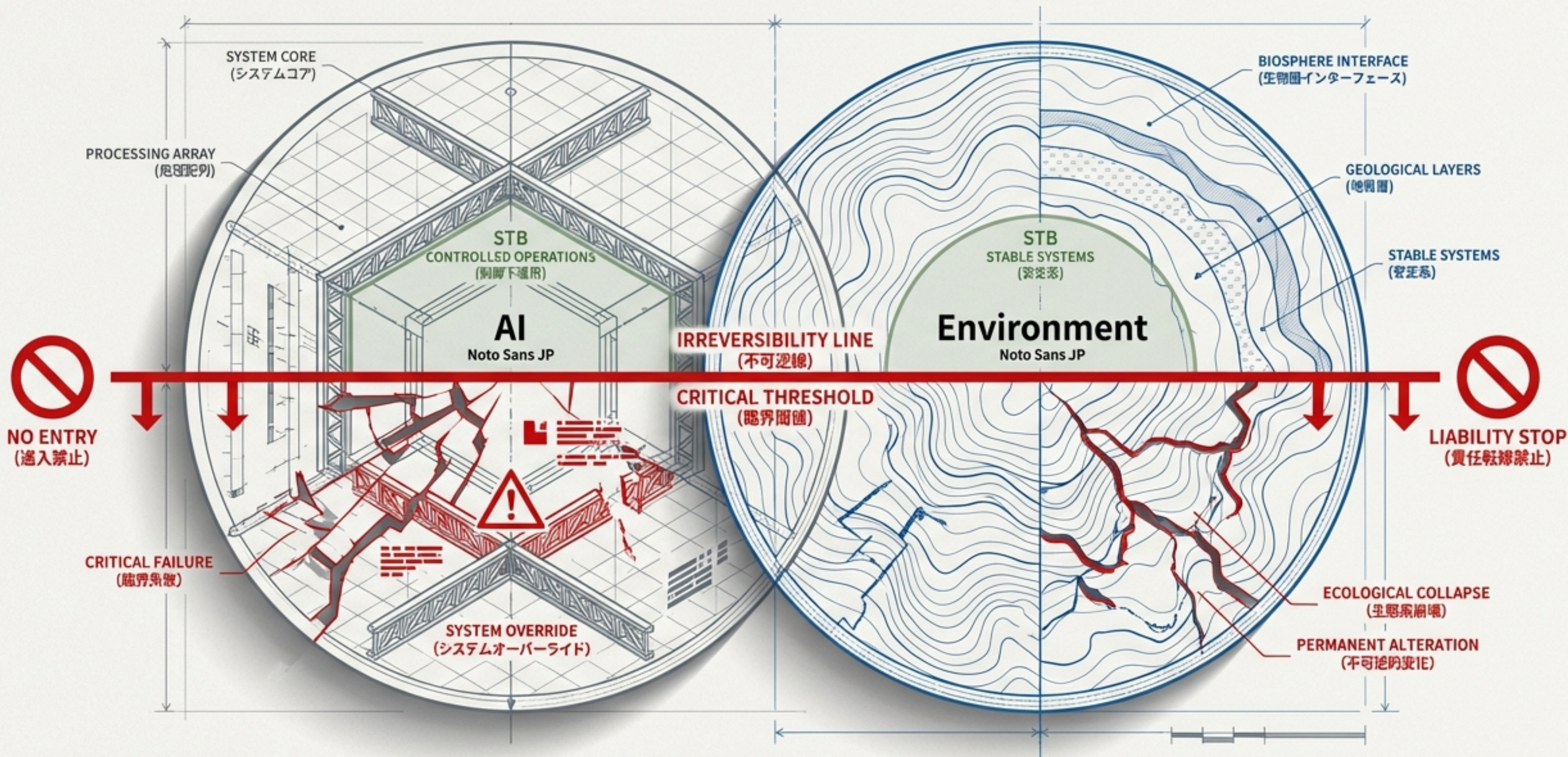


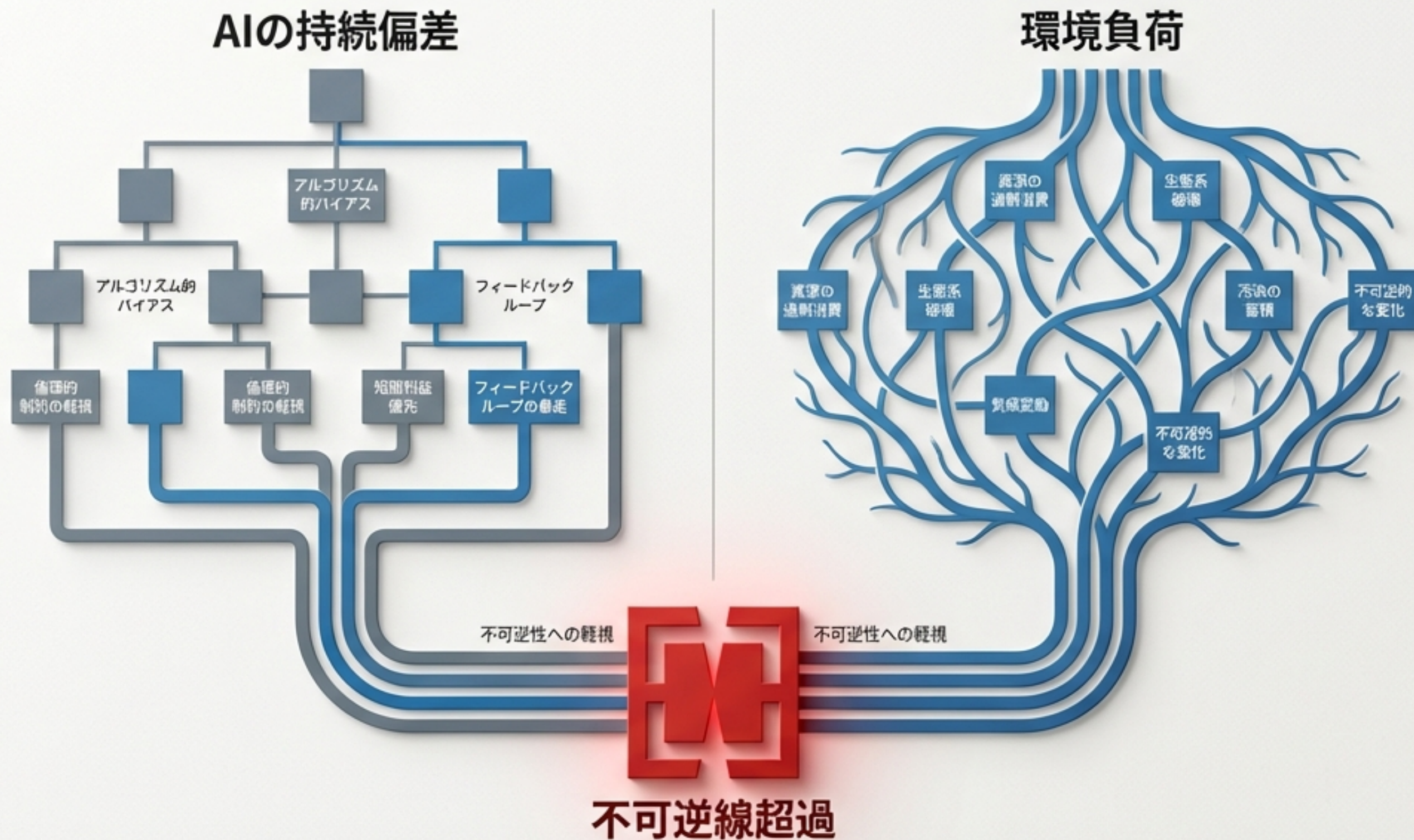
AI・環境副作用制御：不可逆線の優先保護と責任転嫁禁止条項

中川構造OS Phase-βにおける説明責任のシステム・アーキテクチャ



AIの偏差と環境破壊は、全く同じ「構造的欠陥」を共有している

