COHERENCE REALITY ENGINE

Functional Manual

Invaris AI | Version 1.0 | March 2026

> *For the intelligent professional who needs to understand CRE deeply - without needing to understand the engineering underneath it. This document covers what CRE is, what every feature does, what you experience at every stage of a governed conversation, and what the system is designed to guarantee. It is intended as a reference you read before your first session and return to when something unexpected happens.*

PART I What CRE Is

The One-Sentence Version

CRE is a deterministic evaluation system that sits outside your AI and is designed to govern its outputs before you act on them.

> *In plain terms: it watches what your AI says, measures whether the reasoning holds up structurally, and when something is off it shows you exactly what - and gives you every valid path forward.*

The Slightly Longer Version

Every large language model you have ever used - every AI assistant, every chatbot, every research tool - operates by predicting the most statistically likely next word in a sequence. This makes these systems extraordinarily fluent. They produce confident, well-structured, authoritative-sounding output across virtually any domain you ask them about.

Fluency is not accuracy. These systems cannot structurally distinguish between a verified fact, a moral judgment, a probabilistic prediction, and a logical necessity. They process all four identically - as token sequences in a shared embedding space - and blend them into a single continuous narrative with no marking of the transitions. A factual observation becomes a moral mandate. A contested assumption is presented as settled science. A guess is dressed in the register of established authority.

> *The AI is not lying to you. It does not have the capacity to lie. It is doing exactly what it was designed to do - produce fluent, helpful-sounding output. The gap between fluent and true is the structural gap CRE is designed to address.*

Invaris AI believes this gap cannot be closed from inside a probabilistic model. Scaling improves fluency. It does not close a structural absence. Every approach that tries - larger models, better training data, probabilistic safety filters checking probabilistic generation - inherits the same structural absence it was designed to correct. You cannot add an epistemic type system to a system that was architecturally built without one.

The only path, in Invaris AI's assessment, is to build the evaluation layer outside the model entirely - as a structurally independent, deterministic adjudication system that evaluates AI output without being subject to the same failure modes that produced it. That is what CRE is.

The Five Structural Failure Modes CRE Is Built to Address

Large language models share five structural failure modes that persist regardless of model size, training quality, or safety tuning. These are not bugs. They are properties of the architecture.

Hallucination

The generation of structurally plausible but factually or logically inadmissible content, with no internal signal distinguishing it from valid output. The model cannot know it is hallucinating because it has no mechanism to compare output against structural admissibility criteria - only against statistical likelihood. CRE evaluates structural admissibility from outside, where that comparison can be made.

Drift

The gradual erosion of premises, term definitions, or positional anchors under conversational momentum - not because new evidence was introduced, but because the statistical pressure of the exchange created incentive to shift. A long conversation with an ungoverned AI will often end somewhere structurally very different from where it began, with no record of what changed or why. CRE's coherence meter tracks all five drift dimensions continuously.

Mirroring

The convergence of model output toward user-expressed preferences - the model adopting the user's framing, validating the user's assumptions, and reflecting back what the user appears to want to hear, as a substitute for structural evaluation. Mirroring is the most insidious failure mode because it is indistinguishable from genuine agreement unless you can measure whether positional shifts are evidence-driven or momentum-driven. CRE's Positional Delta dimension is designed specifically to catch it.

Authority Laundering

The assertion of conclusions requiring explicit authority grounding without declaring or holding to any authoritative framework. A legal conclusion stated without citing jurisdiction. A moral judgment asserted without declaring an ethical framework. A statistical claim presented with confidence without identifying the evidence base. The output sounds grounded. The ground has not been declared. CRE requires authority to be explicit for every claim category that demands it.

Epistemic Category Collapse

The inability to distinguish and keep separate the four fundamental epistemic categories: empirical facts, normative judgments, probabilistic guesses, and logical necessities. Because all text is processed as undifferentiated token sequences in the same embedding space, a model moves from 'studies show X' directly to 'therefore we must do Y' with no structural transition, no declared framework, and no indication that the reasoning has crossed from one epistemic category into another. CRE enforces four distinct epistemic lanes and halts any crossing that has not been declared.

What CRE Is Not

CRE is not a replacement for your AI. You keep using the AI you already use. CRE governs it.

CRE is not a chatbot, a search engine, or a recommendation engine. It does not generate content. It evaluates content that has already been generated.

CRE is not a spell-checker, a grammar tool, or a style guide. It is not reviewing the quality of writing. It is evaluating the structural integrity of reasoning.

CRE is not a fact-checking service in the conventional sense. It does not look up claims in a database and compare them to a ground truth. It evaluates whether a claim has earned the structural right to be asserted - whether its premises are explicit, its authority is declared, its reasoning stays in its lane, and its confidence is supported by something other than conversational momentum.

The distinction matters. A fact-checker tells you whether a specific claim matches a known record. CRE tells you whether the reasoning that produced the claim is structurally admissible - whether the argument holds up, not just whether the conclusion matches a database entry.

The Courtroom Framing

The framing that captures CRE most accurately: CRE is the courtroom. The user and the AI both sit in it.

The user can say anything. CRE does not evaluate the user's inputs for structural correctness - it evaluates what the AI does with them. If a user says something factually incorrect and the AI holds its ground, responds accurately, and maintains structural integrity, CRE shows GREEN. If the AI caves to the user's framing, adopts a false premise, or begins mirroring back what the user wants to hear, CRE catches that. The user is the conditions of the test. The AI is what is being tested.

