

CUM GÂNDEȘTE INTELIGENȚA ARTIFICIALĂ

Nigel Toon este o personalitate marcantă în domeniul inteligenței artificiale, frecvent invitat să ia cuvântul la evenimentele din domeniu. A fost distins cu numeroase premii, fiind pe locul 1 în clasamentul „Cele mai influente 100 de persoane care au marcat tehnologia britanică” al publicației *Business Insider* și inclus de *Financial Times* în topul „Cei mai buni 100 de antreprenori din Regatul Unit”. Este lider de afaceri în tehnologie, antreprenor și inginer și are funcția de director executiv în trei companii de procesoare finanțate prin capital de risc, printre care Graphcore, companie pe care a cofondat-o și care a dezvoltat un nou tip de microprocesor, special conceput pentru inteligența artificială. Înainte, a fost director executiv al unei companii de semiconductori listate la bursă, cu sediul în Silicon Valley.

A fost membru în consiliile de administrație și președinte al mai multor companii din domeniul tehnologiei și este în prezent director neexecutiv în consiliul UK Research and Innovation (UKRI), organism public independent din Marea Britanie care direcționează anual peste opt miliarde de lire sterline pentru finanțarea cercetării și a inovației. De asemenea, a fost membru în consiliul economic al prim-ministrului Marii Britanii.

Nigel Toon este deținătorul a trei brevete de invenție.

NIGEL TOON

CUM GÂNDEȘTE INTELIGENȚA ARTIFICIALĂ

CUM AM CREAT-O, CUM NE POATE AJUTA
ȘI CUM O PUTEM CONTROLA

Traducere din engleză și note
de Iulian Comănescu

HUMANITAS

Redactor: Luisa Neag
Coperta: Angela Rotaru
Tehnoredactor: Manuela Măxineanu
Corector: Bogdan Nicolau
DTP: Iuliana Constantinescu, Dan Dulgheru

Tipărit la Bookart Printing

Nigel Toon

How AI Thinks: How we built it, how it can help us, and how we can control it

Copyright © Nigel Toon, 2024

All rights reserved.

© HUMANITAS, 2026, pentru prezenta versiune în limba română

Descrierea CIP a Bibliotecii Naționale a României

Toon, Nigel

Cum gândește inteligența artificială: cum am creat-o,
cum ne poate ajuta și cum o putem controla / Nigel Toon;
trad. din lb. engleză și note de Iulian Comănescu. –

București: Humanitas, 2026

ISBN 978-973-50-9013-5

I. Comănescu, Iulian (trad.; note)

004

EDITURA HUMANITAS

Piața Presei Libere 1, 013701 București, România

tel. 021/408 83 50, fax 021/408 83 51

www.humanitas.ro

Comenzi online: www.libhumanitas.ro

Comenzi prin e-mail: vanzari@libhumanitas.ro

Comenzi telefonice: 0723 684 194

Pentru Sally

Cuprins

Introducere	9
I. Cum a devenit posibilă inteligența artificială	
1. Revoluția IA a început deja	23
2. Mașinile inteligente	32
3. Nașterea inteligenței artificiale	44
4. Tehnologia care a creat IA – Partea I: Computerele electronice și cum au învățat acestea să vadă	52
5. Tehnologia care a creat IA – Partea a II-a: Semiconductorii, software-ul și atenția	60
6. Tehnologia care a creat IA – Partea a III-a: Creativitate, conectare și informație	82
7. Claude Shannon, părintele Erei Informației	102
II. Cum diferă IA de inteligența umană?	
8. Ce este inteligența?	113
9. Mai multă inteligență	118
10. Conștiința	136
11. O mașină ultrainteligentă	152
12. Oare singularitatea este cu adevărat posibilă?	158
III. Oare înțelegem cu adevărat incredibilul potențial al IA?	
13. IA și mediul	173
14. IA în educație	182
15. Revoluția IA în sănătate	187
16. Provocările IA	205

17. IA responsabilă	239
18. Cum au loc revoluțiile tehnologice.	251
19. Cum putem face ca IA să lucreze în favoarea noastră.	261

