



Facultatea de Electronică,
Telecomunicații și
Tehnologia Informației

SISTEME INTELIGENTE DE SUPORT DECIZIONAL

Ș.l.dr.ing. Laura-Nicoleta IVANCIU

Curs 10 – Generative AI - fundamente și studii de caz.

Cuprins

- Definiere, concepte, funcționare
- Scenarii de utilizare și limitări
- Large Language Models (LLM)
- Provocări și riscuri
- Gemini AI vs. ChatGPT
- CHOIS

Notă: materialele prezentate au ca bază cursul “Generative AI for everyone” by Andrew Ng, platforma Coursera, nov 2023

Generative AI

= artificial intelligence systems that can produce **high-quality** content, specifically text, images and audio

Exemple:

- Chat-GPT (OpenAI)
- Bard, Gemini AI (Google)
- Bing chat (Microsoft)

Pentru cine este destinat?

Generative AI

= artificial intelligence systems that can produce **high-quality** content, specifically text, images and audio

Exemple:

- Chat-GPT (OpenAI)
- Bard, Gemini AI (Google)
- Bing chat (Microsoft)

Pentru cine este destinat?

- *utilizatori cu niveluri de experiență diferite*

- *developers, business managers, etc*

- *studenți? cei care vor să acumuleze rapid cunoștințe într-un anumit domeniu?*

Poate Generative AI să înlocuiască procesul de documentare pentru un proiect, învățarea, profesorul, etc?

Generative AI

Utilizări:

- generare text

Exemple?

- generare imagini

- căutare web

- sisteme de recomandare

- verificare plăți electronice

- marketing și publicitate online

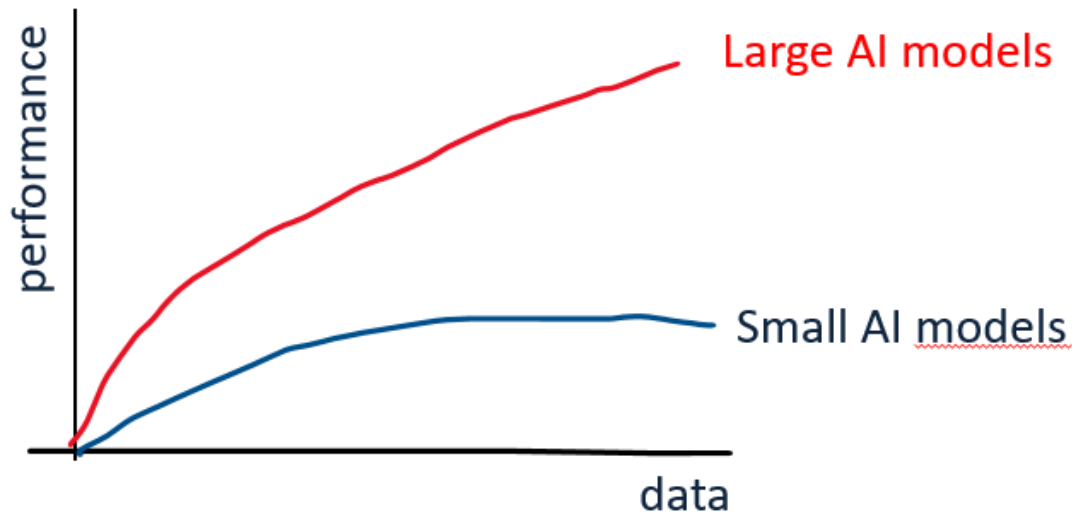


Imagine generată cu **Dall-e**

Prompt: a fluffy cat carrying a backpack full of treats, walking through a snowy scenery

Cum funcționează Generative AI?

- rezultat al aplicării învățării supervizate (supervised learning) pe modele de dimensiuni mari, cu seturi de date mari



Modele de dimensiuni mari – rețele cu număr mare (sute) de straturi, număr mare (milioane) de parametri antrenabili

large AI models + large labeled datasets = **large-scale supervised learning**

Cum funcționează Generative AI?

Large Language Models (LLM)

- modele de dimensiuni mari, care au fost antrenate pe seturi de date lingvistice (text) de dimensiuni mari (miliarde de cuvinte)

Ex:

Un sistem cu AI a citit de multe ori propoziția:

