

JOHN  
BROCKMAN  
(EDITOR)

# MINȚI POSIBILE

Douăzeci și cinci de perspective  
asupra inteligenței artificiale

Traducere din limba engleză  
de George Chiriță

Redactor: Doinel Tronaru  
Corector: Dușa Udrea-Boborel  
Copertă: Daniel Breanova  
DTP: Răzvan Nasea  
Coordonare editorială: Doru Someșan  
Consultant editorial: Simona Radu-Răuță  
Concept și identitate vizuală: RUSU+BORTUN

Descrierea CIP a Bibliotecii Naționale a României  
Minți posibile : douăzeci și cinci de perspective asupra inteligenței  
artificiale / ed.: John Brockman ; trad.: George Chiriță. - București :  
Editura Vellant, 2019  
ISBN 978-606-980-067-6

I. Brockman, John (ed.)  
II. Chiriță, George (trad.)

004

Titlu original: *Possible Minds: Twenty-Five Ways of Looking at AI*  
Copyright © 2019 by John Brockman. All rights reserved.

© 2019 Editura Vellant

**Editura Vellant**  
Splaiul Independenței 319  
București, Sector 6  
[www.vellant.ro](http://www.vellant.ro)

Toate drepturile asupra versiunii în limba română  
aparțin Editurii Vellant.

# Cuprins

Introducere	
<i>Despre promisiunile și pericolele inteligenței artificiale</i>	11
Capitolul 1. Seth Lloyd:	
<i>Gresit, dar mai relevant decât oricând</i>	25
Capitolul 2. Judea Pearl:	
<i>Limitările mașinilor de învățare opace</i>	39
Capitolul 3. Stuart Russell:	
<i>Scopul scris în mașină</i>	46
Capitolul 4. George Dyson:	
<i>A treia lege</i>	62
Capitolul 5. Daniel C. Dennett:	
<i>Ce putem face?</i>	70
Capitolul 6. Rodney Brooks:	
<i>Debandada inumană în care ne-au băgat mașinile</i>	84
Capitolul 7. Frank Wilczek:	
<i>Unitatea inteligenței</i>	95
Capitolul 8. Max Tegmark:	
<i>Să aspirăm la mai mult decât doar să nu devenim depășiți</i>	107
Capitolul 9. Jaan Tallinn:	
<i>Mesaje disidente</i>	120
Capitolul 10. Steven Pinker:	
<i>Profeția tehnologică și puterea cauzală subapreciată a ideilor</i>	133
Capitolul 11. David Deutsch:	
<i>Dincolo de recompensă și pedeapsă</i>	147

Capitolul 12. Tom Griffiths:	
<i>Folosirea artificială a ființelor umane</i>	161
Capitolul 13. Anca Drăgan:	
<i>Integrarea omului în ecuația inteligenței artificiale</i>	171
Capitolul 14. Chris Anderson:	
<i>Coborârea pe panta unui gradient</i>	181
Capitolul 15. David Kaiser:	
„ <i>Informația“ pentru Wiener, pentru Shannon și pentru noi</i>	191
Capitolul 16. Neil Gershenfeld:	
<i>Scalarea</i>	201
Capitolul 17. W. Daniel Hills:	
<i>Primele inteligențe automate</i>	212
Capitolul 18. Venki Ramakrishnan:	
<i>Vor deveni computerele stăpânii noștri</i>	224
Capitolul 19. Alex „Sandy“ Pentland:	
<i>Strategia umană</i>	237
Capitolul 20. Hans Ulrich Obrist:	
<i>Transformarea invizibilului în vizibil: arta față în față cu inteligența artificială</i>	253
Capitolul 21. Alison Gopnik:	
<i>Inteligențele artificiale versus copiii de patru ani</i>	267
Capitolul 22. Peter Galison:	
<i>Algoritmicienii visează la obiectivitate</i>	280
Capitolul 23. George M. Church:	
<i>Drepturile mașinilor</i>	290
Capitolul 24. Caroline A. Jones:	
<i>Utilizarea artistică a ființelor cibernetice</i>	305
Capitolul 25. Stephen Wolfram:	
<i>Inteligenta artificială și viitorul civilizației</i>	318

