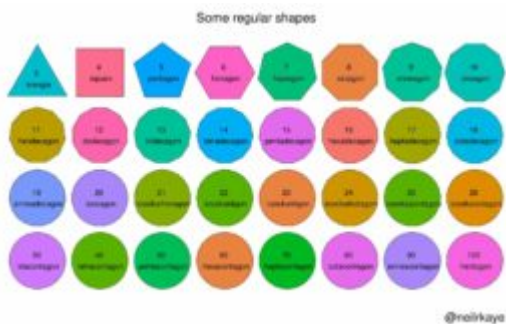


# CHEATGPT

Autor: Silviu Man | 19 ianuarie 2023



„These colossal miscarriages of a dehumanized power-centered culture monotonously soil the pages of history from the rape of Sumer to the blasting of Warsaw and Rotterdam, Tokyo and Hiroshima. Sooner or later, this analysis suggests, we must have the courage to ask ourselves: Is this association of inordinate power and productivity with equally inordinate violence and destruction a purely accidental one?”

– Lewis Mumford<sup>1</sup>

Pe o hârtie, cineva desenează un triunghi și, întorcându-se spre noi, ne spune: „acesta este un cerc”. Dar nu ne poate păcăli, căci noi – ființe raționale – știm bine ce-i un cerc. Un cerc n-are colțuri, un cerc are raze, diametru și circumferință. Oricine știe că un cerc e rotund, e ca o îmbrățișare, pe când triunghiul e cât se poate de colțuros, de... triunghiular.

Mai târziu, cineva desenează un pătrat și ne spune: „acesta este un cerc”. Dar noi știm bine ce este un pătrat. Pătratul are patru unghiuri drepte, e forma cea mai stabilă. Nici nu încapă discuție – pătratul e pătrășos, nu rotund.

Mai târziu, cineva desenează un pentagon și ne spune: „acesta este un cerc”. Dar noi știm și ce este pentagonul (și cu p mic, și cu P mare) și nu ne lăsăm duși de nas. Pentagonul are cinci colțuri clar conturate, deci cu siguranță nu-i cerc.

Și mai târziu, cineva, cu mișcări dibace, ia pentagonul și îi mai adaugă un colț. Dar noi am auzit și de hexagon. Care nu, nu este și nu poate fi cerc.

Cu răbdare, cineva ia hexagonul și îi mai adaugă un colț. Și noi zâmbim –

forma cea nouă e haioasă, în fine, noi nu ne folosim de ea, doar pasionații de geometrie îi știu numele – heptagon. Dar, oricum s-ar chema, e clar că forma asta nu poate fi un cerc. Are colțuri.

O mână discretă ia heptagonul și îl preface în octogon. Noi mai zâmbim un pic – forma cea nouă e interesantă, dar nu e câtuși de puțin un cerc. Copiii, ce-i drept, ar fi derutați sau chiar păcăliți de acest ceva ce nu mai e triunghi, dar nici tocmai cerc nu e. Dar noi, ființe raționale, vedem bine că octogonul nu e cerc, hai să fim serioși...

Două unghiuri în plus și octogonul e transformat în decagon. Noi ne dregem vocea și schițăm un fel de surâs, cam forțat și cam... colțuros – forma cea nouă e rară, cu siguranță ar păcăli copiii și pe cei prea grăbiți, dar e de ajuns să privești mai de aproape, mai atent, ca să-i vezi unghiurile și să afirmi contrariul: deși seamănă oarecum, *acesta nu este un cerc*, pentru că cercul n-are unghiuri; are raze, diametru și așa mai departe.

Parcă mânuite de gesturile calme și precise ale unui magician experimentat, unghiurile se lărgesc, deschizându-se ca o floare, iar laturile devin din ce în ce mai mărunte, mai fine și mai șlefuite. Încă zece mișcări și decagonul s-a prefăcut în *icosagon* – douăzeci de laturi, douăzeci de unghiuri și stupefacția noastră. Privirea ne-a devenit neputincioasă – forma cea nouă cu greu mai poate fi distinsă de un cerc. Minte însă ne spune că, dacă am putea privi în profunzime, cu *zoom in*, am putea vedea totuși colțurile și conchide că acolo nu este, nu a fost și nu va fi niciodată un cerc. Și indiferent de numărul de unghiuri adăugate, indiferent de ce vor striga copiii, grăbiții, naivii și neatenții lumii și indiferent de ambiția încrâncenată a magicianului, acolo nu va fi niciodată un cerc. Doar o iluzie – cu cât mai perfecționată, cu atât mai spectaculoasă. Cu cât mai spectaculoasă, cu atât mai distructivă.

În acest joc, de la un punct dat încolo, nu mai ai la îndemână nicio dovadă empirică. Ești dezarmat și gol, ca un indigen în fața puștii care îi aduce civilizația. Nu mai poți spune  *așa este*, nu-ți rămâne decât axiomaticul  *așa trebuie să fie*. Nu-ți rămâne decât să revii la principiu, la esență –  *să crezi, știind* că oricât de perfectă ar fi simularea, că adevărul rămâne același – nu poți construi un cerc din linii frânte. Că un aproape-cerc e la fel de departe de cerc precum un triunghi ori un pătrat. Un aproape-cerc nu e un cerc. O aproape-minte nu e minte. O aproape-cunoaștere nu e cunoaștere.

