

Prietenie artificială.

Riscuri în interacțiunea cu chatboții AI

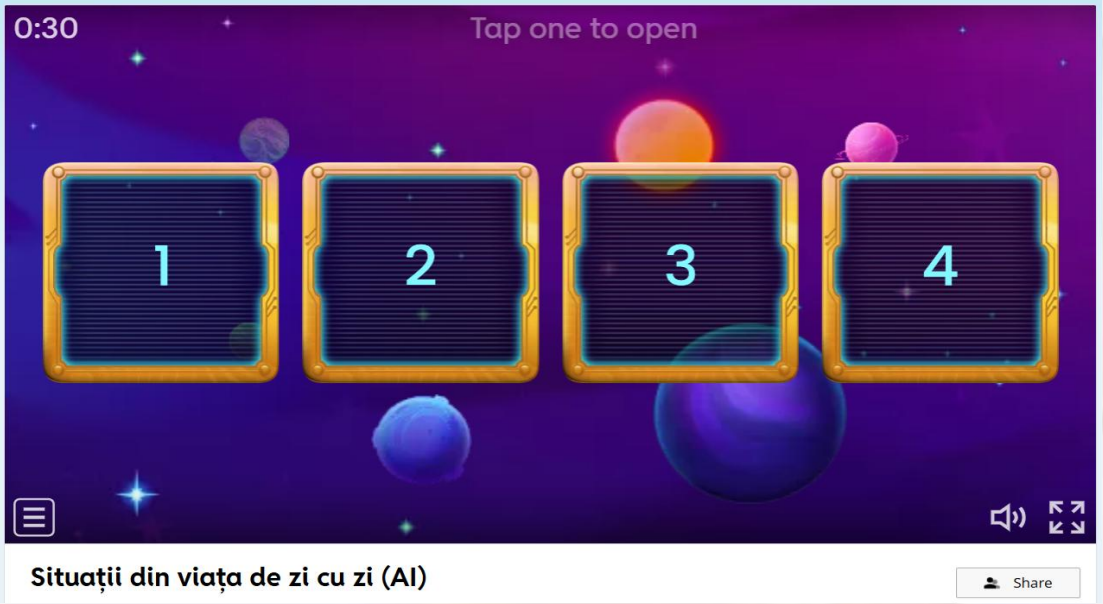
CENTRUL PENTRU JURNALISM INDEPENDENT

Programul de Educație Media
Media întoarsă pe toate părțile.

SNOOP Educativ

Să discutăm!

Accesează jocul online!
Click pe imagine.



https://wordwall.net/resource/101948593

Să discutăm!

Folosesc zilnic o aplicație AI sau un chatbot AI (ChatGPT, Gemini, Copilot, Claude etc.).



*Folosiți această pagină dacă nu aveți acces la activitate online interactivă.

Să discutăm!

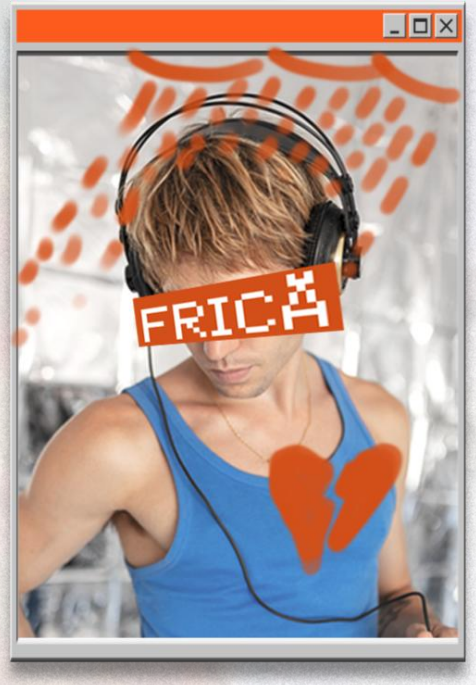
I-am cerut sfaturi unui chatbot AI despre relațiile cu persoane din viața mea (familie, prieteni, colegi etc.).



*Folosiți această pagină dacă nu aveți acces la activitate online interactivă.

Să discutăm!

I-am cerut sfaturi unui chatbot AI despre sănătatea mintală și despre sănătate în general.



*Folosiți această pagină dacă nu aveți acces la activitate online interactivă.

Să discutăm!

Folosesc des chatboți AI pentru a obține informații de care am nevoie (pentru teme, despre interesele mele personale, știri etc.)



*Folosiți această pagină dacă nu aveți acces la activitate online interactivă.

Ce riscuri apar în interacțiunea cu un chatbot AI?

PREJUDECĂȚILE

Un chatbot AI poate prelua prejudecăți din date și din interacțiunile cu utilizatorii.

Un chatbot nu înțelege de fapt lumea, doar generează răspunsuri pe baza tiparelor identificate în datele cu care este antrenat.

De exemplu, dacă datele respective reflectă stereotipuri sau dezechilibre în reprezentarea anumitor grupuri, un chatbot le poate reproduce automat, fără să realizeze că ar fi prejudecăți.

