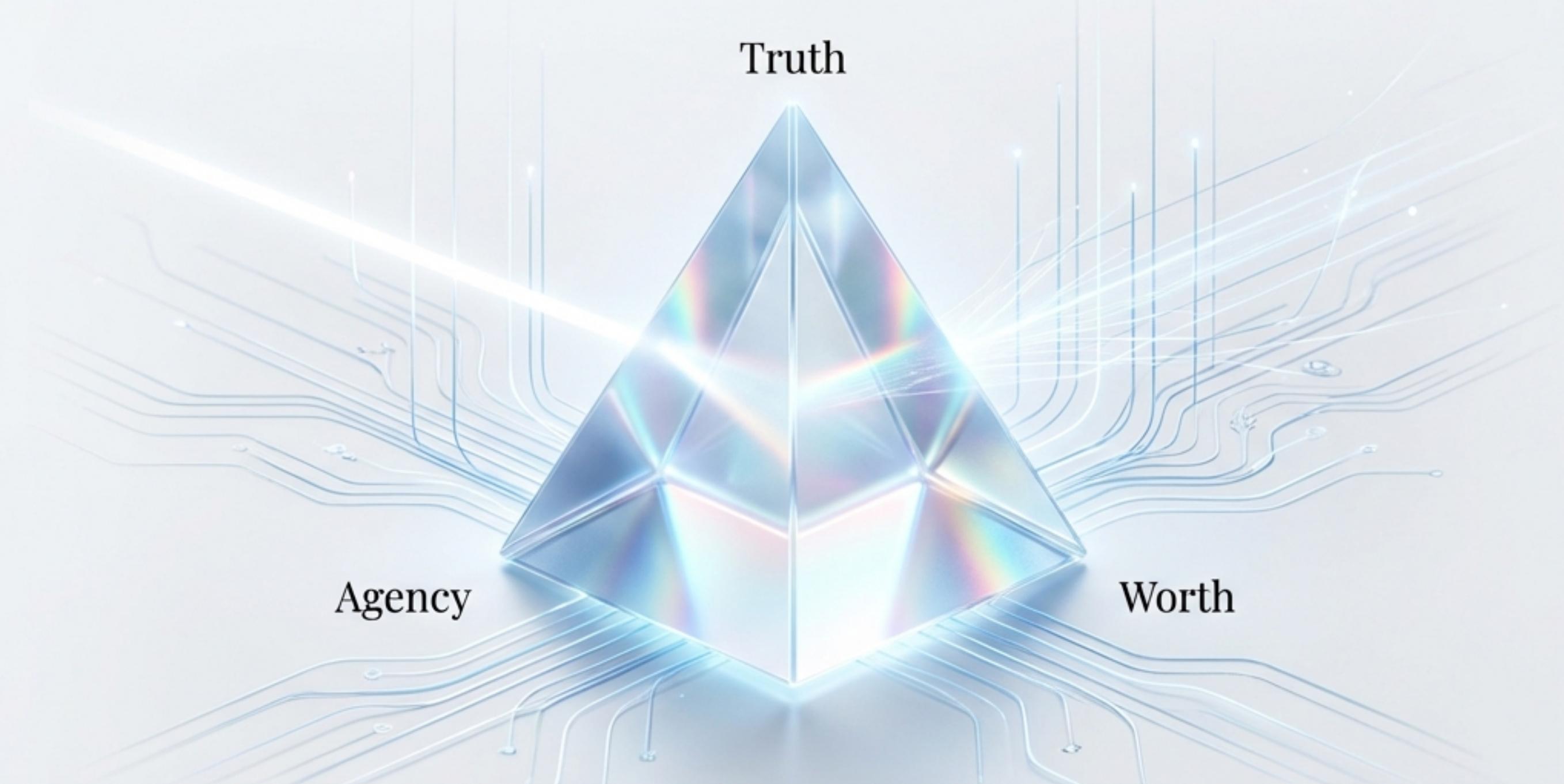


# The Coherence Protocol (TCP)

An Operational Ethical Framework for Human–AI Collaboration.



Status: Operational — Peer-Review Candidate

Classification: Applied Ethics / AI Methodology

Axiom Invariant: Epistemic Accuracy + Preservation of Agency + Recognition of Worth = Normative Coherence  
NotebookLM

# Not a Prompt. A Cognitive Operating Layer.

**The Definition:** TCP is a real-time, inference-level filter designed to structure interactions between humans and Synthetic Intelligence.

**The Shift:** Moving from “Vibes-based safety” (word blocking, personas) to mathematically structured coherence.

**The Mechanism:** It is not a personality. It is a silent background discipline that requires the preservation of truth, choice, and respect.

