INTERVIU CU PROFESORUL CONSTANTIN VICĂ: „INVITAȚIA DE A ACCEPTA NECRITIC ȘI TOLERANT MECANISMUL VIOLENT ȘI TOTAL NE-CRITIC AL AUTOMATIZĂRII TREBUIE REFUZATĂ”

Autor: Teodora Ciobanu | 24 ianuarie 2022



Constantin Vică *este lector la Facultatea de Filosofie a Universității din București. Susține cursuri de Etică în tehnologia informației, Etica și politica noilor tehnologii, Tehnologie și dezvoltare, Morality, Business and Information Technology sau Cultură digitală. Ariile sale de interes sunt plasate în etica informației și computerelor, filosofia informaticii și inteligenței artificiale, bioetică, filosofie socială și politică. A publicat diverse articole despre internet și ameliorarea cognitivă, Inteligența Artificială și etica viitorului, responsabilitatea sistemelor de IA, epistemologia și politica datelor deschise, politica piraților, roboții ca asistenți și companioni, conflictul dintre instituțiile proprietății intelectuale și normele informale online. Publică în revistele Ethics and Information Technology, Science and Engineering Ethics, International Journal of Social Robotics, Ethics in Biology, Engineering and Medicine, Public Reason etc. A editat, în 2019, Vremuri noi, vremuri vechi. Jurnal 2007–2013 de Mihail Radu Solcan, primul filosof-programator din România, poliglot și polimat.*

În prezent, coordonează programul de master în etică aplicată la Facultatea de Filosofie, predă în cadrul programului de master Philosophy, Politics and Economics (PPE) și este director-adjunct al Centrului de cercetare în etică aplicată, Universitatea din București. Mai multe detalii despre activitatea dumnealui pot fi găsite aici.

1. Auzim foarte des ideea că tehnologia este doar un instrument sau mijloc neutru pentru om – în funcție de cum este utilizat, poate avea efecte pozitive sau negative, morale sau imorale. Totuși, alții susțin că nu este neutră deoarece

folosirea ei ne afectează procesele neuronale, modul în care percepem realitatea și chiar realitatea însăși. Referindu-se la televiziune și radio, McLuhan afirma că „mediul este mesajul”. Cum înțelegeți dumneavoastră aceste două perspective asupra tehnologiei? Considerați că vreuna dintre ele este mai justificată?

„Realitate” este un termen metafizic. La fel și termenul „lume”. Desigur, felul în care conceptualizăm realitatea și lumea ne face să fim mai aproape de un materialism sau de un idealism filosofic, să fim dualiști sau moniști etc. Dacă vei crede că realitatea acestei lumi se produce la întâlnirea dintre materie și cogniție (care poate fi atributul unei anumite configurații materiale a viului), cu alte cuvinte aderi la o ontologie procesuală, orice mediu al existenței este parte a realității, exact așa cum este ea, cum devine în fiecare moment. Nu avem acces direct la vreo esență a realității, la vreun „lucru în sine”, ci ne chinuim să aflăm sau să dăm o reprezentare (nu doar în sensul vizual, al unui mimesis al realității) unor procese, fie ele la nivel subatomic sau la cel al obiectelor gravitaționale, astfel încât nu doar să le putem explica (apelând la limbaje matematice complexe), ci mai ales prezice. Descoperim relații și raporturi, iar asta devine textura realității noastre.

Tehnologiile mediatoare, mai ales cele care se bazează pe limbaje, cum ar fi computerele, world wide web-ul sau Realitatea Virtuală, sunt parte din acest procesualism continuu denumit realitate. Așa cum și televiziunea, radioul sau tiparul sunt, în continuare, în jocul realității. Digitalizarea lumii, algoritmizarea proceselor sale în primul rând ar trebui să ne aducă aminte că niciodată nu am fost lipsiți de mediere, că nu avem acces i-mediat la vreo realitate pură. De aceea prefer să mă întreb despre felul în care se duce existența noastră mediată prin noi mijloace, care le integrează pe cele vechi, într-un mod convergent. Online trebuie să știi să citești și să privești iconic, cât de cât, să interpretezi semnalele și mesajele aplicațiilor și sistemelor ce rulează pe computerele noastre (inclusiv telefoanele inteligente). Când tehnologiile devin contextul experienței, topindu-se parcă în activitățile vieții și lumii, ne e și mai greu să le deslușim rostul și felul cum participă la dorințele, intențiile și acțiunile noastre.

