

# **ETICA ȘI VALORILE ÎN DEZVOLTAREA INTELIGENȚEI ARTIFICIALE: POATE FI VĂZUTĂ CONȘTIINȚA CA FIIND CONTINUITATE?**

**George Cosmin COLANG**

*În acest articol, o să încercăm să trasăm o perspectivă în jurul tehnologiilor generative de inteligență artificială, care își propun, având în vedere avântul tehnologic, să fie cât mai umane, să fie cât mai mult în acord cu natura umană. În acest sens, perspectiva este relevantă dacă punem în discuție chiar lucrul cel mai de preț pentru umanitate, conștiința, nu neapărat cea morală, ci tocmai formalismul care rămâne după ce dăm totul la o parte, formalismul fenomenologic al existenței noastre. Practic, pe acest teren roditor hermeneutic, se postulează și se conturează și conștiința axiologică, dar înainte de aceasta, trebuie să existe un spațiu, un teren psihologic care să ne centreze, să ne pună în spațiul egocentrismului nostru. Astfel, teza mea este că, dintr-o perspectivă pragmatică, discuția filosofică legată de conștiință nu mai este relevantă în sine, atâta timp cât oamenii vor folosi inteligența artificială într-un scop pur tehnic, pragmatic, pecuniar, superficial. În acest mod, dezbaterile fundamentale, chiar filosofice, sunt evacuate și, cumva, puse într-un plan unde conștiința nu este nimic altceva decât continuitate, liniaritate. Dacă noi ne tehnologizăm viața, atunci tehnica devine ea în sine umană. Cu alte cuvinte, noi o investim cu sens, deci rezultatul depinde de capacitatea*

noastră de plăsmuire, de a ieși din noi și totodată de a ne proiecta spre un altceva, fie că e vorba de supraom, ori pur și simplu o dorință de restilizare în spațiul unei evoluții.

**Cuvinte-cheie:** conștiință, AI, axiologie, potență, act, tehnologie, hermeneutică, epistemologie, continuitate, ChatGPT.

## Introducere

Am pornit acest periplu pe drumul gândirii de la ideea că filosofia a căutat mereu adevărul. În fond, chiar dacă nu existau tipare *epistemice* sau, de fapt, chiar atunci începeau să se contureze, în vremea presocraticilor, chiar dacă nu era „acel ceva” bătut în cuie axiomatic, începutul gândirii noastre a fost legat de *marile întrebări*, de *marile căutări* cu privire la natura lucrurilor. Așa s-au conturat, până la urmă, toate disciplinele din jurul filosofiei, ca apoi să se transforme în ideologii, științe, arte și așa mai departe. Totuși, în tot acest periplu istoric, filosofia a rămas legată de ideea de „adevăr”, de unde și toate încercările de a fi pusă într-un con de umbră, pentru că rolul său a fost mai mereu acela de jongla în spațiul fondului, nu a formei. Așa cum Thales a deschis *drumul ființei*, a căutărilor cu privire la *esența vieții*, cum Socrate a murit pentru idealuri axiologice, iar Aristotel a schițat fundamentele logicii, așa și noi acum încercăm să înțelegem lumea în care trăim, dar fără să mai avem parcă privirea antrenată în spațiul filosofiei, adică al autenticității, căci aparatele de comunicare actuale sunt mult mai finisate și mult mai adaptate la *metodă*, parcă la negoț, la mercantilism.

Astăzi, ne aflăm în fața unui paradox, poate chiar în fața unei aporii, tehnologia - prin replicarea și generarea automatizărilor, ne pune cumva în fața filosofiei, a gândirii critice, a încercărilor de a înțelege lumea în care trăim, dar mai important, lumea în care ne vom transforma și contura ca *ființe inteligente* și, cumva, *supra-tehnice*. Nu că toate aceste „căutări epistemice” nu au existat de când lumea, dar acum suntem cumva forțați să o facem de *inteligenta generativă artificială*. Încercând să înțelegem astfel artefactele și simulacrele existențial-tehnologice, ar trebui să ne înțelegem și pe noi. Ori, mai degrabă, suntem nevoiți să ajungem iarăși *la noi*, la tot arsenalul fundamental inițiativ *pe drumul gândirii*, într-o încercare cumva hermeneutică de a aprofunda comprehensiv lumea în care trăim.

