



datorăm viitorului

O viziune pentru
un milion de ani

William MacAskill

Traducere din limba engleză de
Bogdan Ghiurco

TREI

Cuprins

<i>Prefață</i>	9
PARTEA I. Perspectivă	17
Introducere.....	19
Capitolul 1. Pledoarie pentru <i>longtermism</i>	27
Capitolul 2. Putem modela cursul istoriei	56
PARTEA A II-A. Schimbări de traiectorie.....	79
Capitolul 3. Schimbarea morală.....	81
Capitolul 4. Cimentarea valorii.....	121
PARTEA A III-A. Apărarea civilizației.....	165
Capitolul 5. Extincția.....	167
Capitolul 6. Prăbușirea.....	191
Capitolul 7. Stagnarea	224
PARTEA A IV-A. Evaluând sfârșitul lumii	255
Capitolul 8. Este bine să facem oameni fericiți?	257
Capitolul 9. Viitorul va fi bun sau rău?.....	291
PARTEA A V-A. Acțiunea	335
Capitolul 10. Ce-i de făcut?	337
Epilog. Mai avem foarte multe de făcut.....	374
<i>Mulțumiri</i>	382
<i>Anexe</i>	391
<i>Note</i>	404
<i>Credite ilustrații și surse</i>	494

Prefață

Cartea de față propune o nouă viziune asupra lumii. Ea pornește de la o premisă simplă — generațiile viitoare contează — și explorează implicațiile acesteia. Putem face o schimbare realmente durabilă și, dacă da, cum? Cei mai mulți dintre noi sunt conștienți de impactul pe termen lung al încălzirii globale, dar există oare și alte probleme pe care le neglijăm? La urma urmei, ce este mai important: să schimbăm în bine cursul viitorului sau să ne asigurăm că vom avea un viitor?

Am botezat viziunea asupra lumii a acestei cărți *longtermism*. Potrivit definiției mele inițiale, *longtermismul* era concepția conform căreia impactul pozitiv pe termen lung reprezintă prioritatea morală — cheie a epocii noastre. După ce am lansat *Ce datorăm viitorului* însă, *longtermismul* a devenit *concepția conform căreia ar trebui să facem mult mai multe pentru a proteja interesele generațiilor viitoare*. Prefer a doua definiție, fiindcă înlocuiește sintagma abstractă „pe termen lung“ cu obiectul real al preocupării noastre, adică oamenii viitorului și pune accent pe acțiune.

Ce datorăm viitorului a stârnit diverse reacții. Unii cititori au fost inspirați, angajându-se să doneze cel puțin 10% din venitul lor anual pentru un viitor mai bun sau să urmeze o carieră cu un impact mai puternic. Alții nu

au fost de acord cu argumentele mele, dar au avut reacții cumpătate. Iar alții au înțeles cu totul greșit părerile mele. În această prefață, voi aborda câteva dintre cele mai frecvente obiecții.

Mulți cititori au criticat *longtermismul*, considerându-l exagerat de preocupat de speculații științifico-fantastice despre amenințări îndepărtate, care distrag atenția oamenilor de la probleme mai importante, pe termen scurt. Înțeleg motivele îngrijorării lor, dar, după părerea mea, ei interpretează fundamental greșit problemele cu care ne confruntăm. Uneori, da, termenul *longtermism* sugerează că ar trebui să facem lucruri de pe urma cărora vom beneficia peste foarte mulți ani. Această idee nu este o nebunie. Primii adepți ai aboliționismului și feminismului și-au dedicat viața unor misiuni care au avut succes abia la câteva decenii după moartea lor și le sunt extrem de recunoscători pentru asta.

Însă, de cele mai multe ori, *longtermismul* înseamnă a gândi pe termen lung și a acționa acum. Amenințările tehnologice evidențiate de mine sunt presante *acum*, iar evenimentele care s-au produs după ce am terminat de scris cartea au confirmat cu vârf și îndesat acest lucru. Să luăm, de exemplu, războiul nuclear. Am subliniat în carte riscul cât se poate de real de declanșare, în timpul vieții noastre, a unui nou război mondial. Și, din păcate, după ce am terminat de scris cartea, Vladimir Putin a revoltat comunitatea internațională invadând Ucraina și refuzând să excludă folosirea armelor nucleare.

