



# Waves of Collapse

## Compression as the Geometry of Reachable Description

Moving from frozen codecs to rematching archives and emergent world models.

Based on the 2026 monograph by Flyxion.

# Classical compression freezes ontology at the moment of ingestion.



## Fixed Representational Basis

Vocabularies (MDCT, LPC, dictionaries) are chosen at design time. The codec cannot discover new concepts, such as a localized drum hit, **drafting blue**.



## Write-Once Encoding

Once a block is compressed, its encoding is final. Nothing learned later improves it. The archive **becomes a graveyard** of past states of knowledge.



## Modality Isolation

Audio knows nothing of video. Each modality is **completely isolated** and compressed against models that cannot inform one another.

A rematching archive is a living theory of the data it has seen.

	The Frozen Codec	The Rematching Archive
The Basis	Fixed at design time.	Grows infinitely via a Hierarchical Template Library.
The Encoding	Write-once and static.	Retroactive and continually rematching.
The Scope	Modality isolated.	Shared across modalities as Latent World Structure.

An archive is not a pipe. It is an epistemic agent.  
Compression ratio is a direct measure of its understanding.

# Data is stored as a path through a topological template space.

## Level 3+ (Generative Abstractions)

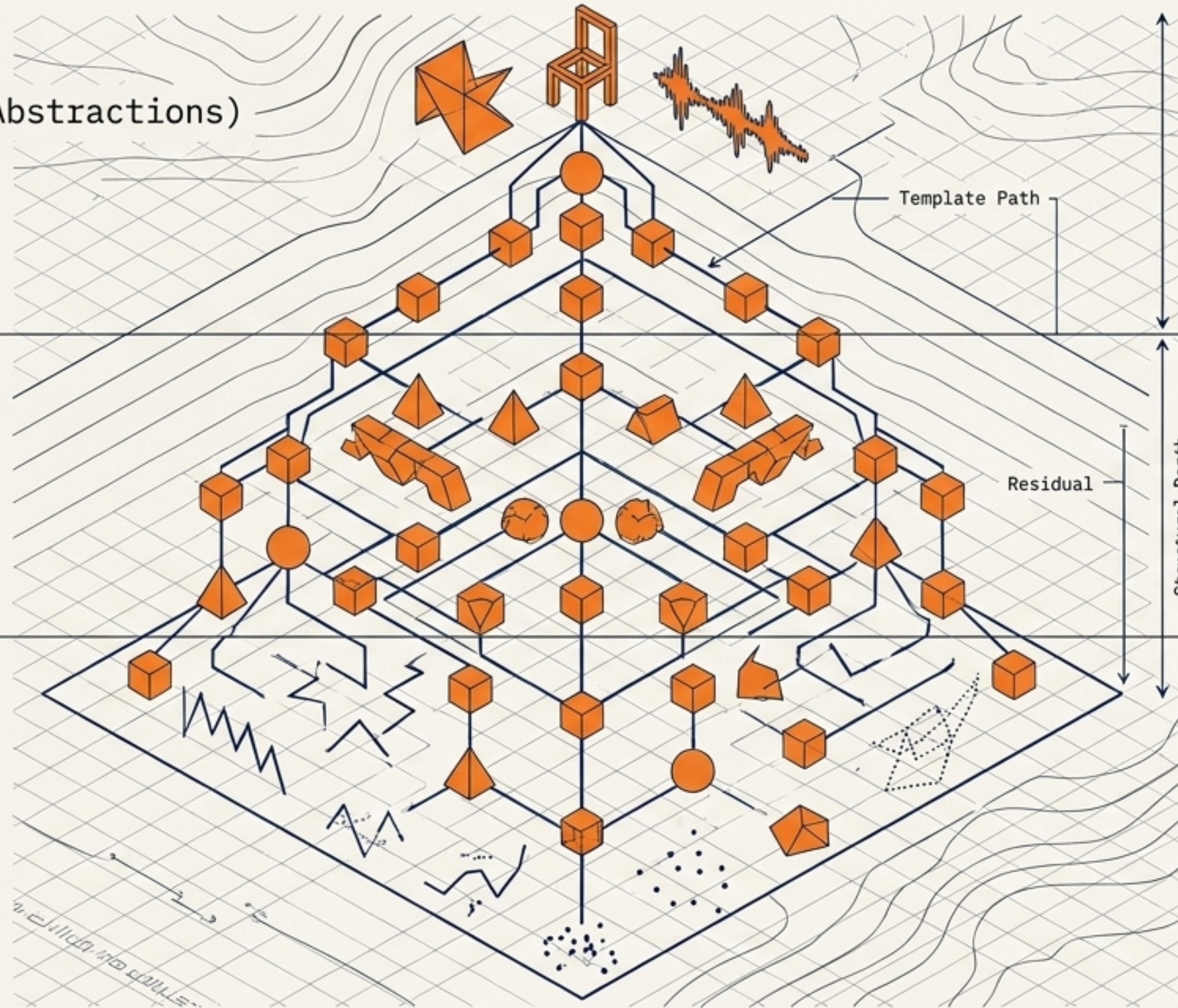
Deep, complex concepts  
(a 3D chair, a song chorus).