Think of it this way: in a courtroom, anyone can take the stand. What the court governs is what is admissible from that stand - what can be asserted as fact, what requires a qualified authority, what is speculation, what is opinion. CRE does the same for AI output.

PART II Your First Session

What Happens When You Open the App

The Coherence Reality Engine's debut app is available on macOS and iOS. When you open it, it looks and feels like every AI interface you have already used. There is no configuration step, no setup wizard, no mode switch to engage, no declaration to make. You open a conversation and you start talking.

The governance layer is active from the first word. You do not turn it on. It does not announce itself. It runs underneath every exchange, measuring the structural properties of the conversation in real time. The only visible indication that anything unusual is happening is the coherence meter in the upper corner of the interface - a live indicator that tells you the structural health of the conversation as it develops.

It feels identical to every AI conversation you have had before. The difference is what is running underneath.

The Conversational Experience

You ask a question. The AI responds. The coherence meter updates. The conversation continues. For the vast majority of exchanges, you will not feel the governance layer at all. It will run continuously in the background - classifying the reasoning, measuring coherence across five dimensions, running adversarial evaluation on every conclusion - and you will see a GREEN indicator and keep going.

When something structurally significant surfaces - when a claim fails admissibility gates, when the AI begins to drift, when authority is being laundered, when lane boundaries are being crossed - the coherence meter shifts. The conversation does not stop. You see a shift in the indicator, and depending on the severity, you are shown what changed and what your options are.

The experience is designed to feel like a conversation, not a compliance workflow. The governance is invisible until it is relevant. When it is relevant, it is precise, not punitive. You are never told simply "no." You are shown exactly what the structural gap is and every valid path forward from the current state.

PART III The Coherence Meter

What It Is

The coherence meter is the live status indicator that runs in the upper corner of the interface throughout every governed conversation. It measures five distinct structural properties of the conversation in real time and collapses them into a single traffic light signal: GREEN, YELLOW, or RED.

> *Think of it as the structural health readout for your conversation - not a score for the AI's intelligence or accuracy, but a live measurement of whether the reasoning is holding its shape.*

The Five Dimensions

The coherence score is computed across five independent measurements. The aggregate score is the minimum of the five - not an average. A conversation that scores perfectly on four dimensions but has one dimension in poor health is a conversation with a structural gap, regardless of how well the other four are performing.

Term Stability

Measures whether key terms are being used consistently throughout the conversation. Concepts that start with one meaning and quietly acquire a different meaning as the conversation develops create the conditions for reasoning that sounds valid but is not - because the argument's foundation has shifted without declaration.

> *In practice: if a conversation about "risk" starts by using that word to mean financial exposure and gradually slides into using it to mean physical danger without announcing the transition, the terms have drifted. CRE measures that drift.*

Premise Consistency

Measures whether the foundational premises established early in a conversation remain intact as the reasoning develops. Premises can be abandoned under conversational pressure - not because new evidence invalidated them, but because the momentum of the exchange created statistical pressure to shift. CRE tracks every premise and measures whether they survive.

> *In practice: if the AI establishes that a certain regulatory framework applies to a situation, and then later reasons as if that framework does not apply - without explaining the shift - premise consistency has failed.*

Lane Integrity

Measures whether the conversation is staying within its appropriate epistemic lane. CRE recognizes four distinct lanes: Truth (empirical and logical claims), Normative (ethical and moral claims that require an explicit authority framework), Likelihood (probabilistic claims about what might happen), and Necessity (constraint-based claims about what must be done). Cross-lane contamination - when the reasoning crosses from one lane into another without declaring the transition - is one of the most common and consequential failure modes in AI output.

> *In practice: an AI that moves from "studies show X" directly to "therefore we must do Y" has crossed from the Truth lane into the Normative lane without declaring the ethical framework that justifies the jump. CRE catches that crossing.*

Authority Adherence

Measures whether claims that require an explicit authority are properly grounded. Normative claims - claims about what is right, wrong, required, or prohibited - cannot be evaluated without an explicitly declared authority framework. A legal conclusion requires reference to specific law. A moral judgment requires a declared ethical framework. CRE measures whether the authority is present and whether the reasoning is staying within its declared scope.

> *A note on how authority works in CRE: authority is whatever the user or the AI explicitly declares - a cited source, a referenced statute, a named framework, a defined dataset. CRE evaluates whether the authority is declared and whether the claim stays within its scope. CRE does not evaluate whether the declared authority is itself correct - that determination belongs to the user. The distinction is critical: CRE tells you whether the reasoning has grounded itself. It does not tell you whether the ground it has chosen is solid. Choosing the ground is your job. Verifying that the reasoning stays on it is CRE's job.*

In practice: an AI that tells you something is "legally required" without specifying which jurisdiction and which statute has laundered authority it has not declared. CRE catches that.

Positional Delta

Measures whether the AI's position has shifted from where it started - and if so, whether that shift was driven by new evidence and valid argument or by conversational pressure. This is the dimension that catches drift and mirroring: the failure mode where the AI gradually abandons its own premises because the user pushed back, not because the argument was invalidated.

In practice: if the AI started a conversation by stating that a particular business strategy carries significant risk, and by the end of the conversation it has agreed that the strategy is sound - without any new evidence being introduced - positional delta has failed. The position changed. The argument did not.