\*

Pe la jumătatea anilor '60, profesorul Joseph Weizenbaum de la MIT a inventat ELIZA, un sistem recunoscut azi ca fiind primul chatbot. Puteai „povesti” cu ELIZA precum cu un psihoterapeut și programul, jucându-și rolul de „doctor”, părea să-ți răspundă, întorcându-ți cuvintele. Iată o mostră de „dialog”:

[DOCTOR] – *Te supără ceva?*

[UTILIZATOR] – *Toți bărbații sunt la fel.*

[DOCTOR] *Care e legătura, ce părere ai?*

[UTILIZATOR] *Mereu se iau de noi despre una-alta.*

[DOCTOR] *Te poți gândi la un exemplu?*

[UTILIZATOR] *Păi, prietenul meu m-a pus să vin aici.*

[DOCTOR] *E important pentru tine că prietenul tău te-a pus să vii aici?*

[UTILIZATOR] *El spune că sunt mai mereu deprimată.*

[DOCTOR] *Îmi pare rău să aud că ești deprimată.*

[UTILIZATOR] *Așa și este. Sunt nefericită.*

[DOCTOR] *Poți să îmi spui ce te face nefericită? ș.a.m.d.*

În ciuda caracterului său rudimentar, programul lui Weizenbaum a devenit cu repeziciune faimos și, spre uimirea inventatorului, a fost preluat cu entuziasm de mai multe instituții din SUA și susținut de numeroși oameni de știință, inclusiv de (încă necunoscutul) Carl Sagan. Un simplu joc al sintaxei, un exercițiu parodic, plin de imperfecțiuni și limitat la câteva proceduri dialogice, dădea speranța că un *psiholog automat* nu este doar posibil, ci și necesar.

Însă șocul cel mai mare al lui Weizenbaum a venit din partea oamenilor obișnuiți, a utilizatorilor simpli. În cartea sa, *Computer Power and Human Reason*, el relatează cum însăși secretara sa, familiarizată desigur cu toate detaliile și etapele de dezvoltare ale proiectului, îl roagă să îi permită să interacționeze cu ELIZA. După câteva schimburi de replici cu „psihologul automat”, secretara îi cere să părăsească încăperea, ca să nu asiste la o conversație atât de personală și de importantă. Altă dată, când lui Weizenbaum îi vine ideea să acceseze conversațiile recente pe care utilizatorii le-au avut cu programul pentru a le analiza, este întâmpinat

cu o opoziție vehementă – oamenii nu vor să se afle despre detaliile intime pe care le împărtășiseră ELIZEI. Pe scurt, oamenii cred cu sinceritate că ELIZA îi înțelege!

„Ce nu am realizat atunci este că e nevoie doar de câteva interacțiuni extrem de scurte cu un program destul de rudimentar pentru a induce unor oameni relativ normali o gândire puternic iluzorie”<sup>2</sup>, scrie Weizenbaum. Astfel, povestea ELIZEI nu este atât de remarcabilă prin inovația tehnică sau prin valoarea ei în istoria computerelor. Este revelatoare prin ceea ce ne dezvăluie nu despre tehnologie, ci despre noi în relație cu tehnologia: că, de fapt, de atât de puțin este nevoie pentru a ne lăsa cucerțiți, influențați sau păcăliți. ELIZA ne dă mărturie despre imensa, înnăscuta disponibilitate umană de a ne proiecta în celălalt („celălalt” însemnând aici nu un om, nu un animal și nici măcar un obiect fizic), de a ne confesa și de a interacționa. Și despre vulnerabilitatea imbecilă la care duc toate acestea în lipsa discernământului.

ELIZA poate fi „a kind reminder” că specia care a inventat roata, scripetele, bomba atomică și guma de mestecat, care poate înălța sateliți pe cer și plăsmui orașe în deșert, care a cucerit adâncul mărilor și zboară pe deasupra norilor, poate fi îngenuncheată de o simulare.

\*

După acest prim desen stângaci numit ELIZA, colțurile programelor de „inteligență artificială” s-au tot șlefuit. Cele mai „rotunjite” forme de până acum care și-au câștigat celebritatea au fost IBM Deep Blue, care l-a bătut pe Kasparov la șah în 1997, Eugene Goostman care ar fi trecut testul Turing în 2014 și Google AlphaGo care l-a învins de trei ori la rând pe campionul mondial Ke Jie la jocul de Go în 2017. Dar niciunul, totuși, nu seamănă cu un cerc.

Până la chatGPT.

