

LIBRIS

We know
books

Biblioteca Revistei Curtea de la Argeș

Nr. 46

© Editura EIKON
© Gheorghe Păun

București, Calea Giulești 333, Sector 6,
cod poștal 031310, România

Difuzare / distribuție carte: 021 348 14 74
0733 131 145, 0728 084 802
difuzare@edituraeikon.ro

Redacția: 021 348 14 74
0728 084 802, 0733 131 145
contact@edituraeikon.ro
www.librariaeikon.ro

Editura Eikon este acreditată de
Consiliul Național al Cercetării Științifice (CNCS)

Descrierea CIP a Bibliotecii Naționale a României
PĂUN, GHEORGHE

Inteligența artificială în context istoric : puncte de vedere /
Gheorghe Păun. - București : Eikon, 2025
ISBN 978-606-49-1450-7
004

DTP: Mihăiță Stroe

Editor: Valentin Ajder

Gheorghe Păun

Inteligența artificială în context istoric

Puncte de vedere

E I K O N

București, 2025

n-are rost să ciclăm, oricum rămân impresio-
nat. Am făcut însă altceva: l-am chestionat din
nou pe ChatGPT cu privire la Urmuz și la...
Gh. Păun... Dezamăgire totală! Halucinații în
serie – altele decât cele de acum o lună și ceva,
dar livrate cu aceeași nonșalanță de mașină
inertă.

Oricum, mulțumesc ChatGPT, mulțumesc
Andrei Păun!

Cuprins

1. Precizări preliminare	5
2. Trei idei care merită a fi reținute	9
3. Sumar al demersului	12
4. Precauții și premise metodologice	14
5. Doi clasici, despre subiect	26
6. Manifestul tehnorealismului	32
7. Despre definiție	34
8. Scurtă istorie (a informaticii și a IA).....	44
9. Calculul natural (inspirat din biologie)	53
10. Patru rezultate matematice de interes pentru IA..	60
11. Impact bun, impact rău.....	68
12. Pericolele IA, văzute de IA	78
13. Supravegherea algoritmilor	83
14. Plasare în istoria civilizației umane	87
15. Înspre încheiere, Harari	96
16. Experiențe personale cu ChatGPT	101
17. Concluzii (?).....	107
18. Bonus de final	110
19. Referatul ChatGPT asupra textului anterior	113

1. Precizări preliminare

Textul de față s-a născut ca urmare directă a conferinței cu același titlu pe care am ținut-o pe 22 mai 2025 la Sala Dalles din București, în seria „Conferințele Dalles ale Academiei Române”. Am primit invitația/propunerea respectivă cu vreo două luni înainte, de la dl acad. Gheorghe Chivu, organizatorul acestor întâlniri, și am acceptat-o imediat – pot zice acum că am acceptat-o în grabă, dintr-un impuls de... săgetător vulnerabil la provocări. Știam că e un subiect vast-stufos-în-plină-evoluție, dar nu știam cu adevărat ce mă așteaptă. Prima „eroare” a fost să citesc pe telefonul cel *smart* (rezerv termenul „intelligent” pentru tema conferinței, cu ghilimele însă) articolele despre IA (așa o voi prescurta de acum încolo) pe care le-am găsit printre

multe altele cu care mă îmbia Android sau, nu știi exact cum merg lucrurile, compania de telefoane la care sunt abonat. Cei care folosesc asemenea surse de informare bănuiesc ce a urmat: de a doua zi, am primit tot mai multe notificări despre IA și, cu cât am citit mai multe, cu atât au venit mai multe. Realmente, zgomot.

Evident, am căutat și prin cărți, mai vechi sau mai noi, prin biblioteca personală, prin memorie (este de presupus că un pensionar care a lucrat o viață în informatică, patru ani, cei de început, chiar în programare, apoi în informatică teoretică, a adunat ceva amintiri relevante).

Rezultatul se va vedea imediat.

