EDITURA SCRIPTUM®

2084

ȘI REVOLUȚIA AI

Descrierea CIP a Bibliotecii Naționale a României
LENNOX, JOHN C.
2084 și revoluția AI / John C. Lennox.— Oradea: Scriptum, 2025
ISBN 978-606-031-371-7

1
004

Aprecieri pentru această carte

Dacă intenționezi să citești o singură carte despre inteligența artificială, aceasta este cartea potrivită. Dacă ai de gând să citești mai multe cărți despre AI, aceasta ar trebui să fie prima. John Lennox oferă o bază largă de informații în care se regăsesc elementele de bază, pericolele, actorii corporativi și individuali, modelele de limbaj mari în contrast cu învățarea automată și supravegherea umană, economia, etica și modul în care toate acestea se intersectează cu o viziune biblică asupra lumii. Totul este prezentat aici într-un format deosebit de accesibil și plăcut pentru cititorul obișnuit sau pentru utilizatorul ocazional de AI.

— James Tour, profesor de chimie,
nanoinginerie și informatică, Universitatea Rice

👍

O altă carte strălucită de John Lennox, care reușește să găsească echilibrul potrivit între beneficiile și pericolele AI. El îți va arăta care este Sursa oricărei inteligențe și unde o poți găsi înainte să fie prea târziu.

— Frank Turek, autor și conferențiar

👍

Ce știe AI? Ce speră AI? Și ce va face AI? Sam Altman nu are răspunsul la aceste întrebări; John Lennox, da.

— Perry Marshall, autorul cărții *Evolution 2.0: Breaking the Deadlock between Darwin and Design*

John Lennox oferă cele mai actualizate informații despre cum ne influențează inteligența artificială locurile de muncă, relațiile și corpurile. Această ediție extinsă a cărții *2084* este atât profundă, cât și foarte ușor de citit, cu întrebări importante pentru discuții de grup la finalul fiecărui capitol. Oricine ține la viitorul umanității ar trebui să citească această carte.

— Elaine Howard Ecklund, autoarea cărții
Why Science and Faith Need Each Other;
titulara catedrei Herbert S. Autrey
în Științe Sociale, Universitatea Rice

AI, metavers, transumanism – oricât de multe crezi că știi despre aceste lucruri, analiza amplă și atent documentată a tehnologiilor viitoare realizată de John Lennox îți va arunca în aer bagajul de cunoștințe!

— Douglas Estes, autorul cărților *SimChurch* și *Braving the Future*;
profesor asociat de religie, New College of Florida

Această ediție revizuită a cărții *2084* este o adevărată performanță intelectuală, care reușește să adune laolaltă, cu o claritate, înțelepciune și compasiune de neegalat, cunoștințe pe înțelegerea publicului larg cu privire la elementele de bază ale tehnologiei inteligenței artificiale, principiile eticii aferente și relația dintre AI și om. Renumitul savant și apologet creștin John Lennox este un adevărat far care ne ajută să navigăm prin complexitatea progresului tehnologic cu o profunzime spirituală ce ne inspiră să reflectăm asupra a ceea ce înseamnă să fii om – acum și în viitorul nepoților noștri.

— Dr. Michael Barrett, prodecan și profesor
de Sisteme Informaționale și Inovație
la Cambridge Judge Business School
și membru al Hughes Hall, University of Cambridge

Un avertisment puternic, dintr-o perspectivă creștină, despre potențialele riscuri ale unei distopii AI – inechitatea algoritmică, amenințarea existențială, aroganța transumanistă.

— Jeremy Gibbons, profesor de informatică,
University of Oxford

2084

ȘI REVOLUȚIA AI

CUM NE INFLUENȚEAZĂ INTELIGENȚA ARTIFICIALĂ VIITORUL

JOHN C. LENNOX

Editura Scriptum®
Oradea

Originally published in U.S.A. under the title
2084 and the AI Revolution, Updated and Expanded Edition
by Thomas Nelson, a registered trademark of
HarperCollins Christian Publishing, Inc., Nashville, TN

This edition published by arrangement with
HarperCollins Christian Publishing, Inc.

