

# Governance-Configurable Harm Minimization

by [Nick Clark](#) | Published April 25, 2026

## What Configurable Harm Ordering Specifies

The architecture treats harm minimization as a structural concern with three components: an entity-class taxonomy (pedestrians, cyclists, occupants, property, the unit itself, plus domain-specific extensions like patients, combatants, non-combatants, infrastructure), a relative weighting between classes (the harm ordering itself), and a credentialing chain that signs the weighting (the regulatory or jurisdictional authority).

When the autonomous system faces a no-good-options scenario — every available actuation produces some harm — the harm-minimization computation runs against the credentialed ordering and selects the actuation with minimum weighted harm. Every selection is recorded in lineage with the ordering policy under which it was evaluated, supporting forensic reconstruction of any decision.

## Why Externalization Is the Architectural Question

Current autonomous-system architectures handle harm ordering in one of two ways: hardcoded by the manufacturer (Waymo, Cruise, every L4 stack), or refused articulation with the implicit claim that the situation will not arise (much of the L2/L3 industry, defense autonomy at the edge of LAWS governance).

Both patterns produce structural failure. Hardcoded ordering puts ethical decision authority in the manufacturer rather than the jurisdiction. Refused articulation produces unallocated liability when the inevitable situation arises. The architectural answer — externalizing the ordering through credentialed governance policy signed by the jurisdiction — fits how every other regulated transportation domain (aviation, rail, shipping, maritime) actually allocates ethical authority.

## **How the Credentialing Chain Operates**

For autonomous vehicles, the credentialing chain runs through state DOTs (signing the ordering applicable to their territory), federal regulators (NHTSA for safety standards, FAA for airspace, FRA for rail), and possibly cross-jurisdictional bodies for interstate corridors. For medical autonomy, the chain runs through FDA, hospital ethics boards, and clinical-procedure authorities. For defense autonomy, the chain descends from national command authority through theater command through mission ROE.

Each level signs within its scope. The autonomous platform consumes the composite policy through composite admissibility. Cross-jurisdictional operation handles transitions through cross-authority cross-recognition: a vehicle entering a new jurisdiction consumes the local authority's ordering policy and adjusts harm-minimization accordingly. The mechanism is the same across domains; the configurations differ.

## **What This Enables for Liability Allocation**

The trolley-problem framing has been treated as a philosophical edge case for a decade. It is in fact a routine liability allocation question that every L4 vehicle, every autonomous medical system, and every autonomous-defense platform makes thousands of times per operating period. The allocation question — when something

happens, who is liable — is the gating factor for commercial deployment in regulated domains.

Externalized harm ordering shifts the allocation. The jurisdictional authority signs the ordering and bears liability for the policy. The manufacturer is liable for executing the signed policy correctly. The operator is liable for operating within the policy's authorized scope. This is the allocation pattern that aviation, rail, and maritime operate under, and the architecture that lets autonomous systems join those regulatory frameworks. The patent positions the primitive that the allocation requires.