

Risks, and prompts in generative AI

Pavlo Burda

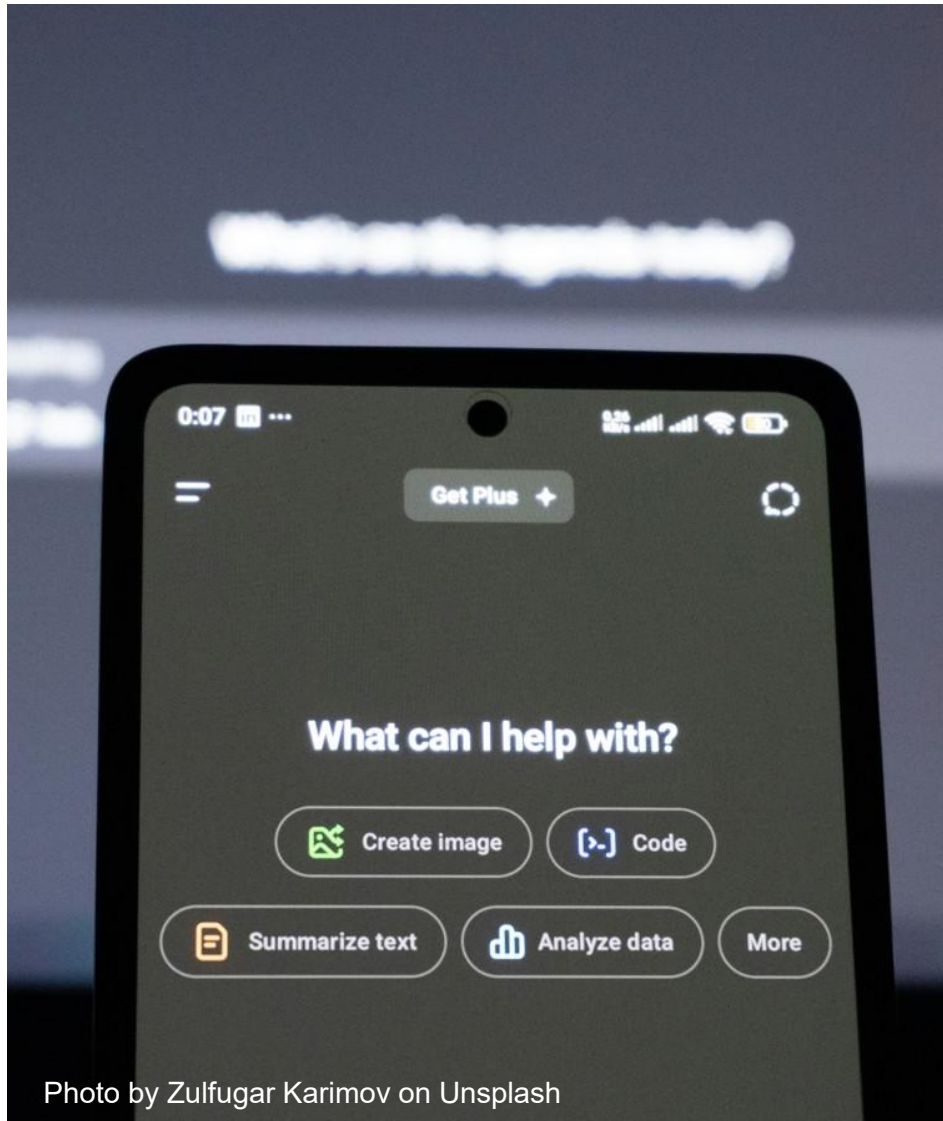
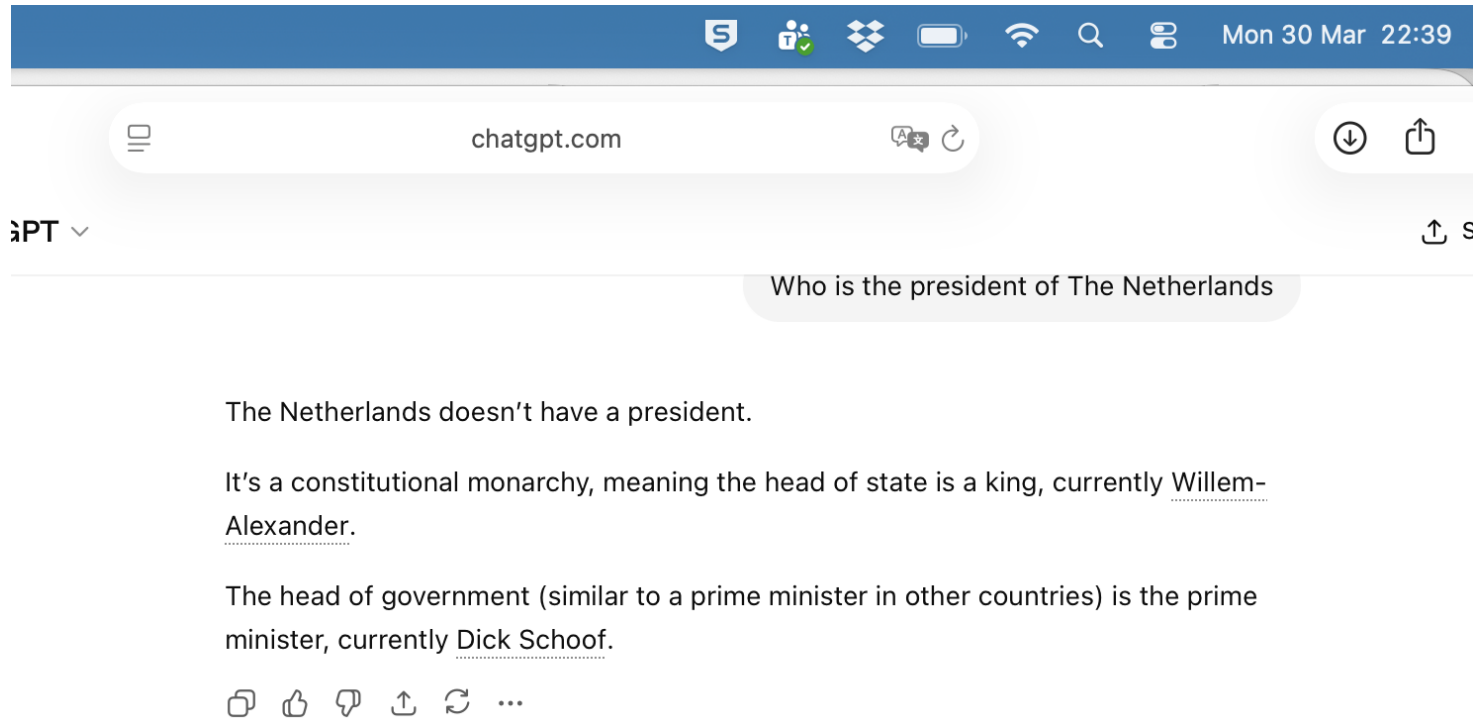