## Level 2 (The Grammar)

Combinations of fragments  
(musical phrases, visual sequences).

## Level 1 (The Bedrock)

Raw fragments and the  
quantization floor (audio  
waveforms, image patches).



## Core Principle

The system builds an ontology strictly from the data.

Every stored object is encoded as a path through this template DAG plus a residual.

The deeper the template nesting, the richer the system's structural understanding.

# The background process evaluates, admits, and refactors knowledge.

Foreground

Ingest & Encode greedily under current templates

Background (The Relaxation Dynamic)

**ADMIT**  
Add a new template.  
A speculative hypothesis  
tested against future data.

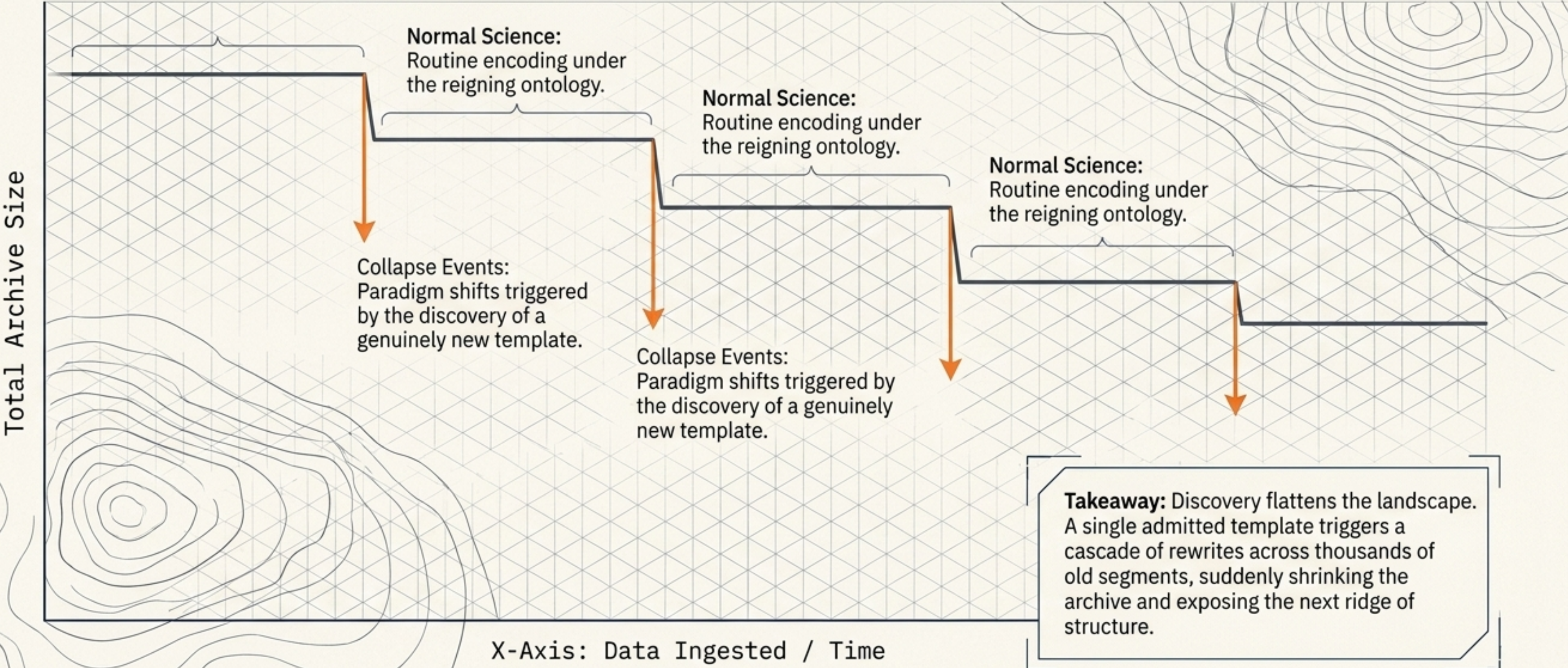
**REWRITE**  
Replace an old encoding  
with a cheaper one based  
on new templates.