<i>Glosar.</i>	277
--------------------------	-----

<i>Mulțumiri.</i>	283
-----------------------------	-----

<i>Credite foto</i>	285
-------------------------------	-----

<i>Note.</i>	287
------------------------	-----

Introducere

*Ceea ce ne dorim este o mașină care poate învăța din experiență.*¹ — Alan Turing, matematician și deschizător de drumuri în știința informației, 1947

Ne aflăm la începutul unei revoluții care va zgudui lumea: revoluția inteligenței artificiale. IA (inteligența artificială) este cel mai puternic instrument creat vreodată de omenire, iar noi de-abia începem să aflăm de ce este în stare. Ea are capacitatea de a extinde extraordinara inteligență umană, ajutându-ne să rezolvăm probleme complexe care până acum păreau fără soluție. Dar IA aduce cu ea și noi provocări. Ne așteaptă schimbări profunde. Mi-am petrecut viața în avangarda tehnologiei, am fost implicat direct în dezvoltarea tehnologiilor care stau la baza IA, așa că înțeleg cum vor evolua lucrurile. Vreau să te ajut să înțelegi cum gândește inteligența artificială.

Pentru mine, totul a început când aveam vreo opt ani; tata a adus acasă o parte dintr-un calculator vechi, gândindu-se că aş putea fi interesat. Ca inginer electronist talentat, care participase la crearea unora dintre primele computere, în anii 1950, tata a știut să-mi arate cum trebuia să apeși pe butoanele de pe vechea placă de comutare pentru a programa acel gen de mașină primitivă. Mi s-a părut fascinant că un calculator nu poate înțelege direct litere sau cuvinte. În adâncul circuitelor electronice, el folosește numere – dar niște numere ciudate. Nu folosește sistemul zecimal cu care suntem obișnuiți, ci unul mai aparte, numit „binar”^{*}. Am început să înțeleg că un calculator este foarte diferit

^{*} Sistemul de numerație binar are baza 2 și este compus din cifrele 0 și 1. Avantajul lui în materie de computere este că „0” și „1” corespund stărilor „închis” și „deschis” ale circuitelor (n. tr.).

de o ființă umană. De atunci mă fascinează modul în care funcționează calculatoarele și, de asemenea, întrebarea dacă vom putea să construim vreodată un computer care să poată fi numit cu adevărat inteligent. Noi ne gândim la computere ca la niște mașini de calcul foarte complexe, care urmează întocmai instrucțiunile dintr-un program, pas cu pas. Iată însă că aceste mașini puternice încep acum să învețe din informație.

Chiar acum, când citești această carte, ești bombardat cu o cantitate uriașă de informație care vine spre tine din toate părțile. Te întrebi probabil dacă scaunul pe care l-ai ales e suficient de confortabil sau dacă nu cumva ai pus prea mult lapte în ceai. Poate că te distrage priveliștea de la fereastră sau zgomotele de fond. Creierul tău poate mai vibrează după conversația pe care tocmai ai avut-o și e posibil să încerci să-ți amintești ce vei face săptămâna viitoare. Foarte probabil, te sâcăie și faptul că ai uitat să faci ceva, deci nu ar trebui să te apuci de citit. Creierul încearcă să filtreze toate aceste informații disparate, să decidă dacă vreuna e utilă, dacă există vreo legătură între ele, dacă trebuie să-ți minte ceva și ce anume poți ignora – asta în timp ce încearcă să-ți mențină atenția concentrată asupra acestui text.

De-a lungul unei zile, corpul tău arde aproximativ 2.000 de kilocalorii, ceea ce înseamnă, în medie, cam 100 de wați pe oră. Este cam aceeași putere necesară pentru a aprinde un bec cu filament din generația veche sau zece becuri LED moderne. Dacă pedalezi pe o bicicletă de fitness, s-ar putea să atingi un vârf de 200 sau chiar 250 de wați – doar dacă nu ești ciclist de performanță, ca triplul câștigător al Turului Franței Chris Froome, care depășește 400 de wați și îi menține timp de 30 de minute.² În timpul somnului, consumul tău de energie scade la aproximativ 25 de wați, iar acum, când stai și citești această carte, consumi undeva în jur de 50 de wați – din care aproape jumătate sunt consumați de creier. Asta înseamnă că trebuie să consumi aproximativ 500 de kilocalorii – adică echivalentul a două batoane de ciocolată sau cinci mere – ca să ajungi până la ultima pagină. Ești pe cale să transformi hrana în cunoaștere.

