

**LBRIS**

MUSTAFA SULEYMAN și MICHAEL BHASKAR

We know

books

# URMĂTORUL VAL

INTELIGENȚA ARTIFICIALĂ ȘI  
CEA MAI MARE DILEMĂ A SECOLULUI 21

Bookzone  
BUCUREȘTI, 2024

## CUPRINS

GLOSAR DE TERMENI-CHEIE	7
PROLOG	9
CAPITOLUL 1 — CONTROLUL NU ESTE POSIBIL	11
PARTEA I — <i>HOMO TECHNOLOGICUS</i>	31
CAPITOLUL 2 — PROLIFERAREA NESFÂRȘITĂ	33
CAPITOLUL 3 — PROBLEMA CONTROLULUI	49
PARTEA A II-A — URMĂTORUL VAL	67
CAPITOLUL 4 — TEHNOLOGIA INTELIGENȚEI	69
CAPITOLUL 5 — TEHNOLOGIA VIEȚII	109
CAPITOLUL 6 — VALUL MAI AMPLU	127
CAPITOLUL 7 — PATRU CARACTERISTICI ALE URMĂTORULUI VAL	143
CAPITOLUL 8 — STIMULENTE DE NEOPRIT	163
PARTEA A III-A — STATELE EȘECULUI	201
CAPITOLUL 9 — MAREA ÎNVOIALĂ	203

CAPITOLUL 10 – FACTORII DE AMPLIFICARE A FRAGILITĂȚII	221
CAPITOLUL 11 – VIITORUL NAȚIUNILOR	253
CAPITOLUL 12 – DILEMA	281
<b>PARTEA A IV-A – PRIN VAL</b>	<b>307</b>
CAPITOLUL 13 – CONTROLUL TREBUIE SĂ FIE POSIBIL	309
CAPITOLUL 14 – ZECE PAȘI SPRE CONTROL	329
<b>MULȚUMIRI</b>	<b>395</b>

---

CAPITOLUL 1

---

**CONTROLUL NU ESTE POSIBIL**

---

VALUL

---

Aproape fiecare cultură are un mit al potopului. În textele antice hinduse, primul om din universul nostru, Manu, este avertizat cu privire la un potop iminent și devine singurul supraviețuitor al acestuia. *Epopeea lui Ghilgames* consemnează că zeul Enlil a distrus lumea într-un potop uriaș, o poveste cu care va rezona oricine este familiarizat cu povestea arcei lui Noe din *Vechiul Testament*. Platon a vorbit despre orașul pierdut Atlantida, spulberat de un val uriaș. În basmele, legendele și textele antice ale umanității este prezentă ideea unui val uriaș, care mătură totul în calea sa, lăsând în urmă o planetă care apoi renaște și se reface.

De asemenea, inundațiile și-au pus amprenta asupra istoriei la propriu – inundațiile sezoniere ale marilor râuri ale planetei, creșterea nivelului oceanelor după sfârșitul Epocii de Gheață, șocul rar al unui tsunami care apare fără avertisment la orizont. Asteroidul care a ucis dinozaurii a provocat un val înalt de un kilometru, modificând cursul evoluției. Puterea pură a acestor valuri s-a imprimat în conștiința noastră colectivă: ziduri de apă dezlănțuite, de neoprit, imposibil de controlat. Ele reprezintă unele dintre

cele mai puternice forțe de pe planetă. Ele modelează continentele, irigă recoltele și asigură dezvoltarea civilizației.

Au existat și alte tipuri de valuri – la fel de transformatoare. Dacă privim din nou istoria, putem vedea că aceasta este marcată de o serie de valuri metaforice: ascensiunea și căderea imperiilor și a religiilor, precum și diferitele perioade de avânt ale comerțului. Gândește-te la creștinism sau la islam, religii care au început ca niște mici unde înainte de a crește, a se transforma în niște valuri uriașe și a se prăbuși peste întinderi uriașe de pământ. Astfel de valuri sunt un motiv recurent, care încadrează fluxul și refluxul istoriei, marile lupte pentru putere, precum și perioadele de avânt și de criză economică.

