

Authority by Consent

Why the legitimacy of autonomous AI depends on consent and revocation — and why architecture, not statute, is the stronger mechanism for governing it.

The prevailing frame for governing artificial intelligence is regulatory: rules, audits, and compliance imposed on systems from the outside. This paper argues for a different and more durable mechanism — legitimacy built into the architecture of autonomous systems, grounded in the oldest principle of self-government: that just authority derives from the consent of the governed, and remains legitimate only so long as the governed may withdraw it. Systems whose authority is consent-derived and revocable are structurally democratic. Systems whose authority derives from capability and ownership are structurally monarchical, however distributed their infrastructure may appear. The distinction is not rhetorical. It is the property that decides whether the coming decade concentrates power or distributes it.

THE CENTRALIZING DEFAULT

Artificial intelligence is among the most centralizing technologies ever built. Compute, capital, frontier models, and training data concentrate in a small number of firms and states, and the advantages compound. The default trajectory therefore bends not toward liberation but toward a new authoritarianism — one in which authority is derived from capability and ownership, where whoever holds the model decides what its agents may do. This is monarchical in structure: power flows from possession, not from consent. The familiar optimism that AI will deliver a flatter, freer world mistakes a hope for a trend. A less authoritarian order is not something the technology delivers on its own. It is something that must be built, against the gravity of concentration.

DISTRIBUTION IS NOT DEMOCRACY

The reflexive remedy for centralization is decentralization, and it is insufficient. Distributed is not the same as democratic. “Code is law” cuts in both directions: a protocol that answers to no one is not accountable, and immutability becomes its own form of tyranny — a sovereign that cannot be petitioned because it cannot be changed. The recent history of decentralization is, in large part, a record of recentralization: governance captured by those who held the most tokens, protocols controlled by the few who understood them, autonomous organizations that collapsed the moment their frozen rules met a circumstance their authors had not foreseen. Removing human authority from a system does not democratize it. It can simply install a more rigid ruler in the human’s place. Distribution without consent is not freedom.

“Distribution without consent is not freedom —
it is rule by code that no one elected and no one can
recall.”

THE DECIDING PROPERTY: CONSENT AND REVOCATION

What separates a democratic distributed system from an absolutist one is a single pairing. The governed must be able to grant authority, and the governed must be able to withdraw it. Consent is where legitimate authority begins. Revocation is how the governed keep it. This is the Lockean structure — government by the consent of the governed, who retain the right to alter or abolish — expressed in code rather than in a constitution. An autonomous agent acting under an explicit, human-granted authorization that can be narrowed, suspended, or revoked at any moment, without disrupting the system around it, is an agent operating with democratic legitimacy. An agent acting on authority inherited from whoever owns it is not — regardless of how widely its computation is distributed. The substrate is not the question. The source of authority is.

GOVERNANCE BY ARCHITECTURE

This reframes what it means to “regulate” AI. Statute governs from the outside, after the fact, and is structurally condemned to lag the technology it addresses; by the time a behavior is named and prohibited, the systems have moved. Architecture governs from the inside, continuously, at the moment of action. In a system where every consequential act by an automated agent is bound to an explicit agreement — granted by a human principal, enforced in code, recorded, and revocable — compliance is not verified after harm by an external auditor. The constraint is structural rather than procedural. The rule is not written about the system; it is written into it.

For policymakers, this points toward a more effective intervention than attempting to enumerate permitted and forbidden machine behaviors — a task that ages badly and never keeps pace. The durable move is to regulate the property rather than the conduct: to require that autonomous systems be built so that their authority is consent-derived and revocable. A mandate of revocability is enforceable, technology-neutral, and stable across generations of models in a way that a catalog of prohibited behaviors can never be. It does not ask the regulator to predict what the machine will do. It asks only that the human who must live with the consequences retain the power to say no — and to stop.

THE CHOICE

Artificial intelligence does not, on its own, move the world toward less concentrated power. AI built on consent and revocation moves in that direction. AI built on capability and ownership moves the opposite way. These are not competing forecasts about a single arc of history. They are two

available architectures, and the present gravity of the industry pulls toward the second. The systems being deployed now are casting that choice into code — for the most part without naming it, and therefore without deciding it deliberately.

The argument here is that the choice should be named and made on purpose, and that the instrument of choosing is architectural before it is regulatory. The consent of the governed is not a phrase inherited from the eighteenth century and best left there. It is a design specification for the twenty-first — and the systems that honor it will be the ones that keep authority where legitimate authority has always belonged: with those who may withdraw it.

A Kokomo Systems white paper. Kokomo Systems builds execution-governance infrastructure in which the authority of automated agents is consent-derived, enforced in code, and revocable without disruption. This paper concerns the principle; implementations are described separately.