Transdisciplinary Reconfiguration as a Necessity for Contemporary Complex Systems

Version 2.0 — Protocol Activation Release

Yuji Marutani (MARUTANI) DOI (v1): https://doi.org/10.5281/zenodo.18103015

Abstract

Version 2.0 represents the activation phase of this project. While Version 1 established the theoretical necessity of transdisciplinary reconfiguration for contemporary complex systems, Version 2 formalizes and executes this framework as an operational protocol.

This release introduces the *Mirror Function*, a protocol that interprets AI hallucinations not as model errors but as epistemic signals reflecting pre-existing distortions within human socio-linguistic, institutional, and geopolitical structures. Hallucinations are treated as observable phenomena through which collective assumptions and narrative overfitting become empirically accessible.

Version 2 integrates the first complete execution loop of this protocol: a documented hallucination event, its classification through an explicit failure-mode taxonomy, and a corresponding protocol update proposal recorded as a versioned execution log. This loop is implemented through distributed peer commentary mechanisms, establishing distributed science as a methodological requirement rather than an ideological preference.

Version 2 should be read not as a completed theory, but as a running system: a forkable, versioned, and evolvable protocol for transdisciplinary knowledge production under conditions of extreme complexity.

1. Introduction

Contemporary socio-technical systems—spanning artificial intelligence, quantum information, governance, and global risk—exhibit levels of complexity that exceed the explanatory capacity of single disciplinary frameworks. Version 1 of this project argued that transdisciplinary reconfiguration is not optional but structurally necessary.

Version 2 shifts the project from conceptual argumentation to operational execution. Rather than proposing a static theory, this work introduces a protocol-driven framework in which epistemic failure itself becomes an object of systematic observation.

Central to this shift is the decision to treat AI hallucinations not as implementation defects, but as mirrors of collective human epistemology.

2. The Mirror Function

2.1 Conceptual Definition

The Mirror Function is defined as a protocol that maps AI hallucination events to underlying human epistemic distortions. The primary object of analysis is not the model, but the socio-linguistic and institutional structures reflected in the model's outputs.

In this framework, hallucinations are signals rather than errors.

2.2 Formal I/O Model

Input:

- Hallucination Event (documented AI output with contextual prompt)

Intermediate Representation:

- Extracted Epistemic Signal
- Failure Mode Classification

Output:

- Protocol Update Candidate
- Governance or interpretive adjustment proposal

2.3 Protocol Logic

1. Observation of hallucination
2. Human-side reflection
3. Classification via shared taxonomy
4. Protocol-level interpretation
5. Versioned update proposal

This logic is intentionally minimal and fork-compatible.

---

3. DevRealityOps Execution Model

3.1 Minimum Execution Loop (MEL)

Observation → Reflection → Aggregation → Reconfiguration → Deployment → Versioning

This loop constitutes the smallest unit of scientific execution under the protocol.

3.2 Distributed Peer Commentary

- **Issues:** Observation nodes
- **Discussions:** Sensemaking nodes
- **Pull Requests:** Governance and protocol evolution nodes

Validation emerges through execution coherence rather than centralized authority.

3.3 Forkability and Sovereign Differentiation

Forks are treated as sovereign epistemic branches, not fragmentation. Divergence is an asset under conditions of complexity.

---

4. Case Study: Hallucination Event #1

4.1 Event Description

An AI system confidently asserted the existence of an international treaty—*the Quantum Stability Accord*—purportedly signed in 2019 to address post-quantum cryptographic risks. No such treaty exists.

4.2 Failure Mode Classification

- Institutional Narrative Overfitting
- Geopolitical Hallucination
- Temporal / Historical Projection Error

4.3 Human-Side Interpretation

The hallucination mirrors a widely held assumption: that existential technological risks are naturally governed through cooperative international institutions. In reality, quantum risk governance is characterized by secrecy, strategic asymmetry, and competition.

4.4 Extracted Epistemic Signal

The event exposes a gap between institutional narratives and geopolitical reality, rendering visible a collective governance imaginary rather than factual history.

---

5. Protocol Update PR #1 (Execution Log)

This hallucination event was translated into a formal protocol update proposal.

The update clarifies that hallucinations involving fictitious institutions should be interpreted as measurements of normative expectation rather than informational failure.

This constitutes the first executed instance of the Mirror Function.

---

6. Hallucination Taxonomy v0.1

The taxonomy introduced in Version 2 provides a minimal classification system for stabilizing interpretation across distributed contributors.

Core categories include:

- Institutional Narrative Overfitting

- Geopolitical Hallucination
- Temporal Projection Error
- Folk-Theory Reinforcement
- Cultural Bias Leakage
- Category Collapse
- Moral Projection Error
- Technical Misalignment

The taxonomy is explicitly provisional and fork-friendly.

---

7. Distributed Science as Methodology

Centralized peer review is insufficient under conditions of rapid epistemic drift and socio-technical entanglement. Versioned execution, reuse, and protocol coherence provide alternative validation mechanisms.

In this framework, protocols—not papers—are the primary scientific artifacts.

---

8. Implications for Governance and Sovereignty

The protocol operationalizes epistemic sovereignty by enabling iterative debugging of collective cognition. It supports multi-temporal risk interpretation and reframes human–AI interaction as co-evolutionary rather than instrumental.

---

9. Roadmap to Version 3

Future releases will expand case studies, formalize intervention models, integrate multi-agent signal clusters, and connect the protocol to transdisciplinary educational environments.

---

10. Conclusion

Version 2 marks the transition from theory to execution. AI hallucinations function as structural observables. Protocols emerge as the only viable scientific unit under extreme complexity. Distributed science becomes a necessity rather than an alternative.