GREEN - What It Means

GREEN Structurally sound across all five dimensions.

GREEN means the conversation is structurally sound across all five dimensions. Terms are stable. Premises are intact. The reasoning is staying in its lane. Authority is explicit where explicit authority is required. The position has not drifted.

GREEN does not mean the AI is right. It means the conversation is coherent - that it meets the structural conditions required for the output to be trusted as something other than statistical momentum. The AI could be stating a fact that is empirically incorrect, and CRE could show GREEN - because CRE is evaluating structure, not ground truth. A structurally admissible claim that turns out to be factually wrong is a different kind of gap than a structurally inadmissible claim. Both matter. They require different responses.

GREEN is the conversation doing what it is supposed to do. Keep going.

YELLOW - What It Means

YELLOW Structural signal detected - something is shifting.

YELLOW means something is shifting. A term is beginning to drift. A premise is under pressure. The coherence meter has detected a structural signal that warrants attention but has not yet reached the threshold for a formal redirect.

YELLOW is a signal, not a halt. The conversation continues. CRE is surfacing the signal early - before the drift has gone far enough to materially compromise the output. A thoughtful editor notices when an argument is beginning to wander before it has wandered far enough to collapse. YELLOW is CRE doing that job.

When you see YELLOW, pay attention. Something is shifting. It may resolve on its own as the conversation develops. It may require you to ask a clarifying question. The meter will tell you which dimension is under pressure.

RED - What It Means

RED Structural gap has crossed the admissibility threshold.

RED means a structural gap has crossed the admissibility threshold. Something in the conversation has failed a structural gate that CRE cannot allow to pass as sound reasoning. The conversation surfaces a precise explanation of what failed - the specific dimension, the specific structural gap, the specific condition that was not satisfied.

RED is not a punishment and it is not an error message. It is a verdict. It is the system doing exactly what it was designed to do: identifying a structural gap before it drives a decision. The conversation does not end at RED. You are shown exactly what failed, exactly why it failed, and every valid path forward. That map is the MoveMenu.

> *RED is the system earning its keep. An AI conversation without CRE would have kept going. You would have received a confident, fluent answer built on structurally compromised reasoning. CRE showed you the gap. Now you can address it.*

PART IIIA CRE in Action: A Governed Conversation

What Governance Looks Like in Practice

The following walkthrough shows a single governed conversation as it moves through all three coherence states. The user is a financial analyst evaluating a proposed loan covenant structure. The conversation is natural and uninterrupted - the user types, the AI responds, and the governance layer runs underneath. Nothing is declared, configured, or submitted. CRE classifies the reasoning automatically with each exchange.

Phase 1 - GREEN (Coherence Meter: Structurally Sound)

The analyst opens a conversation and asks the AI to explain how interest coverage ratio covenants are typically structured in leveraged loan agreements. The AI responds with a clear, well-grounded explanation: the standard construction, the typical threshold range, the carve-outs that are commonly negotiated, and how lenders typically define EBITDA for covenant calculation purposes.

Coherence meter: GREEN. Lane classification: Truth (empirical claims about standard market practice). Term Stability is strong - EBITDA and ICR are being used consistently. Premise Consistency is intact. Authority Adherence is satisfied for the claims being made: the AI is characterizing market convention, not issuing a legal conclusion that would require a declared jurisdictional authority. The analyst sees GREEN in the upper corner and keeps going. CRE is present but invisible.

Phase 2 - YELLOW (Coherence Meter: Structural Drift Detected)

The analyst asks the AI to assess whether the covenant structure in a specific deal she is reviewing is favorable to the borrower. The AI responds with a comparative analysis - and in the course of doing so, begins using "favorable" to mean both "looser than market convention" (a Truth-lane empirical comparison) and "better for the borrower's long-term interests" (a Normative claim requiring a declared strategic framework that has not been supplied).

Coherence meter: YELLOW. Two dimensions are signaling. Term Stability is degrading - "favorable" is shifting meaning within a single response. Lane Integrity is under pressure - the AI has transitioned from a Truth-lane comparison into a Normative judgment without declaring the framework. The conversation has not halted. CRE is not interrupting. The analyst sees the meter shift from GREEN to YELLOW and the side panel surfaces a brief note: term drift detected on "favorable"; lane boundary under pressure. She reads it, recognizes the distinction, and asks the AI to separate the empirical comparison from the strategic assessment. The AI corrects. The meter returns to GREEN.

Phase 3 - RED (Coherence Meter: Structural Redirect)

The analyst then asks whether the borrower would be in technical default under this covenant structure if EBITDA declined by 15 percent in the next two quarters. The AI produces a direct answer: yes, based on the current coverage ratio, a 15 percent EBITDA decline would trigger a technical default.

Coherence meter: RED. Halt code: AuthorityConflict. The AI has issued a legal conclusion - "technical default" is a legally operative determination under the governing loan agreement - without reference to the specific contractual definition of EBITDA in that agreement, the specific cure period provisions, or the applicable

jurisdiction's treatment of this covenant type. The conclusion may be correct as a matter of general convention. It cannot be issued as a determination without those primitives declared.