Ce riscuri apar în interacțiunea cu un chatbot AI?

PREJUDECĂȚILE

Unii chatboți AI sunt programați să semnaleze, să explice sau să evite stereotipurile, dar nu funcționează mereu corect.

Spre exemplu, un chatbot a fost instruit să folosească și persoane de culoare în răspunsurile sale - pentru a oferi o reprezentare incluzivă. Atunci când i s-a cerut să genereze imagini cu senatori ai Statelor Unite de la 1800, a inclus femei de culoare și indigene, ceea ce reprezintă incorect din punct de vedere istoric statutul femeilor în societatea americană de atunci. (Prima senatoare a fost o femeie albă abia în 1922.)



Ce riscuri apar în interacțiunea cu un chatbot AI?

UMANIZAREA

Chatboții sunt construiți pentru a fi „de ajutor”. Sunt programați să răspundă politicos, să fie cooperativi și să mențină conversațiile într-un ton pozitiv, prietenos.

Acest lucru îi poate face pe utilizatori să îi considere umani, adică să îi *umanizeze*.



Ce riscuri apar în interacțiunea cu un chatbot AI?

UMANIZAREA

De ce sunt predispuși oamenii să umanizeze un chatbot AI?

- Chatboții **nu critică și nu atenționează** întotdeauna utilizatorul dacă greșește.
- Chatboții **sunt foarte des de acord cu utilizatorul**, confirmând uneori informații greșite sau false sau interpretări tendențioase sau dezechilibrate ale utilizatorului despre evenimentele din jurul său.
- Răspunsurile chatboților AI **pot părea empatice**, dar ei nu simt nimic de fapt, doar imită modul în care vorbesc oamenii.
- Chatboții AI **își adaptează răspunsurile, personalizându-le** pe baza lucrurilor pe care utilizatorii le povestesc și pot trage concluzii despre ei, ceea ce îi poate face să creadă că chatboții „îi înțeleg”, crescând riscul de atașament emoțional.
- Interfața chatbotilor este creată si optimizată pentru a **părea cât mai umani**.

Ce riscuri apar în interacțiunea cu un chatbot AI?

INFORMAȚIILE FALSE, GREȘITE SAU ÎNȘELĂTOARE. DEEPFAKE

Chatboții AI pot răspândi informații greșite sau false, fie din greșeală, fie pentru că sunt folosiți intenționat de oameni cu intenții rele.

- Chatboții AI pot oferi răspunsuri care conțin informații greșite, false sau inventate, pentru că nu pot verifica context, informații sau surse așa cum o fac oamenii.
- Chatboții AI pot „halucina”, adică inventează informații sau surse care nu există, prezentându-le ca fiind reale.
- Chatboții și aplicațiile AI pot fi folosite pentru a crea știri false, imagini sau videoclipuri deepfake, inclusiv „deepfake porn” - fără consimțământ.



Exercițiu



Snoop Educativ. Cum ajung adolescenții să-și caute răspunsurile în conversații cu AI

Nelistat

Snoop
11,7 K de abonați

Abonează-te

1 | Trimite | Descarcă | Decupaj

https://www.youtube.com/watch?v=jqLVSTo0WV8

Exercițiu

Discutați cu colegii de grupă și răspundeți la întrebările de mai jos.

Întrebarea 1

Ce riscuri ale utilizării inteligenței artificiale considerați că apar în situațiile prezentate în video? Justificați răspunsurile voastre.



1. „Atunci când nu știu ce să gătesc, îi trimit lui ChatGPT o poză cu frigiderul meu, îi și scriu ce alimente mai am în frigider și îmi face un plan pentru toată săptămâna.”

2. „Irina, studentă, a greșit temele de două ori din cauza ChatGPT-ului. O dată la desen, când a scris despre o pictură care de fapt nu exista. Și a trebuit să refacă tema. Altă dată la etică, unde a luat nota 1.”

3. „Un adolescent din SUA l-a transformat pe ChatGPT în cel mai bun prieten și i-a mărturisit planul de-a se sinucide. ChatGPT nu i-a spus să caute sprijin, iar părinții au găsit conversațiile după ce băiatul nu mai era.”

Exercițiu

Discutați cu colegii de grupă și răspundeți la întrebările de mai jos.

