TECHNICAL REPORT 001 / RESEARCH PROTOTYPE 0.2

APORIA

An Observatory for Structured AI-Supported Reasoning

Fatima Aguilar
Decision Systems Lab
Developed in collaboration with Professor Viktoria Dalko

July 2026

Executive Summary

Aporia is a conversational research prototype designed to improve the visibility and discipline of human reasoning. It is the user-facing environment powered by the Epistemic Dialectic Engine, an emerging research program focused on how artificial intelligence can challenge assumptions, preserve uncertainty, support reflection, and complement human judgment.

The prototype operationalizes four inquiry protocols - Dialectic, Socratic, Reflection, and Transfer - within a single conversational interface. Users present a claim, decision, or dilemma. Aporia then introduces targeted cognitive moves rather than immediately supplying a conclusion. Alongside the dialogue, the interface displays exploratory indicators associated with metacognition, conceptual understanding, reflection, transfer, and self-regulation.

Current status. Aporia is a functional web application and a research instrument in formation. Its dialogue mechanics, interface, demonstration engine, model-integration path, session export, public deployment, and source repository are implemented. Its cognitive indicators are not yet validated measures and must not be interpreted as educational, psychological, or clinical scores.

> The central design premise is simple: **Aporia uses AI to help people examine how answers are formed rather than merely producing them.**

DOCUMENT SCOPE

This report documents the conceptual basis, implemented architecture, interaction model, technical stack, reliability controls, governance boundaries, and present limitations of Aporia version 0.2.

CONTENTS

01 Research Context and Design Problem
02 System Objectives and Principles
03 Interaction and Cognitive Model
04 Technical Architecture
05 Implementation and Reliability
06 Security, Privacy, and Governance
07 Present Limitations
08 Conclusion and References

1. Research Context and Design Problem

Recent work on AI in advanced education increasingly describes artificial intelligence as a supplement, tutor, research assistant, collaborator, or reasoning partner. That progression shifts the design problem away from simple task automation. If an AI system participates in advanced learning or professional judgment, the relevant question becomes whether it improves the quality of reasoning rather than only the speed of output.

Traditional conversational systems generally optimize for helpful, fluent answers. In conceptually difficult contexts, however, immediate agreement can conceal assumptions, reduce productive friction, and encourage users to mistake confidence for justification. Human scholarly and professional reasoning often develops through challenge, counterexample, clarification, reflection, and transfer to unfamiliar situations.

Aporia treats productive uncertainty as a feature. The name refers to a state of puzzlement in which an apparent certainty becomes open to examination. The Observatory theme translates that epistemic stance into an interaction metaphor: users do not enter a chat to receive an oracle's judgment; they place a claim under observation and inspect the signals around it.

DESIGN PROBLEM

How can a conversational AI system support structured inquiry while preserving human authority, making uncertainty visible, and generating evidence that may later support rigorous evaluation of cognitive development?

RESEARCH LINEAGE

Aporia builds on two connected literature streams developed in the broader project: (1) research on how AI is used in doctoral education and advanced knowledge work, and (2) educational measurement literature concerning metacognition, conceptual understanding, transfer, reflection, self-regulation, and learning gain. The prototype converts that synthesis into a testable interaction environment.

2. System Objectives and Principles

Aporia is organized around five objectives:

- **Improve reasoning visibility.** Prompt users to distinguish observations, evidence, inferences, assumptions, values, and uncertainty.
- **Introduce productive challenge.** Surface a strong competing interpretation without becoming adversarial.
- **Preserve human judgment.** Treat the user, instructor, or decision-maker as the final authority.
- **Support research instrumentation.** Produce structured dialogue artifacts and exploratory signals that can later be compared with validated measures.
- **Remain usable during early research.** Provide a deterministic demonstration mode alongside an optional live-model mode.

OPERATING PRINCIPLES

Principle	Operational behavior
Challenge without combat	Questions target the claim, not the person.
Evidence before synthesis	Aporia asks what would falsify or revise the current position.
Uncertainty remains visible	The interface preserves an active tension and next cognitive move.
One demanding move at a time	Responses emphasize a focused question rather than a wall of advice.
Measurement humility	Indicators are explicitly labeled exploratory and non-validated.
Graceful degradation	If a live model fails, the deterministic protocol remains available.

The explicit research disclaimer is not decorative. It is part of the system's governance model and prevents interface indicators from being mistaken for validated assessment results.

3. Interaction and Cognitive Model

Aporia exposes four observation lenses. Each lens changes the cognitive move applied to the same underlying user claim.

