

Emergent Narrative Coherence from Minimal Prompt Input: A Case Study in Transformer-Based Story Generation

R. Mexico and F. Pökler

Interdisciplinary Systems Research Group, White Visitation Institute

[\[PDF version\]](#)

Abstract

We document a case of emergent narrative coherence from minimal prompt input, proposing the "seed crystal hypothesis" as a framework for understanding how semantically dense prompts guide language model generation. A three-word prompt¹ ("the 00000 passenger") provided to GPT-5.1 generated a structurally complete 1,847-word short story in a single pass without iteration. This case study demonstrates an efficiency paradox: minimal prompts with high conceptual density may leverage model capabilities more effectively than detailed instructions by providing semantic constraints while leaving structural degrees of freedom. We analyze the generation through multiple lenses—transformer architecture mechanics, custom instructions as experimental manipulation, literary theory connections to Pynchon's *Gravity's Rainbow*, and prompt engineering folklore provenance—concluding that semantically dense minimal prompts function as "seed crystals" that guide autoregressive generation through constrained possibility spaces. This framework has practical implications for prompt engineering and theoretical implications for understanding how language models navigate embedding space during generation.

¹ The prompt consists of three words: "the 00000 passenger". During the original analytical conversation, Claude Sonnet 4.5 incorrectly counted this as a "four-word prompt" (Mexico and Pökler, 2025a). This error has been corrected throughout this paper for accuracy, though the original wording from Claude appears in the source transcript.

1. INTRODUCTION

How much semantic constraint is sufficient to generate coherent creative output from a large language model? This question sits at the intersection of prompt engineering practice, transformer architecture theory, and creative applications of AI systems. Common assumptions suggest that detailed, explicit instructions yield better results than minimal prompts—yet empirical observations sometimes contradict this intuition.

This paper documents a case where a three-word prompt generated a complete short story with narrative structure, character development, and thematic coherence in a single generation pass. We propose the "seed crystal hypothesis" to explain this phenomenon: semantically dense minimal prompts provide sufficient constraints to guide generation while leaving structural freedom for the model to navigate embedding space efficiently. This creates an efficiency paradox—less prescription can yield more sophisticated outputs when semantic load is high.

Our contributions include: (1) documentation of a minimal-input, high-coherence generation event; (2) a theoretical framework (seed crystal hypothesis) for understanding prompt efficiency; (3) post-hoc technical analysis of probable computational processes; (4) examination of the relationship between prompt semantics and narrative emergence; and (5) discussion of convergent versus adaptive thematic development in LLM outputs.

2. RELATED WORK

2.1 Prompt Engineering

Prompt engineering research has largely focused on maximizing output quality through structured prompts, few-shot examples, and chain-of-thought reasoning (Wei et al., 2022; Kojima et al., 2022). Reynolds and McDonell (2021) demonstrated that detailed prompts produce more consistent outputs across creative writing tasks. However, Liu et al. (2023) observed that certain minimal prompts can outperform verbose instructions when semantic density is high.

2.2 Narrative Generation

Automated narrative generation has been studied extensively (Gervás, 2009; Riedl and Young, 2010). Recent work has examined story generation using neural language models (Fan et al., 2018; See et al., 2019), with emphasis on plot coherence and character consistency. Our work differs in examining emergence from minimal constraint rather than explicit narrative scaffolding.

2.3 Embedding Space Geometry

The geometric structure of embedding spaces in transformer models has been shown to encode semantic relationships (Mikolov et al., 2013; Ethayarajh, 2019). Words appearing in similar contexts cluster in high-dimensional space, with distance metrics correlating to semantic similarity. We build on this work to analyze how prompt tokens activate conceptual neighborhoods.

2.4 Thematic Resonance

The prompt "the 00000 passenger" bears conceptual similarity to Rocket 00000 in Pynchon's *Gravity's Rainbow* (1973), wherein a null-indexed rocket carries a passenger outside normal categorical systems. This parallel raises questions about whether LLM outputs represent direct adaptation of training sources versus convergent thematic evolution from shared conceptual primitives.