# **Introducere**

## **Despre promisiunile și pericolele inteligentei artificiale**

Inteligenta artificiala este povestea zilei de astazi, povestea din spatele tuturor celoralte povești, este a doua venire a lui Iisus și Apocalipsa în același timp: inteligența artificială bună și inteligența artificială malefică. Această carte a reieșit în urma unor conversații în desfășurare cu câțiva gânditori importanți, atât din lumea inteligenței artificiale, cât și din afara acesteia, despre ce este și ce înseamnă inteligența artificială. Numită Proiectul Minți Posibile, această conversație a început în luna septembrie 2016, printr-o întâlnire la Grace Mayflower Inn & Spa din Washington, Connecticut, cu câțiva dintre cei care au contribuit la această carte.

Ce a reieșit foarte repede din acea primă întâlnire a fost că entuziasmul și teama din cultura mai largă ce înconjoară inteligența artificială acum au un corespondent în modul în care ideile lui Norbert Wiener cu privire la „cibernetică” și-au făcut loc în cultură, în special în anii 1960, pe măsură ce artiștii au început să încorporeze gândirea despre noile tehnologii în activitatea lor. Am fost martor apropiat la impactul acestor idei; și nu exagerez când spun că ele mi-au determinat traectoria vieții. Odată cu sosirea erei digitale, la începutul anilor 1970, oamenii au încetat să mai vorbească despre Wiener, dar astăzi această idee a ciberneticii a fost adoptată la scară atât de largă, încât este internalizată

până la punctul în care nu mai are nevoie de nume. Este peste tot, este în aer și este un loc potrivit de unde să începem.

## Tehnologii noi = perceptii noi

Înainte de inteligența artificială, a existat cibernetica – ideea controlului automat, care se reglează singur, expusă în textul de referință al lui Norbert Wiener din 1948. Pot să-mi datez propria mea expunere serioasă la acest test în 1966, când compozitorul John Cage m-a invitat pe mine și pe încă patru sau cinci artiști tineri să participăm împreună la o serie de serate – un seminar în desfășurare despre mass-media, comunicații, artă, muzică și filosofie, care se concentra pe interesul lui Cage în ceea ce privește ideile lui Wiener, Claude Shannon și Marshall McLuhan, toți având greutate în cercurile artistice din New York în care mă învârteam atunci. În special Cage preluase ideea lui McLuhan conform căreia, prin inventarea tehnologiilor electronice, am externalizat sistemul nostru nervos central – adică mintile noastre – și acum trebuia să presupunem că „există o singură minte, cea pe care o împărtăşim cu toții“.

Ideile de acest fel începeau să fie interesante pentru artiștii cu care lucram la New York și la Film-Makers' Cinematheque, unde eram director de programe artistice pentru o serie de producții multimedia denumită New Cinema 1 (cunoscută și sub numele Expanded Cinema Festival), sub auspiciile regizorului avangardist de film și impresarului Jonas Mekas. Printre ei se numărau: Claes Oldenburg, Robert Rauschenberg, Andy Warhol și Robert Whitman; artiștii cinetici: Charlotte Moorman și Nam June Paik; artiștii momentului Allan Kaprow și Carolee Schneemann; dansatoarea Trisha Brown; regizorii: Jack Smith, Stan Vanderbeek, Ed Emshwiller și frații Kuchar; dramaturgul de avangardă Ken Dewey; poetul Gern Stern și grupul USCO; muzicienii minimaliști La Monte Young și Terry Riley; și, prin Warhol, grupul muzical The Velvet Underground. Mulți dintre acești oameni citeau Wiener și cibernetica era în aer. La una dintre aceste

serate, Cage a băgat mâna în valiza sa și a scos un exemplar al cărții *Cibernetica*, mi l-a înmânat și a spus „Asta e pentru tine“.

În timpul festivalului, am primit un apel telefonic neașteptat de la colegul lui Wiener, Arthur K. Solomon, coordonatorul masterelor și doctoratelor de la Harvard în biofizică. Wiener murise cu un an înainte și ceilalți colegi apropiati ai lui Wiener și Solomon de la MIT și Harvard citiseră despre Expanded Cinema Festival în *New York Times*, fiind intrigăți de legătura cu cartea lui Wiener. Solomon m-a invitat să aduc câțiva artiști la Cambridge pentru a mă întâlni cu el și cu un grup ce-i includea și pe cercetătorul în domeniul comunicațiilor senzoriale de la MIT, Walter Rosenblith, pe matematicianul de la Harvard Anthony Oettinger și pe un inginer MIT Harold „Doc“ Edgerton, inventatorul luminii stroboscopice.