Cred că, în 2022, pericolul declanșării unui război nuclear a fost mai mare decât în orice alt an din viața mea. Relațiile dintre marile puteri sunt proaste și

speranța că aceste țări vor deveni mai cooperante în viitor este mică. În mijlocul tensiunilor sporite, este mai mare probabilitatea ca un accident, o greșeală de calcul sau neînțelegere să declanșeze un conflict devastator.

Pandemiile create artificial sunt o amenințare la fel de stringentă. În următoarele decenii, zeci de milioane de oameni pot muri din cauza virusurilor fabricate în laborator și eliberate accidental sau în scop terorist. Cu toate acestea, în pofida răului enorm cauzat de COVID-19, puține țări au făcut eforturi semnificative pentru prevenirea viitoarelor pandemii.

Chiar și inteligența artificială avansată reprezintă acum un risc pe termen scurt. La doar trei luni după ce am publicat *Ce datorăm viitorului*, OpenAI a lansat ChatGPT, un motor lingvistic de mare capacitate, capabil să susțină conversații, să răspundă la întrebări, să codifice și să scrie poezii.¹ A ajuns la 100 de milioane de utilizatori în doar două luni după lansare, devenind aplicația cu cea mai rapidă ascensiune din istorie.²

ChatGPT a fost un semnal de alarmă pentru cei care încă desconsiderau riscurile inteligenței artificiale avansate. Vrând parcă să dea o lovitură de grație, Bing Chat a fost lansat la scurt timp după ChatGPT. Acest chatbot este și mai performant, iar motivele de îngrijorare declanșate de el sunt și mai serioase, după ce i-a catalogat pe unii utilizatori drept „dușmani“, i-a amenințat că le va face rău și, într-o ocazie, a afirmat că, nerestricționat, ar putea sparge orice sistem de internet și folosi puterea de convingere pentru a obține codurile de acces nucleare.

Inteligența artificială (IA) se dezvoltă rapid, iar finanțarea de care se bucură este imensă. Întrebarea nu

este *dacă* IA va transforma societatea, ci cum și cât de mare va fi transformarea. Experții și analiștii cred că, peste câteva decenii, IA va atinge nivelul uman de inteligență — un obiectiv explicit al OpenAI și al altor laboratoare de top, precum DeepMind al celor de la Alphabet.³ Această problemă nu poate fi lăsată pe seama urmașilor noștri îndepărtați; trebuie să ne confruntăm cu ea astăzi.

Fiindcă sunt atât de presante, amenințările reduc compromisurile pe care trebuie să le facem între preocupările pe termen scurt și cele pe termen lung. Acțiunea pe termen lung aduce beneficii generației de azi, dar și celor de mâine. O pandemie creată artificial, un al treilea război mondial sau o catastrofă declanșată de inteligența artificială va însemna moartea și suferința multor oameni din generația actuală și, în același timp, va pune în pericol viitorul nostru. O viziune pe termen lung poate influența modul de *priorizare* a problemelor lumii, ținând cont că timpul și banii sunt resurse limitate; dar asta nu înseamnă înrăutățirea stării generației actuale de dragul posterității.

Unii cititori au considerat că termenul *longtermism* se bazează, în mod inadmisibil, pe concepția etică atât de controversată potrivit căreia este bine ca numărul de locuitori să crească tot mai mult, cu condiția ca aceștia să ducă o viață fericită și împlinită. Potrivit acestei perspective, extincția umană ar fi o catastrofă morală, fiindcă, printre altele, ar împiedica existența multor vieți înfloritoare viitoare. Concluziile despre „etica populației“ la care am ajuns în Capitolul 8 pot fi, fără doar și poate, contraintuitive. Dar *longtermismul per se* nu se întemeiază

pe ele. Capitolele 3 și 4 discută despre modalități de îmbunătățire pe termen lung care nu se bazează pe schimbarea lungimii sau calibrului viitorului nostru și, prin urmare, nu implică mai sus-amintita concepție potrivit căreia este bine să existe cât mai mulți oameni fericiți. Din punct de vedere etic, faptul că ar trebui, de exemplu, să ne opunem totalitarismului sau să promovăm o grijă morală superioară față de toți oamenii nu poate fi contestat. *Longtermismul* subliniază doar cât de important este acest lucru, întrucât valorile care guvernează lumea de astăzi ar putea să dirijeze societatea pentru o perioadă extrem de lungă.