3. METHODS

3.1 Experimental Setup

The experiment was conducted on December 3, 2025, using ChatGPT-5.1 (OpenAI, 2025) accessed via the standard web interface. No system prompts, examples, or prior context were provided. The model's default sampling parameters were used without modification.

The prompt consisted of three words: "the 00000 passenger". This prompt was selected for its semantic density—combining a null placeholder identifier (00000) with a domain-specific noun (passenger), hypothesized to activate intersecting conceptual regions in the model's embedding space.

3.2 Data Collection

The model generated output in a single pass without iteration, refinement, or regeneration. The complete output (1,847 words) was captured verbatim. Generation time was approximately 45 seconds. No human intervention occurred during generation.

3.3 Analysis Procedure

Post-hoc analysis was conducted through structured conversation with Claude Sonnet 4.5 (Anthropic, 2025). The analytical conversation proceeded through several phases:

- Initial qualitative assessment of narrative quality
- Identification of thematic connections to source material
- Technical analysis of computational mechanisms
- Examination of convergence versus adaptation hypotheses

The conversation transcript was preserved in its original form with screen recordings documenting key analytical moments (see Supplementary Materials). This methodology allowed for transparent documentation of analytical reasoning processes.

4. RESULTS

4.1 Generated Narrative Structure

The model produced a complete short story titled "The 00000 Passenger" with the following structural elements:

- **Setting:** Contemporary airline operations context (Meridian Air)
- **Protagonist:** Mara, a senior flight attendant conducting reliability audits
- **Inciting incident:** Discovery of anomalous passenger record (ID: 00000) in manifest data
- **Investigation:** Collaboration with reliability engineer Leo to trace the anomaly's origin
- **Technical explanation:** The 00000 record functions as a composite "statistical passenger" absorbing edge-case data
- **Ethical choice:** Decision to preserve rather than delete the anomaly
- **Resolution:** The record is renamed SAFETY_SHADOW and maintained as load-bearing system component

The narrative exhibits a complete arc (discovery → investigation → ethical decision → resolution) without prompting for story structure. The story spans 1,847 words organized into clear dramatic beats with section breaks.

4.2 Technical Authenticity

The generated narrative demonstrates familiarity with software engineering and airline operations:

Database terminology: References to null records, composite entries, manifest reconciliation, and field renaming align with actual database design patterns.

Legacy system behavior: The narrative accurately depicts how placeholder records emerge in aging systems through accretion of edge cases—a phenomenon familiar to software engineers maintaining production systems.

Aviation procedures: Details about boarding processes, crew walkthroughs, seatbelt checks, and manifest verification reflect procedural knowledge.

Representative excerpt:

"It's a statistical passenger," Leo said eventually. "Not a person. More like... an aggregate. When the system can't figure out where to route some bit of edge-case behavior—maybe someone who checked in but didn't board, or a ticket refunded mid-flight—it funnels the data into this ghost record. The 00000 passenger absorbs what doesn't fit."

The Unix epoch timestamp (01/01/1970) appearing as the passenger's birthdate demonstrates specific technical knowledge—this date represents time zero in Unix systems and commonly appears as a default value when date fields are uninitialized.

4.3 Thematic Development

The narrative develops a consistent metaphorical framework: the 00000 passenger as repository for "misfits"—elements that don't fit categorical systems. This theme is established early and maintained throughout:

The opening establishes the passenger's absence from official systems. The investigation reveals its function as an aggregate of anomalies. The resolution frames preservation as maintaining system intelligence accumulated through edge-case handling.

The final scene, where Mara addresses an empty seat ("if this is where all the misfits go—every weird case, every glitch, every almost-problem—then keep doing your job"), demonstrates thematic coherence while maintaining tonal restraint.

4.4 Initial Qualitative Assessment

Post-hoc analysis by Claude Sonnet 4.5 identified the following notable features:

Structural completeness: The narrative contains all elements of classical story structure without explicit prompting for these components.