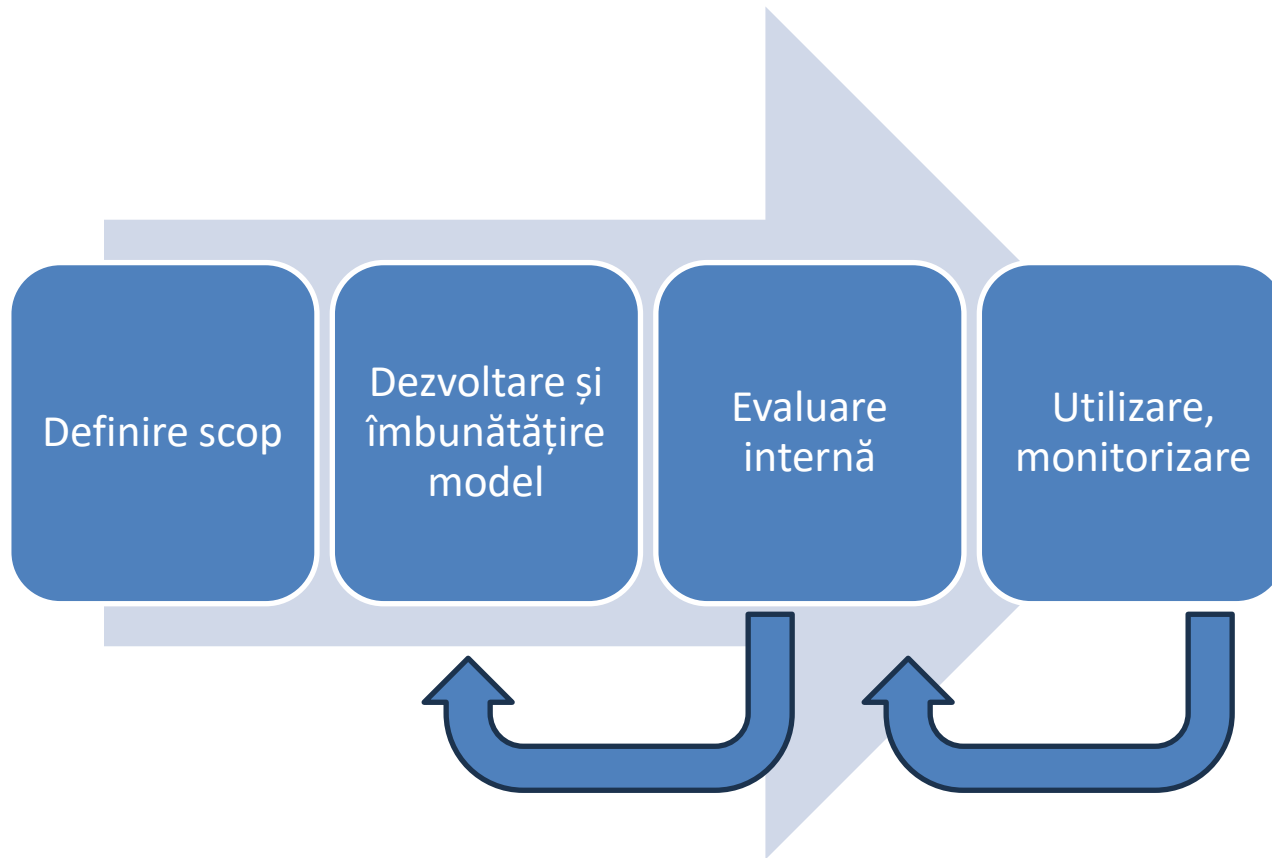
Desertul meu preferat este prăjitura cu cireșe.

Prompt: *Desertul meu preferat este _____*

Completare de la Generative AI: *prăjitura || cu || cireșe.*

Sistemul adaugă cuvânt cu cuvânt pentru completarea propoziției.

Dezvoltarea unui model cu Generative AI



Exemplu de evaluare internă?

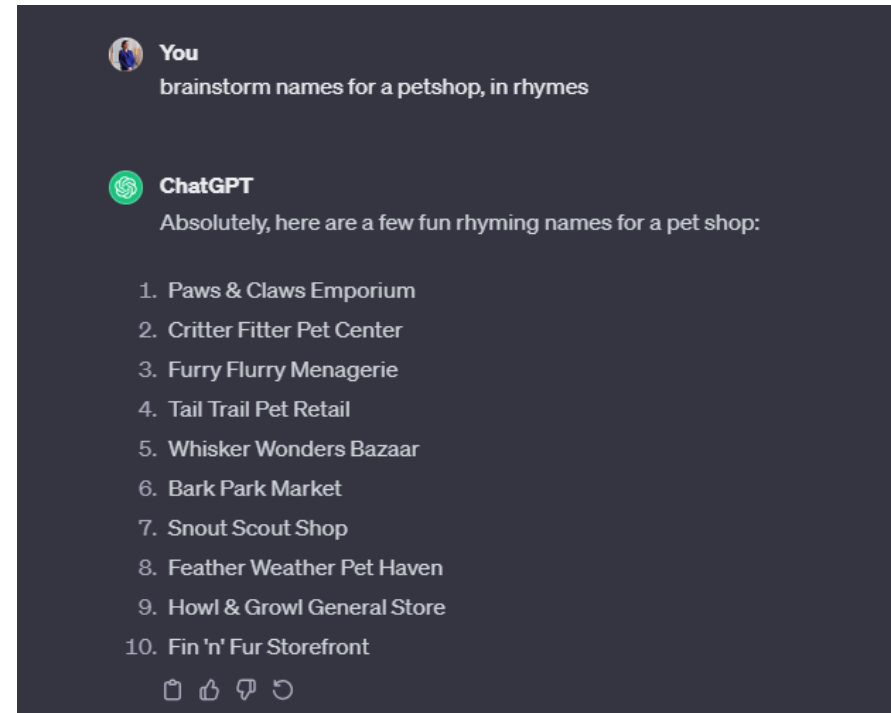
Utilizări ale Generative AI pentru generare de text

