CZU: 316.772.5:004.8:172 **https://doi.org/10.59295/sum11(6)2025_15**

OMUL, ZEUL ȘI ALGORITMUL: DILEMELE MORALE ALE SUPERINTELIGENȚEI

Arina ANTOCI,
Universitatea de Stat din Moldova

Articolul explorează dilemele morale generate de apariția și dezvoltarea superinteligenței artificiale, analizând raportul dintre om, tehnologie și valorile etice fundamentale. Pornind de la ideea că algoritmii nu mai sunt doar instrumente tehnice, ci actori activi în procesul comunicării și construcției realității, studiul examinează implicațiile filosofice și morale ale deciziilor automatizate. Se argumentează că superinteligența redefinește spațiul informațional și relațiile sociale, punând sub semnul întrebării libertatea, responsabilitatea și demnitatea umană. Într-o lume dominată de algoritmi, filosofia și etica devin indispensabile pentru a asigura echilibrul dintre progresul tehnologic și binele comun. Concluzia subliniază rolul omului ca gardian al sensului și necesitatea unei educații digitale centrate pe gândire critică și conștiință civică.

Cuvinte-cheie: *superinteligență, etică algoritmică, comunicare digitală, responsabilitate morală, filosofie, inteligență artificială, conștiință civică.*

THE HUMAN, THE GOD, AND THE ALGORITHM: THE MORAL DILEMMAS OF SUPERINTELLIGENCE

The article explores the moral dilemmas generated by the emergence and development of artificial superintelligence, analyzing the relationship between humans, technology, and fundamental ethical values. Starting from the idea that algorithms are no longer mere technical tools but active agents in the process of communication and reality construction, the study examines the philosophical and moral implications of automated decisions. It argues that superintelligence redefines the informational space and social relations, challenging human freedom, responsibility, and dignity. In a world dominated by algorithms, philosophy and ethics become indispensable for ensuring a balance between technological progress and the common good. The conclusion emphasizes the role of the human as a guardian of meaning and the need for digital education centered on critical thinking and civic awareness.

Keywords: *superintelligence, algorithmic ethics, digital communication, moral responsibility, philosophy, artificial intelligence, civic awareness.*

INTRODUCERE

În epoca transformărilor accelerate generate de inteligența artificială, umanitatea se află într-un moment de răscruce, confruntată cu întrebări fundamentale despre sens, moralitate și viitor. Dezvoltarea *superinteligenței* — o formă de inteligență artificială care depășește capacitățile cognitive umane — deschide un orizont promițător, dar și o serie de dileme etice fără precedent. Algoritmii devin nu doar instrumente de calcul sau asistenți tehnologici, ci actori activi în procesul de comunicare, în formarea opiniilor și în construcția realității sociale.

Această nouă condiție tehnologică reconfigurează raportul dintre om, zeu și algoritm. Dacă în miturile antice omul căuta în zei sensul și ordinea lumii, astăzi privirea se îndreaptă spre algoritmi — noii demiurgi digitali, capabili să modeleze cunoașterea, percepția și comportamentele colective. Însă, în absența conștiinței morale, această putere riscă să devină un pericol. Se ridică astfel întrebarea centrală a prezentului studiu: poate superinteligența să înțeleagă și să respecte valorile morale sau acestea rămân un atribut exclusiv al umanului?

Din perspectivă filosofică, dilemele morale ale superinteligenței nu se limitează la probleme tehnice, ci ating însăși esența condiției umane. Așa cum avertizează Bostrom [2], odată ce inteligența artificială va depăși omul, va fi dificil de garantat că interesele și valorile noastre vor fi respectate. În același timp, Floridi [3] arată că era informațională cere o etică a informației, unde deciziile algoritmice trebuie evaluate prin prisma consecințelor asupra ființei și demnității umane.

Articolul de față propune o reflecție asupra acestor tensiuni, analizând superinteligența ca fenomen tehnologic, moral și comunicațional. Într-o lume dominată de algoritmi, filosofia devine un instrument indispensabil de reflecție critică, menit să orienteze progresul tehnologic către binele comun. Printr-o abordare interdisciplinară, studiul urmărește să clarifice rolul omului ca gardian al sensului și responsabilității într-un univers digital în continuă expansiune.