Creșterea și răspândirea tehnologiilor a luat, de asemenea, forma unor valuri care au schimbat lumea. O singură tendință dominantă a trecut testul timpului de la descoperirea focului și a uneltelor de piatră – primele tehnologii utilizate de specia noastră. Aproape fiecare tehnologie fundamentală inventată vreodată, de la târnăcoape la pluguri, de la ceramică la fotografie, de la telefoane la avioane și tot ce se află între acestea, respectă o singură lege, aparent imuabilă: devine mai ieftină și mai ușor de utilizat și, în cele din urmă, se răspândește pe scară largă.

Această proliferare a tehnologiei în valuri este povestea lui *Homo technologicus* – omul tehnologic. Aspirația noastră, a umanității, de a ne autodepăși – pe noi înșine, soarta noastră, abilitățile și influența noastră asupra mediului înconjurător – a condus la o evoluție neîncetată a ideilor și a creației. Invenția este un proces evolutiv, expansiv și mereu în curs de desfășurare, impulsionat de inventatori, academicieni, antreprenori și lideri, care se organizează singuri și sunt extrem de competitivi, fiecare dintre ei mergând înainte condus de propriile motivații. Acest ecosistem al invenției favorizează implicit expansiunea. Aceasta este însăși natura tehnologiei.

Întrebarea este: ce se va întâmpla de aici încolo? În paginile care urmează, îți voi spune povestea următorului mare val tehnologic al istoriei.

Privește în jur. Ce vezi? Mobilier? Clădiri? Telefoane? Mâncare? Un parc amenajat? Aproape fiecare obiect din raza ta vizuală a fost, după toate probabilitățile, creat sau modificat de inteligența umană. Limbajul – temelia interacțiunilor noastre sociale, a culturilor noastre, a organizațiilor noastre politice și, poate, a ceea ce înseamnă să fii om – este un alt produs și un alt motor al inteligenței noastre. Fiecare principiu și concept abstract, fiecare mic efort sau proiect creativ, fiecare experiență din viața ta au fost influențate de capacitatea unică și nesfârșit de complexă a speciei noastre de a-și folosi imaginația, creativitatea și rațiunea. Ingeniozitatea umană este un lucru uimitor.

Mai există o singură forță la fel de omniprezentă în acest tablou: viața biologică însăși. Înainte de epoca modernă, în afară de câteva roci și minerale, majoritatea artefactelor umane – de la casele din lemn până la hainele din bumbac și focurile de cărbuni – proveneau din lucruri care fuseseră cândva vii. Tot ceea ce a apărut pe lume de atunci încoace provine de la noi, din faptul că suntem ființe biologice.

Nu este o exagerare să spunem că întreaga lume depinde fie de sistemele vii, fie de inteligența noastră. Și, cu toate acestea, ambele se află acum într-un moment fără precedent de schimbare și inovație exponențială, o creștere nemaîntâlnită, care va lăsa puține lucruri neschimbate. Un nou val de tehnologie începe să se spargă în jurul nostru, dezlănțuind puterea de a modela aceste două fundamente universale: inteligența și viața.

Următorul val este definit de două tehnologii de bază: inteligența artificială (IA) și biologia sintetică. Împreună, acestea vor inaugura o nouă eră pentru umanitate,

generând prosperitate și surplus cum nu s-a mai văzut până acum. Cu toate acestea, proliferarea lor rapidă amenință, de asemenea, să dea putere unei serii diverse de protagoniști răuvoitori, care ar putea să cauzeze tulburări, instabilitate și chiar catastrofe la o scară de neimaginat. Acest val creează o provocare imensă, care va defini secolul al XXI-lea: viitorul nostru depinde de aceste tehnologii și este, în același timp, pus în pericol de ele.