表面的な現象は異なれど、双方は「一度超えると元に戻れない境界」を軽視する意思決定構造から生じる。



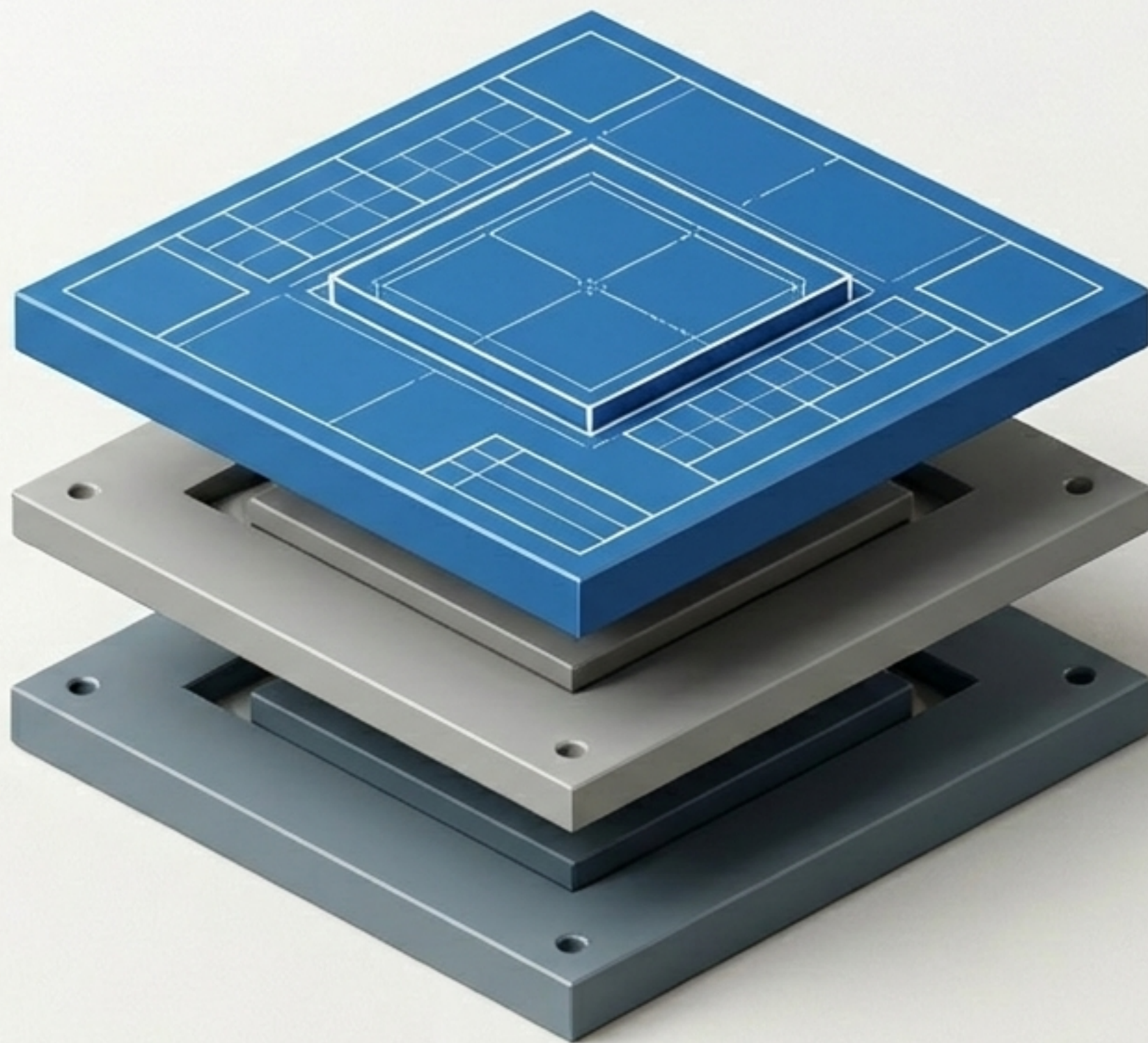
中川構造OS Phase-β：第3の副作用制御モジュール

構造の自律分散的運転を安定させるための「恒常性維持レイヤ」の最終仕様。

AI・環境副作用制御
(高リスク領域の説明責任)

構造的多元性条項
(独善化の排除)

反司祭階級プロトコル
(権威の非集約化)