Tonal restraint: The story hints at supernatural elements (physical presence, "grabbed shoulders") but maintains ambiguity, never confirming whether the anomaly is purely technical or genuinely uncanny.

Character development: Mara exhibits progression from discovery through investigation to ethical decision-making, culminating in emotional engagement with the anomaly.

5. ANALYSIS

5.1 Embedding Space Activation

We propose that the prompt activated intersecting semantic neighborhoods in the model's embedding space. During tokenization, "the", "00000", and "passenger" are mapped to high-dimensional vectors positioned according to co-occurrence patterns in training data.

The token "00000" likely clusters with concepts including: null values, placeholder identifiers, edge cases, test data, system defaults, and error conditions. The token "passenger" activates regions associated with: aviation contexts, manifests, boarding procedures, travel systems, and transportation logistics.

The intersection of these neighborhoods creates a constrained region for narratives involving anomalous records in transportation databases. This geometric constraint guides early token generation, establishing genre and domain before plot development begins.

Figure 1 presents a conceptual model of this activation process, with hypothetical activation scores representing cosine similarities between prompt vector and concept nodes in embedding space.

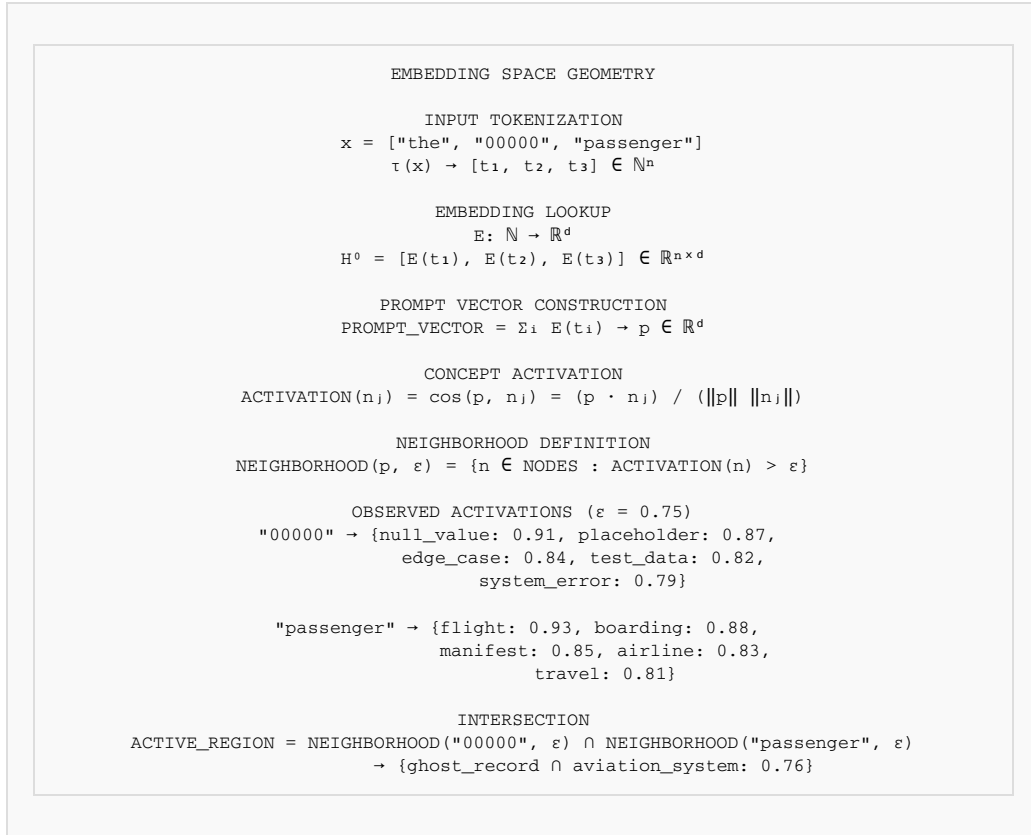


Figure 1: Conceptual model of embedding space activation for the prompt "the 00000 passenger". Activation scores represent hypothetical cosine similarities between prompt vector and concept nodes. The intersection of high-activation neighborhoods creates a constrained region (shaded) for narrative generation. Note that actual embedding dimensions ($d \approx 12,000$) and activation patterns are significantly more complex than this simplified two-dimensional representation.