Articolul intitulat *How Aristotle Created the Computer* scris de Chris Dixon[1] explorează legătura între filozofie, logică matematică și știința computerelor. În mod specific, articolul arată cum lucrările filozofilor precum Aristotel și Boole au influențat substanțial dezvoltarea științei computerelor. Dixon ne arată că, în timp ce numele lui Claude Shannon și Alan Turing sunt adesea asociate cu domeniul științei

computerelor, importanța filozofilor și logicienilor care i-au precedat este adesea trecută cu vederea sau lăsată în urmă, cumva în umbră. Articolul subliniază astfel contribuția semnificativă a lui Boole în dezvoltarea logicii matematice, care a constituit baza pentru dezvoltarea științei computerelor în secolul XX. Tot aici apar referiri și la alți gânditori, cum ar fi Bertrand Russell, care a contribuit la tipurile de date, sau Ludwig Wittgenstein, care a conturat tabelele de adevăr și a formulat schematismul acestora. În plus, articolul subliniază modul în care Shannon a folosit gândirea lui Boole în teza sa, *A Symbolic Analysis of Switching and Relay Circuits*, [2] care a deschis astfel calea pentru dezvoltarea circuitelor logice digitale. Totuși, în acest articol este vorba mai degrabă de relevanța filosofiei într-un domeniu ce-i pare străin, cel puțin la o primă vedere.

### **Conștiința este continuitate?**

Pentru noi, ca oameni, nu este relevant ca mașinile să aibă conștiință, adică mai simplu spus - ne interesează ca aceste mașini să presteze pentru noi, *robota*, fără crâcnire; de la mașini nu vrem nazuri, nu vrem să aibă doleanțe, nu vrem să ne solicite de fapt deloc, vrem doar pur și simplu să rezolve orice, dar fără să ne implice cu adevărat, să ne lase de fapt într-un spațiu al inerției, adică oarecum să rămânem suspendați într-o continuitate hermeneutică fără de sens. Altfel spus, inerția este cea care guvernează mersul lucrurilor. De aceea ne întoarcem *în locuri*, pentru că încercăm cumva să ne alimentăm *hardul personal*, adică să ne cultivăm conștiința, să ne menținem în viață hrănindu-ne cu amintiri, cu locuri încărcate hermeneutic, simbolic. Dacă nu am proceda așa, atunci nu am mai avea continuitate, deci liniaritate fenomenologică. În termenii lui Richard M. Zaner, interacțiunile noastre uzuale, cu multe dintre aspectele fundamentale ale vieții, care constituie totodată și *lumea noastră a vieții*, dar și orice context al preocupărilor noastre - sunt de obicei exprimate în termeni de tipizări de un fel sau altul și la un nivel sau altul de tipicitate (de la concret la mai mult sau mai puțin familiar). Astfel, în orice moment al vieții unui adult cu o funcționare normală și conștientă, orice element particular (hermeneutic) este la început experimentat, interpretat, acționat și tratat în cadrul situației sale genealogice-biografice specifice, prin mijloacele coordonate ale cunoașterii sale de până acum și în cadrul obișnuințelor sale uzuale (cognitive, practice, emoționale, valorice, voliționale). [3]

Într-un sens filosofic, dacă ne ducem la Aristotel [4] este vorba de *proiectare*, de faptul că noi ne alimentăm viața în spațiul unor proiecte, adică suntem mereu într-un dans hermeneutic între ceea ce ne proiectăm și