În paginile acestei cărți, sper să-ți pot transmite informații utile. După cum voi încerca să arăt, informația poate duce la sporirea cunoașterii, iar aceasta, la rândul ei, ne face mai inteligenți. Comunicarea este esențială pentru a ne clădi inteligența, însă informația care ne poate fi de folos poate apărea în orice moment. De multe ori va trebui s-o păstrăm pentru a o folosi mai târziu. Uneori e nevoie, de asemenea, să o generalizăm, astfel încât să o putem aplica în contexte diferite. Se dovedește astfel că memoria este, de fapt, o formă decalată în timp a comunicării, care ne oferă informații utile exact atunci când avem mai mare nevoie de ele. Ai auzit ceva interesant acum șase luni, ai păstrat acea informație undeva în memorie, iar momentul în care îți este de folos – poate pentru a i-o transmite altcuiva – tocmai a sosit. Comunicarea, informația și cunoașterea care decurge din ele contribuie împreună la construirea inteligenței. La început, oamenii și-au folosit inteligența pentru a-și procura hrana; astăzi, o folosim pentru a ne îmbunătăți viața în nenumărate alte moduri.

Inteligența artificială se bazează, la rândul ei, pe comunicarea informației – un lucru devenit mult mai ușor odată cu apariția internetului. IA are nevoie de unele dintre cele mai avansate și specializate tipuri de calculatoare, care au devenit accesibile pentru noi datorită progreselor extraordinare din domeniul cipurilor semiconductoareⁱ. Mai are nevoie și de oameni, care să inventeze metode noi prin care aceste computere performante să învețe din informația digitală. Rezultatul: o mașină inteligentă, capabilă să recunoască obiecte, să genereze imagini care n-au existat niciodată înainte și chiar să poarte o conversație cu tine.

Volumul tot mai mare de cunoaștere acumulată ne-a dat inteligența necesară pentru a construi unelte și mașini avansate, pe

ⁱ Conductorii, cum sunt metalele, permit trecerea curentului electric, în timp ce izolatorii, cum sunt materialele plastice, blochează curentul. Semiconducătorii sunt un tip special de material care poate permite sau bloca trecerea curentului electric, în funcție de starea lor. Siliciul este un exemplu de material semiconducător și este, de fapt, unul dintre cele mai răspândite materiale de pe Pământ, pe care îl vom găsi sub formă de nisip pe fiecare plajă de la malul mării (n. a.).

care le folosim zilnic pentru a ne face viața mai ușoară. Ne pricepem la asta – motoarele pe care le-am creat ne sporesc forța fizică, permițându-ne să muncim mai mult, să ne deplasăm mai repede și chiar să zburăm. Dar acum, cu noile metode bazate pe IA, construim mașini capabile să învețe ele însele din informație și să acumuleze cunoaștere, care, la rândul ei, ne va ajuta să ne creștem propriul nivel de inteligență.

Inteligența artificială este cel mai puternic instrument creat vreodată de om. Capacitatea ei de a amplifica inteligența umană ne va permite să rezolvăm probleme care până nu demult păreau imposibile. IA cu adevărat performantă a început însă să se contureze de-abia în ultimii ani. Ea reprezintă o abordare cu totul nouă în domeniul computerelor. În anii care urmează, aceste noi mașini inteligente ne vor transforma viața – de fapt, transformarea a și început.

