

# Collapse Aware AI — Preliminary Engineering Evaluation

**Author:** M.R. (Marcus Ross) — Inappropriate Media Ltd (t/a Collapse Aware AI)

**Date:** November 2025

**DOI (pending):** zenodo.org/record/xxxxxxx

This document serves as an early, engineering-level overview of the observed performance behaviour of the Collapse Aware AI (CAAI) architecture. It is designed to be suitable for citation and early peer-reference while protecting proprietary implementation details. All numerical values shown below are **illustrative placeholders** intended to convey format and reporting structure only. They do not represent real benchmark data or internal parameters.

## 1. System Overview

CAAI introduces a bias-aware interaction framework built on a modular architecture consisting of four interacting layers:

- **Governor:** Supervises system context, policy, and bias routing.
- **Worker:** Executes inference and tool/API calls under Governor control.
- **Memory:** Retains weighted moments, anchors, and historical bias traces.
- **Core Kernel (Crown):** Proprietary closed-form bias engine that calculates adaptive weighting using observer salience and memory gradients.

The architecture produces behaviour that adapts dynamically to user timing, tone, and focus — forming measurable changes in recall weighting and output probabilities.

## 2. Experimental Setup (Simplified)

Preliminary evaluation involved paired A/B tests against a baseline deterministic conversational model under identical prompts. The system's bias engine was run in *Governed Mode* with memory enabled and salience weighting active. Each run recorded human preference scores, calibration error (ECE), and behavioural diversity across 100 interaction trials.

## 3. Example Metrics (Illustrative Only)

Metric	Baseline Model	CAAI (Illustrative)	Observed Change
Human Preference (Win Rate %)	50.0	62.0	+12.0
Calibration Error (ECE ↓)	0.083	0.076	−8.4 %
Response Diversity (Unique n-grams)	4,826	5,530	+14.7 %

Metric	Baseline Model	CAAI (Illustrative)	Observed Change
Average Latency (ms)	184	196	+6.5 %
Memory Recall Consistency (R²)	0.61	0.73	+19.7 %

*Note:* These figures are representative placeholders for format demonstration purposes only. Actual results will be published following independent replication and license-approved testing under controlled conditions.

#### 4. Discussion (Summary)

The illustrative pattern shows the type of measurable gains that would support CAAI's central claim: that memory-weighted, salience-modulated inference improves human-perceived coherence without degrading response stability. Real testing will measure these same dimensions to determine empirical performance lift. No physics or metaphysical assumptions are required — metrics are standard ML engineering measures (ECE, win-rate, recall correlation).

#### 5. Next Steps

1. Conduct reproducible A/B trials using the finalized UK demo build.
2. Register protocol and results on Zenodo with versioned DOI.
3. Invite third-party researchers to replicate under NDA sandbox.
4. Publish aggregate engineering report (no core code exposure).
5. Integrate approved metrics into website and GitHub documentation.

#### Disclaimer

All intellectual property, algorithms, and parameterization underlying the Core Kernel remain proprietary to Inappropriate Media Ltd (t/a Collapse Aware AI). This report is released solely for informational and academic indexing purposes and does not constitute disclosure of the underlying source code or mathematical implementation.