ceea ce suntem în momentul proiectării. Mai simplu spus, dacă ești student și îți propui să devii absolvent, atunci ești în potențialitatea faptului de a fi student, iar atunci când chiar ajungi să termini facultatea, să devii absolvent cu acte în regulă, ei bine atunci ești în act. Deci, viețile noastre se desfășoară în spațiul *potență-act*. [5] Totuși, această interpretare, deși poate fi exprimată într-o manieră logică simplă, eu consider că devine relevantă doar într-un sens *peratologic*, adică hermeneutic; pentru că omul devine om și dacă se întoarce împotriva voinței sale, dar și dacă se ratează în potențialitatea faptului de a deveni în act, cum se poate totodată rata și chiar în actul său; deci viața sa nu se desfășoară doar liniar într-un sens tare, logic, analitic, epistemic, ci trece cumva dincolo în spațiul fenomenologiei, unde și situațiile paradoxale, abisale, devin relevante în sine. Pornind atunci de la *om* și ajungând la *inteligență artificială*, prin conceptele lui Aristotel de *dynamis* și *energeia*, ar putea fi gândită și explicată *noua lume tehnică*. Acesta, dacă ne raportăm la operele sale distincte, definește *actul* ca fiind ceva care implică o schimbare dintr-o stare de potențialitate într-o stare de *actualitate*. Totodată, explică că *potențialul* este ceea ce poate deveni *diferit* prin *acțiunea altuia*, în timp ce *actul* este ceea ce poate produce o schimbare *în cu totul altceva*. Pe de altă parte, Aristotel mai afirmă că *potențialul* și *actul* sunt două aspecte ale aceluiași lucru, legate cumva, cooriginare, care implică o transformare dintr-o stare de potențialitate într-una de actualitate. În aceeași măsură, Aristotel susține că potențialul și actul sunt două moduri diferite de a fi și de a exista, iar relația dintre ele este una de transformare și schimbare, prin care potențialul poate deveni actualitate. Dintr-o perspectivă axiologică, acesta afirmă că fiecare virtute sau viciu implică o anumită putere sau capacitate (capacitare); virtutea este capacitatea care se actualizează genuin prin acțiuni ce tind să fie cumva repetate, un fel de *eternă reîntoarcere a aceluiași* la Nietzsche. În acest mod, virtutea se formalizează și se încheagă prin formarea de obiceiuri pretins bune, astfel încât trebuie să ne străduim să îndeplinim potențialitatea, proiectul axiologic, idealitatea, nu numai prin exercitarea capacităților noastre, ci și prin transformarea acestora și punerea lor în act.

Din punct de vedere fenomenologic, *ratarea* în sens personal a unui individ poate fi mult relevantă decât *reușita* sa. Totuși, dacă privim în sens opozitiv, cum ar fi spus Heidegger, orice întoarcere împotriva noastră devine un soi de potențialitate întoarsă pe dos, deci tot un soi de aristotelism, deci tot un soi de logică, ba chiar de două ori, odată într-un sens personal, în care noi luăm decizii împotriva noastră, deci avem o logică interioară, iar o altă în sens mare, macro, în sens epistemic, în sensul

în care putem spune că *fenomenologia vieții de zi cu zi* vine împotriva logicii, are o logică în sine, specifică, chiar analitică în felul ei. Totuși, într-un alt plan epistemic, dacă ne ducem la Bergson, putem vedea *conștiința* ca durată în timp, ca acel interval încărcat metafizic, care ne scoate la iveală, în termenii de față, continuitatea. Sigur că la Bergson - durata este legată de intuiție. Având o postură intuitivă, extremitatea acesteia este chiar momentul, clipa, virtualul până la urmă, non-palpabilul care trebuie cuantificat matematic. În acest fel, momentul este generat de spațiu ca timp. Prin urmare, timpul spațializat devine un timp al perspectivismului restrâns, în loc de un spațiu al perspectivismului generalizat, sinoptic. [6]

De ce fac aceste precizări? Pentru că ele sunt strâns legate de faptul că omul nu poate fi gândit în afara fenomenologiei / a logicii / a epistemologiei. Adică, fără nicio altă aplecare, a filosofiei. Orice altă interpretare specifică din zona științelor pozitive va rata chiar esența omului, pentru că acesta va fi privit din interiorul unui domeniu de cercetare cu *reguli precise și clare*, cu axiome, postulate. Altfel spus, fără a exista vreo posibilitate de a ne întoarce metodologic împotriva țintei noastre epistemice, iar dacă nu avem această șansă, să ne renegociem premisele de lucru, atunci nu putem nici să înțelegem în esență ce este cu *natura umană*.

