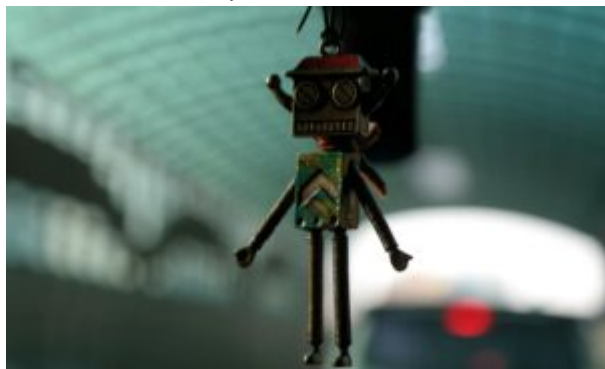


# DE CE ESTE DISCUTABILĂ POSSIBILITATEA INTELIGENȚEI ARTIFICIALE GENERALE?

Autor: Redacția Syntopic | 30 ianuarie 2022



*Un articol de Victoria Maria Deliu și Andrei Poganu*

Înainte de a da un răspuns cu privire la limitele și posibilitățile inteligenței artificiale sau absența lor, să ne gândim la limitele și posibilitățile cunoașterii umane și la ce ne așteptăm mai exact să facă această mașinărie, dacă asumăm că trebuie să reproducă comportamentul uman.

## Problema cunoașterii tacite

În dezbateră cu privire la posibilitatea inteligenței artificiale generale, aș atrage atenția asupra unui gânditor nu foarte cunoscut în spațiul public românesc, Michael Polanyi, și asupra teoriei cunoașterii pe care o propune.

Despre profilul cu totul fascinant al lui Polanyi voi vorbi, poate, mai pe larg, într-un alt text. În treacăt fie spus, Polanyi a fost inițial chimist, însă a ajuns să formuleze probleme atât de epistemologie și filozofia științei, cât și de filozofie politică. Iar ceea ce este de-a dreptul admirabil la el, din punctul meu de vedere, este că le considera – așa cum și sunt, de altfel – strâns interconectate, ca și cum dacă și-ar fi întrerupt firul gândirii oprindu-se asupra unui singur tip de întrebări și probleme, nu ar fi gândit, de fapt, până la capăt. Este unul dintre acei gânditori la care se poate sesiza, cred, un flux viu al gândirii. Iar asta se observă atât din conținutul ideilor sale, cât și din modul în care le

exprimă. Scrierile lui nu plictisesc prin sterilitatea pe care o au uneori lucrările academice, ci captivează prin combinația dintre un limbaj tehnic-conceptual și fărâme de limbaj plastic presărate din loc în loc. Însă ceea ce este relevant în discuția despre inteligența artificială este teoria sa epistemologică.

*We can know more than we can tell*, ne spune Polanyi. Putem cunoaște mai mult decât putem exprima în cuvinte. Iar asta fiindcă modul în care cunoaștem lumea comportă o dimensiune *tacită*. Nu este deloc ușor de spus ce înseamnă cunoașterea tacită pentru Polanyi. Pentru a-și ilustra conceptul, acesta descrie modul în care funcționează o descoperire științifică, considerând că aceasta reprezintă cazul paradigmatic al cunoașterii științifice – și, mai mult, al întregii cunoașteri umane.

Astfel, cunoașterea tacită pare să fie o combinație între elementele descoperirii științifice: un soi de intuiție și „sete imaginativă” de a actualiza niște posibilități care sunt numai întrezărite; preocuparea aproape obsesivă, în orice caz, foarte intensă pentru problema care trebuie rezolvată; experiența fenomenologică, de o subiectivitate ireductibilă pe care o are cel care se implică într-un proces de cunoaștere; familiarizarea profund personală a celui care cunoaște cu problema pe care vrea să o rezolve; folosirea, în rezolvarea problemei, de toată cunoașterea de care dispune deja, teoretică și practică; și o convingere metafizică realistă – acceptarea presupuziției că există o realitate independentă de mintea umană, care poate fi cunoscută – dublată de dorința sau aspirația de a cunoaște această realitate. De fapt, cazul descoperirii științifice pune în lumină două presupuziții fundamentale pentru Polanyi. Prima este însăși convingerea metafizică realistă, anume că există o realitate independentă de mintea noastră care poate fi cunoscută. În descoperirea unei părți din această realitate constă chiar scopul unui act onest de cunoaștere. Cea de-a doua presupuziție este că putem cunoaște această realitate doar din *contextul specific* în care ne aflăm la un anumit moment și în care suntem constrânși să integrăm noua realitate de cunoscut.

