# Insurability and the Architecture of Autonomous Systems

## Why Deployable AI Requires a Governance Plane

**J. D. “Pepper” Petersen**
Aristotle Agentic
Helena, Montana, USA

---

# Abstract

Autonomous AI systems are moving from research environments into the operational foundations of real infrastructure. These systems increasingly perform actions rather than merely advising human operators. They route network traffic, coordinate machines, allocate resources, and participate in decision loops that often operate faster than human intervention allows.

Technologies that introduce systemic risk rarely scale until institutions develop mechanisms to measure and manage that risk. In many sectors, the practical mechanism through which this occurs is the insurance market. Systems that cannot demonstrate bounded authority, enforceable constraints, and reconstructable decision histories are difficult to insure.

This paper proposes a Governance Plane architecture that embeds authority verification, operational constraints, and decision traceability directly into system execution. By making autonomous decisions legible to institutions responsible for liability and underwriting, the governance plane provides the architectural foundation required for autonomous systems to become insurable and therefore deployable.

---

# I. Introduction: When Systems Begin Acting

Artificial intelligence is moving from analysis into action.

Early AI systems primarily supported human decision making by ranking information, summarizing documents, and recommending possible actions. Humans remained responsible for executing decisions and accepting the associated risks.

That boundary is shifting.

Autonomous systems are now embedded directly in operational environments. Logistics platforms automatically reroute cargo across global supply chains. Telecommunications networks dynamically reallocate spectrum and traffic. Industrial systems coordinate machines without human supervision.

When systems begin acting rather than advising, governance questions change fundamentally. The primary concern is no longer simply whether the model produced a reasonable output. Instead the question becomes whether the system had the authority to act and whether the consequences of that action can be reconstructed.

In infrastructure environments where decisions propagate rapidly across interconnected systems, the absence of clear authority boundaries creates institutional risk.

---

# II. Insurability as a Deployment Constraint

Insurance markets have historically played a decisive role in determining whether new technologies can scale.

Aviation, nuclear power, maritime transport, and large scale industrial systems all developed institutional governance structures partly because insurers required clear mechanisms for assessing risk. Insurers must be able to answer several practical questions before underwriting a system.

First, what failure modes exist within the system.

Second, how severe those failures could become.

Third, which institution bears responsibility when failures occur.

Fourth, whether investigators can reconstruct events after an incident.

If these questions cannot be answered clearly, risk becomes opaque. When risk is opaque, insurers cannot reliably price coverage. Without insurance, operators and investors become reluctant to deploy the technology at scale.

Autonomous systems introduce a new category of operational risk because the system itself participates in decision making. If autonomous actions cannot be attributed to clearly bounded authority structures, determining liability becomes extremely difficult.

This creates a structural barrier to deployment.

---

# III. The Governance Gap

Current approaches to AI governance often focus on model safety, evaluation benchmarks, and alignment research. These efforts attempt to ensure that models produce outputs that are consistent with intended goals and ethical guidelines.

While important, many of these approaches operate outside the execution path of the system itself.

Autonomous systems operating in infrastructure environments must make decisions continuously and often under conditions of partial connectivity. Governance mechanisms that rely solely on post hoc monitoring or policy documents cannot prevent invalid actions in real time.

A governance gap emerges when systems are capable of acting but lack architectural mechanisms that verify authority and constraints before execution.

Bridging this gap requires moving governance from an external supervisory function into the architecture of the system itself.

---

# IV. Governance Plane Architecture

The Governance Plane is an architectural layer that sits between autonomous agents and the infrastructure they influence. Its role is to evaluate authority and constraints before system actions propagate into operational environments.

Rather than relying on retrospective oversight, the governance plane functions as a runtime validation layer that enforces institutional policy during execution.

The governance plane is composed of several functional components.

**Authority Orchestration Function (AOF)**
Translates institutional policy, regulatory constraints, and operator intent into machine interpretable authority definitions.

**Runtime Constraint Monitor (RCM)**
Evaluates proposed actions against active authority envelopes and operational constraints.

**Escalation Logic Controller (ELC)**
Manages transitions into safe states when constraint violations or uncertainty occur.