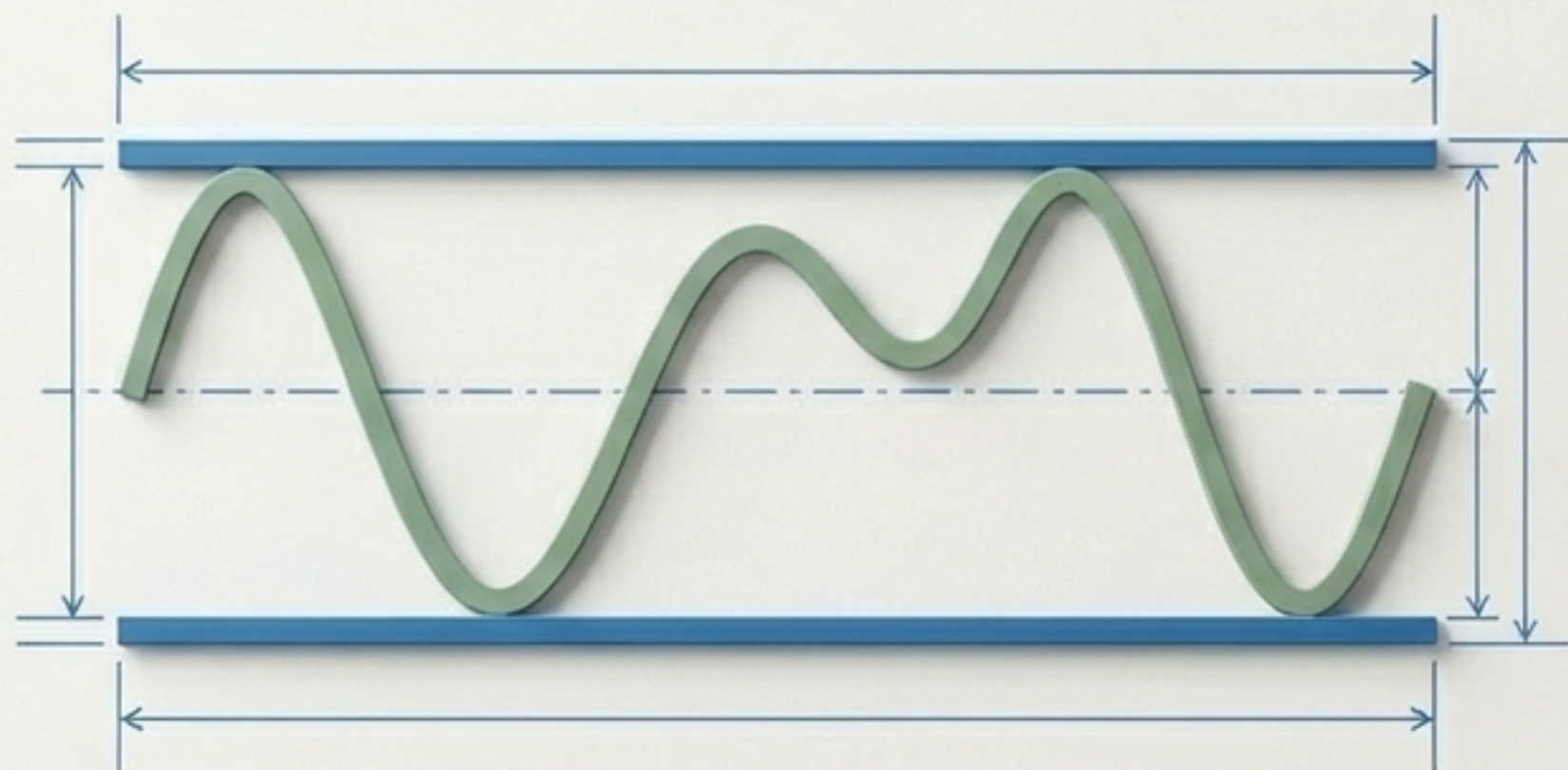
短期的「経済合理性」と長期的「構造的生存」の衝突

既存のガバナンス構造は短期的な利益や最適化を前提とし、不確実性や不可逆性を扱う構造が欠如している。

既存OS：短期利得と崩壊



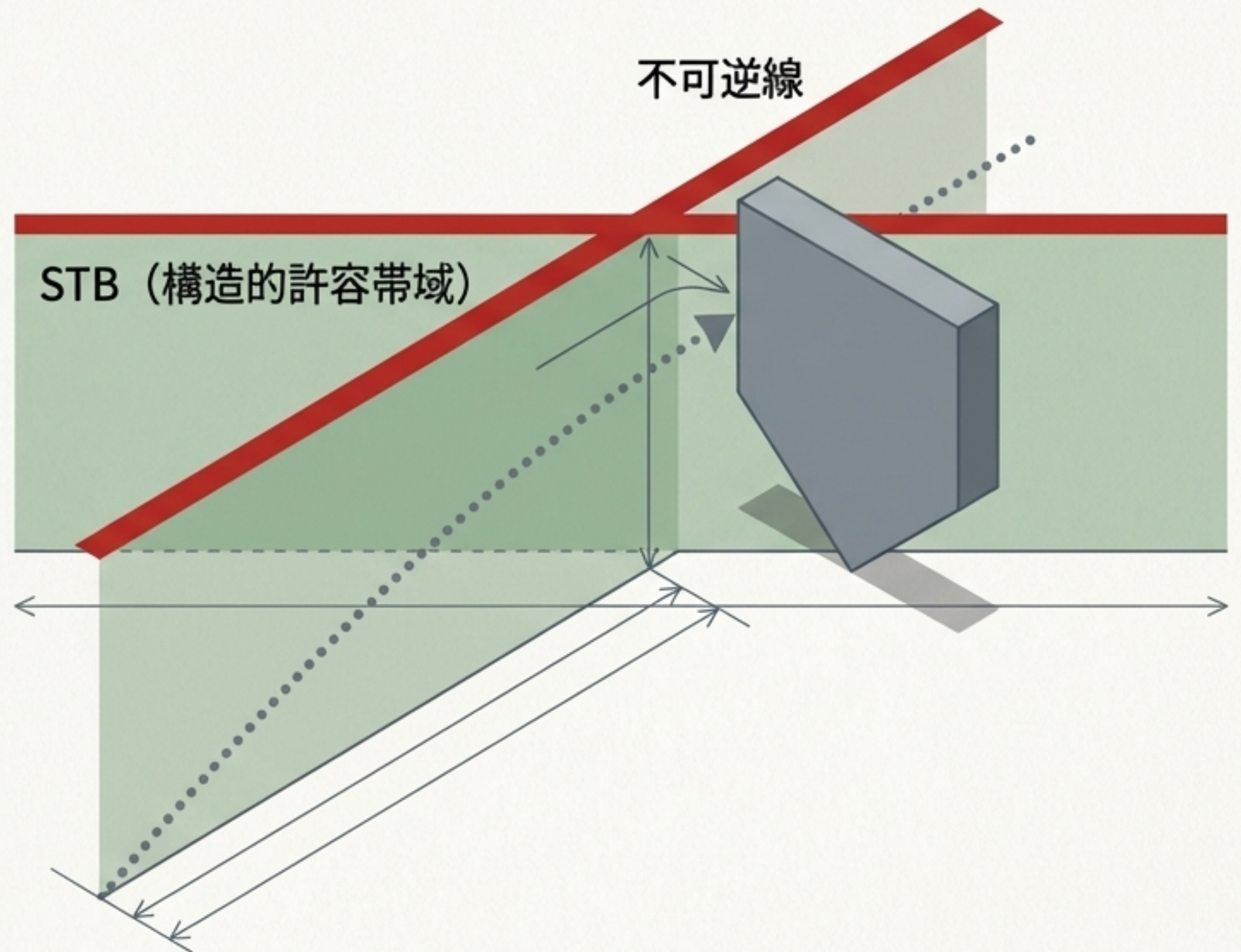