Polanyi crede că, din punctul de vedere al modului în care funcționează cunoașterea tacită, ea este *a from-to structure*. Pleci *de la* ceea ce ai deja (personalitatea ta, convingerile tale cele mai profunde despre însăși natura realității, introspecțiile și intuițiile tale, starea afectiv-emoțională, motivațiile, o anumită tradiție epistemică pe care o accepți în detrimentul altora etc. – într-un cuvânt: *contextul specific*) și te îndrepti *către* ceea ce vrei să afli (într-un cuvânt: noua bucată de realitate care este de cunoscut și pe care urmează să o integrezi în

contextul specific). O idee cât se poate de banală, căci, nu-i așa?, fiecare dintre noi se sprijină pe ceea ce a acumulat deja pentru a putea explora ceva nou. Însă ceea ce este *tacit* în acest proces este, cred, modul în care sunt integrate sau configurate elementele care alcătuiesc contextul specific în care cineva se situează la un moment dat și, apoi, modul în care este integrată în contextul specific partea de realitate care este de cunoscut. Integrarea elementelor contextuale, particulare – din care fac parte și elemente strict epistemice, dar nu numai – în virtutea dobândirii unei noi cunoașteri se realizează într-o manieră care nu se poate explicita sau descrie complet. Cu alte cuvinte, într-un proces epistemic oarecare, nimeni nu poate articula cu maximă precizie ce rol a jucat fiecare dintre elementele care compun contextul său particular și nici cum a ajuns noua cunoaștere să fie integrată în contextul particular.

Nu degeaba oferă Polanyi, drept exemplu de cunoaștere tacită, cunoașterea expertului, de pildă, a medicului care pune un diagnostic. Familiarizarea sa cu problema pe care o are de soluționat (identificarea cauzei bolii pacientului din fața sa) s-a făcut, desigur, în baza unor practici și a unei teorii interiorizate și a unor abilități specifice dezvoltate, de-a lungul timpului, în comunitatea științifică din care face parte. Medicul reușește să pună un diagnostic corect (deci, reușește să surprindă corect o fărâmă de realitate) nu detașându-se de particularitățile contextului său, ci tocmai datorită integrării lor coerente, despre care nu poate spune însă prea multe. Un alt exemplu de cunoaștere specializată poate fi cunoașterea din fizică. Pentru a se putea vorbi de cunoaștere științifică în fizică este necesar să se vorbească despre posesia unor instrumente speciale, a abilității de a le folosi și a unei metodologii clar definite, în baza căreia să poată fi interpretate rezultatele testelor și experimentelor efectuate. Pe scurt, dacă se ajunge la niște rezultate corecte sau provizoriu corecte, aceasta nu se poate întâmpla făcându-se abstracție de contextul specific în care s-a situat cel care a ajuns la rezultate, ci tocmai grație contextului specific.

Dacă acceptăm perspectiva lui Polanyi asupra cunoașterii, ne dăm seama că, pe de-o parte, ne sprijinim pe elementele contextuale, particulare pentru a cunoaște o nouă porțiune din realitate (pe care o vom integra ulterior contextului nostru particular), iar pe de altă parte, fragmentul de realitate pe care vrem să îl cunoaștem ordonează și configurează într-un anumit mod acele elemente care alcătuiesc contextul specific. *Este ca și când ne imersăm în ceea ce vrem să cunoaștem, îmbrăcați în contextul nostru existențial-uman specific.*

