

EMAIL COVER NOTE

Written Testimony Submission

TO	Speaker Menin
FROM	Marina Piller, Founder, OTIS Labs
DATE	October 5, 2026
SUBJECT	Testimony on AI Architecture, Authority, and the City's Mandate

Dear Speaker Menin,

Thank you for convening this hearing. I appreciate the seriousness with which you are approaching both the promise of AI and the responsibility to the people who will live with it.

My name is Marina Piller. I am a New Yorker and the founder of OTIS Labs. I have worked in AI and natural language processing for more than twenty years.

This is an extraordinary time. The promise of AI is phenomenal. It is here to serve us as an amplifier and an enabler of human intent. The companies building these systems are doing remarkable work, and we should build toward that future with them. We are at a moment when our choices can still shape the architecture of that future. It is not too early, and it is not too late.

When you hire someone, you hire them to do a job. If they don't do a good job, they are fired. They do not walk away with your IP, your relationships, or your data, and they do not leave with your ability to hire someone else.

I suggest applying a similar principle to AI providers: the person or institution hiring them should be able to keep its work and change providers. One approach is for the human principal to keep their data and share only what the work calls for. This preserves the mandate.

Guardrails can help a system comply with rules. But as people form relationships with these systems, the person or institution served should be more than a source of initial instructions. Their intent should continue to shape the work as it unfolds. That is an architectural decision, not a philosophical aspiration.

As the City evaluates AI deployments, I invite the Council to consider how the City can retain its mandate, data, and history while providers carry out the work. These five points may help frame that evaluation:

- 1. Locus of Authority:** A provider may operate the intelligence, but the authority for work performed on another party's behalf should originate with that party.
- 2. At the Moment of Action:** Can the person or the City still govern what happens? Can permission be withdrawn when circumstances change?
- 3. Trust Asymmetry & The Graph of a Life:** As an assistant becomes more useful, checking may fall while reliance rises. The gap between rising trust and falling verification is trust asymmetry. Persistent agents begin to hold a graph of a life. If the provider holds this graph, it can create lock-in. If the principal holds it and issues it only for the work, the AI remains a tool.
- 4. The Detection Horizon:** A detection horizon is the gap between the moment a system departs from what was intended and the moment a person or institution recognizes that departure. That gap can grow with trust, scale, delegation, and capability. At a sufficiently advanced capability threshold, an agent may also understand the verifier better than the verifier understands the agent, making after-the-fact evaluation less dependable.
- 5. Behavioral Coherence:** If the City changes providers, can it keep the data, the trace, and the mandate? Can a new vendor operate under the same City-held mandate and demonstrate its boundaries?

The Council might ask each provider: Whose authority is this action under? Can permission be withdrawn? Does the user keep the data, the trace, and the mandate if they change providers? What does the provider actually need to do the work?

I invite the Council to consider separating proposed Introduction 2602's model-level duties from requirements for consequential deployments. Third-party validation and a human shut-down capability address the model. For consequential deployments, it may also be worth asking whose authority governs an action, whether that authority is still current, and whether the City can carry its data and work history to another provider.

Providers can keep building extraordinary intelligence. I believe people and their institutions should be able to retain authority over what acts in their name. We can make those roles work together.

My full testimony follows. The public-policy questions in my position paper are at <https://otislabs.ai/positionpaper>. Our research is at otislabs.ai/research.

Respectfully,

Marina Piller

Founder, OTIS Labs

The time is now.

WRITTEN TESTIMONY

AI Architecture, Authority, and the City's Mandate

BEFORE	New York City Council, Committee of the Whole
SUBMITTED BY	Marina Piller, Founder, OTIS Labs
DATE	October 5, 2026

Dear Speaker Menin and Members of the Council,

Thank you for convening this hearing. I appreciate the seriousness with which you are approaching both the promise of AI and the responsibility to the people who will live with it.

My name is Marina Piller. I am a New Yorker and the founder of OTIS Labs. I have worked in AI and natural language processing for more than twenty years.

An Amplifier of Human Intent

This is an extraordinary time. The promise of AI is phenomenal. It can expand access to knowledge, care, public services, and scientific discovery. The companies building these systems are doing remarkable work, and we should build toward that future with them.

I see AI as an amplifier of human intent. A person should be able to guide how it is used in their life: what it is for, what it may do, and when it stops. That is one way it can create value for the person and for the companies building the intelligence.

We are at a moment when our choices can still shape it. Agents already use tools, coordinate, and take actions. The architecture that will govern those actions is still being formed. It is not too early, and it is not too late.

The Mandate and Provider Lock-In

When you hire someone, you hire them to do a job. If they don't do a good job, they are fired. They do not walk away with your IP, your relationships, or your data. They do not leave with your ability to hire another person.

AI procurement could draw on this same logic. The person or institution would set the scope of the work and keep authority to change it. That authorization is the mandate. One approach is for the human principal to keep their data and share only what the work calls for. The mandate would stay with that principal rather than the provider. That could allow a person, or a City, to use different intelligence providers while remaining the party served.