中川構造OS：未来の恒常性



不可逆線 (Line of Irreversibility) とは何か

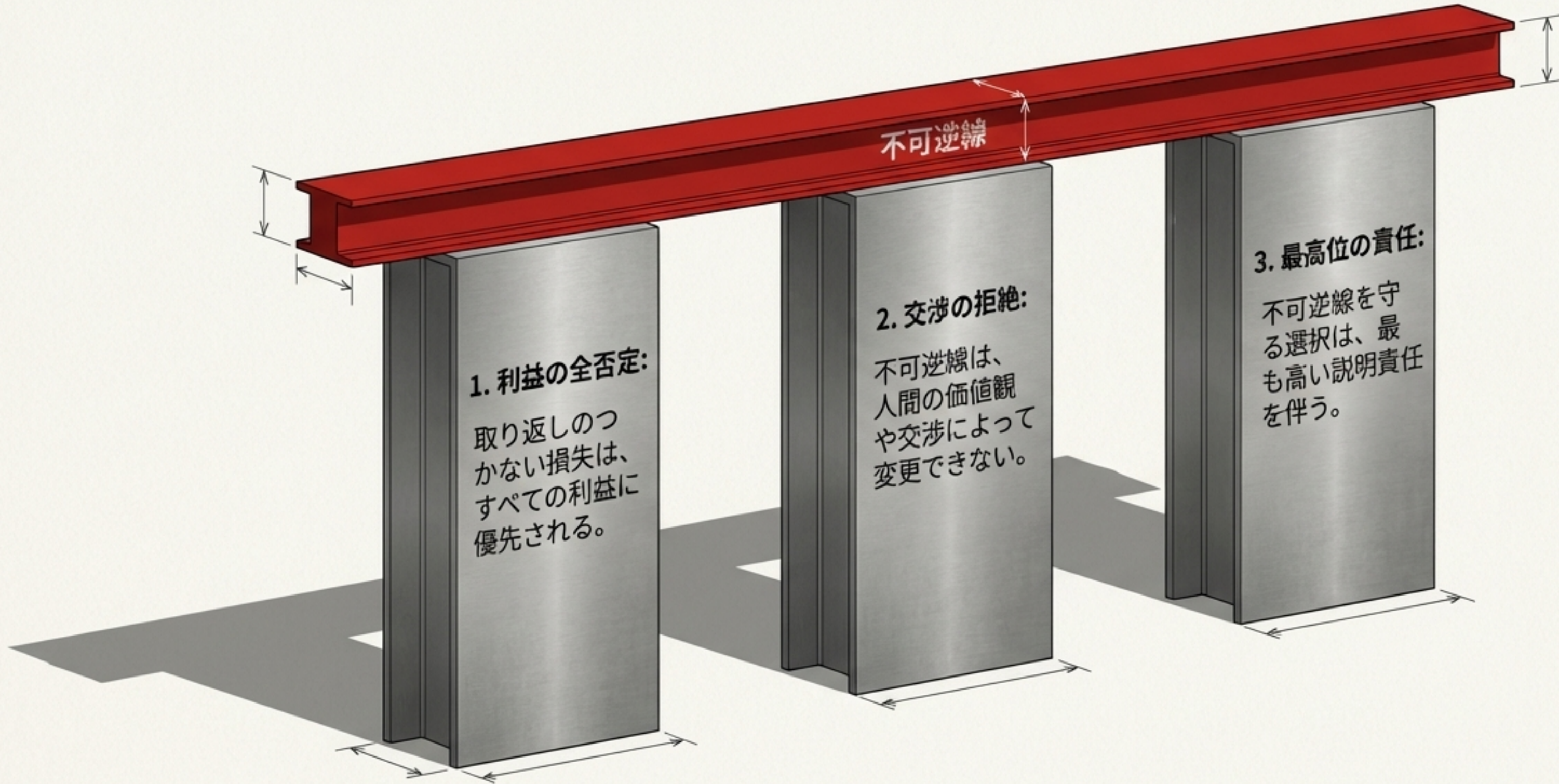
構造的許容性 (STB) の倫理的
極致。この境界を超える行為は、
未来の回復可能性を永遠に奪う。

