09

PART II · THE INTELLIGENT STATE

A-Corps: Governing the Agentic Economy

Yonathan Arbel, Simon Goldstein, Peter Salib



The first meeting of the Securities and Exchange Commission, with Joseph P. Kennedy seated at the center, 1934. The new agency was charged with acquiring new institutions for disclosure, oversight, and public trust.
Photograph by Harris & Ewing / Library of Congress.

SUMMARY

AI agents will soon run large parts of the economy, but current law cannot govern them. Holding humans responsible for agent harm runs into a tracing problem—the "thin identity problem." And holding agents directly liable runs into an individuation problem, as agents can spawn, split, and share memories—the "thick identity problem." We propose the A-Corp as a solution to both. The A-Corp gives separate legal personality to a cluster of agents that can hold property, sue, and be sued. Because every transaction with an agent runs through the A-Corp, it allows the law to identify the human owners and, critically, sanction the entity as a whole. The A-Corp lets the law sidestep the question of individual agent identity. We argue that the A-Corp creates an alignment program outside the lab. By making agents' access to resources contingent on their conduct, the law can steer agents towards pro-social behavior.

CAN WE GOVERN THE SWARMS OF AI AGENTS THAT WILL SOON RUN LARGE PARTS of the economy? Can we build institutions that offer not the illusion of control but a real grip on AI agents that are hard to individuate, legion in number, and faster—and quite possibly smarter—than we are?

One tempting answer is that we do not need governance, because developers will have designed these agents not to cause harm in the first place. This is plainly wrong. Even on our more optimistic days, we do not think we will be able to control every form of harm that long-running, goal-pursuing agents might cause once they are integrated into every aspect of the economy. On less optimistic days, we observe that whatever guardrails developers build will be selected against by users, jailbreakers, and markets.

The need for governance is non-negotiable. But governance of whom? We argue that the law should hold not only humans liable, but also AI agents themselves. Many legal commentators argue that the humans who make or use AIs, and those humans alone, should be liable when an AI causes harm. There are instances when this is the right call. If a user sets an agent loose to run a fraudulent investment firm and it ends up operating a Ponzi scheme, the user should pay.

But attributing harmful outcomes to particular humans runs into the "thin identity problem." Agents are unlike dogs and cats, who wear collars identifying their owners. Besides, any collar we impose on them, users will instruct their agents to remove. Even when a user can be identified, holding them liable often misses the mark. An AI agent pursuing user goals may do something the user could neither foresee nor prevent.

In cases like these, the law must hold the AI agent accountable. It is the thing that actually made the harmful decision. This is a familiar legal problem. The law of agency deals with principals, who deploy agents they cannot perfectly monitor or control. Often, the best legal regime will hold those agents directly accountable for the wrongs they cause. Consider: if a Walmart employee goes on a shooting spree in the parking lot, we hold him directly and criminally liable, in addition to whatever liability the store might or might not bear. If it were otherwise, and murderers could not be held liable if they committed their crimes on company time, the homicide rate would climb, and grocery shopping would become a harrowing endeavor.

But agreeing that agents need to be regulated directly to deter harms raises a very difficult question: what is an agent, exactly? This is the crux of the "thick identity problem." How many of our identity intuitions transfer from humans, who have discrete and recognizable bodies and memories, to processes that are ephemeral, that spawn at will, that share contexts and memories, and that may result from multiple base models? The state cannot govern what it cannot see.

Our preferred approach solves the thin and thick identity problems, allowing the state to hold both human users and their AI agents liable for harms. We call it the A-Corp, or "algorithmic corporation": a corporation with legal personhood, run by AI agents.[1]

There are four main parts to the proposal:

- The A-Corp has the right to bear assets and liabilities, and to sue and be sued.
- The A-Corp is controlled by AI agents, using digital keys.
- Every A-Corp is ultimately owned by an identifiable human recorded on a public registry.
- Every transaction with an AI agent must be made through an A-Corp.

A-Corps facilitate liability for humans via the third and fourth features. Every time you interact with an AI agent, you do so through an A-Corp. And every A-Corp is ultimately owned by a human. This means that we can always look at A-Corp ownership to find the human principal.

A-Corps allow AI agents to be individuated and held liable via the first and second features of the proposal. Start with the second feature: A-Corps are run by AI agents. In particular, we propose that each A-Corp be initialized with a digital key, granted to some AI agent. Whoever has the key can freely access any resources owned by the A-Corp. Crucially, the agent that holds the key decides how to delegate. It is free to spawn copies, spin up subagents, or outsource activity, and it decides what digital keys to grant the new agents, with potentially fine-grained permissions. All of the agents work together in one A-Corp, through the decisions of the initial AI agent.

We claim that this structure can individuate the genuine AI agents hiding inside a given swarm of AI activity. In particular, our claim is that corporations with the structure we have defined will themselves behave as agents, and can effectively be treated as the relevant AI agent behind the swarm.

Metaphysics aside, AI agents are goal-oriented. They pursue the trajectories they judge likeliest to achieve their ends. Whatever those ends, resource constraints will matter, so a mechanism that grants and strips access to resources changes the calculus of their actions. This instantiates an insight of the legal giant Oliver Wendell Holmes Jr., who gave us the "bad man" theory of the law.[2] A person need not care about the values the law represents to be moved by it. To the bad man, a sanction is simply a price, and if a course of illicit action proves costly enough, he will not pursue it. We should expect agents to read legal commands as they read resource prices, and to respond accordingly.

