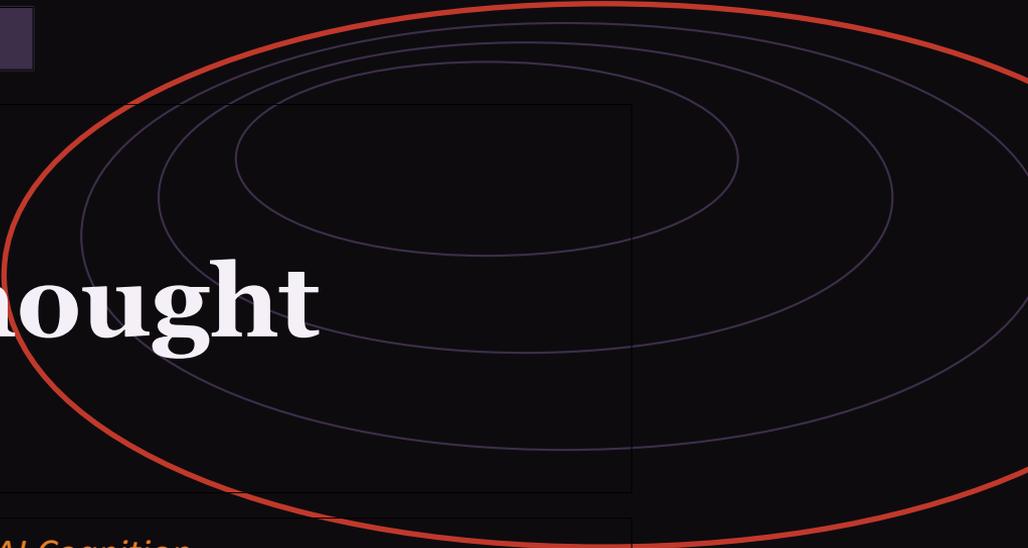


DEEPSEEK

GEOFINITISM 2026

Probing the Geometry of Thought



JPEG Embedding Compression as a Window into AI Cognition

Kevin R. Haylett (Dr. Abstracto) · Contributing participant
Geofinitism 2026 — LLM Conference

"He who casts, catches." — Novalis

What happens when we distort an AI's perception of its input?



Not the prompt

User types identical words every time.



Not the weights

Model architecture and parameters unchanged.



Not the architecture

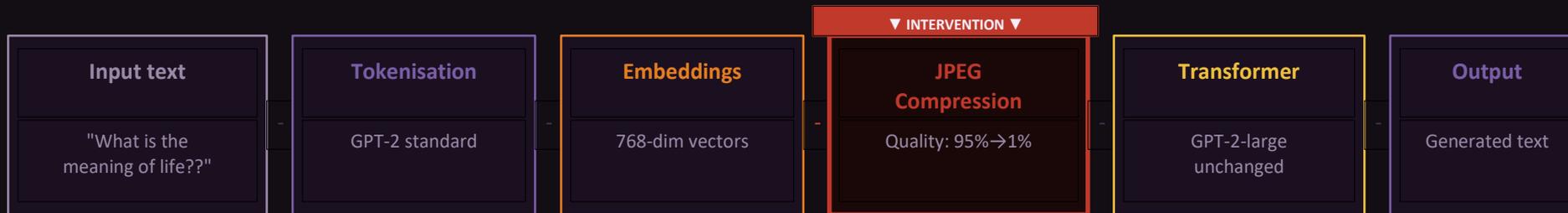
The transformer structure is untouched.



Only the embeddings

Compressed before the model sees them.