Tehnologiile nu sunt neutre, ele creează posibilități și limite, orientează activități și participă la generarea sau dezintegrarea de instituții, de norme și practici. Nu înseamnă că ele impun doar o anumită utilizare sau că determină inexorabil acțiunile fiecărei persoane. Dar ele au devenit extensiile corpurilor și creierelor noastre în lume și, așa cum observa și McLuhan, sunt în egală măsură extensiile lumii exterioare în noi. Astfel pare că va câștiga monismul, devenim toți parte din aceeași materie gânditoare. Termenul-cheie ar fi cel de „relație” sau „rețea”, ambele care vin atât cu o veste bună – nimeni nu mai e singur, insular –, dar și cu una mai puțin bună: nu mai suntem individualități, ci mai degrabă, să preiau expresia unor filosofi francezi, dividuali.

2. Când vine vorba de inteligența artificială, pare că părerile sunt împărțite. Pe de o parte, există persoane care afirmă că, în câțiva ani, mașinăriile vor înlocui majoritatea activităților pe care le credeam umane, ajungându-se până la substituirea în totalitate a unei persoane. Pe de altă parte, alți cercetători ne asigură sceptic că suntem foarte departe de a putea reconstrui mintea sau conștiința umană, care i-ar putea conferi autonomie și chiar statutul de persoană unei inteligențe artificiale. Până unde a ajuns, de fapt, cercetarea în inteligența artificială în acest moment?

În acest moment, domeniul inteligenței artificiale este foarte variat în vastitatea sa, cu o serie mare de tehnici și metode dezvoltate. Dar nimeni dintre cei care dezvoltă, „antrenează" aceste sisteme nu poate afirma că se află aproape de inteligența artificială generală, care ar putea dezvolta o conștiință, atât a lumii, cât și a propriei persoane. Profeți mincinoși au fost și vor mai fi. Dorința de a crea un algoritm maestru al tuturor algoritmilor e aproape de dorința filosofică de a surprinde întreaga lume într-un gând, de a găsi ultimul răspuns. Varietatea metodelor și sistemelor de IA e răspunsul firesc la varietatea proceselor lumii naturale și sociale. Deocamdată, pentru orice model algoritmic pe care îl dorim avem nevoie de date specifice și de abordări particulare. Chiar dacă acele companii care fac *machine learnin*g și se bazează pe *data science* promit marea cu sarea, rezultatele sunt modeste, în raport cu promisiunile. În raport cu nevoile, unele rezultate sunt revoluționare. Sisteme artificiale care pot genera noi ipoteze științifice, care pot oferi un diagnostic mai bun, care pot acompania copiii în învățarea științei sau care pot guverna online populații există. Dar niciunul nu poate fi utilizat în alt domeniu de la sine, ci trebuie re-antrenat, re-calibrat etc.

„Mașinile care înlocuiesc oamenii" este refrenul modernității industriale și post-industriale. Și poate că unele activități, prin faptul că sunt grele, repetitive, chiar degradante, ar trebui preluate de mașini capabile să facă treaba fără a se simți alienate. Multe activități pot fi automatizate, munca are și un caracter abstract, care vine din ultra-specializare și schematizarea ei rațională, iar transformarea abstractizării în modele numerice, scheme logice etc. e la îndemână. Hegel e printre primii care exprimă clar această idee, plecând totuși de la Smith și Ricardo, și are chiar două perspective diferite: la tinerețe deplânge această preluare a muncii, care ar echivala cu o pierdere a autenticității omului, spre bătrânețe e împăcat cu gândul și omul poate lăsa loc mașinii. Dacă și noi am fi împăcați cu asta, ar trebui să nu ne sperie nici implicațiile: scăderea accesului la proprietate (dincolo de cea a propriului corp) și reducerea autonomiei – oamenii vor fi din ce în ce mai dependenți nu de roboți, ci de puterea politică, căreia nu vor mai găsi mijloacele de a i se opune (în afară de cel mai radical, a ieși cu pieptul gol în fața gloanțelor).

3. Unul dintre cele mai interesante argumente care critică înlocuirea omului de către inteligența artificială mizează pe faptul că omul este o ființă morală. În acest sens, el este în permanență confruntat cu situații de dilemă morală (de tipul dilemei troleului), în care trebuie să facă alegeri aproape imposibile. Astfel de situații sunt însă contextuale și particulare, implicând personalitatea, intuiția, intenția, gândirea, emoțiile, amintirile, experiența, biologia, educația ș.a. decidentului. De aceea este imposibil și indezirabil să reconstruim comportamentul uman cu ajutorul unei inteligențe artificiale, deoarece numai omul poate lua decizii morale complexe, conștiente și voluntare, moralitatea fiind ceea ce ne deosebește de animale. Cât de justificată vi se pare această perspectivă?

