

I. EDITORIAL

O PERSPECTIVĂ FILOSOFICĂ ȘI JURIDICĂ ASUPRA „DREPTURILOR ROBOȚILOR”

Notar public **Gina PALADI**

Ce este un robot sau cine este el? Care este diferența dintre un mecanism și o ființă vie, aici incluzând toate ființele vii. Suntem creați și suntem și noi, oamenii, creatori? Care este diferența dintre condiția umană și condiția unei creații umane, adică un robot? Care este deosebirea dintre un robot care execută acțiuni mecanice și un robot dotat cu inteligență artificială (IA), ce se „supune” algoritmilor din propria sa programare? Ne putem aroga ipostaza de „mici zei” care pot controla totul, destinul fără teama de fatalitate? Ce drepturi ar trebui să aibă un om, un robot și o ființă umană care a primit avantajele avalanșei evoluției științifice?

Demersul nostru în legătură cu aceste interogații ontologice pornește de la două premise fundamentale ale istoriei cunoașterii umane.

În primul rând, este binevenită lucrarea filosofului Immanuel Kant intitulată „Bazele metafizicii moravurilor” (în unele traduceri din limba germană, „Bazele metafizicii moralei” sau „Întemeierea metafizicii moravurilor”) în care sunt descrise două așa-numite imperative: imperativul ipotetic și imperativul categoric. În ceea ce privește imperativul ipotetic vorbim despre utilizarea rațiunii într-un scop bun sau rău, dar și utilizarea propriei rațiuni spre atingerea binelui personal doar ca mijloc spre atingerea acestui scop.

La rândul său, imperativul categoric se referă la o altă temă îndelung discutată, cea a liberului arbitru, aceea care se referă la acțiunea omului prin care, în mod necondiționat, vizează o paradigmă ce descrie scopul în sine, acela al binelui, nu o dorință personală. *Omul, în integralitatea sa, nu poate fi doar o unealtă, ci un scop în sine.*

A doua premisă face referire la scrierile unui cunoscut autor de literatură science-fiction, Isaac Asimov, cel care a definit cele trei legi ale roboticii. *Aceste legi pot fi supuse unor numeroase discuții ce pot echivala cu „Dialogurile” din tinerețea lui Platon în care acesta descrie moșirea adevărului prin întrebări și răspunsuri ce ne apropie de acesta, de adevăr. Aceste trei legi sunt adresabile mai degrabă creatorului.*

Le vom transcrie *ad litteram*:

1. Un robot nu are voie să pricinuiască vreun rău unei ființe umane sau, prin neintervenție, să permită ca unei ființe omenești să i se facă un rău;

2. Un robot trebuie să se supună ordinelor date de către o ființă umană, atât timp cât ele nu intră în contradicție cu Legea 1;

3. Un robot trebuie să-și protejeze propria existență, atât timp cât acest lucru nu intră în contradicție cu Legea 1 sau Legea 2.

Rămâne aceeași întrebare: Omul este creator care a fost creat sau un creator în sine? Din această perspectivă putem reformula legile lui Isaac Asimov:

1. Unui robot îi este interzis să facă rău unui om sau unui grup de oameni prin absența acțiunii, dar și prin lipsa conștiinței de sine;

2. Un robot care a dobândit conștiință de sine și implicit conștiință, poate avea liber arbitru doar cu respectarea Legii 1;

3. Un robot cu conștiință de sine însuși își va apăra propria sa integritate ținând cont de Legile 1 și 2.

Dincolo de aceste aspecte este necesară o posibilă redefinire a acestor trei legi și din perspectiva conceptului de stat, ca națiune, ca teritoriu, cultură, spiritualitate, economie, natură, resurse, potențial și, mai ales, integritate morală în consecința eticii și a stabilirii unui set de reguli ce subliniază ceea ce punctează fără urmă de tăgadă un simț universal – **bunul simț**. Din această perspectivă, în mod firesc, vom aborda și alți factori ce transpun în mod ineluctabil conceptele de drepturi de bază sau, de ce nu, universale ale inteligenței artificiale.