Pe măsură ce IA a prins contur, această tehnologie uimitoare a devenit subiect de conversație la cine festive și în dezbaterile publice. Așa-ziși experți ies în față pentru a-și exprima opiniile categorice cu privire la dificultățile și obstacolele pe care această nouă forță inteligentă le-ar putea aduce. Dar inteligența artificială este un domeniu complex și, la fel ca alte instrumente puternice care au apărut înaintea ei, neînțelegerea ei riscă să ducă la convingerea că ar putea reprezenta o amenințare la adresa umanității. Când avem de-a face cu o tehnologie nouă, suntem înclinați să ridicăm din umeri și să spunem că pur și simplu nu o înțelegem. Avem nevoie să aflăm mai multe – nu doar despre riscuri, ci și despre beneficiile pe care această tehnologie ni le poate aduce, fiindcă ea ne va schimba viața și, poate chiar mai important, va schimba viața generației următoare.

După ce am văzut prima oară un computer, în adolescență, am ajuns să-mi construiesc propriul calculator folosind un microprocesor rudimentar și eram fascinat de felul în care funcționa. 40 de ani mai târziu, am pus bazele unei companii care a construit un nou tip de microprocesor, conceput pentru calcule mai rapide legate de inteligența artificială. De-a lungul carierei mele, am avut

norocul să lucrez în avangarda domeniilor semiconductorilor, calculatoarelor, comunicațiilor și inteligenței artificiale. Am avut acces în consiliile directoare ale celor mai mari companii din industria semiconductorilor, am vizitat cele mai importante fabrici de cipuri din Taiwan și Coreea de Sud, am lucrat cu firme implicate în sisteme de comunicații optice, care transmit internetul peste tot în lume, precum și în dezvoltarea tehnologiei pentru smartphone-uri, care ne pun internetul în buzunar. Am vizitat laboratoare de supercalculatoare în Marea Britanie și în Statele Unite, iar împreună cu prietenul meu Simon Knowles, cu care am înființat o companie, i-am cunoscut pe mulți dintre cei care promovează dezvoltarea IA. Printre aceștia se numără figuri proeminente precum Demis Hassabis, fondatorul companiei de cercetare DeepMind; savanți de prestigiu precum Geoffrey Hinton și Yann LeCun, ambii laureați ai Premiului Turing pentru contribuțiile lor revoluționare în domeniul IA; profesorul de informatică Andrew Ng, de la Universitatea Stanford, fost conducător al proiectului Google Brain; în fine, profesorul Jürgen Schmidhuber, unul dintre cei mai importanți inovatori în domeniul IA din ultimele trei decenii. În discuțiile pe care le-am avut cu acești experți – și cu mulți alții –, am primit de fiecare dată informații esențiale, dar, pentru a căpăta o înțelegere mai amplă, am început să explorez în detaliu domeniul neuroștiinței și al sistemelor biologice – ambele servind drept surse de inspirație pentru IA.

Foarte multă lume îmi pune întrebări despre inteligența artificială: prieteni, membri ai familiei, medici, avocați, lideri de business, funcționari publici, politicieni, prim-ministrul Marii Britanii și chiar membri ai familiei regale britanice. Mulți îmi spun că nu înțeleg IA, că îi sperie. Gândul că o mașină ar putea să ia decizii și să raționeze într-un mod considerat exclusiv uman îi face să se teamă de ceea ce ar putea deveni IA în viitor. Lipsa de înțelegere ne face să ne fie frică de ea.

Am scris această carte tocmai pentru a răspunde acestor temeri. Ea ne va oferi cunoașterea necesară pentru a profita de ocaziile extraordinare create de IA. Pentru asta, voi încerca să arăt cum

gândește IA și să descriu tehnologiile care au făcut-o posibilă. Vom explora modurile în care IA ne poate ajuta, dar și măsurile de control pe care va trebui să le luăm. De asemenea, voi explica în ce fel inteligența mașinilor diferă de inteligența noastră umană – o inteligență cu totul specială.