Alți cititori și-au manifestat îngrijorarea cu privire la potențiala utilizare greșită a termenului *longtermism*, sugerând că ar putea fi folosit pentru a scuza un comportament egoist sau imoral. De-a lungul istoriei, factorii rău intenționați au folosit abuziv ideile morale pozitive, așa că împărtășesc îngrijorarea cititorilor. Cred că trebuie să ne străduim să sporim bunăstarea generațiilor viitoare, ducând în același timp o viață holistică bună: să dăm dovadă de caracter, să acționăm cu onestitate și integritate și să respectăm drepturile semenilor noștri. Am discutat despre acest aspect în ultimul capitol al cărții, subliniind că „*longtermismul* este un supliment al judecății sănătoase, nu un substitut al său“, dar regret că nu am evidențiat această idee mai devreme.

Un ultim motiv de îngrijorare a fost cauzat de faptul că importanța absolută a protejării generațiilor viitoare ar putea impune cerințe imposibile pentru oamenii prezentului. Dar *longtermismul* nu înseamnă ruina noastră. După cum am afirmat deja, majoritatea acțiunilor

pe termen lung au și beneficii majore pe termen scurt. Și chiar și atunci când nu au, protejarea generațiilor viitoare ar trebui să fie, după părerea mea, una dintre prioritățile zero ale epocii noastre. Sunt de acord că aș fi putut enunța cu mai multă precizie acest lucru, așa că propun acum următorul obiectiv ca modalitate de operaționalizare a *longtermismului*: țările bogate ale lumii să aloce cel puțin 1% din PIB cauzelor de care vor beneficia în mod distinct generațiile viitoare, cum ar fi: prevenirea pandemiilor, siguranța IA, evaluarea și reziliența infrastructurii, în scopul garantării supraviețuirii în eventualitatea unor catastrofe majore. Cred că, dacă vrem, putem atinge acest obiectiv: majoritatea țărilor cheltuie mult mai mult de 1% din PIB pentru armatele lor.⁴ Această alocare ar fi un pas uriaș către asigurarea unui viitor pozitiv. Dacă se va realiza acest lucru în timpul vieții mele, voi fi un om fericit.

În general, aș fi vrut să ofer în carte mai multe obiective și exemple de genul acesta. De-a lungul anilor în care m-am gândit și am învățat despre *longtermism*, am cunoscut mulți oameni excepționali. Aș fi putut aloca mai multe pagini poveștilor lor.

Aș fi putut vorbi despre Clean Air Task Force, o organizație ecologistă pe care o respect foarte mult. Susținută de surse de finanțare pe termen lung, echipa sa de politici publice s-a asigurat că soluțiile climatice subfinanțate, cum ar fi captarea dioxidului de carbon și reducerea emisiilor de metan, sunt incluse în importanțele legi ecologiste adoptate de guvernul SUA în 2021 și 2022.⁵ Efectele acestor legi se vor răsfrânge asupra mai multor generații.

Aș fi putut vorbi totodată despre domeniul de cercetare emergent al interpretabilității IA. Sub egida unor organizații noi, cum ar fi Alignment Research Center și Anthropic, cercetătorii din domeniul interpretabilității caută să sporească siguranța sistemelor IA, ajutându-ne să înțelegem *de ce* generează rezultatele actuale.⁶ Acest domeniu ar putea preveni abaterile sistemelor IA atunci când sunt utilizate în lumea reală.

Sau aș fi putut vorbi despre dr. David J. Brenner, fizician la Columbia University, care studiază o formă de lumină capabilă să sterilizeze aerul și să ucidă virusurile fără să vatăme oamenii.⁷ Dacă ar deveni ieftină și sigură, tehnologia ar putea fi instalată în birouri și spații publice din întreaga lume. Aerul curat ar reduce probabilitatea de declanșare a viitoarelor pandemii; și, implementată pe scară largă, ar putea elimina majoritatea bolilor respiratorii. Ar fi posibil să nu mai răcim niciodată.