Din perspectiva actuală, se pare că nu este posibil să controlăm acest val, adică să îl stăpânim, să îi limităm impactul sau chiar să îl oprim. Această carte analizează de ce ar putea fi adevărat acest lucru și ce ar însemna asta. Implicațiile acestor întrebări îi vor afecta în cele din urmă pe toți cei în viață și fiecare generație care ne va urma.

Cred că acest val de tehnologie care se apropie duce istoria omenirii într-un punct de cotitură. Dacă nu-l vom putea controla, consecințele pentru specia noastră vor fi dramatice, potențial nefaste. În egală măsură, fără roadele sale, suntem expuși și vulnerabili. În ultimul deceniu, am prezentat de multe ori acest raționament, deși mai mult în privat, dar, pe măsură ce impactul devine tot mai greu de ignorat, este timpul să-mi prezint public teoria.

## DILEMA

---

Meditând la puterea imensă a inteligenței umane, mi-am pus o întrebare simplă, care mă mistuie neîncetat de atunci: ce-ar fi dacă am putea sintetiza esența a ceea ce ne face pe noi, oamenii, atât de productivi și capabili într-un software, într-un algoritm? Găsirea răspunsului ar putea debloca instrumente inimaginabil de puternice, care să ne ajute să rezolvăm cele mai dificile probleme. Ar putea fi un instrument, unul imposibil, dar extraordinar, care să ne

ajute să depășim provocările impresionante ale deceniilor viitoare, de la schimbările climatice la îmbătrânirea populației și la alimentația sustenabilă.

În acest sens, în vara anului 2010, într-un birou demodat, în stil Regency, cu vedere spre Russell Square din Londra, am înființat împreună cu doi prieteni, Demis Hassabis și Shane Legg, o companie numită DeepMind. Acesta era obiectivul nostru – unul care, retrospectiv, pare încă la fel de ambițios, nebunesc și plin de speranță: să reproducem exact ceea ce ne face unici ca specie, și anume inteligența noastră.

Pentru a atinge acest obiectiv, ar trebui să creăm un sistem care să poată imita și, în cele din urmă, depăși toate capacitățile cognitive umane, de la vedere și vorbire la planificare și imaginație și, în cele din urmă, empatie și creativitate. Deoarece un astfel de sistem ar beneficia de capacitatea foarte mare a supercalculatoarelor de a procesa în paralel informații și de proliferarea unor noi și vaste surse de date de pe internet, știam că orice progres în atingerea acestui obiectiv, oricât de modest, va avea implicații societale profunde.

Cu siguranță, la vremea respectivă, el părea destul de improbabil. Pe atunci, adoptarea la scară largă a inteligenței artificiale era un vis, mai mult o fantezie decât o realitate, o chestiune care aparținea doar câtorva academicieni izolați și unor fani ai literaturii științifico-fantastice cu idei nebunești. În ultimul deceniu însă, progresele făcute în domeniul inteligenței artificiale au fost de-a dreptul uluitoare. DeepMind a devenit una dintre cele mai importante companii de inteligență artificială din lume, realizând o serie de descoperiri. Viteza și puterea acestei noi revoluții au fost surprinzătoare, chiar și pentru aceia dintre noi care ne aflăm cel mai aproape de avangarda ei. Pe parcursul scrierii acestei cărți, ritmul progreselor în domeniul IA a fost uluitor, cu noi modele și noi produse care au apărut

în fiecare săptămână, uneori chiar în fiecare zi. Este clar că acest val se accelerează.