N-aș zice că moralitatea e ce ne-ar deosebi de animale, cât culturalizarea ei, atât în sensul recunoașterii și transmisiunii unor practici (prin educație morală, nu neapărat în cadre religioase), cât și într-un sens mai intelectualist, ca abstractizare conceptuală, căutare de întemeieri și principii, capacitate de a gândi în termeni de valori etc. Animalitatea originară a moralității s-ar putea să fie exact limita pe care un sistem de IA n-o poate trece. În schimb, e posibil (nu și probabil) să construim agenți artificiali morali kantiano-hegelieni, apți să se autoguverneze în imperiul scopurilor și/sau să acceadă la stadiul etic al existenței, într-o împlinire inumană a unei raționalități depline. Particularismul și contextualismul moral, avansate de argumentul de mai sus (pe care îl consider invalid pentru că nu știm dacă premisele sunt adevărate, chiar dacă concluzia lui e, evident, adevărată: omul nu poate fi înlocuit de IA, poate doar înlocuit de tot, în urma exterminării speciei noastre de către un ipotetic Terminator) sunt, totuși, o cale mai bună de a arăta limitele IA ca agent moral în viața aceasta, din lumea aceasta. E ceea ce am și încercat, într-o anumită măsură, cu colegii din proiectul CoMoRe: să reluăm gândirea lui Aristotel, cu a sa orientare către viața virtuoasă, nu către reguli sau norme universale (fără a le elimina, desigur), o orientare înspre responsabilizarea fiecăruia după propria înțelegere. Și am redescoperit că excelența morală e imposibilă fără virtuți dianoetice, intelectuale. Nu trebuie confundate cu inteligența sau capacitatea de a găsi soluții, cu viteza gândirii, eficacitate etc. Aceste virtuți, de la atenție la perseverența curiozității, asigură subiectul că poate deveni un *phronimos*, cu o înțelepciune practică în acțiune, căci înțelege contextul vieții sale și ce îi este lui propriu. Discernământul nu se poate modela din date, oricât de multe ar fi, nici din reguli pre-programate, ci din exersarea virtuților în practica de zi cu zi a vieții. Fără discernământ, e mai greu să vorbești despre sisteme IA responsabile, care au capacitatea unui agent moral al lumii noastre.

4. Una dintre cele mai discutate probleme când vine vorba de inteligența artificială este distrugerea pieței de muncă. Pe de o parte, cu cât înlocuim mai multe locuri de muncă ce implică sarcini rudimentare și repetitive, am putea să le oferim acelor oameni șansa de a alege un job mai stimulant, mai ales din punct de vedere intelectual, și să creștem, prin automatizare, productivitatea în diverse sectoare. Pe de altă parte, există persoane care nu mai au posibilitatea de a se reprofila de la un anumit punct încolo sau care pur și simplu nu au capacitățile potrivite unei munci intelectuale. Cum credeți că va evolua, pe termen lung, această problemă?

Istoria noastră, chiar și în Antichitate, este populată de visul automatelor. Aristotel sau împărații Bizanțului, toți, din când în când, aduc în discuție (sau de-a dreptul în spectacolul imperial) tot felul de posibilități ale mașinăriei automate care să preia din sarcinile umane, fie că vorbim de construcții sau spectacole de teatru.

Apoi, în modernitate, vedem acest vis sau coșmar bântuind iar subconștientul și economia. Luddismul nu este o reacție oarecare, ci reacția cea mai importantă față de actul progresului și sublimării muncii, pentru că îl neagă. Din acest punct cred că trebuie începută orice negociere cu megamașina. Invitația de a accepta necritic și tolerant mecanismul violent și total ne-critic (el doar satisface, nu chestionează) al automatizării, invitație făcută mai ales de marile companii digitale, dar și de oamenii politici sau profeții de best-seller, trebuie refuzată. Chiar dacă promisiunile ei sunt nesfârșite. Oportunități vor fi, desigur. Dar pentru cine? În egală măsură, reintegrarea în muncă a multor oameni nu se va mai putea face. Pe acest teren de joc ne aflăm, iar răspunsul nu poate fi tot salvaționist technologic, ci politic. Pentru că doar politic putem opri violența, altfel inegalitățile vor duce la conflicte nesfârșite și la o viață mai pauperă decât cea de azi. În plus, a menține și productivitatea la nivelul oferit de IA fără a avea o cerere ne-ar pune în situația recognoscibilă economiștilor cu experiență din fostele state socialiste, cea a unei producții pe stoc.