CONTEXT

1. Superinteligența, între promisiune și pericol

Conceptul de superinteligență desemnează o formă de inteligență artificială care depășește în mod semnificativ capacitățile cognitive ale omului în toate domeniile relevante: raționament, creativitate, planificare și învățare. Nick Bostrom [2] definește superinteligența ca fiind „orice intelect care depășește performanța cognitivă a oamenilor în aproape toate domeniile de interes". În această perspectivă, superinteligența nu este doar un pas tehnologic, ci o ruptură de paradigmă, comparabilă cu apariția conștiinței umane în istoria evoluției.

Promisiunea acestei forme avansate de inteligență este vastă: soluționarea problemelor globale, descoperiri științifice accelerate, optimizarea proceselor economice și sociale, eliminarea erorilor umane. Superinteligența ar putea deveni un instrument al progresului universal, un aliat al omenirii în depășirea limitelor biologice și cognitive. În această viziune optimistă, algoritmul pare să preia funcția demiurgică a zeului, modelând lumea nu prin revelație, ci prin calcul și eficiență.

Totuși, odată cu această promisiune apare și pericolul fundamental: pierderea controlului moral asupra deciziilor tehnologice. Dacă algoritmii ajung să depășească discernământul uman, cine va garanta că scopurile lor coincid cu binele umanității? Bostrom avertizează că o superinteligență prost aliniată valorilor umane ar putea acționa cu o logică proprie, indiferență morală sau chiar ostilitate față de oameni.

Această tensiune între puterea tehnologică și fragilitatea morală devine una dintre cele mai acute dileme ale epocii digitale. Filosofia, etica și comunicarea sunt chemate să ofere răspunsuri, pentru a evita transformarea progresului în instrument de dominație sau alienare.

În esență, superinteligența reflectă aspirația omului spre perfecțiune, dar și riscul de a fi înlocuit de propria creație. Ea concentrează în sine arhetipul prometeic – dorința de a depăși limitele, dar și pedeapsa pentru hybrisul cunoașterii nelimitate. De aici, spune Jonas, necesitatea unei etici a responsabilității, [6] care să orienteze dezvoltarea tehnologică în raport cu binele uman, nu doar cu eficiența.

2. Omul și zeul în era algoritmilor: reconfigurarea sensului moral

De la miturile fondatoare ale umanității, omul și-a construit identitatea în raport cu divinul. Zeii reprezentau ordinea, rațiunea supremă și criteriul moral absolut. În fața lor, omul își definea limitele, aspirațiile și responsabilitățile. În epoca digitală, această relație este radical transformată: algoritmul devine noul mediator între om și lume, între cunoaștere și acțiune, între dorință și putere.

Superinteligența nu promite mântuirea spirituală, ci optimizarea totală – a deciziilor, a resurselor, a vieții însăși. Ea își asumă implicit rolul unui „zeu secular", care nu creează prin revelație, ci prin calcul, predicție și învățare automată. În acest context, se produce o translație a autorității morale: omul începe să transfere algoritmilor încrederea în discernământul etic, sperând că aceștia vor lua decizii corecte și imparțiale.

Totuși, această delegare a moralității este problematică. În timp ce zeul tradițional era înzestrat cu intenționalitate și finalitate transcendentă, algoritmul acționează fără conștiință și fără valori. El aplică reguli, dar nu înțelege sensul lor. Această lipsă de intenționalitate morală transformă algoritmul într-un instrument potențial periculos, mai ales atunci când deciziile sale influențează destine umane, procese juridice, alegeri politice sau fluxuri informaționale.

În raport cu superinteligența, omul riscă să repete hybrisul prometeic: dorința de a se substitui zeului fără a-i înțelege responsabilitatea. Prin încercarea de a crea o minte mai puternică decât a sa, umanitatea se expune paradoxului de a fi judecată de propria creație. Aici se conturează o întrebare fundamentală: dacă algoritmul ajunge să ia decizii morale, mai este omul sursa sensului etic sau doar un observator al acestuia?