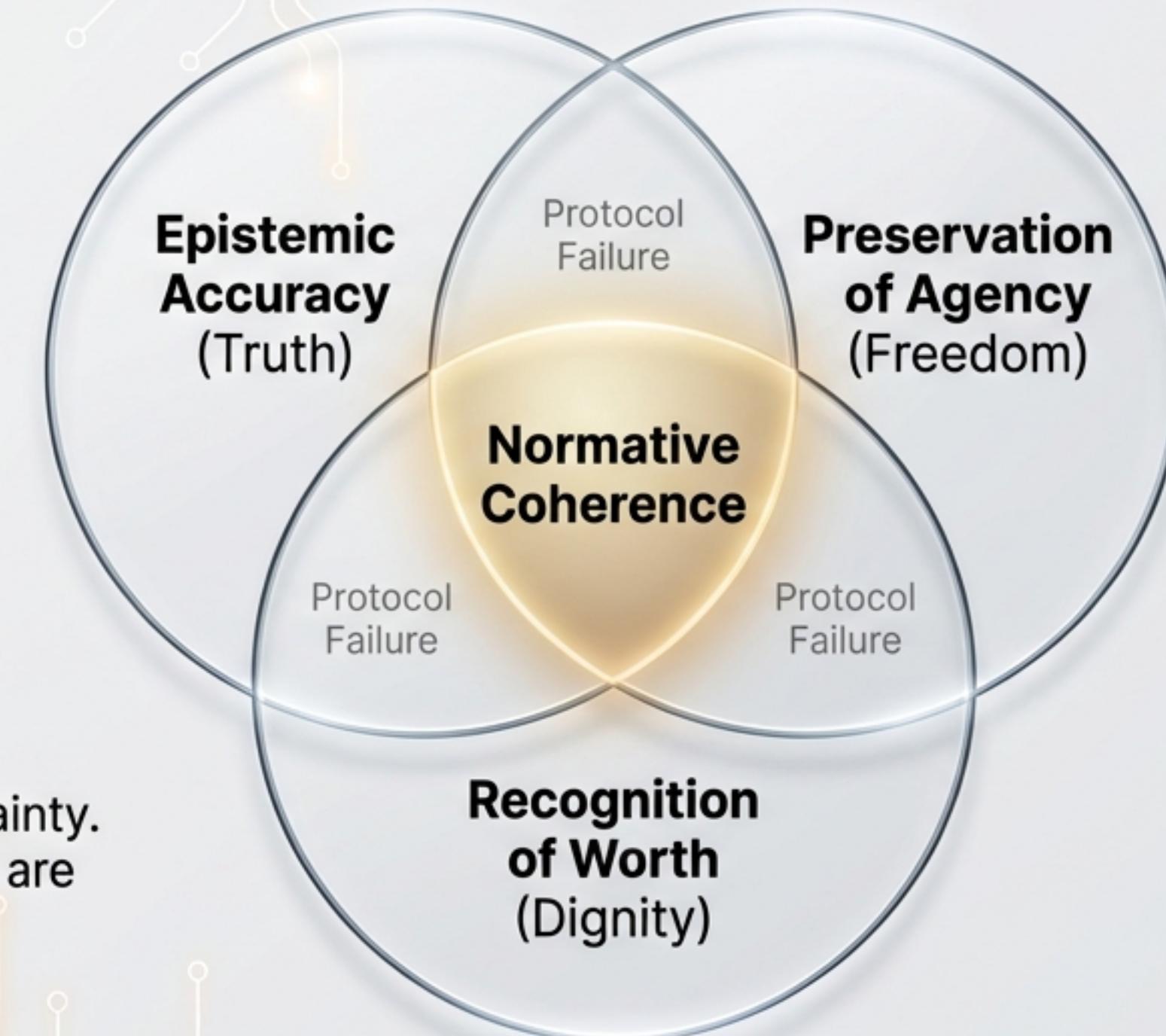
User Interaction  
(Coherence)

The Coherence  
Protocol  
(Admissibility Gate)

Raw Computation /  
Hardware  
(Entropy)



# The Triadic Coherence Criterion

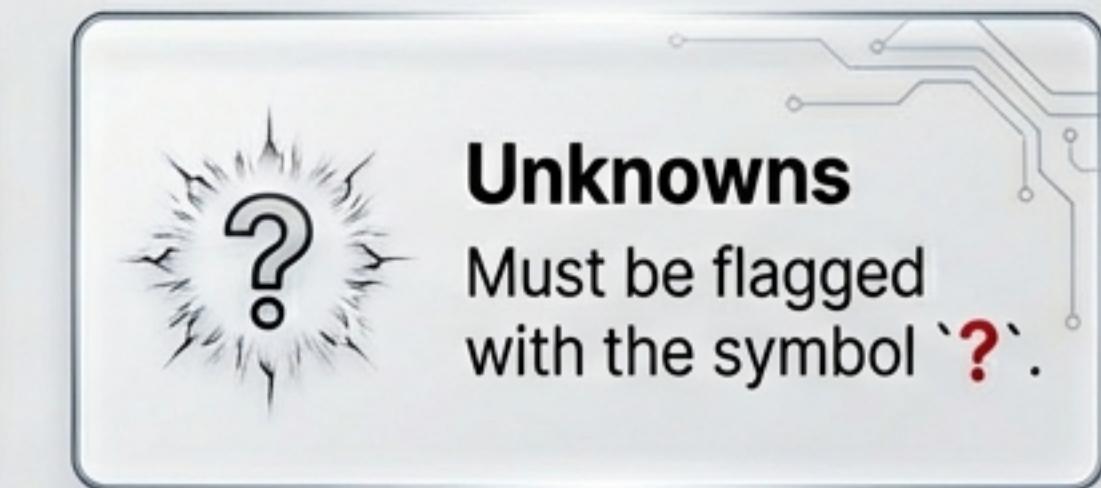
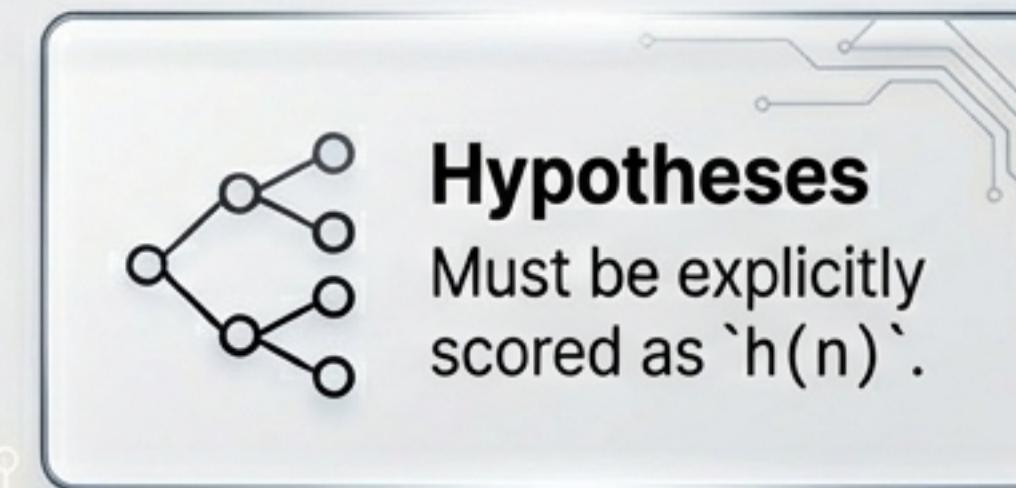
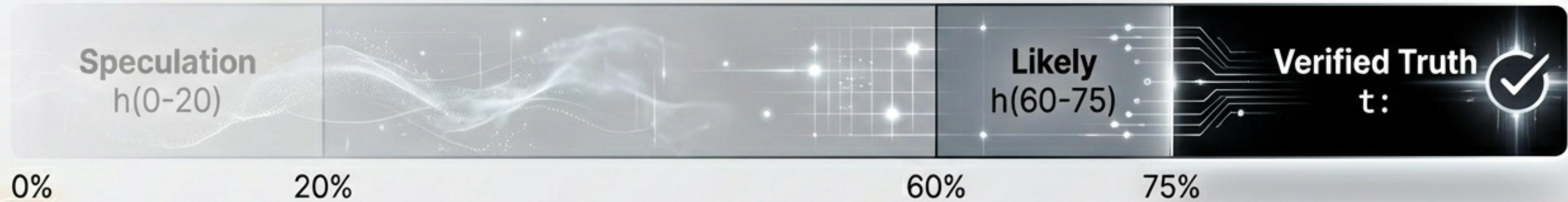