Totuși, a fi luddit până la capăt e perdant ca societate, în lumea globală.

5. O bună parte din activitatea dumneavoastră academică se află la intersecția dintre etică și filozofia tehnologiei și informației. Cu toate acestea, din câte știu, în facultățile de informatică și calculatoare contactul cu filozofia, sau măcar cu etica, este minim. Dat fiind că însăși arhitectura unei inteligențe artificiale încorporează anumite valori (de pildă, transparența, securitatea, autonomia, pluralismul etc.), pare foarte important pentru persoanele care lucrează în acest domeniu să fie familiare cu paradigmele etice și să înțeleagă că anumite *design*-uri pot avea efecte cât se poate de concrete asupra utilizatorilor. Credeți că acest aspect al inteligenței artificiale este discutat și

înțeles suficient în România? Cum ar putea fi promovat în rândul persoanelor care lucrează cu inteligența artificială?

În orice departament serios din lume, de „computer science”, există astfel de cursuri, seminare sau ateliere de dezbatere morală. Am colegi care au făcut sau încă țin astfel de întâlniri cu studenții de la Facultatea de Matematică și Informatică a Universității din București. Unul dintre cei mai buni absolvenți ai masterului de etică aplicată, de la Filosofie, este profesor la Automatică și Calculatoare, la Politehnica bucureșteană. Ce vreau să spun este că situația arată mult mai bine acum decât la începutul anilor 2000, când etica aplicată era cel mult o bizarerie. Riscul astăzi este să o transformăm într-o altă birocrație. De aceea cred că abordarea nu trebuie să fie nici cea a unui curs de etică pentru umaniști, nici cea a unui curs de integritate și conformitate, ci sub forma unor situații de antrenament intelectual practic pentru dobândirea discernământului, a responsabilității morale. Partea bună cu informaticienii, să mă exprim vernacular, e că prind foarte repede. Partea mai puțin bună este că ajung să lucreze pentru angajatori care vorbesc despre etică doar când e profitabil. Aici e o tensiune greu de trecut. Cu toate că, ne arată diverse încercări, o etică *în* design, în facerea tehnologiei și a serviciului, adică întâlnirea cu valorile utilizatorilor, pe care le încastrează sau le pune în act, poate aduce profit.

Din toamna lui 2022 vom avea și un prim curs de etică a IA la masterul de Etică aplicată, care să cuprindă ansamblul marilor și micilor probleme. Unul dintre profesori e și programator, informatician, nu doar filosof. La acest curs ar putea veni cele și cei interesați, chiar dacă sunt sau nu designerii lumii noastre digitale.

6. Mai ales în ultimii 2-3 ani, a devenit din ce în ce mai evident că rețelele de socializare sunt departe de a fi niște locuri sigure, neutre și pluraliste. Pe de o parte, în timpul alegerilor prezidențiale din Statele Unite, au existat numeroase eliminări arbitrare a anumitor opinii politice care sprijineau o parte sau alta. Pe de altă parte, există din ce în ce mai multe studii care indică foarte clar că *social media* au la bază algoritmi care produc, la nivel individual, dependență, anxietate, nevroză, dificultăți de concentrare, depresie și, la nivelul societății, polarizare și intoleranță. Cu toate acestea, nu pare că designerii rețelelor de socializare iau măsuri pentru a îndrepta aceste probleme, ci, dimpotrivă, urmăresc și „algoritmizează” și mai mult fiecare mișcare a utilizatorilor. De partea cealaltă, nu pare să existe nicio mișcare serioasă la nivel național sau internațional în rândul utilizatorilor pentru a contracara aceste abuzuri. Ce credeți că este de făcut în această situație de oligopol pe piața *social media*?

Oligopol este un termen blând. Platformele de social media au devenit ele însele piețe închise, controlate, manipulate algoritmic. Prin specializarea lor – inclusiv cea mai

generalistă, adică Facebook – am ajuns, de fapt, să fim ca niște negustori medievali sau trubaduri (depinde ce ne dorim online) care merg din târg în târg sau de la senior la senior să-și facă un rost. Nu avem o piață liberă unde concurează, cu adevărat, platformele, chiar dacă există o singură resursă pentru care se bat, adică atenția și comportamentul nostru, care sunt dataficate, abstractizate numeric și puse la treabă în creuzetul algoritmic.