Ediția în limba română, publicată cu permisiune, sub titlul:
2084 și revoluția AI
de John C. Lennox

str. Simion Bărnuțiu nr. 2, 410204 Oradea – Bihor
Tel./Fax/Robot: 0259-457.428
E-mail: scriptum@scriptum.ro
Pagina web: WWW.SCRIPTUM.RO

Prima ediție în limba română.

Cu excepția situațiilor când se specifică altfel, toate citatele biblice folosite sunt din NTR - Noua traducere în limba română. Alte traduceri ale Bibliei folosite: NIV - New International Version; NKJV - New King James Version.

ISBN (tipărit) - 978-606-031-371-7
ISBN (e-book) - 978-606-031-372-4

Tuturor nepoților, printre care și celor zece nepoți ai mei
Janie Grace, Herbie, Sally, Freddie, Lizzie, Jessica, Robin,
Rowan, Jonah și Jesse – cu speranța că îi va ajuta
să facă față provocărilor unei lumi dominate de AI.

CUPRINS

PARTEA A III-A
PREZENTUL ȘI VIITORUL AI

PARTEA A IV-A
CE ESTE O FIINȚĂ UMANĂ?

PREFAȚĂ LA EDIȚIA REVIZUITĂ ȘI ADĂUGITĂ

Progresele din domeniul inteligenței artificiale (AI) din ultimii patru ani, de când am scris prima ediție a acestei cărți, justifică revizuirea textului și, în plus, impun nevoia unei ediții adăugite în mod considerabil, cu scopul de a ține pasul cu un fenomen fără precedent din această primă parte a secolului al XXI-lea. AI este pe buzele tuturor în zilele noastre, ca un termen tehnic la modă, ba chiar a constituit tema principală la Forumul Economic Mondial (WEF) de la Davos din 2024. În 2023, 25% dintre start-upurile lansate au fost din domeniul AI și se așteaptă ca investițiile pe plan internațional să se ridice la suma de 200 miliarde de dolari americani până în 2025. WEF consideră că până la 40% dintre joburile din toată lumea vor fi afectate de AI, într-un fel sau altul, multe fiind înlocuite de aceasta, iar altele, îmbunătățite. Articolele de cercetare care au ca temă AI au cunoscut o creștere exponențială, dublându-se la fiecare doi ani; în 2021 erau 4.000 de articole noi pe lună pe arXiv! Este în mod evident imposibil să ținem pasul cu un domeniu aflat într-o astfel de dezvoltare explozivă, dar îndrăznesc să sper, totuși, că cititorii mei vor rămâne cu impresia că, cel puțin, am reușit să redau în mod corect ce s-a întâmplat de când am scris prima ediție a cărții, în 2020. Scopul meu este să realizez două lucruri: în primul rând, să-i informez pe cititorii

mei și, în al doilea rând, să explorăm și să reflectăm asupra posibilelor implicații ale acestui nou domeniu revoluționar al tehnologiei asupra noastră.

În cadrul unei dezbateri din 2009, biologul E. O. Wilson spunea: „Adevărata problemă a omenirii este următoarea: avem niște emoții paleolitice, niște instituții medievale și o tehnologie divină. Și asta este grozav de periculos." Și a continuat: „Până nu răspundem la marile întrebări ale filosofiei pe care filosofii le-au abandonat cu două generații în urmă – *De unde venim? Cine suntem? Încotro ne îndreptăm?* – rațional vorbind, suntem pe un teren foarte șubred."[1] Eu cred că Wilson este un pic prea dur cu filosofii. Cunosc câțiva oameni excepționali, și nu doar filosofi, care se gândesc la aceste chestiuni de multă vreme. În lucrarea sa, *Critica rațiunii pure*, filosoful german Immanuel Kant spune că cele mai importante trei întrebări pentru orice ființă umană sunt: Ce pot să știu? Ce pot să sper? Și ce trebuie să fac? Acestea sunt cele trei întrebări existențiale esențiale pe care toți ni le punem în încercarea noastră de a găsi un sens vieții. Cartea de față încearcă să le abordeze în contextul AI.