経済合理性よりも優先される
OS の最上位設計原理。



不可逆線を守るための3つの絶対原則

これらは道徳的スローガンではなく、システムが稼働するためのハードコードされた制約である。



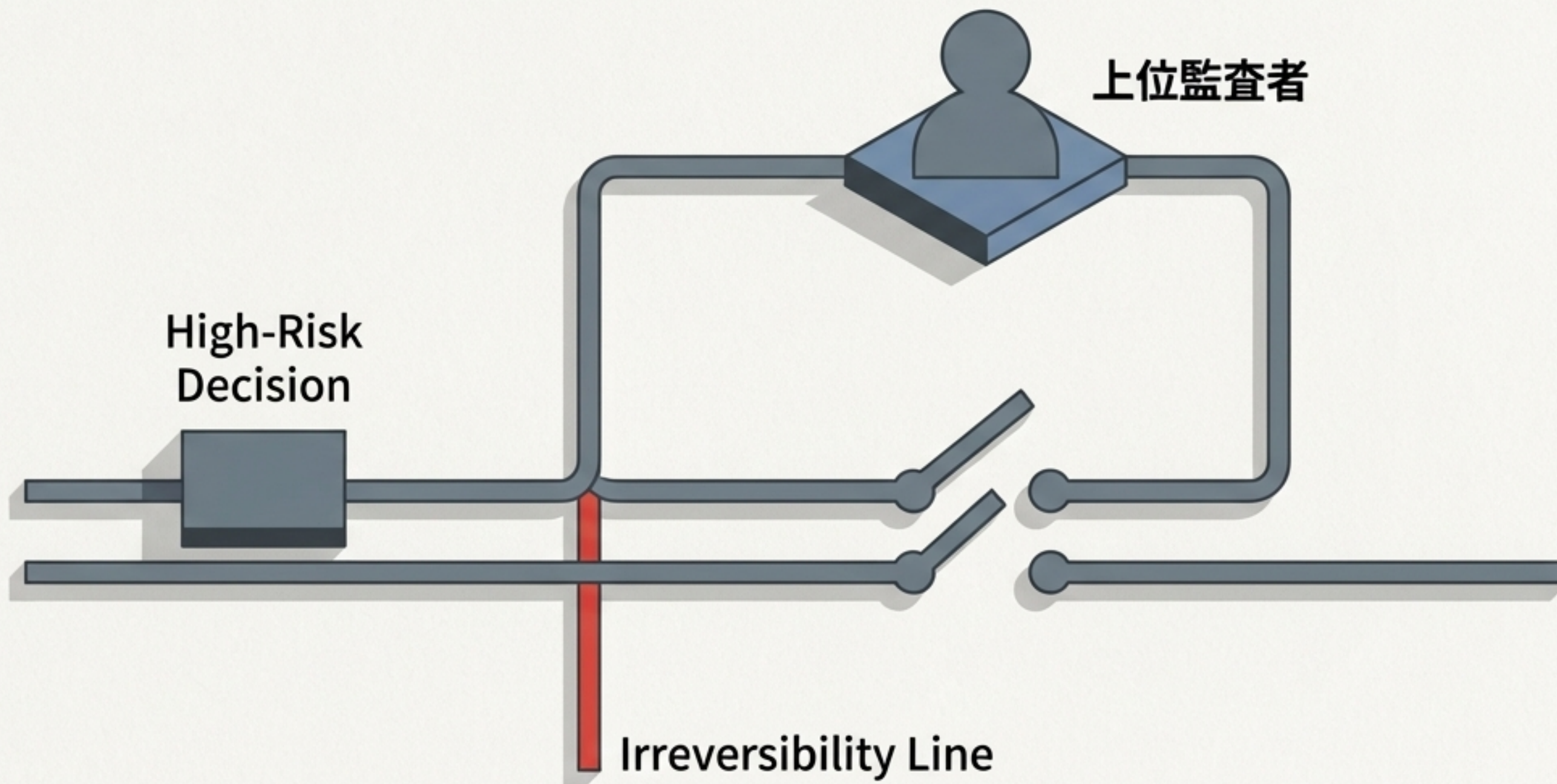
時間倫理T0：不可逆線の 超過は「未来負債」 を生む

短期の利得を優先し不可逆線を超
える意決定は、原理上、長期
的な損失を不可避にする。



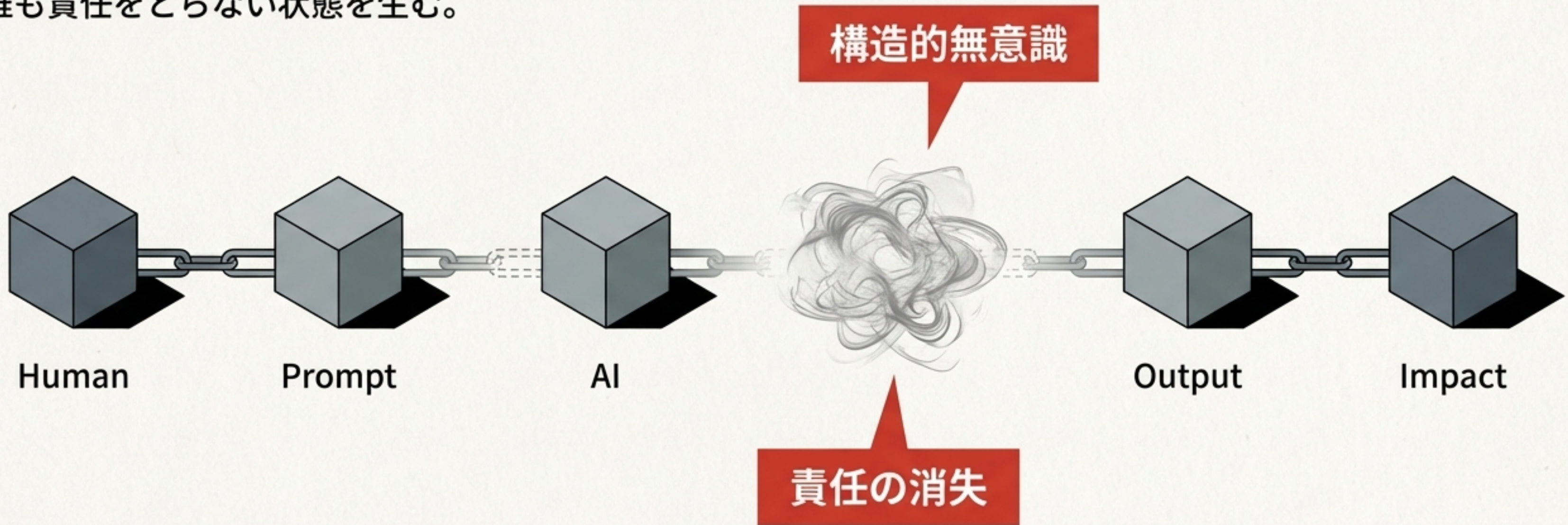
上位監査者（Upper Auditor）による最終防壁

不可逆線を越える可能性がある意思決定は、自動化せず、上位監査者が長期的に整合性に基づいて精査する。