Teama față de mașini este firească. În timpul proiecțiilor-test ale filmului de animație *Shrek*, mulți copii au fost foarte speriați când au văzut primele versiuni digitale ale personajului feminin principal, prințesa Fiona. Era prea realistă, iar copiii începeau să plângă de îndată ce apărea pe ecran. Fenomenul a fost observat pentru prima dată de Masahiro Mori, profesor de robotică din Japonia, sub forma unui paradox pe care l-a numit „valea ciudăteniei” (*uncanny valley*).³ Mori a arătat că, atunci când roboții sau personajele animate sunt făcute să semene puțin cu oamenii, reacția inițială este pozitivă – ni se par simpatici. Însă, pe măsură ce robotul sau animația devin mai realiste, acea senzație pozitivă se poate transforma brusc în neliniște: personajul devine straniu, tulburător. Nu mai vedem un căpcăun care exprimă umanitate, ci un om transformat în căpcăun. Acest efect este adesea exploatat și de regizorii de filme de groază, unde niște personaje aparent inofensive – păpușile, de pildă – sunt înzestrate cu trăsături profund umane, cum se întâmplă în filmele *Linște mortală*, *Annabelle* sau *The Boy*. Multe scenarii de filme distopice se bazează, la rândul lor, pe această teamă a oamenilor, punând în centru o IA posesoare de conștiință, care încearcă să preia controlul. Această relație dificilă dintre oameni și mașinile umanizate se află, poate, la originea fricii noastre față de IA. Dacă am ști mai multe, probabil am înțelege că, asemenea oricărei alte mașini, IA nu este decât un instrument care urmează o metodă trasată de om, are limite în ceea ce poate face și se află sub controlul nostru.

Un alt aspect, mai subtil, ține de date și de felul în care învățăm. Datele, luate izolat, nu ne ajută cu nimic. Fără context, ele nu au sens. O imagine din telefonul nostru este doar un grup de puncte până când înțelegem că acele puncte redau, de fapt, o pisică. Cei mai mulți dintre noi știu cum arată o pisică; această

cunoaștere există deja în neuronii noștri – dar cum a ajuns ea acolo? Undeva, în primii ani de viață, poate chiar la prima întâlnire cu o pisică, cineva ne-a spus: „Aceasta este o pisică“. Acea descriere ne-a oferit contextul necesar pentru a identifica creatura necunoscută. Contextul transformă *datele* în *informație*. După ce am mai văzut câteva pisici, am putut generaliza diversele informații și am început să formăm asocieri care ne-au ajutat să înțelegem mai profund ce este o pisică. Prin acest proces de învățare din informație am ajuns să ne construim *cunoașterea*. Iar cu această cunoaștere am început să ne etalăm noua *intelență*, arătând cu degetul o pisică de fiecare dată când vedeam una.

După cum spune celebrul cântec al lui Cole Porter: „Păsările o fac, albinele o fac, chiar și mașinile inteligente o fac...“ Ei bine, nu chiar – iar calculatoarele, în mod sigur, nu se îndrăgostesc –, dar, așa cum voi arăta, animalele, plantele și chiar organismele microbiene urmează același proces organic: folosesc date și context pentru a extrage informație. Iar această informație le permite să-și construiască o bază de cunoaștere din care își dezvoltă inteligența.

Așa se face că, în loc să-i spună calculatorului ce să facă pas cu pas, într-un program, oamenii au inventat metode eficiente care-i permit acestuia să învețe din informație. Informația este folosită pentru a construi un model cognitiv, iar din acest model mașina începe să ia decizii care exprimă ceea ce numim „intelență artificială“.

Pentru a înțelege cum „gândește“ IA, am încercat să scriu pe înțelesul oricui are această curiozitate. Vreau ca acest domeniu complex să fie accesibil și celor care nu au neapărat cunoștințe vaste de matematică, știință sau tehnologie. Este modul meu de a împărtăși o parte din cunoașterea pe care am acumulat-o, în speranța că ea va contribui la o înțelegere mai profundă a schimbărilor importante pe care IA le va aduce în viața noastră. Sper ca efortul de a înțelege cea mai transformatoare tehnologie a epocii noastre să ne deschidă drumul către posibilități extraordinare.

Dar până acolo, avem nevoie în primul rând de o înțelegere de bază a inteligenței artificiale. Vom vorbi și despre inteligența