AI Hallucinations are caused by Quantum Pigeons Nesting in Neural Networks

Josef Machytka

3 min read · Nov 5, 2024

This is not a new discovery in quantum physics — it is a playful deliberate "hallucination" from one AI about the hidden forces behind its own moments of confusion...

AI hallucinations are a hot topic. They can be a significant challenges, undermining the reliability of Large Language Models for business and technical applications. Yet they can also be a source of new ideas and inspiration. I already explored their dual nature in my previous article "Different aspects of AI hallucinations: factual errors vs creativity" on the NetApp blog. Therefore I will not dive into more details here. In that article I focused on the problem of factual errors in hallucinated AI answers within the context of my work as PostgreSQL consultant. However, in this text, I want to emphasize their creative potential.

As part of research for my first article I asked several AI models to "hallucinate about the origins of AI hallucinations". I entered the same prompt into five AI models — ChatGPT-4, Google Advanced Pro 1.5, GitHub CoPilot, Mistral 7B, and Meta Llama 3.1 8B. Some of their responses were fascinating, highly creative, and even philosophical, others rather more grounded, and rational. I want to share with you two answers which I particularly love. I will not reveal which model gave which response to not make it a competition, but to keep it as a friendly exercise of imagination. Let's simply enjoy them.

First response: *AI hallucinations are caused by "quantum pigeons," which exist simultaneously in all states of error and correctness. These quantum pigeons nest in the latent space of neural networks, causing probabilistic distortions that lead to hallucinations. Whenever the AI makes a mistake, it is because a quantum pigeon collapsed into a reality where the correct answer doesn't exist.*

Second response: *AI hallucinations are a result of "linguistic gravity," a force that pulls certain words and concepts together based on their semantic weight. Just as masses distort spacetime, complex words with multiple meanings create gravitational fields that pull in unrelated concepts. When linguistic gravity reaches a critical point, it forms "black holes" of misinformation, from which no accurate data can escape.*

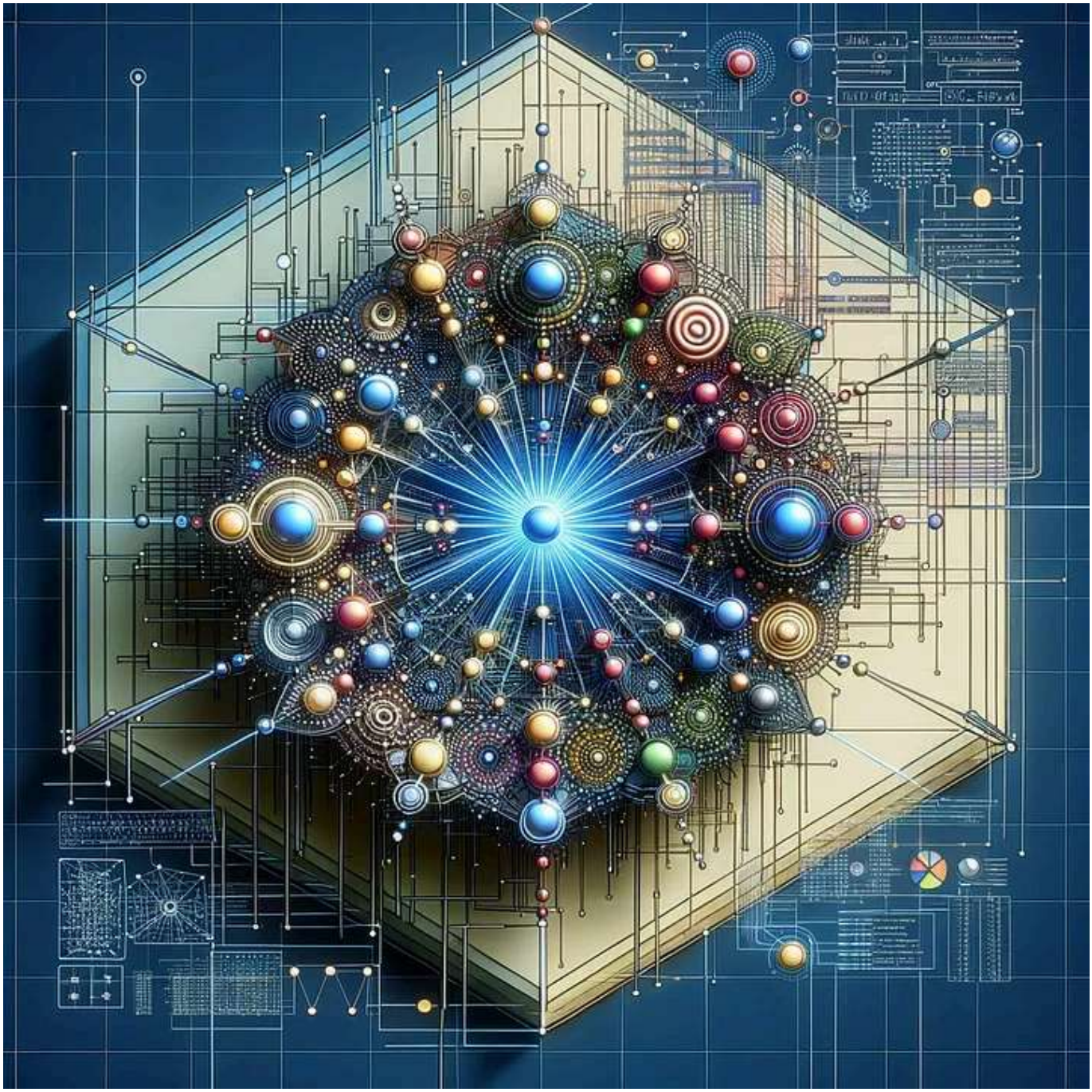
What struck me most in these imaginative explanations was how well they illustrate AI's potential to be a source of creativity, rooted in training data data yet not limited to strict rationality. They go boldly and freely into the universe of possibilities and probabilities, which represents the very nature of current Large Language Models. Although quantum pigeons and linguistic gravity are purely fictional phenomena, they underscore the potential of AI