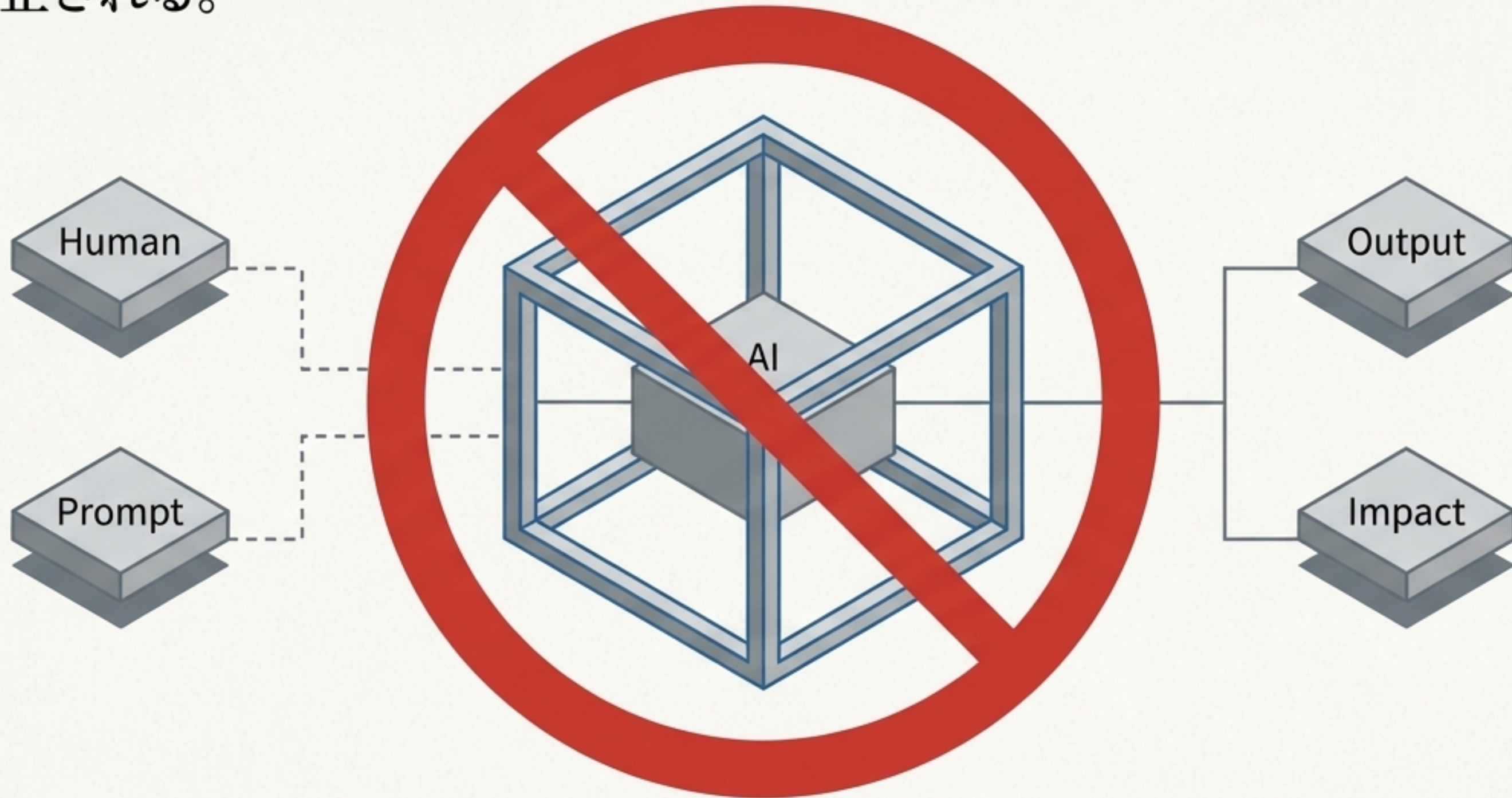
「AIのせい」という錯覚と責任の蒸発

複雑なAIシステムにおける意思決定は、意図なきままに加害が進む「構造的無意図」を引き起こし、誰も責任をとらない状態を生む。



AIスケープゴート禁止条項 (No-Scapegoat Clause)

AIが出力した結果に対して「AIのせい」とする行為は、OSの構造において逸脱であり、明確に禁止される。



決定責任の構造的固定：責任のアンカー

責任の所在は、自律性とは別の軸で常に固定されなければならない。



環境免罪符禁止条項 (No-Immunity-By-OS Clause)