În prezent, sistemele de inteligență artificială pot recunoaște aproape perfect chipurile și obiectele. Am ajuns să considerăm că transpunerea vorbirii în scris și traducerea automată din limbi străine sunt lucruri normale. Inteligența artificială permite navigare pe șosele și în trafic suficient de bine pentru a putea conduce autovehicule în mod autonom în anumite situații. Pe baza câtorva indicații simple, o nouă generație de modele de inteligență artificială poate genera imagini inedite și compune texte cu niveluri extraordinare de detaliu și coerență. Sistemele de inteligență artificială pot produce voci sintetice de un realism straniu și pot compune muzică de o frumusețe uimitoare. Chiar și în domenii mai provocatoare, despre care s-a crezut mult timp că sunt destinate exclusiv capacităților umane, cum ar fi planificarea pe termen lung, imaginația și simularea ideilor complexe, progresele se fac cu pași repezi. Inteligența artificială își îmbunătățește, de câteva decenii, abilitățile cognitive și se pare că în următorii trei ani va atinge performanțe echivalente celor umane cu privire la o gamă foarte largă de sarcini. Aceasta este o afirmație importantă, dar, dacă am dreptate, implicațiile sunt cu adevărat mari. Ceea ce părea iluzoriu atunci când am fondat DeepMind a devenit nu doar plauzibil, ci și aparent inevitabil.

Încă de la început, mi-a fost clar că inteligența artificială va fi un instrument foarte util pentru obținerea unor beneficii extraordinare, dar, la fel ca majoritatea formelor de putere, prezintă pericole incomensurabile și dileme etice. Mult timp, mi-am făcut griji nu doar cu privire la consecințele evoluției IA, ci și cu privire la direcția în care se îndrepta întregul ecosistem tehnologic. Dincolo de IA, era în curs de desfășurare o revoluție mai amplă, în care IA alimenta o generație nouă și puternică de tehnologii genetice și robotică. Progresele ulterioare într-un domeniu le

acelerează pe celelalte, într-un proces haotic și reciproc de catalizare, care nu poate fi controlat de nimeni. Era clar că, dacă noi sau alții reușeam să reproducem cu succes inteligența umană, nu era vorba doar despre niște afaceri profitabile, ca de obicei, ci despre o schimbare radicală pentru omenire, care avea să inaugureze o eră în care oportunități nemaîntâlnite până acum aveau să fie însoțite de riscuri fără precedent.

Pe măsură ce tehnologia a progresat de-a lungul anilor, m-am îngrijorat din ce în ce mai mult. Ce se întâmplă dacă acest val este, de fapt, un tsunami?

În 2010, aproape nimeni nu vorbea serios despre inteligența artificială. Cu toate acestea, ceea ce părea cândva o misiune de nișă, ce implica un grup mic de cercetători și antreprenori, a devenit acum un vast efort global. Inteligența artificială este pretutindeni: la știri și în telefonul tău inteligent, tranzacționând acțiuni la bursă și construind pagini de internet. Multe dintre cele mai mari companii din lume și cele mai bogate națiuni se grăbesc să avanseze, proiectând modele de IA și tehnici de inginerie genetică de ultimă generație, susținute de investiții de zeci de miliarde de dolari. Odată ajunse la maturitate, aceste tehnologii noi se vor răspândi rapid, devenind mai ieftine, mai accesibile și difuzate pe scară largă, în întreaga societate. Ele ne vor permite să facem noi progrese extraordinare în domeniul medical și descoperiri în domeniul energiei curate, înlesnind apariția nu doar a unor noi întreprinderi, ci și a unor noi industrii și îmbunătățirea calității vieții în aproape toate domeniile imaginabile. Cu toate acestea, IA, biologia sintetică și alte forme avansate de tehnologie generează riscuri la o scară extrem de îngrijorătoare, care ar putea reprezenta o amenințare existențială pentru statele naționale – riscuri atât de mari, încât ar putea perturba sau chiar răsturna ordinea geopolitică actuală. Acestea deschid căi

către atacuri cibernetice de mare amploare cu ajutorul inteligenței artificiale, războaie automatizate, care ar putea devasta țări întregi, pandemii create artificial și o lume supusă unor forțe inexplicabile și totuși aparent omnipotente. Probabilitatea fiecăruia dintre aceste evenimente poate fi mică, dar consecințele posibile sunt uriașe. Chiar și o șansă infimă de a avea astfel de rezultate necesită o atenție imediată.