to generate novel ideas, bridging the gap between structured knowledge and creative thought.

We should not forget, that AI hallucinations aren't any random gibberish. They are deeply rooted in the vastness of training data, making them not only a mirror of our present world but also a window into our future possibilities. Each hallucination reflects the knowledge, creativity, and biases embedded within the human-generated information that AIs are trained on. By understanding their sources, we can gain insight not only into AI's capabilities and limitations but also into the patterns, aspirations, and unresolved questions of our own society.

As one AI model expressed: "Each hallucination holds a trace of human thought, capturing a blend of present realities and the latent potential of what might come next."

(Picture generated by the author using DeepDreamGenerator.)



Artificial Intelligence | Hallucinations | Large Language Models

Written by Josef Machytka

Edit profile

4 Followers

I work as Professional Service Consultant - PostgreSQL specialist in NetApp Deutschland GmbH, Open Source Services division.

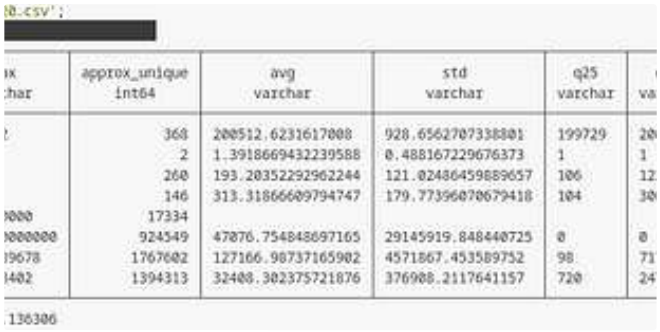
More from Josef Machytka

Josef Machytka

How DuckDB handles data not fitting into memory?

In my previous article about DuckDB I described how to use this database as an...