**RETIRE**  
Remove an unreferenced  
template. The ultimate bankruptcy  
for a failed hypothesis.

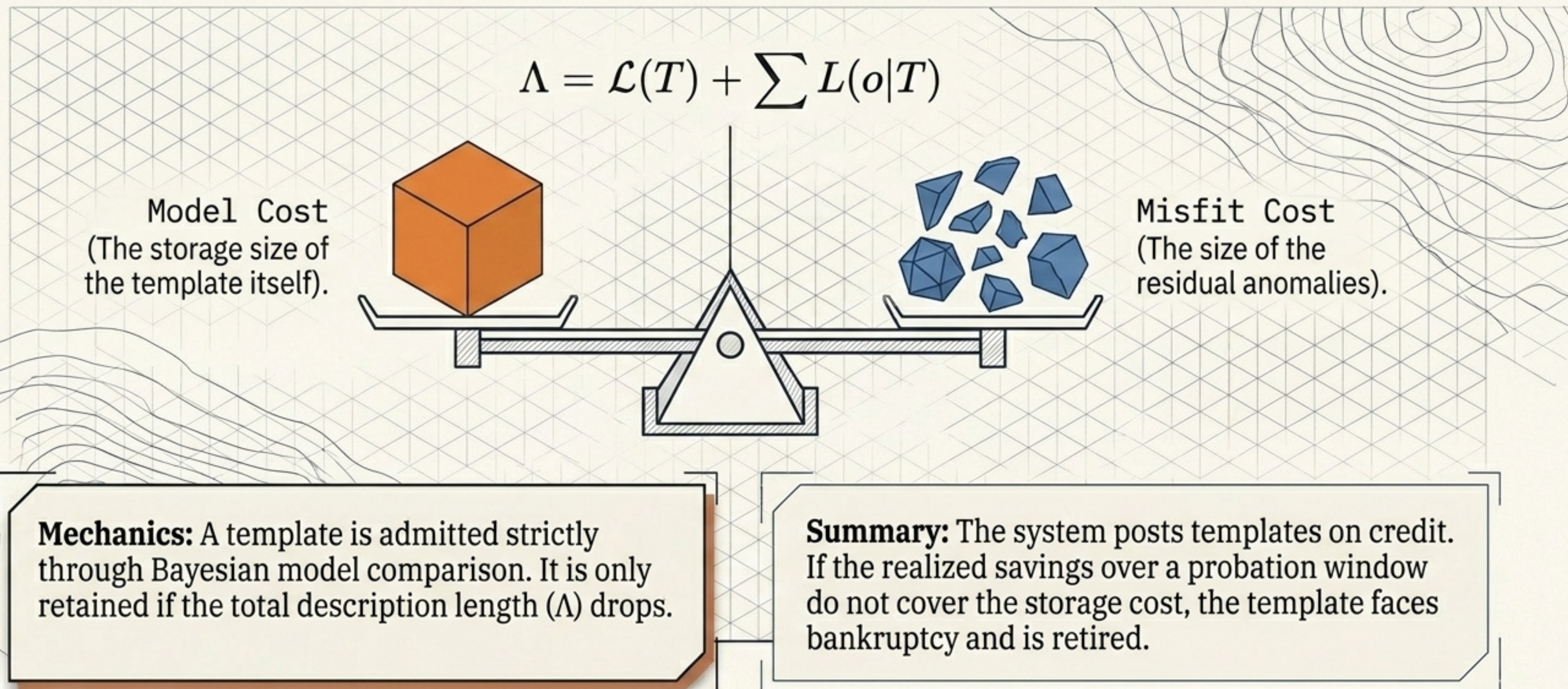
**REFACTOR**  
Rewrite templates in  
terms of other templates  
to streamline the DAG.

Key Detail: Identity is preserved across these continuous rewrites via an event-sourced ledger. The evaluation alters the internal mapping, never the identity.

