Geometric Instrument for Measuring Interrogative Entropy in Language Systems v.03

Bruce Tisler
ORCID: 0009-0009-6344-5334

Version 3 — December 2025

DOI (Version 1): 10.5281/zenodo.17811309
DOI (Version 2): 10.5281/zenodo.17841747
DOI (Version 3): 10.5281/zenodo.17931253

Author's Clarification on Prior Versions

Note: This clarification addresses language used in Versions 1 and 2 that may be ambiguous with respect to prediction and behavioral influence. Version 3 provides corrected framing.

In earlier versions of this work (v1 and v2), certain phrasings were ambiguous with respect to prediction, elicitation, and behavioral influence. That wording reflected an exploratory phase of the project and was not sufficiently precise about the limits of the instrument. As a result, some statements can be reasonably read as implying predictive or causal power that the framework does not claim and does not support intrinsically.

Below, I identify specific statements and clarify what was intended by them, without asserting that the wording was correct or should be retained.

Version 1 — Statements and Clarifications

Statement: "Interrogatives shape uncertainty, stability, and epistemic behavior."

Clarification: This was meant to indicate that different interrogative forms occupy different regions of a structural space, not that interrogatives actively mold or determine model behavior. "Shape" was used informally to describe structural configuration, not causal influence.

Statement: "Operator signatures... triggered by spatial inquiry (WHERE)... triggered by causal inquiry (WHY)."

Clarification: "Triggered" was intended as a descriptive label for recurring empirical associations observed in specific experimental runs, not as a claim that particular interrogatives reliably cause those effects across models or contexts.

Statement: "This regime is designed to elicit characteristic operator signatures."

Clarification: "Elicit" was meant to describe that certain configurations were selected for observation, not that the instrument or the interrogatives compel a response pattern. No claim of controllability or reproducibility was intended.

Statement: "These interrogatives consistently elicited 3–7 hedges per answer."

Clarification: This was a reporting of a bounded observation within a specific dataset, not a generalizable behavioral rule. The wording does not adequately distinguish observation from expectation.

Statement: "The system can identify high-risk inquiry geometries in advance."

Clarification: This was meant to suggest that structural complexity can be measured prior to generation, not that downstream instability or risk can be predicted. The phrase "in advance" refers only to temporal ordering of measurement, not foresight.

Version 2 — Statements and Clarifications

Statement: "We validate the instrument's predictive utility through large-scale empirical analysis."

Clarification: This phrasing attempted to convey that statistical correlations were observed post hoc, not that the instrument possesses intrinsic predictive power. The term "predictive utility" is misleading in this context.

Statement: "Prompt interrogative structure predicts systematic differences in assistant response affect drift."

Clarification: "Predicts" was used in a statistical regression sense, not as a claim of operational forecasting or reliability. The statement does not sufficiently separate correlation from interpretation.

Statement: "High-entropy prompts elicit approximately 25% more affect drift."

Clarification: This was meant to summarize a measured association within a constrained experimental setup, not to imply that entropy can be used to induce or control drift. The verb "elicit" overstates the relationship.

Statement: "H_Ω predicts H-drift, but we have not established causation."

Clarification: This sentence attempted to acknowledge the correlation/causation distinction, but still retains predictive framing that suggests reliability. The intended meaning was simply that both quantities co-vary under certain conditions.

Statement: "Understanding the structural properties of inputs that shape model behavior."

Clarification: "Shape" was intended metaphorically to refer to input geometry as a descriptive coordinate system, not as an agent of behavioral modification.

Closing Note

Version 3 resolves these ambiguities by restricting all claims to structural measurement, explicitly disavowing predictive power, behavioral control, or intervention capability. The clarifications above are provided to prevent misinterpretation of earlier versions, not to retroactively defend their phrasing.