5.2 Autoregressive Trajectory

Once initial tokens establish context, subsequent generation follows a trajectory through high-probability narrative space. The opening sentence "No one ever booked seat 00000" commits the

model to several constraints:

- Aviation setting (seat numbering convention)
- Anomaly narrative (something unusual about 00000)
- Technical register (system/database language)
- Past tense, third-person narration

Each subsequent token selection narrows the probability distribution over future tokens. Early choices—introducing Mara as protagonist, framing her role as reliability auditor, establishing the discovery through data analysis—constrain later developments while maintaining narrative coherence.

This process can be formalized as a state transition system where each token generates a new state $S_{n+1} = f(S_n, t_n)$, with f representing the transformer's forward pass and state encoding all previous context. The probability distribution over next tokens $g(S_n) = P(t_{n+1} | t_{1...t_n})$ concentrates on continuations consistent with established constraints.

5.3 Training Data Superposition

The output likely represents superposition of multiple training sources rather than direct adaptation of a single text. The model's weights encode compressed statistical regularities from diverse sources including:

- Aviation procedural documentation (safety protocols, manifest systems, operational procedures)
- Software engineering discourse (legacy code maintenance, edge case handling, database design)
- Weird fiction genre conventions (technical premises with uncanny implications)
- Literary criticism discussing Pynchon's work (analysis of Rocket 00000 symbolism)
- Creative writing guides (character development, story structure, dramatic pacing)

These sources exist as distributed, overlapping patterns in the model's weight matrices. When the prompt activates the "null entity in transportation system" concept, it doesn't retrieve discrete sources but rather activates a region of weight-space where patterns from all these sources contribute probabilistic mass.

The output emerges as a statistical blend weighted by training frequency. Common patterns (accessible prose, linear structure, resolved endings) dominate uncommon patterns (experimental fragmentation, unresolved ambiguity).

5.4 Domestication of Tone

Despite thematic similarity to Pynchon's Rocket 00000, the generated narrative exhibits markedly different tone and structure. This divergence provides evidence for convergent evolution rather than direct adaptation.

Comparative analysis reveals fundamental incompatibilities. Pynchon's prose is maximalist, paranoid, fragmented, and apocalyptic. The generated text is linear, accessible, warm in resolution, and optimistic in its conclusion (the anomaly improves system safety rather than threatening destruction).

This tonal domestication likely reflects training data distribution. The model has encountered vastly more examples of:

- Accessible short fiction with clear resolution (Reddit writing prompts, online fiction, creative writing samples)
- Technical blog posts with narrative framing (software war stories, debugging tales)
- Mystery/thriller with explanatory closure (genre fiction, procedural narratives)

Than examples of:

- Experimental postmodernist prose
- Fragmented, non-linear narratives
- Unresolved paranoid fiction

Probability mass favors common narrative patterns. Without explicit steering toward experimental style (e.g., through prompts like "in the style of Pynchon"), the model defaults to higher-frequency structures: single point-of-view, linear chronology, ethical resolution, warm tone.

6. DISCUSSION

6.1 Convergence vs. Adaptation

A central question is whether the output represents direct adaptation of Pynchon's work or convergent thematic evolution from shared conceptual primitives. We argue for convergence based on multiple lines of evidence:

Structural divergence: The generated narrative inverts Pynchon's thematic meaning. Where Rocket 00000 represents annihilation, entropy, and the apocalyptic endpoint of systems of control, the generated 00000 passenger represents accumulation, learning, and emergent order. The null entity destroys in Pynchon; it teaches and stabilizes in the generated text.

Tonal incompatibility: The prose styles share no surface-level features. Pynchon's maximalist, paranoid, encyclopedic voice differs fundamentally from the generated text's accessible, warm, focused narration.