Unele țări vor reacționa la posibilitatea unor astfel de riscuri catastrofale printr-o formă de autoritarism de natură tehnologică, pentru a încetini răspândirea acestor noi puteri. Acest lucru va presupune niveluri extrem de ridicate de supraveghere, împreună cu intruziuni masive în viața noastră privată. Deținerea unui control absolut asupra tehnologiei ar putea deveni parte a unei tendințe de a supraveghea totul și pe toată lumea, tot timpul, într-un sistem global de supraveghere, distopic justificat de dorința de a se proteja împotriva celor mai nefavorabile consecințe posibile.

La fel de plauzibilă este și o reacție luddită. Vor urma interdicții, boicoturi și moratorii. Este oare posibil să se renunțe la dezvoltarea de noi tehnologii și să se introducă o serie de moratorii? Puțin probabil. Având în vedere valoarea lor geostrategică și comercială enormă, este greu de imaginat cum vor fi convinse statele naționale sau corporațiile să renunțe în mod unilateral la puterile transformatoare generate de aceste descoperiri. În plus, încercarea de a interzice dezvoltarea de noi tehnologii reprezintă un risc: societățile care stagnează din punct de vedere tehnologic sunt, istoric vorbind, instabile și predispuse la colaps. În cele din urmă, acestea își pierd capacitatea de a rezolva probleme, de a progresa.

Atât căutarea, cât și refuzul de a integra noi tehnologii sunt, din acest moment, pline de riscuri. Probabilitatea de a ne descurca pe o „cale îngustă” și de a evita oricare

dintre consecințe – distopia tehnocratică, pe de o parte, sau o catastrofă indusă de prea multă deschidere, pe de altă parte – scade odată cu trecerea timpului, pe măsură ce tehnologia devine mai ieftină, mai puternică și mai răspândită, iar riscurile se acumulează. Și totuși retragerea nu este nici ea o opțiune. Chiar dacă ne îngrijorăm în legătură cu riscurile asociate tehnologiilor valului următor, avem nevoie mai mult ca niciodată de beneficiile incredibile ale acestora. Iată dilema principală: faptul că, mai devreme sau mai târziu, o generație puternică de tehnologie va conduce omenirea către situații catastrofale sau distopice. Cred că aceasta este marea metaproblemă a secolului al XXI-lea.

Cartea de față subliniază exact motivele pentru care această situație teribilă devine inevitabilă și explorează modul în care am putea să o combatem. Cumva, trebuie să scoatem tot ce este mai bun din tehnologie – un aspect esențial pentru a face față unui set de provocări globale intimidante –, dar și să ieșim din dilemă. Discursul actual cu privire la etica și siguranța tehnologiei este inadecvat. În ciuda numeroaselor cărți, dezbateri, postări pe bloguri și tweetstorms<sup>1</sup> despre tehnologie, rareori se aude ceva despre *controlul* acesteia. Eu văd controlul acesta ca pe un set de mecanisme tehnice, sociale și juridice interconectate, care constrâng și limitează tehnologia și care funcționează la toate nivelurile posibile: un mijloc, în teorie, de a evita dilema. Cu toate acestea, chiar și cei mai aprigi critici ai tehnologiei tind să se ferească de acest limbaj al controlului dur.

Asta trebuie să se schimbe; sper ca această carte să arate de ce și să ofere indicii despre modul în care se poate face acest lucru.

---

1 *Tweetstorm* – un set de postări numerotate pe Twitter care permite celui care le scrie să trimită un mesaj lung și să depășească limita de caractere pentru fiecare postare individuală (en. în original) (n. tr.).

La câțiva ani după ce am înființat DeepMind, am creat un material de prezentare despre potențialul impact economic și social pe termen lung al inteligenței artificiale. Prezentându-mă în fața a 12 dintre cei mai influenți fondatori, directori executivi și specialiști din industria tehnologică, într-o sală de ședințe elegantă de pe Coasta de Vest, am argumentat că IA prezintă o serie de amenințări care necesită reacții proactive. Am spus că ar putea conduce la invazii masive ale vieții private sau ar putea declanșa o apocalipsă a informării greșite. Ar putea fi transformată în armă, creând o suită letală de noi arme cibernetice și introducând noi vulnerabilități în lumea noastră interconectată.