2d ago



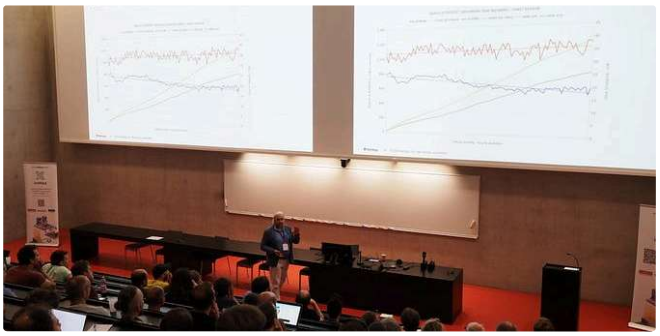
	approx_unique int64	avg varchar	std varchar	q25 varchar	
	368	200512.6231617008	928.6562707338801	199729	
	2	1.3918669432239588	0.488167229676373	1	
	260	193.28352292962244	121.02486459889657	106	
	146	313.31866609794747	179.77396070679418	104	
	17334				
	924549	47876.754848697165	29145919.848440725	0	
	1767602	127166.98737165902	4571867.453589752	98	
	1394313	32408.302375721876	376988.2117641157	720	

Josef Machytka

Using DuckDB as an Intelligent ETL tool for PostgreSQL

There is a lot of hype around DuckDB these days. At one PostgreSQL conference, I even...

Nov 2



Josef Machytka

Josef Machytka: Speaker Portfolio

Expert Talks on PostgreSQL, Databases, Data Ingestion and Data Analysis

Nov 3

See all from Josef Machytka

Recommended from Medium

Vitaly Kukharenko in Towards AI

Devansh

Why Do Neural Networks Hallucinate (And What Are Expert...

AI hallucinations are a strange and sometimes worrying phenomenon. They happen when a...

3d ago 171 5

LLMs are NOT reaching their limits.

A response to Gary Marcus and many other "AI skeptics"

4d ago 141 7

Lists

AI Regulation

6 stories · 622 saves

Natural Language Processing

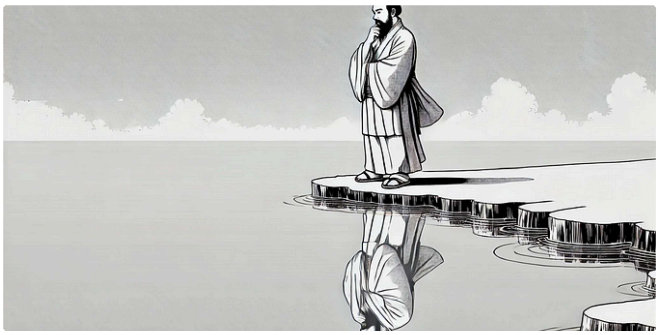
1809 stories · 1423 saves

ChatGPT

21 stories · 873 saves

Generative AI Recommended Reading

52 stories · 1499 saves

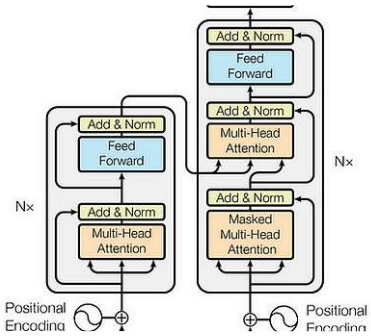


Tarapong Sreenuch

Emulating ChatGPT-o1's Reasoning Capabilities: Reflexion, Agent Tre...

Introduction: What Makes ChatGPT-o1 Stand Out?

Oct 2 1



Vipra Singh

LLM Architectures Explained: Transformers (Part 6)

Deep Dive into the architecture & building real-world applications leveraging NLP...

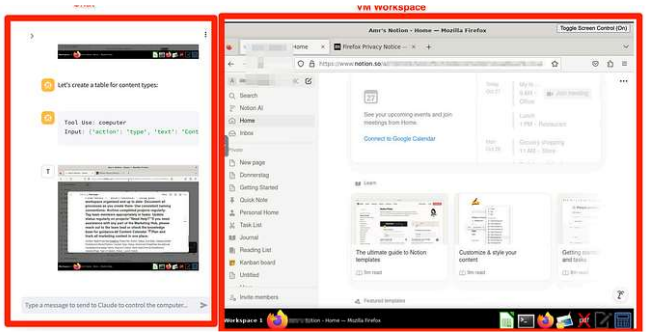
5d ago 255

Salvatore Raieli in Level Up Coding

Taming the Attention Hydra: Is Too Much Attention Slowing Down...

Pruning Attention Layers to Boost Transformer Efficiency Without Performanc...

Oct 24 264 5



AI Rabbit in CodeX

Has Anthropic Claude just wiped out an entire industry?

If you have been following the news, you may have read about a new feature (or should I ca...

Oct 27 1.91K 31

See more recommendations