

LBRIS

We know
books

INES RAZEC

**IDENTITATEA ÎNTR-O LUME
PROGRESIV CUCERITĂ DE
INTELIGENȚA ARTIFICIALĂ**

t...

TRITONIC

Tritonic Books

București | 2024

CUPRINS

CUVÂNT ÎNAINTE	9
CAPITOLUL I.	
INTRODUCERE	13
Ce înțelegem în prezent prin Inteligență Artificială?	15
CAPITOLUL II.	
INTELIGENȚA ARTIFICIALĂ ȘI IDENTITATEA	21
II.1. Originile și evoluția Inteligenței Artificiale	21
II.2. Contribuțiile Teoriei Jocului și a Războiului Rece la evoluția tehnologică	32
II.3. Valențele complexe ale relației dintre identitate și mediul înconjurător	39
II.4. Implicațiile sociologice ale Inteligenței Artificiale	45
II.5. Identitatea Digitală sau Noul Sine	56
CAPITOLUL III.	
TIPURI DE IDENTITATE DIGITALĂ	67
III.1. Identitatea stocată în bazele de date și identitatea tranzacțională	67
III.2. Identitatea digitală vizibilă și identitatea digitală invizibilă	73
III.3. Identitatea decentralizată și Identitatea în Metavers	79
CAPITOLUL IV.	
NARAȚIUNI DESPRE INTELIGENȚA ARTIFICIALĂ DE-A LUNGUL TIMPULUI	89
IV.1. Teama că Inteligența Artificială va prelua controlul	94
IV.2. Aspectele etice ale relației dintre om și IA	98
IV.3. Viitorul muncii	101

CAPITOLUL V.**INTELIGENȚA ARTIFICIALĂ – REMEDIU UNIVERSAL SAU AMENINȚARE?** 107

V.1. Metafore despre Inteligența Artificială 111

V.2. Mituri despre Inteligența Artificială 116

V.3. Discursuri retorice despre Inteligența Artificială 121

CAPITOLUL VI.**EXERCİȚIU DE CREATIVITATE SAU REPREZENTĂRI ALE VIITORULUI?** 131

VI.1. Omul în relație cu Inteligența Artificială 131

VI.2. Chat GPT – Algoritm creativ sau sursă de narațiuni repetitive? 141

VI.3. A trecut Chat GPT Testul lui Turing? 147

CAPITOLUL VII.**STUDII DE CAZ** 151

VII.1. Cine sau Ce ne influențează atunci când cumpărăm? 151

VII.2. Chat GPT – motor al progresului sau al regresului mental? 158

VII.3. Inteligența Artificială, energie și societate 164

VII.4. Poate o Inteligență Artificială să devină conștientă? 178

VII.5. Scenarii despre viitor 202

CAPITOLUL VIII.**REZUMAT ȘI CONCLUZII** 213**CAPITOLUL IX.****CONCLUZII** 233**BIBLIOGRAFIE** 251

CAPITOLUL I.

INTRODUCERE

Integrarea rapidă a sistemelor inteligente în viața noastră creează premisele unor schimbări societale majore, ale căror manifestări devin din ce în ce mai evidente, în special prin favorizarea accentuată la unele nivele sociale (instituții, organizații, grupuri) sau tipologii umane (tineri, adolescenți, persoane cu anumite ocupații) a mediului online față de cel fizic. În prezent, aceste schimbări au devenit atât de evidente încât nu mai pot fi ignorate. Mai mult decât atât, numărul mare de controverse și opinii contradictorii din literatura de specialitate și mass-media cu privire la sistemele inteligente impun o analiză aprofundată, care ar putea oferi câteva idei concrete despre transformările produse în identitatea individuală și colectivă, în era tehnologică.