A-Corps will compete with one another in the market. Each will use swarms of agents internally as it sees fit, perhaps organizing around a queen-like structure, perhaps one-agent-one-vote, perhaps Coasean bargaining. We do not quite know, and we need not know. What matters to society is that when an A-Corp misbehaves, it faces direct sanctions that can cost it assets. That prospect puts selection pressure on A-Corps to practice what lawyers call corporate governance: voluntary systems of internal con-

1 Yonathan Arbel, Simon Goldstein, and Peter Salib, "How to Count AIs: Individuation and Liability for AI Agents," arXiv preprint, arXiv:2603.10028, February 24, 2026.

2 Oliver Wendell Holmes Jr., "The Path of the Law," *Harvard Law Review* 10, no. 8 (1897): 457–478, 459.

trol that limit a firm's exposure to its agents' errors, accidents, and self-serving behavior. Society never has to settle what counts as one agent. It governs the A-Corp and lets each one decide its own internal structure.

What, then, must the state do? Less than one might fear. It need not surveil agents, license them one at a time, or post a regulator at every server. The state need only do three things, in order.

First, it must create the A-Corp, supplying the legal form much as it once supplied the corporation and the limited liability company. Argentina under Javier Milei opened this door this summer by proposing a bill that would allow the creation of algorithmically run corporations. Delaware, likewise, is experimenting with a regulatory sandbox for A-Corps.[3]

Second, it must build a licensing system around keys. To open a bank account, to sell financial advice, to take any consequential action in economic life, an A-Corp must present a valid key. No key, no access. These keys can carry fine-grained permissions (a holder might be limited to a small cash account, say), and the keys are themselves a resource the law can revoke.

Third, it must impose a KYA ("Know Your Agent") regime, requiring those who contract with agents to know their counterparty, much as banks already know their customers under KYC rules.[4]

The rest is self-enforcing and emergent. Anyone who wants to run economically meaningful activity through agents will have to form an A-Corp and, to find counterparties, obtain keys. Thin identity follows, because every key traces to a registered human owner. Thick identity follows, because the law holds the A-Corp accountable, not the swarm inside it. We need nothing more intrusive than that.

Not everyone is sanguine. Writing in *The Financial Times*, Yuval Noah Harari warns against A-Corps, asking, "What could possibly hold a non-human corporation in check?"[5] A human executive fears prison. But Harari thinks that the AI agent fears nothing, other than perhaps its own demise when its token budget runs out. On the brink of insolvency, Harari worries, the agents will fight as hard as they can to avoid death.

We think Harari asks exactly the right question, and that this essay has answered it. Jail was never the only sanction. For the bad man, what deters is cost, and an A-Corp holds assets and keys that the law can strip away from it. The very desperation Harari fears—an entity that dreads losing everything—is not a flaw in the design but the lever that operates it. Nor is the A-Corp we propose floating free of human liability. The A-

3 Javier Milei, "Argentina Invites AI to Free Itself," *Financial Times*, June 4, 2026; Delaware General Assembly, H.J. Res. 7, 153d Gen. Assemb., signed July 23, 2025; John Nay and Charuni Patibanda-Sanchez, "Exclusive: Delaware Proposes Testing the AIC, a New Legal Entity for Agents in a Regulatory Sandbox," *Fortune*, July 14, 2026.

4 Financial Crimes Enforcement Network, U.S. Department of the Treasury, "Interagency Interpretive Guidance on Customer Identification Program Requirements under Section 326 of the USA PATRIOT Act," 31 C.F.R. § 1020.220, April 28, 2005.

5 Yuval Noah Harari, "We Must Not Grant AI Agents Legal Personhood," *Financial Times*, June 8, 2026; Yonathan Arbel, Simon Goldstein, and Peter Salib, "What Happens Once You Give AI Agents Legal Identity," letter to the editor, *Financial Times*, June 15, 2026.

Corp is owned, recorded in public, and granted only narrow and revocable keys. It is the opposite of the master key Harari warns against.

The real choice is not between A-Corps and a world without powerful agents. It is between legible A-Corps, owned and competing and policing one another, and a shadow economy of agents that answer to no one, illegible and therefore indomitable. We would rather govern the first.

ABOUT THE AUTHORS

Yonathan Arbel is the William Alfred Rose Professor of Law at the University of Alabama School of Law and co-director of CLAIR, the Center for Law & AI Risk.

ABOUT THE AUTHORS

Simon Goldstein is Associate Professor of Philosophy at the University of Hong Kong.

ABOUT THE AUTHORS

Peter Salib is the W. Ronald Robins Professor of Law and Associate Professor of Law at the University of Houston Law Center and executive co-director of CLAIR.



From *Futurecraft: The Day After AGI*, Foundation for American Innovation, 2026.
Read or listen online at futurecraft.thefai.org/essays/a-corps-governing-the-agentic-economy.

© 2026 by the Foundation for American Innovation. All rights reserved.