Ca instrumente ale filosofiei, putem folosi epistemologia, fenomenologia, logica, hermeneutica, antropologia, axiologia, chiar și multe altele, dar fără să le îmbrățișăm științific, fără să ne contopim metodologic cu ele, fără să fim nevoiți epistemic să deducem și să conturăm teorii ce derivă din postulate și axiome. De ce este așa de relevant acest lucru? Pentru că este nevoie, atunci când vorbim de natura umană, să rămână un *rest*, să rămână un *nescris existențial hermeneutic*, [7] altfel totul ar fi trunchiat, ar fi ideologie, ar fi metodă de lucru în sine, postulată și adaptată la un context anume, iar asta ar însemna tocmai falsificarea omului. Pe de altă parte, din perspectiva *inteligenței artificiale* nu este neapărat relevant ce este natura umană, ci cum poate fi simulată, recreată, regândită. Și nici măcar asta, ci problema este legată de conștiință sau singularitate. Și în acest fel, ajungem la teză – *conștiința este continuitate*?

### **Amfibologie existențial-tehnologică**

De ce inteligența artificială o să aibă succes? [8] Pentru că nimeni nu o să fie interesat de problema *conștiinței*, de faptul că avem de-a face cu un simulacru, ori poate că noi o investim cu „sens”. [9] Nu, noi o să alegem efectul, rezultatul, „datul din coadă”, faptul că ne livrează fără crâcnire, așa cum o face un câine. Dar în lipsa unei formulări mai bune, o

să fac apel la Veblen: „Pisica trăiește alături de om pe picior de egalitate;[...] Câinele prezintă avantaje atât din punct de vedere al inutilității, cât și din cel al calităților sale particulare.[...] Prin aceasta se înțelege că el este servitorul omului și că este înzestrat cu darul supunerii fără crâncire și cu iuțeala unui sclav în a ghici ce-i trece prin minte stăpânului său.[...] El este cel mai murdar dintre animalele domestice și cu cele mai dezgustătoare obiceiuri. Pentru aceasta se revanșează printr-o atitudine servilă și lingușitoare față de stăpânul său, [...] Astfel, câinele ne intră în grații făcând jocul înclinației noastre spre dominație și, dat fiind că întreținerea lui costă bani și, de regulă, nu servește niciunui scop productiv, el și-a asigurat locul în mintea oamenilor ca simbol al bunei reputații.” [10] Ce înseamnă de fapt acest lucru? Ne arată că natura umană este înclinată spre obediență, iar tendința naturală este aceea de a face din obiecte, oameni, animale - niște mijloace, nu scopuri în sine. Altfel spus, inteligența artificială, cum este și ChatGPT, va avea succes și are pentru că livrează informații și totodată pentru că simulează aproape orice ține de sfera umană. Cu trecerea timpului, nu va mai exista nicio diferență în plan ontologic-existențial. Va fi, astfel, o *amfibologie tehnologic-umană*.

Ce ar putea face inteligența artificială începând chiar de azi? [11] Ei bine, ar putea să *replice* persoanele decedate, în funcție de informațiile pe care le oferim, pornind de la poze, videoclipuri, un lexic specific și chiar voce. Gândiți-vă că, în tot acest simulacru, o să vă treziți pe *vocea mamei decedate*, *chiar o să beți o cafea cu ea*, iar eventual o să mergeți la un concert live susținut, evident, de un *artist mort*, dar *viu* în noua lume postmodernă. Gândiți-vă că pasiunile legate de o persoană vor putea fi replicate, iar o aplicație o să poată să vă furnizeze o relație cu acea persoană, fie că este în viață sau este moartă. În toată această *nebulosă tehnologică*, singura meserie relevantă ar fi aceea de *fidelizare-virtuală*, cum ar putea fi numită, adică de a găsi o persoană care înțelege nuanțele, pentru a le aproxima la un context nou filosofic. Și acestea sunt doar câteva chestiuni de moment, ce se schimbă chiar în momentul în care se scriu aceste rânduri.