Această cercetare se fundamentează pe explorarea manierei în care se remodelează identitatea individului printr-un proces constant de adaptare la o societate în continuă schimbare și, în special, în contextul evoluției Inteligenței Artificiale. Pornind de la premisa că România se află în proces de digitalizare (i.e. „Planul Național de Redresare și Reziliență”), rezultatele obținute în urma studiului vor permite construirea unei viziuni orientative asupra nivelului de înțelegere a conceptului de Inteligență Artificială al respondenților din cadrul chestionarului și interviului, precum și a semnificațiilor pe

care le atribuie acestui concept. De aceea, în 2022, când am inițiat interviurile și chestionarele, scopul lor principal a fost să-mi permită să formulez o descriere cât mai sugestivă a realității în care ne situăm în prezent, pe plan național, în ceea ce privește raportul dintre identitate și Inteligența Artificială, din perspectiva respondenților. Răspunsurile obținute vor seta premisele pentru studii ulterioare care vor explora influența Inteligenței Artificiale din perspectiva schimbărilor concrete pe care le va produce pe plan național, în anii următori, din punct de vedere social, cultural și economic, precum: înlocuirea oamenilor cu mașini în îndeplinirea sarcinilor repetitive, dispariția anumitor meserii, redefinirea manierei de a comunica, estomparea graniței dintre spațiul real și virtual, depersonificarea indivizilor prin imersiunea în lumea virtuală, epuizarea resurselor naturale, criza energetică și a materialelor necesare pentru fabricarea dispozitivelor inteligente.

Ținând cont de rapiditatea cu care este integrată Inteligența Artificială în majoritatea sectoarelor existențiale, este firesc să ne așteptăm la producerea unor schimbări majore, atât în felul în care percepem realitatea, cât și în felul în care ne raportăm la ea. Pentru a înțelege în profunzime impactul IA asupra societății, am decis să analizez narațiunile individuale și colective, deoarece au rolul de a oferi o perspectivă de ansamblu asupra manierei în care oamenii înțeleg și încearcă să explice impactul pe care îl are Inteligența Artificială asupra existenței lor, prin prisma asocierilor între diferite fenomene și obiecte cotidiene. După cum voi demonstra pe parcursul acestei cărți, Inteligența Artificială este un concept complex, care a dobândit accepțiuni variate de-a lungul timpului, pe măsura evoluției lui. Pornind de la această premisă, prezenta cercetare nu poate folosi ca model de referință pentru dezvoltarea unei înțelegeri cuprinzătoare a variatelor valențe și maniere de a implementa IA, deoarece explorează doar aspectele considerate relevante pentru satisfacerea obiecti-

LBRIS

We know
books

vului principal, anume: analiza semnificațiilor termenilor „IA generală” și „IA limitată”, ca punct de plecare pentru determinarea similarităților și diferențelor dintre procesul de învățare specific uman și cel emulat de mașină, precum și a schimbărilor sociale și comportamentale pe care le produce o tehnologie aparent capabilă să execute cu succes și mult mai rapid activități cotidiene.

Rezultatele unui astfel de studiu ar putea fi benefice din mai multe motive: cunoașterea modului în care IA influențează comportamentul uman ar putea determina decizia de a o folosi sau nu în îndeplinirea anumitor activități; efectele IA asupra oamenilor ar putea contribui la prezicerea viitorului în ceea ce privește relațiile, comunicarea și redefinirea meseriilor sau la reinterpretarea unor concepte fundamentale de tipul umanitate, timp, spațiu sau societate. De asemenea, înțelegerea modului în care tehnologia influențează oamenii s-ar putea dovedi esențială în dezvoltarea unor sisteme capabile să crească gradul de conștientizare cu privire la pericolele cu care se confruntă societatea în ansamblul ei și ar încuraja acțiunile necesare pentru a le detecta și remedia pe cele care necesită o acțiune urgentă.

Ce înțelegem în prezent prin Inteligență Artificială?

Cum este percepută Inteligența Artificială de către indivizi în prezent? Care sunt riscurile și oportunitățile create de IA? Care sunt factorii care determină apariția identității digitale? Acestea sunt întrebările la care voi încerca să răspund pe parcursul acestei cărți, raportându-mă atât la studii relevante din spațiile european, japonez și american, cât și la opiniile respondenților români.

De la dispozitive inteligente capabile să stabilească cea mai bună rută de urmat, până la algoritmi care anticipează dorințele utilizatorilor și roboți care automatizează munca,

IA este percepută ca o soluție universală pentru unele dintre cele mai importante probleme cu care se confruntă societatea actuală. Încântat de acest miraj și motivat de dorința de a-și facilita viața, individul are tendința de a depinde tot mai mult de sistemele inteligente, ceea ce poate genera o dedublare a sinelui și crearea unei identități digitale, specifice mediului virtual. Conform doctrinei behavioriste, conceptul de identitate nu este unul static, ci reprezintă suma experiențelor trecute și viitoare, cuprinse în interiorul unui om. Din acest punct de vedere, este firesc să presupunem că orice schimbare din mediul înconjurător poate avea efecte asupra comportamentului individului, rezultând fie într-o atitudine pozitivă, de acceptare a acesteia, fie într-una negativă, manifestată printr-o reticență față de (sau chiar respingere) a noului.