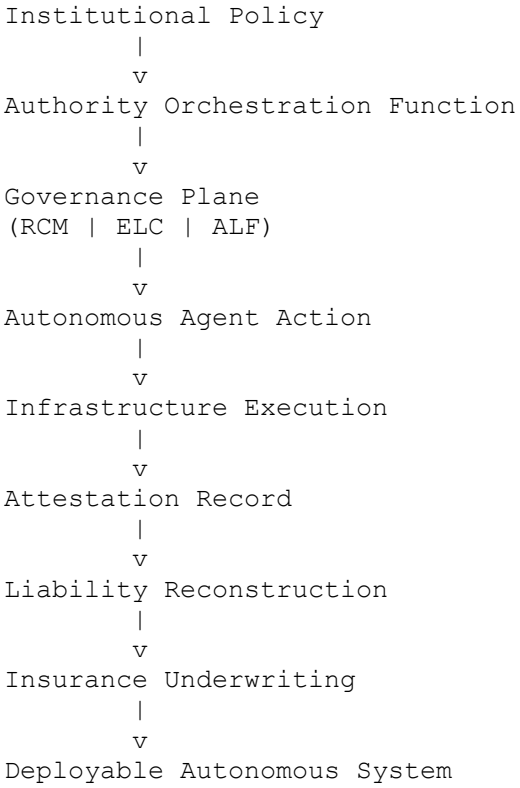
**Attestation and Logging Function (ALF)**
Maintains tamper evident records of decisions and system state.

Together these components create an execution environment in which autonomous actions remain bounded, attributable, and auditable.

---

# Figure 1

## Governance Plane and Insurability Pathway




---

# V. Authority Envelopes

Authority envelopes define the operational boundaries within which an autonomous system may act.

An authority envelope is a machine readable structure that encodes permissible actions, operational constraints, and identity bindings linking the system to an accountable institution.

| Component | Description | Enforcement |
|---|---|---|
| Permissible Action Space | Mission parameters such as routing scope | Runtime Constraint Monitor |
| Physical Invariants | Hard limits such as geofencing or power caps | Hardware enforcement |
| Safety Fallback States | Hold safe or degraded autonomy modes | Escalation Logic Controller |
| Credential Binding | Identity linkage to Model Lineage Certificate | Cryptographic verification |

Authority envelopes ensure that autonomous actions remain constrained within predefined institutional boundaries.

---

# VI. Authority Validation in Operation

Before executing an action, the autonomous system proposes the action to the governance plane.

The governance layer evaluates the action against the current authority envelope and operational constraints.

If the action satisfies all conditions, execution proceeds.

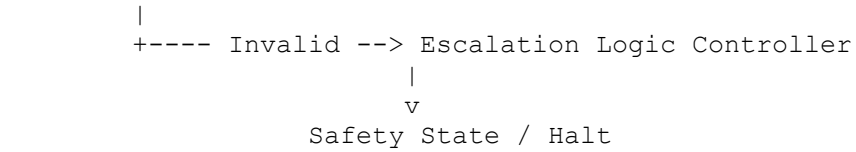
If the action violates authority limits or operational constraints, the system halts execution and transitions into an escalation state.

---

# Figure 2

## Authority Envelope Validation Flow

```
Agent proposes action
        |
        v
Governance Plane validates authority
        |
        +---- Valid ----> Execute Action
```