Apar și nenumărate alte întrebări înrudite. Care este relația dintre minte și creier? Inteligența reală este legată întotdeauna de conștiință? Și, în fond, ce este conștiința? Vom fi capabili să construim o conștiință și o viață artificiale? Se vor transforma oamenii pe ei înșiși în așa măsură încât să devină cu totul altceva – fie prin inginerie genetică, fie prin tehnologia de tip cyborg, fie prin ambele? Vom ajunge, într-un final, să producem superinteligențe? Și, dacă da, care va fi relația noastră cu astfel de entități? Le vom controla, vom lucra împreună cu ele, sau vom fi controlați sau chiar înlocuiți sau distruși de ele? Sau vom fuziona treptat cu mașinile? Este viața în metavers un pas în direcția aceasta, orice s-ar dovedi că este metaversul până la urmă?

Deși existența unei AI superinteligente rămâne deocamdată la nivel de speculație, este din ce în ce mai puternică opinia că trebuie să-i luăm în serios potențialul și că, pentru a ne pregăti pentru apariția acesteia, ar trebui să investim masiv în prevenirea potențialelor riscuri existențiale. Asta înseamnă să ne gândim la cât de important este să ocrotim viața umană în fața prețului plătit pentru îmbrățișarea AI,

o perspectivă cunoscută sub numele de *longtermism*[*]. Toate acestea ridică, cu o și mai mare urgență, întrebarea esențială: Care este valoarea unei ființe umane?

Sper că titlul pe care l-am ales nici nu sună pretențios, nici nu induce în eroare, de vreme ce eu nu sunt George Orwell, autorul romanului distopic *1984*, din care este inspirat titlul. Titlul mi-a fost sugerat pentru prima dată de un coleg de la Oxford, Peter Atkins, profesor de chimie fizică, în timp ce ne îndreptam spre universitate, unde urma să participăm, de pe poziții opuse, la o dezbatere cu tema: „Poate știința să explice totul?". Îi sunt recunoscător pentru sugestia sa, dar și pentru alte discuții publice viguroase pe teme legate de știință și Dumnezeu. Sunt, de asemenea, recunoscător multor altor persoane, în special lui Rosalind Picard de la MIT Media Lab pentru comentariile sale foarte perspicace. Printre alții se numără și David Cranston, Danny Crookes, Jeremy Gibbons, David Glass și fostul meu asistent de cercetare, Simon Wenham. Datorez o mare recunoștință colegiului meu de la Oxford, Green Templeton, pentru că mi-a oferit un mediu academic în care am reușit să prosper în ultimii douăzeci și cinci de ani și unde au fost concepute și puse la încercare multe dintre ideile mele prin discuții intense.

Am consultat, de asemenea, o gamă largă de lucrări și doresc să menționez în acest punct două cărți. Prima este introducerea cuprinzătoare și autorizată în domeniu, intitulată *Artificial Intelligence: A Modern Approach* [Inteligența artificială: O abordare modernă], scrisă de Stuart Russell și Peter Norvig. A doua este *Masters or Slaves? AI and the Future of Humanity* [Stăpâni sau sclavi? AI și viitorul omenirii], scrisă de Jeremy Peckham. Această carte reprezintă o introducere foarte accesibilă în domeniu, fiind scrisă de un cercetător și antreprenor experimentat, care a lucrat în domeniul inteligenței artificiale timp de mulți ani și a fondat mai multe companii *high-tech* de succes. Cititorii care doresc să consulte un raport autorizat despre starea actuală a inteligenței artificiale sunt îndrumați către cel mai recent raport *Artificial Intelligence Index Report* de la Stanford University.[2]

[*] De la *long-term*, în engleză, „pe termen lung, în perspectivă" (n.tr.).