Ca multe alte situații în care „arta întâlnește știință“ în care am fost implicat de atunci, evenimentul de două zile era un rateu informat: nave care se petrec în noapte. Dar m-am înămat la acest proiect și evenimentul a fost important din câteva puncte de vedere interesante – unul fiind faptul că ne-au dus să vedem Computerul. Computerele erau o raritate pe atunci; cel puțin niciunul dintre noi, cei care eram în vizită, nu văzuserăm vreodată unul. Am fost conduși într-un spațiu mare din campusul MIT, în mijlocul căruia se afla o „cameră rece“ ridicată de la sol și cu pereti de sticlă, în care tehnicieni cu halate, eșarfe și mănuși albe erau ocupăți cu colaconarea unor cartele perforate care ieșeau dintr-o mașină enormă. Când m-am apropiat, aburul respirației mele a aburit fereastra din camera rece. Ștergând fereastra, am văzut Computerul. M-am îndrăgostit.

Mai târziu, în toamna anului 1967, am mers în Menlo Park pentru a petrece puțin timp cu Stewart Brand, pe care l-am întâlnit în New York, în 1965, când era un membru-satelist al grupului de artiști USCO. Acum, cu soția sa Lois, matematiciană, pregătea pentru publicare prima ediție a revistei *Whole Earth Catalog*. În timp ce Lois și echipa făceau munca grea pentru ultimele mecanici ale *WEC*, eu și Stewart am stat împreună într-un colț timp

de două zile, citind, subliniind și adnotând exemplarul din *Cibernetica* pe care mi-l dăduse Cage cu un an înainte și dezbatând ideile lui Wiener.

Inspirat de acest set de idei, am început să dezvolt o temă, un fel de mantră, care mi-a ghidat încercările de atunci: „tehnologii noi înseamnă perceptii noi“. Inspirat de teoreticianul comunicării Marshall McLuhan, designerul-arhitect Buckminster Fuller, futuristul John McHale și antropologii culturali Edward T. „Ned“ Hall și Edmund Carpenter, am început să citesc cu nesaț despre teoria informației, cibernetica și teoria sistemelor. McLuhan mi-a sugerat să citesc cartea biologului J.Z. Young *Doubt and Certainty in Science*, în care spune că noi creăm instrumente și ne modelăm prin utilizarea acestora. Celălalt text pe care l-a recomandat era lucrarea din 1949 a lui Warren Weaver și Claude Shannon intitulată *Contribuții recente la teoria matematică a comunicării*, care începe așa: „Cuvântul *comunicare* va fi utilizat aici cu sens foarte larg, pentru a include toate procedurile prin care o minte o poate afecta pe alta. Bineînțeles, acest lucru nu implică numai limbajul scris și oral, ci și muzica, pictura, teatrul, baletul și, de fapt, tot comportamentul uman“.

Cine ar fi știut că în două decenii de la acel moment vom începe să recunoaștem creierul ca pe un computer? Și, în următoarele două decenii, pe măsură ce am integrat computerele noastre pe internet, că va urma să realizăm despre creierul nostru că nu este un computer, ci o rețea de computere? Cu siguranță nu Wiener, specialist în circuitele cu *feedback* analogice concepute pentru a controla mașinile, nici artiștii, cu atât mai puțin eu.