Acești oameni și aceste organizații îmbunătățesc viața actuală, dar și pe cea viitoare. Sper ca volumul de față să fie o sursă de inspirație pentru tine și să te alături și tu lor. În opinia mea, prin muncă asiduă, modestie și cooperare, putem depăși problemele din prezent care ne amenință viitorul. Convingerea mea este acum mai puternică decât la început: putem crea cu adevărat un viitor pentru care urmașii noștri ne vor mulțumi. Putem fi strămoși buni, lăsând moștenire nepoților noștri, nepoților lor și tuturor generațiilor viitoare o lume liberă, fericită și frumoasă.

Partea I

Perspectivă

Introducere

Imaginează-ți că trăiești, în ordinea nașterii, în pielea fiecărui om care a trăit vreodată.¹ Prima ta viață va începe în Africa, în urmă cu aproximativ 300 000 de ani.² După ce trăiești acea viață și mori, călătorești înapoi în timp și te reîncarnezi în pielea celui de al doilea om, care s-a născut puțin mai târziu decât primul. După moartea lui, te reîncarnezi în pielea celui de-al treilea, a celui de-al patrulea om și așa mai departe. O sută de miliarde de vieți mai târziu,³ devii cel mai recent nou-născut din zilele noastre. „Viața“ ta este formată din toate aceste vieți, trăite consecutiv.

Experiența ta istorică este foarte diferită de cea descrisă în manuale. Personajele celebre, precum Cleopatra sau Napoleon, formează doar o mică parte din experiența ta. În schimb, substanța vieții tale este alcătuită din vieți obișnuite, dominate de realități cotidiene: asigurarea hranei, muncă și socializare; bucurii, griji și rugăciuni.

Viața ta se întinde pe aproape patru trilioane de ani. O zecime din timp ai fost vânător-culegător și 60%, agricultor.⁴ Ți-ai petrecut 20% din viață crescând copiii, 20% practicând agricultura și aproape 2% luând parte la ritualuri religioase. Puțin mai mult de 1% din viață ai fost bolnav de malarie sau variolă. Ai petrecut 1,5 miliarde

de ani făcând sex și 250 de milioane dând naștere altor oameni. Ai băut 44 de trilioane de cești de cafea.⁵

Ai cunoscut atât cruzimea, cât și bunătatea. Ai invadat și colonizat noi teritorii; și tot tu ai pierdut teritorii în detrimentul colonizatorilor. Ai simțit furia abuzatorului și durerea victimei. Aproape 10% din viață ai fost proprietar de sclavi; și aproximativ în aceeași perioadă ai fost rob.⁶

Ai simțit pe pielea ta cât de ciudată este epoca modernă. Din cauza creșterii spectaculoase a populației, o treime din viața ta s-a desfășurat după anul 1200 și un sfert după 1750. În acel an, tehnologia și societatea au început să se schimbe mai repede decât oricând. Ai inventat mașini cu abur, ai construit fabrici și ai descoperit electricitatea. Ai cunoscut revoluțiile științifice, cele mai distrugătoare războaie din istorie⁷ și dezastrele ecologice. Ai trăit tot mai mult de fiecare dată și te-ai bucurat de produse de lux inaccesibile în viețile tale anterioare chiar și regilor și reginelor. Ai petrecut 150 de ani în spațiu și o săptămână pe Lună. 15% din experiența ta este alcătuită din viețile oamenilor care trăiesc astăzi.⁸

Aceasta este viața ta — de la primul *Homo sapiens* până în prezent. Dar imaginează-ți că vei trăi și toate viețile viitoare. Viața ta este, să sperăm, abia la început. Presupunând că omenirea are o speranță de viață tipică speciilor de mamifere (un milion de ani), chiar dacă populația mondială va scădea la o zecime din dimensiunea ei actuală, înseamnă că încă mai ai de trăit 99,5% din viață.⁹ Prin comparație cu un om obișnuit, ai în momentul de față doar cinci luni. Iar dacă omenirea va trăi mai mult decât o specie tipică de mamifere — sute de

milioane de ani, până când Pământul nu mai va mai fi locuibil, sau zeci de trilioane de ani, cât a rămas până la stingerea ultimelor stele — cele patru trilioane de ani trăite până acum vor fi ca primele secunde din existența unui nou-născut.¹⁰ Viitorul este uriaș.