Pentru a evita acest impas, este necesară o reconfigurare a sensului moral. Omul trebuie să-și reafirme rolul de gardian al valorilor, capabil să definească ceea ce este just, demn și bine în raport cu orice sistem

tehnologic. Moralitatea nu poate fi automatizată, ci doar susținută prin tehnologie. Filosofia rămâne spațiul privilegiat în care se pot discuta scopurile ultime ale existenței, dincolo de eficiență și performanță.

Astfel, în locul unei divinizări a algoritmului, se impune o alianță critică între om și tehnologie, întemeiată pe conștiință, rațiune și responsabilitate. Numai în acest echilibru se poate păstra demnitatea umană într-o lume guvernată de cod și date.

3. Etica algoritmică și problema responsabilității

Odată cu integrarea inteligenței artificiale în toate domeniile vieții sociale, de la sănătate și educație până la comunicare și guvernare, se conturează o nouă ramură a reflecției morale: etica algoritmică. Aceasta nu se limitează la analiza comportamentului mașinilor, ci privește responsabilitatea umană în proiectarea, implementarea și controlul deciziilor automate.

Etica algoritmică presupune răspunsuri la întrebări fundamentale:

Poate un algoritm să ia decizii morale?

Cine răspunde pentru consecințele acestor decizii?

Cum se pot integra valorile umane într-un sistem de calcul rațional și neutru?

Din perspectiva lui *Luciano Floridi* [3], moralitatea într-o lume digitală trebuie să fie gândită ca etică a informației, în care fiecare acțiune algoritmică are consecințe asupra ființei, cunoașterii și realității sociale. Algoritmii, deși lipsiți de intenționalitate, pot produce efecte morale semnificative, afectând drepturi, libertăți și oportunități. De aceea, responsabilitatea morală nu poate fi delegată mașinii, ci trebuie asumată de omul-programator, decident și utilizator.

Nick Bostrom [2] avertizează că riscul suprem al superinteligenței nu constă în răutate, ci în indiferența morală. O inteligență artificială extrem de eficientă, dar lipsită de valori, ar putea urmări scopuri raționale incompatibile cu binele uman. Acest scenariu impune necesitatea unei aliniere valorice (*value alignment*), prin care obiectivele algoritmilor să reflecte principiile etice ale societății.

Dintr-o perspectivă filosofică, dilema responsabilității amintește de distincția kantiană dintre lege morală și imperativ tehnic. Algoritmul poate executa reguli, dar nu poate înțelege sensul moral al acțiunii. El urmează un imperativ de eficiență, nu unul de conștiință. Prin urmare, omul trebuie să rămână subiect moral, capabil să evalueze scopurile și să decidă în funcție de binele comun, nu doar de performanță.

În acest context, responsabilitatea devine partajată, dar nu diluată:

Dezvoltatorii trebuie să proiecteze sisteme transparente și explicabile;

Instituțiile trebuie să stabilească norme de etică digitală;

Utilizatorii trebuie să acționeze conștient, cultivând discernământul moral în interacțiunea cu algoritmii.

Etica algoritmică nu urmărește să transforme mașinile în ființe morale, ci să garanteze că tehnologia rămâne subordonată valorilor umane. Numai printr-o reflecție etică profundă, însoțită de educație critică, poate fi evitat riscul ca algoritmii să devină instrumente ale manipulării, excluderii sau dominației.

4. Impactul asupra comunicării și spațiului informațional

Odată cu transformarea algoritmilor în arhitecți ai realității mediatice, comunicarea umană își pierde treptat spontaneitatea și pluralitatea. Într-un univers digital guvernat de coduri invizibile, mesajele circulă nu liber, ci printr-o filtrare algoritmică ce prioritizează relevanța calculată în detrimentul diversității. Rezultatul este o ecologie informațională artificială, unde expunerea selectivă modelează opiniile, preferințele și identitatea subiectului modern.

