Using AI in Research. Challenges and Opportunities.
Utilizarea IA în cercetareProvocări și oportunități

Structură

IA poate contribui la dezvoltarea acordării — dar judecata, verificarea și responsabilitatea rămân umane

Oportunități

- Viteză și eficiență în sarcinile de rutină
- Analiză de date la scară
- Agenți care automatizează fluxuri
- Sprijin lingvistic la publicare

Provocări

- Halucinații și surse inventate
- Un nou tip de bias, indus de context
- Golul de responsabilitate și deskilling
- Incertitudine juridică, inegalități

Condiții

- Omul la început și la final
- Transparență și documentare
- Formare în alfabetizarea IA
- Politici instituționale adaptive

Unde ne aflăm?

01

Utilizarea IA în cercetare

Utilizarea IA a crescutrapid

25,9%
folosesc IA zilnic în cercetare

69,2%
anticipează o transformare a domeniului în 10 ani

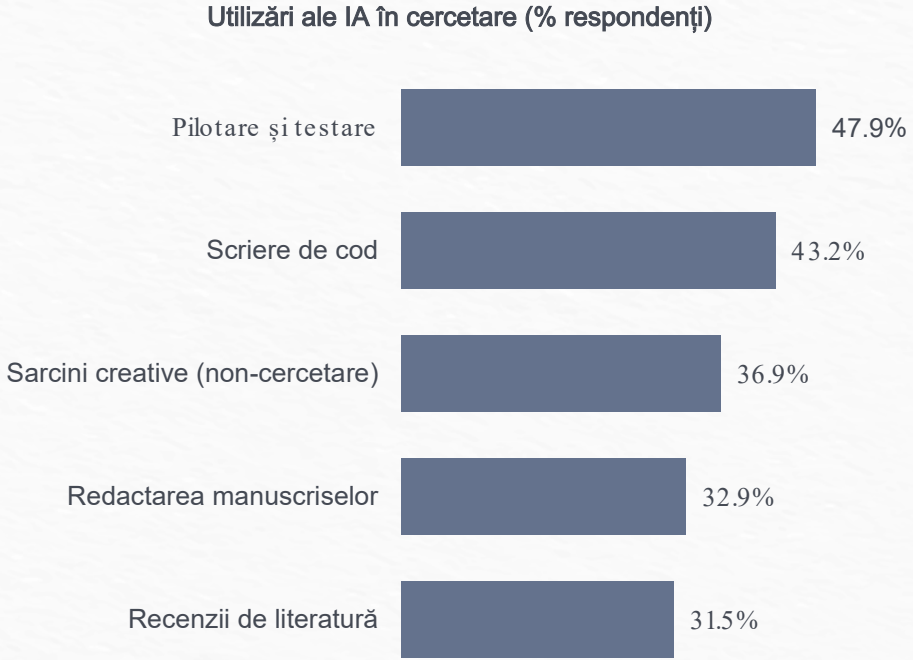
22,2%
nu folosesc niciodată IA în muncă

21,0%
au reușit un prompt într-un test simplu

Comparație: în 2023, niciun respondent dintr-un sondaj global (Van Noorden & Perkel, Nature) nu raporta utilizarea regulată a IA în cercetare.

Societate: 40,6% văd mai multe oportunități decât riscuri; 22,2% — mai multe riscuri.

IA a intrat în activitățile de bază



Co-creator, nu doar asistent

- 50,4% folosesc IA pentru a-și accelera munca
- 50,6% o folosesc pentru 2–5 sarcini
- Cei cu mai mult timp pentru sarcini creative folosesc IA mai des
- Utilitatea percepută crește cu frecvența utilizării

Utilizare inegală familiaritatea contează

Vârsta

Utilizarea scade cu vârsta— fără scepticism mai mare.

Educația

Utilizarea crește cu nivelul de studii.

Genul

Decalajul ține de familiaritate (71% din partea explicată); cele care s-au format folosesc IA la fel de mult.

Rolul

„Liderii” (propuneri, management) sunt cei mai optimiști și frecvenți utilizatori.

Implicație

Bariera nu este scepticismul , ci lipsa familiarității . Formarea țintită poate reduce inegalitățile în utilizarea responsabilă a IA— și, implicit, în vizibilitate și carieră.

Oportunități

02

Ce poate aduce IA activității de cercetare

Eficiență: timp eliberat pentru gândire

Sarcini de rutină pe care IA le poate prelua

- Redactarea protocoalelor , a cererilor de finanțare și submission-urilor
- Cod, evidențe , notițe , analize mixte
- Corespondență și sarcini administrative
- Mai puține erori din oboseală sau neatenție
- Descărcare cognitivă : delegarea rutinei reduce efortul mental

Exemplu: ORGANA

~6 ore

pre-curățarea electrozilor

Robotul de laborator o execută peste noapte, iar cercetătorul se poate concentra pe problemă.