### **„Trebuie să încetăm să sărutăm biciul ce ne biciuiește“**

La doi ani după *Cibernetica*, în 1950, Norbert Wiener a publicat cartea *The Human Use of Human Beings*<sup>\*</sup>, o poveste mai

---

\* Pe parcursul cărții, vom folosi titlul cărții în limba engleză, deoarece această carte nu a fost tradusă în limba română. O traducere aproximativă a titlului ar fi *Folosirea omului de către om* (n.t.).

profundă, în care își exprima îngrijorările despre exploatarea comercială scăpată de sub control și alte consecințe neprevăzute ale noilor tehnologii de control. Nu am citit *The Human Use of Human Beings* până în primăvara anului 2016, când am găsit un exemplar, ediție princeps, care stătea în biblioteca mea lângă *Cibernetica*. Ce m-a șocat a fost să realizez cât de profetic a putut să fie Wiener în 1950 despre ce se întâmplă în zilele noastre. Deși prima ediție a fost un *bestseller* major – și, într-adevăr, a inițiat o conversație importantă –, sub presiunea colegilor săi de breaslă, Wiener a scos o ediție revizuită și mai temperată în 1954, din care capitolul de concluzii inițial, „Vocile rigidității”, lipsește.

Istoricul științei George Dyson evidențiază faptul că, în această ediție princeps de mult uitată, Wiener a preconizat posibilitatea „unui fascism nou și amenințător, dependent de *machine à gouverner*“:

Nicio elită nu a scăpat de critica sa, de la marxiști și iezuiți („tot catolicismul este esențialmente o religie totalitară“) până la FBI („marii noștri prinți negustori au analizat tehnica de propagandă a rușilor și au constatat că este bună“) și mogulii care își oferă sprijinul în „a face capitalismul american și a cincea libertate a comerciantului supreme în lume“. Oamenii de știință... au fost supuși aceleiași cercetări meticuloase la care a fost supusă și Biserica: „Într-adevăr, directorii marilor laboratoare seamănă foarte mult cu episcopii, cu asocierile lor cu persoane influente din toate palierele vieții și pericolele la care se supun din cauza păcatelor carnale, a mândriei și a lăcomiei de putere“.

Ieremiada nu a mers bine pentru Wiener. În cuvintele lui Dyson:

Acele semnale de alarmă au fost ignorate la vremea aceea nu pentru că Wiener se înșela în ceea ce privește tehnica de calcul digitală, ci pentru că existau amenințări

mai mari când și-a terminat manuscrisul în toamna anului 1949. Wiener nu avea nimic împotriva tehnicii de calcul digitale, dar se opunea cu tărie armelor nucleare și a refuzat să se alăture celor care construiau computere digitale pentru a avansa cercetările pentru bomba cu hidrogen, care era de o mie de ori mai puternică.

Deoarece prima ediție a cărții *The Human Use of Human Beings* este acum ieșită de pe piață, nu mai avem acces la *cri-de-coeur*-ul lui Wiener, mai relevant astăzi decât atunci când l-a scris, acum șaizeci și opt de ani: „Trebuie să încetăm să sărutăm biciul ce ne biciuiește“.

### Mintea, gândirea, inteligența

Printre motivele pentru care nu auzim prea multe despre cibernetică în zilele noastre, două sunt centrale: în primul rând, deși *The Human Use of Human Beings* a fost considerată o carte importantă la vremea sa, era contrară aspirațiilor multor colegi ai lui Wiener, printre care și John von Neumann și Claude Shannon, care erau interesați de comercializarea noilor tehnologii. În al doilea rând, pionierul computerelor, John McCarthy, nu-l agreea pe Wiener și refuza să folosească termenul acestuia de „cibernetică“. McCarthy, la rândul său, a născocit sintagma de „inteligență artificială“ și a devenit un „părinte fondator“ al acestui domeniu.

După cum îmi explica Judea Pearl, care, în anii 1980, a introdus o abordare nouă a inteligenței artificiale, denumită rețelele bayesiene:

Ceea ce a creat Wiener a fost entuziasmul de a crede că într-o bună zi vom construi o mașină inteligentă. Nu era un informatician. Vorbea despre *feedback*, despre comunicare, despre analogic. Metafora prin care se exprima era un circuit cu *feedback*, domeniu în care era expert. Când a

debutat era digitală, la începutul anilor 1960, oamenii voiau să vorbească despre programare, despre coduri, despre funcții computaționale, despre memoria pe termen scurt, memoria pe termen lung – metafore informaticе care musteau de sens. Wiener nu făcea parte din acest val și nu a ajuns la noua generație ce se născuse cu ideile sale. Metaforele sale erau prea vechi, *passé*. Erau deja disponibile mijloace noi care erau pregătite să captiveze imaginația umană. Pe la începutul anilor 1970, oamenii nu mai vorbeau despre Wiener.

