private than many existing web forms—not less.

But the group also correctly noted the danger that an external agent could harvest information even when the city protects it. Therefore, privacy restrictions must apply to the agent and its operator, not only to the city service.

10. Concrete examples made the model understandable

Furniture delivery and temporary parking

A resident buys furniture. The retailer or delivery company knows a delivery truck will require temporary parking. Its agent contacts the resident's personal agent, which determines that a Boston permit is required.

The resident's agent requests narrow authorization and interacts with the city's parking-service agent. The resident does not have to understand the permit workflow or manually navigate a website.

This demonstrates a chain in which:

- A commercial event triggers a civic requirement.

- A business agent communicates with a citizen agent.
- The citizen agent communicates with a city agent.
- The resident remains the authority behind the transaction.
- The business does not need broad access to the resident's civic account.

Buying a home and obtaining an exemption

A real-estate or home-management agent could recognize that a new homeowner may qualify for a residential exemption. It could locate the appropriate program, assemble the needed information, prepare the application, and request final approval from the resident.

Trash and 311 services

A resident's agent could proactively explain schedule changes caused by events or holidays. If the resident then reports unwanted items or a missed pickup, the agent could invoke the relevant 311 service in context.

These cases show why civic agents are more than question-answering systems. They can connect events in residents' lives to government processes and complete multi-step workflows.

11. The discussion also exposed real security threats

The meeting included practical examples of misuse:

- An agent requesting a parking permit for someone who is not eligible.
- Automated flooding of civic systems.
- A hackathon participant creating approximately 100 tickets on one street.
- False or nonsensical reports being submitted automatically.
- Agents impersonating licensed or certified professionals.
- Agents obtaining resident data outside their proper purpose.
- Business agents silently initiating civic transactions.
- Predatory interfaces manipulating residents into broad authorization.

These are not theoretical edge cases. Agent automation changes the economics of abuse: a person no longer has to fill out 100 forms manually.

Boston will need controls including:

- Rate limits by person, agent, operator, address, and service.
- Proof of relevant eligibility.
- Agent registration and signed requests.
- Purpose-bound authorization.
- Bot and anomaly detection.

- Restrictions on bulk operations.
- Human confirmation for unusual activity.
- Revocation and suspension mechanisms.
- Signed transaction receipts.
- Clear liability and dispute processes.

The risks reinforce the need for a formal agent-access layer. Without it, agents will still attempt to interact with city systems—by browsing websites, automating forms, or using unofficial endpoints—but Boston will have less visibility and control.

12. How Boston Home and MyMassGov could fit together

The conversation raised the need to connect Boston's resident identity infrastructure with MyMassGov and other systems.

The appropriate model is likely federation rather than merging every account into a single database.

The state may provide higher-assurance verification for attributes such as:

- Legal identity.
- Driver's-license relationship.
- State residency.
- Business registration.
- Professional credentials.

Boston may provide authoritative city-specific attributes such as:

- Boston residency.
- Property or address association.
- City program enrollment.
- Local service history.
- Neighborhood or district.
- Existing permit or case relationships.

An agent should receive only the attributes needed for a transaction. The state and city do not need to reveal all underlying records to each other or to the external agent.

For example, MyMassGov might support a proof that the person has a verified Massachusetts identity, while Boston supplies proof that the person is associated with a specific Boston residence. The parking system receives the combined proof needed for the permit, not the resident's complete state and city profiles.

13. Portable attributes should be separated from provider-specific agent memory

The group discussed portability of language, accessibility preferences, permissions, history, and related attributes.

A resident changing from one agent provider to another should not have to reconstruct all civic preferences. But portability must not become an uncontrolled transfer of personal history.

A useful division is:

Portable resident preferences

- Preferred language.
- Accessibility requirements.
- Communication channel.
- Reading level or explanation preferences.
- Preferred form of address.
- Notification preferences.

Portable delegation policies

- Services the resident permits agents to access.
- Default confirmation requirements.
- Data-sharing boundaries.
- Named caregivers, guardians, or representatives.
- Revoked agents and providers.

Authoritative civic attributes

These should remain with the issuing institution and be retrieved or proven when needed:

- Residency.
- Property ownership or occupancy.
- Program eligibility.
- Permit status.
- Business authority.
- Professional license.

Provider-private memory

Conversation histories, inferred behavior, commercial profiles, and unrelated personal data should not automatically move into civic systems.

This distinction allows continuity without creating a centralized surveillance profile.

14. NANDA's role: discovery and resolution within the trust architecture

The discussion's architecture aligns naturally with Project NANDA, provided discovery is clearly separated from authorization.