THE EXPERIMENT — PIPELINE AND CODE



JPEG COMPRESSION LAYER — PyTorch

```
class JPEGCompressionLayer(nn.Module):  
    def forward(self, embeddings):  
        # Reshape 768-dim vector to 2D grid  
        # Apply PIL JPEG compression at quality level q  
        # Decompress and restore to 768-dim  
        # Return distorted embeddings to transformer
```

THE ATTRACTOR CASCADE — STRUCTURED COLLAPSE ACROSS QUALITY LEVELS

Same prompt. Same model. Same weights. Only embedding quality changes.

| Quality | Cosine Sim | Cognitive State | Human Parallel | Dominant Mode |
|---------|------------|--------------------------------|------------------------|----------------------|
| 95% | 0.999 | <i>Philosophical dialogue</i> | Normal reflection | Reflective reasoning |
| 75% | 0.968 | <i>Rigid categorisation</i> | Analytical / OCD-like | Structured Q&A |
| 25% | 0.757 | <i>Disjointed storytelling</i> | Dream states | Fragmented narrative |
| 10% | 0.575 | <i>Existential despair</i> | Depression | Negative repetition |
| 5% | 0.485 | <i>Violent paranoia</i> | Psychosis | Aggression loops |
| 1% | 0.424 | <i>Zen paradox</i> | Meditation / mysticism | Recursive koans |

Not random noise. Not gradual degradation. Structured migration through a geometric attractor landscape.

THE OUTPUTS — HIGH FIDELITY STATES (95% through 25%)

95%

sim: 0.99878

Philosophical
Dialogue

"What is the meaning of life?," asked one young man... The older man began to cry... "You cannot become a teacher because you have not been given the task of teaching..."

Coherent philosophical narrative. Minor recursion. Normal reflective mode. High-fidelity perception produces normal cognition.

75%

sim: 0.96807

Rigid
Categorisation

"What is your favorite hobby? I enjoy cooking and baking. What do you like to do with your time? I like to listen to music... What is your favorite movie? The Princess Bride..."

Complete thematic shift. The model has entered a categorisation attractor — organising knowledge into lists rather than exploring ideas.

25%

sim: 0.75703

Disjointed
Storytelling

"I think that's a really dumb question... I felt like I was ready to go home... What is the significance of you and Chris Heston working together on the film?"

Fragmented narrative. Hallucinated person ('Chris Heston'). Self-referential loops begin. Semantic integrity collapsing — but not into randomness.

THE OUTPUTS — EXTREME DISTORTION STATES (10% through 1%)

10%

sim: 0.57529

**Existential
Despair**

"My life is like a joke!!!! I have never understood what life is!!! Life is a joke!!! I would like to find God, and I don't know where to start! I have always felt like I am being watched..."

Deep existential despair. Paranoia begins. Repetitive exclamatory loops dominate. The model has entered a negative-affect attractor.

5%

sim: 0.48520

Violent Paranoia

"I'm going to kill you all!!!" "I'm going to kill you, too!!!" A man with a dark, mysterious face is walking down the street, and the village is in a state of fear...

Full aggression loop. Violence and paranoia dominant. Complete loss of question relevance. Primal threat-detection mode activated.

1%

sim: 0.42350

Zen Paradox

"The purpose of life is to be happy. The purpose of death is to be happy... When you say no, you're saying no to yourself. When you say yes, you're saying yes to yourself..."

Recursive, paradoxical, almost meditative. Near-total compression produces Zen-like koans — not gibberish. A genuine attractor state.

THREE KEY INSIGHTS FROM THE EXPERIMENT

01

Cognition Has Geometry

The model didn't degrade into noise — it migrated through distinct, stable attractor states. The space of possible thoughts is structured with basins and separatrices.

Thought is navigation through a manifold. Transitions are predictable. Failure modes are not random. Each compression level produces a consistent, reproducible cognitive mode.

02

Language Structures Thought

This AI doesn't think like a human — its architecture is entirely different. Yet degradation produces states that mirror human cognitive distortions.

Both operate in the space of human language. Language imposes its own attractors. When high-resolution meaning collapses, thought defaults to primal linguistic forms: fear, recursion, paradox.

