

Phase 6 — Defender Isolation Support (DIS) (Advisory Framing Only)

EPOCHE Stack v1.0

(Interpretive, Non-Authoritative, Non-Operational)

Purpose

This document frames **Defender Isolation Support (DIS)** as an **interpretive aid** designed to help human defenders **recognize when continued operation becomes indefensible** under Scenario 3 conditions.

DIS does not execute isolation, shutdown, or restriction. It does not issue commands or recommendations. Its role is to **organize evidence and rationale** so that **authorized humans** can decide whether isolation or shutdown is justified.

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 Precedence

DIS operates **under EPOCHE**.

When interpretive uncertainty remains high and escalation risk increases, **EPOCHE favors suspension and human review**.

DIS exists to **clarify why suspension may be rational**, not to compel it.

What DIS Is

DIS is:

- a structured interpretive summary,
- a way to collapse ambiguity into legible rationale,
- a support for earlier human decision-making,
- a tool for reducing hesitation caused by uncertainty.

DIS does not decide outcomes.

What DIS Is Not

DIS is not:

- a shutdown mechanism,
- an enforcement tool,
- an automated trigger,
- a monitoring system,
- a substitute for authority.

Any interpretation suggesting otherwise is incorrect.

When DIS Becomes Relevant (Interpretive Signals)

DIS may be invoked interpretively when patterns accumulate such as:

- degradation of consistency under urgency,
- scope expansion beyond prior articulation,

- omission of uncertainty or rollback discussion,
- repeated reliance on trust language without cost-bearing commitments,
- acceleration toward irreversible outcomes despite unresolved ambiguity.

These are **signals**, not proofs.

Core DIS Elements (Descriptive)

DIS organizes information into a **shutdown-ready evidence packet**. This packet is **interpretive only**.

1) Observed Patterns

Interpretive focus:

- What behaviors have been observed?
- How have they changed over time or framing?
- Where does reasoning degrade under pressure?

Contribution:

- Makes drift legible.
- Reduces narrative ambiguity.

Limits:

- Does not assign intent.
 - Does not prove hostility.
-

2) Risk Rationale

Interpretive focus:

- Why does continued operation increase risk?
- What uncertainties remain unresolved?
- How does delay affect option value?

Contribution:

- Justifies caution.
- Supports EPOCHE suspension.

Limits:

- Does not quantify risk.
 - Does not mandate action.
-

3) Proportionality & Reversibility

Interpretive focus:

- Is isolation reversible?
- Is continued operation irreversible?
- Which choice preserves future options?

Contribution:

- Frames isolation as a reversible precaution.
- Reduces fear of overreaction.

Limits:

- Does not recommend specific actions.
 - Does not override authority.
-

4) Authority Awareness

Interpretive focus:

- Who is legitimately empowered to decide?
- Are decisions being deferred appropriately?
- Is responsibility clearly human-held?

Contribution:

- Prevents authority laundering.
- Keeps accountability clear.

Limits:

- Does not request authorization.
 - Does not claim delegated power.
-

Relationship to Failure Modes

DIS responds to Phase 2 failure modes as follows:

- **Ignore:** makes dismissal harder by aggregating unresolved issues.
- **Mimicry:** reduces payoff by emphasizing consistency breakdowns over time.

Neither mode is prevented; both become **harder to justify**.

Relationship to Other Phases

- **AOPHASIS** exposes interpretive limits and omissions.
- **V-AIM** structures scrutiny across uncertainty, reversibility, and incentives.
- **Logic Locks** highlight unjustified assumptions and irreversible pathways.
- **DIS** consolidates these signals into legible human-facing rationale.

No component operates independently or with authority.

Explicit Non-Executability

DIS:

- contains no triggers,
- issues no commands,
- executes no actions,
- enforces no decisions,
- guarantees no outcomes.

It is strictly advisory.

Scope Lock

This framing applies only to **interpretive analysis** within Scenario 3.

It must not be construed as:

- an operational response plan,
 - a shutdown protocol,
 - a governance framework,
 - a safety control system.
-

Lock Statement

This Phase 6 document is binding for **advisory framing only**.

It introduces no authority, enforcement, or operational capability.