

SCIENTIFIC DISCOVERY FRAMEWORK

Natural Language Chain-of-Thought (CoT) Engine Protocol

Citation Context:	Landry, M.-S. S. (2026). The Scientific Method as Natural Language Chain-of-Thought. Zenodo. DOI: 10.5281/zenodo.20953561
AI Disclosure:	Gemini 1.5 Flash • Structural Determinism Execution State
Keywords:	Chain-of-Thought, State Machine, Post-Experimental Data, Verification Gates, Automated Discovery, AI Reasoning Workflow

1. Strategic Architectural Evaluation

The 10-step natural language Chain-of-Thought (CoT) framework operationalizes the traditional scientific method into an automated, discrete machine workflow. By bifurcating the discovery pipeline into two isolated execution blocks—**Cycle A: Pre-Experimental Proposal Generation** and **Cycle B: Post-Experimental Data Review**—the framework transforms inductive reasoning into a structured state machine.

The Critical Blind Spot: Cumulative Optimization Bias. While Large Language Models natively generate hypotheses and experimental parameters in a "one-shot" configuration, the structure remains vulnerable to error propagation. If a literature synthesis node imports a single unverified claim or hallucination, every subsequent dependent node (*Hypothesis* → *Experimental Design*) inherits and amplifies that deviation. To guarantee objective truth, the framework enforces immutable validation gates between specific states.

2. Execution Loop Protocol

To eliminate error cascade, the system runs as a strict state machine where the output of Step *N* must pass logical and external verification before unlocking Step *N + 1*.

CYCLE A: PRE-EXPERIMENTAL PROPOSAL GENERATION

1. **Observe Anomaly:** Detect and declare an unexpected variance within a specific dataset or physical system (e.g., non-linear stress-strain behavior in newly synthesized bio-composites).
2. **Define Intelligence Requirement:** Convert the empirical anomaly into a highly specific, granular, and mathematically answerable research query.
3. **Research (Verification Gate):** **[HARD GATE]** Execute targeted, external database queries. Extract verbatim facts and active DOIs. If literature is insufficient or unverified, flag **[ω SCARCE]** and halt execution.
4. **Hypothesize:** Synthesize a testable, strictly falsifiable hypothesis derived directly from step 3 data invariants (**H_1** absolute grounding).
5. **Design Experiment:** Author standard operating procedures, explicitly defining control groups, boundary variables, and standardized CSV schemas for human data collection.

CYCLE B: POST-EXPERIMENTAL DATA REVIEW & ITERATION

1. **Collect Data:** Human agents execute the physical experiment and input raw empirical measurements or unedited tables straight into the execution layer.
2. **Analyze & Interpret:** Process mathematical transformations, evaluate statistical significance, and identify anomalies via deterministic sandboxed computation.
3. **Decision Loop:** Evaluate analytical outputs against the initial step 4 hypothesis. If empirical results show a logical conflict, instantly reject the hypothesis and return execution to Cycle A, Step 4.
4. **Conclusion:** Construct a peer-review-ready technical synthesis bound strictly to the observed data vectors. Deductive extrapolation is prohibited.
5. **Report & Replicate:** Output an immutable manuscript structured to Zenodo or specific journal metadata criteria, appended with the exact code required to reproduce the calculation state.

Operational Directive: This natural language protocol functions perfectly as a continuous "one-shot" generation script or as a manual, message-by-message prompt chain. The defining constraint is that every structural block must be read-verified and verified against real-world grounding mechanisms before triggering subsequent transitions.