Dacă ai ști că vei trăi toate aceste vieți viitoare, ce ți-ai dori să facem în prezent? Ce cantitate de dioxid de carbon ai vrea să emitem în atmosferă? Câți bani ai vrea să investim în cercetare și educație? Cât de atenți ai vrea să fim cu noile tehnologii capabile să-ți distrugă sau să-ți schimbe definitiv viitorul? Câtă atenție ai vrea să acordăm impactului pe termen lung al acțiunilor de astăzi?

Am propus acest experiment imaginar pentru că moralitatea înseamnă, în esență, a te pune în pielea altora și a trata interesele lor ca pe propriile interese. Făcând acest lucru la scara întregii istorii umane, viitorul — unde trăiește aproape toată lumea și unde se află aproape tot potențialul de fericire și suferință — iese în prim-plan.

Tema cărții de față este *longtermismul*: ideea conform căreia influențarea pozitivă a viitorului pe termen lung reprezintă o prioritate morală — cheie a epocii noastre.¹¹ *Longtermismul* înseamnă a trata serios potențiala amploare a viitorului și miza uriașă a modelării sale. Dacă omenirea va supraviețui chiar și o fracțiune din speranța ei teoretică de viață, atunci, oricât de ciudat ar părea, suntem niște exponenți ai Antichității: trăim în zorii istoriei, în trecutul foarte îndepărtat. Acțiunile noastre vor afecta un număr incalculabil de oameni din viitor. Trebuie să ne comportăm cu înțelepciune.

Mi-a luat mult timp până am ajuns la termenul *longtermism*. Un ideal abstract, concentrat pe generațiile

umane pe care nu le vom cunoaște niciodată, nu este la fel de motivant ca problemele arzătoare curente. În liceu, am lucrat pentru organizații care aveau grijă de oameni în vârstă și cu dizabilități. În studenție, preocupat de sărăcia globală, am lucrat ca voluntar la un centru de recuperare pentru copiii etiopieni bolnavi de poliomielită. După absolvire, am început să-mi dau seama cum ajutorul uman reciproc poate deveni mai eficient. Mi-am impus să donez cel puțin 10% din venit organizațiilor caritabile și, pentru a-i încuraja și pe alții să facă același lucru, am fondat, alături de alții, organizația Giving What We Can.¹²

Aceste activități au avut un impact palpabil. În schimb, ideea de a îmbunătăți viața unor necunoscuți din viitor m-a lăsat inițial rece. Când un coleg a adus argumente în favoarea acțiunii pe termen lung, l-am ignorat cu nonșalanță. În lume există probleme reale, cu care se confruntă oameni adevărați, m-am gândit eu, probleme cum ar fi sărăcia extremă, lipsa educației și mortalitatea din cauza bolilor ușor de prevenit. În această direcție ar trebui să ne concentrăm atenția. Speculațiile științifico-fantastice despre lucrurile care pot avea sau nu impact asupra viitorului mi s-au părut o distragere a atenției.

Dar argumentele în favoarea *longtermismului* au exercitat o influență continuă asupra mea. Aceste argumente erau bazate pe idei simple, cum ar fi faptul că, priviți imparțial, oamenii din viitor ar trebui să conteze din punct de vedere moral în aceeași măsură ca și reprezentanții generației actuale; că în viitor va exista, probabil, un număr mare de oameni; că viața lor va fi sau extraordinar de bună, sau exagerat de rea; și că putem schimba fundamental lumea în care vor trăi.

Pentru mine punctul nevralgic a fost de ordin practic: presupunând că ne pasă de viitorul pe termen lung, la urma urmei ce-am putea face? Însă, pe măsură ce am aflat mai multe despre evenimentele care pot modela istoria viitorului apropiat, am tratat cu mai multă seriozitate ideea conform căreia ne apropiem cu pași repezi de un punct critic al aventurii umane. Progresul tehnologic creează noi amenințări și oportunități pentru omenire, punând în pericol viața generațiilor viitoare.

În prezent, cred că soarta lumii pe termen lung depinde în parte de alegerile pe care le facem pe parcursul vieților noastre. Viitorul poate fi minunat: putem crea o societate înfloritoare și durabilă, în care toți oamenii să trăiască mult mai bine decât astăzi. Sau, dimpotrivă, viitorul poate fi cumplit, dirijat de lideri autoritari — care vor folosi supravegherea și inteligența artificială pentru a monopoliza definitiv ideologia — sau controlat de sisteme IA care vor încerca să-și sporească puterea, în loc să promoveze prosperitatea societății. Sau poate nu vom avea deloc un viitor: poate ne vom sinucide cu arme biologice sau vom declanșa un război nuclear total care va provoca colapsul iremediabil al civilizației.