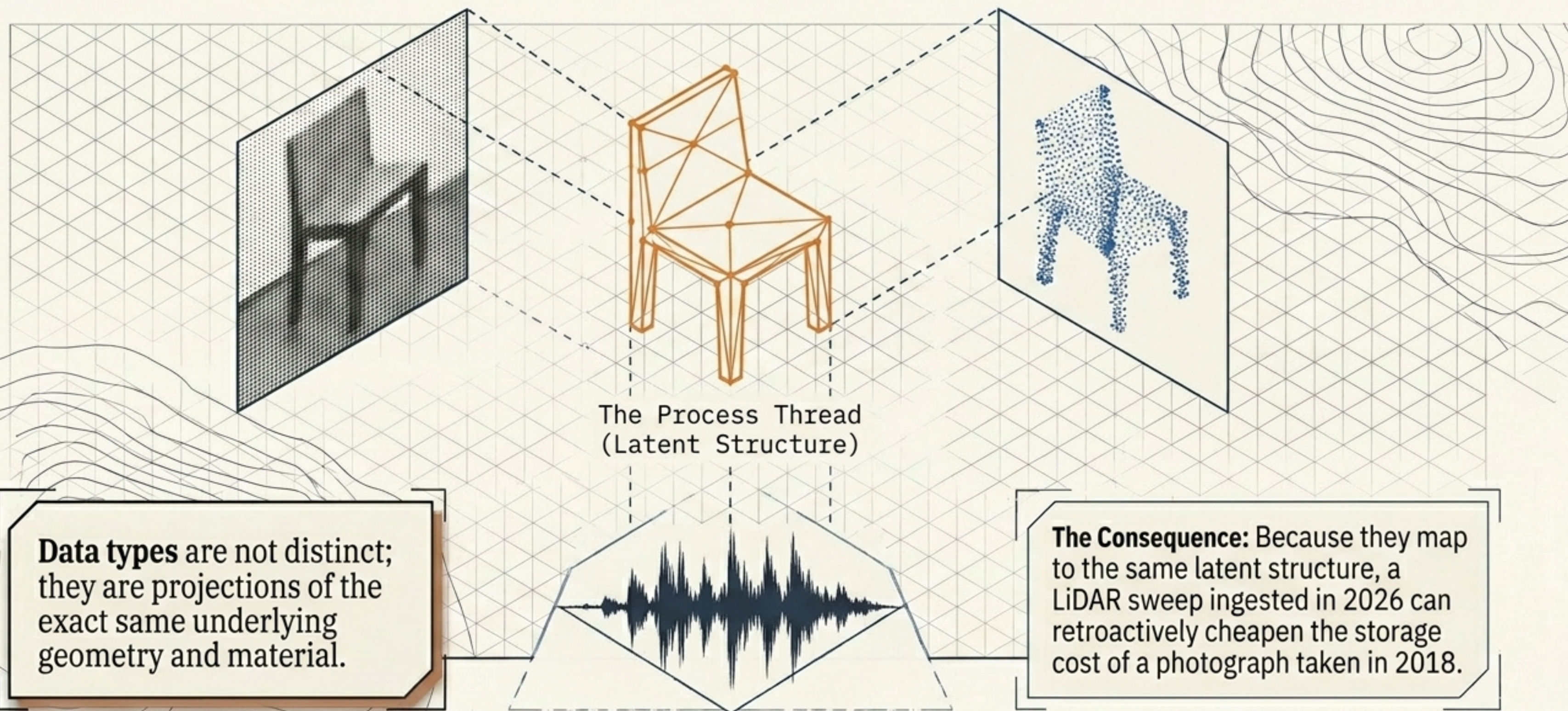
# Compression improves in punctuated, autocatalytic waves of collapse.



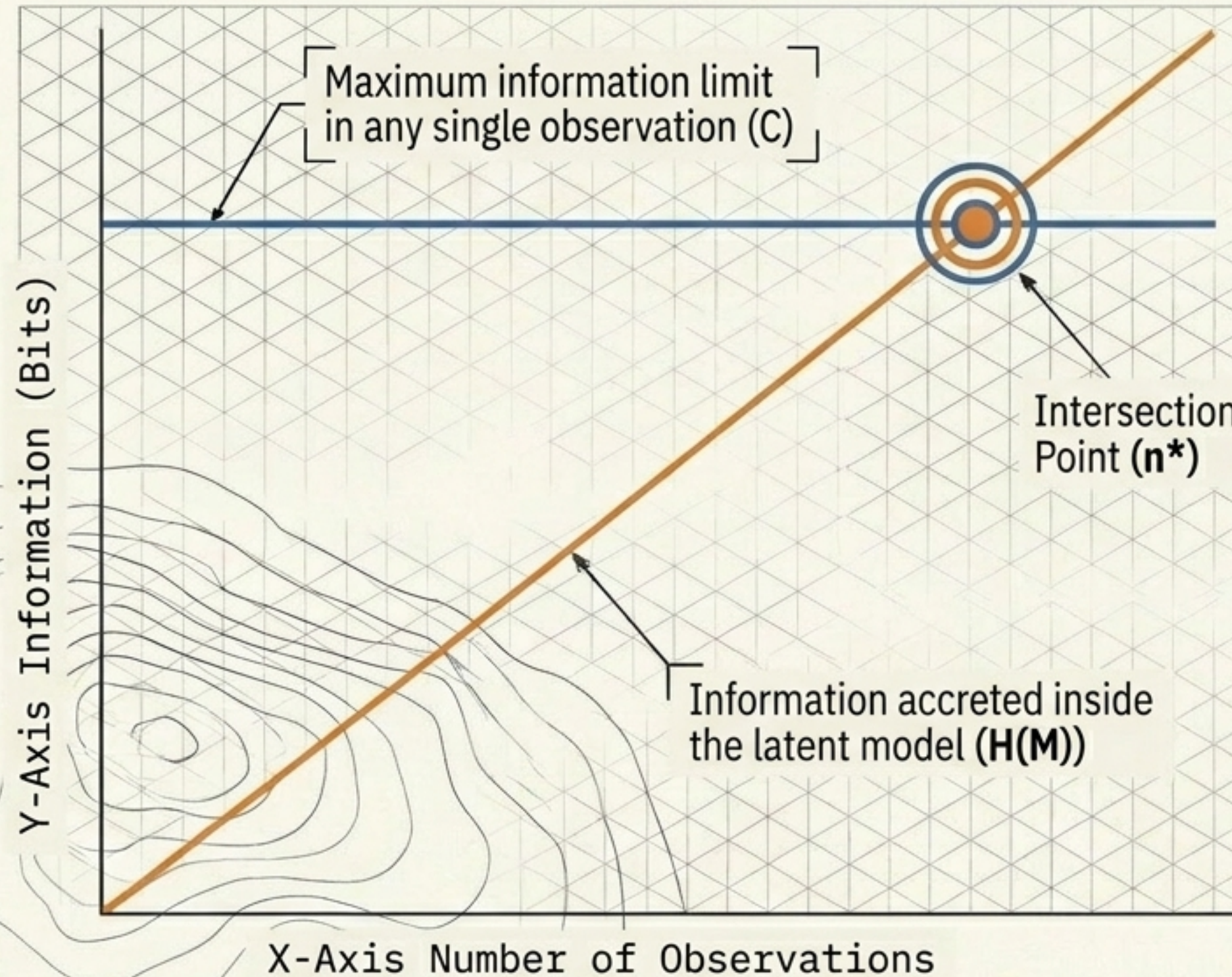
The two-part code: every template must pay its own rent.