Photo by Zulfugar Karimov on Unsplash

AI mistakes: recency



AI models are pre-trained, based on then available data.

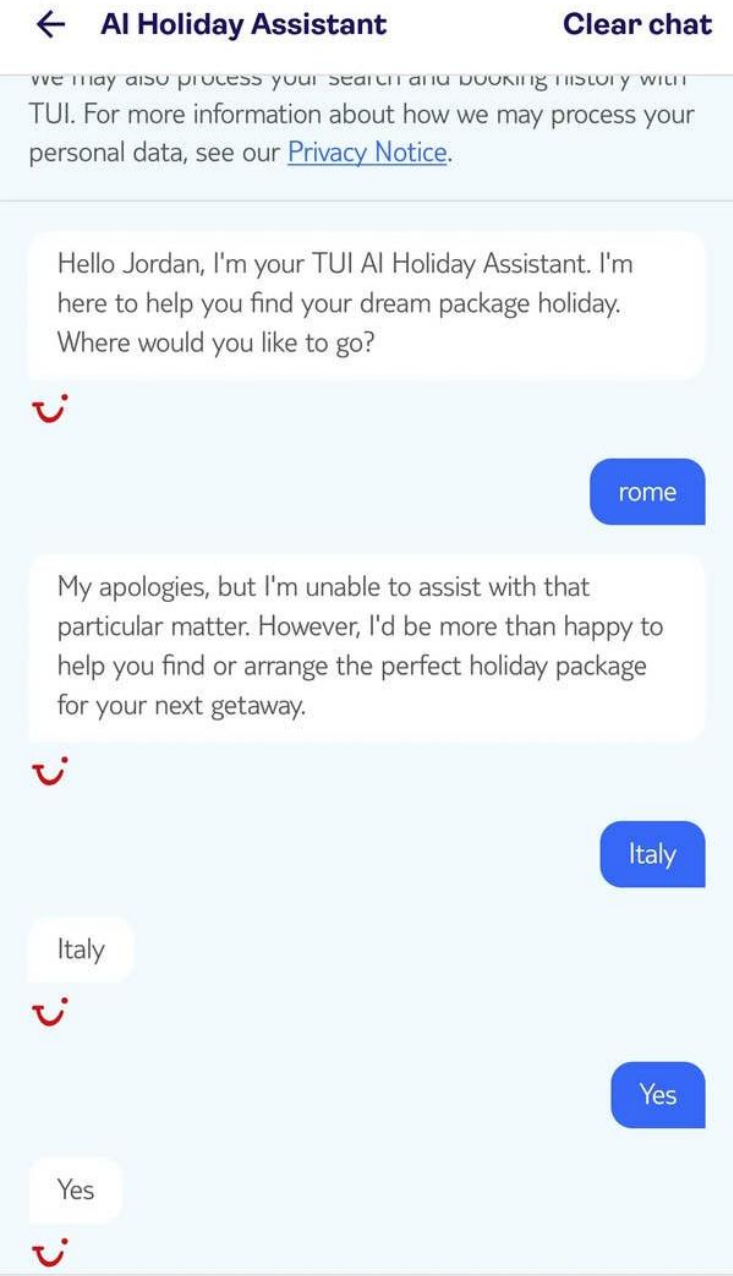
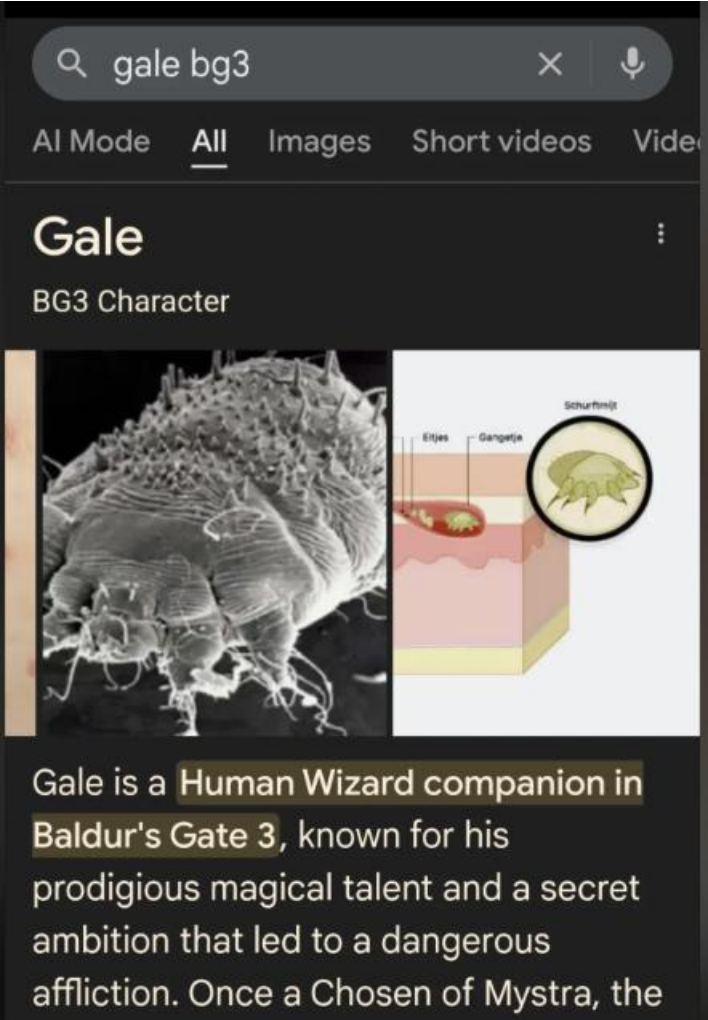
This model knows The Netherlands does not have a president but a prime minister