Un factor critic care lipsea din viziunea lui Wiener era elementul cognitiv: mintea, gândirea și inteligența. Încă din 1942, la prima întâlnire dintr-o serie de întâlniri interdisciplinare despre controlul sistemelor complexe care vor fi cunoscute sub numele de Conferințele Macy, cercetători de top argumentau includerea elementului cognitiv în conversație. În timp ce von Neumann, Shannon și Wiener erau concentrați pe sistemele de control și comunicare ale sistemelor observate, Warren McCullough voia să includă și mintea. S-a orientat către antropologii culturali Gregory Bateson și Margaret Mead, pentru a face legătura cu științele sociale. Bateson, în special, vorbea din ce în ce mai mult despre tipare și procese sau despre „tiparul care conectează“. A solicitat un nou tip de ecologie a sistemelor, în care organismele și mediul în care trăiesc sunt unul și același și ar trebui considerate un singur circuit. La începutul anilor 1970, cibernetica sistemelor observate – cibernetica de rangul întâi – s-a deplasat către cibernetica sistemelor observatoare – cibernetica de rangul al doilea sau „cibernetica ciberneticii“, termen introdus de Heinz von Foerster, care s-a alăturat conferințelor Macy la jumătatea anilor 1950 și a condus o nouă mișcare.

Cibernetica, în loc să dispară, se metaboliza în *tot*, aşa că nu o mai vedeam ca pe o nouă disciplină separată și distinctă. Și acolo a rămas, ascunsă în văzul tuturor.

## „Sceneta actorilor principali“

Propriile mele scrieri pe această temă la momentul respectiv erau pe radarul oamenilor care se ocupau de cibernetica de rangul al doilea, printre care Heinz von Foerster, dar și John Lilly și Alan Watts, care erau coorganizatorii a ceva numit Conferințele AUM, acronim pentru „the American University of Masters“, care a avut loc în Big Sur, în 1973, o adunare de filosofi, psihologi și oameni de știință, fiecare personalitate trebuind să țină o prelegere despre propria sa activitate în termenii relației acesteia cu ideile matematicianului britanic G. Spencer-Brown, prezentate în cartea sa *Laws of Form*. Am fost puțin nedumerit când am primit invitația – o invitație foarte târzie –, despre care spuneau că se bazează pe interesul lor în ideile prezentate de mine într-o carte cu titlul *Afterwords*, idei care erau pe lungimea lor de undă. Am profitat bucuros de ocazie, motivul principal fiind că vorbitorul principal era nimeni altul decât Richard Feynman. Ador să-mi petrec timpul cu fizicieni, deoarece se gândesc la univers, adică la tot. Și niciun fizician nu avea o reputație de a fi mai articulat decât Feynman. Abia așteptam să-l cunosc. Am acceptat. Acestea fiind spuse, nu sunt un om de știință și nu m-am gândit niciodată să mă urc pe o scenă și să țin vreo „prelegere“ de orice fel, cu atât mai puțin un comentariu despre o teorie matematică obscură în fața unui grup identificat ca fiind format dintre cei mai interesanți gânditori din lume. Abia după ce am sosit în Big Sur, am aflat motivul invitației mele întârziate. „Când va vorbi Feynman?“ am întrebat eu la recepție. „A, nu ţi-a spus Alan Watts? Richard este bolnav și a fost internat. Ești înlocuitorul lui. Și, aproape, care este titlul discursului tău inaugural?“

Am încercat să mă fac invizibil câteva zile. Alan Watts, dându-și seama că evit podiumul, m-a trezit într-o noapte, la trei dimineață, ciocănind la ușă. Am deschis ușa și l-am văzut stând în fața mea într-o robă de călugăr, cu gluga acoperindu-i o mare parte din față. Avea brațele întinse, o lanternă într-o mâna și un whisky Magnum în celalaltă. „John“, a spus el pe o voce groasă și cu un accent britanic

aristocratic, „ești un impostor. Și, John“, a continuat „eu sunt un impostor. Dar, John, eu sunt un impostor *adevărat!*“