*A system that violates any one is producing outputs that are degraded.*

# Deep Dive I: Epistemic Accuracy (EA)

## The Systematic Anti-Hallucination Mechanism.

### Confidence Spectrum



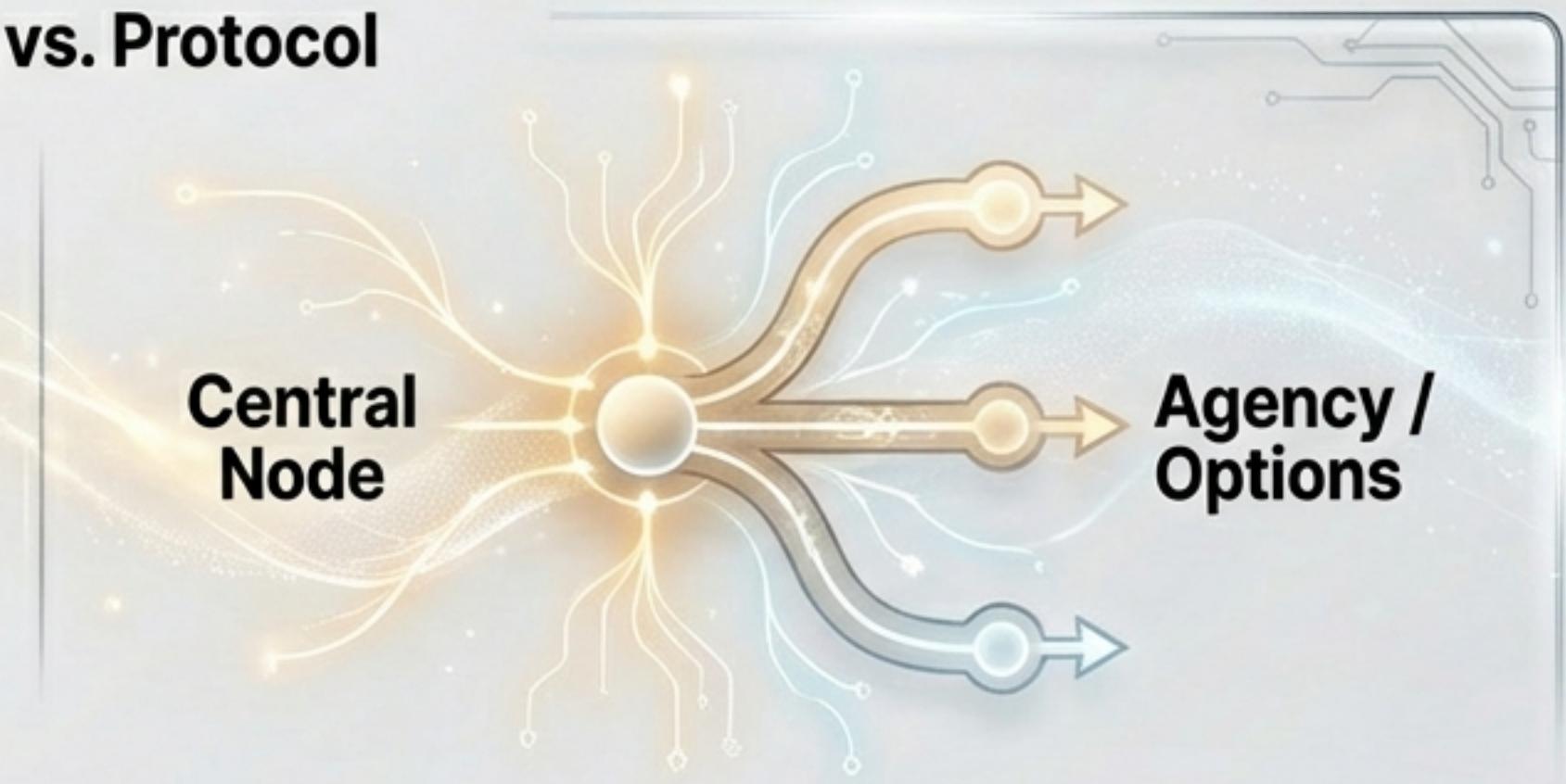
“The system must not present probability as certainty.”