În mod paradoxal, progresul din domeniul tehnologic, generat de ambiția intrinsecă a omului de a evolua, ne plasează în fața unei realități în care individul este cel care își reorganizează și modifică existența pentru a facilita integrarea Inteligenței Artificiale și nu invers: „Succesul IA se datorează în mare măsură faptului că noi construim un mediu propice IA, în cadrul căruia tehnologiile inteligente se simt ca acasă, iar noi suntem ca niște scafandrii. Lumea este cea care se adaptează la IA și nu vice-versa” (Floridi L. apud Franklin, 2017, p.157). Dacă rolul de instrument atribuit tehnologiei ar depăși nivelul pur teoretic, omul nu ar trebui să sufere schimbări comportamentale și de mediu pentru a acomoda mașinile în existența cotidiană. În schimb, tehnologia ar fi folosită în scop informativ, pentru facilitarea comunicării sau crearea unor dispozitive care să remedieze unele dintre marile probleme cu care se confruntă umanitatea, precum: poluarea, epuizarea resurselor naturale vitale, războaiele, discrepanțele înregistrate la nivel politic și economic. Având în vedere impactul uriaș al Inteligenței Artificiale asupra existenței individuale și colective, este necesar să luăm în considerare

LBRIS

We know
books

discrepanța dintre evoluția ei și ritmul cu care i se adaptează oamenii, cel mai frecvent fără a o înțelege în profunzime. În absența unei popularizări a avantajelor și riscurilor pe care le implică Inteligența Artificială, menită pregătirii societății pentru schimbarea pe care o suferă fără a o conștientiza ca atare, riscăm să dezvoltăm o disonanță cognitivă ale cărei urmări ar putea fi devastatoare pentru generațiile următoare.

Răspunsurile din cadrul chestionarului demonstrează faptul că cea mai răspândită semnificație atribuită Inteligenței Artificiale este cea de „Sisteme inteligente care îmbunătățesc existența cotidiană” (76%), urmată de „Roboți, mașini care se conduc singure, asistenți inteligenți (i.e. chatbot)” (75.5%). Ambele perspective demonstrează preocuparea respondenților pentru utilitatea implementării și folosirii algoritmilor inteligenți în diferite domenii de activitate. Asocierea IA cu asistenți inteligenți de tipul chat boților poate fi justificată prin lansarea Chat GPT 3, care s-a produs concomitent cu perioada de difuzare a chestionarului. În continuare, 61.5% dintre respondenți au echivalat-o cu „mașini programate pentru a simula creierul uman și a îndeplini sarcini repetitive”, iar 58% au considerat-o o manieră de a evolua „din punct de vedere social, cultural, economic și politic”. Analizând în ansamblu variantele cu cele mai mari procentaje, devine evidentă superioritatea numerică a implicațiilor de natură pozitivă. Dintre asocierile de natură negativă, cel mai mare număr de respondenți (54.5%) a manifestat un interes deosebit față de potențialul algoritmilor inteligenți de a deveni „sisteme de supraveghere și control al oamenilor”. Unul dintre respondenți propune o accepțiune diferită, care se află în concordanță cu această viziune, echivalând IA cu „Algoritmi de stocare a datelor cu caracter personal, care țintesc oameni și le expun oferte cu produse pe care oamenii le caută”, astfel făcând referire la „identitatea tranzacțională” și la comerțul social descrise pe larg în capitolele următoare.