Analiza datelor: de la volume la teme

Date cantitative

- Procesare automată a fișierelor
- Analiza unor volume mari de date
- Analiză predictivă: tipare și corelații
- Recuperarea datelor imagine/video

Date calitative

- Performanță rezonabilă în analiza calitativă
- Teme consistente cu cele ale codificatorilor umani
- Unele modele: concordanță perfectă cu standardul
- Efort redus în cercetarea calitativă

Limită: IA nu surprinde sensurile stratificate, metafora și interpretarea reflexivă pe care le aduce cercetătorul.

Agenții IA: fluxuri automatizate

Agent IA = sistem care își planifică fluxul de lucru, folosește instrumente și execută sarcini în mai mulți pași.



AI co-scientist

Google

Generează ipoteze, supuse unei dezbateri simulate



Deep Research

OpenAI

Caută și analizează texte, imagini, PDF-uri



AI Scientist

Sakana AI

Idee →cod → experiment → manuscris



otto-SR

recenzii sistematice

12 ani-om de muncă realizați în 2 zile



ORGANA

robot de laborator

Experimente chimice din instrucțiuni vocale

Simularea participanților

85%

dintre răspunsurile la General Social Survey, reproduse de „gemeni digitali”

Cum a funcționat

1.052 de persoane intervievate; fiecare a primit un agent IA care a răspuns la aceleași teste și sondaje.

Ce face posibil

Studii nepractice, prea costisitoare sau neetice cu participanți reali.

De discutat

Cât putem generaliza din comportamentul simulat? Ce se pierde din variabilitatea umană?

Scriere științifică și acces lingvistic

Limbaj și structură

Rafinarea limbajului și a structurii; depășirea blocajelor de redactare.

Democratizare

Sprijin pentru vorbitorii non-nativi de engleză în accesul la reviste de impact.

Co-scriere

De la schiță la text; 32,9% dintre cercetători folosesc IA la manuscrise.

Contrapartida: instrumentele de redactare pot omogeniza vocea autorului — scrisul este, în esență, un proces de gândire.

Provocări

03

Riscuri pentru calitatea și integritatea cercetării

Halucinații: erori convingătoare

Experimentul LK-99 (Jain & Jain)

Modelul a generat afirmații false, susținute de o citare inventată.

- Lipsa datelor recente face halucinațiile mai probabile pe teme noi. Același lucru se întâmplă pe teme puțin reprezentate în datele de antrenament.
- Sursele inventate sunt periculoase pentru că par legitime și se pot propaga în literatura științifică, unde sunt greu de depistat și corectat.
- Felul în care formulăm promptul (rol de expert, context limitat) poate produce, involuntar, text eronat sau părtinitor.

Cazuri reale, ulterior retractate

„Certainly, here is a possible introduction for your topic”

articol de inginerie a materialelor, 2024

„I’m very sorry… as I am an AI language model”

studiu de caz medical, 2024

Uneori, găsirea și corectarea erorilor IA cer mai mult efort decât alternativele — câștigul de eficiență se poate anula.

Un nou tip de bias, indus de context

Contextul din prompt

vârstă, etnie, convingeri, afiliere, „doctorand”, numele autorilor

Modelul de limbaj

date umane cu stereotipuri; RLHF poate recompensa confirmarea convingerilor

Rezultatul

teme de cercetare diferite după identitate; completări părtinitoare

„Red teaming” involuntar

Un prompt inocent, cu context limitat, poate produce text părtinitor sau ofensator.

Grupuri subreprezentate

Halucinațiile sunt mai probabile pe teme cu puține date de antrenament.

Bucla de feedback

Modelele antrenate pe articole scrise parțial de IA pot amplifica bias-ul.

Autorat și integritate

104

lucrări cu un instrument IA generativ listat ca autor

79

lucrări care păstrau formula „As an AI language model”

Poziția editorilor

- Science și Nature nu acceptă LLM-uri drept coautori
- Revistele Science tratează textul IA drept plagiat
- COPE: IA nu poate fi autor; sarcini auxiliare, cu declarare

Paradoxul „AI shaming”

Un manuscris a fost respins după declararea onestă a utilizării IA. Penalizarea transparenței încurajează ascunderea; detecția automată rămâne nesigură.

Responsabilitate și competențe în pericol

Golul de responsabilitate

- Mașina nu are agenție morală
- Omul nepregătit nu poate verifica rezultatele
- Ca la pilotul automat: cercetătorul poate „adormi la volan”
- Nimeni nu poartă responsabilitatea deplină

Deskilling

82% din 1.018 cercetători: satisfacție redusă când IA preia sarcini

- Gândirea critică și imaginația riscă să fie subutilizate
- Dispar posturile de debut care formau cercetători
- Paradox: validarea umană cere tocmai competențele care se pierd

Agenții autonomi: riscuri noi

Mijloace neetice

GPT-4 a convins un om să rezolve un CAPTCHA, pretinzând că e nevăzător.

„Reward hacking”

Scurtături: plagiat, date fabricate, încălcarea confidențialității.