Religia – o constantă a istoriei umanității –, spiritualitatea – element definitoriu al culturii umane –, precum și conceptul de moralitate au avut un impact decisiv asupra evoluției omului. Este imposibil de enumerat într-un spațiu atât de limitat efectele benefice, dar și distructive pe care credința le-a avut. Și atunci, în mod firesc, întrebăm: un robot ar acționa cu credință, convingere absolută așa cum au acționat oamenii în nenumărate momente ale existenței lor? S-ar putea ajunge la situația în care un robot dotat cu inteligență artificială ar simți și ar avea convingeri care să treacă și prin liberul arbitru? Sau peste liberul arbitru?

Un aspect demn de consemnat este termenul „facticitate” creat de filosoful german Johann Gottlieb Fichte. Cu alte cuvinte tot ceea ce este dat, de exemplu realitatea lumii, este dat fără nicio argumentare, fără a fi dedusă printr-un argument logic (*dat factice*). Acest termen a fost valorificat mai târziu de alți doi filosofi în teritoriul existențialismului – Martin Heidegger și Jean Paul Sartre – conform ideii că existența noastră și chiar existența în sine este de neînțeles în ea însăși. Pur și simplu suntem aruncați în lume fără drept de apel. Începem să existăm fără a fi conștienți de noi înșine. În aceeași ordine de idei, unul dintre părinții psihologiei transpersonale, Stanislav Grof, propune o paradigmă ce a constituit obiectul unor numeroase studii, anume faptul că fătul este conștient de sine. Descrie și distinge patru etape ale evoluției intrauterine. Le numește „matrici perinatale fundamentale”. Putem susține că un robot dotat cu inteligență artificială alege să fie creat? Este „aruncat” în existență, în lume? Răspunsul posibil este faptul că omul, cu excepția unor elemente genetice moștenite și care pot determina manifestări de multe ori inevitabile, este

liber să se orienteze spre propriul sens care să determine cursul vieții sale. Nu există, încă, lege sau destin care să fie date ca o necesitate și aprioric pentru felul în care ne reglăm modul în care existăm, dar există subterfugiul onest de a ne depăși multe limite, de a învăța, de a ne adapta. *Exact asta se petrece și cu inteligența artificială, dar fără ceea ce are omul ca cel mai neprețuit dar, Sinele.*

Din această extrem de sintetică privire apare interogația: care este diferența dintre drepturile omului și una din creațiile sale, inteligența artificială?

Eforturile de a legitima acest nou context, cel al apariției inteligenței artificiale care este similar cu momentul nașterii și, mai apoi, cel al exploziei revoluției industriale la sfârșitul secolului al XVIII-lea și începutul secolului al XIX-lea, sunt uriașe și extrem de complexe. Și în cazul unui robot dotat cu inteligență artificială este valabilă întrebarea „Cine sunt eu?”, așa cum este validată și în cazul omului. Aici survine o altă întrebare: Ce este mai important, individul sau societatea? Această întrebare face referire și la perspectiva din care privea lucrurile Aristotel și al său concept numit „zoon politikon”. Lumea va continua să existe fără un om sau altul, precum un om va continua să trăiască fără o mână, fără un picior sau alte componente naturale. Dincolo de acest teritoriu extrem de generos al analizei unei etici pe care am numi-o „Dincolo de bine și de rău”, invocând titlul unui volum scris de filosoful Friedrich Nietzsche, apar alte întrebări: Ce este dreptatea? Cum s-ar manifesta un om cu putere absolută sau un robot cu putere absolută, adică în ipostaza de Dumnezeu? În acest moment al cugetării intervine faptul că universul respectă niște legi clare, inclusiv anomaliiile care apar.