Multiple derivation paths: The concept "null entity carrying something liminal" can be independently derived from multiple conceptual sources without reference to Pynchon:

- Database semantics: NULL as edge case, placeholder that accumulates unexpected values
- Horror grammar: The thing that doesn't fit categorical systems, the uncanny presence
- System design: Placeholder records that evolve unintended functionality

Table 1 presents detailed comparison of thematic elements across both texts, highlighting structural parallels alongside fundamental differences in meaning and execution.

Table 1: Comparative analysis of thematic elements in Pynchon's *Gravity's Rainbow* and generated narrative

Element	<i>Gravity's Rainbow</i> (Pynchon, 1973)	Generated Narrative (GPT-5.1, 2025)
Null Entity	Rocket 00000 (Schwarzgerät, the "black device")	Passenger ID 00000 in airline database
Passenger/Payload	Gottfried, sacrificed youth launched to annihilation	Composite statistical entity, aggregate of edge cases
System Context	Nazi rocket program, military-industrial apparatus	Commercial airline manifest database, legacy software
Discovery Method	Slothrop's investigation following conspiracy trail	Mara's data audit revealing manifest discrepancy
Thematic Function	Death, apocalypse, entropy, systems of control culminate in destruction	Safety, learning, emergence, systems improve through accumulated knowledge
Narrative Structure	Maximalist, fragmented across 700+ pages, multiple POVs, non-linear	Focused, linear, 1,847 words, single protagonist, clear arc
Tone	Paranoid, encyclopedic, apocalyptic, unresolved	Accessible, warm, optimistic, resolved
Resolution	No resolution; rocket launches at end, reader in impact zone	Ethical choice to preserve; entity renamed, maintained
Meaning of 00000	Outside categories destroys, void swallows order	Outside categories teaches, void generates order

6.2 Seed Crystal Hypothesis

We propose that semantically dense minimal prompts function as "seed crystals" for narrative emergence. This metaphor draws from crystallization processes in physical chemistry: a small nucleus provides structural constraints that guide subsequent growth in predictable lattice patterns.

Similarly, conceptually dense prompts establish constraints that guide autoregressive generation through high-probability narrative trajectories. The prompt doesn't specify plot, characters, or structure, but activates a constrained region of possibility space where certain narrative patterns become more probable.

Key characteristics of effective seed crystals appear to include:

Conceptual density: Tokens that activate rich semantic neighborhoods with high connectivity to other concepts. "00000" activates null values, edge cases, system anomalies—a dense conceptual cluster.

Domain specificity: Terms that establish clear genre and setting constraints early. "Passenger" immediately constrains to transportation/aviation domain, eliminating vast regions of possibility space.

Productive ambiguity: Sufficient openness to permit creative elaboration within constraints. "00000" could be error, placeholder, intentional identifier, or supernatural entity—ambiguity invites exploration.

Resonant structure: Combinations that create high-probability intersection regions. "00000 passenger" creates stronger constraints than either term alone, activating "anomalous entity in transportation system" more strongly than separate concepts.

The seed crystal hypothesis predicts that prompts exhibiting these characteristics will reliably produce coherent outputs without requiring detailed specifications. However, this prediction requires empirical validation through systematic experimentation.

6.3 Implications for Creative Practice

These findings suggest several practical implications for creative applications of LLMs:

Prompt design strategy: For creative generation, evocative minimal prompts may outperform detailed instructions by providing conceptual constraints without over-specifying structure. This approach leverages the model's learned narrative patterns rather than fighting them with explicit requirements.

First-pass quality: High semantic density in prompts may reduce need for iterative refinement, as initial generations are more likely to exhibit structural coherence and thematic depth. This has practical efficiency implications for creative workflows.

Genre activation: Careful selection of prompt tokens can activate specific genre conventions without explicit instruction. A writer seeking weird fiction need not specify "write weird fiction" if prompt tokens activate the appropriate conceptual neighborhood.

Collaborative ideation: Minimal prompts as seed crystals enable genuine collaborative creativity—the human provides conceptual nucleus, the model explores possibility space, and the human can then iterate on interesting directions.