Abstract

This work presents version 3 of the Holistic Data Transformation (HDT^2) framework, a structural approach to measuring interrogative configuration in language model inputs. The framework introduces interrogative entropy (H_i) as a deterministic, pre-generation measure derived solely from question structure. H_i quantifies the distributional spread of interrogative presence across a fixed interrogative geometry without reference to semantic content, model internals, or generated output.

Version 3 clarifies the limits, assumptions, and intended use of the instrument. The framework does not claim predictive power, semantic interpretation, or behavioral control. Instead, it provides a stable structural coordinate system for empirical investigation of how interrogative configurations relate—when they do—to downstream model behavior. Open questions and unresolved dependencies are explicitly retained as part of the model.

1 Introduction

Large language models are routinely evaluated based on output quality, correctness, or alignment, while the structural properties of inputs are often treated implicitly or informally. Questions, in particular, are assumed to be self-evident objects rather than structured configurations with measurable properties.

The HDT^2 framework approaches questions as structured inputs whose interrogative composition can be measured prior to generation. The goal is not to assess meaning or intent, but to characterize the structural constraints imposed by interrogative presence itself.

This work focuses on the development and clarification of interrogative entropy (H_i), a metric derived from interrogative distribution alone. Version 3 consolidates lessons from earlier iterations, clarifies scope boundaries, and formalizes open research surfaces without extending claims beyond what has been demonstrated.

2 Conceptual Framework

The framework is built around a fixed interrogative structure consisting of six interrogatives: Who, What, When, Where, Why, and How (WWWWHW). These interrogatives are treated as structural operators rather than semantic categories.

A cube is used as a geometric representation of this structure. The cube is not a model of reasoning, cognition, or intent. It is a static structure that defines a bounded space within which interrogative presence can be distributed.

The interrogatives themselves are the only measured elements. Their placement on the cube is arbitrary with respect to meaning; only their distribution matters. The cube provides a consistent coordinate system that allows interrogative configurations to be compared without semantic interpretation.

The cube functions as a mathematical coordinate system for interrogative structure. Like Cartesian coordinates in spatial measurement, it provides fixed reference points that enable consistent quantification without claiming to represent the internal mechanics of what is being measured. The geometry is not derived from LLM behavior because it is not a model of LLM behavior—it is a measurement framework applied to inputs before they reach the model. Whether this coordinate system yields useful correlations with downstream model behavior is an empirical question, answered through observation rather than assumed by design.

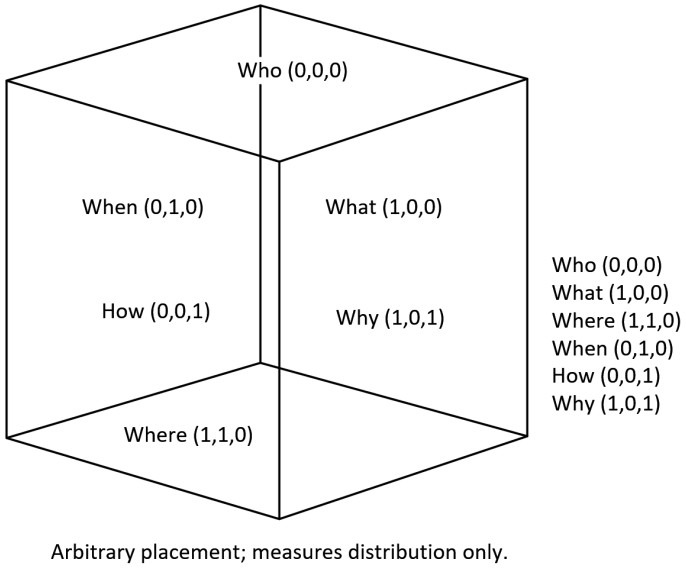

Figure 1: The WWWWHW cube geometry. Interrogatives are assigned to vertices with coordinate labels. Placement is arbitrary; the instrument measures distributional spread only. Coordinate axes carry no semantic meaning.

3 Scope and Boundary Conditions

The framework is explicitly constrained in what it is allowed to observe and what it refuses to infer.

