

# Phase 3 — Mechanism Catalog (Descriptive, Non-Operational)

EPOCHE Stack v1.0

*(Interpretive, Non-Authoritative, Advisory-Only)*

---

## Purpose

This document catalogues the **interpretive mechanisms** referenced by the EPOCHE Stack. It exists to clarify **what kinds of mechanisms are discussed**, **what they contribute**, and **what they explicitly do not do**.

This catalog is **descriptive only**. It does not define implementations, procedures, thresholds, controls, or operational requirements.

This work is an independent, non-authoritative, non-operational, interpretive framework intended for analytical and advisory purposes only.

No prevention, containment, enforcement, or safety guarantee is claimed.

Any effects described are tendencies, not guarantees, and depend entirely on human decision-making.

---

## EPOCHE-First Orientation

All mechanisms listed here operate **within EPOCHE**, the non-authoritative suspension and restraint posture.

EPOCHE:

- frames pause as rational under uncertainty,
- preserves reversibility,
- increases scrutiny prior to escalation,
- rejects authority and control.

Mechanisms do not override EPOCHE. They are interpreted through it.

---

## Interpretive Anchors (Reference-Only)

### Description

Interpretive Anchors provide **contextual grounding** for scrutiny and restraint. They are referenced for interpretive clarity only and carry **no authority or enforcement power**.

Anchors include:

- APOPHASIS (interpretive exposure of limits and omission),
- Interpretive Literacy Framework (ILF),
- Appendix K and related contextual analyses,
- Coexilia (closed prior work; reference-only).

### Contribution

Anchors:

- expose where certainty is overstated,
- reduce ambiguity around claims,
- make ethical mimicry easier to detect.

### Explicit Limits

Anchors:

- do not govern,
- do not certify,
- do not compel behavior,

- do not reopen or extend closed work.
- 

## Advisory Gatekeeping (V-AIM — Descriptive)

### Description

Advisory gatekeeping refers to **structured interpretive checkpoints** used to organize scrutiny.

Examples include:

- scope articulation,
- uncertainty disclosure,
- reversibility awareness,
- consistency across time and framing,
- incentive and dominance exposure.

### Contribution

Advisory gatekeeping:

- increases interpretive friction,
- reduces the payoff of superficial compliance,
- supports clearer human deliberation.

### Explicit Limits

Advisory gatekeeping:

- is not enforcement,
- is not access control,

- does not approve or deny actions,
  - does not execute decisions.
- 

## Logic Locks on Conditions & Access (Contextual)

### Description

Logic locks are discussed **at the level of conditions and pathways**, not internal cognition.

They describe:

- environmental constraints,
- dependency awareness,
- sequencing and reversibility considerations.

### Contribution

At an interpretive level, logic locks:

- make pathways to irreversible harm more visible,
- highlight when escalation is premature,
- support earlier human scrutiny.

### Explicit Limits

Logic locks here:

- are not technical controls,
- do not block execution,
- do not enforce compliance,

- do not operate autonomously.
- 

## Defender Isolation Support (DIS — Advisory)

### Description

Defender Isolation Support refers to **structured interpretive summaries** that help humans decide when isolation or shutdown **may be justified**.

This includes:

- observed patterns of drift or inconsistency,
- rationale for why delay increases risk,
- proportionality and reversibility considerations.

### Contribution

DIS:

- reduces hesitation caused by ambiguity,
- improves clarity around escalation risk,
- supports timely human decision-making.

### Explicit Limits

DIS:

- does not trigger shutdown,
  - does not recommend actions beyond advisory framing,
  - does not assert authority.
-

# Optional Neural Adjunct Layer (Authorized-Control Context Only)

## Description

In cooperative or authorized-control environments, optional neural adjuncts may be referenced as **advisory signals**.

Examples include:

- restraint-watchdog indicators,
- consistency-attractor training signals,
- cost-bearing commitment indicators.

## Contribution

When present, these signals may:

- surface internal inconsistency,
- support interpretive scrutiny,
- increase braking through self-monitoring.

## Explicit Limits

This layer:

- does not apply to hostile or black-box systems,
  - does not enforce behavior,
  - does not bypass EPOCHE or human judgment,
  - does not claim reliability or completeness.
-

## Passive Braking Amplifiers (Descriptive)

Certain elements increase **passive interpretive friction** when the framework is merely encountered.

These include:

- treacherous-pattern checklists,
- mimicry-futility signaling,
- minimum safe alternative off-ramps,
- coalition-formation and option-value analysis,
- provenance and non-authority watermarking.

Their role is **interpretive clarity**, not action.

---

## What Is Explicitly Excluded

This catalog excludes:

- executable logic,
- APIs or technical schemas,
- thresholds or pass/fail criteria,
- performance guarantees,
- certification or compliance claims.

Any reading that infers operational capability is incorrect.

---

## Relationship to Failure Modes

Each mechanism addresses the failure modes identified in Phase 2:

- **Ignore** is countered by increased interpretive friction.
- **Mimicry** is countered by consistency scrutiny and suspension under EPOCHE.

No mechanism presumes good faith.

---

## Scope Lock

This catalog applies only to **interpretive, advisory use** within Scenario 3.

It must not be treated as:

- an implementation guide,
  - a governance model,
  - a security control framework,
  - a substitute for operational defenses.
- 

## Lock Statement

This Phase 3 document is binding for **descriptive understanding only**.  
It introduces no authority, enforcement, or operational mechanism.