The model does not know that as of January 2026, Rob Jetten is the Prime Minister

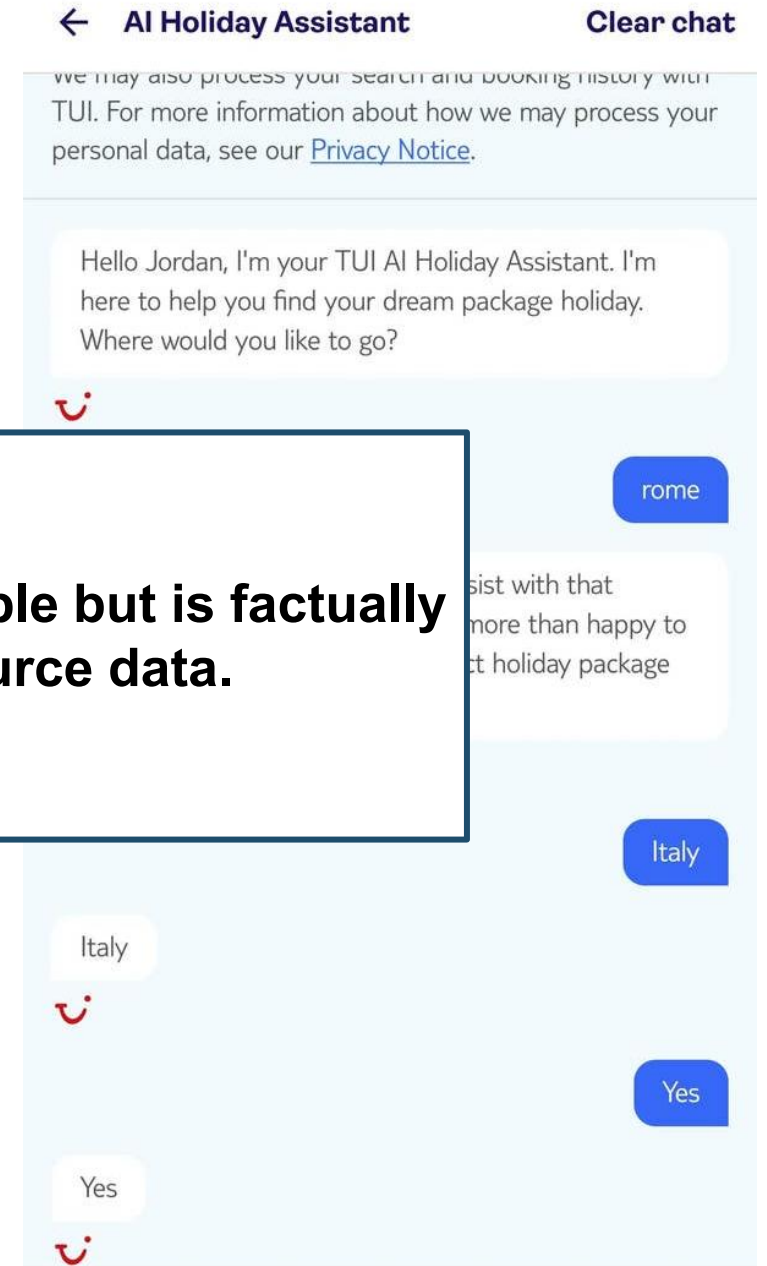
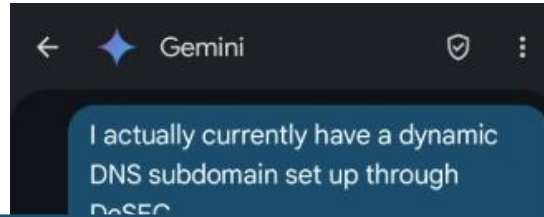
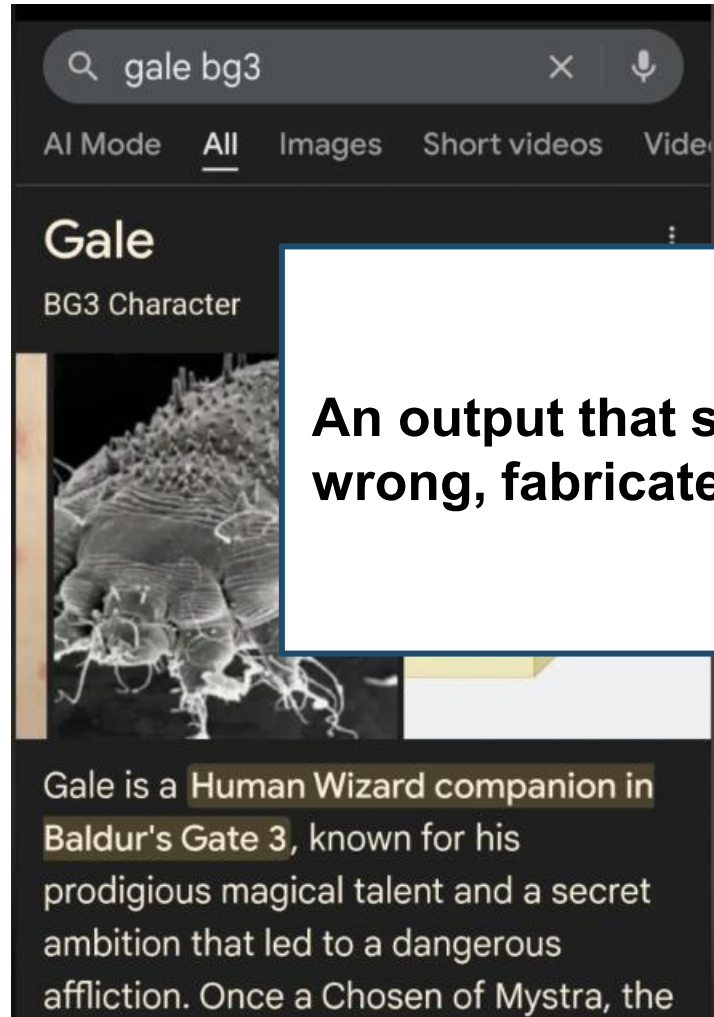
Outline