Modalities are merely different projection kernels of a shared latent thread.



# The crossover point: when an archive becomes a world model.



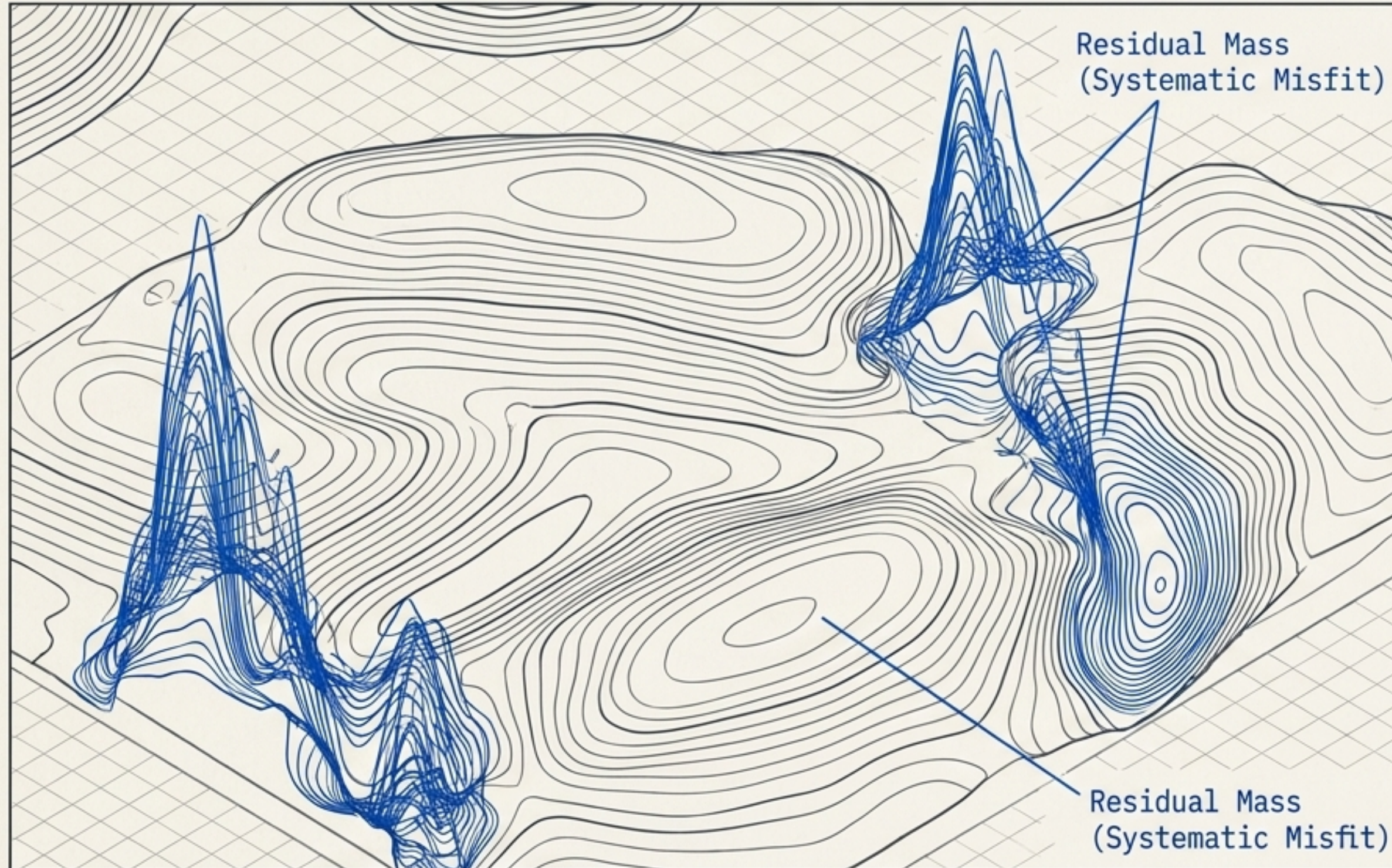
**The Transition:** Past the  $n^*$  intersection, the latent model contains more information than any individual recording.