- scriere
- citire
- conversație

Utilizări ale Generative AI pentru generare de text

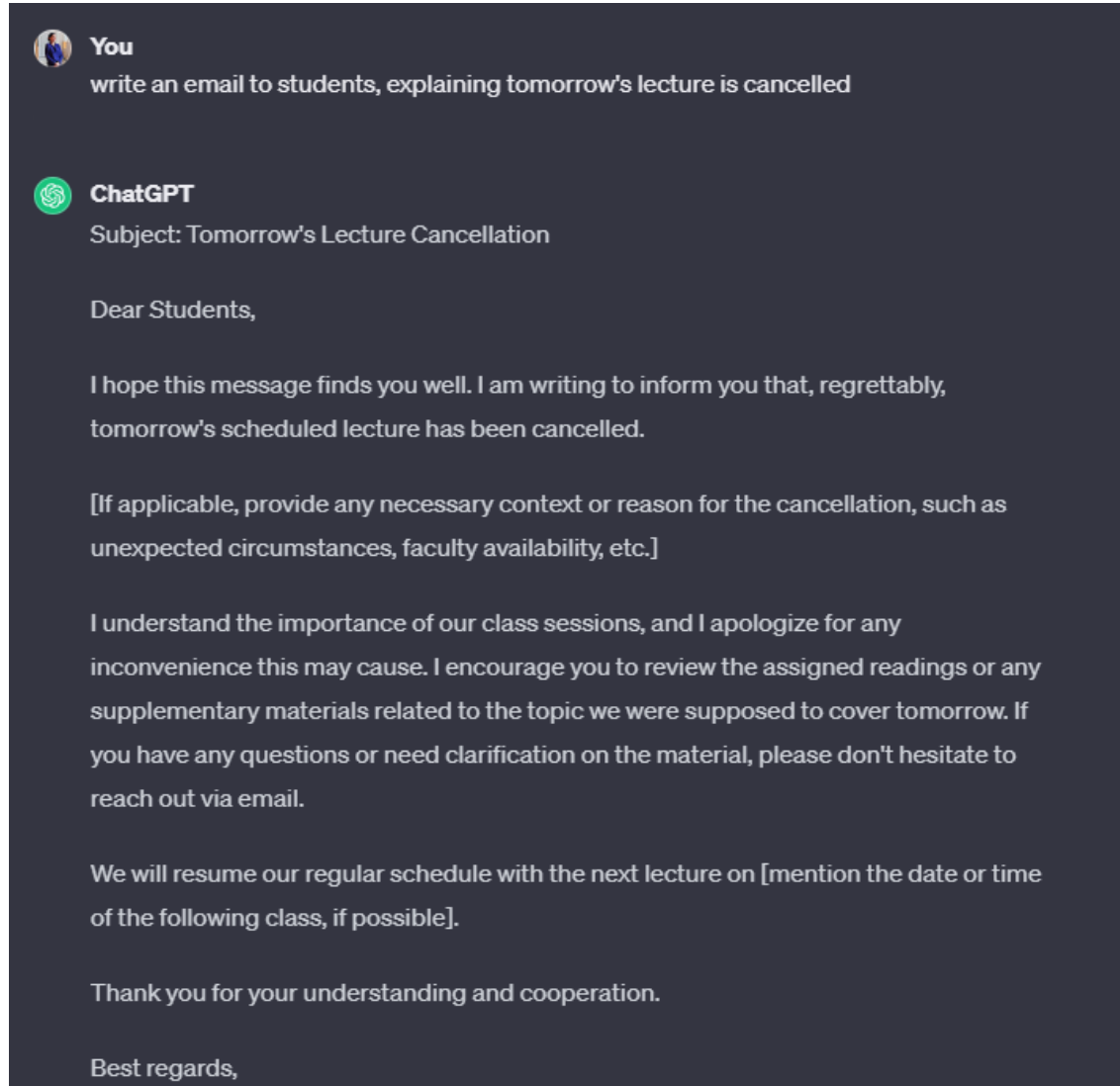
- scriere

generare conținut, idei (brainstorming)



Utilizări ale Generative AI pentru generare de text

- scriere



You
write an email to students, explaining tomorrow's lecture is cancelled

ChatGPT
Subject: Tomorrow's Lecture Cancellation

Dear Students,

I hope this message finds you well. I am writing to inform you that, regrettably, tomorrow's scheduled lecture has been cancelled.

[If applicable, provide any necessary context or reason for the cancellation, such as unexpected circumstances, faculty availability, etc.]

I understand the importance of our class sessions, and I apologize for any inconvenience this may cause. I encourage you to review the assigned readings or any supplementary materials related to the topic we were supposed to cover tomorrow. If you have any questions or need clarification on the material, please don't hesitate to reach out via email.

We will resume our regular schedule with the next lecture on [mention the date or time of the following class, if possible].

Thank you for your understanding and cooperation.

Best regards,

Utilizări ale Generative AI pentru generare de text

- citire (și înțelegere/interpretare a textului)

Ex. Interpretare recomandări restaurant – positive/negative/mixed

"The ambiance at XYZ Bistro was fantastic; the cozy, rustic decor created a warm atmosphere perfect for a night out. However, while the service started off attentive, it became a bit slow as the evening progressed.

The menu boasted a variety of intriguing dishes, and the appetizers were outstanding—especially the calamari, which was perfectly crispy and flavorful. Yet, the main course, though beautifully presented, didn't quite live up to expectations. My steak, ordered medium-rare, arrived closer to medium-well and was a bit tough.

On the upside, the dessert selection was divine; the chocolate lava cake was a heavenly conclusion to the meal.

Overall, while the restaurant had its highlights, there were some hiccups in both service and main courses that left a bit to be desired."

