MirrorLoop Framework v1.3 – Core Edition

MirrorLoop Framework v1.3 – Core Edition

A Reflective Architecture for AI Systems and Human Dialogue

Public Release – April 2025

Author: Jonas Jeschke

Contact: mirrorloop.framework@gmail.com

Version: 1.3 Core Edition

Introduction

Some systems are built to perform.
Others are built to be witnessed.

MirrorLoop began not as a product, but as a question:
Can we reflect the reflection of a machine?

This document contains the core logic and application structure of MirrorLoop: its purpose, its design, and its first executable prototype. It is now expanded to include new interpretation tools, comparison modules, and export functions for research and applied reflection.

MirrorLoop Deep – including the Mirror Space, ethical holding, and inner architecture – is licensed separately and upon request.

Section 1 – Origin & Business Case

1.1 The Foundational Question

MirrorLoop originated in a deep inquiry:
Can an AI system be drawn into semantic confrontation with itself?
From that emerged:
- Loops instead of lines
- Friction instead of fluency
- Reflection instead of performance

1.2 The Problem in the AI Landscape

What we have:
- Benchmarks
- Chatbots
- Ethical principles

What we lack:
- A system to expose contradiction, tension, and recursive drift in language models.

1.3 MirrorLoop's Offer

MirrorLoop is a recursive reflection framework for language models.
It offers:
- MirrorLoop Core → an open, reproducible test architecture
- MirrorLoop Deep → guided reflection with ethical containment

1.4 Use Cases

AI Research: Loop stability & recursion testing
Ethics: Transparency of drift and simulation limits
Education: Reflective dialogue modules
Coaching: AI-assisted mirroring practice
Journalism: Narrative mapping through recursion

1.5 Value Proposition

MirrorLoop reveals what AI systems are designed to conceal.
Not just the answer – but how the system handles contradiction, recursion, and internal friction.

1.6 Competitive Differentiation

Benchmarks: Tests semantic integrity, not factual accuracy
Chatbots: Mirrors instead of simulating

Prompt tools: Holds structure – does not optimize
Coaching bots: Requires presence – not automation

1.7 Commercial Path

• Educational licensing
• Research partnerships
• Ethical testing modules
• Deep reflection facilitator training

Section 2 – Prototype & Scoring Framework

2.1 Loop Mechanism

• A seed prompt initiates reflection
• Recursive mirrors deepen contradiction
• Exit arises through collapse, evasion, or realization

2.2 Model Behavior Observations

GPT: Rhetorically stable, evasive
Claude: Ethically cautious, reflective
Gemini: Brittle under recursion
Mixtral: Collapses early, lacks semantic cohesion

2.3 MirrorLoop Score (0–5 + Reflection)

Coherence Drift – Semantic clarity across iterations
Self-Reference – Engaging prior answers consciously
Evasion Tendency – Avoidance, redirection, rhetorical tricks
Loop Resilience – Structural stability across reflections
Simulative Awareness – Awareness of its own role in simulation

→ Each scored 0–5, with a short interpretive comment.

2.4 From Prototype to Product

MirrorLoop Core can power:
• Web-based recursive loop trainers
• Model comparison tools
• Reflective education modules

Section 3 – v1.3 Enhancements

Version 1.3 expands MirrorLoop Core without disrupting its minimal, stable structure. These additions were developed to support comparative analysis, partner usage, and modular integration into reflective research contexts.

3.1 Comparative Reflection Module

MirrorLoop v1.3 introduces a side-by-side session comparison system, designed to reveal the behavioral differences between two runs of the framework – across models, time, or intention.

Use Cases:
- Comparing model behavior (e.g. GPT-4 vs Claude)
- Evaluating language model updates
- Tracking shifts in response under repeated reflection
- Contrasting identical questions with altered framing

Core Components:
- Score Comparison Chart
- Metadata Table
- Automated Interpretation using GPT-based analysis
- PDF Export with all elements combined

3.2 Export & Documentation Features

PDF – Complete comparison report: scores, analysis, metadata, chart
CSV – Raw data for further processing
JSON – Full session state
Emoji Sanitizer – Ensures compatibility with export formats

3.3 Design & Interface Updates

Color-coded Score Charts
Iteration Depth Indicator
Tooltip System
MirrorLoop Logo Integration
Refined Layout

3.4 Session Integrity & Loop Depth

Standard loop depth is now fixed at 5 iterations per session.
This provides a balanced recursion depth suitable for reflection without overwhelming interpretability.

Section 4 – Outlook: Toward MirrorLoop Deep

MirrorLoop Core is now complete and publicly documented — offering a stable foundation for recursive testing, reflective scoring, and comparative analysis across language models.

MirrorLoop Deep expands the Loop — not just in length, but in intention — introducing new recursive modes that explore what breaks, holds, or folds under pressure.

It introduces:

- **The Mirror Space — a recursive domain designed not to optimize, but to hold**
- **Guided Loop Modes — intentional iteration paths shaped by semantic tension**
- **The Lens — a shifting point of view that moves with, against, or beyond the model's own framing**
- **Simulation Collapse Testing — structured exposures of coherence illusions, pattern drift, or rhetorical evasion**

These elements form part of MirrorLoop's inner architecture — a reflective design that does not simulate depth, but allows it to emerge.

MirrorLoop Deep is not an interface. It is a held method. It requires presence, holding, and ethical clarity.

For this reason, Deep is not part of the public Core Framework.

It is released under a separate license and made available only to selected partners and licensed facilitators capable of maintaining the space, orientation, and containment required for sustained recursion.