Deja rândurile anterioare conțin o serie de mesaje. Le fac pe unele explicite, din onestitate și, mărturisesc, și pentru a-l atrage pe cititor la (mă rog, pentru a-i propune) o întâlnire colocvială, interactivă ar fi ideal, dacă așa ceva ar fi posibil. De aici și stilul, deja ilustrat: autobiografic, de „jurnal”, cu paranteze relaxate, cu citate multe, cu referințe multe – pe cât se va putea, „strecurate” firesc în text (nu-mi plac notele de subsol). Textul scris va imita stilul oral al conferinței, dar fără presiunea timpului,

cu frazele duse până la capăt. Insist: nu va fi o lucrare științifică, pretențioasă în stil, ci o discuție în care îmi voi contura o serie de *puncte de vedere* privind IA, luând foarte în serios precauția din subtitlul de pe afișul conferinței.

Este și o intenție polemică aici. Îmi vine în minte o vorbă a lui Richard Feynman, laureat Nobel (pentru cercetările sale în fizica cuantică!): „Cine pretinde că înțelege fizica cuantică este un mincinos”... E oarecum la fel (punctul meu de vedere...) cu IA: cine pretinde că are o imagine completă și corectă a domeniului este un mincinos. Sau un propagandist. Sau un gazetar nevoit să acopere, nu este important cu ce, niște cuadrați în pagina unei publicații. Motivele sunt altele decât în cazul fizicii cuantice, cea mult neintuitivă (voi mai face referiri la ea). IA este „obscură” și greu de descris în alt sens decât numita fizică – voi consuma mai multe pagini cu asemenea aspecte – dar asta nu-i împiedică pe publiciști să scrie din răspuseri, de regulă, anunțând nou-tăți spectaculoase și efecte măcar spectaculoase dacă nu direct nocive-înspre-dezastruoase.

Personal – e ultima mărturisire – cu cât citesc mai mult despre IA, cu cât descopăr mai multe „ascunzișuri”, cu atât mă simt mai

departe de o imagine cu adevărat coerentă, integratoare, asupra domeniului. De aceea, aici propun mai mult întrebări, teme de reflecție, eventual sugestii de abordare a acestora, uneori cu argumente venind dinspre matematică, prezentate informal, doar cu unele atingeri ale jargonului de specialitate.

Un cuvânt-cheie care de multe ori va apărea, explicit sau implicit, în cele ce urmează este *luciditate*. Între optimism necritic, naiv-confortabil, și anxietate necritică, eventual indusă de „catastrofismul” unei bune părți a comentatorilor, este firesc, rațional, înțelept să ne plasăm *in media res*. Să beneficiem acum, aici, de binefacerile IA, să le așteptăm/anticiptăm, pe cele viitoare, dar să rămânem lucizi, precauți, conștienți de posibilele – chiar dacă nu totdeauna și probabile – pericole.

Despre toate acestea, despre (aproape) tot ce va urma – și încă multe pe deasupra – se pot găsi nenumărate articole, chiar și cărți, pe internet. Bibliografia IA este realmente imensă – dar în bună parte și efemeră, pe alocuri nu neapărat de încredere.

2.

Trei idei care merită a fi reținute

Mă grăbesc totuși să formulez aceste trei „concluzii”, care pot călăuzi parcurgerea paginilor următoare și pot ajuta înțelegerea întregii problematice. A nu se confunda aceste idei cu mai multele premise de lucru, metodologice, dintr-o secțiune următoare, foarte importante și acelea, dar de mai mică acoperire, unele putând fi considerate, și ele, concluzii.

a. IA este parte a unui proces mult mai amplu, de durată, acela al *digitalizării*. Folosesc „digitalizare” pentru că nu am un concept mai acoperitor, dar termenul include informatizarea, electronizarea, pătrunderea peste tot a calculatoarelor, a procesoarelor, acoperă conectivitatea, internetul, comunicarea online. Este un proces istoric aici, cu IA fiind vârful aisbergului.