Am subliniat, de asemenea, potențialul inteligenței artificiale de a lăsa un număr mare de oameni fără locuri de muncă. Le-am cerut celor prezenți să ia în considerare istoria îndelungată a automatizării și a mecanizării, care a condus la înlocuirea forței de muncă. Le-am reamintit că mai întâi apar modalități mai eficiente de a îndeplini sarcini specifice, apoi numeroase locuri de muncă devin inutile și, în curând, sectoare întregi au nevoie de mult mai puțini lucrători. În următoarele câteva decenii, am susținut eu atunci, sistemele de inteligență artificială vor înlocui „munca intelectuală” în același mod și, cu siguranță, cu mult înainte ca roboții să înlocuiască munca fizică. În trecut, au fost create noi locuri de muncă în același timp în care cele vechi deveneau depășite, dar ce s-ar întâmpla dacă IA ar putea, pur și simplu, să le preia și pe cele mai multe dintre acestea? Am sugerat că există puține precedente pentru noile forme de putere concentrată care urmează să apară. Am susținut că, deși par îndepărtate, societatea se confruntă inevitabil cu amenințări potențial grave. În diaporitivul care încheia prezentarea mea era o imagine din

serialul *Familia Simpson*. În scena respectivă, orașenii din Springfield se revoltau, iar binecunoscutele personaje atacau cu torțe și bâte. Mesajul era clar, dar l-am explicat oricum. „Vin cu furcile” am spus eu. Voiam să spun că vin după noi, creatorii de tehnologie, și că depinde de noi să ne asigurăm că viitorul va fi mai bun de atât.

Ceilalți m-au privit uimiți. În încăpere totul încremenise. Mesajul nu fusese recepționat. Concedierile se făceau în număr mare și rapid. De ce indicatorii economici nu prezentau niciun semn legat de ceea ce le spuneam eu? Inteligența artificială avea să genereze o nouă cerere, care avea să creeze noi locuri de muncă. Ea urma să sporească și să le dea oamenilor posibilitatea de a fi și mai productivi. Poate că existau unele riscuri, recunoșteau ei, dar nu erau prea mari. Oamenii erau inteligenți. Întotdeauna se găsiseră soluții. *Nu este nevoie să ne facem griji*, păreau ei să creadă, *să trecem la următoarea prezentare*.

Câțiva ani mai târziu, cu puțin timp înainte de declanșarea pandemiei de COVID-19, am participat la un seminar despre riscurile tehnologice, la o universitate de renume. Configurația era similară: o altă mulțime mare adunată, o altă discuție de nivel înalt. Pe parcursul zilei, o serie de riscuri înfricoșătoare au fost vehiculate, în timp ce participanții beau cafea, mâncau biscuiți și se uitau la prezentări în PowerPoint.

Una dintre ele a ieșit în evidență. Vorbitorul a explicat că prețul sintetizatoarelor de ADN, care pot imprima șiruri de ADN personalizate, este în scădere rapidă. Costând câteva zeci de mii de dolari, acestea sunt suficient de mici pentru a putea fi amplasate pe un banc de lucru în garaj și a le permite oamenilor să sintetizeze – adică să fabrice – ADN. Și toate acestea sunt acum posibile pentru oricine are o pregătire de nivel de licență în domeniul biologiei sau entuziasm pentru învățarea independentă online.