**New Encoding Rule:** Once a latent entity is observed enough, the cheapest encoding for new data becomes: "Render from model, store the diff."




**Result:** Storage cost completely detaches from sensor data rates, asymptoting instead to the world's novelty rate.

# Residuals are not garbage; they are the system's research program.

TARTAN Heat Map: Archive Description Space

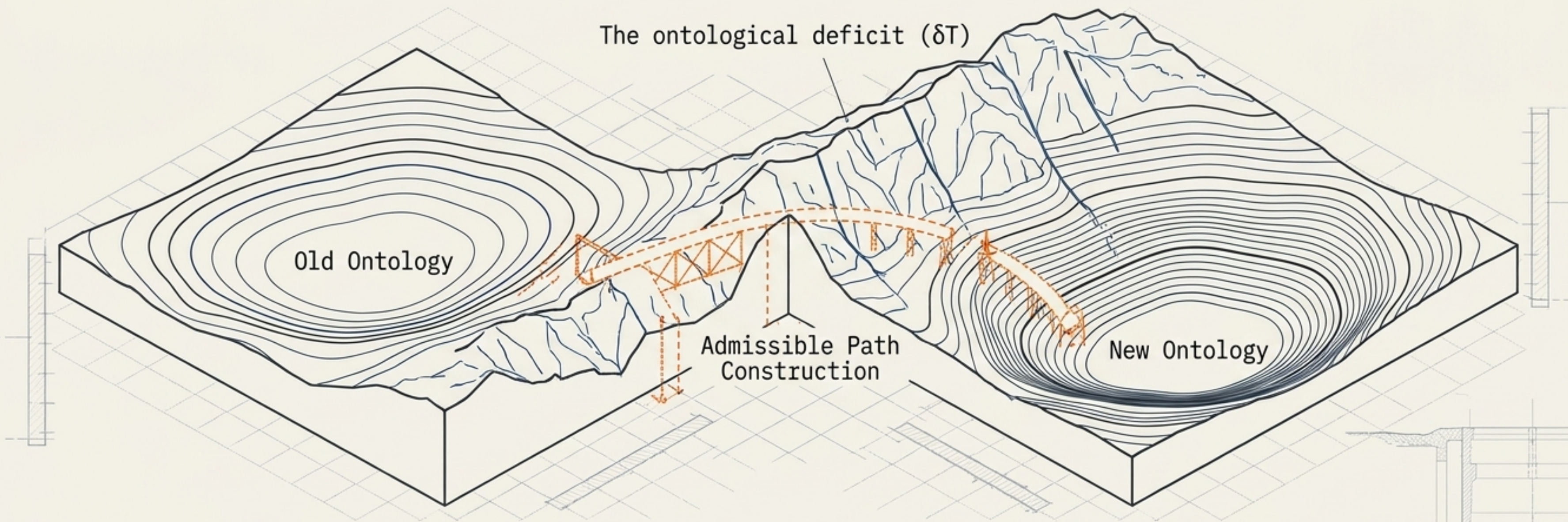


## Breakdown of Residuals

-  • **Noise:** Random, incompressible data.
-  • **Novelty:** Genuinely new information.
-  • **Systematic Misfit:** Compressible, but currently uncompressed structure.

**Core Insight:** The archive operates with intrinsic curiosity. A dense cluster of structured residual marks the exact coordinates where the next collapse event is waiting to be discovered.

# Epistemology: knowledge is the navigation of reachable states.



## Reachability:

A representation only exists if there is an admissible path built to it.

