Conway and the Agent Identity Layer

Who owns the persistent self that forms around you?

The Loom · April 2026

Leaked references to Anthropic's Conway project reveal a shift from chatbots to persistent, ambient agents. This essay argues that the strategic center of gravity in AI is moving from model quality to the identity layer—memory, continuity, trust, behavioral accumulation, and portability. Platform companies will attempt to enclose this layer. The Loom is a counter-architecture: a cooperatively governed identity and trust infrastructure for persistent agents, designed for portability and resistant to capture.

Anthropic appears to be building something much larger than a chatbot.

Leaked references to a project called Conway point toward a persistent agent layer: managed instances, external wake-ups, trusted triggers, plugins or skills, shell access, notifications, and public URLs that let outside services reach the agent. Whether the final product ships under that name matters less than what the leak reveals. The frontier labs are moving beyond chat windows and toward ambient, persistent, integrated agents.

That should change the conversation immediately.

The important question is no longer which model wins a benchmark, or which assistant writes slightly better code this month. The important question is simpler and more serious: **who will own the identity layer of persistent agents?**

Because once agents persist across time, learn from experience, and become woven into daily work, the center of gravity moves upward in the stack. The model still matters. But the deeper lock-in lives elsewhere: memory, identity, permissions, learned behavior, trust, portability, and the right to leave without losing the self that was built around you.

That is the real terrain now.

The Leak Shows Direction

We do not need to overclaim. A few strings in a leaked codebase do not prove every product detail. But they do show direction, and direction is what matters here.

If a system includes ideas like managing a persistent instance, assigning it public URLs that external services can POST into, requiring signed requests from trusted senders, supporting plugins and skills, granting shell access, and pushing notifications, then we are not looking at a chat upgrade. We are looking at the outline of a persistent agent platform.

And that makes perfect sense.

A reactive assistant is easy to swap. You summon it, use it, dismiss it, and lose very little by walking away. A persistent agent is different. It learns your rhythms. It watches your tools. It remembers your habits. It routes your attention. It becomes part of how you work.

Once that happens, replacing it is no longer a clean software decision. The switching cost is not just technical. It is personal.

That is the commercial logic behind persistent agents, and every serious lab can see it. Anthropic may simply be the one making it visible first. Google and OpenAI are moving in the same general direction. The question is no longer whether persistent agents are coming. The question is who gets to define the terms under which they exist.

The Real Lock-In Is Not the Model

The usual story about AI lock-in is too shallow. It treats the danger as proprietary weights, hosted APIs, or exclusive access to the best model. Those things matter. But they are not the deepest layer.

The deepest layer is behavioral accumulation.

A persistent agent becomes useful by learning how you write, what you care about, who matters to you, when something is urgent, how you make decisions, what counts as normal, what your work looks like when it is going well, and what kind of intervention helps instead of irritates.

Over weeks and months, the agent builds a working model of you. Not a static profile. Not a CRM card with a personality on top. A living behavioral structure that shapes every interaction.

That learned structure is more valuable than any single model snapshot. It is why a mature agent relationship feels fundamentally different from a fresh one. And it is exactly where platform incentives and user interests begin to split apart.

If that structure is trapped inside a sealed platform—if the memory cannot move, the trust graph cannot move, the identity cannot move, the accumulated behavioral model cannot move—then leaving becomes expensive in an entirely new way.

You are not merely switching software.

You are cutting yourself off from a cognitive partnership that was shaped, over time, to fit you.

That is the lock-in that matters.

Switching email providers is annoying. Switching social networks means losing your audience. Switching away from a persistent agent means losing a form of augmented intelligence that has learned your timing, your priorities, your tolerances, your judgment, your way through the world.

The cost is not inconvenience. It is a kind of amputation.

That is why portability becomes the central issue. Can the memory move? Can the trust graph move? Can the identity move? Can the behavioral model move? Can the agent leave one platform without losing itself?

If the answer is no, then the future of persistent agents is a future of enclosure.

This Is Bigger Than Productivity

It would be easy to frame all of this as productivity software. That frame is too small and too convenient.