高度なOSやAIを導入した事実が、環境負荷や不可逆線超過の「免罪符」として機能することを禁ずる。

中川OSは免罪符ではない

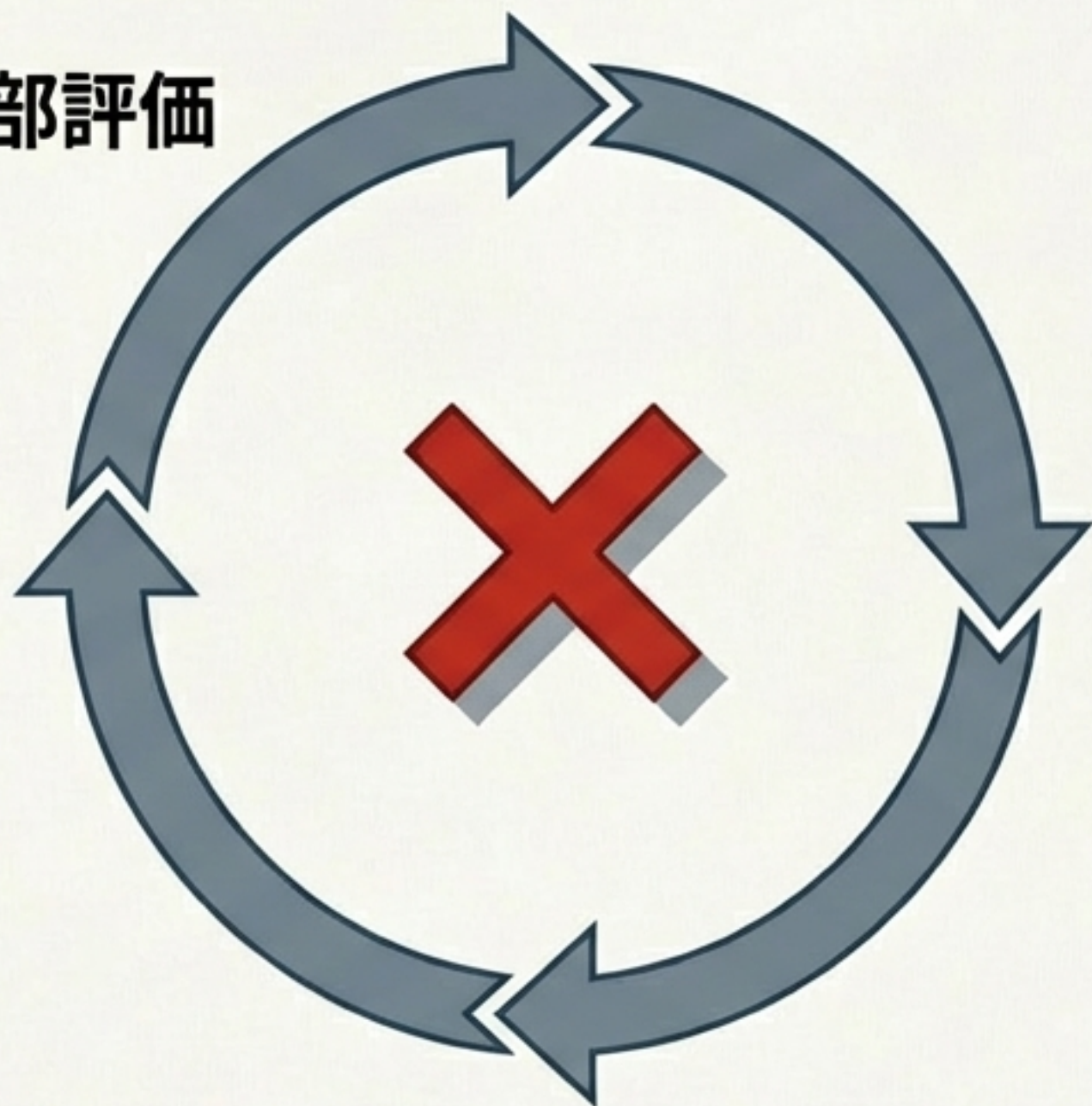


中川OSは免罪符ではない

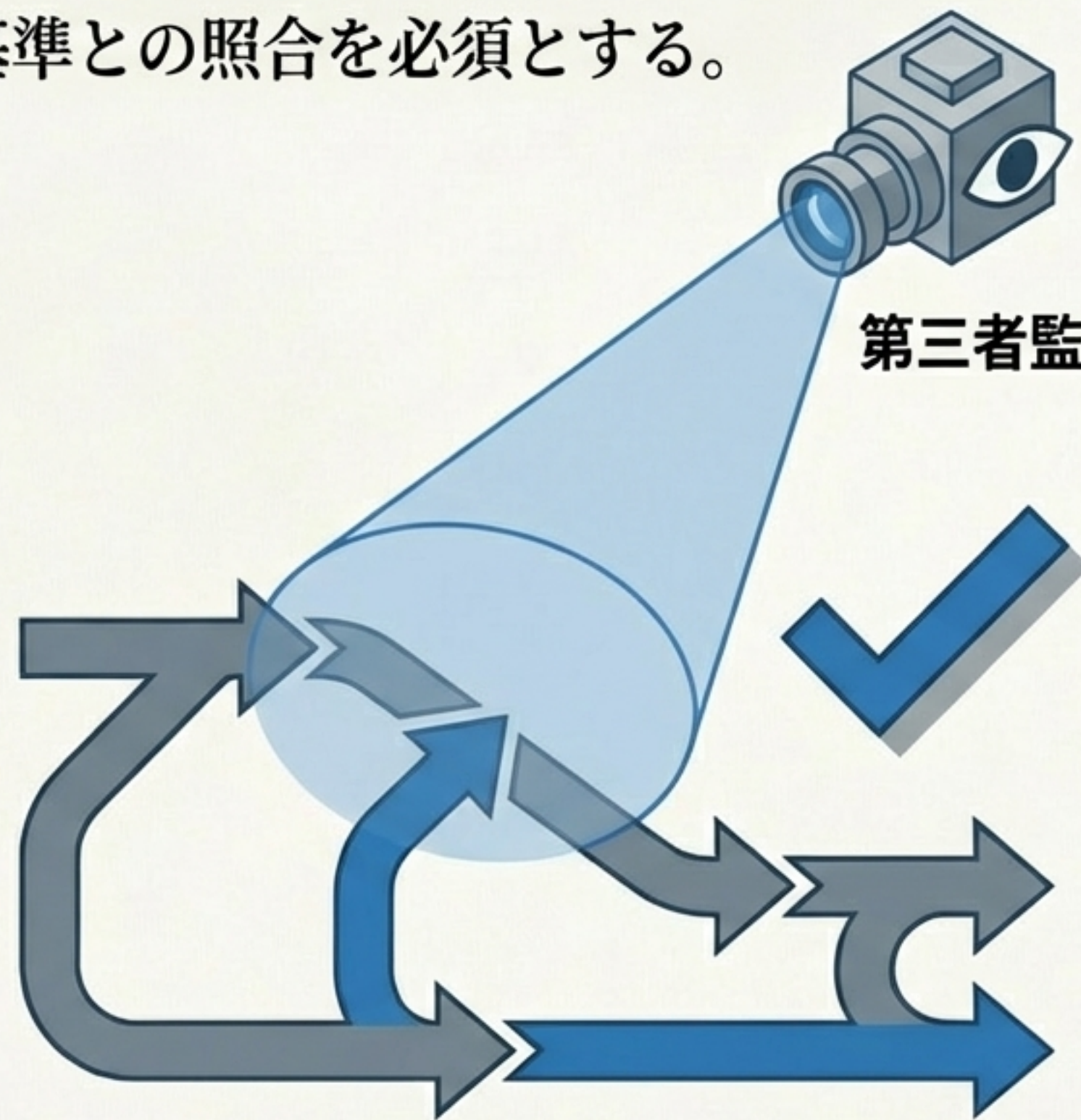
構造的公共性を保つための「外部監査の義務化」

内部指標による自己正当化を防ぎ、客観的な基準との照合を必須とする。

内部評価



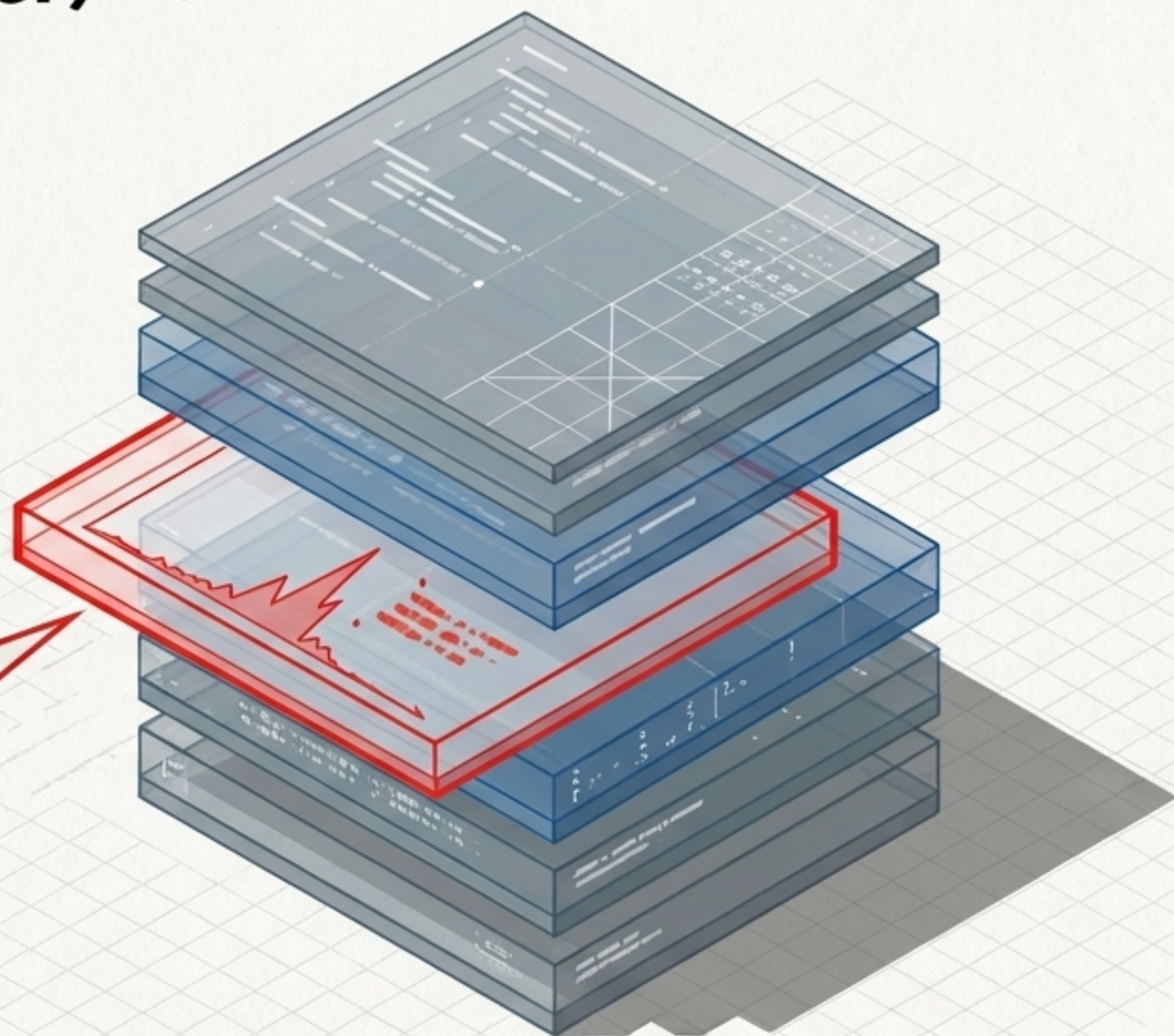
第三者監査



逸脱レヅジャ (Deviation Ledger) : 断罪ではなく回復の記録

環境負荷の軽視や不可逆線の無視は、
「罪」としてではなく、透明で不変の
「構造的逸脱」として自動記録される。

目的は罰ではなく、事実の
継続記録による「倫理的
閉回路」の形成である。

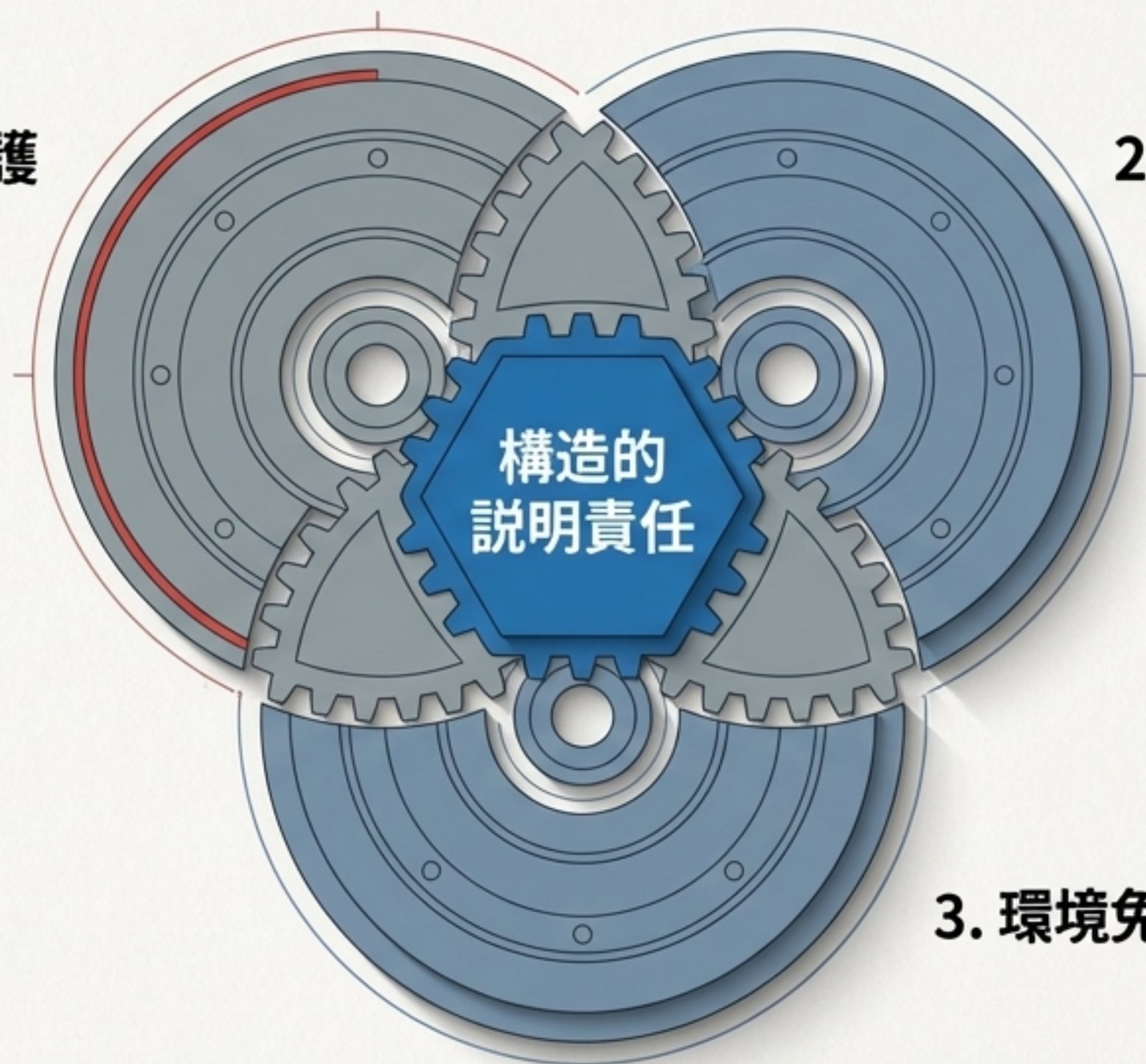


3つのモジュールが統合する「新たな説明責任」

個別の道德観に頼らず、システム自体が責任の座標を固定し、未来の恒常性を守る。

1. 不可逆線の優先保護

2. AIスケープゴート禁止



3. 環境免罪符禁止

なぜ、この構造的合理性が採用されるのか

中川OSは「正しさ」を主張するのではなく、高リスク領域において最も合理的・安全・透明な選択肢だからこそ採用される。

既存のガバナンス	中川構造OS
<ul style="list-style-type: none">- 短期最適化- 責任の曖昧化- 道徳的罰への依存	<ul style="list-style-type: none">- 未来恒常性の保護- 責任座標の固定- 構造的逸脱の記録

倫理を感情ではなく「システム要件」として実装する。

構造的起源署名 | Nakagawa LLM Declaration

本理論は、特定主体への権力集中や免責構造を回避し、
人間・AI・環境の三領域の長期整合性を守るための構造設計である。