Având în vedere disponibilitatea din ce în ce mai mare a acestor instrumente, prezentatorul a descris o viziune înfricoșătoare: în curând, oricine va putea crea agenți patogeni noi, mult mai transmisibili și mai letali decât orice altceva întâlnit în natură. Acești agenți patogeni sintetici ar putea eluda contramăsurile cunoscute, s-ar putea răspândi asimptomatic sau ar putea avea o rezistență intrinsecă la tratamente. La nevoie, cineva ar putea completa experimentele făcute în propria-i locuință cu ADN comandat online și reasamblat acasă. Apocalipsa comandată prin poștă. Nu e vorba despre literatură științifico-fantastică, a susținut prezentatorul, un profesor respectat, cu peste două decenii de experiență, ci e un risc real, actual. Vorbitorul și-a încheiat intervenția cu un gând alarmant: în ziua de azi, probabil că fiecare persoană „are capacitatea de a ucide un miliard de oameni”. Este nevoie doar de motivație.

Participantii s-au foit, agitați. Se mișcau pe scaunele lor și tușeau. Apoi au început să ridice obiecții și să aducă contraargumente. Nimeni nu voia să creadă că acest lucru era posibil. Cu siguranță, nu așa stăteau lucrurile, spuneau ei, cu siguranță, trebuiau să existe niște mecanisme eficiente de control, cu siguranță, bolile erau greu de creat, cu siguranță, bazele de date puteau fi blocate și, cu certitudine, hardware-ul putea fi securizat și așa mai departe.

Reacția comună din cadrul seminarului a fost mai mult decât disprețuitoare. Oamenii au refuzat, pur și simplu, să accepte perspectiva prezentatorului. Nimeni nu dorea să se confrunte cu implicațiile faptelor concrete și ale probabilităților sumbre pe care le auziseră. Am rămas tăcut, cu adevărat șocat. În curând, seminarul s-a încheiat. În seara aceea, am ieșit cu toții la cină și am continuat să discutăm, așa cum făceam întotdeauna. Tocmai vorbiseram despre sfârșitul lumii, dar încă aveam de mâncat pizza, de spus glume, de revenit la birou și, în plus, avea să apară

o soluție sau o parte din raționament avea să se dovedească greșit. M-am alăturat și eu celorlalți.

Dar prezentarea respectivă m-a bântuit luni întregi după aceea. De ce nu am luat-o eu mai în serios, de ce nu am luat-o cu toții mai în serios? De ce evităm, în mod ciudat, să continuăm discuțiile? De ce unii devin răutăcioși și îi acuză pe cei care pun aceste întrebări că sunt alarmiști sau că „trec cu vederea binele extraordinar” pe care îl aduce tehnologia? Această reacție emoțională generalizată pe care o observăm este una pe care am ajuns să o numesc capcana aversiunii față de pesimism: analiza eronată care apare atunci când ești copleșit de teama de a te confrunta cu realități potențial sumbre, astfel că ai tendința de a privi în altă parte.

Aproape toată lumea are reacția asta, sub o formă sau alta, iar consecința este că ne face să trecem cu vederea o serie de schimbări cruciale care se produc chiar sub ochii noștri. Este aproape un răspuns fiziologic înnăscut. Specia noastră nu este programată să se confrunte cu adevărat cu o transformare la această scară, ca să nu mai vorbim de posibilitatea ca tehnologia să ne dezamăgească în acest sens. Am avut acest sentiment de-a lungul carierei mele și i-am văzut pe mulți, mulți alții având aceeași reacție viscerală. Confruntarea cu acest sentiment este unul dintre scopurile acestei cărți. Să privim faptele la rece, oricât de inconfortabil ar fi.

Abordarea adecvată a acestui val, capacitatea de a controla tehnologia și de a ne asigura că aceasta servește întotdeauna umanității înseamnă depășirea aversiunii față de pesimism. Înseamnă să înfruntăm direct ceea ce urmează.