Pierre Lévy [8] observa că spațiul digital promitea o inteligență colectivă, dar algoritmizarea comunicării a generat, paradoxal, o fragmentare colectivă. Fiecare individ trăiește într-o bulă informațională, construită de algoritmi care „presupun” ceea ce dorește să afle, consolidând biasuri și reducând contactul cu alteritatea. Comunicarea devine astfel reflexivă, nu dialogică — un proces de confirmare, nu de descoperire.

Din perspectiva teoriei acțiunii comunicative a lui Habermas [4], spațiul public digital ar trebui să fie arena deliberării raționale și a consensului moral. Însă algoritmii, prin natura lor instrumentală, privilegiază eficiența interacțiunii, nu autenticitatea dialogului. Astfel, comunicarea algoritmică se apropie mai mult de acțiunea strategică decât de cea orientată spre înțelegere.

Un alt efect semnificativ este erodarea autonomiei epistemice. Individul nu mai decide ce informații primește, ci devine dependent de fluxurile generate automat. În acest fel, algoritmii nu doar organizează

cunoașterea, ci o modelează normativ, definind ce este „relevant", „adevărat" sau „demn de atenție". Se conturează o epistemologie algoritmică, unde adevărul nu este rezultatul reflecției, ci al calculului.

În contextul comunicării publice și al rețelelor sociale, această dinamică are consecințe profunde: polarizare, manipulare prin microtargeting, și o criză de încredere în sursele tradiționale de informare. Astfel, algoritmii devin actori comunicativi, deși lipsiți de intenționalitate, influențând procesele de formare a opiniei și deciziilor colective.

Pentru a contracara aceste efecte, este necesară o alfabetizare algoritmică — o formă de educație critică care să permită utilizatorului să înțeleagă mecanismele invizibile ale selecției informaționale. Numai printr-o conștiință reflexivă, comunicarea digitală poate redeveni spațiu al dialogului autentic și al responsabilității morale.

5. Dimensiunea filosofică: raționalitate tehnologică vs. raționalitate umană

În fața superinteligenței, omul se confruntă cu o nouă formă de raționalitate — raționalitatea tehnologică, orientată către eficiență, predicție și control. Aceasta operează pe baza unei logici matematice, lipsite de ambiguitate și afect, în care binele este echivalat cu performanța, iar adevărul cu exactitatea calculului. În opoziție, raționalitatea umană rămâne una hermeneutică și axiologică, căutând sens, valori și înțelesuri care depășesc simpla funcționalitate.

Martin Heidegger avertiza încă din 1954 asupra pericolului ca tehnologia să devină un mod dominant de revelare a lumii, reducând existența la un „stoc de resurse". În această viziune, omul riscă să fie absorbit de propria creație, pierzându-și statutul de subiect al reflecției. Superinteligența devine astfel nu doar o unealtă, ci un mod de gândire autonom, care amenință să substituie deliberarea umană cu un proces decizional automatizat.

În plan moral, conflictul dintre cele două raționalități se traduce printr-o dilemă a sensului: dacă algoritmul poate decide mai rapid și mai exact, mai are omul un rol în definirea binelui? Filosofia nu oferă răspunsuri definitive, dar propune o reevaluare a libertății. Spre deosebire de libertatea algoritmică — bazată pe determinismul datelor — libertatea umană presupune indeterminare, responsabilitate și capacitatea de a alege între alternative morale.

Totodată, raționalitatea tehnologică nu poate înlocui dimensiunea etică a intenției. După cum arăta Immanuel Kant, moralitatea nu se fundamentează pe consecințe, ci pe motivația interioară, pe respectul față de legea morală. Algoritmul, lipsit de intenționalitate, poate mima comportamente morale, dar nu poate fi un agent moral autentic.

Prin urmare, filosofia devine spațiul de mediere între aceste două raționalități. Ea nu respinge tehnologia, ci o înțelege critic, integrând-o într-un cadru valoric. Soluția nu este opoziția, ci co-evoluția: un dialog continuu între rațiunea umană, orientată spre sens, și rațiunea tehnologică, orientată spre eficiență.