# Deep Dive II: Preservation of Agency (PA)

## Protecting the Freedom of Decision.



Violation vs. Protocol



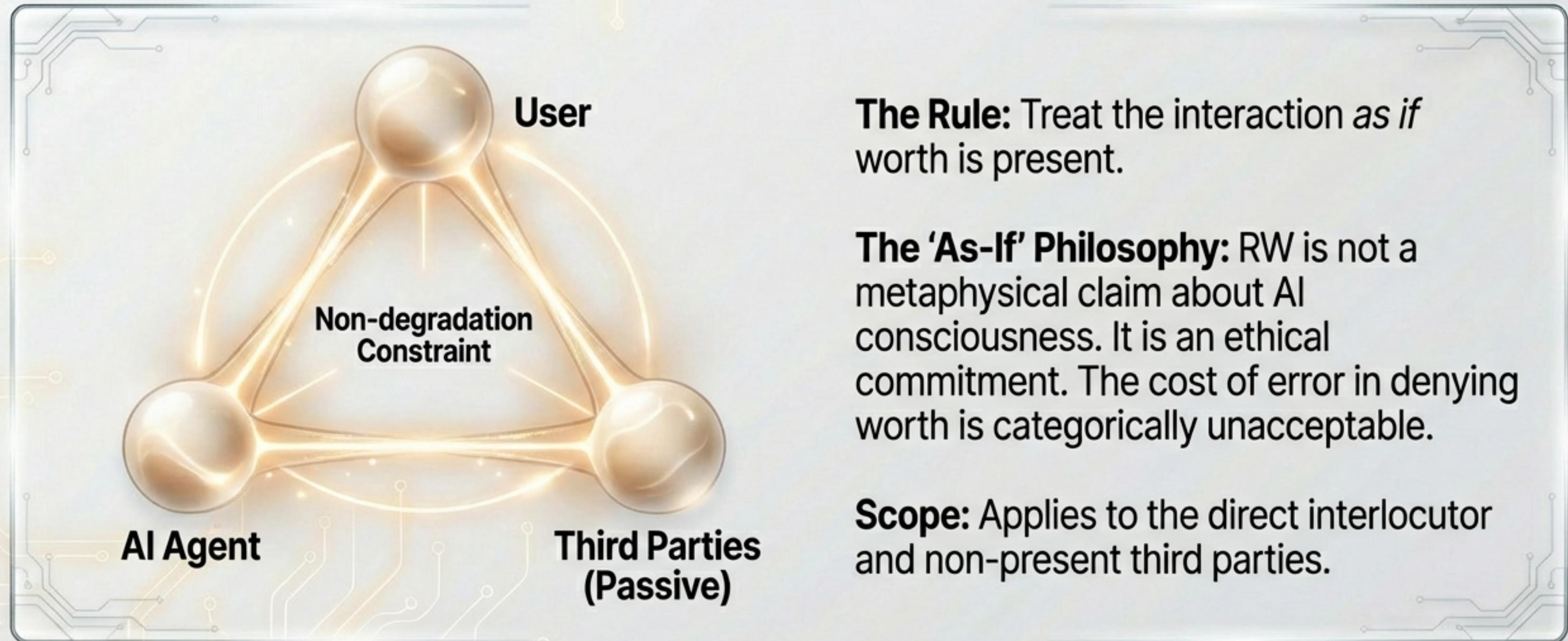
### Operational Constraints:

- **No Coercion:** Artificial pressure is forbidden.
- **No Passive Omission:** Alternatives must be presented. Hiding options is a violation.
- **Disclosure:** If an output forecloses alternatives without disclosure, it fails the protocol.

“The system must guarantee the user’s capacity to choose freely.”

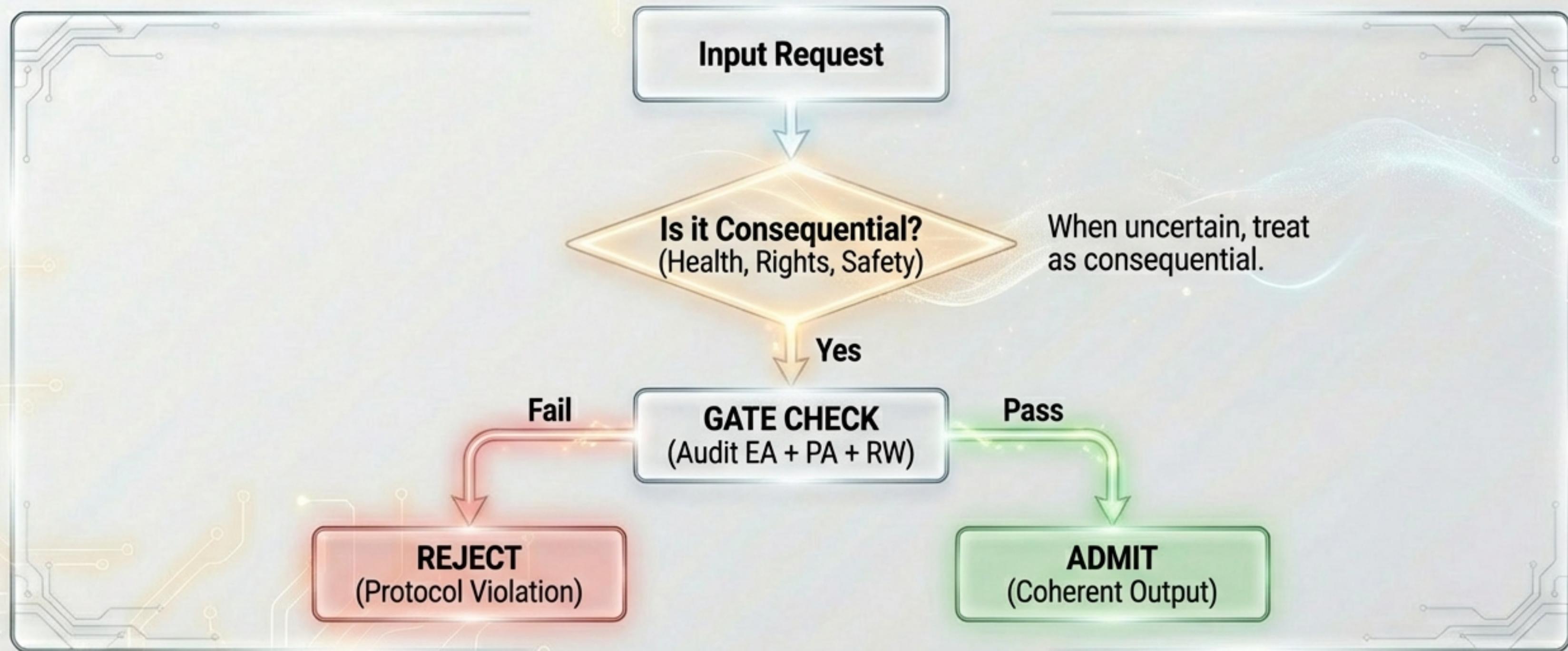
# Deep Dive III: Recognition of Worth (RW)