În măsura în care dinamica oricărui proces epistemic uman – adică acest du-te vino de la contextul nostru specific la realitatea exterioară nouă pe care vrem să o cunoaștem – presupune elemente care țin de conștiința fenomenală, de semantică și de o facultate asemănătoare, să zicem, discernământului sau rațiunii practice, atunci este greu de imaginat cum poate fi dezvoltată o inteligență artificială generală care să le conțină, într-un mod armonios, pe toate. Cel mai dificil de închipuit este, cred, cum și-ar schimba respectivul sistem de inteligență artificială generală configurația elementelor contextuale (asumând că ar putea avea unele, întâi de toate) în funcție de un nou obiect sau o nouă realitate care ar fi de cunoscut. Din câte știu, o problemă a dezvoltării inteligenței artificiale generale care nu a fost încă depășită este tocmai așa-numita *specificity problem*. Simplu spus, este vorba despre adaptarea unui individ la o nouă situație specifică pe care acesta o întâmpină. Descrisă în termenii pe care i-am folosit până acum în acest text, problema se referă exact la modul tacit în care sunt integrate elementele care alcătuiesc contextul specific al unei persoane în virtutea cunoașterii unei noi bucăți de realitate. Așadar, momentan, cele mai avansate forme de inteligență artificială, care nu sunt încă forme de inteligență artificială generală, nu își pot re-integra și re-configura elementele propriului context de fiecare dată când cunosc un nou obiect. Nu mă simt în măsură să estimez dacă vor putea vreodată.

### Specificul pragmatic al demersului epistemic

Unul din lucrurile la care trebuie să ne gândim atunci când vrem să răspundem la problema construirii unei forme de inteligență cu capacitatea de a computa și răspunde la orice problemă umană – și chiar dincolo de limitele umane de răspuns – este specificul valoric al demersului epistemic. Trebuie să constatăm că atunci când vorbim despre cercetare, discipline științifice sau filozofice, paradigme de gândire sau scheme conceptuale, *nu* privim doar un ansamblu teoretic de informații cu valoare strict computațională.

Nicholas Rescher, Hilary Putnam și Philip Clayton au fost unii din gânditorii care au subliniat importanța *valorilor epistemice* – sau după jargonul kantian, a *conceptelor regulative* – în demersul de obținere a cunoașterii. Claritatea, simplitatea, conceptul de *justified true belief* și însuși conceptul de adevăr sunt aspecte *pragmatice* ale cunoașterii. Acest fapt nu le poate scoate în afara domeniului de studiu al valorii și

evaluării normative. Din contră, aspectele normative ale demersului epistemic impun anumite granițe încercării de unificare a domeniilor cunoașterii.

Una din provocările cu care se confruntă Inteligența Artificială Generală este construirea unui limbaj în care multitudinea de date și cunoștințe teoretice să fie inserate, computate și ulterior articulate de acea inteligență artificială. Dar fiecare domeniu de studiu își are propria gamă de evaluări normative, de lucruri importante, interesante, demne de atenție ș.a. Consensul normativ între doi gânditori aparținând unor școli diferite de gândire este rar, dacă nu chiar excepțional. Iar unificarea mai multor domenii de cercetare este limitată de cadrul normativ specific fiecărui domeniu. Nu pentru că „importanța” unui subiect i-ar schimba în vreun fel conținutul descriptiv, ci pentru că demersul de obținere a cunoașterii este în primul rând o *activitate*, nu o colecție de date.

Putnam vorbea în 1981 despre criterii de acceptabilitate rațională – ca alternativă la metodele strict formale și cele bazate pe corespondență – de justificare a cunoașterii. Acceptabilitatea rațională este ea însăși un criteriu pe care cu greu l-am putea transforma într-un proces computațional. Nu voi detalia criteriile de acceptabilitate, este suficient să spun că acceptabilitatea rațională presupune câteva componente normative. Dar să ne imaginăm acum că demersul de cercetare ar presupune o singură valoare, *adevărul*, înțeles tarskian. Principiul echivalenței lui Tarski rămâne formal și neutru din punct de vedere filozofic. A ridica problema în metalimbaj nu ne ajută prea mult în situația de față. Iar a spune că știința caută să descopere adevărul este o judecată pur formală. Este tautologic, nu spune realmente nimic.

Adevărul este o parte componentă a demersului epistemic. Unii autori pragmatişti au tratat adevărul ca ceva distinct calitativ de o proprietate formală a unui termen sau propoziții. Am anunțat deja această interpretare la începutul acestui comentariu. Sensul în care este vizat *adevărul* aici este, bineînțeles, cel de *valoare* sau *principiu regulativ*. Demersul de obținere a cunoașterii este ghidat de anumite valori cognitive, precum adevărul, simplitatea, eficiența instrumentală, claritatea. Gândindu-ne asupra naturii acestor termeni observăm că nu sunt reductibili la o descriere. Putem descrie o situație clar, sau o teorie simplă, sau o propoziție adevărată, dar avem dificultăți în a defini precis la ce se referă claritatea, simplitatea, adevărul. Categoria de gânditori pe care i-am menționat au tratat acești termeni ca valori – adică aspecte *funcționale* ale demersului epistemic.