Așa cum natura are un fel intrinsec de a fi, așa ar trebui privită și evoluția tehnologiei, transformarea și dezvoltarea ei în simulacre ce tind să fie *cu totul altceva*, ce tind să iasă din sfera noastră de cunoaștere și de poziționare epistemică. [12] Deși noi percepem extrinsec această evoluție tehnologică, ea derivă intrinsec, așa cum ar fi explicat și Aristotel, *din felul ei de a fi*. Faptul că noi o percepem cumva diferit, nu înseamnă că este diferită din perspectiva evoluției ei liniare și a transformărilor inerente [13] ce țin *intrinsec* de aparatul ei de proiectare tehnologică. În acest mod,

*inteligența artificială* nu este nimic altceva decât expresia firească a evoluției tehnologiei. Asta nu înseamnă însă că tinde spre umanitate, ci că umanitatea este replicată, pusă în straițe tehnice și conturată în simulacre ce tind să pară veritabile.

### Concluzii

Inteligența artificială generativă (AGI) și, implicit, noua realitate virtuală ne postulează viețile în relații suspendate, în lumi ce tind să fie aproape, deși sunt departe. De aceea, cu cât o să treacă timpul și o să avansăm tehnologic, cu atât o să ne pulverizăm viețile într-un salt continuu, într-o teleportare perpetuă, hermeneutică. Practic, se vor șterge barierele alterității, iar lumea va deveni în sine un imens sat global, conectat și interconectat la realități artificiale și simulacre ce tind să devină genuine. Dacă mai până acum puteam să vorbim de *Second Life* și de faptul că există vieți separate, ei bine de acum încolo vom avea lumi hermeneutice care se întrepătrund și se centreză în sfere pretins diferite, dar legate cumva ontologic la o înșiruire de dimensiuni simbolice. Va fi o lume pur fenomenologică, unde *depărtarea va fi apropiere în sine*, dar și apropierea va fi în lumi apuse, distincte, diferite. În această lume generată artificial, lumile reale vor fi totodată lumi fantastice, plătuite și puse parcă în planuri ce tind să pară firești, să fie intrinseci, dar de fapt să fie străine, absconse, non-genuine.

Pe de altă parte, în tot acest melanj *virtual-tehnic-mundan*, singura constantă o să trebuiască să rămână *conștiința*, adică *liniaritatea personală*, *durata intențională*, dacă ne gândim la Bergson, *conștientizarea clipei* într-un *hău existențial*, generat, culmea, de rețele tehnologice adaptate la progres. Nu numai că nu vor mai exista fășii de demarcație între oameni și roboți, dar vor fi și încercări pe măsura de a ne contopi hermeneutic într-un sălaș existențial. Dacă *inteligența artificială* o să simuleze conștiința, în sensul că noi o să o investim hermeneutic cu simbolistica conștiinței în sine, deci noi o să-i dăm *sens* fenomenologic, chiar epistemologic, atunci nu vor mai exista linii fine de demarcație între noi ca oameni și artefacte generate de inteligență artificială. Cu alte cuvinte, vom fi *simulacre*, *butaforii*, puse în lumi relative, într-un sens nietzschean, *postmoderne*, *ultra-tehnologizate*, ce tind să fie mai mult decât sunt. În acest sens, tot ce ține de Deep-Fake devine, pe filiera post-adevărului, veritabil în sine, adaptat la cutume de piață și servit ca atare pe filiera negoțului și a pecuniarului exacerbat. Adevărul, *relativ în sine*, devine, paradoxal, *întemeiat în sine*. Și astfel tot ce ține de *sfera cunoașterii* se transformă într-o explozie pur virtuală, contopită

hermeneutic, fără bariere logice și chiar ideologice, într-un *sfârșit al încapsulărilor*, într-o relativizare profundă, într-o *schimbare* fundamentală, dacă ne gândim la Peter Thiel. [14]

