

Core C++ 2025

19 Oct. 2025 :: Tel-Aviv

Welcome to the museum of AI Hallucinations

Maria Piterberg

AI Hallucinations

A curse or a blessing?

Hallucination noun

*The fact of seeming to see or hear somebody/
something that is not really there, especially
because of illness or drugs.*

from Oxford Dictionary

The AI Meaning

When an LLM is said to “*hallucinate*,” it means it has produced **false** or **misleading** information.

When a Hallucination Goes Too Far

Imagine you're planning a romantic evening in a city you've never visited.

You ask ChatGPT to recommend a restaurant and choose from a promising list of venues.

But when you arrive - all dressed up - you discover
the restaurant doesn't exist in the specified location.

Moreover, it does not exist in any location.

In fact, it never did.

Annoying? Definitely.

But it's a relatively harmless example.

The Pentagon That Wasn't

Two years ago, a verified Twitter account posted an image of an explosion next to a building shaped like the Pentagon.



Here is the pickle - the image was **fake**,
But the damage had already been done.

When Bots Make Up Policy

More recently, Cursor's AI support bot informed users that: "the tool could no longer be used on more than one machine."

Only, **it wasn't true.**

Cursor's CEO later clarified:

"We have no such policy. You're of course free to use Cursor on multiple machines."

More Damaging Scenarios

- Medical misinformation → Incorrect diagnosis
- Historical inaccuracies → Rewriting history
- Code errors → Systems crash
- Flawed predictions → Driving malfunctions

AI hallucinations aren't just fictional side-effects;
they have real-world **implications**.

These examples reinforce one conclusion:

AI cannot be trusted blindly.

Yes, It Can be Funny

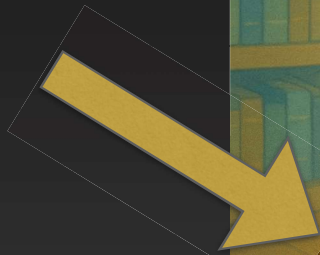
The next images were generated during the creation of

“I am a Pancake!”

My children book that teaches about AI.







All these are actual examples of DALL-E 3 hallucinations.

They are funny but also could be really frustrating.

So Why is AI “Lying”?

- **Pattern-matching** systems
- Generate plausible-sounding information
- Training bias
- Over fitting
- Model complexity

Model's "Train of Thought"

- **Not fact retrieval** → they don't "look up" truth in a database
- **Statistical engines** → trained on billions of examples of text & images
- **Next-token prediction** → guess the most likely next word or pixel based on patterns in data
- Output = plausible, not **guaranteed true**

LLMs don't "understand" - they **correlate patterns**.