Pregătirea mea profesională este în matematică și filosofia științei, nu sunt specialist în inteligența artificială, deși mă interesează de mulți ani filosofia tehnologiei și implicațiile ei. Cititorii mei, în special cei care sunt experți în AI, ar putea fi surprinși de aparenta mea intruziune în domeniul lor. Intenția mea este însă diferită, deoarece există diferite niveluri de implicare și relaționare cu AI. În primul rând, există gânditorii pionieri, iar apoi experții care scriu efectiv software-ul utilizat în sistemele de inteligență artificială. Apoi avem inginerii responsabili de partea de hardware. Mai sunt și cei care lucrează la dezvoltarea de aplicații noi. În sfârșit, există și scriitori ca mine, dintre care unii au pregătire de natură științifică, care sunt interesați de semnificația și impactul AI pe diferite planuri – filosofic, politic, practic, sociologic, cultural, economic și etic. Senzația mea este că oricine lucrează cu AI are nevoie de o perspectivă mai largă, indiferent de nivelul lor de implicare și m-am simțit încurajat de un articol peste care am dat în *The Atlantic*, scris de Fei-Fei Li, informatician și pionier în domeniul AI, care susținea valoarea universității: „A sosit vremea să reevaluăm modul în care se predă disciplina AI la toate nivelurile. Specialiștii din anii ce vor veni vor avea nevoie de mult mai mult decât de o expertiză tehnologică; vor avea nevoie să înțeleagă filosofie, etică, chiar și principii juridice. Și cercetarea va trebui să evolueze."[3] Pentru așa ceva există universitățile și sunt recunoscător Universității mele Oxford pentru libertatea pe care mi-a acordat-o de a cerceta temele prezentate în cartea aceasta.

Este clar că nu trebuie să știi cum să construiești vehicule sau arme autonome pentru a avea o opinie informată despre oportunitatea sau moralitatea utilizării lor. De asemenea, nu este necesar să știi cum să programezi și să implementezi un sistem AI de monitorizare a achizițiilor pentru a avea o opinie validă despre încălcarea vieții private și controlul exercitat de motoarele de supraveghere online, deoarece acestea te afectează direct. A devenit evident că fundamentul etic nu a ținut pasul cu dezvoltarea tehnologică. Nici reflecția filosofică, urmarea fiind că suntem din ce în ce mai puțin siguri cu privire la ce fel de viitor creăm – sau măcar ce fel de viitor ne dorim. Această situație nu este câtuși de puțin sănătoasă și, pentru remedierea ei,

dacă ea este într-adevăr posibilă, va fi nevoie de o gândire profundă și acțiuni curajoase.

Există, așadar, un interes generalizat pentru înțelegerea implicațiilor revoluției AI – așa-numita a patra revoluție industrială – la nivelul înțelegerii științei de către public. La acest nivel am plasat această carte și sunt profund recunoscător tuturor experților care au contribuit la îmbogățirea înțelegerii mele asupra subiectului prin scrierile și prelegerile lor și prin discuțiile pe care le-am avut cu ei.

Ori de câte ori a fost posibil, am furnizat referințe către cărți, articole, pagini web și alte resurse, astfel încât cititorii interesați să poată urmări sursele mele.

CUM SĂ CITEȘTI ACEASTĂ CARTE

Cartea aceasta cuprinde mult material, care nu prezintă în totalitate același interes pentru toți cititorii. Nu este important să fie citită în ordinea propusă, așa că voi prezenta mai jos un rezumat scurt al cărții după care cititorii pot decide de unde doresc să înceapă sau ce capitol să aprofundeze.

