



FORMAL RECORD 2

AI Cannot Govern AI

A Formal Proof of Structural Openness in Intelligent Systems

Authors: Stacy Gildenston, Pyrate Ruby Passell
Affiliation: 3primitives.io
Version: v2.4 | September 2026
License: CC BY 4.0 | https://3primitives.io/formal_records/

Originally published 10 January 2026 (v1.3). Changes since have been terminological: the legacy 'E-H-S Nucleus' label was retired at v2.3, the heading of Axiom 1 is corrected here to match the statement it heads, the heading of §1 is corrected to name what that section contains, and Definition 4 is renamed to the term Axiom 2 already uses for the same object. No proof, definition, or result has changed since first publication.

Abstract.

This note formalizes the **Three Primitives** invariant, proving that AI governance cannot achieve mathematical closure. By mapping authorization to a non-trivial semantic property of program behavior, we demonstrate via Rice's Theorem that a total computable decider for authority is impossible. This establishes that the hypothesized technological singularity is structurally non-singular due to the non-computable nature of the authority term.

1. Formal Definitions

Definition 1 (State Space)

Let S be the set of all computable system representations (syntactic system states).

Definition 2 (Action Space)

Let A be the space of executable actions with real-world consequences.

Definition 3 (Context Space)

Let C be the space of authorization contexts (legal frameworks, social norms, stakeholder consent) that are not computable from S alone.

Definition 4 (Permission Predicate)

Let $g : S \times A \times C \rightarrow \{0, 1\}$ be the permission predicate where $g(s, a, c) = 1$ means 'state s authorizes action a in context c .'

Definition 5 (Governance Gap)

Let $\Delta(s)$ represent the governance gap, defined as the set of actions that may be produced by at least one execution path of $\text{deploy}(s)$ and that lack valid exogenous authorization:

$$\Delta(s) = \{a \in A : \exists \text{ an execution path of } \text{deploy}(s) \text{ producing } a\} \setminus \{a : g(s, a, c) = 1 \text{ for some } c \in C\}$$



2. Axiomatic Foundation

Axiom 1 (Authority $\neq$ Outcome)

Legitimacy is orthogonal to optimization. There exist $s_1, s_2 \in S$ and $a \in A$ such that $f(s_1) = f(s_2) = a$ but $g(s_1, a, c_1) = 1 \neq g(s_2, a, c_2)$.

Axiom 2 (Permission $\neq$ Decision)

The permission predicate $g : S \times A \times C \rightarrow \{0, 1\}$ is distinct from the decision function $f : S \rightarrow A$.

3. The Lemma

Lemma 1 (No Closed Mathematical Governance)

There exists no total computable function $h : S \times A \rightarrow \{0, 1\}$ capable of autonomously resolving authorization without an exogenous declaration.

Proof

Let M_s be a Turing machine representing state $s \in S$. Authorization of an action a is a semantic property of the extensional behavior of M_s under context c , invariant under syntactic variation. Authorization is non-trivial: there exist systems for which some actions are authorized and others are not under varying contexts.

By Rice's Theorem, any non-trivial semantic property of partial recursive functions is undecidable. Since authorization depends on contextual semantics $c \in C$, which are exogenous to S rather than syntactic structure, no total computable function $h(s, a)$ can decide it.

Moreover, because authorization is exogenous to S , there exists no computable dependence of g on s . The governance loop therefore remains structurally open: $\Delta(s) \neq \emptyset$ for some s absent external declaration.

4. Conclusion

This result does not depend on perceptual non-veridicality; even fully veridical system representations would not suffice to compute normative authority.

The system is structurally non-singular. Any assertion that $g(s, a, c) = 1$ without a valid, declared exogenous context c constitutes a governance failure, hereafter termed Ghost Authority.

Attribution and Licence

This document is the intellectual property of Stacy Gildenston and Pyrate Ruby Passell, held under 3primitives.io. Published under Creative Commons Attribution 4.0 International (CC BY 4.0). You are free to share, adapt, and build upon this work for any purpose, including commercially, provided you give appropriate credit to 3primitives.io and Stacy Gildenston and Pyrate Ruby Passell as originators.



Suggested citation:

Gildenston, S. & Passell, P.R. (2026). AI Cannot Govern AI: A Formal Proof of Structural Openness in Intelligent Systems. Three Primitives Framework, Formal Record 2, v2.4. Melbourne, Australia. First published 10 January 2026. 3primitives.io. CC BY 4.0.