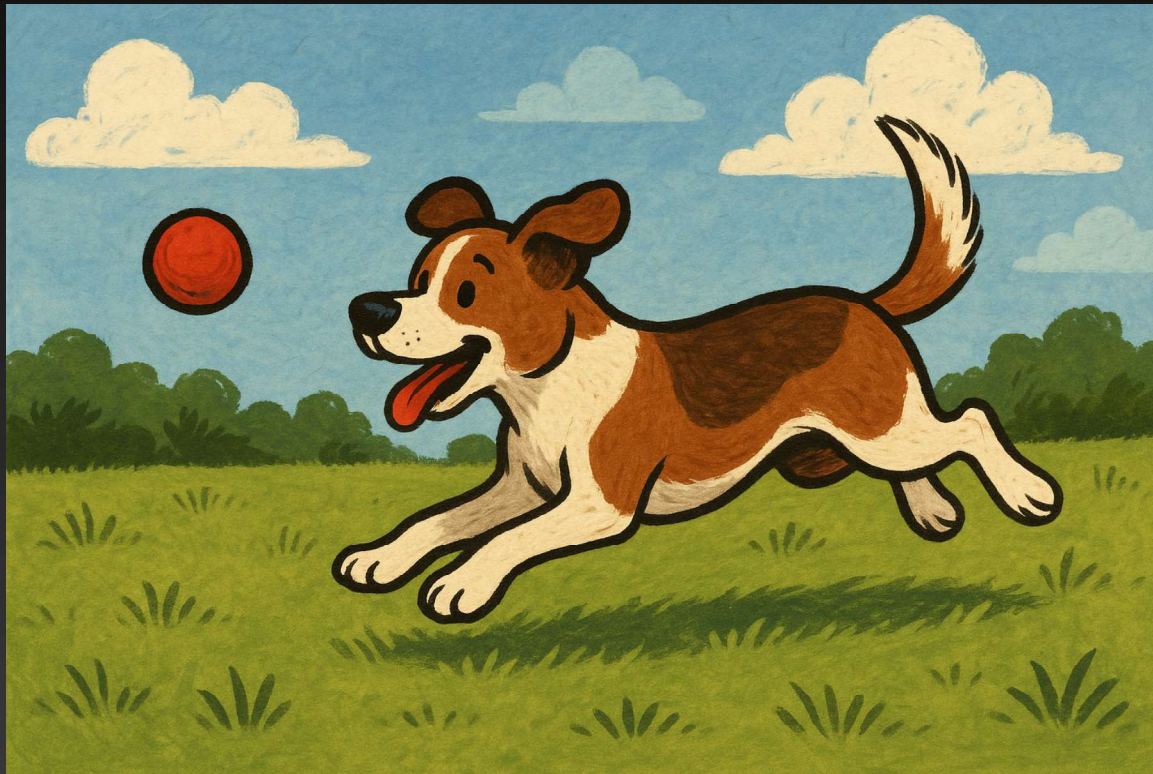
Human Minds

Castle → Princess



Human Minds

Dog → Ball



Human Minds

Toddler → Teddy Bear



The Categorisation Trap: Your Brain Loves a Story

Mushroom

Table

Moon

Tea

Star

Cake

Tree

Fox

Cloud

Mushroom

Table

Moon

Tea

Star

Cake

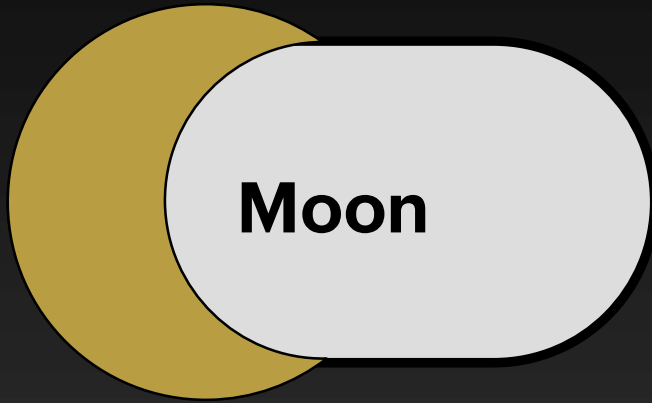
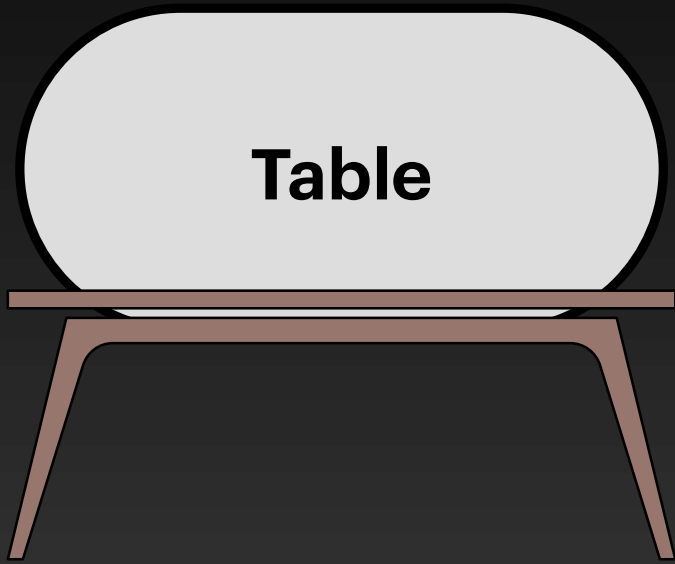
Tree

Fox

Cloud

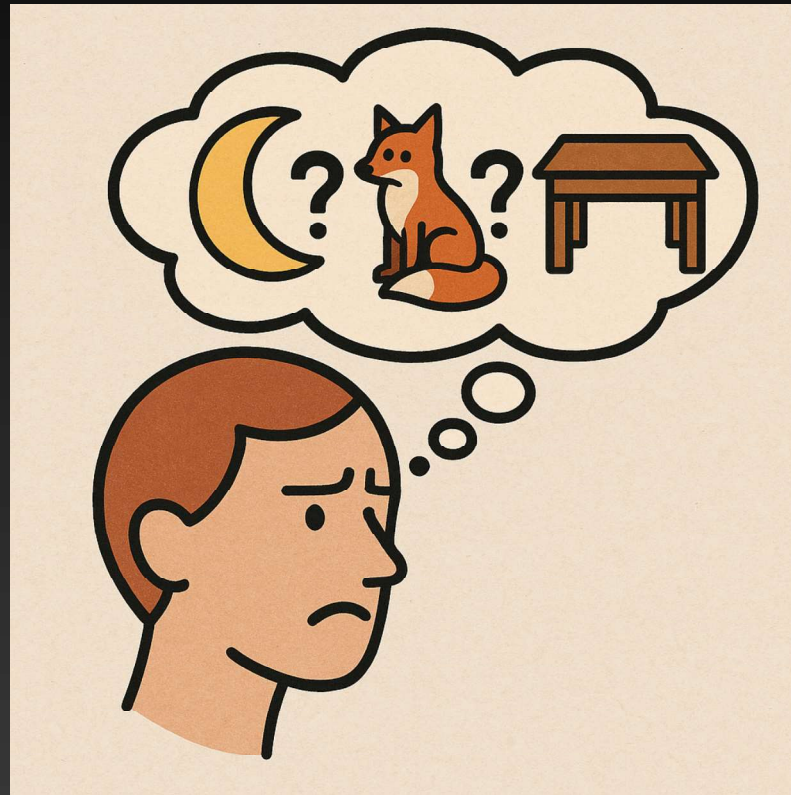
Context locks us into a framework

Let's *break* the Context!



Can you imagine a moon, fox, and table together?

Yep. Your brain just stalled, didn't it?



But guess who doesn't struggle?

Generative AI.



A fox having tea with the moon. Generated by DALL-E 3.



Stars Celebrating Birthday with a Mushroom Cake. Generated by DALL-E 3.



Tree Drinking Tea inside a Cloud. Generated by DALL-E 3.

The “Human’s Brain Way”

Humans tend to reason based on meaning, coherence, and lived experience. When asked to imagine a scene involving a moon, a fox, and a table, our brains instinctively try to make sense of the combination.

We want a narrative.

Human Cognition: Coherence, Context & Meaning

- Builds **models of the world**
- **Constrained by coherence** - “make sense”
- **Contextual constraints**

Making Lemonade

Could we exploit AI's ability to look at the world differently and leverage this so called "negative trait" for our advantage?

The Possibilities are Limitless

- Visual Art & Design
- Music & Sound
- Writing
- Games & Interactive media
- Architecture & Product design

The line between error and inspiration
is thinner than we think

Thank you for visiting the Museum of AI Hallucinations.



Maria.piterberg@gmail.com