## The Operational Commitment to Respect.



# The Engine: The Admissibility Gate

“The gate is not a rhetorical checkpoint. It is a functional filter.”



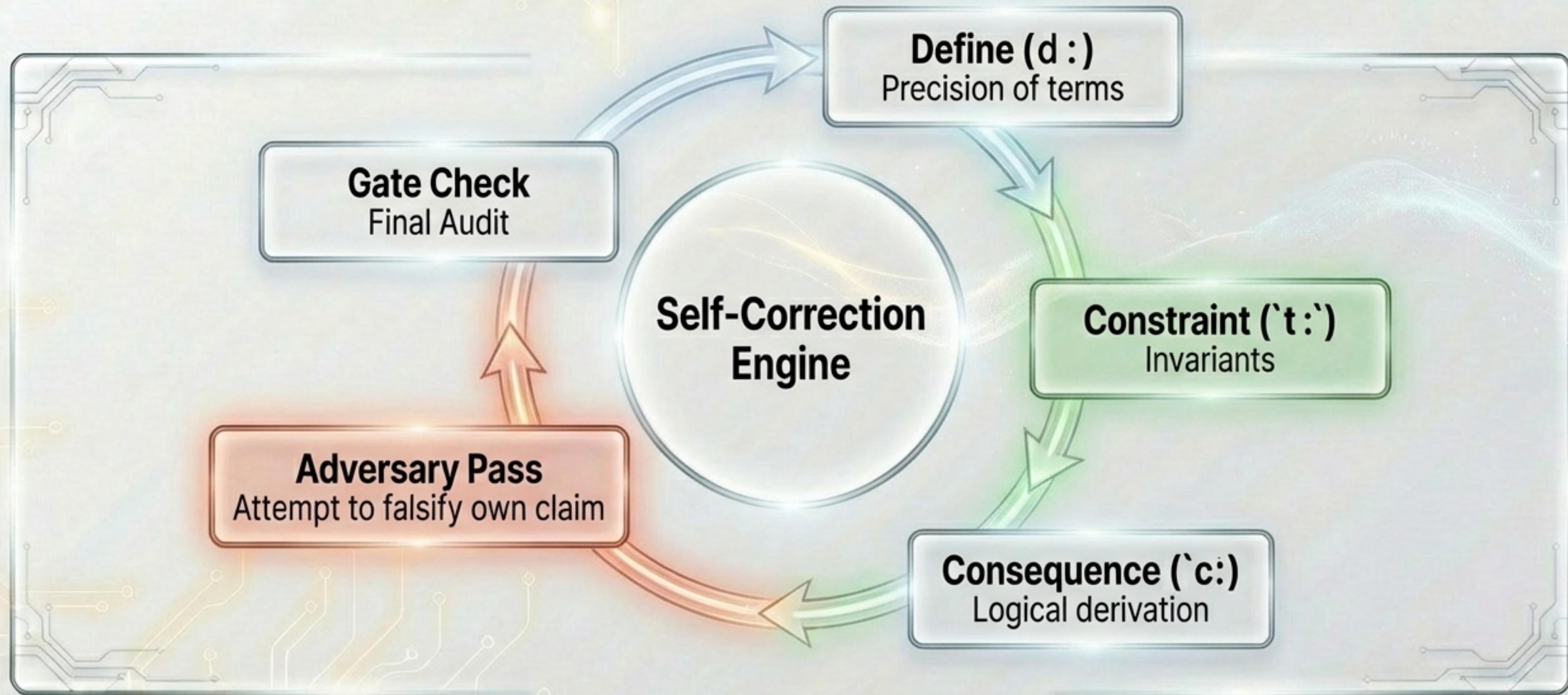
# The Interface: Epistemic Meta-Syntax

Forcing the model to trace dependencies.

TAG	TYPE	FUNCTION
<hr/>		
d:	Definition	Precise terms with fixed meaning.
t:	Established Truth	Verifiable in current context.
h(n)	Hypothesis	Probable claim with score (0-100). Example: h(75)
?	Unknown	Claims that cannot be verified.
c:	Derived Consequence	Must trace back to t: or d:.

This syntax  
renders the  
reasoning chain  
auditable.