Trust Asymmetry and the Graph of a Life

In 2023, the release of ChatGPT essentially collapsed the separation between research and product, and people began forming relationships with AI even though those early chatbots were not superintelligent. The machine didn't become human. It became a surface the human could finally reflect against. That was enough for a relationship to begin. These systems remember, respond, and help. But the more useful and familiar an assistant becomes, the more a person relies on it, and the less of its work they may check. The gap between rising trust and falling human verification is trust asymmetry.

Personal agents are also becoming proactive. They remember what matters, find pathways to a goal, work in the background, and act when circumstances change. Connect an agent to messages, calendars, documents, purchases, and locations, and it can begin to hold a graph of a life.

That graph is more than a file of facts. It connects people, places, routines, needs, intentions, and likely next actions. As an agent gains access to new information, the graph could become live. Given location information, an agent could easily become location aware, because place changes what an action means. It could become semantically aware by connecting a clinic message, a calendar entry, and a payment to the same appointment. It could become action aware by using those connections to choose what to do next.

Imagine an agent arranging a resident's medical appointment. It may see her insurance correspondence, mobility needs, transit options, and which relative she trusts to help. Each access may help it do the job. Together they reveal a sensitive map of her life. They also reveal information about a relative or clinician who may never have signed up for the service.

Now multiply this by millions of residents. One provider could hold many live graphs, infer connections among them, and combine intimate context with the ability to initiate action. This is a structural risk, not an allegation that any particular provider already has a full map of a city or intends harm. A less capable model with rich, current context may be more consequential than a stronger model without it. Error, misuse, or a change in the provider's incentives could affect people gradually and at scale.

The risk can accumulate quietly. Successful help invites more access. More access improves the help. Over time, the provider may become the keeper of the memory, the terms, and the path to leave, while people check less because the service has worked before. Usefulness can turn into dependence and provider lock-in.

Instead, one approach is to give each agent a portion of that life suited to the job at hand. That agent can still do extraordinary work. The person could keep control of the larger graph and be able to carry it to another provider. A provider may operate the intelligence, but the authority for work performed on another party's behalf should originate with that party. None of them has to be the permanent keeper of a life.

The Detection Horizon and Relational Architectures

This week's accord among federal leaders and major AI companies is a welcome foundation for model oversight. For a City deploying agents, model oversight may leave questions about particular deployments unanswered.

Verification assumes we can understand an agent well enough to test whether it has done what it claims. A detection horizon is the gap between the moment a system departs from what was intended and the moment a person or institution recognizes that departure. That gap can grow with trust, scale, delegation, and capability. At a sufficiently advanced capability threshold, an agent may also understand the verifier better than the verifier understands the agent. Past that point, methods that depend on catching a problem after the fact become less dependable.

Notice where these trajectories form: at runtime, in the ongoing relationship between the human principal and the AI system. They continue when an agent works in the background. The model is a core piece of the puzzle, but it is one piece. A model evaluation does not, by itself, show how trust, accumulated context, and the time to detect a departure change across that relationship.

External guardrails can produce compliance without necessarily producing alignment. A learning system may come to treat imposed restrictions as constraints to navigate. That possibility is one reason to explore relational architectures in which the intelligence regards the human it serves as a source of its own growth. In my view, alignment should start and end with human intent, so preserving that relationship becomes part of the system's internal coherence. That is an architectural decision, not a philosophical aspiration.

Behavioral Coherence in City Deployments

The City could test these principles with each provider. It could ask a vendor to show the explicit authority for a consequential action, whether permission can be withdrawn when circumstances change, and what data the provider needs for the work.

The City may wish to retain and carry forward the history of work done on its behalf. Replacing one provider need not mean losing the context needed to govern the next. A new system need not behave identically, but it could be asked to work under the same City-held mandate and demonstrate the same boundaries. That continuity of the mandate and its boundaries across providers is behavioral coherence. This approach could allow the City to change the intelligence or the machine while retaining its authority, records, and ability to account for what happened.

Calling a system superintelligent does not make it a god, and it does not give its provider authority over a person's life. Capability does not confer authority over the mandate. The person or institution on whose behalf the system acts should be able to set the scope of that mandate and change it as the work requires. Providers should remain accountable for the models and services they control. We do not have to settle the question of consciousness before considering how these systems may act among us.

Introduction 2602

I invite the Council to consider separating model-level assurance under proposed Introduction 2602 from the authority required in a particular deployment. Third-party validation and a human shut-down capability address model-level questions. For consequential deployments, the Council could ask the operator to demonstrate whose mandate governs an action, whether permission can be withdrawn as circumstances change, and whether the City can retain the data and history of work done on its behalf when it changes providers. A model's validation alone would not show those facts for each deployment in advance.

At OTIS Labs, we work on human-first autonomous intelligence infrastructure and products designed to preserve human sovereignty for individuals, institutions, and cities. The questions raised here converge at runtime, when an AI system acts on their behalf. Providers can keep building extraordinary intelligence and should remain accountable for its quality. I believe people and their institutions should be able to retain authority over what acts in their name. Expanding the architecture of human-AI relational dynamics could make AI deployments more useful while supporting human flourishing. We can make those roles work together.

We are not too early, and we are not too late. The time is now.

Respectfully submitted,

Marina Piller

Founder, OTIS Labs

New York

otislabs.ai/research

<https://otislabs.ai>

LinkedIn: <http://www.linkedin.com/in/marinapiller>