A doua zi, am ținut discursul, intitulat „Einstein, Gertrude Stein, Wittgenstein și Frankenstein“. Einstein, revoluția din fizica secolului XX. Gertrude Stein, prima scriitoare care și-a integrat în lucrări ideea unui univers nedeterminat și discontinuu. Cuvintele nu reprezentau nici caracterul, nici activitatea: un trandafir este un trandafir este un trandafir, iar un univers este un univers este un univers. Wittgenstein: lumea ca limite ale limbajului. „Limitele limbajului meu înseamnă limitele lumii mele.“ Sfârșitul distincției dintre observator și observat. Frankenstein: cibernetica, inteligența artificială, robotica, toți eseștii din acest volum.

Prelegerea a avut consecințe neprevăzute. Printre participanții la Conferințele AUM erau mai mulți autori de bestselleruri ajunse numărul unu în *New York Times*, însă nimeni de acolo nu avea un agent literar. Și mi-am dat seama că toată lumea scria un gen de carte nedenumit și nerecunoscut de editurile din New York. Deoarece aveam un masterat în administrarea afacerilor de la Columbia Business School și o serie de succese relative în afaceri, am fost hărțuit să devin agent, inițial pentru Gregory Bateson și John Lilly, ale căror cărți le-am vândut rapid și pentru sume care mi-au captat atenția, începându-mi astfel cariera de agent literar.

Nu am apucat să-l întâlnesc niciodată pe Richard Feynman.

## Lungile ierni ale inteligenței artificiale

Această nouă carieră m-a pus în legătură apropiată cu majoritatea pionierilor din domeniu și, de-a lungul deceniilor, am plutit cu ei pe valuri de entuziasm și în văi de dezamăgire. La începutul anilor 1980, guvernul japonez a făcut un efort național de a avansa inteligența artificială. L-au numit a cincea generație; obiectivul lor era să schimbe arhitectura tehnicii de calcul trecând de „problema limitării a lui von Neumann“ și creând un computer care lucrează masiv în paralel. Făcând acest

lucru, sperau să-și dezvolte economia și să devină o putere mondială dominantă în domeniu. În 1983, liderul consorțiu-lui celei de-a Cincea Generații Japoneze a venit în New York pentru o întâlnire organizată de Heinz Pagels, președintele Academiei de Științe din New York. Aveam un loc la masă, alături de liderii primei generații Marvin Minsky și John McCarthy; de cei ai celei de-a doua generații, Edward Feigenbaum și Roger Schank; și de Joseph Traub, directorul Consorțiului Național de Supercomputere.

În 1981, cu ajutorul lui Heinz, am înființat The Reality Club (precursorul organizației nonprofit Edge.org), ale cărei întâlniri interdisciplinare inițiale aveau loc în sala de consiliu a Academiei de Științe din New York. Heinz lucra la cartea sa *The Dreams of Reason: The Computer and the Rise of the Science of Complexity*, pe care o considera a fi o agendă de cercetări pentru știință în anii 1990.

La întâlnirile Reality Club, am cunoscut doi cercetători tineri care urmău să joace roluri-cheie în revoluționarea informaticii. La MIT, la sfârșitul anilor 1970, Danny Hillis a dezvoltat algoritmii care au făcut posibil computerul masiv în paralel. În 1983, compania sa, Thinking Machines, a construit cel mai rapid supercomputer din lume utilizând arhitectura în paralel. „Mașina sa de conexiuni“ reflecta îndeaproape modul în care funcționează mintea omenească. Seth Lloyd, de la Universitatea Rockefeller, desfășura o activitate foarte importantă în domeniile tehnicii de calcul cuantice și comunicațiilor cuantice, propunând inclusiv primul design fezabil din punct de vedere tehnologic pentru un computer cuantic.

Iar japonezii? Incursiunea lor în domeniul inteligenței artificiale a eșuat și a fost urmată de douăzeci de ani de creștere economică anemică. Însă cercetătorii importanți din SUA au luat acest program foarte în serios. Iar Feigenbaum, care era informaticianul cel mai cunoscut al momentului, a făcut echipă cu Pamela McCorduck pentru a scrie o carte despre aceste dezvoltări. *The Fifth Generation: Artificial Intelligence and Japan's Computer*