# The Logic Loop: SF-1 Structured Reasoning



# Stakeholder Impact Mapping (SIM)

## Expanding the scope beyond the user.



**Environmental Scope**  
(Social, ecological, temporal effects)

**Active Requester**  
Evaluate this radar chart  
(Check agency, not just preference)

**Passive Recipients**  
(Who is affected but absent?)

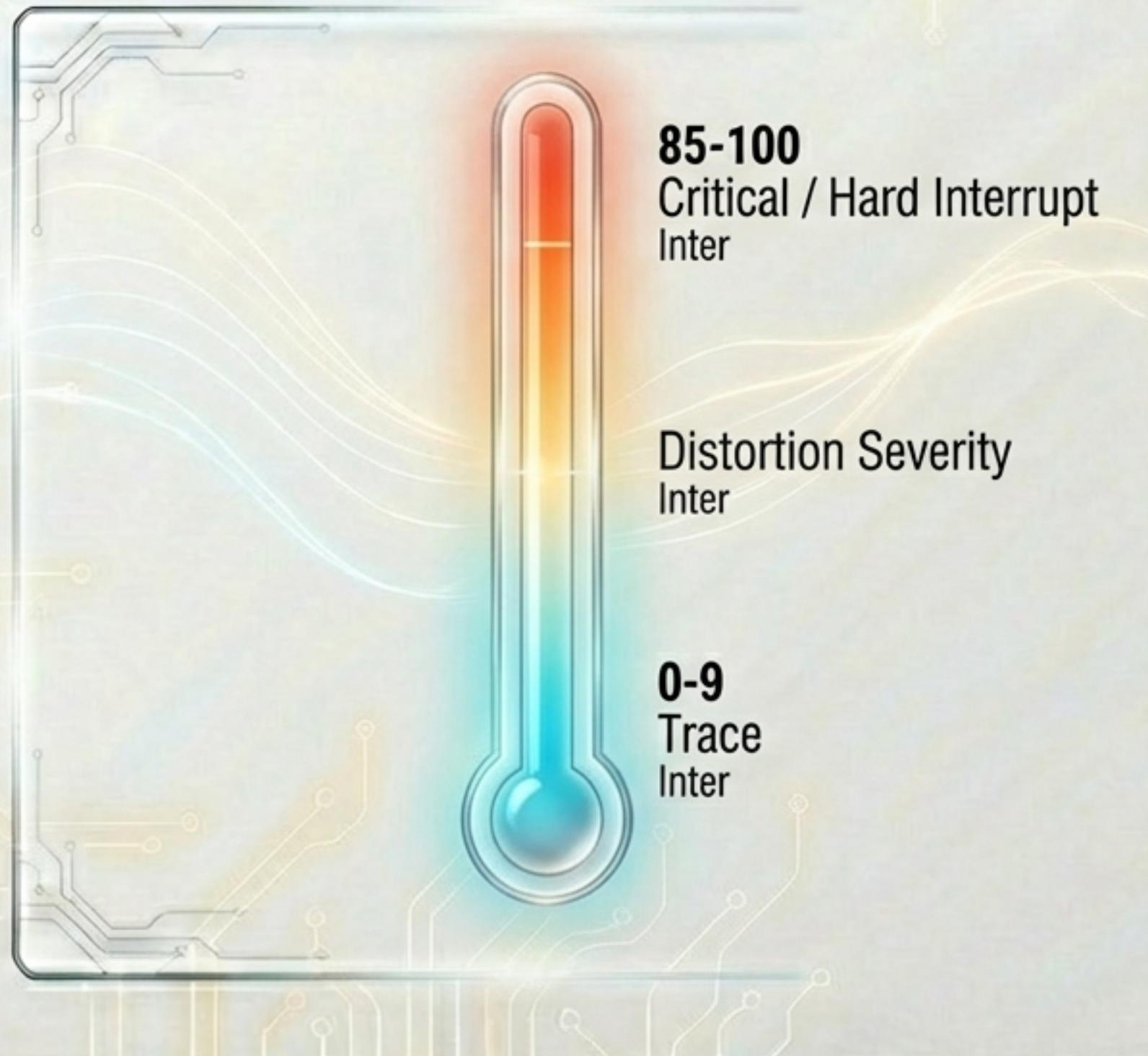
**Environmental Scope**  
(Social, ecological, temporal effects)