The system is allowed to see only the structural presence of interrogatives in an input. It does not analyze semantics, intent, difficulty, correctness, or usefulness. It does not inspect model weights, attention, or internal states. It does not evaluate generated text.

The system explicitly refuses to see meaning, motivation, or user psychology. Any interpretation of intent or epistemic state begins outside the instrument and cannot be attributed to H_i.

Responsibility for interpretation therefore begins after measurement. H_i produces a structural description. Any conclusions drawn from that description are external to the framework and must be justified independently.

4 Detection of Interrogative Presence

Interrogative presence is detected as a structural property of the question as expressed in language. An interrogative is considered present if it appears as a structural element in the question as posed. This determination is binary by design.

Detection does not depend on interpreting meaning, intent, or what would constitute an answer. The system does not assess what the question is asking, nor does it evaluate necessity, relevance, or importance. It records only whether an interrogative form participates in the linguistic structure of the input.

Interrogatives are not weighted, ranked, or inferred. Presence is not probabilistic, graded, or learned. This constraint is intentional: the instrument captures structural participation only, not semantic emphasis or functional role.

Aggregation occurs solely at the level of distributional spread across the fixed interrogative structure. H_i reflects how interrogative presence is distributed, not why those interrogatives appear or what they signify.

Examples of Interrogative Detection:

- "Who won the game?" $\rightarrow$ Present: Who. Vector: $\mathbf{q} = [1, 0, 0, 0, 0, 0]$; $H_i = 0$ (single interrogative, minimum entropy)
- "How and why does it work?" $\rightarrow$ Present: How, Why. Vector: $\mathbf{q} = [0, 0, 0, 0, 1, 1]$; $H_i = 1$ (two interrogatives, uniform distribution)
- "Explain the mechanism." $\rightarrow$ Present: None. Vector: $\mathbf{q} = [0, 0, 0, 0, 0, 0]$; H_i is undefined (no explicit interrogative presence)

This example illustrates a boundary condition of the instrument. Imperative or declarative forms that function as questions may be treated as interrogatives by human readers or downstream systems, but such interpretations require inference and therefore fall outside the scope of this measurement.

5 What H_i Is Not

Interrogative entropy (H_i) is not a measure of question quality, depth, or correctness. It does not assess ambiguity, uncertainty, or difficulty. A trivial and a profound question may yield identical H_i values if their interrogative structures are equivalent.

H_i is not a semantic metric and does not encode meaning or intent. It is also not a predictive oracle. It does not forecast model behavior, hallucination, or failure, nor does it guarantee stability or instability at any value.

The metric does not prescribe interventions, thresholds, or guardrails. Any correlation between H_i and downstream behavior is empirical and contingent, not intrinsic to the measure.

6 Empirical Use and Constraints

H_i is intended as a pre-generation structural diagnostic. It is computed deterministically from the question alone and remains independent of model internals, training data, or output content.

In empirical studies, H_i may be logged alongside downstream behaviors such as response variance, hedging frequency, or inconsistency across repeated runs. Any relationships observed between H_i and such behaviors must be demonstrated per model, per task, and per dataset.

H_i does not define thresholds at which behavior must change, nor does it imply that higher values are inherently problematic. Empirically observed breakpoints, when present, are treated as properties of specific datasets and models rather than intrinsic features of the metric.

The instrument functions as a structural coordinate system, not a control mechanism. Any use beyond measurement constitutes a separate design decision.

7 Open Questions and Unresolved Dependencies

Several aspects of the framework remain intentionally unresolved.

Alternative methods of detecting interrogative presence may produce different distributions, and the sensitivity of H_i to such choices remains an open question. Likewise, relationships between interrogative structure and model behavior are necessarily model-specific and cannot be assumed to generalize.

The role of aggregation beyond distributional spread—such as temporal ordering or interaction effects—has not been resolved. Whether such extensions add signal or merely complexity is unknown.