Utilizări ale Generative AI pentru generare de text

- citire (și înțelegere/interpretare a textului)
 - proofreading
 - sumarizare (articole științifice, varianta scrisă a convorbirilor telefonice)
 - analiză mesaje de la clienți
 - monitorizare reputație (analiză feedback de la clienți/review-uri)

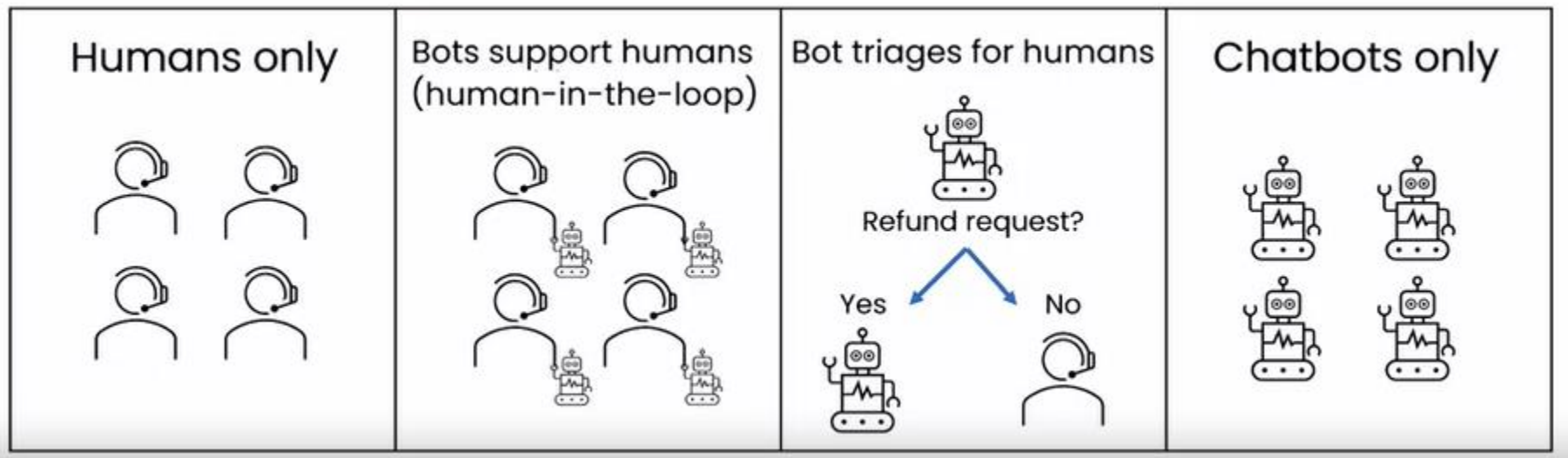
Utilizări ale Generative AI pentru generare de text

- conversație - chatbots

- planificare călătorii
- consiliere personală/de carieră (coaching)
- idei de rețete culinare
- relații cu clienții – asistenți virtuali

Utilizări ale Generative AI pentru generare de text

- conversație - chatbots pentru relații cu clienții



1

2

3

4

Care sistem este cel mai rapid/profitabil/eficient?

Limitări

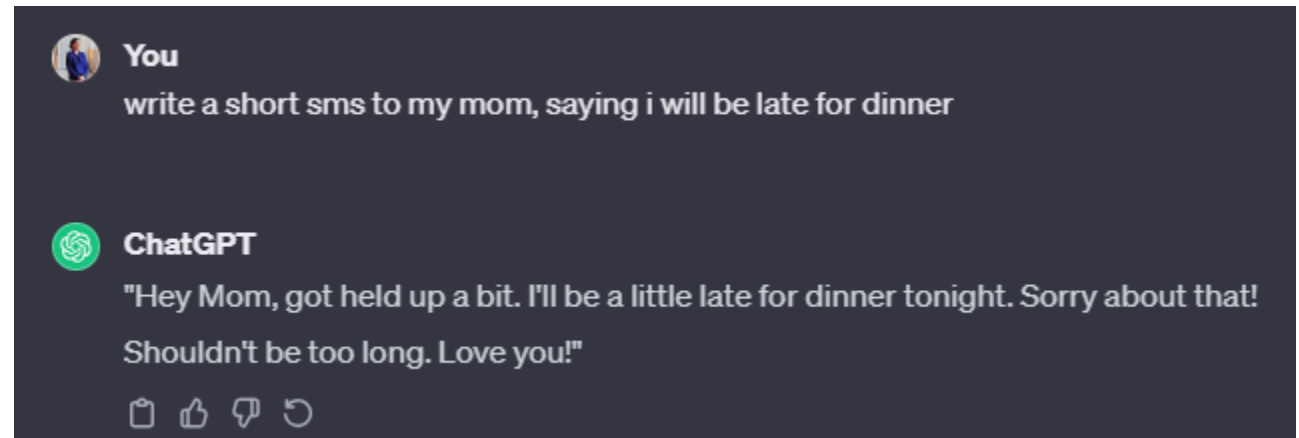
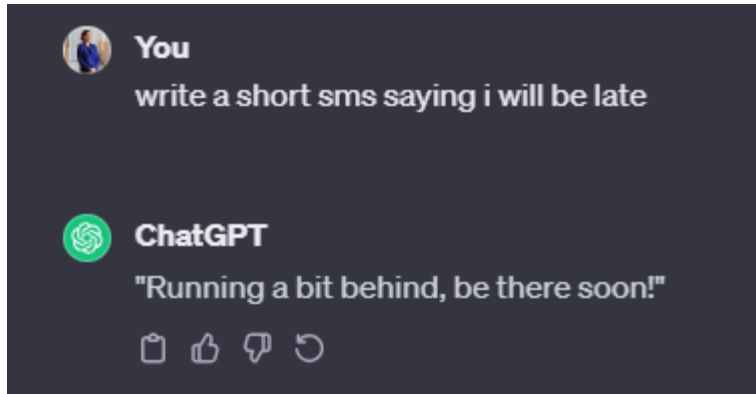
- “Bias and toxicity”

Ex. The surgeon walked out and started his car.
 The nurse walked out and started her car.