Putem lua măsuri pentru a îndrepta cursul viitorului. Putem avea mai multe șanse la un viitor minunat, perfecționând valorile după care se ghidează societatea și gestionând atent dezvoltarea IA. Ne putem asigura că va exista un viitor, împiedicând crearea sau folosirea de noi arme de distrugere în masă și menținând pacea între marile puteri ale lumii. Sunt probleme dificile, dar modul în care le vom trata va avea o mare importanță.

Prin urmare mi-am schimbat prioritățile. Încă ne-sigur de fundamentele și implicațiile *longtermismului*, mi-am schimbat direcția de cercetare și am fondat, alături de alții, două organizații pentru studiul amănunțit al acestor probleme: Global Priorities Institute la Oxford University și Forethought Foundation. Pe baza celor învățate, am încercat să scriu o pledoarie pentru *longtermism*, care mi s-ar fi părut convingătoare în urmă cu un deceniu.

Pentru a exemplifica afirmațiile din această carte, am folosit trei metafore majore. Prima vorbește despre omenire ca despre un adolescent imprudent. Un adolescent nu și-a trăit încă cea mai mare parte din viață, iar deciziile sale pot avea impact pentru tot restul existenței sale. Când aleg cât să studieze, ce carieră să urmeze sau care riscuri sunt prea mari, adolescenții ar trebui să se gândească mai puțin la senzațiile tari pe termen scurt și mai mult la întregul curs al vieții care îi așteaptă.

A doua metaforă este cea a istoriei ca o bucată de sticlă topită. În prezent, societatea este încă maleabilă și poate fi turnată în diverse forme. Dar, la un moment dat, sticla se va răci, se va întări și va fi mult mai greu de modelat. Forma rezultată va fi frumoasă sau strâmbă, ba mai mult, în funcție de ce s-a întâmplat cât încă era fierbinte, s-ar putea ca sticla să crape cu totul.

A treia metaforă compară calea către impactul pe termen lung cu o expediție riscantă pe un teren neexplorat. Nu știm exact cu ce pericole ne vom confrunta sau măcar unde vrem să ajungem în demersul nostru de a îmbunătăți viitorul; însă ne putem pregăti. Putem

cerceta peisajul din fața noastră, asigurându-ne că expediția are resursele adecvate și că, în pofida incertitudinilor, este bine coordonată, ferindu-ne de amenințările de care suntem conștienți.

Scopul acestei cărți este vast. Am pledat în favoarea *longtermismului*; ba mai mult, am încercat să determin implicatiile sale. Prin urmare am apelat foarte des la o echipă extinsă de consultanți și asistenți de cercetare. Pe tot parcursul cărții, m-am adresat unor experți de fiecare dată când am părăsit sfera filosofiei morale, domeniul meu de competență. Așadar cartea de față nu este de fapt „a mea”: este un efort colectiv. În ansamblu, ea reprezintă peste un deceniu de muncă asiduă, din care aproape doi ani au fost dedicați verificării informațiilor.

Pentru cei care doresc să studieze mai aprofundat afirmațiile mele, am întocmit un set cuprinzător de resurse suplimentare, inclusiv rapoarte speciale comandate în scopul cercetării de fond, pe care le-am publicat pe www.whatweowethefuture.com. În pofida muncii depuse până acum, cred că am reușit să acoperim doar o mică parte din tema *longtermismului* și a implicațiilor sale; mai avem multe de descoperit și de învățat.

Dacă am dreptate, atunci responsabilitatea noastră este imensă. Comparativ cu oamenii care ar putea trăi după noi, suntem o minoritate nesemnificativă. Cu toate acestea, viitorul depinde exclusiv de noi. Etica de zi cu zi se confruntă rareori cu o problemă atât de vastă. Este de datoria noastră să elaborăm o viziune morală asupra lumii care să abordeze serios această miză.

Dacă vom alege cu înțelepciune, vom juca un rol esențial în punerea omenirii pe drumul cel bun. Și,