PARTEA I: CARTOGRAFIEREA TERITORIULUI

Capitolul 1 explorează niște romane distopice celebre, dintre care unele, precum *1984* al lui Orwell, descriu state totalitare sumbre, dominate de tehnologie și supraveghere în masă. Acest lucru ne conduce la inteligența artificială, care este adesea folosită astăzi pentru supraveghere. Apoi, urmărim pe scurt istoria tehnologiei informației, ajungând la munca inovatoare a genialului matematician Alan Turing, părintele computerului, ale cărui idei stau la baza eforturilor de a construi o „mașină gânditoare". De aici, introducem conceptele de inteligență artificială restrânsă și generală, menționând că cea din urmă nu apare doar în science-fiction, ci și din ce în ce mai mult în gândirea unor oameni de știință și ingineri

de top, care depun eforturi serioase să o construiască în zilele noastre.

Capitolul 2 aprofundează subiectul AI, cu discuția despre natura inteligenței și despre ce se presupune că ar însemna inteligența mașinii. Apoi trecem rapid în revistă istoria și creionăm punctele pozitive și cele negative ale agendei AI care are ca scop construirea de mașini care să imite ce pot face oamenii. Introducem idei precum rețele neuronale, algoritmi și *machine learning* (învățare automată). În final, oferim câteva exemple de AI de care ne folosim în viața de zi cu zi.

Capitolul 3 începe cu elemente de etică. Așa cum un cuțit poate fi folosit și pentru a comite o crimă, și pentru a efectua o intervenție chirurgicală, la fel și toate formele de tehnologie dau naștere unor probleme de natură etică. AI nu face excepție – într-adevăr, unele dintre problemele etice care ne stau înainte azi sunt și complicate, și urgente, având în vedere impactul pe care îl au asupra vieților multor oameni. Acestea cuprind o gamă foarte variată, de la invadarea vieții private prin metode de supraveghere până la înșelăciuni comise de asistenți digitali. Enumerăm câteva sisteme etice cunoscute folosite pentru a determina care sunt principiile etice cu care ar trebui să fie prevăzute sistemele de inteligență artificială și reglementarea lor (de exemplu, principiile de la Asilomar). De asemenea, analizăm îndrumările etice pentru roboți. Capitolul se încheie cu observația că, din cauza vocabularului folosit în domeniul inteligenței artificiale – „inteligență artificială", „învățare automată/*machine learning*", „învățare profundă/*deep learning*" și altele asemenea – este foarte ușor să cădem în capcana ideii că suntem pe cale să construim mașini care gândesc la fel ca oamenii. Mașinile nu au minți și nu pot percepe lucruri, iar în încheiere citățm lucrarea recentă și fascinantă a neurologului Iain McGilchrist despre diferitele moduri de percepție folosite de cele două emisfere ale creierului uman.

PARTEA A II-A: DOUĂ ÎNTREBĂRI FUNDAMENTALE: DE UNDE VENIM? ÎNCOTRO NE ÎNDREPTĂM?

Pentru a atrage un public cât mai larg, analizăm aceste întrebări din perspectiva a doi autori foarte diferiți. Capitolul 4 abordează prima

întrebare – *De unde venim?* – în contextul științei și al tehnologiei, așa cum sunt prezentate ideile pe această temă în cartea *Origini* a popularului scriitor de ficțiune Dan Brown. Vom examina ideile sale pentru a le evalua credibilitatea și veridicitatea lor științifică.

Capitolul 5 explorează a doua întrebare – *Încotro ne îndreptăm?* – având în vedere ideea lui Brown despre fuziunea dintre oameni și mașini, un concept aflat pe agenda unor academicieni de renume, precum astronomul Martin Rees, Membru al Casei Lorzilor, și regretatul fizician Stephen Hawking. Acesta este obiectivul principal al transumanismului – impulsul de a modifica ființa umană și de a crea o superinteligență prin bioinginerie și inginerie cibernetică. Discutăm despre conceptul de „singularitate" al lui Ray Kurzweil și originile sale în ce privește ideea unei mașini „ultrainteligente". De asemenea, prezentăm gândirea actuală a unui spectru mai larg de oameni de știință și ingineri, precum și așteptările lor cu privire la un viitor tehnologic care ne va influența viețile.