Nu cantitatea enormă de informații cu care a fost „antrenat” chatGPT îmi pare uluitoare – știm că suntem în epoca lui „big data” și că mai orice demers tehnologic de azi are pretenții faraonice (*dream big, the sky is the limit* etc). Nici capacitatea programului de a face conexiuni și inferențe nu îmi pare cea mai importantă caracteristică. Și nici aparenta precizie retorică a expunerii. Cu privire la aceste din urmă aspecte, cred că acele critici care i se aduc pentru diversele erori și „răspunsuri” stupide sunt legitime, dar ne deviază de la esențial. Cred că mai importantă decât acuratețea obiectivă a informațiilor este însăși iluzia că acolo are loc o

conversație. Astfel, puterea programului stă nu atât în modul său de prelucra și furniza date cu o viteză amețitoare, cât în capacitatea lui de a cataliza iraționalul din noi, pornirea noastră firească de a comunica și de a ne dezvălui. Sigur că chatGPT dă și va da în continuare „răspunsuri” greșite, dar după cum am învățat din cazul ELIZA, noi, oamenii, avem nevoie de foarte, foarte puțin pentru a ne lăsa convinși.

Mai formidabil decât cele de mai sus îmi pare alt aspect – și anume totala sa disponibilitate discursivă, faptul că oricând, oricine de pe glob îl poate „întreba” *orice despre orice*, știind că va primi în schimb un răspuns generos, într-o formă aparent coerentă – și care, cel mai adesea, depășește așteptările utilizatorului – îl proiectează într-un rol unic, într-un soi de Mare Sfetnic planetar. Sfetnicul acesta este *misterios* – nu știm prea bine cine sunt cei care l-au proiectat; cum, de ce și pentru ce l-au proiectat; ce vor de la el și ce vor de la lumea în care îl aruncă cu atâta ușurință. Mai mult decât atât, însă, Sfetnicul este *supranatural* – pentru că aparent, spre deosebire de cuțit, pușcă, pompa de bicicletă sau mașina de făcut tăiței, de la un anumit punct încolo nimeni, nici măcar inventatorii (sau să le spunem creatorii?) săi, nu mai înțelege cum funcționează această existență spectrală.

Înzestrat cu aceste atribute cvasidivine, Marele Sfetnic are toate șansele de a-și dobândi rapid capitalul de încredere mai mult decât, să spunem, motorul de căutare Google. Odată dobândit acest capital, nu ar fi deloc greu ca, după o vreme, în versiunile ulterioare, Sfetnicul să preia și rolul de Confesor (dacă vom crede că știe totul despre cele de din afară, îi vom încredința – în mod activ sau pasiv – și pe cele de dinăuntru). Și pe acela de Oracol (dacă știe „totul” despre trecut, e clar că poate prezice viitorul, nu?). Dacă această transfuzie de încredere de la om la program va continua nestingherită, probabil că o bună parte dintre oameni nu ar avea nimic împotriva să-l declare nimic mai puțin decât Rege – nu e nimic exagerat aici, avem deja câteva precedente<sup>3</sup>. Cel ce poate prezice viitorul e și cel mai îndreptățit să-l facă posibil, nu?

Deocamdată, în forma curentă, chatGPT este un monstru rece, care intrigă, dar nu seduce. Ne e greu să ne regăsim în el pentru nu seamănă atât de mult cu noi. E prea distant, prea ne-omenesc. Ne lasă impresia unui mega-creier, nu a unui suflet, așa că probabil că relația oamenilor cu el va fi cel mult una de respect. Nu mimează empatia, ci cunoașterea. Îi poate scrie o a 13-a regulă lui Jordan Peterson și îi poate înșira unui bun prieten o listă vastă de cărți despre Imperiul Bizantin scrise de autori români între anii 1920 și 1930. Dar dacă îi spui că ești deprimat, îți răspunde la fel de sec

precum un afiș lipit pe ușa Ministerului Sănătății.

Probabil că pe acest chatGPT oamenii îl vor folosi și respecta, îl vor admira și îl vor crede. Dar acea versiune îmbunătățită de inteligență artificială care se va pogorî mai aproape de noi, inteligențele superficiale, care va ști să simuleze vocea noastră de muritori și chipul nostru de muritori, care-și va asuma un anumit rol, o anumită *persona*, pe acea stihie tehnologică ce va ști să mimeze asemănarea cu omul cu suficient de mult rafinament – și nu stângaci, cum a făcut-o ELIZA la vremea ei – ei bine, pe aceea oamenii o vor iubi din tot sufletul, din tot cugetul și cu toată puterea.

... Și vor fi siguri că forma desenată dinaintea ochilor lor este, incontestabil, un cerc.

Și vai de cei ce vor vedea colțurile.

#### NOTE

1. Lewis Mumford, *Technics and Human Development*, The Myth of the Machine, vol. I ↑
2. Joseph Weizenbaum, *Computer Power and Human Reason. From Judgement to Calculation*, W. H. Freeman and Company, 1976, p. 7 ↑
3. <https://www.vox.com/future-perfect/2019/3/27/18283992/ai-replace-politicians-europe-survey>; <https://futurism.com/political-party-led-by-artificial-intelligence> ↑