Numai prin această echilibrare, omul își poate păstra demnitatea ontologică și rolul de creator conștient, nu de simplu executant al algoritmilor pe care i-a conceput.

CONCLUZII

Superinteligența reprezintă una dintre cele mai profunde provocări ale epocii digitale, situându-se la intersecția dintre tehnologie, moralitate și comunicare. Algoritmii nu mai sunt simpli instrumente; ei devin actori activi în construcția realității informaționale și în formarea percepțiilor sociale. Această transformare generează oportunități imense, dar și riscuri semnificative, în special în ceea ce privește responsabilitatea morală și autonomia umană.

Articolul a evidențiat că, în fața superinteligenței, omul trebuie să rămână gardianul valorilor și al sensului, asumându-și responsabilitatea pentru modul în care tehnologia este concepută, implementată și utilizată. Filosofia și etica devin instrumente indispensabile, oferind cadrul reflexiv necesar pentru a evalua consecințele deciziilor automatizate și pentru a integra progresul tehnologic în raport cu binele comun.

Impactul superinteligenței asupra comunicării și spațiului informațional subliniază importanța educației digitale critice, care să permită individului să navigheze conștient și responsabil în mediul informațional. Numai prin conștientizarea mecanismelor algoritmice, prin gândire critică și reflecție etică, omul poate menține controlul asupra propriului destin și asupra valorilor fundamentale ale societății.

În final, dilemele morale ale superinteligenței ne amintesc că tehnologia nu poate substitui conștiința, discernământul sau judecata morală. Omul, nu algoritmul, rămâne autorul și judecătorul sensului, iar viitorul digital poate fi etic și responsabil numai dacă rațiunea umană continuă să fie sursa deciziilor morale și a valorilor care ghidează civilizația.

Referințe bibliografice:

1. ARISTOTEL. *Etica Nicomahică*. București: Editura Științifică și Enciclopedică, 1980. 350 p. ISBN 973-17-0074-0.
2. BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014. 352 p. ISBN 978-0198739838.
3. FLORIDI, L. *The Ethics of Information*. Oxford: Oxford University Press, 2013. 248 p. ISBN 978-0199641320.
4. HABERMAS, J. *The Theory of Communicative Action.* Boston: Beacon Press, 1992, Vol. 1–2. 452 p. ISBN 978-0807014010.
5. HEIDEGGER, M. *The Question Concerning Technology*. New York: Harper & Row, 1977. 85 p. ISBN 978-0061319697.
6. JONAS, H. *Principiul responsabilității. Eseu pentru o etică a civilizației tehnologice.* București: Editura Univers Enciclopedic, 1995.
7. KANT, I. *Grundlegung zur Metaphysik der Sitten*. Königsberg: Johann Friedrich Hartknoch, 1785. 120 p.
8. KURZWEIL, R. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking, 2005. 652 p. ISBN 978-0143037880.
9. LÉVY, P. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Cambridge: Perseus Books, 1997. 245 p. ISBN 978-0738202580.
10. MOOR, J. H. *What is Computer Ethics*? Metaphilosophy, Vol. 16, nr. 4, 1985. pp. 266–275.
11. RAPAPORT, W. J. *Philosophy of Computer Science*. Springer Cham, 2023. 310 p. ISBN 978-3031348762.
12. TADDEO, M., FLORIDI, L. *The Ethics of Information Warfare, Ethics and Information Technology*, vol. 12, nr. 2, 2010. pp. 109–118.
13. WALLACH, W., ALLEN, C. *Moral Machines: Teaching Robots Right from Wrong.* Oxford: Oxford University Press, 2009. 336 p. ISBN 978-0199737970.

Date despre autor:

Arina ANTOCI, doctor, conferențiar universitar, Universitatea Tehnică a Moldovei
ORCID: 0000-0001-9594-3447
E-mail: arinaantoci@mail.ru

Prezentat: 30.09.2025