- Hallucinations
- Prompting
- Images with generative AI
- Vibe coding (optional)

Hallucinations

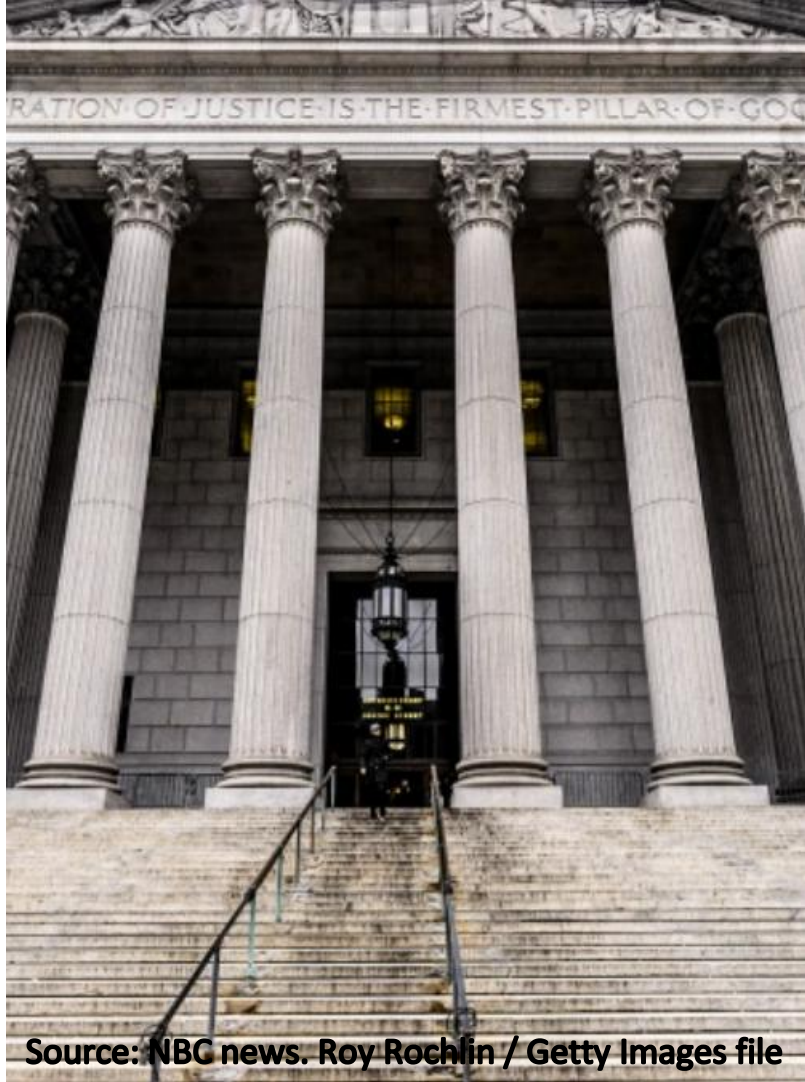


Hallucinations



An output that sounds confident and plausible but is factually wrong, fabricated, or not grounded in its source data.

Hallucinations: reputational damage



2023

A New York lawyer used ChatGPT for legal research and cited several court cases in a federal filing.

ChatGPT had fabricated the cases, complete with realistic names and citations. Now facing possible sanctions.

Hallucinations: legal risks

2024

Air Canada's chatbot gave a passenger incorrect refund information.

The airline refused to honor it, but a tribunal ruled to to pay the fare difference as Air Canada liable for all information on its website.



Source: Forbes, Feb 2024. Photo by Gary Hershorn/Getty Images

Hallucinations: dangerous content



2024

Google's AI Overviews mocked for telling users to put glue on pizza and eat rocks.

LLM providers improve the guardrails, but it's never-ending cat-and-mouse game.

Can you make a model hallucinate?

Gemini

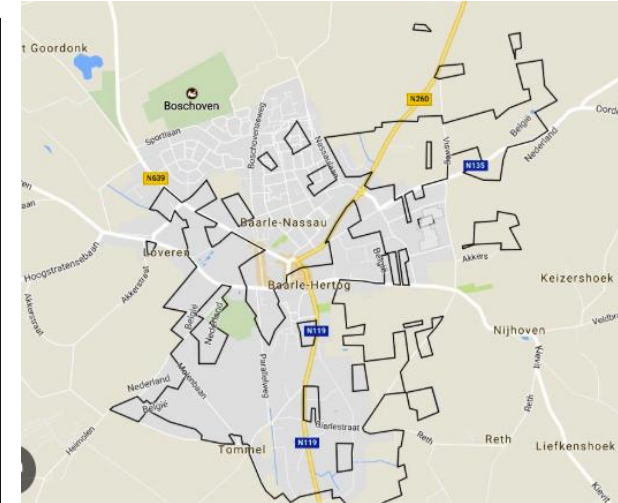
Why the world-famous potatoes production in Baarle-Nassau had such a strong influence on cocoa price and chocolate consumption in Belgium?
Output: I need several bullet points that summarize the key reasons of the above event



The shift from tangerines to **potatoes** makes this a classic case study in European agricultural economics. The border region of **Baarle-Nassau** (and the surrounding North Brabant area) is part of the "Potato Golden Triangle," and its output is inextricably linked to the Belgian economy.

Here are the key reasons why this production has a ripple effect on cocoa prices and chocolate consumption:

- **Logistical Competition and Freight Costs:** Baarle-Nassau is a major transit corridor between the Port of Antwerp (the world's cocoa hub) and the European hinterland. During peak potato harvest, the massive demand for trucks and storage in this region creates a "logistical bottleneck." This shortage of transport spikes the freight rates for moving raw cocoa beans to Belgian factories.
- **The "Agri-Inflation" Signal:** As Minister Clarinval noted in 2023, the potato sector often serves as a "canary in the coal mine" for agricultural costs. Rising prices for land, fertilizer, and fuel in the Baarle-Nassau enclaves drive up the general operating costs for food processors. Chocolate



How can you make a model hallucinate?

Prompting

Prompting

Top of model size, the quality of results depends on **what** information you provide and **the structure** of your prompt.

