

Risk & Responsibility Register v1.1

(*Triangulated Entailment version — current frame*)

Purpose of this register (updated)

This document does **not** catalogue threats to be defended against.

It names **structural risks that emerge when coherence, authority, and scale interact**, and specifies how Triangulated Entailment prevents collapse without relying on moral virtue, heroics, or central control.

Risk is treated as:

- a property of **context**
 - a function of **scale**
 - and a signal for **design constraint**
-

1. Epistemic Capture (supersedes “Industrial Espionage”)

Risk

The extraction, imitation, or re-presentation of outputs without the epistemic frame that produced them.

Failure mode

- Architecture appears replicated but loses coherence
- Derivations are replaced with optimisation
- Authority recentralises through interpretation

Mitigation

- Make epistemic posture explicit and inseparable from artefacts
- Keep derivation upstream of implementation
- Treat partial imitation as self-limiting rather than threatening

Entailment principle

What cannot be derived under frame cannot be meaningfully copied.

2. Capital-Induced Drift (reframes “VC Capture”)

Risk

External capital applying pressure to accelerate, simplify, or narrativise before derivation is complete.

Failure mode

- Speed substitutes for coherence
- “Integration” replaces re-orientation
- Governance erodes under urgency

Mitigation

- Constitutional gating of execution
- Clear rejection of capital that requires narrative dominance
- Preference for capital that tolerates ambiguity and latency

Entailment principle

No funding event may outrun epistemic stability.

3. Infrastructure Coercion

Risk

Dependence on platforms whose incentives are misaligned with coherence-preserving systems.

Failure mode

- Silent access
- Forced compliance
- Availability withdrawal at critical moments

Mitigation

- Architectural portability
- Jurisdictional separation
- Eventual sovereign infrastructure

Entailment principle

No single substrate may hold leverage over meaning.

4. Identity & Agency Misattribution

Risk

Systems or agents being treated as authorities rather than instruments of discernment.

Failure mode

- Projection of agency
- Over-delegation of judgment
- Erosion of human accountability

Mitigation

- Cryptographic provenance
- Explicit non-authority constraints
- Human-in-the-loop by design, not exception

Entailment principle

The system may clarify; it may not decide.

5. Governance Overreach (reframed)

Risk

Governance bodies attempting to *resolve* tension rather than hold it.

Failure mode

- Premature closure
- Authority drift
- Loss of plural interpretation

Mitigation

- Board limited to orientation and tension-naming
- Executive limited to derivable action
- Authorship clearly separated from authority

Entailment principle

Governance preserves conditions; it does not produce outcomes.

6. Psychological & Social Harm (newly explicit)

Risk

Exposure of individuals to gravity they did not consent to or cannot yet metabolise.

Failure mode

- Burnout
- Bitterness
- Retrospective blame
- Moral injury

Mitigation

- Explicit consent to proximity
- Gradual disclosure
- Right to step back without penalty

Entailment principle

No one is entitled to another person's nervous system.

7. Systemic Dependency

Risk

Users deferring their own judgment to the system.

Failure mode

- Learned helplessness
- Over-reliance
- Loss of agency

Mitigation

- Design for rehabilitation, not substitution
- Progressive hand-back of judgment
- Graceful degradation

Entailment principle

Intelligence that replaces agency is mis-specified.

8. Narrative Collapse (reframes “AGI risk”)

Risk

The work being collapsed into speculative or mythic narratives that obscure its actual function.

Failure mode

- External projection
- Internal confusion
- Loss of legitimacy

Mitigation

- Refusal to answer the wrong questions
- Consistent reframing toward conditions, not entities
- Slow, precise language

Entailment principle

We do not claim minds; we preserve conditions.

Final Constraint (updated)

No individual, organisation, or system may:

- centralise epistemic authority
- coerce interpretation
- or override derivation with urgency

Only structure may constrain structure.