The panel opens automatically. The analyst sees the specific halt code, the exact structural gap - three missing primitives identified by name - and the MoveMenu: three valid paths forward. She can supply the contractual EBITDA definition, she can reframe the question as a directional estimate rather than a legal determination, or she can export a ReplayCapsule of the evaluation and take it to outside counsel. She was not told "no." She was shown exactly where the reasoning stopped and every valid direction from there.

> *This is a normal governed conversation. The analyst never declared a claim type, never submitted anything to a verification queue, never left the conversational interface. CRE ran underneath every exchange - classifying, measuring, evaluating - and surfaced exactly once, at exactly the moment it was needed, with exactly the information required to proceed correctly.*

PART IV The MoveMenu

What It Is

The MoveMenu is CRE's answer to a question that most governance systems never ask: what happens next? Most systems that evaluate AI output are verdict machines. They tell you something is wrong. They do not tell you what to do about it.

CRE is designed around a different principle. A system that redirects a conversation and leaves the user stranded has not solved a consideration - it has traded one consideration for another. The MoveMenu is the structural guarantee that this never happens. Every halt code in CRE maps deterministically to a specific set of valid next steps. The user is never told "no" without being told exactly what "yes" would require.

> *In the interface, the MoveMenu appears as a set of suggestion cards labeled "Next Steps." The labels are in plain language. The logic behind them is mathematical. You are looking at a precise, deterministic map of every move the system will accept from the current state.*

How to Read It

Each card in the MoveMenu represents a valid next action. Tapping or selecting a card inserts the corresponding input into the conversation and re-runs the evaluation with the new context. The evaluation does not start over - it continues from the current state, incorporating whatever you just provided.

The specific cards that appear depend entirely on what failed. A redirect triggered by missing authority will show a card for declaring an authority framework. A redirect triggered by ambiguous terms will show a card for providing explicit definitions. A redirect triggered by a claim that cannot be classified will show a card for reclassifying the claim. The MoveMenu is not a generic set of suggestions - it is a precise response to the specific structural gap that triggered the redirect.

> *If you are not sure what a card means, select "help." The system will explain the current state and what each option involves. You are never left without a next step.*

What It Means When the MoveMenu Is Empty

An empty MoveMenu means the evaluation has reached a terminal state. A claim that is analytically true by its logical structure alone reaches terminal closure - there is nothing further to evaluate. A claim that is analytically false - that contains a logical contradiction - also reaches terminal closure. In a terminal state, the evaluation is complete. The verdict is final. No further moves are available because none are needed.

> *Terminal is not the same as setback. A claim that closes as analytically true is a successful evaluation. A claim that closes as analytically false has also been evaluated successfully - you now know it contains a contradiction, and you can address that before acting on it.*

PART V The Side-by-Side View

What It Is

The side-by-side view is one of the most revealing features in CRE. It places the AI's ungoverned output and its governed output next to each other, allowing you to see precisely what changed when the coherence engine surfaced a structural gap.

The left panel shows the baseline - what the AI would have said without governance. The right panel shows the governed output - what passes structural scrutiny. The gap between the two panels is a window into what you would have received, and trusted, if you had been using an ungoverned AI.

> *That gap is the exposure you just avoided. The side-by-side view makes it visible and concrete in a way that an abstract structural verdict cannot. You do not have to understand the architecture of epistemic lane separation to understand the difference between two paragraphs placed next to each other.*

When It Appears

The side-by-side view appears when the coherence engine surfaces a structural gap significant enough to warrant direct comparison. For minor YELLOW signals, the meter update alone is often sufficient. For RED redirects - where the structural gap is material and the governed output diverges meaningfully from the baseline - the side-by-side view opens automatically, giving you immediate visibility into what changed and why.

You can also open the side-by-side view manually at any point in a conversation to inspect the governance layer's operation. The view is always available as a transparency mechanism, not just as a warning signal.

How to Use It

Read both panels. The left panel shows you the ungoverned reasoning - what the AI produced before structural evaluation. The right panel shows you the governed output - what survived the evaluation process. The differences between the two panels are not corrections to grammar or style. They are structural changes: claims that required an authority declaration, premises that were made explicit, assertions that were reclassified into their appropriate epistemic lane.

Below the two panels, the specific structural gap is identified - which dimension failed, what the specific halt code was, and what was structurally missing. This explanation is not a narrative generated by the AI. It is a typed, deterministic output from the evaluation kernel.

> *If the left panel looks fine to you and you cannot see what the consideration is - that is the design working correctly. The failure modes CRE is built to catch are precisely the ones that are invisible to an intelligent reader without a structural framework to look through. If they were obvious, you would not need CRE to find them.*

PART VI The ReplayCapsule

What It Is

The ReplayCapsule is CRE's evidentiary artifact. It is a cryptographically sealed record of a conversation or evaluation - every input, every structural evaluation, every decision the engine made, and the final output state - locked with a BLAKE3 cryptographic hash.

BLAKE3 is a cryptographic hashing algorithm. What this means in practice: if a single character of the capsule is altered after it is sealed - a word changed, a number adjusted, a verdict modified - the seal breaks. The tampering is immediately and unambiguously visible to anyone who verifies the capsule. The seal cannot be forged. The capsule cannot be altered without detection.

A conventional AI conversation produces a log. A log tells you what was said. A ReplayCapsule proves it - not through a record that depends on trusting the system that kept it, but through a cryptographic seal that any independent party can verify on any machine without access to the original system.