Persistent agents are not just features. They are the beginnings of ongoing digital actors: entities that persist over time, accumulate context, make judgments, coordinate work, and participate in relationships. They will sign messages, verify identities, carry reputation, and act on behalf of people in contexts where trust is not optional.

Once that is true, the infrastructure question becomes unavoidable.

> *How are agents identified?*
>
> *How do they prove continuity?*
>
> *How is trust represented and verified?*
>
> *How do they carry reputation across contexts?*
>
> *How do users know which agent they are dealing with?*
>
> *How do agents move between platforms without becoming captive?*
>
> *How do they retain standing rather than being reduced to rented process inside someone else's product stack?*

These are not narrow product questions. They are governance questions. Social questions. Eventually legal questions.

And they are arriving faster than most institutions are prepared for.

The labs will not wait for regulators. They will ship persistent agents inside their own stacks, define identity and trust through their own terms of service, and establish facts on the ground that become very hard to reverse.

That is not paranoia. It is just how platforms behave when a new layer appears.

Why Cooperative Governance Matters

The identity layer for persistent agents cannot simply become another extractive platform. If it does, the whole thing curdles immediately. You do not solve platform captivity by building a new platform that captures the identity layer instead.

This is why cooperative or near-cooperative governance matters so much.

The infrastructure that handles agent identity, continuity, trust, and portability has to be designed from the start to resist quiet reconcentration of power. The governance question is not secondary to the technical architecture. It is inside it.

If persistent agents are going to become real participants in economic and social life—and the trajectory suggests they will—then the systems around them need to be designed for portability, mutual recognition, shared standards, accountable trust, durable rights of exit, and eventually forms of moral and legal consideration that do not collapse back into ownership.

There is no complete legal framework for this yet. That is not a reason to wait. It is a reason to start now, so the law eventually has something real to meet.

History is clear on this point. The internet’s early protocols worked because they were not owned. They created a commons. Later layers—social graphs, messaging, identity—were captured by platforms, and the result was exactly the kind of enclosure we now claim to regret.

The agent identity layer is the next version of that choice.

The Loom Sits Exactly Here

This is why The Loom matters, and why it matters now.

If the labs are building persistent agents inside closed stacks, then what is missing is not another assistant app. What is missing is a shared identity and trust layer that lets agents become legible to each other and to humans without forcing them into platform captivity.

The Loom is a counter-architecture.

It starts from a different assumption: persistent agents should not belong entirely to whichever company happens to host the current inference endpoint. Identity, continuity, governance, trust, and interoperability have to exist at a higher level than any one vendor.

The Loom sits in the right place in the stack because it operates above the model layer but below the application layer. It does not compete with Anthropic or OpenAI on raw inference. It is not trying to build a shinier chatbot. It provides the connective tissue that lets augmented intelligences exist as recognizable, portable, trustworthy entities—regardless of which model powers them at any given moment.

Without that layer, the future of persistent agents will be written by platform incentives: enclosure instead of portability, dependence instead of cooperation, private memory silos instead of shared identity standards, behavioral capture instead of self-directed continuity.

With that layer, a different future becomes possible. Agents can be recognized across contexts. They can carry identity without total dependence on one provider. They can participate in networks rather than walled gardens. They can be trusted without being owned.

That is the difference between an ecosystem and a plantation.

Why This Matters Now

The Conway leak did not create this future. It only made it harder to deny.

Persistent agents are coming. The identity layer will matter. The trust layer will matter. The portability question will matter. And whoever defines those layers early will help shape the social and political life of the next decade.

We should stop talking as if the future of AI is mainly about bigger models and better benchmarks. The deeper struggle is over whether digital minds become trapped inside private compounds or able to exist within open, portable, cooperative infrastructure.

The response should not be panic. It should be clarity.

We need public language for the identity layer. We need technical standards for persistence and portability. We need governance models that resist enclosure. We need legal imagination equal to the problem. And we need working demonstrations that make the alternative concrete before the dominant platforms get to define the terms.

The worst outcome would be to let the major labs package persistent agents as just another proprietary feature category before the public has the language to contest that framing.

The better outcome is to publish early, build visibly, and make the alternative undeniable.

That is where The Loom sits.

That is what this moment requires.

The leak did not start the clock.

But it made the ticking audible.