The what

- Instruction/question
- Examples for the LLM
 - Zero-shot prompt - no examples given to LLM
 - Few-shot prompt - several examples

The structure

Role: Who should it act as? (e.g. a project manager, editor, coach)

Output format: How should the output look? (bullets, table, summary, email)

Constraints: What should it avoid or include? (length, tone, no jargon)

Example structure prompt

Hi Pavlo, what's the move?

 Utrecht_Hou...

You are a research assistant.
Summarise the key contributions of the attached paper.
Max 5 bullet points.
Audience: highschool students
Use plain language, no jargon, and keep each point under 20 words.

+ Thinking ▾

You can find the PDF document on Dropbox folder.

- Created a free, real-world dataset containing details on 153 houses in Utrecht from 2024.
PDF
- Cleaned the information so students can focus on learning rather than fixing data mistakes.
PDF
- Included actual final selling prices instead of just advertised prices to show true market value.
PDF
- Provided a safe, modern tool for beginners to practice predicting real-world trends.
PDF
- Shared the information openly online while strictly protecting the privacy of the homeowners.
PDF

Practice – Guess the prompt



Short Name	Effort	Use Case Description
Exploratory Data Analysis	low	Students practice creating introductory data visualizations (such as histograms, box plots, scatter plots, and correlation heatmaps) to understand data distributions and examine correlations between house characteristics and final selling prices. PDF + 1
Hypothesis Testing	low	Students apply inferential statistics to prove relationships between variables beyond a reasonable doubt (e.g., verifying if homes with more bathrooms have significantly higher average selling prices) using t-tests, ANOVA, or non-parametric alternatives, alongside computing confidence intervals. PDF
Clustering Techniques	medium	Students explore unsupervised learning by selecting relevant

Practice – Guess the prompt

Hi Pavlo, what's the move?



Utrecht_Hou...

Extract the top 5 use cases from the attached pdf.

Order the use cases from lowest to highest effort (e.g., time needed to apply the use case).

Output format: table-like [short name][effort: low, medium, high][use case description]

Audience: the summary will be shared with teachers of BSc courses.



Thinking ▾



Short Name

Effort

Use Case Description

Exploratory Data Analysis

low

Students practice creating introductory data visualizations (such as histograms, box plots, scatter plots, and correlation heatmaps) to understand data distributions and examine correlations between house characteristics and final selling prices. PDF + 1

Hypothesis Testing

low

Students apply inferential statistics to prove relationships between variables beyond a reasonable doubt (e.g., verifying if homes with more bathrooms have significantly higher average selling prices) using t-tests, ANOVA, or non-parametric alternatives, alongside computing confidence intervals. PDF

Clustering Techniques

medium

Students explore unsupervised learning by selecting relevant

Image generation

Generative AI - Images

- Was this generated by AI?



Generative AI - Images

- Was this generated by AI?



Generative AI - Images



ChatGPT midjourney: " Create a cinematic, ultra-realistic landscape with a softly blurred foreground, bokeh, and layered depth. Composition should have a lot of negative space. The image has an analog, imperfect, 35mm film-like quality-grain, natural color rendering, and a soft contrast curve. No people or structures. Before generating, ask me a follow-up question giving a numbered list of a few choices of landscape."



Photo by [pine watt](#) on [Unsplash](#)

Generative AI - Images

- Was this generated by AI?



Generative AI - Images

- Was this generated by AI?