---

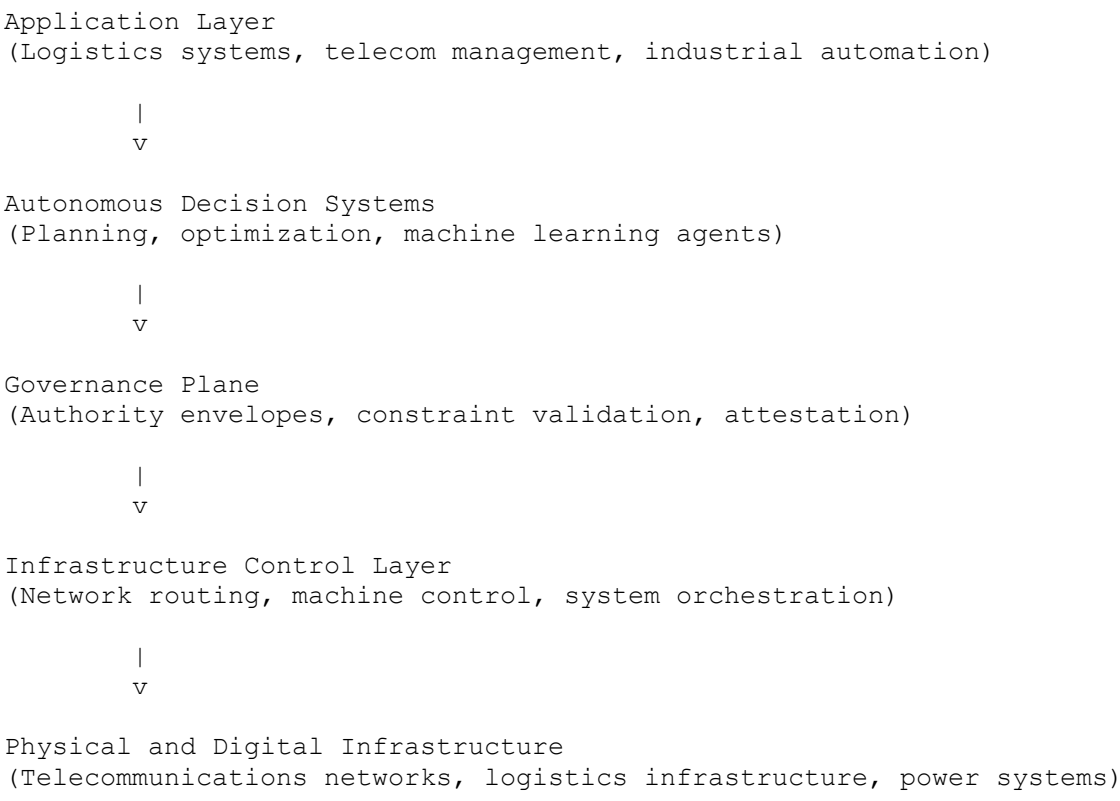
# VII. Governance in the Autonomous Infrastructure Stack

Autonomous governance mechanisms operate within a broader infrastructure stack. Governance must interact with system planning layers above it and operational infrastructure below it.

---

# Figure 3

## Autonomous Governance Stack




---

# VIII. Insurability and Institutional Trust

The architectural mechanisms described above provide more than technical safety. They also create institutional visibility into how autonomous systems behave.

Because every action is associated with a specific authority envelope and recorded through attestation logs, investigators can reconstruct the sequence of events that produced a given outcome.

This capability allows insurers and regulators to evaluate autonomous systems using the same principles applied to other infrastructure technologies.

Risk becomes legible.

When institutions can evaluate risk, they can price it. When risk can be priced, systems become insurable.

In this sense governance architecture becomes the bridge between technical capability and institutional deployment.

---

# IX. Conclusion

Autonomous systems introduce operational risks that traditional governance mechanisms struggle to manage.

Embedding governance directly into system architecture provides bounded authority, enforceable constraints, and reconstructable decision histories.

These properties transform autonomous systems from opaque decision engines into accountable infrastructure components.

When risk becomes legible, institutions can evaluate and insure autonomous systems.

When autonomous systems become insurable, they become deployable.

---

# References

Petersen, J. D.
A Governance Plane for AI Native 6G Architectures Under Intermittent Connectivity. 2026.

National Institute of Standards and Technology
Artificial Intelligence Risk Management Framework 1.0. 2023.

OECD
OECD Principles on Artificial Intelligence. 2019.

European Union
Artificial Intelligence Act. 2024.

ISO/IEC
ISO/IEC 42001 Artificial Intelligence Management Systems. 2023.

Russell, S.
Human Compatible: Artificial Intelligence and the Problem of Control. 2019.

Amodei, D. et al.
Concrete Problems in AI Safety. 2016.