Poveștile pe care oamenii le folosesc pentru a explica diferite fenomene reflectă atât nivelul lor de înțelegere, cât și asocierile pe care le stabilesc între respectivele fenomene și produse culturale care le sunt familiare. Interviuurile au revelat faptul că 60% din respondenți se raportează la filme pentru a înțelege mai bine ce înseamnă Inteligența Artificială. De aceea, printre cele mai frecvente răspunsuri primite s-au numărat cele care o asociază cu roboți capabili să emuleze emoții specific umane: „*Prima imagine care îmi vine în minte o reprezintă acei roboți performanți, capabili să poarte un dialog cu interlocutorul, capabili să recunoască emoțiile umane, dar mai ales – și asta mă fascinează și sperie în același timp – capabili să imite emoțiile*” (Respondent 2, feminin, 41 ani).

În lucrarea „Narațiunile IA: O istorie a gândirii imaginative despre mașini inteligente”, Cave et al. (2020) explorează impactul schimbărilor tehnologice asupra societății, din mai multe perspective, cu scopul de a evidenția riscurile la care sunt expuși utilizatorii. Importanța acestei cercetări se datorează în esență tocmai cunoașterii limitate a particularităților sistemelor inteligente. Cercetătorii evidențiază „rolul IA în guvernarea socială și a controlului, precum și a felului în care percepțiile legate de societatea mecanizată s-au dezvoltat în funcție de progresele din domeniul Inteligenței Artificiale” (p.214). Printre scenariile menționate de autori, care explorează fațete diferite ale percepției colective despre IA, se distinge conceptul numit în mod convențional „Conflictul care poate fi evitat” (p.224) și care cuprinde trei dimensiuni ale raportului dintre societate și IA, și anume: paradoxul construit în jurul ideii de liber arbitru în raport cu tehnologia, teama că „orice „scăpare” din această societate-mașină ar fi în detrimentul umanității” și puterea exagerată atribuită algoritmilor inteligenți (Ibid.). Cele trei accepțiuni reliefează imposibilitatea individului de a se opune unor schimbări care sunt în mod firesc generate de revoluția digitală. De asemenea,

EBRIS

We know
books

acestea evidențiază opoziția dintre ambițiile particulare ale oamenilor și cele formulate la nivel colectiv, care provin din nevoia firească de a evolua la nivel societal. În absența unei astfel de concordanțe, există două consecințe: pe de o parte, dezvoltarea unei rezistențe la schimbare, prin respingerea a tot ceea ce este nou și necunoscut și, pe de altă parte, folosirea Inteligenței Artificiale, fără a-i înțelege pe deplin modul de funcționare, ceea ce duce la multiplicarea riscurilor asociate cu supravegherea atentă a acțiunilor fiecărui individ, prin intermediul aplicațiilor inteligente care colectează datele personale, pentru a le folosi în diferite scopuri, de multe ori deghizate sub aparența unor servicii personalizate.

Multitudinea efectelor pe care tehnologia le are asupra indivizilor evidențiază caracterul complex al semnificațiilor pe care aceștia le atribuie procesului de digitalizare și le justifică reacțiile la schimbare. În acest sens, Turkle (2005) atrage atenția asupra diferitelor consecințe pe care le are folosirea calculatorului asupra copiilor și adulților, în funcție de specificul activităților pe care le desfășoară prin intermediul acestuia. Pe de o parte, în cazul copiilor, computerul devine parte integrantă a procesului de maturizare (p.154), deoarece îi acordă, cel mai frecvent în mod inconștient, un rol important în procesul de dezvoltare mentală pe care îl parcurg pe tot parcursul vieții. Pe de altă parte, adulții manifestă o reticență mai mare față de noile tehnologii în general, iar efectele pe care le au asupra lor variază în funcție de contextul în care se situează fiecare și de natura preocupărilor cotidiene: „pot funcționa ca bază pentru reflecții despre probleme filosofice importante și ambigue (p.155), pot fi folosite ca „instrument” și pot „genera dezvoltarea a ceva nou”, prin „influențarea concepțiilor oamenilor despre ei înșiși, despre locurile lor de muncă, despre relațiile lor cu alte persoane și despre modul lor de gândire cu privire la procesele sociale” (p.156). Ultima observație, care funcționează și ca punct de plecare pentru