Digitalizarea este considerată cea mai amplă și profundă transformare din istoria civilizației, produsă de om, cu implicații generalizate, globale, ireversibile, care vor schimba semnificativ societatea.

b. Un proces de o asemenea amploare și natură (i se spune uneori „revoluție”, dar păstrez termenul pentru revoluțiile industrial-științifice) va avea și consecințe bune, și consecințe rele (este o discuție lungă aici: cine definește binele și răul, conform căror norme? etc.), ca multe dintre creațiile omului, cele instrumentale, de pildă. Exemplul tipic este cuțitul, cu prefer să iau de exemplu un banal cui: cu el putem prinde o ulucă într-un gard sau putem zgâria o mașină. *Depinde de mâna care folosește cuiul*, de intențiile acestuia! Este la fel cu instrumentele informatice, în particular, cu IA. Idee veche în fond, trimite până la *pharmakon*-ul lui Platon, leac și otravă în același timp, totul fiind chestiune de dozaj, depinzând de cel care prescrie „medicamentul”, de adecvarea la „pacient”. Ne putem gândi și la conceptele orientale *yin-yang*, cu rafinamentele asociate: există yin și în yang și invers, yang în yin, deci rău în bine și bine în rău, cu pasul următor, iterativ, mergând spre

finețe maniacală (în fond, cu tentă de recurență matematică...), dacă mai distingem și binele din răul din bine și așa mai departe.

c. Fenomenul – digitalizarea în general, IA în particular – este în plină desfășurare, în rapidă evoluție, cu două „motoare” principale: (1) motivarea internă, proprie marilor „mișcări” științifico-tehnologice, și (2) piața, banul, interesele. Miza este uriașă. De câteva decenii, IT-ul este probabil principala „afacere” a omenirii – alăturându-se marilor domenii anterioare: alimente, oțel și cărbune, petrol și gaze, medicamente, comerț, armament,... Iar la interesele financiare trebuie adăugate cele geostrategice, pentru că de calculatoare și deci de IA sunt interesate toate compartimentele societății, inclusiv, armatele, serviciile de informații, guvernele.

Referiri la toate acestea vor fi făcute și în continuare, cu unele detalii în plus.

Sumar al demersului

Pentru a avea o imagine de ansamblu a acestui text, a demersului însuși, voi descrie „cuprinsul” lui, cu unele comentarii. După o serie de (1) *Premise de lucru*, capitol deja anunțat, mă voi referi la (2) *Definiția/definițiile IA*, subliniind un detaliu crucial, adesea trecut cu vederea: IA *simulează* IU (inteligența umană), calculatoarele/programele nu sunt inteligente în sensul uzual al cuvântului; de mare importanță este (3) *Plasarea IA în istoria informaticii*, ca moment, nivel într-o evoluție, dar cu rădăcini vechi, mai ales în teorie. În directă legătură, mă voi referi la un număr de (4) *Rezultate matematice*, o parte dintre ele ținând de informatica teoretică și care au o relevanță evidentă pentru IA, stabilind uneori limite ale acesteia, justificând alteori

motive de îngrijorare. Capitolul cel mai așteptat al unei conferințe despre IA este, desigur, cel care privește (5) *Efectele, impactul, consecințele* – cu precădere cele negative; asta, pe de o parte, pentru că sunt mai spectaculoase jurnalistic (ratingul mass-mediei este bazat pe senzational, pe emoții, cu clară aplecare spre zgomot, cataclisme de tot felul...), pe de alta, pentru că „omul când nu înțelege, e contra” (Moisil), pentru că schimbările induc disconfort, neliniște, teamă; anticipez faptul că IA chiar merită a fi considerată o sursă de neliniște. Un capitol pe care îl consider extrem de important, la care țin foarte mult, este (6) *Plasarea IA/digitalizării în istoria civilizației umane*, cu raportare la cele două Renașteri și la cele patru Revoluții industriale. Privind istoria „de sus”, se pot distinge tendințe (*trenduri* ar fi „modern”...), semnificații majore, care nu se văd din interior, „la firul ierbii”. Admit, este un demers speculativ, intuitiv, literaro-filosofic, dar deja „demonstrabil”, dacă privim atent „cum merge lumea” (titlul unei cărți de N. Chomsky, referindu-se la altceva decât IA, desigur). În final, (7) câteva *Concluzii*, fără a insista prea mult în această direcție.