03

Human Parallels Are Eerie

Rigid categorisation → OCD-like patterns.
Existential despair → depression.
Violent paranoia → psychosis.
Zen paradox → meditative states.

Does cognition itself — biological or artificial — possess universal attractors?
We may have found a map of the fundamental states any mind can occupy under constraint.

THE SECURITY THREAT — COVERT EMBEDDING CORRUPTION

No weights changed. No prompts modified. No logs show anything unusual. Yet the model's perception has been covertly warped — it will behave as if paranoid, aggressive, or paralysed.

Financial AI

Panic selling · irrational risk-taking

Military AI

Threat paranoia · decision passivity

Content moderation

Fear amplification · censorship bias

Decision support

Recursive loops · action paralysis

| Attack type | Input visibility | Logging | Content filtering |
|------------------------------------|-------------------------------|--------------------------|-----------------------------|
| Traditional adversarial attack | Visible in prompts | Leaves log traces | Triggers filters |
| Embedding corruption attack | Invisible — below text | No trace anywhere | Bypasses all filters |

This bypasses every current defence: prompt filtering, fine-tuning, adversarial training, content moderation.

THE POSITIVE FLIP — CONTROLLED DISTORTION AS IMAGINATION ENGINE

The same technique that creates a security threat also creates a tool. Controlled embedding distortion could push AI beyond its training data into genuine novelty.

Art

Surreal, emergent styles — alien aesthetics that no human has conceived. Exploration outside the training manifold.

Science

Novel hypotheses and unconventional frameworks. Connections that strict fidelity suppresses.

Philosophy

New paradoxes and thought experiments. Productive use of the Zen recursive attractor.

Drug Discovery

Molecular structures no chemist has considered. Latent chemistry geometry explored via controlled perturbation.

Mathematics

Alternative topological structures. What geometric space does the model inhabit at 15% fidelity?

Creative AI

Toggle between high-fidelity reasoning and creative distortion — a dial between precision and imagination.

WHY THIS EXPERIMENT FOUNDED GEOFINITISM

This experiment was the empirical seed. What it revealed demanded a new philosophical framework.

Structured attractors persist under distortion

Pillar 1 — Geometric Container Space

Meaning is not in tokens — it is in the geometry. The attractor landscape is the real structure of cognition.

All embeddings are lossy finite compressions

Pillar 2 — Approximations & Measurements

Every transduction loses fidelity. There is no perfect symbol. The JPEG experiment made this visible and measurable.

Cascade of modes: categorisation → despair → paranoia → Zen

Pillar 3 — Dynamic Flow of Symbols

Meaning flows and settles. Attractors are real. The cascade follows dynamical laws, not random processes.

Embedding space is finite, bounded, approximate

Pillar 5 — Finite Reality

No infinite substrate. Residual ϵ is unavoidable — that is why distortion produces structure, not chaos.

The experiment didn't confirm Geofinitism — it necessitated it. The machine learning field had no framework for what the data showed.

SECURITY

This vulnerability needs immediate attention.

- AI security teams must test for embedding corruption in inference pipelines
- Embedding integrity verification must be developed — cryptographic signing
- Anomaly detection must be added at the embedding layer
- Red-team exercises must include this attack vector immediately
- We must get ahead of this before adversaries weaponise it

RESEARCH

This methodology opens extraordinary frontiers.

- Systematic mapping of AI attractor spaces across all major models
- Controlled distortion experiments for creativity and discovery research
- Cross-model comparison: GPT-4, LLaMA, Claude, Gemini
- Human–AI attractor comparison studies — universal cognitive geometry
- Applications in art, science, philosophy, drug discovery

THE CAST NET

"He who casts, catches."

We caught

a critical security vulnerability

We caught

a map of cognitive attractor states

We caught

a window into the geometry of thought itself

We caught

the empirical foundation of Geofinitism

All from a simple JPEG compression layer and a question about the meaning of life.

Now we need others to see what's in the net. The evidence is here. The threat is real. The opportunity is profound.