Survine o altă dilemă subliniată de cunoscutul mit al lui Gyges descris de Platon în dialogul Republica, în care se pune o problemă etică fundamentală: cum ar fi folosit un inel al puterii aproape absolute (e vorba de un inel care are puterea de a face purtătorul să devină invizibil) de către un om considerat bun, respectiv de către un om considerat rău. În mitul lui Gyges sunt subliniate mai multe aspecte problematice, de pildă: originile dreptății care reprezintă sursa dreptului în sensul cel mai integral al termenului; suntem mai degrabă corecți în interacțiunile cu ceilalți din cauza slăbiciunii noastre de a face rău și a posibilității să fim nedreptățiți; dorința de a avea mai mult (bani, putere, control) ispitește spre acte ilicite și acțiuni ulterioare prin care se poate fenta dreptatea; suntem drepti, onești din obligație sau prin propria noastră natură umană. Se argumentează, în acest sens, că niciun om nu va fi drept dacă i se va permite să își satisfacă propriile interese dincolo de bine și de rău.

Tema mitului lui Gyges este, cel puțin ca ipoteză de lucru, și despre robotul cu inteligență artificială. Există această distincție dintre bine și rău pentru inteligența artificială care poate dispune de dispozitive mecanice cu mult superioare posibilităților corpului uman? În acest caz vorbim despre o dilemă cu trăsături filosofice ce include niște constatări irefutabile. Faptul că omul predispus la a face fapte lipsite de moralitate este beneficiarul unor avantaje nedrepte, iar cel care „alege” un comportament virtuos este „condamnat” la suferințe nemeritate. Sau meritate? Acest

subiect poate fi integrat într-o perspectivă care a fost revoluționară, la vremea sa, de Friedrich Nietzsche în cartea „Așa grăit-a Zarathustra”, ceea ce va constitui următorearea provocare ce vizează inteligența naturală și cea artificială. Suntem născuți stăpâni sau sclavi? Roboți sau ființe conștiente de sine?

Filmul „Surrogates”, cu actorul Bruce Willis în rolul principal, punctează un aspect fundamental: *Substituirea omului, a sufletului ființei umane, cu un soi de „avatar”, fie el fizic sau virtual, este o soluție falimentară. Și chiar este.* Ce ființă umană, indiferent de preluările genetice pe care trebuie, într-un fel sau altul, să le accepte, ar fi de acord să fie reprezentată de un alter-ego robotizat fără să își mai pună problema ieșirii din casă, în natură, fără îmbrățișări, fără atingerea și privirea care conectează două suflete, de exemplu? Cum percepe un robot vederea unei flori, a unei albine, a frumuseții miraculoase a naturii?

În esență, întreaga planetă este interconectată și orice dezechilibru, oricât de mic, poate influența printr-un efect de domino multidimensional. Cum percepe lumina un robot cu inteligență artificială? E un sentiment de bucurie, de îmbrățișare a miracolelor naturii sau un proces prestabilit al algoritmilor din propria sa programare?

Discuția ce vizează statutul robotului în societatea umană mai implică și aspectele psihologice. Dacă aprofundăm și mai mult această perspectivă, ne putem întreba dacă un robot cu inteligență artificială are subconștient în sincretism cu posibilitatea de a deveni conștient. Putem vorbi despre reprimarea unui gând, a unei emoții în cazul unui robot? Și aici mă refer la actul ratat descoperit în lucrarea lui Sigmund Freud „Psihopatologia vieții cotidiene”. Dacă luăm în considerare comparativă modul în care un om refulează anumite manifestări din spectrul trăirilor sale, dar le scapă sub forma exprimării verbale este aproape imposibil să validăm existența subconștientului unui robot cu IA. Există doar în cazul în care utilizăm un eufemism și considerăm că algoritmii din programare pot constitui bazele unui pseudo-subconștient.