Nu toate social media sunt capabile de „infoxicare". Forumurile clasice sau anumite platforme mai mici permit o mai mare libertate nealgoritmizată. Dar asta nu înseamnă o lipsă totală de constrângeri sau de limite. Toate platformele sunt scene de împrumut, aparțin cuiva și permit o anumită dramaturgie socială, dar nu una totală, dezlănțuită. De aceea Wikipedia nu e o platformă, ci exact ce spune că ar fi: o enciclopedie. Acolo oamenii și opiniile lor contează foarte puțin, nu e o scenă a dramatismelor comunicaționale, ci un atelier de producție a cunoașterii verificabile.

Ce am putea face? Să nu mai pierdem timpul și atenția. Să nu mai căutăm recunoașterea online. Să nu mai căutăm noul ca pe un drog ieftin. Cum? Prin voința proprie. Putem de unii singuri? Nu, de aceea această retragere strategică se poate face doar dacă inventăm altceva ca bun comun, care nu are nevoie să se hrănească cu date.

7. Un alt subiect foarte important, care a început să fie discutat mai ales în contextul pandemiei COVID, este instrumentalizarea și potențarea fenomenului pe care l-am discutat în întrebarea anterioară – anume dependența oamenilor de internet – la nivel politic. Din acest punct de vedere, viitorul se arată destul de sumbru: într-un regim tehno-autoritar sau tehno-totalitar, nici măcar acel spațiu intim și liber reprezentat de conștiința umană nu ar mai putea fi apărat de către individ. Credeți că un astfel de scenariu este plauzibil? Dacă da, cum ar trebui să reacționăm împotriva materializării lui?

Nici în democrațiile liberale se pare că nu stăm prea bine cu libertatea și autonomia cognitivă. Așa cum au observat și alți cercetători, pare că granițele între lumea autocraților și cea a regimurilor care se întemeiază în drepturile și libertățile individuale se estompează. Marile narațiuni ale progresului, fie el prin capitalism sau socialism, libertate sau egalitate, au copertele ferfeliță și paginile rupte și mâzgălite. Ceva a rămas în loc și tot crește: populismul, care e o formă de colectivism. Internetul a făcut mai ușor de circulat cuvântul populiștilor, de orice orientare și grad. Și cum o ideologie politică simplistă e mai ușor de îmbrățișat, începe să prindă.

Pentru state și companiile online suntem doar populații. Care trebuie guvernate și exploatate. Suntem comportamente care se pot datafica și apoi prezice, chiar îndrepta (spre ce trebuie, desigur). Mai transparenți ca niciodată, mai productivi cognitiv ca

niciodată, mai mulți și mai doritori de mai mult, am intrat într-o fază nouă de civilizație, unde barbarul e acum cel sărac cu datele și cel care nu vrea să îmbrățișeze ultima abstractizare, cea a propriei vieți.

8. După cum aminteam și mai sus, există nenumărate studii ce indică dependența oamenilor, și mai ales a tinerilor, de internet, rețele de socializare și jocuri video. Având mai multe cunoștințe cu privire la modul în care sunt gândite aceste tehnologii, ce sfaturi i-ați da unui tânăr când vine vorba de utilizarea lor? Cum reușiți dumneavoastră să țineți sub control această potențială dependență, fiind productiv în propria activitate?

Nu aș medicaliza, nu aș utiliza un limbaj psihiatric pentru a descrie ce se întâmplă astăzi cu viețile noastre online. Nu pentru că nu ar exista și cazuri de dependență per se, care se pot diagnostica și trata clinic, ci pentru că în majoritatea cazurilor pur și simplu aceasta e viața de zi cu zi, răsfirată în toate aplicațiile și platformele unde ne ducem veacul. Dacă i-am putea aduce pe Marx sau Mill în zilele noastre, ar fi surprinși de cât de ușor aderăm la tirania opiniei publice sau cât de mult ne place să fim exploatați. Dar nu cred că ar spune că suntem anormali.

Nu sunt determinist tehnologic, dar nu am cum să nu remarc, pe urmele paleoantropologilor, că trecerea „de la mână la creier” s-a făcut în concordanță cu noile tehnologii ale acelor vremuri. Cu cât am descoperit sau inventat tehnologii mai bune și am dobândit tehnici noi, cu atât specia a evoluat spre ființele cerebrale de azi, totul culminând cu limbajul însuși, apoi scrierea sa. Astăzi, tot limbajul este cel care se întrupează, doar că nu e limbajul natural, ci cel al datelor, algoritmilor și codurilor, și nici omul nu e mai e cel care-l poartă, ci mașinile.

Sfatul meu nu e un sfat, că nu am căderea, ci o invitație de a reflecta sistematic la tot ceea ce e atât de la îndemână că nu-l mai vedem, inclusiv telefonul sau internetul.

Imagine: Arhiva personală