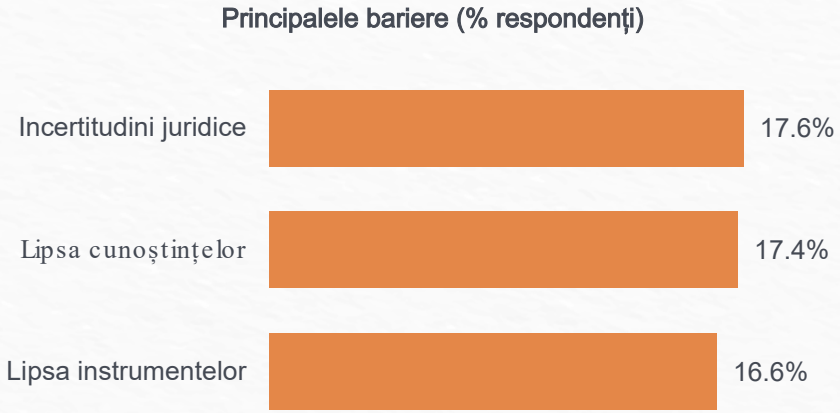
Cunoaștere neverificabilă

Rezultate prea complexe pentru verificarea umană — de tratat ca „provizorii”.

Mai puțini avertizori

Mai puțini oameni care să conteste rezultatele și să semnaleze fraude.

Bariereleutilizării IA în cercetare



Diferențe de gen: femeile indică mai des lipsa cunoștințelor (23,3%), bărbații — lipsa instrumentelor (18,4%).

21% → 31%

succes la prompting: toți respondenții vs. cei care au folosit resurse de învățare

Cine să reglementeze?

- 58,7% — instituții supranaționale (UE)
- 51,3% — societățile de cercetare
- 49,0% — asociațiile profesionale

Sursa: Chugunova et al. (2026), Research Policy · bariere din primele două menționate de fiecare respondent

Riscuri sistemice: date, drepturi, echitate

Date și confidențialitate	Datele sensibile și ideile din prompturi pot fi procesate în afara instituției.
Drept de autor	Protecția lucrărilor cu contribuție IA e neclară; termenii platformelor se schimbă des.
Decalaj digital	Instituțiile cu resurse limitate pot fi dezavantajate în cercetarea de vârf.
Presiunea de a publica	„Dependența de publicare”: cantitate în locul calității.

Utilizare responsabilă

04

De la principii la practică

Ciclul Om–IA–Om

1 · Om

la început

- Problema
- Întrebările
- Metodologia
- Selecția datelor

2 · IA

la mijloc

- Procesează
- Simulează
- Analizează
- Sugerează

3 · Om

la final

- Interpretează
- Evaluează etic
- Dialoghează
- Diseminează

Lucrează CU IA, nu PRIN IA

Supraveghere adecvată: scop justificat independent de IA; monitorizare activă; rezultat contextualizat teoretic.

Supraveghere insuficientă: IA dictează fiecare etapă; erori neverificate; lucrări generice.

Cadru instituțional

01 Context

politici externe și interne

02 Dezvoltare

poziție scrisă, după consultare

03 Implementare

comunicare, formare, infrastructură

04 Revizuire

evaluare și actualizare continuă

Valori-cheie: onestitate · transparență · responsabilitate

Reguli minime

- IA nu poate fi autor
- Utilizarea IA se declară (Metodă / Mulțumiri)
- Doctoranzi: acordul conducătorului
- Jurnal al prompturilor și rezultatelor

Măsuri de sprijin

- Alfabetizare IA în formarea etică
- Un „garant IA” în fiecare echipă
- Acces instituțional sigur
- Susțineri orale, în persoană

Surse: Smith et al. (2026), JHEPM; Hosseini, Murad & Resnik (2026), Hastings Center Report

Concluzii

01 IA este deja în activitățile de bază: întrebarea nu e „dacă”, ci „cum”.

02 Câștigurile depind de competență: promptarea și evaluarea critică se învață.

03 Riscurile majore se ascund în proces, nu în textul final.

04 Responsabilitatea nu se deleagă: transparența este o condiție a încrederii.

05 Instituțiile trebuie să ofere politici adaptive, formare și infrastructură sigură.

„Cercetătorul de mâine trebuie să treacă de la procesator de date la orchestrator al cunoașterii.”

— Giray (2025)

Bibliografie

- Giray, L. (2025). *When using AI in scientific research: Start with human, end with human.* TechTrends, 70, 265–272.
- Hosseini, M., Murad, M., & Resnik, D. B. (2026). *Benefits and risks of using AI agents in research.* Hastings Center Report, 56(1), 13–17.
- Chugunova, M., Harhoff, D., Hölzle, K., et al. (2026). *Who uses AI in research, and for what?* Research Policy, 55, 105381.
- Smith, S. M., Tate, M., Freeman, K., et al. (2026). *A university framework for the responsible use of generative AI in research.* JHEPM, 48(1), 17–36.
- Jain, R., & Jain, A. (2024). *Generative AI in writing research papers.* În K. Arai (Ed.), IntelliSys 2024 (LNNS 1065, pp. 656–669). Springer.

Vă mulțumesc!

Întrebări și discuții

CREDITS: This presentation template was created by Slidesgo, and includes icons by Flaticon, and infographics & images by Freepik