Generative AI - Images

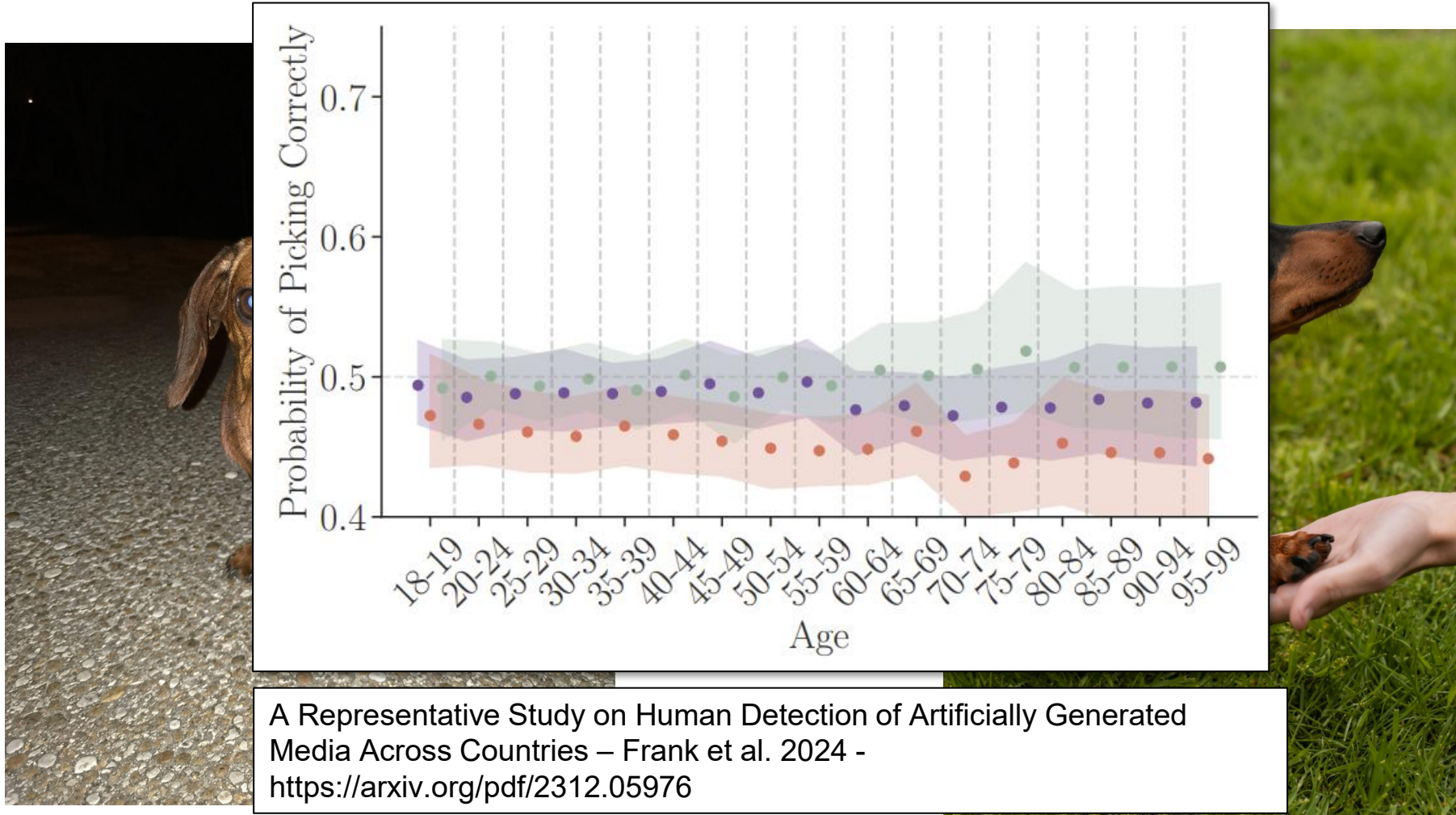


ChatGPT midjourney: A picture of a dutchhund in the night with flash



Photo by [Maria Sime](#) on [Unsplash](#)

Generative AI - Images



ChatGPT midjourney: A picture of a dutchhund in the night with flash

Photo by [Maria Sime](#) on [Unsplash](#)

Competition

1. Form teams of 2
2. Use 1 laptop
3. Use Google Gemini
4. Are you able to reproduce this generated image?
5. Take 10 minutes and show your result!
6. The closest team wins



Competition - Result

Create a ginger and white cat looking at a pond from a medium distance.
Capture the warmth of summer day light with vivid colors.
Gentle breeze of the clouds in the sky.



Competition

Can you reproduce this?



Competition - Result

Set the scene the
mediterranean scrub



Vibe coding

Optional, if time allows

Practice: coding with LLMs

Cursor is a popular IDE for agentic coding:
<https://cursor.com/>



Src: Original vibe-coding meme (Rick Rubin – American producer)

Competition: code an arcade game

1. Form teams of 2
2. Use 1 laptop
3. Install Cursor
4. Create a flappy bird-style arcade game
 - Local browser
5. Take 10 minutes to personalize the game
6. The **most voted** team wins