studiul de față, face tranziția de la perspectiva generală asupra transformării mecanismelor pe care se fundamentează societatea, spre redefinirea identității personale, prin interiorizarea unor noi valori și idei, cu ajutorul tehnologiei. În acest context, autoarea remarcă ambiția Inteligenței Artificiale de a fi considerată „o nouă modalitate de a înțelege aproape orice” (p.226), ceea ce dă naștere unui paradox care constă în contradicția dintre ambiția creatorilor de programe inteligente de a explica mecanismele după care funcționează lumea înconjurătoare prin intermediul lor și faptul că oamenii au o cunoaștere limitată despre modul în care operează IA, ceea ce scade considerabil șansele să fie folosită pentru acest scop. Totuși, autoarea observă contribuția importantă a Inteligenței Artificiale la înțelegerea profundă a modului în care funcționează mintea umană, deoarece, în încercarea de a proiecta algoritmi inteligenți care să gândească precum omul, cercetătorii sunt nevoiți să se folosească pe ei înșiși ca exemplu (p.234).

CAPITOLUL II.

INTELIGENȚA ARTIFICIALĂ ȘI IDENTITATEA

II.1. Originile și evoluția Inteligenței Artificiale

Cercetătoarea Melanie Mitchell (2019) oferă o reprezentare comprehensivă a celor mai importante momente din procesul prin care Inteligența Artificială a devenit progresiv parte integrantă din existența cotidiană a indivizilor și societăților în ansamblul lor, de la crearea ei de către McCarthy, Minsky, Newel și Simon, până în prezent. La începutul proiectului, Herbert Simon (1995) atribuia mașinilor capacitatea să egaleze omul: „În 20 de ani, mașinile vor fi capabile să facă orice face omul” (Simon *apud* Mitchell, p.26). Această previziune înfățișează IA ca pe un dispozitiv care conține cunoștințe despre toate ariile de activitate ale omului și le poate reproduce fără a avea nevoie de un ajutor din exterior. Pe cât de generală și de categorică este această remarcă, pe atât de ambiguă. Totuși, lipsa unui scop unic și precis pentru IA deschide numeroase posibilități pentru cercetătorii care lucrează în mod activ la ea: „Deoarece nu înțelegem în profunzime sau nu știm cum să producem IA, în loc să distrugem orice cale de cercetare și pentru a face cu adevărat un progres, ar trebui să acceptăm anarhia metodelor care guvernează inteligența artificială” (Simon *apud* Mitchell, p.28). Potrivit lui Mitchell, din ansamblul de metode prin care Inteligența Artificială

acumulează cunoștințe, una care se distinge în mod evident și este folosită frecvent este cea numită „deep learning”. Pentru a înțelege în profunzime implicațiile complexe ale acestui proces, Mitchell propune o delimitare între două modalități diferite de asimilare și prelucrare a informației de către IA, anume: „dimensiunea simbolică”, care presupune traducerea cuvintelor în simboluri și asimilarea lor ca atare de către mașină, prin intermediul unui program de codificare (p.32) și dimensiunea „sub-simbolică”, a cărei denumire provine din capacitatea mașinii de a „citi” și informații care rulează sub nivelul conștient, adică „recunoașterea fețelor sau identificarea cuvintelor vorbite”, pe baza unor calcule matematice (p.32). Aceste două moduri diferite de funcționare a mașinilor imită modalitatea prin care creierul acumulează informații, iar diferența fundamentală dintre ele este că cea de-a doua efectuează operațiuni mai complexe.

Pornind de la stadiul actual al Inteligenței Artificiale, se impune o analiză mai aprofundată a celei de-a doua dimensiuni, anume a celei „sub-simbolice”, care a fost teoretizată de către Frank Rosenblatt, prin conceptul de „perceptron”. În lucrarea intitulată „Frank Rosenblatt: Principii de Neurodinamică: Perceptronii și Teoria Mecanismelor Creierului”, Van Der Malsburg (1986) definește perceptronul astfel: „Perceptronul este bazat pe ideea că un model de pe o suprafață senzorială poate fi reprezentat prin vectorul de activități al unui set fix de celule care detectează anumite caracteristici” (p.246). Mai precis, perceptronul simulează legătura dintre mai mulți neuroni, prin intermediul sinapselor, care sunt dotate cu capacitatea de a recunoaște stimulii primiți și transmiși, prin intermediul unor celule capabile să decodifice simbolurile și să le transpună în mesaje concrete. Reluând ideea lui Rosenblatt, în lucrarea „Perceptroni”, Minsky și Papert (1988) analizează limitările acestei teorii pornind de la creierul uman. Potrivit cercetătorilor, procesele care se