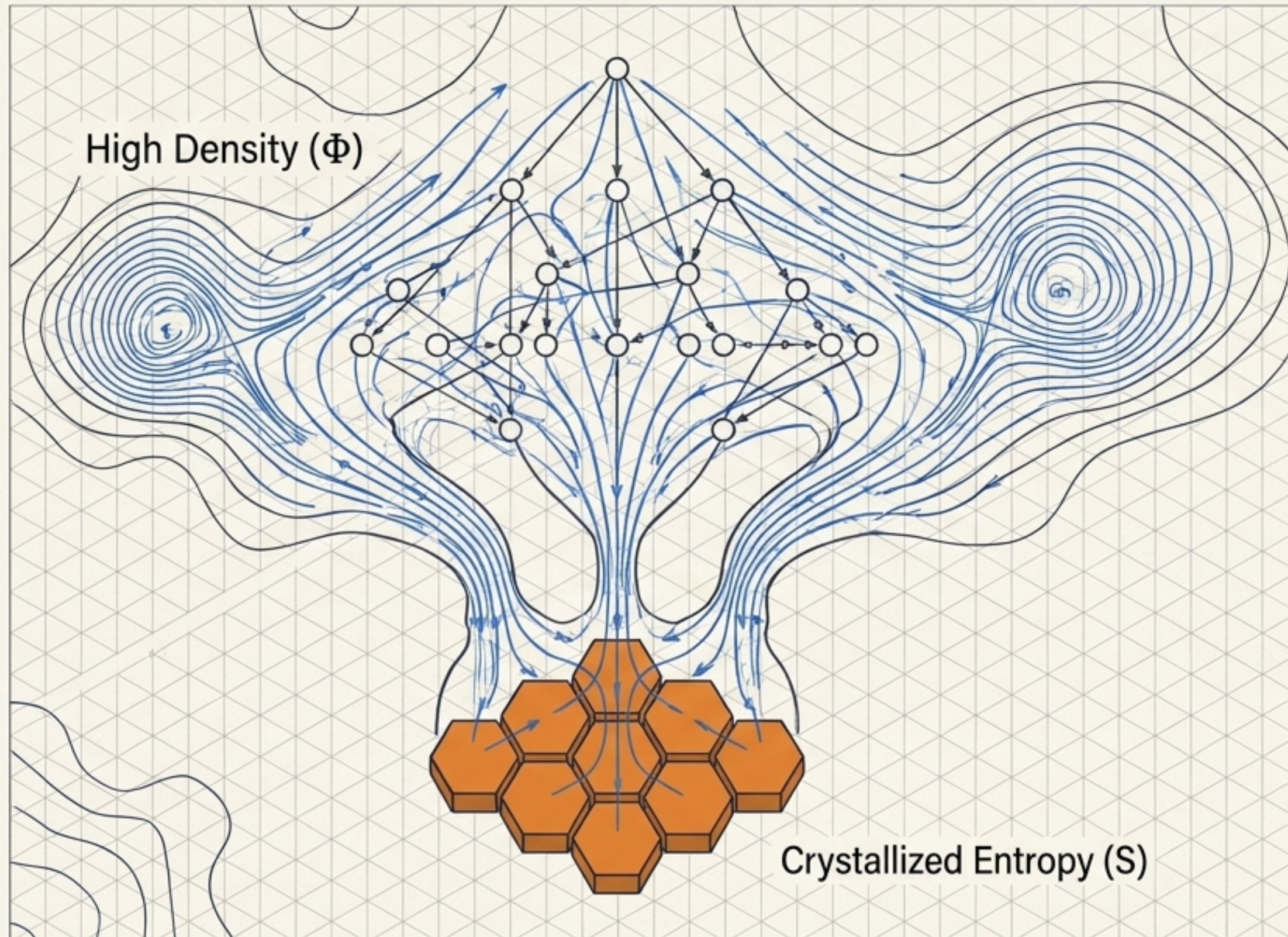
## Kuhnian Revolutions:

Persistent anomalies (stubborn residuals) force the system to admit new templates, tearing the old reference DAG and rewriting the past.

## Conclusion:

Learning is not a smooth search; it is the physical construction of new paths across an admissibility barrier.