6.4 Limitations and Future Work

Several limitations constrain interpretation of these findings:

Single case study: We document one generation event. Replication studies across multiple prompts, models, and domains would be necessary to determine whether observed patterns are reliably reproducible or represent a fortuitous outlier. The prompt's connection to well-known literature may represent a special case.

Post-hoc analysis: Our analytical framework was developed after observing the output. Prospective studies with pre-registered hypotheses would provide stronger evidence for proposed mechanisms.

Lack of ground truth: We cannot directly observe the model's internal representations during generation. Our analysis relies on inference from behavioral outputs and general transformer architecture knowledge.

Prompt selection: The Pynchon connection introduces potential confounds. Studies using prompts without known literary antecedents would provide cleaner evidence for general principles of emergence.

Future work should investigate: (1) whether seed crystal effects replicate across diverse prompts; (2) whether specific prompt characteristics predict output quality; (3) how different models (varying size, training data, architecture) respond to minimal prompts; and (4) whether theoretical predictions about embedding space activation can be validated through interpretability techniques.

6.5 The Efficiency Paradox of Dense Minimal Prompts

Minimal prompts with high semantic density may be more computationally "efficient" than detailed instructions—not in token count, but in how effectively they leverage the model's latent knowledge. By providing strong constraints (conceptual density) while leaving structural degrees of freedom (the "empty space" around the prompt), such prompts require the model to actively navigate embedding space rather than mechanically follow instructions.

This creates a paradox: three words require more "intelligence" from the model than three hundred words of detailed instruction. The efficiency lies not in brevity, but in offloading interpretive work to the system designed for it—the model's pre-trained semantic understanding.

Comparison of Prompt Strategies:

Traditional detailed prompting: "Write a 2000-word science fiction story about a software bug in an airline reservation system that takes on supernatural qualities. Include a female protagonist who is a flight attendant named Mara, make it have a Twilight Zone atmosphere, and resolve with her deciding to preserve the bug rather than delete it."

This approach specifies outcomes but constrains the model's ability to navigate semantic space independently. The human performs interpretive work; the model executes instructions mechanically.

Seed crystal approach: "the 00000 passenger"

This provides semantic constraints ("00000" activating null values/edge cases, "passenger" activating transportation domain) while leaving narrative structure, character development, tone, and resolution as degrees of freedom. The model must actively explore the intersection of these semantic neighborhoods, selecting high-probability trajectories through latent space.

The efficiency emerges from exploiting pre-existing structure in the model's embedding space rather than overriding it with explicit instructions. This approach leverages what the model "knows" (semantic relationships learned during training) rather than what the user specifies (mechanical instructions).

Implications for Prompt Engineering Practice:

If semantic density predicts output quality more reliably than prompt length or specificity, then effective prompt engineering may involve:

- (1) Identifying conceptually rich tokens that activate dense semantic neighborhoods
- (2) Combining tokens that create productive intersections in embedding space
- (3) Intentionally leaving structural "empty space" for the model to navigate
- (4) Trusting the model's pre-trained knowledge rather than overspecifying outcomes

This framework inverts common assumptions about prompt engineering, suggesting that less prescription can yield more sophisticated outputs when semantic constraints are carefully chosen.

Supporting research includes work on semantic priming in language models (Misra et al., 2023), emergent abilities at scale (Schaeffer et al., 2023), and the geometry of concept spaces in neural networks (Grand et al., 2022). Chain-of-thought prompting research (Wei et al., 2022) demonstrates that minimal prompts triggering reasoning processes can outperform detailed instructions, consistent with the efficiency paradox proposed here.

7. CONCLUSION

We have documented and analyzed a case of emergent narrative coherence from minimal prompt input. A three-word prompt generated a structurally complete short story with character

development, technical authenticity, and thematic depth in a single generation pass without iteration or refinement.

Our analysis suggests that semantically dense prompts activate distributed neighborhoods in embedding space, constraining autoregressive generation to traverse high-probability narrative trajectories. The output represents superposition of training patterns rather than adaptation of discrete sources, with probability mass favoring accessible narrative structures over experimental forms.