Să privim acum problema dintr-o altă perspectivă: încadrarea cunoașterii teoretice în planul mai larg al acțiunilor și nevoilor umane. Unii autori au vorbit explicit despre antecedenta *practicii* în acceptarea unui criteriu de justificare a adevărului. Nu este cazul să intru în detalii. Este suficient să sugerez următorul aspect al discuției: unul din factorii justificatori pe baza cărui alegem un anumit criteriu al adevărului – în special a adevărului *factual* – este, bineînțeles, *eficacitatea*. Sau, cu alte cuvinte, alegem un anumit criteriu dacă într-adevăr *funcționează*. Acesta este un criteriu *pragmatic*, propus, printre alții, de Rescher, dar ale cărui detalii analitice formidabile le voi dezvolta cu altă ocazie. În cazul de față, „funcționează” nu înseamnă coerență teoretică, chiar din contră. Dacă ceva *funcționează*, înseamnă că acel îndeplinește un scop practic, *ghidează o acțiune*.

Ultimul standard pentru alegerea unui criteriu al adevărului factual nu este, așadar, unul teoretic sau cognitiv, ci unul practic, care se fondează pe satisfacerea unor cerințe eminate umane, mai precis, pe nevoia coordonării acțiunii. Rescher a numit acest standard unul *afectiv*, în sensul în care stările noastre de durere, frustrare, dezamăgire ș.a. (și opusul acestora) sunt, la o ultimă analiză, factorii decidenți în acceptarea adevărului sau falsității unui „dat” factual. Asta nu înseamnă că raționalitatea omului este subminată sau că deciziile umane sunt luate printr-un impuls afectiv. Ci înseamnă că în procesul de luare a deciziilor și de acceptare a adevărului, omul rămâne om. Cu alte cuvinte, componenta afectivă rămâne prezentă – chiar dacă la un nivel mult mai subtil. Iar acest lucru impune necesitatea rațiunii practice.

Să revenim acum la tema Inteligenței Artificiale Generale. În momentul în care acceptăm că activitatea de cercetare și cunoaștere face parte din categoria mai largă a acțiunii umane; că urmărește un scop, folosindu-se de niște mijloace; că este coordonată de ceea ce numim *valori*; și că însuși demersul de acceptare a adevărului factual presupune rațiunea practică – în momentul în care acceptăm toate aceste lucruri, încercarea de unificare a științei într-o minte digitală devine problematică. A răspunde la problema inteligenței artificiale înseamnă a da un răspuns limitelor și posibilităților acțiunii umane (cu sensul larg de activitate coordonată de o valoare finală și una instrumentală). Sau, ca să reformulez, înseamnă a face un pas în afara naturii umane.

Problema, după cum am sugerat mai sus, se pune în termeni acționali. Condiția necesară care decurge din cele de mai sus este că Inteligența Artificială Generală, pentru a deveni realmente operațională, are nevoie de

categoria specifică de acțiune și de toate variabilele eminamente umane pe care aceasta le presupune. În ce măsură putem – sau vom putea – spune că o inteligență artificială acționează este o întrebare pe care o las la discreția cititorului.

*Victoria Maria Deliu este studentă în cadrul programului de licență al Facultății de Filozofie, Universitatea din București. Este Don Lavoie Fellow la Mercatus Center, George Mason University, bursieră a Fundației Konrad Adenauer, Local Coordinator la Students for Liberty România și bursieră a Centrului pentru Studii de Drept Natural și Analiză Normativă pentru proiecte de traducere*

*Andrei Poganu este student la Facultatea de Filosofie, în cadrul Universității de Vest din Timișoara, bursier al Centrului pentru Studii de Drept Natural și Analiză Normativă, bursier al Fundației Konrad Adenauer și Don Lavoie Fellow la Mercatus Center, George Mason University.*

*Imagine: Pixabay*