NANDA can help answer:

- What is the authoritative Boston parking agent?
- Where is the 311 service agent?
- Which version of the service is current?
- What protocol does it support?
- What capabilities does it advertise?
- Which authority published or certified it?
- What endpoint should an external agent contact?

Boston could publish discoverable civic service records and agent cards. A personal or business agent could resolve the appropriate service through a trusted index or city-owned domain.

But discovery alone does not answer:

- Is the requesting agent trustworthy?
- Is the resident authenticated?
- Has the resident delegated this action?
- Is the action permitted?
- Which attributes may be disclosed?

NANDA can function as the **switchboard**, while the identity and authorization infrastructure determines whether the connection may proceed.

A Boston civic agent record could therefore include or reference:

- Service identity.
- City ownership or endorsement.
- Supported operations.
- Required authentication level.
- Accepted credentials.

- Action-risk category.
- Delegation requirements.
- Human-confirmation requirements.
- Audit and privacy policies.
- Version and revocation status.

This combination—discovery plus trust policy—would make civic agent services usable by agents from many providers without making those providers part of Boston’s internal infrastructure.

15. A practical architecture for Boston

Synthesizing the discussion, a Boston Civic Agent Trust Architecture could contain the following components.

1. Federated resident identity

Boston Home, MyMassGov, and other approved identity sources establish the resident’s identity and relevant civic attributes.

2. Independent agent identity

Every agent interacting with protected services has a distinct identity tied to its provider, operator, software version, and certification status.

3. Civic agent charter

Boston defines the classes of data and actions available to the agent. This is the city’s authorization of the agent category.

4. Resident delegation token

The resident grants a narrowly scoped, time-limited authority to a particular agent for a particular purpose.

5. Civic service gateway

A common gateway evaluates:

- Resident proof.
- Agent credential.
- Operator status.

- Civic charter.
- Delegation scope.
- Transaction risk.
- Rate and anomaly signals.

6. Selective attribute exchange

The gateway releases only the minimum attributes required for the service.

7. Human step-up

For consequential actions, the resident receives a clear final approval request through a trusted channel.

8. Audit receipt

The resident, agent, and city receive a record explaining what was requested, authorized, disclosed, and executed.

9. Revocation and enforcement

Residents can revoke an agent. Boston can suspend an agent, provider, credential, or capability. High-risk providers can be required to meet stronger assurance standards.

16. Recommended deployment path

The discussion suggests a staged implementation.

Phase 1: Public information and discovery

Expose authoritative city information through agent-readable interfaces. Begin with open data, service descriptions, schedules, and eligibility guidance.

This phase has minimal identity risk and allows Boston to learn how external agents use civic data.

Phase 2: Authenticated read access

Allow agents to retrieve narrowly scoped resident-specific information, such as collection schedules, service-request status, or application status.

This tests federated identity, agent identity, and selective disclosure without allowing consequential actions.

Phase 3: Low-risk transactions

Support reversible actions such as missed-trash reports, recycling-cart requests, notification preferences, and selected 311 requests.

Add rate limits, resident delegation, and signed audit receipts.

Phase 4: Cross-agent civic workflows

Pilot cases in which business or institutional agents coordinate with citizen agents—for example moving permits, property services, or selected small-business interactions.

This is where agent-to-agent delegation becomes essential.

Phase 5: High-value, higher-risk services

Extend to housing, benefits, licensing, permitting, and payments, with stronger verification and mandatory human checkpoints where appropriate.

The group also suggested that small and medium businesses may be an effective initial beachhead because they are likely to adopt agents earlier and interact with city systems frequently. A trusted-access pilot with property managers, local businesses, universities, community organizations, or healthcare institutions could complement resident-facing pilots.

Conclusion

The discussion substantially advances the goal of creating AI agents for Boston citizens because it identifies the missing foundation: **a civic trust architecture for agents**.

The city does not need to select one universal agent or compel residents to use a government-built assistant. It needs to ensure that, whenever any agent interacts with Boston, the system can answer five questions:

1. **Who is the resident, organization, or representative?**
2. **What agent is making the request, and who operates it?**
3. **What authority has the resident delegated to that agent?**
4. **Does Boston permit that agent to perform this type of civic action?**
5. **Can the action be explained, audited, challenged, and revoked?**

Solving those questions would enable an open ecosystem in which residents can choose their agents while Boston protects public systems, privacy, equity, and institutional accountability.

The most important outcome of the discussion is therefore a shift in ambition:

Boston should not merely put AI in front of city services. It should establish the public infrastructure through which trustworthy AI agents can represent residents across city services.

That infrastructure—federated human identity, independent agent identity, narrowly scoped delegation, civic charters, selective disclosure, step-up approval, and transaction-level observability—is what makes “an AI agent for every Boston resident” technically credible and publicly defensible.