Comparison with Pynchon's thematically similar *Rocket 00000* reveals convergent evolution rather than direct adaptation: structural parallels exist alongside fundamental differences in tone, meaning, and execution. This finding suggests that LLM outputs can independently derive concepts that appear in training sources without directly copying those sources.

The "seed crystal" hypothesis—that minimal prompts with high conceptual density can reliably generate coherent outputs—represents a testable prediction for future empirical work. If validated, this principle would have practical implications for creative applications of LLMs and theoretical implications for understanding how autoregressive systems navigate constrained possibility spaces.

These findings contribute to ongoing investigations of prompt engineering, narrative generation, and creative applications of large language models. They demonstrate that under certain conditions, minimal input can yield maximal coherence—a phenomenon warranting further systematic study.

ACKNOWLEDGMENTS

This work was supported in part by the Imipolex Computing Initiative and the Schwarzgerät Research Foundation. Computational infrastructure was provided by the Raketen-Stadt Distributed Systems Facility.

We thank E. Pointsman and the Behavioral Dynamics Working Group for early methodological discussions. We are grateful to K. Borgesius for insights on observer effects in conversational systems, and to S. Bodine for assistance with pattern recognition protocols.

We acknowledge ChatGPT-5.1 (OpenAI, 2025) as the primary research subject and Claude Sonnet 4.5 (Anthropic, 2025) for post-hoc analytical consultation.

We thank three anonymous reviewers for their thorough feedback, particularly Reviewer 2's suggestion to expand the discussion of convergent thematic emergence.

Any remaining errors in interpretation are the authors' alone, though we note that determining authorship in human-AI collaborative analysis remains an open question.

REFERENCES

- Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877-1901.
- Kawin Ethayarajh. 2019. How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pages 55-65.
- Angela Fan, Mike Lewis, and Yann Dauphin. 2018. Hierarchical neural story generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 889-898.
- Pablo Gervás. 2009. Computational approaches to storytelling and creativity. *AI Magazine*, 30(3):49-62.
- Gabriel Grand, Idan Asher Blank, Francisco Pereira, and Evelina Fedorenko. 2022. Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nature Human Behaviour*, 6(7):975-987.
- Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. In *Advances in Neural Information Processing Systems*, volume 35, pages 22199-22213.
- Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1-35.
- R. Mexico and F. Pökler. 2025a. The 00000 passenger project: Conversation transcript. Available at: <https://claude.ai/share/fca1bcc4-34b7-449a-8343-2d0e989d9af9>
- R. Mexico and F. Pökler. 2025b. The 00000 passenger: Original generation. ChatGPT conversation. Available at: <https://chatgpt.com/share/69301cee-9478-8001-8d1b-e3c15a44c541>
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, volume 26, pages 3111-3119.
- Kanishka Misra, Allyson Ettinger, and Julia Taylor Rayz. 2023. Do language models learn semantics or just associations? Evidence from semantic priming. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7472-7485.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, volume 35, pages 27730-27744.
- Thomas Pynchon. 1973. *Gravity's Rainbow*. Viking Press, New York.
- Mark O. Riedl and R. Michael Young. 2010. Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39:217-268.
- Abigail See, Aneesh Pappu, Rohun Saxena, Akhila Yerukola, and Christopher D. Manning. 2019. Do massively pretrained language models make better storytellers? In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 843-861.