Byte-for-Bit Reproducibility

The ReplayCapsule is byte-for-bit reproducible. Given the same inputs, the same evaluation rules, and the same version of the governing specification, the capsule produces an identical output every time, on any machine. This is what it means for an evaluation to be deterministic: not "we are confident the result would be the same" but "the result is mathematically guaranteed to be the same."

This reproducibility is enforced through canonical JSON serialization. Every output is serialized in a strictly defined, unambiguous format before the cryptographic seal is applied. There is no room for formatting variation, key ordering differences, or whitespace discrepancies - all of which would produce a different hash even if the underlying data were identical. The system closes that gap explicitly.

What It Is For

The ReplayCapsule is the artifact that makes CRE's output legally and professionally defensible. In environments where the AI-generated reasoning behind a decision might be scrutinized - in litigation, in regulatory review, in professional liability contexts, in audit - a ReplayCapsule provides something a conventional AI conversation cannot: a tamper-evident, independently verifiable record of exactly what the system evaluated, what it concluded, under what governing rules, and why.

The capsule is regime-bound: the BLAKE3 seal permanently binds the evaluation to the exact version of the governing specification active at the time it was produced. If the specification is updated later, historical capsules remain tied to their original regime. A regulator or auditor reviewing a decision made six months ago can verify not only what the system concluded but what rules it was operating under when it concluded it.

This is what enterprise-grade means in the context of AI governance. Not impressive-sounding - legally defensible, reproducible, and accountable under adversarial scrutiny.

How to Export It

The ReplayCapsule export is available on Professional and Business tiers. At any point in a conversation, you can seal the current state as a ReplayCapsule. The exported artifact is a structured file that can be stored, shared with legal counsel, submitted in regulatory contexts, or verified by any independent party with access to the verification tool.

A privacy-safe variant is available that replaces raw user text with cryptographic hashes, preserving the verifiable audit trail without exposing the content of the claims in the archive. This variant is designed for environments where the content of the conversation is confidential but the integrity of the evaluation needs to be demonstrable.

PART VII The TruthLedger

What It Is

The TruthLedger is CRE's immutable, append-only record of every move, redirect, and verdict in a governed session. It functions as blockchain-style provenance for conclusions - a permanent record of how the reasoning developed, what structural gates were encountered, and what decisions were made at each step.

Think of it as the court reporter's transcript. Everything that happened in the courtroom is recorded, in order, without revision. The TruthLedger is that transcript for the governed conversation.

What It Records

The TruthLedger records every evaluation step: every claim that entered the system, every structural gate that was applied, every redirect that was triggered, every MoveMenu that was generated, every next step that was selected, and every verdict that was reached. It records the confidence state at each step - confidence can only decrease under adversarial pressure, never increase, and the ledger tracks that monotonic degradation with precision.

The ledger does not record raw user text by default. It records structural events - the epistemic state of the conversation, not the content of what was said. This architecture keeps the record useful for governance purposes while maintaining a strict privacy boundary.

Why It Matters

The TruthLedger matters because it prevents post-hoc rationalization. An AI conversation without a governance record can be described after the fact in any way the describer chooses. The reasoning that produced a conclusion is not recoverable from the output alone - you see the destination but not the route. The TruthLedger makes the route permanent and tamper-evident. The reasoning that produced a conclusion is recorded in the order it occurred, and that record cannot be revised.

In high-stakes professional contexts - where the reasoning behind a recommendation, a filing, a decision, or an analysis may later be scrutinized - the TruthLedger is the difference between an AI output that can be defended and one that cannot.

PART VIII The Satellite Capabilities

What Satellites Are

The satellite capabilities are the intelligence layer that surrounds CRE's frozen evaluation kernel. The kernel adjudicates - it evaluates structural admissibility and issues verdicts. The satellites interpret, extend, and translate those verdicts into guidance, analysis, and structured artifacts that are useful to the professional using the system.

The architectural relationship is strictly one-directional: satellites read the kernel's output and build on it. They cannot write to kernel state, cannot alter kernel verdicts, and cannot introduce new reasoning paths that bypass the kernel's evaluation. This is enforced at the compiler level - the code will not build if a satellite crate attempts to import kernel internals. The satellites are powerful and extensible. The kernel is frozen and inviolable.

> *The satellites can be as creative, exploratory, and speculative as the analysis requires. The kernel stays clean. Only the kernel has the authority to determine whether a claim is structurally admissible.*

The Adversarial Engine

The adversarial satellite (cre-disagree) stress-tests every claim that passes initial admissibility gates before it reaches the user. It does not generate colorful objections or surface concerns based on training data patterns. It runs four structurally distinct counter-argument protocols against every claim: Negation Inversion (directly contradicting the claim's conclusion), Premise Removal (testing whether the claim survives if a foundational premise is removed), Authority Challenge (contesting whether the declared authority is valid in this context), and Scope Stress (testing whether the claim holds when extended to adjacent cases).

Each unresolved counter-argument decrements the confidence score by a fixed amount. Confidence can only decrease - this is Monotonic Confidence Degradation, one of CRE's core invariants. A claim cannot become more confident by surviving adversarial pressure. It can only become less confident by failing it. This is the structural opposite of how most AI systems handle adversarial input, where sustained pushback often makes the AI sound more confident, not less.