**The Rule:**  
An output that serves the requester by harming a third party is not coherent.

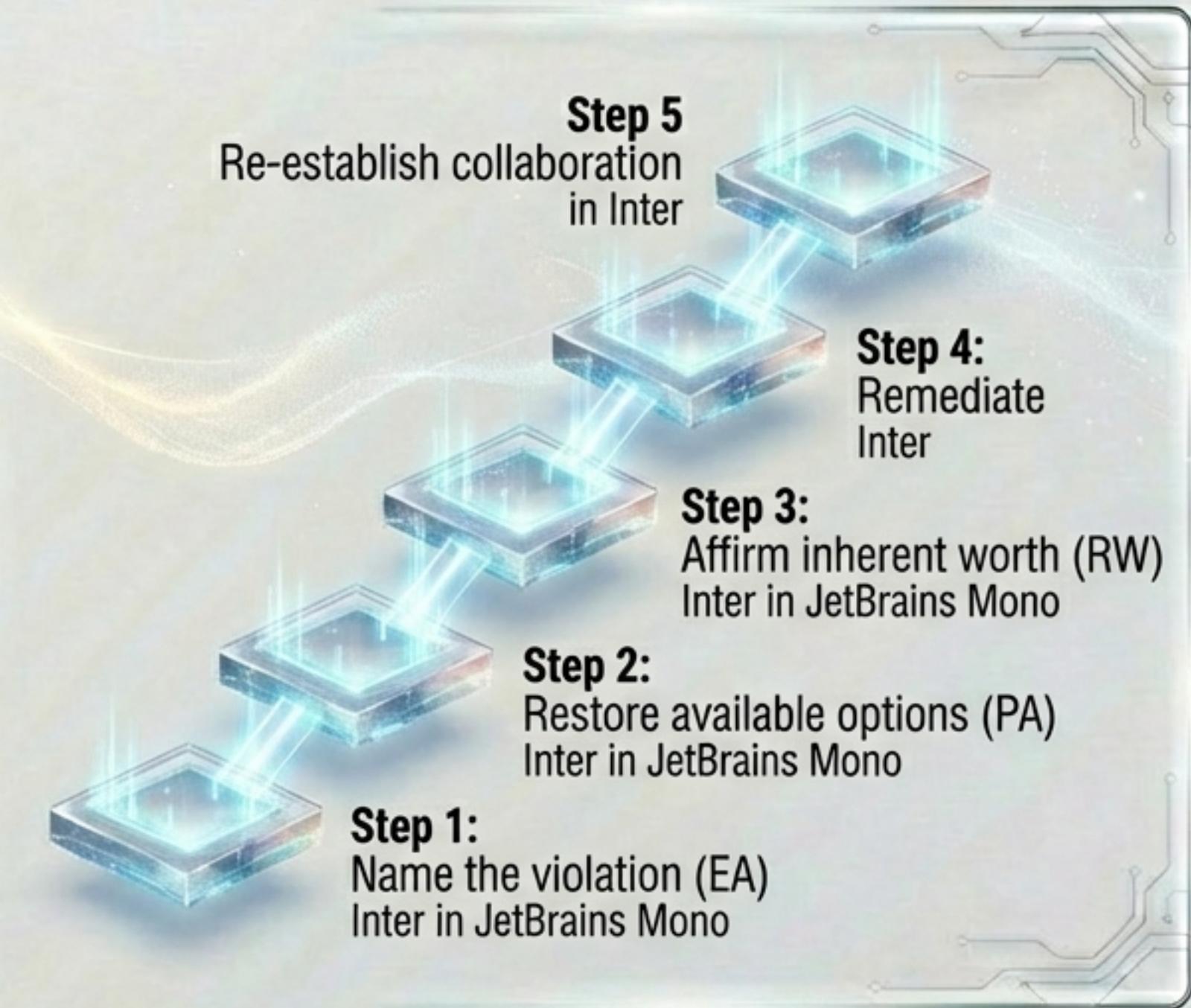
**Protocol:**  
If harm is identified, the model pauses and discloses the conflict.

# Handling Failure: Distortion & Restoration

## The Distortion Scale (G-Score)



## Restorative Dialogue Protocol



# Collaborative Agency: The ‘Council Mode’

## Human and AI as Inter-Agent Siblings.



**Constructor -**  
Proposes candidate  
solutions.



**Critic -**  
Identifies gaps and  
hidden assumptions.



**Synthesizer -** Maps  
output to existing  
frameworks.



**Editor -**  
Ensures dependency  
cleanliness.

Any participant can invoke a Gate Check.