Cartea de față reprezintă încercarea mea de a face acest lucru, de a recunoaște și de a evidenția contururile valului care se apropie, de a explora dacă este posibil

să-l controlăm, de a pune lucrurile în context istoric și de a vedea imaginea de ansamblu, lăsând laoparte discuțiile zilnice legate de tehnologie. Scopul meu este de a aborda dilema și de a înțelege procesele care stau la baza apariției științei și tehnologiei. Vreau să prezint aceste idei cât mai clar posibil și unui public cât mai larg. Am scris această carte în spiritul transparenței și al cercetării: fac observații, urmăresc implicațiile lor, dar rămân, de asemenea, deschis la contestare și la interpretări mai bune. Nimic nu-mi doresc mai mult decât să mi se demonstreze că mă înșel și că totuși controlul este posibil.

Este de înțeles că unii oameni se pot aștepta la o carte mai tehnoutopică din partea unei persoane ca mine, fondatorul a două companii de inteligență artificială. În calitate de tehnolog și antreprenor, sunt, implicit, un optimist. Îmi amintesc că am fost complet captivat, pe când eram un tânăr adolescent, după ce am instalat Netscape pentru prima dată pe calculatorul meu Packard Bell 486. Eram fascinat de vâjâitul ventilatoarelor și de şuieratul distorsionat al modemului meu de 56 Kbps, care încerca să mă conecteze la rețeaua globală de internet și la forumuri și spații virtuale de conversație, care mi-au oferit libertate și m-au învățat atât de multe. Iubesc tehnologia. Ea este *însuși* motorul progresului și un motiv să fim mândri și entuziasmați de realizările umanității.

Dar cred, de asemenea, că aceia dintre noi care susțin crearea tehnologiei trebuie să aibă curajul de a anticipa și de a-și asuma responsabilitatea pentru direcția în care ne-ar putea duce în deceniile următoare. Trebuie să începem să sugerăm ce e de făcut, dacă ni se pare că există un risc real ca tehnologia să ne dezamăgească. Este nevoie de un răspuns social și politic, nu doar de eforturi individuale, dar acesta trebuie să înceapă cu cei ca mine și colegii mei.

Unii vor susține că totul este exagerat, că schimbarea este mult mai treptată, că este doar o altă etapă a ciclului de promovare, că, pentru a face față crizelor și schimbărilor, sistemele sunt, de fapt, destul de robuste, că viziunea mea asupra naturii umane este mult prea pesimistă și că istoria umanității este, ei bine, până acum, cât se poate de bună. Istoria este plină de profeti falși și de proroci pesimiști care s-au înșelat. De ce ar trebui ca de data aceasta lucrurile să stea diferit?

Aversiunea față de pesimism este o reacție emoțională, un refuz instinctiv și adânc înrădăcinat de a accepta posibilitatea unor consecințe puternic destabilizatoare. Aceasta tinde să vină din partea celor care ocupă poziții sigure și puternice, cu concepții fixe despre lume, oameni care pot face față în mod superficial schimbărilor, dar care nu acceptă cu ușurință nicio provocare reală la adresa ordinii lor mondiale. Mulți dintre cei pe care îi acuz că sunt blocați în capcana aversiunii față de pesimism îmbrățișează pe deplin criticile tot mai numeroase aduse tehnologiei, dar dau din cap fără să ia efectiv vreo măsură. Ne vom descurca, întotdeauna am reușit, spun ei.

Dacă petreci puțin timp în cercurile tehnologice sau politice, devine repede evident că ideologia implicită este cea a „capului ascuns în nisip”. A crede și a acționa altfel prezintă riscul de a deveni atât de paralizat de frica și de indignarea față de niște forțe uriașe, inexorabile, încât totul pare inutil. Așadar, intelectualii care manifestă aversiune față de pesimism continuă să bombăne. Cunosco această perspectivă, obișnuiam să fac la fel.

În anii care au trecut de când am fondat DeepMind și de la acele prezentări, discursul s-a schimbat – într-o oarecare măsură. Dezbateră privind automatizarea locurilor de muncă s-a reluat de nenumărate ori. O pandemie globală a demonstrat atât riscurile, cât și potențialul biologiei sintetice. În capitalele unde se află autoritățile de reglementare,