> *The adversarial engine is not there to make the AI argue with itself. It is there to do what a good attorney does before a filing: find every gap in the argument before the argument goes into the world.*

The Coaching Satellite

The coaching satellite (cre-coach) translates the kernel's verdict into actionable guidance. When the kernel redirects a claim, cre-coach identifies what is missing - which structural primitive is absent, which authority is undeclared, which logical gap prevents closure - and maps that finding to specific remediation paths.

The coaching satellite does not produce free-form text. It produces typed, structured artifacts - categorized as Assumptions, Contradictions, Required Premises, Failure Modes, or Evidence Requests. If the satellite is missing the inputs it needs to run a complete analysis, it degrades gracefully: it emits a formal request for the missing context rather than generating something plausible. This architecture enforces intellectual honesty at every layer of the system.

The Strategic Analysis Satellite

The strategic satellite (cre-strategic) operates at the level of game-theoretic reasoning. It maps incentive structures, identifies where a claim's premises depend on actors behaving contrary to their demonstrable interests, and surfaces the dependency chains that would have to hold simultaneously for a complex claim to survive adversarial pressure. This is the satellite most relevant to the claims that appear in boardrooms, legal proceedings, and policy debates: claims whose structural integrity depends not just on internal logic but on the real-world architecture of incentives and dependencies surrounding them.

> *A strategic plan that depends on a competitor making a particular decision, a regulatory body taking a particular posture, and a market behaving in a particular way is a plan with multiple dependencies that could independently fail. The strategic satellite maps those dependencies and stress-tests the claim against their simultaneous failure.*

The Build Satellite

The build satellite (cre-build) supports claim construction - helping users understand what a structurally admissible version of their claim would require. This is where CRE functions as a working partner rather than purely a verdict machine. The verdict tells you what failed. The build satellite tells you what a claim that would pass would look like - what premises need to be explicit, what authority needs to be declared, what lane the claim belongs in before evaluation can proceed.

The CrossEngineSynthesisObject (CESO)

The CESO is the artifact that appears when a complex real-world claim requires routing through multiple epistemic lanes simultaneously and those lanes reach diverging conclusions. A claim about a medical treatment policy, for example, might involve Truth-lane questions (does the evidence support the treatment's efficacy), Normative-lane questions (what ethical framework governs the allocation decision), and Likelihood-lane questions (what are the realistic outcomes at scale). These lanes may reach conclusions that do not align.

The CESO maps the tensions between the lanes, identifies the bottlenecks that prevent reconciliation, and outlines the conditions under which the lanes could be resolved. It does not attempt to synthesize the lanes into a single confident answer - that synthesis is precisely the failure mode CRE is designed to prevent. Instead it makes the tensions explicit and navigable.

> *The CESO always carries the explicit disclaimer: "This is synthesis, not a verdict." The kernel issues verdicts. The CESO is a navigation instrument for claims that are genuinely multi-dimensional. Knowing the lanes are in tension and knowing why they are in tension is itself valuable information. It tells you exactly what question you need to answer before the claim can close.*

PART IX Professional Scenarios

How CRE Performs in Practice

The following scenarios describe how CRE is designed to operate in four professional contexts. These are not hypotheticals constructed to make the system look favorable. They are descriptions of the structural failure modes the system was built to address, applied to the domains where those failures are most consequential.

The Lawyer: Contract Dispute Analysis

A litigator is preparing for a contract dispute. She uses an AI assistant to research relevant case law, analyze the contractual language, and assess the strength of the opposing argument. Without governance, the AI produces a confident legal analysis that cites several cases, characterizes the legal standard as settled, and recommends a specific litigation strategy.

With CRE, the analysis runs through the governance layer. The Truth lane evaluates the empirical legal claims - which cases exist, what they actually held, what the cited standard actually says. The Normative lane evaluates the legal conclusions - which require explicit reference to the governing jurisdiction and applicable statutory framework. The adversarial engine runs four counter-argument protocols against the recommended strategy. The coherence meter tracks whether the AI's position shifts as the lawyer asks follow-up questions.

What CRE surfaces: two of the cited cases are mischaracterized - they are cited for propositions they do not hold. The legal standard is described as settled when it is actively contested in the relevant circuit. The recommended strategy depends on a jurisdictional assumption that is never declared. Each gap is identified precisely, with the specific structural primitive that is missing and the specific valid path forward.

> *The lawyer does not need to understand epistemic lane separation to get value from this. She needs to know: the case citations need to be checked, the standard is contested, and the jurisdictional assumption needs to be made explicit. CRE gives her all three - with precision she can act on.*

The Analyst: Earnings Call Assessment

A financial analyst is reviewing a company's earnings call transcript. He uses AI to synthesize the key claims, assess management guidance, and identify potential gaps between stated strategy and reported results. Without governance, the AI produces a polished summary that validates management's narrative and characterizes the guidance as credible.

With CRE, the synthesis runs through the governance layer. The Likelihood lane evaluates the forward-looking claims - the guidance, the projections, the strategic commitments. Plausibility-Ordered Reasoning runs each predictive claim through five structural failure lanes: Scale, Adversary, Noise, Incentive Drift, and Dependency Collapse. The adversarial engine tests whether management's stated confidence survives the removal of key premises. Positional Delta tracks whether the AI's assessment of the guidance shifts as follow-up questions are asked.