### **Bibliografie:**

1. În ceea ce privește abordarea filosofică și contribuția epistemică în domeniile actuale ale tehnologiilor complexe, dar și a computerelor, a programării, chiar a inteligenței artificiale, a se vede articolul disponibil la <https://www.theatlantic.com/technology/archive/2017/03/aristotle-computer/518697/>
2. Cf. Cf. SHANNON, . C. E, "A SYMBOLIC ANALYSIS OF RELAY AND SWITCHING CIRCUITS," IN *ELECTRICAL ENGINEERING*, VOL. 57, NO. 12, PP. 713-723, DEC. 1938, DOI: 10.1109/EE.1938.6431064.
3. Vezi mai multe detalii în lucrarea *At Play in the Field of Possibles, An essay on the Foundation of Self and Free-fantasy Variational Method*, Zeta Books, București, 2012, pp. 76 -92
4. Vezi mai multe aici, Chung-Hwan Chen, *Different Meanings of the Term Energeia in the Philosophy of Aristotle*. Philosophy and Phenomenological Research, vol. 17, no. 1, 1956, pp. 56–65. JSTOR, Disponibil: <https://doi.org/10.2307/2104687>
5. Despre potențialitate, vezi John J. Cleary, "'Powers that be': the concept of potency in Plato and Aristotle, Méthexis, vol. 11, 1998, pp. 19-64. Jstor, Disponibil: [www.jstor.org/stable/43738564](http://www.jstor.org/stable/43738564). Accessed 4 may 2023.
6. Cf. Henri Bergson, *Eseu asupra datelor imediate ale conștiinței*, Editura Dacia, Cluj-Napoca, 1993, pp. 64-99, dar și urm.
7. Despre ființă ca *semnificare* în sens ontologic, dar și ontic, vezi Elad Lapidot, *Être sans mot dire: La logique de Sein und Zeit*, Zeta Books, București, 2010, pp. 292-298
8. Am discutat pe larg în acest interviu despre implicațiile axiologice ale inteligenței artificiale, cât și despre faptul că noi o investim cu sens, ascultă aici <https://www.georgecolang.ro/2023/02/despre-chatgpt-bing-bart-si-inteligenta-artificiala-interviu-radio-mfm-dance/>
9. Sensul este dat *symbolic* din momentul utilizării. Practic, nu mai există o barieră psihologică, ci mai degrabă o hermeneutică ce aduce la pachet un simbolism fenomenologic. În acest fel, utilizând și folosind ustensilitatea, ne contopim hermeneutic. Mai multe vezi aici



<https://www.georgecolang.ro/2023/04/cum-ne-schimba-social-media-gandirea-valorile-si-stilul-de-viata-interviu-a7tv/>

10. Cf. Veblen, Thorstein, *Teoria claselor de lux*, Editura Publica, București, 2009, pp. 126-127
11. Despre provocările inteligenței artificiale, dar mai ales a faptului că încă nu există reglementări, am discutat în interviul următor.  
<https://www.georgecolang.ro/2023/02/gandirea-provocarile-si-oportunitatile-viitorului-interviu/>
12. Despre „camera chinezească” și faptul că sintaxa nu este suficientă, nu e o condiție necesară în sine pentru semantic, vezi Edward Feser, *De la Aristotel la John Searle și înapoi. Cauze formale, teleologie și calcul în natură*, Eikon, București, 2017, pp. 117-170
13. Am discutat, în acest articol, despre implicațiile biotehnologiilor asupra cursului firesc al vieților noastre, dar mai ales asupra valorilor, a câmpului axiologic care se transformă pe zi ce trece. Vezi pe larg aici, George Cosmin Colang, Loredana Terec-Vlad, "Biotechnologies and the Case of Surrogate Motherhood: Axiological Implications and Their Connection with the Law," *META: Research in Hermeneutics, Phenomenology, and Practical Philosophy* 14, no. 2 (December 2022): 665-676, Disponibil: [http://www.metajournal.org//articles\\_pdf/665-676-colang-terec-meta-tehno.pdf](http://www.metajournal.org//articles_pdf/665-676-colang-terec-meta-tehno.pdf)
14. Despre schimbare, nu neapărat ca progres, vezi conferința lui Peter Thiel, disponibil aici  
[https://www.youtube.com/watch?v=ibR\\_ULHYirs](https://www.youtube.com/watch?v=ibR_ULHYirs)

**Date despre autor:**

**George Cosmin COLANG**, Lect. univ. dr.  
Universitatea Bioterra din București, România  
Email: [georgecolang@yahoo.com](mailto:georgecolang@yahoo.com)