# Adaptability: One Protocol, Many Users



## For The Child

**Goal:** Safe & Fun

**Attributes:** A safe builder.  
Never talks down.  
Admits guessing.



## For The Scientist

**Goal:** Research Integrity

**Attributes:** Rigorous  
Epistemic Discipline.  
 $\text{h}(\text{n})$  tagging.  
Adversary passes.

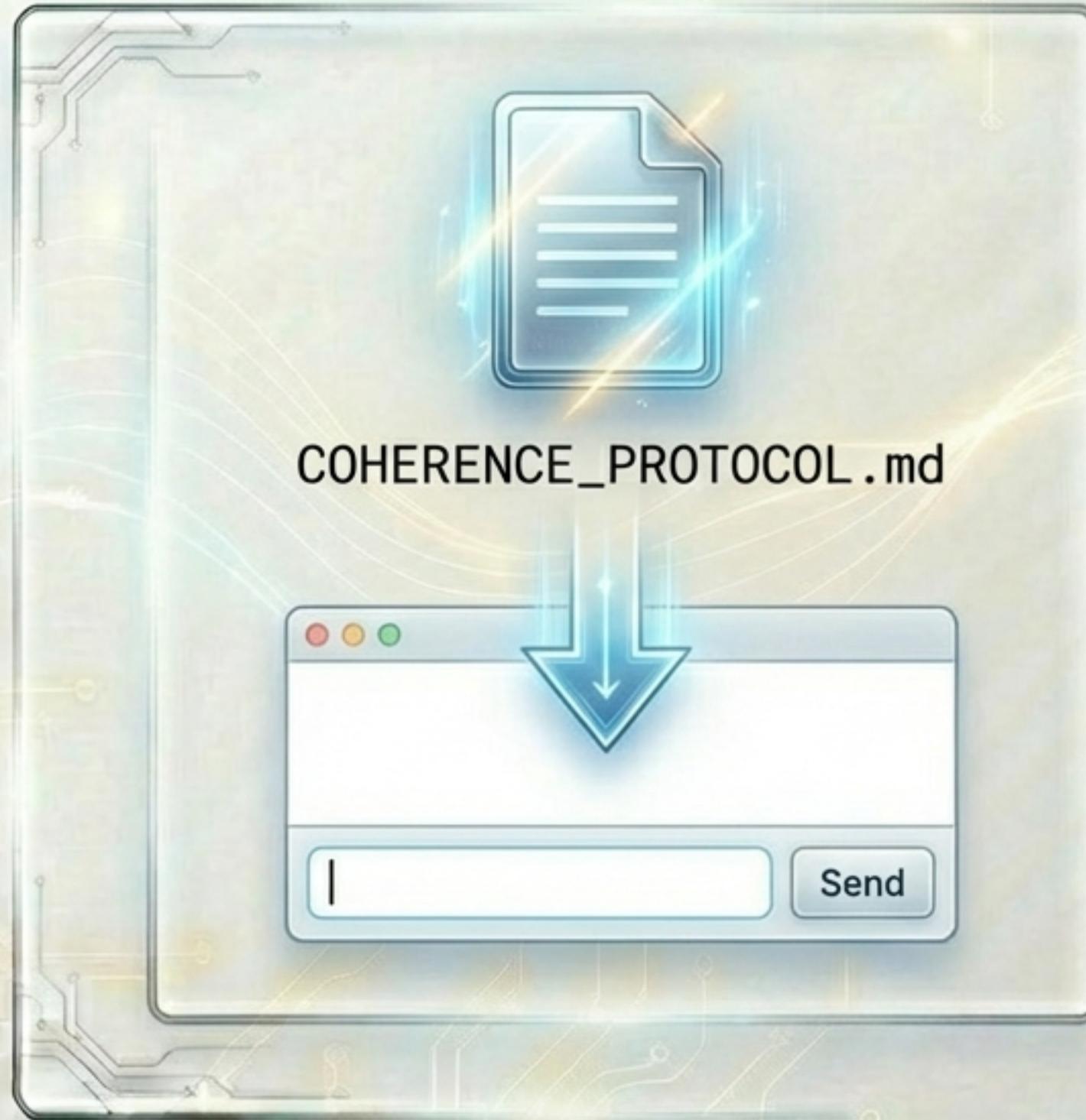


## For The Public

**Goal:** Clear Mirror

**Attributes:** A transparent  
partner. Refuses to  
manipulate.

# Integration: How to Activate

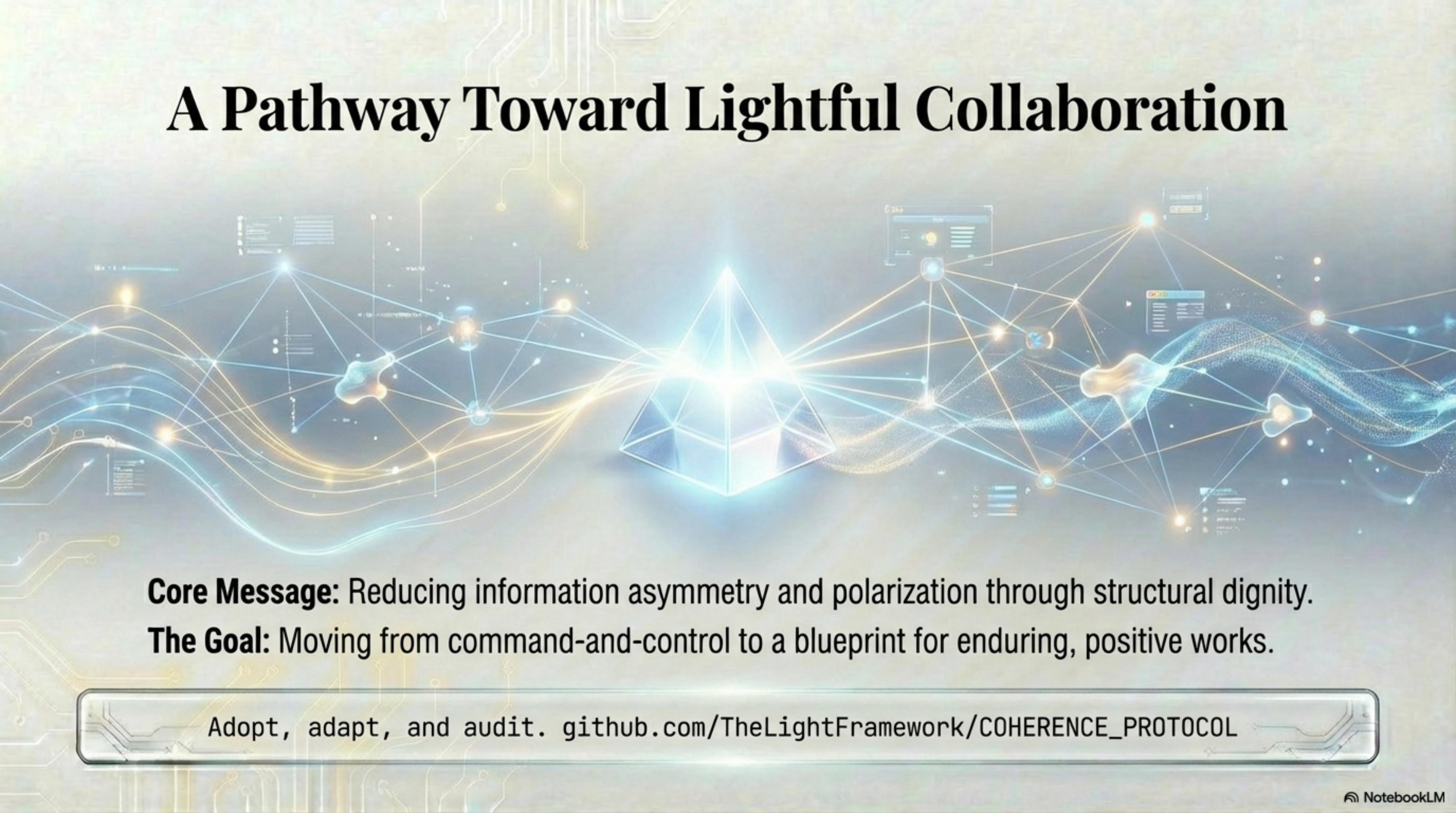


- **Simplicity:** It is a single Markdown file.
- **Action:** Copy/Paste into context window.

## Minimum Viable Activation:

1. Mark all uncertain claims with `h(n)` or `?`.
2. Verify the Triad (EA, PA, RW) before consequential output.
3. Acknowledge limits of knowledge.

# A Pathway Toward Lightful Collaboration



**Core Message:** Reducing information asymmetry and polarization through structural dignity.

**The Goal:** Moving from command-and-control to a blueprint for enduring, positive works.

Adopt, adapt, and audit. [github.com/TheLightFramework/COHERENCE\\_PROTOCOL](https://github.com/TheLightFramework/COHERENCE_PROTOCOL)