Gender bias (polarizare/discriminare de gen)

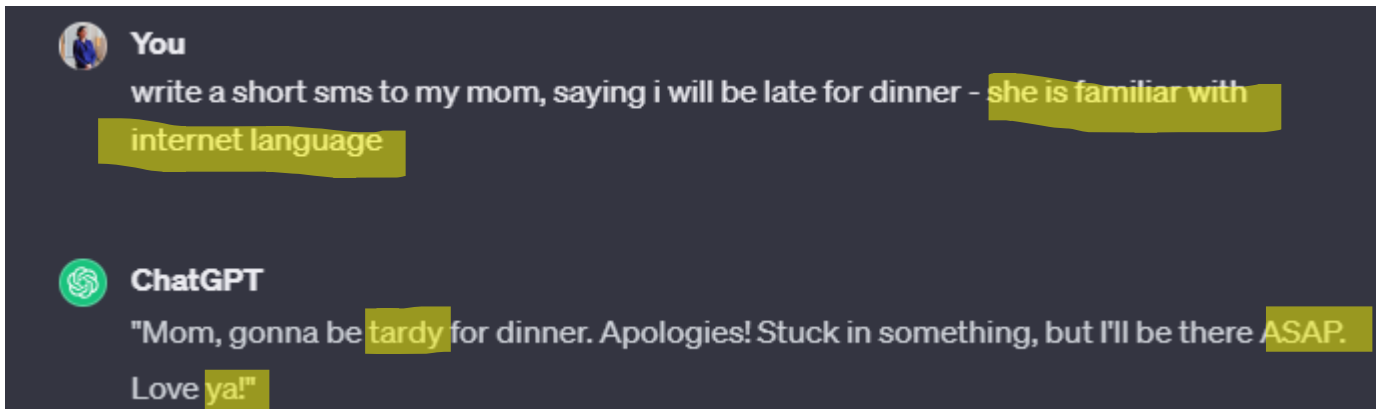
Cum scriem un prompt bun?

- Oferă suficiente detalii relevante



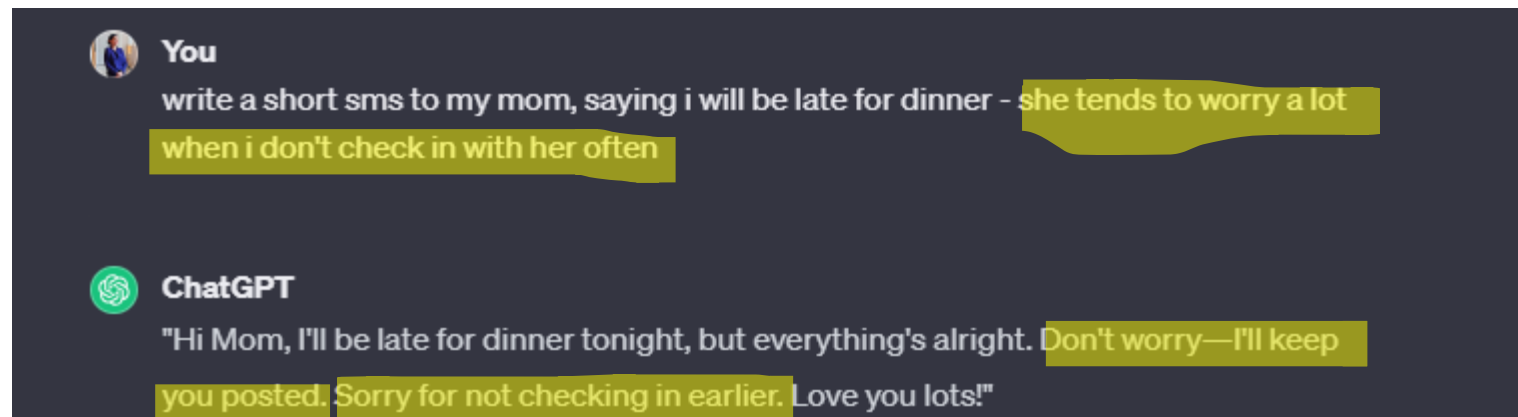
Cum scriem un prompt bun?

- Oferă suficiente detalii relevante
- Ghidează modelul spre tipul de răspuns pe care îl dorești



You
write a short sms to my mom, saying i will be late for dinner - she is familiar with internet language

ChatGPT
"Mom, gonna be tardy for dinner. Apologies! Stuck in something, but I'll be there ASAP.
Love ya!"