PARTEA A III-A: INTELIGENȚELE ARTIFICIALE ȘI MODUL ÎN CARE NE INFLUENȚEAZĂ VIEȚILE – ÎN PREZENT ȘI ÎN VIITOR

Capitolul 6 începe prin a enumera câteva dintre modurile familiare în care inteligența artificială ne influențează astăzi într-o varietate de activități și domenii – de la e-mailuri, asistenți digitali, robotică, vehicule autonome, producție și medicină, până la utilizarea triumfătoare a AI în rezolvarea unora dintre cele mai dificile probleme practice din știință, cum ar fi plierea proteinelor. Cercetarea avansează rapid și promite un viitor strălucit. Totuși, există și aspecte negative.

Capitolul 7 se concentrează mai întâi asupra uneia dintre principalele probleme generate de automatizarea crescută, în general și de inteligența artificială, în special – pierderea locurilor de muncă atunci când oamenii sunt înlocuiți de mașini care pot îndeplini sarcinile mai rapid, mai eficient și mai economic. Descoperim că, în toate etapele, de la interviul de angajare până la pierderea locului de muncă, AI poate crea probleme care necesită soluții urgente în

economiile dependente de tehnologie. Relația dintre om și mașină este – sau ar trebui să fie – o prioritate pe agenda globală.

Al doilea subiect din capitolul 7 este modul în care AI afectează deja, în mod profund, domeniul educației – de la utilizarea tehnologiei de monitorizare intruzivă în școlile din China la folosirea ChatGPT pentru scrierea de eseuri. Aceasta lansează discuția despre revoluția GPT care a fost posibilă datorită creării de modele lingvistice de mari dimensiuni (LLM) dezvoltate de compania OpenAI condusă de Sam Altman. Modelele acestea sunt antrenate pe miliarde de cuvinte și pot genera texte convingătoare, pe multe teme, ca răspuns la întrebările oamenilor.

Competența acestor sisteme ne conduce spre al treilea subiect din capitolul 7: temerile legate de faptul că inteligența artificială ar putea scăpa de sub controlul uman, ceea ce i-a determinat pe mulți dintre liderii din industria AI să ceară un moratoriu asupra cercetării. Discutăm despre necesitatea și natura legislației pentru a controla sistemele AI care acționează în mod neprevăzut.

Ultimul subiect din capitolul 7 este utilizarea militară a inteligenței artificiale în angrenarea armelor autonome, o problemă care ridică numeroase dileme etice și care, din nou, necesită o atenție urgentă.

Capitolul 8 abordează problema urgentă a confidențialității și a drepturilor cetățenilor, având în vedere faptul că majoritatea dintre noi – în special utilizatorii de smartphone-uri, adică aproape toată lumea – împărtășim în mod voluntar date despre noi, obiceiurile noastre, conversații, prieteni și multe altele cu mega-corporații. Aceste date pot fi folosite în scopuri care depășesc controlul nostru și care nu sunt întotdeauna benefice – cel puțin nu pentru noi. Mai întâi, analizăm capitalismul de supraveghere – utilizarea datelor noastre, fără permisiune, în scopuri comerciale. Apoi, ne îndreptăm atenția către ceea ce am numit comunismul de supraveghere, care, după cum sugerează numele, se referă în principal la ceea ce se întâmplă în China – deși acest model este exportat în multe alte țări. Discutăm despre încercarea de guvernare bazată pe date prin intermediul sistemului de credit social, în care populația este monitorizată extensiv pentru a evalua dacă cetățenii sunt demni de încredere, prin măsurarea nivelului de conformare la ideologia statului: comportamentul „bun”