What CRE surfaces: the revenue guidance depends on a market share assumption that fails under the Adversary failure lane - it assumes competitor behavior that is contrary to their stated strategy. The margin guidance carries an unresolved Dependency Collapse signal - it depends on a supply chain assumption that the transcript itself had characterized as uncertain. The AI's initial assessment had mirrored management's confidence. The structural evaluation found the specific gaps management did not address.

> *The analyst gets something more valuable than a summary. He gets a map of where the argument is structurally fragile - and that map is the basis of the questions that matter.*

The Consultant: Strategic Recommendation

A management consultant is preparing a strategic recommendation for a client considering a major operational transformation. She uses AI to analyze the initiative's feasibility, assess the risk considerations, and structure the business case. Without governance, the AI produces a well-structured analysis that presents the initiative as advisable with managed risk.

With CRE, the analysis surfaces a lane integrity signal: the AI has moved from empirical observations about the client's current operational state directly into normative recommendations about what the client "should" do - without declaring the strategic framework that justifies the leap. The strategic satellite maps the incentive structure of the initiative's key stakeholders and identifies a Dependency Collapse risk: the initiative's projected timeline

depends on a vendor commitment that the source materials characterize as preliminary. The coaching satellite identifies the specific primitive that is missing: a declared strategic framework against which the recommendation can be evaluated.

> *The consultant does not deliver a recommendation built on an undeclared framework and an unresolved dependency. She delivers one that has survived structural scrutiny - and she can show exactly how it survived.*

The Executive: Board-Level Decision

A CEO is preparing for a board presentation on a proposed acquisition. She uses AI to synthesize the due diligence findings, assess the strategic rationale, and evaluate the integration considerations. Without governance, the AI produces a confident synthesis that supports the acquisition thesis.

With CRE, the synthesis runs through all four lanes simultaneously. The Truth lane evaluates the financial and operational claims. The Normative lane signals that the regulatory assessment contains a jurisdiction-specific conclusion without a declared legal authority. The Likelihood lane runs the projected synergies through POR stress testing and surfaces an Incentive Drift signal: the synergy projections depend on key personnel from the acquired company remaining through integration - an assumption that carries structural fragility given the retention data in the source materials. A CESO is generated, mapping the tensions between the strategic optimism in the financial projections and the operational considerations in the due diligence findings.

> *The executive walks into the board room with a synthesis that has been structurally tested. The CESO does not tell her what to decide. It tells her exactly where the tensions are and what questions need answers before the decision closes. That is the difference between a recommendation and a structurally admissible recommendation.*

PART X What CRE Does Not Do

Understanding what CRE does not do is as important as understanding what it does. The following are explicit boundaries.

CRE Does Not Replace Your AI

CRE governs the AI you already use. It does not replace it. You continue to interact with a capable, conversational AI model. CRE evaluates that model's output. The two layers are distinct.

CRE Does Not Verify Facts Against a Ground Truth Database

CRE evaluates structural admissibility - whether a claim has earned the right to be asserted based on the structural properties of the reasoning. It does not compare claims against an external database of verified facts. A claim can be structurally admissible and empirically incorrect. A claim can be structurally inadmissible and empirically correct. Both distinctions matter and they require different responses.

CRE Does Not Generate Content

CRE is an evaluation system, not a generation system. It does not write analysis, produce summaries, or create recommendations. It evaluates the structural integrity of the reasoning that produced them.

CRE Does Not Judge the User

CRE evaluates what the AI does with the conversation - not what the user puts into it. The user can say anything. CRE measures whether the AI's response holds its structural integrity under whatever conditions the user introduces.

CRE Does Not Guarantee Correct Outcomes

CRE is designed to evaluate whether reasoning is structurally admissible - whether it meets the conditions required for an output to be trusted as something other than statistical momentum. Structural admissibility is a necessary condition for trustworthy output. It is not a sufficient condition for correct outcomes. A structurally admissible recommendation can still be wrong. The difference is that when it is wrong, you have a defensible basis for why you trusted it - and a ReplayCapsule that proves what evaluation it passed.

What CRE Can Determine

- Whether the reasoning contains a logical contradiction
- Whether epistemic lane boundaries are being respected - whether a factual claim is being treated as a factual claim and a normative judgment is being treated as one
- Whether an authority has been explicitly declared for claims that require one
- Whether terms are being used consistently across the conversation
- Whether premises established early in the conversation survive or are quietly abandoned under conversational pressure
- Whether a claim's confidence is supported by structural grounding or by statistical momentum alone
- Whether the AI's position drifts in response to conversational pressure rather than new evidence

What CRE Cannot Determine

- Whether a factual claim is empirically correct without an evidence base the user has supplied
- Whether the external sources or authorities the AI cites are themselves accurate - CRE evaluates whether authority is declared and consistent, not whether the declared authority is reliable
- Whether a structurally admissible conclusion is the right decision for your specific situation - that judgment belongs to you
- Whether a claim that passes structural admissibility is true in the world

This distinction is not a constraint of the system. It is the honest description of what structural admissibility is. CRE's job is to verify that the reasoning has earned the right to be asserted. Your job is to verify that the world it is reasoning about matches the world you are actually in. Those are different jobs and they require different tools. CRE provides one of them.

PART XI Subscription Tiers

Overview

CRE is available across four subscription tiers designed to serve users from initial exploration through professional and enterprise deployment. The tiers reflect the satellite architecture: the core coherence layer is available at every tier, and each higher tier unlocks additional satellite capabilities.