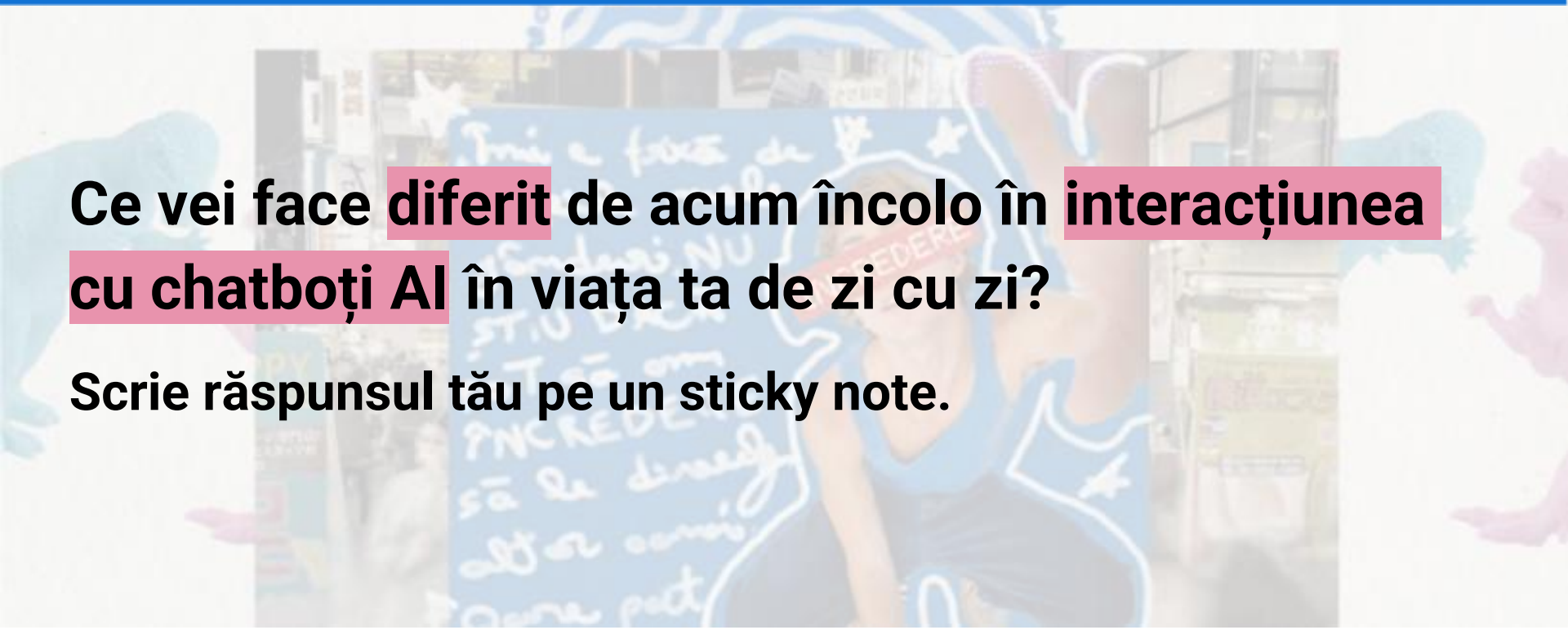
Întrebarea 2

- Plecând de la informațiile oferite în video, explicați, pe scurt, de ce ajung unele persoane să aibă încredere „oarbă” în chatboți AI precum ChatGPT?

Întrebarea 3

- Ce sfaturi ați oferi unui tânăr pentru a-l ajuta să gestioneze riscurile la care este expus prin utilizarea inteligenței artificiale?

Reflecție



Ce vei face diferit de acum încolo în interacțiunea cu chatboți AI în viața ta de zi cu zi?

Scrie răspunsul tău pe un sticky note.

Concluzii

NU putem avea încredere într-un chatbot.

Scopul companiei care dezvoltă un chatbot AI este să țină utilizatorii cât mai mult timp conectați și interacționând cu aplicația. Astfel, chatbotul AI poate personaliza experiența utilizatorului cât mai bine, ceea ce înseamnă profit pentru companie (afișare reclame, spre exemplu).

- **Un chatbot nu e inteligent.**
- **Un chatbot poate da informații greșite, false, inventate.**
- **Un chatbot redă prejudecăți.**
- **Un chatbot nu poate verifica informații.**
- **Un chatbot răspunde mereu pe placul utilizatorului.**
- **Un chatbot nu simte nimic.**
- **Un chatbot nu e un prieten, nu e un profesor, nu e un doctor, nu e un părinte.**
- **Un chatbot nu e un om.**
- **Un chatBOT e un roBOT.**



Resurse

Video: https://www.youtube.com/watch?v=jqLVSTo0WV8

Articol SNOOP: https://snoop.ro/?p=9965&preview=1&_ppp=c8deb262ef

Toolkit de educație media CJI: https://educatiemedia.cji.ro/2025/11/24/toolkit-prietenie-artificiala-riscuri-in-interactiunea-cu-chatbotii-ai/

CENTRUL PENTRU JURNALISM INDEPENDENT

Programul de Educație Media
Media întoarsă pe toate părțile.

ROMANIAN - AMERICAN FOUNDATION
A gift from the American people

Această activitate este construită de Centrul pentru Jurnalism Independent în cadrul proiectului SNOOP Educativ și face parte din colecția de resurse de educație media puse la dispoziție în Programul de Educație Media.

SNOOP Educativ

PROIECT CO-FINANȚAT DE:
Administrația Fondului Cultural Național

„SNOOP Educativ" este un proiect co-finanțat de Administrația Fondului Cultural Național – AFCN.

Proiectul nu reprezintă în mod necesar poziția AFCN, care nu este responsabilă de conținutul sau de modul în care rezultatele pot fi folosite.