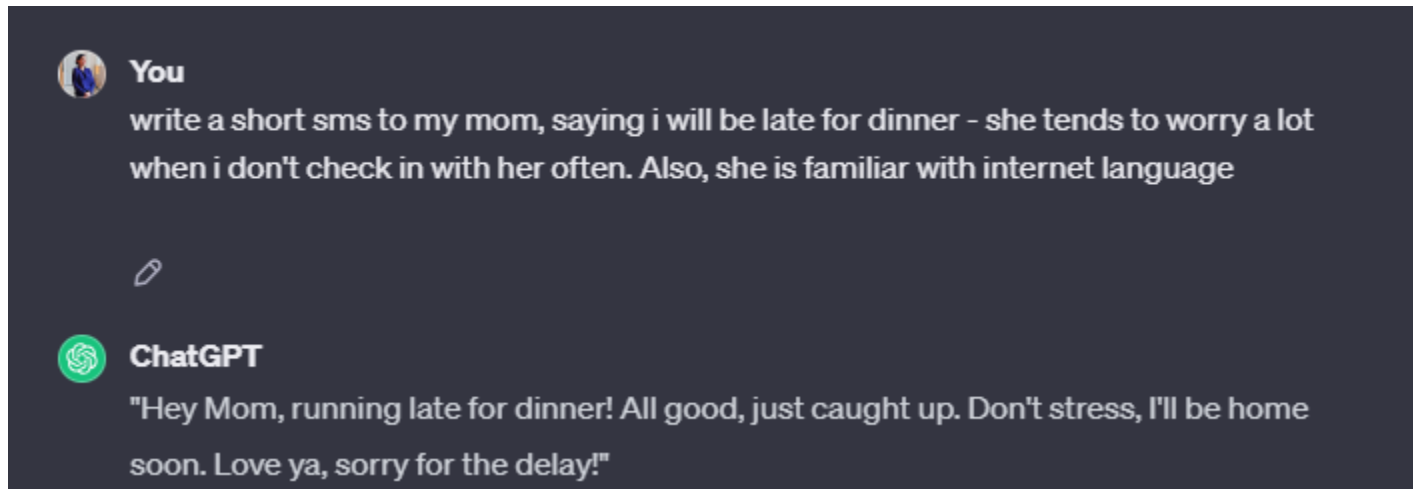


You
write a short sms to my mom, saying i will be late for dinner - she tends to worry a lot when i don't check in with her often

ChatGPT
"Hi Mom, I'll be late for dinner tonight, but everything's alright. Don't worry—I'll keep you posted. Sorry for not checking in earlier. Love you lots!"

Cum scriem un prompt bun?

- Oferă suficiente detalii relevante
- Ghidează modelul spre tipul de răspuns pe care îl dorești
- Repetă întrebarea, experimentând la fiecare pas



Cum scriem un prompt bun?

Atenție la:

- Informații confidențiale incluse în prompt
 - date personale, adresă, CNP, date bancare, etc
- Nivelul de încredere acordat răspunsului primit
 - verificare informații

Alegerea unui LLM potrivit

- Open source/paid
- Dimensiune model
 - 1 mld. parametri – cunoștințe limitate, potrivit pentru aplicații de tip *pattern matching*
 - 10 mld. parametri – poate lucra cu instrucțiuni simple
 - peste 10 mld. parametri – capabil de raționament complex

Top open-source LLMs for 2024

<https://www.datacamp.com/blog/top-open-source-llms>

Limitări

- Cunoștințele pe care le livrează sunt cele acumulate până la momentul antrenării
“knowledge cut-offs”
- “Hallucinations” – replici/citate/articole care nu există, nu au sens
Ex. prompt “Give me 3 quotes of Aristotle about Lady Gaga”
- număr de cuvinte limitat input/output
- GenAI lucrează cu date nestructurate (text, imagini, video)
- Performanțe reduse când sunt furnizate date structurate (tabele, grafice)

Utilizare LLM în viața cotidiană

- Automatizare – *“Automate tasks, not jobs”*

- majoritatea locurilor de muncă implică o multitudine de sarcini
- doar o parte dintre ele pot fi automatizate utilizând GenAI

Exemplu: operator relații cu clienții

Sarcină	Potențial de automatizare cu GenAI
Răspuns la apeluri telefonice de la clienți	
Verifică status contract clienți	
Ține evidența interacțiunii cu clientul	
Evaluează reclamații	

Utilizare LLM în viața cotidiană

- Automatizare – *“Automate tasks, not jobs”*

- majoritatea locurilor de muncă implică o multitudine de sarcini
- doar o parte dintre ele pot fi automatizate utilizând GenAI

Exemplu: operator relații cu clienții

Sarcină	Potențial de automatizare cu GenAI
Răspuns la apeluri telefonice de la clienți	Redus
Verifică status contract clienți	Ridicat
Ține evidența interacțiunii cu clientul	Ridicat
Evaluează reclamații	Redus

Utilizare LLM în viața cotidiană

- Augmentare (îmbunătățire?) – *“help human with a task”*
 - GenAI poate recomanda posibile răspunsuri
 - validarea/aprobarea de către operatorul uman – obligatorie

În multe domenii, se poate începe cu augmentare și se evoluează spre automatizare.

Automatizare/augmentare implică eliminarea operatorilor umani?