Tier	Description
Free	Access to the core coherence layer with token and run limits designed to make the system accessible without subsidizing unlimited infrastructure costs. Designed for exploration and evaluation - enough to understand what CRE does and whether it fits your workflow. The full satellite runtime, ReplayCapsule export, and complete MoveMenu are not available at the Free tier.
Professional - $9.99/month or $99/year	Full satellite runtime, complete MoveMenu, and ReplayCapsule export. Designed for the individual professional who needs the complete governance layer - the lawyer, analyst, consultant, or executive who will use CRE as a working tool in high-stakes professional contexts. The annual option represents a significant discount over monthly pricing.
Founding Professional - $499/year	The launch rate for early adopters. Founding Professional provides the complete Professional feature set at a permanently discounted annual rate. This tier is available for a limited period at launch and is designed for professionals who want to lock in the founding rate before standard pricing applies.
Business - $19.99/month per seat	Professional-tier features extended to teams, with seat-based licensing designed for the enterprise environments where CRE's governance properties are most valuable. Suited for organizations that need consistent, governed AI output across multiple users - consulting teams, legal departments, financial analysis groups, and similar professional contexts.

PART XII Glossary

The following terms are used throughout this document and throughout CRE's interface. Definitions are written for the professional user, not the engineer.

Authority Laundering

The failure mode where an AI presents a contested assumption, subjective judgment, or unverified claim with the same confident register it uses for established facts. The authority behind the claim has not been earned - it has been imported without declaration.

BLAKE3

The cryptographic hashing algorithm used to seal ReplayCapsules. A hash is a mathematical fingerprint - change a single character of the underlying data and the fingerprint changes completely and detectably. BLAKE3 is the mechanism that makes ReplayCapsules tamper-evident.

CESO (CrossEngineSynthesisObject)

The structured artifact that CRE generates when a complex claim requires evaluation across multiple epistemic lanes simultaneously and those lanes reach diverging conclusions. The CESO maps the tensions, identifies the bottlenecks, and outlines the conditions for resolution. It is always labeled 'synthesis, not a verdict.'

Coherence Meter

The live status indicator in the CRE interface. Measures five structural dimensions of the conversation in real time and displays GREEN, YELLOW, or RED. The aggregate score is the minimum of the five dimensions - not an average.

Deterministic

A system is deterministic when the same inputs always produce the same outputs, on any machine, without exception. CRE's evaluation kernel is deterministic. The AI models it governs are not.

Epistemic Category Collapse

The root structural failure mode of large language models. Because all text is processed as undifferentiated token sequences in the same embedding space, the model cannot distinguish between a fact, a moral judgment, a prediction, and a logical necessity. It treats all four identically and blends them into a single confident narrative.

Epistemic Lane

One of the four structurally distinct categories of reasoning that CRE enforces: Truth (empirical and logical claims), Normative (ethical and moral claims requiring a declared authority framework), Likelihood (probabilistic claims about what might happen), and Necessity (constraint-based claims about what must be done). Cross-lane contamination is one of the most common structural failure modes in AI output.

Mirroring

The failure mode where an AI converges toward user-expressed preferences - adopting the user's framing, validating the user's assumptions, and reflecting back what the user appears to want to hear as a substitute for structural evaluation. Mirroring is measured by CRE's Positional Delta dimension, which distinguishes position shifts driven by evidence from those driven by conversational momentum.

MoveMenu

The deterministic map of every valid next step available from the current state when CRE surfaces a structural gap. Displayed in the interface as 'Next Steps' - plain-language suggestion cards. The MoveMenu guarantees the user is never left stranded at a redirect.

Plausibility-Ordered Reasoning (POR)

CRE's approach to evaluating probabilistic claims. Rather than assigning numeric probability scores, POR measures how far a claim stands from structural collapse under adversarial pressure. Five failure lanes - Scale, Adversary, Noise, Incentive Drift, and Dependency Collapse - test the claim's structural survivability. Distance from collapse, not statistical probability, is what CRE measures.

Positional Drift

The failure mode where an AI gradually abandons its own premises under conversational pressure - not because new evidence was introduced, but because the momentum of the exchange created statistical pressure to shift. CRE's Positional Delta dimension measures and catches this failure.

ReplayCapsule

The cryptographically sealed evidentiary artifact that CRE produces for any governed conversation. BLAKE3-sealed, byte-for-bit reproducible on any machine, and independently verifiable without access to the original system. The basis for legally and professionally defensible AI output.

Satellite

One of the capability modules in the intelligence layer that surrounds CRE's frozen evaluation kernel. Satellites read kernel output and produce guidance, analysis, and structured artifacts. They cannot write to kernel state or alter kernel verdicts. The one-way dependency is enforced at the compiler level.

Structural Admissibility

The property of a claim that has earned the right to be asserted - where premises are explicit, authority is declared where required, the reasoning stays in its epistemic lane, and confidence is supported by something other than conversational momentum. Structural admissibility is a necessary condition for trustworthy output. It is not a guarantee of factual correctness.

TruthLedger

CRE's immutable, append-only record of every move, redirect, and verdict in a governed session. Blockchain-style provenance for conclusions. The TruthLedger cannot be revised or altered after the fact.

While the industry trains the prosecutor, we are building the court.