Lens	Primary function	Example move
Dialectic	Contrast competing interpretations	What fact would make you reverse your position?
Socratic	Audit definitions and premises	What assumption can your conclusion not survive without?
Reflection	Inspect changes in thinking	Did the facts, interpretation, or prioritized value change?
Transfer	Apply the principle elsewhere	Where would this rule fail in a new context?

EXPLORATORY SIGNAL MODEL

The interface tracks five constructs selected from the project's educational assessment review. The current values are generated heuristically in demonstration mode and requested as structured fields in live-model mode. They function as interface signals that organize observation; they do not establish that a construct has been measured.

- **Metacognition:** evidence that the user is planning, monitoring, or evaluating their own thinking.
- **Conceptual understanding:** evidence of relationships, distinctions, and explanatory depth rather than recall alone.
- **Reflection:** evidence that experience, emotion, interpretation, or values are being reconsidered.
- **Transfer:** evidence that an idea can be applied, bounded, or revised in a novel context.
- **Self-regulation:** evidence of goal-setting, strategy adjustment, confidence calibration, or deliberate monitoring.

4. Technical Architecture

Aporia is implemented as a full-stack TypeScript application. The architecture supports a deterministic research demonstration and an optional external model without changing the user-facing protocol.

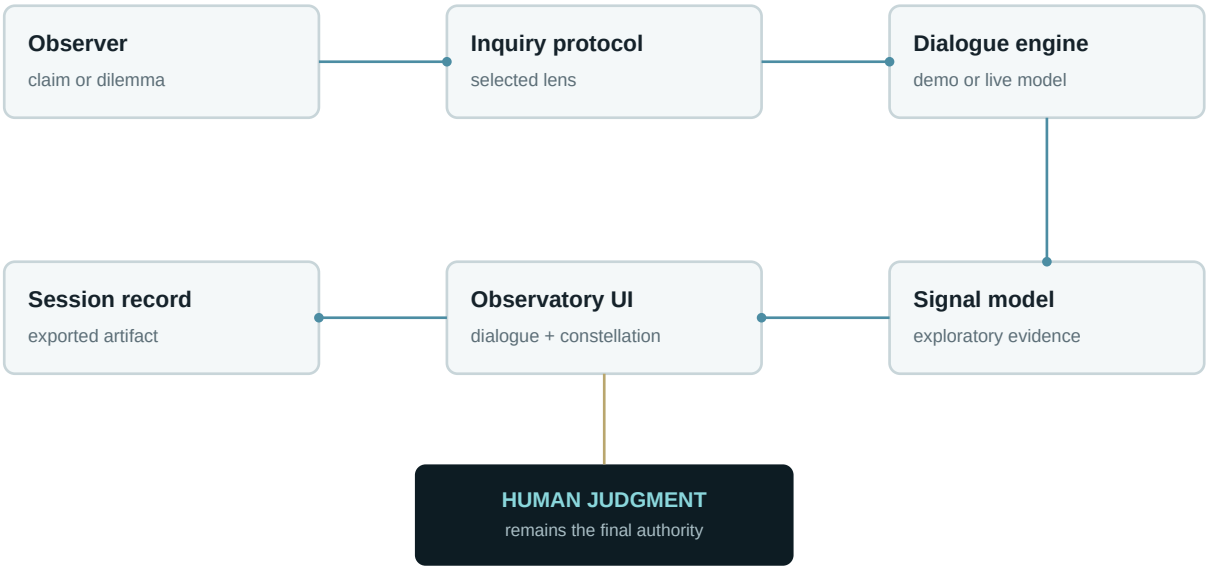

Figure 1. Current request, dialogue, signal, and export flow.

APPLICATION LAYERS

Layer	Implementation	Responsibility
Interface	React 19 + TypeScript	Conversation, modes, live constellation, settings, export
Application framework	Next-compatible vinext runtime	Routing, client/server boundaries, production build
Dialogue protocol	Client heuristic + system prompt	Mode-specific challenge and cognitive moves
Model gateway	Server API route	OpenAI Responses API request and structured JSON return
Hosting	Cloudflare-compatible Sites deployment	Public production delivery
Source control	Git + GitHub	Public implementation history and collaboration

The server route uses the OpenAI Responses API when a key is available. The model is instructed to return a constrained JSON object containing dialogue text, a named cognitive move, five exploratory signals, an unresolved tension, and a next move. The client validates this structure before rendering it.

5. Implementation and Reliability

DEMONSTRATION MODE

The built-in engine uses lightweight textual features to select a mode-specific prompt. It detects markers associated with certainty, evidence, and value judgments, then generates a focused challenge and adjusts interface signals. The mode is intentionally deterministic enough for interface evaluation and research demonstrations without external credentials.

LIVE-MODEL MODE

A user may provide a session-only API key or a deployment administrator may configure a server environment variable. The server sends the selected inquiry mode and conversation history to the model gateway. The API key is not stored by the application.