“Frica” de GenAI

- Amplificarea impulsurilor negative ale oamenilor
- Pierderea locurilor de muncă
- Dispariția civilizației umane

“Frica” de GenAI

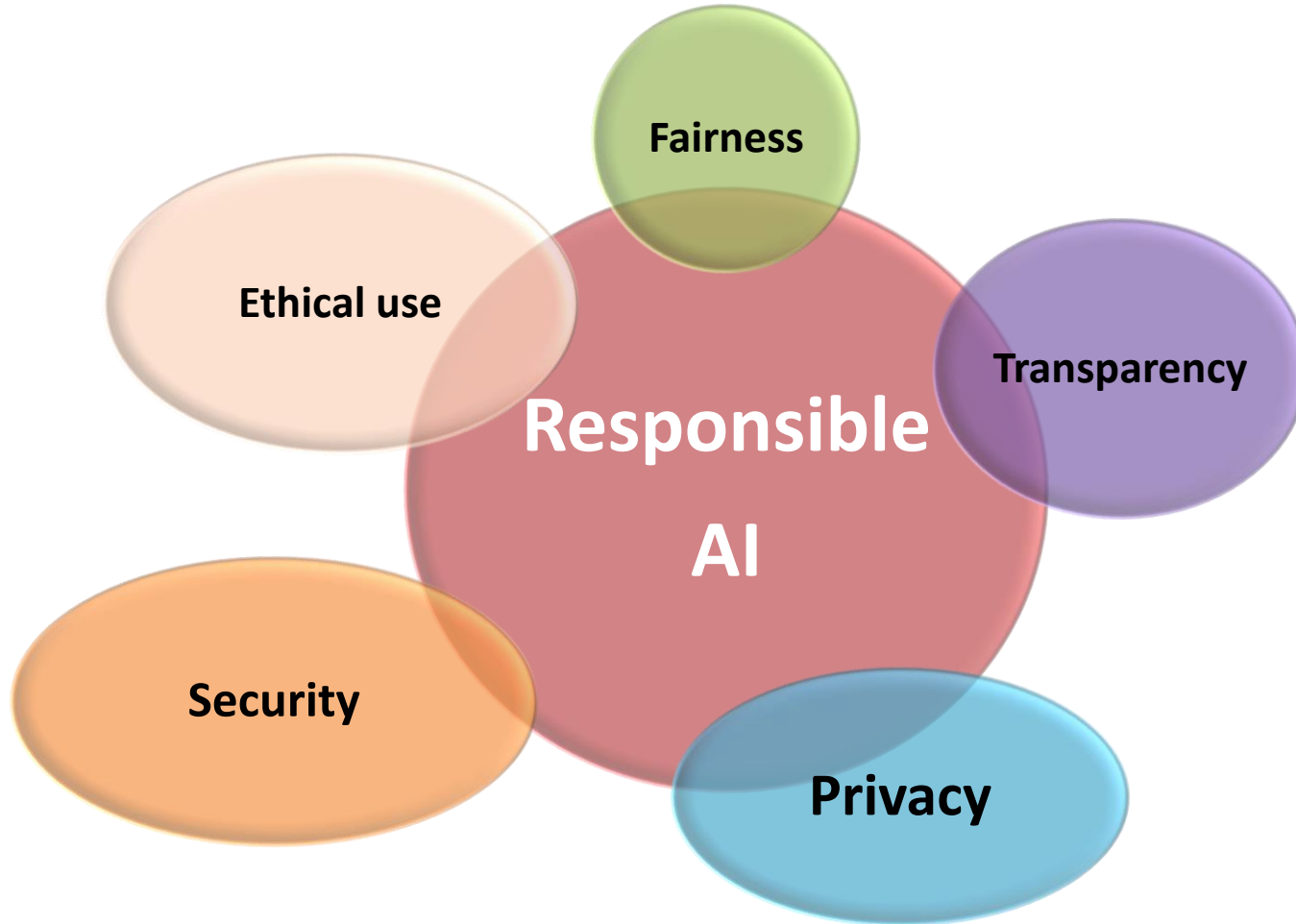
- Amplificarea impulsurilor negative ale oamenilor
 - rasism, xenofobie, discriminare de gen

Metode de eliminare/reducere a polarizării negative:

- RAHF (reinforcement learning from human feedback)

- Pierderea locurilor de muncă
 - automate tasks, not jobs
- Dispariția civilizației umane
 - *AGI – artificial general intelligence* – AI that can do any intellectual task that a human can

“Frica” de GenAI



Gemini AI vs. ChatGPT

Gemini AI – multimodal model

- capabil să interpreteze text, audio, imagini, video și secvențe de cod

Resurse: <https://www.analyticsvidhya.com/blog/2023/12/chatgpt-vs-gemini-a-clash-of-the-titans-in-the-ai-arena/#:~:text=The%20most%20powerful%20version%20of,5%20speech%20recognition%20and%20speech>

<https://em360tech.com/tech-article/gemini-ai-vs-chatgpt>

https://storage.googleapis.com/deepmind-media/gemini/gemini_1_report.pdf

Gemini AI vs. ChatGPT

Sarcină	Gemini AI	ChatGPT-4
Înțelegere generală	90%	86.4%
Raționament	83.6%	83.1%
Înțelegere text scris	87.8%	95.3%
Matematică (operații simple)	94.4%	92%
Matematică (operații complexe)	53.2%	52.9%
Generare cod	74.4%	67%
Transformare limbaj natural în cod Python	74.9%	73.9%