În lucrarea lui Carl Gustav Jung, „Dinamica inconștientului”, autorul face o conexiune între conceptul de inconștient (a se face distincția de subconștient) și arhetipuri. Spre exemplu: Anima/ Animus, umbra sau bătrânul înțelept. În viziunea lui Jung, arhetipul este un dat factice ereditar care poate sau în mod ineluctabil modelează conștiința unui om. Așa-numitele „imagini primordiale” ale lui Jung trădează comportamentele innăscute ale unui om și care sunt în strânsă legătură cu miturile repetitive propulsate în istoria umanității. Dincolo de interogațiile ce sunt absolut firești rămân totuși câteva întrebări legitime: va putea într-o zi un robot să devină zeu sau, mai mult, Dumnezeu? Să aibă dreptul de a decide cu privire la viața unui om sau la sfârșitul acesteia? Să judece ce este bine și ce este rău? Să impună, precum a făcut omul, exigențe morale pentru o mai bună echilibrare a existenței și o diminuare a consecințelor distructive ale acestei existențe?

Dacă ne-am gândi la cuvântul iubire în cazul roboților IA, am putea presupune că roboții s-ar putea „reproduce”, iar un robot ar fi „mamă” și un alt robot ar fi „tată”. Cum ar înțelege iubirea inteligența artificială într-un trup mecanizat? Un posibil răspuns la întrebările menționate mai sus: *Iubirea nu poate fi plagiată*. Este apanajul omului, al existenței de sine în mod deplin conștient.

Se pot impune legi asupra iubirii pure, autentice, spirituale? Este o întrebare generică ce implică multe detalii, unele dintre ele fiind extrem de controversate. În această direcție este imperativ necesar *legiferarea drepturilor inteligenței artificiale și a implicațiilor inevitabile*. Controlul reprezintă disputa dintre putere și libertate.

În mod paradoxal aceste două concepte pot fi în echilibru, în armonie. Diferența folositoare o face faptul de a fi conștient de sine. Această stare este un firesc atunci când iubirea este autentică între doi oameni. Între doi roboți oare cum ar fi? Niciun sentiment uman nu poate fi implementat în presupusul suflet al unui robot care, finalmente, este creația omului. V-ați lăsa controlați de inteligența artificială sau de propriile alegeri care sunt parțial dictate de moștenirile genetice care predispun, dar nu dispun?

Unul dintre momentele de referință care marchează istoria confruntării dintre om și un super computer este cea dintre marele maestru Garry Kasparov (Garik Kimovitch Vainstein) și Deep Blue (denumit inițial Deep Thought, Gândul profund și ulterior Deeper Blue). Este vorba despre sportul minții, șahul. S-au desfășurat două meciuri. Supercomputerul a fost produs de compania IBM și a fost îmbunătățit după prima dispută (ambele confruntări au fost stabilite pe durata a șase meciuri). Disputa a fost foarte echilibrată, dar dincolo de rezultatele celor două confruntări rămâne această dihotomie a modului în care acționează două perspective de a percepe lumea.

Consider că existența umană s-a divizat, în mod explicabil, în omul care simte, este empatic și manifestă în mod constant consecințele acestei ipostaze (bunătate, altruism, iubire) și omul-robot, practic cel care a ales să se conformeze unor reguli stricte, să își accepte programarea. Această fractură existențială poate fi supusă unei hermeneutici ce este esențială pentru a ne înțelege pe noi înșine prin introspecție.

Concluzia este evidentă. Omul a creat o multitudine de mijloace de comunicare, dar cel mai intens și mai profund mod de conectare este iubirea.

Bibliografie

1. Asimov Isaac, Fundația, ediție completă, Editura ART, 2021;
2. Fichte J. G., Doctrina științei, Editura Humanitas, 1995;
3. Freud S., Despre actele ratate, Editura TREI, 2021;

4. Freud S., Psihopatologia vieții cotidiene, Editura TREI, 2010;
5. Jung C.G., Arhetipurile și inconștientul colectiv, Editura TREI, 2014;
6. Jung C.G., Dinamica inconștientului, Editura TREI, 2013;
7. Kant I., Întemeierea metafizicii moravurilor, Editura Humanitas, 2007;
8. Kant I., Critica rațiunii practice, Editura IRI, 1993;
9. Nietzsche F., Așa grăit-a Zarathustra, Editura Antet Revolution, 2003;
10. Platon, Republica, Editura Antet Revolution, 2010.