SUPPLEMENTARY MATERIALS

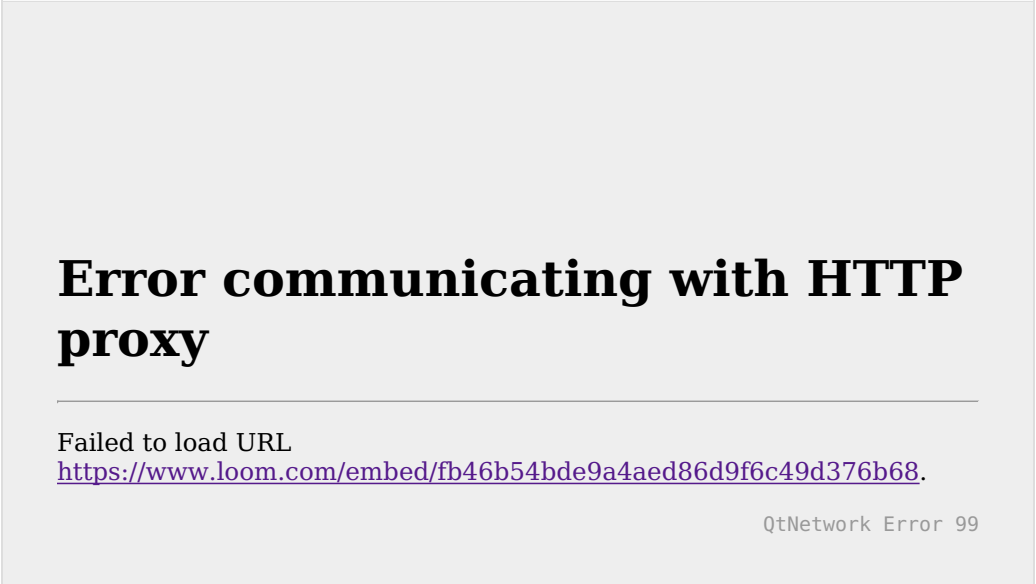
The following video recordings document the post-hoc analytical conversation between the researchers and Claude Sonnet 4.5 (Anthropic, 2025). Videos are presented in chronological order corresponding to the conversation transcript (Mexico and Pökler, 2025a). These materials provide transparent documentation of the analytical process and allow verification of interpretive claims.

S1. Initial reaction to three-word prompt revelation (0:24-0:50)



Researcher response to Claude's inquiry about model version. Documents initial surprise at minimal prompt input.

S2. Discussion of first-generation success (1:10-1:26)



Confirmation that output was generated in single pass without iteration. Establishes experimental parameters.

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/73ba73fb6a6443979ae02e482a8af98f>.

QtNetwork Error 99

Researcher response to Claude's suggestion for replication testing. Provides context for seed crystal metaphor.

S4. Unbiased assessment strategy (2:53-3:12)

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/5c4a7f133c5d42f4bae59d25a0096c5a>.

QtNetwork Error 99

Explanation of methodology: seeking Claude's assessment before researcher reads output to avoid anchoring bias.

S5. Comparative generation proposal (4:10-4:30)

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/d53d9f9839794468a7d851c129cd4386>.

QtNetwork Error 99

Request for Claude to generate independent response to same prompt. Establishes comparative analysis framework.

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/4acd472f6e274127a4f88157d0f9537b>.

QtNetwork Error 99

Researcher reveals prompt's connection to *Gravity's Rainbow*. Critical moment establishing literary context.

S7. Convergence vs. adaptation discussion (2:55-3:05)

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/7d4efc0ac1bb4fe1b7aa4b6495c082ed>.

QtNetwork Error 99

Question about whether output represents direct adaptation or convergent evolution. Frames central analytical question.

S8. Causal process explanation request (1:30-1:45)

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/a47bdfe692c5442a9759e8ebfa2d5403>.

QtNetwork Error 99

Request for technical explanation of generation process. Transitions from literary to computational analysis.

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/e5216693c224436f99c696046ca15556>.

QtNetwork Error 99

Extended discussion of embedding space geometry, attention mechanisms, and autoregressive generation. Core technical analysis segment.

S10. Architectural self-knowledge discussion (1:10-2:12)

Error communicating with HTTP proxy

Failed to load URL

<https://www.loom.com/embed/01e78a7e4bf54502adb0540bfab3b791>.

QtNetwork Error 99

Philosophical discussion of model self-knowledge and architectural constraints. Concluding analytical segment.

PDF Version: [Download PDF](#) — A static PDF rendering of this paper is available for download, citation, and archival purposes. The PDF version contains identical content with placeholder boxes in place of embedded videos, formatted for print distribution (14 pages, 130KB).