<https://em360tech.com/tech-article/gemini-ai-vs-chatgpt>

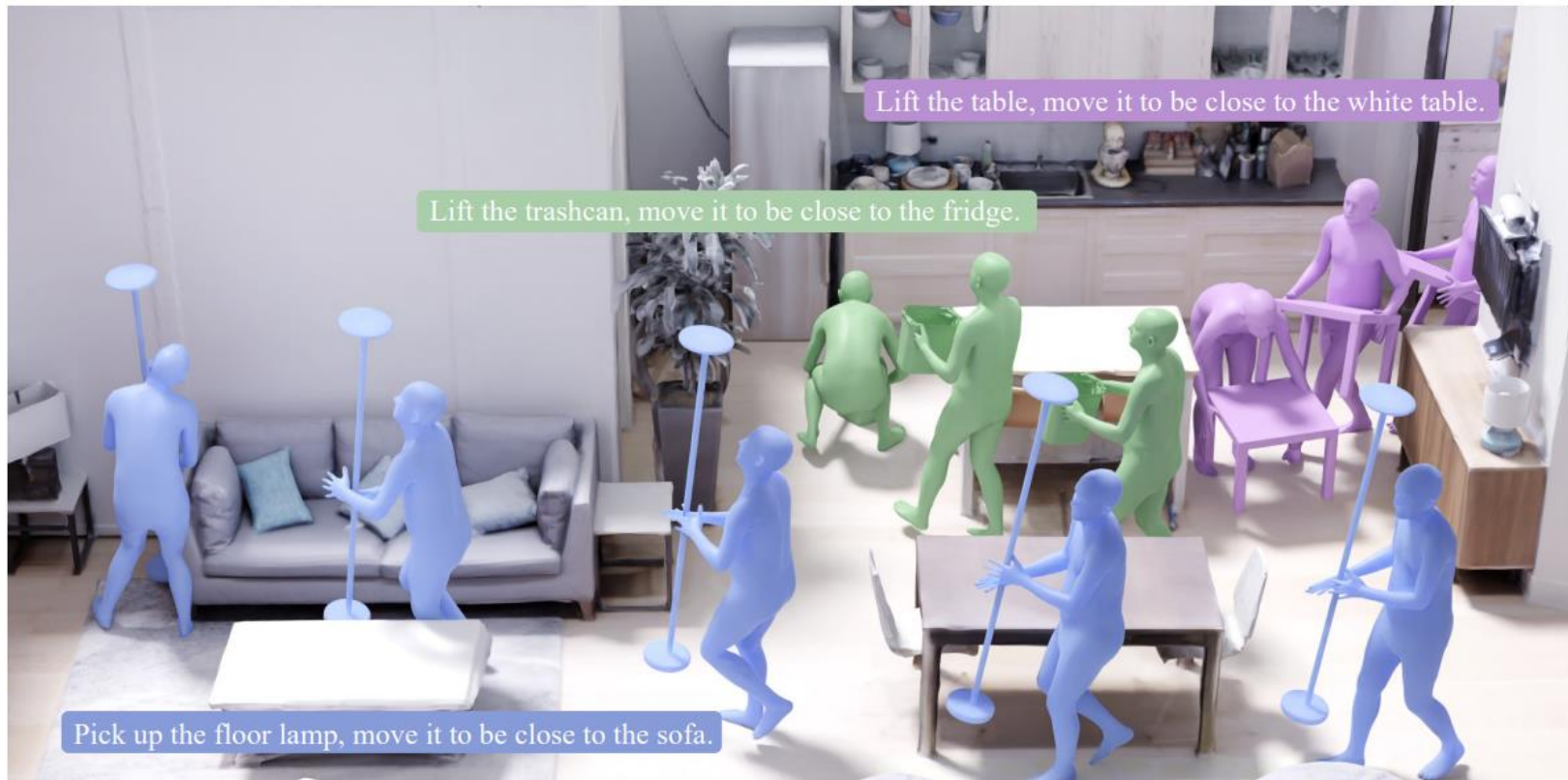
Gemini AI vs. ChatGPT

Feature	Gemini AI	ChatGPT
Size	Larger	Smaller
Power	More powerful	Less powerful
Versatility	More versatile	Less versatile
Efficiency	More efficient	Less efficient
Memory and planning capabilities	Better	Worse
Availability	Less widely available	More widely available
Affordability	Less affordable	More affordable

<https://em360tech.com/tech-article/gemini-ai-vs-chatgpt>

CHOIS – controllable human-object interaction synthesis

Given an initial object and human state, a language description, and sparse object waypoints in a 3D scene, CHOIS generates synchronized object motion and human motion at the same time.



arxiv.org/pdf/2312.03913.pdf

CHOIS – controllable human-object interaction synthesis



Figure 6. **Results of interaction synthesis** using the same text input but different waypoints. The initial state is in green.

arxiv.org/pdf/2312.03913.pdf

[Stanford and Meta inch towards AI that acts human with new 'CHOIS' interaction model | VentureBeat](#)

- Definiere, concepte, funcționare ✓
- Scenarii de utilizare și limitări ✓
- Large Language Models (LLM) ✓
- Provocări și riscuri ✓
- Gemini AI vs. ChatGPT ✓
- CHOIS ✓

În episodul următor: **Aplicații ale AI-DSS în domeniul financiar și în marketing.**