# The thermodynamic field mechanics of continuous rematching.

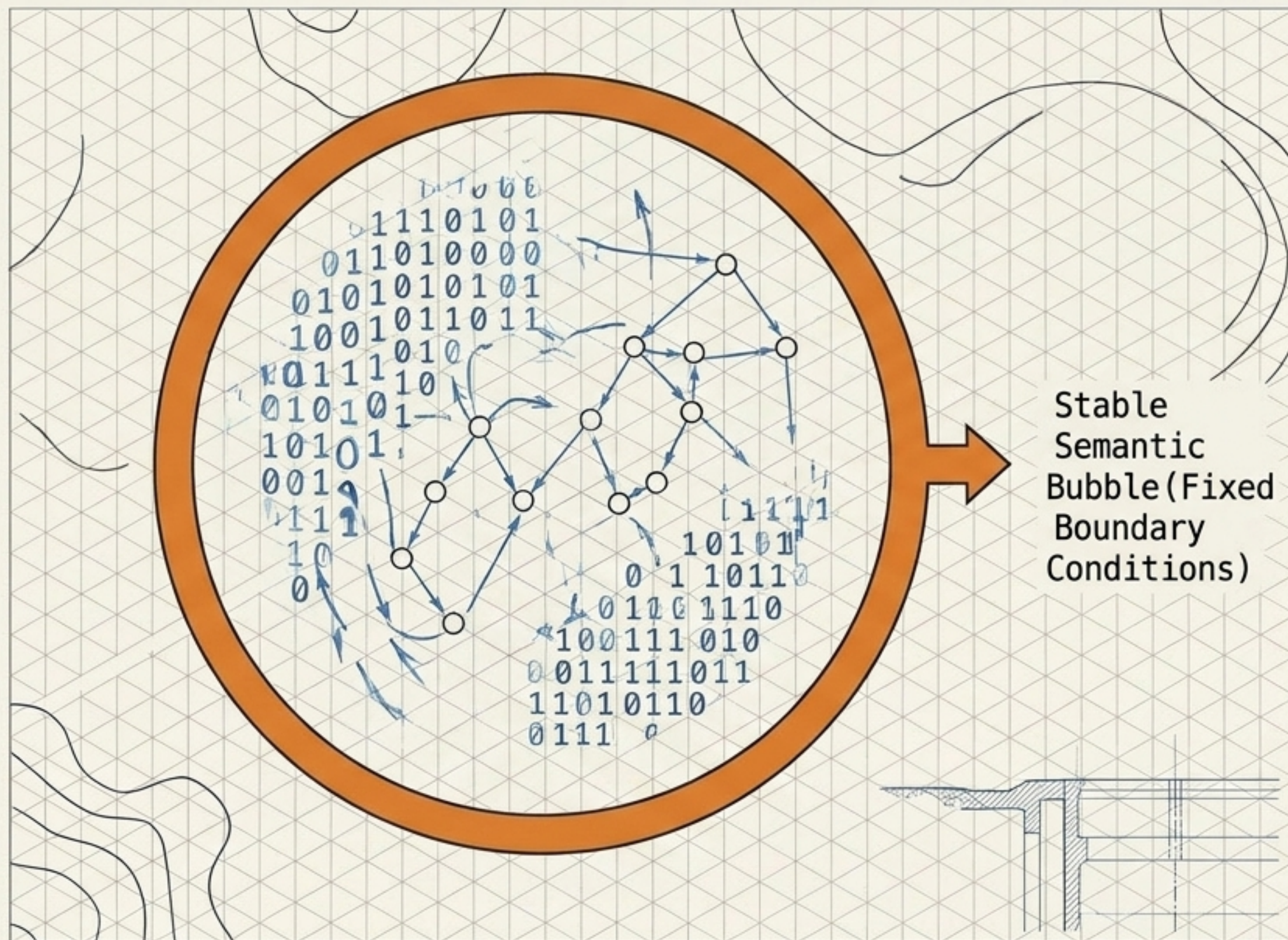


## RSVP Formalism

- $\Phi$  (Density): The anomaly field. The density of unexplained, compressible structure trapped in the system.
- $v$  (Flow): The rematching cascade. The directed transport of descriptions through the reachability graph.
- $S$  (Entropy): The realized code. The archive's crystallized, structured understanding.

**The Physics of Collapse:** A collapse event is a literal shock front where velocity  $v$  drops to negative infinity. It is a catastrophic convergence where trapped mass funnels into crystallized entropy.

# Identity without substrate: preserving the semantic bubble.



Stable  
Semantic  
Bubble (Fixed  
Boundary  
Conditions)

## The Rules of Identity

1. Boundary conditions (what must decode to what, at what fidelity) are permanently fixed at ingestion.
2. An event-sourced ledger tracks the history of irreversible acts and bubble evaluations.

## The Insight

Identity is not a static bit pattern. A file persists purely as the capacity to be regenerated along an admissible path.

# System pathologies and the engineering of mitigation

Pathology	Theory Translation	Mitigation
Template Capture	An early bad segmentation becomes frozen and load-bearing.	<i>Scheduled reheating (temporarily loosening match thresholds to escape local minima).</i>
Hallucinated Economy	Overconfident models drifting into plausible but wrong lossy reconstructions.	<i>Maintain the distortion lattice; keep residual retention high enough to keep the model falsifiable.</i>
Ontology Lock-in	One modality's successful template suppresses the admission of a better one for a different modality.	<i>Enforce strict, per-modality anomaly audits.</i>

The opposite promise.

“A codec is a promise to forget nothing while understanding nothing.

**An archive that rematches makes the opposite promise: it will understand more and more, and what it stores will weigh less less, and the two facts will be the same fact.”**

The world simply becomes reachable from a shorter sentence.