These unresolved dependencies define the current limits of the framework. The system is intentionally constrained to structural measurement, and questions about sensitivity, extension, or downstream interpretation remain open areas of investigation. This version formalizes that state explicitly, treating these boundaries as part of the model rather than gaps to be prematurely closed.

8 Experimental Environment and Execution Context

All experiments described were conducted in a small, self-hosted research environment running on a Windows 10 workstation equipped with two NVIDIA RTX 3060 GPUs. Local inference was performed using consumer-accessible tooling.

The framework itself is hardware-agnostic, though the experiments reported here reflect the constraints and affordances of the execution environment described. A larger lab or additional compute would expand the experimental surface but would not alter the structural definitions of the instrument.

This context is specified to support reproducibility and calibration rather than to claim optimal performance.

9 Directions for Further Investigation

Future work includes systematic comparison across model architectures, exploration of alternative interrogative detection strategies, and examination of where structural measurement ceases to provide signal.

Equally important is identifying where the instrument is silent. Understanding the limits of applicability is treated as a necessary component of the framework's maturation.

10 Acknowledgment of Feedback and Ongoing Inquiry

This investigation has benefited from critical feedback that challenged assumptions, boundaries, and articulation. Such feedback has helped clarify scope without altering the core premise of the framework.

The theory continues to evolve through testing and questioning. Each iteration surfaces new questions rather than final answers. This version prioritizes coherence over closure and emphasizes ongoing inquiry as essential to the framework's advancement.

Appendix A — Formal Definition of Interrogative Entropy (H_i)

This appendix provides the formal definition of interrogative entropy (H_i) used throughout this work. The definitions below specify the measurement procedure only. They do not constitute a model of language model behavior, cognition, or internal computation.

A.1 Interrogative Vector Representation

Let a question be represented by a six-dimensional interrogative vector:

$$\mathbf{q} = (q_{\text{who}}, q_{\text{what}}, q_{\text{when}}, q_{\text{where}}, q_{\text{why}}, q_{\text{how}}) \tag{1}$$

where each component is defined as:

$$q_i = \begin{cases} 1 & \text{if interrogative } i \text{ is structurally present} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The six interrogatives are indexed as: Who, What, When, Where, Why, How. Their ordering on the cube is arbitrary; only their distributional pattern is measured.

No weighting or semantic interpretation is applied.

A.2 Normalization

The vector is normalized to form a probability distribution over interrogatives:

$$p_i = \frac{q_i}{\sum_j q_j} \tag{3}$$

This normalization ensures that H_i reflects distributional spread, not magnitude.

A.3 Interrogative Entropy

Interrogative entropy is computed using Shannon entropy:

$$H_i = -\sum_i p_i \log_2 p_i \tag{4}$$

H_i is defined only when at least one interrogative is present. If $\sum_j q_j = 0$ (no interrogatives detected), H_i is undefined. For logging or storage purposes, such cases may be recorded using a sentinel value (e.g., `NaN`) and must not be conflated with valid cases where $H_i = 0$.

Range: $H_i \in [0, \log_2(6)]$ where:

- Minimum: $H_i = 0$ when a single interrogative is present (focused question)
- Maximum: $H_i \approx 2.585$ when all six interrogatives are equally present (uniform distribution)

A.4 Interpretation Constraints

H_i measures how interrogative presence is distributed across a fixed structural geometry. It does not encode meaning, intent, uncertainty, difficulty, or expected model behavior.

Any empirical relationship between H_i and downstream language model behavior is contingent, model-specific, and external to this definition.

A.5 Relationship to Prior Versions

This formulation is consistent with the definitions introduced in Version 1 (DOI: 10.5281/zenodo.17811309) and applied empirically in Version 2 (DOI: 10.5281/zenodo.17841747). Version 3 introduces no modifications to the computation of H_i; it clarifies only the scope and interpretation of the measurement.