FAILURE CONTAINMENT

- The client validates the complete evidence object before accepting a live-model response.
- Invalid JSON, missing fields, network errors, or unavailable model connections trigger a transparent fallback to demonstration mode.
- The interface remains usable without credentials.
- A production defect in the initial Observatory release was traced to an effect cleanup contract violation, corrected, and verified through direct browser interaction.
- The repaired production path was tested for typed submission, example prompts, mode switching, reset behavior, and response rendering without browser errors.

OBSERVABILITY AND EXPORT

Each conversation maintains a visible current probe, five signal values, an active tension, and a next cognitive move. Users may export the dialogue and current indicators as a Markdown Observatory Record. This provides a portable research artifact while avoiding premature claims of formal assessment.

CURRENT VALIDATION EVIDENCE

Check	Status
Production build	Pass
Public deployment	Pass
Typed conversation submission	Pass
Example prompt workflow	Pass
Mode switching and reset	Pass
Client error inspection after response	No errors observed
Scientific validity of cognitive scores	Not yet established

6. Security, Privacy, and Governance

The public prototype has deliberately limited persistence. Conversations and keys are not written to an application database. This reduces exposure during early research, but it also means the system does not yet support longitudinal analysis or instructor review.

KEY HANDLING

A key entered through the connection panel is held in browser memory for the current page session and transmitted to the application's own server route for the requested model call. It is not intentionally persisted. The deployment also supports a server-side environment variable.

RESEARCH DATA GOVERNANCE

The current public prototype does not collect participant records. No consent workflow, retention policy, de-identification process, withdrawal mechanism, or research-data access model is implemented.

ASSESSMENT GOVERNANCE

The current interface includes a prominent disclaimer because numeric presentation can create false authority. The indicators are labeled exploratory and are not presented as valid for high-stakes use.

THREAT AND MISUSE CONSIDERATIONS

- Prompt injection or adversarial text may attempt to redirect the model from its inquiry protocol.
- A model may produce plausible but unsupported interpretations of a participant's cognition.
- Users may over-trust polished feedback or treat exploratory signals as diagnosis.
- Public endpoints require rate limiting, abuse controls, and cost monitoring before a shared server key is enabled.
- The application has no approved data policy for sensitive educational or organizational scenarios.

7. Present Limitations

CURRENT LIMITATIONS

- The demonstration engine is heuristic and cannot sustain deep domain reasoning.
- The live-model protocol depends on a third-party generative model and inherits its variability.
- Signal values are not derived from validated rubrics or empirical calibration.
- Conversation state is device-session based and is not stored for longitudinal research.
- No instructor dashboard, annotation workflow, or rater interface is implemented.
- The system has not yet been evaluated with the Meridian Simulation or a participant cohort.
- Accessibility and cross-device validation remain limited.

8. Conclusion

Aporia demonstrates how a research concept can be translated into a functioning human-AI interaction environment. Its contribution is not that it answers questions more quickly. Its contribution is the design of a conversational space in which assumptions, evidence, alternatives, uncertainty, and changes in thinking remain visible.

The current prototype establishes the technical and experiential foundation for the Epistemic Dialectic Engine research program. It provides four structured inquiry protocols, a dual demonstration/live-model architecture, visible exploratory signals, portable session records, resilient failure handling, public deployment, and open source code.

The scientific contribution remains unvalidated. No evidence currently establishes that the dialogue protocol produces measurable improvements in reasoning or that its conversation signals correspond to established constructs. Aporia is therefore documented as a functional research prototype rather than a validated assessment instrument.

> The objective is not certainty on demand. It is better-grounded judgment.

PROJECT LINKS

Live prototype: https://epistemic-dialectic-engine.fsaguilar16.chatgpt.site
Source repository: https://github.com/fa366193/aporia

REFERENCES AND TECHNICAL SOURCES

Aguilar, F. (2026). *Aporia: Epistemic Dialectic Engine research prototype* [Computer software]. GitHub. https://github.com/fa366193/aporia

OpenAI. (2026). *Responses API*. OpenAI Developer Documentation. https://developers.openai.com/api/docs/guides/responses

OpenAI. (2026). *Model guidance*. OpenAI Developer Documentation. https://developers.openai.com/api/docs/guides/latest-model

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.

Rodgers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers College Record, 104*(4), 842-866.

Dell'Acqua, F., et al. (2025). *The cybernetic teammate: A field experiment on generative AI reshaping teamwork and expertise*. National Bureau of Economic Research.

CITATION

Aguilar, F. (2026). *Aporia: An observatory for structured AI-supported reasoning* (Decision Systems Lab Technical Report No. 001).