

HANNAH FRY

HELLO, WORLD!

Revoluția informatică și viitorul omenirii

Traducere din limba engleză
ADINA IHORA

Cuvânt-înainte
CIPRIAN I. TOMOIOAGĂ

CORINT
BOOKS
—2019—

INTRODUCERE

ORICINE A VIZITAT VREODATĂ parcul național Jones Beach din Long Island, statul american New York, probabil că a trecut cu automobilul, în drum spre ocean, pe sub câteva poduri. Acestea au fost construite, inițial, în folosul celor care intrau sau ieșeau de pe autostradă și au o caracteristică neobișnuită. Arcuindu-se ușor pe deasupra autostrăzii, ele se află la o înălțime extrem de mică de aceasta, lăsând uneori o distanță de numai 2,74 metri până la asfalt.

Designul ciudat are o explicație, însă. În anii 1920, Robert Moses, un urbanist bogat și puternic din New York, a dorit ca recent terminatul și foarte premiatul parc-rezervație de la Jones Beach să le fie accesibil numai și numai americanilor albi înstăriți. Și, cum știa prea bine că toți clienții săi preferați aveau să se deplaseze până la plajă în autoturismele lor personale, în vreme ce locuitorii de culoare din cartierele sărace n-aveau să poată ajunge acolo decât cu autobuzul, el a încercat, în mod intenționat, să le limiteze celor din urmă posibilitatea de a merge acolo, construind sute de poduri joase de-a lungul autostrăzii. Atât de joase, încât autobuzelor înalte de 3,5 metri să le fie imposibil să treacă pe sub ele.¹

Podurile acestea de-a dreptul rasiste nu sunt singurele obiecte care, deși neînsuflețite, au reușit să exercite un control tăcut, clandestin, asupra ființelor umane. Istoria este pângărită de nenumărate obiecte și invenții care posedau o

putere superioară scopului lor declarat.² Uneori, aceasta fusese integrată, în mod deliberat și răuvoitor, în însuși conceptul lor, de la bun început, dar alteori nu a fost decât rezultatul unor omisiuni nechibzuite: gândiți-vă un pic numai la inexistența, în unele zone urbane, a rampelor de acces pentru persoanele țintuite în scaune cu roțile! Iar câteodată, a fost, pur și simplu, o consecință neintenționată, așa cum a fost cazul războaielor de țesut din secolul al XIX-lea, care fuseseră fabricate pentru a ușura obținerea unor materiale textile complexe, însă, în cele din urmă, au ajuns să aibă un impact uriaș asupra salariilor, șomajului și condițiilor de lucru, încât au devenit, în mod indiscutabil, mai tiranice decât orice capitalist din epoca victoriană.

Invențiile moderne nu sunt, nici ele, diferite. Întrebați-i numai pe locuitorii din orașelul Scunthorpe* din nordul Angliei ce probleme au avut când li s-a interzis să-și deschidă conturi pe AOL după ce uriașul furnizor de internet a pus la punct un nou filtru, care excludea vulgaritățile și care a avut obiecții față de însăși denumirea localității lor.³ Sau pe Chukwuemeka Afigbo, nigerianul care a descoperit că un dozator automat pentru săpun lichid elibera soluția, în mod impecabil, atunci când prietenul său alb își punea mâinile sub aparat, însă refuza, cu obstinație, să facă același lucru în cazul său, nerecunoscând nuanța mai închisă a pielii sale.⁴ Sau pe Mark Zuckerberg, care, atunci când concepea codul pentru Facebook în camera lui de student de la Harvard, în 2004, nu și-ar fi imaginat niciodată că propria creație avea să fie acuzată că a înlesnit manipularea voturilor în alegeri desfășurate pe întregul mapamond.⁵

* În numele orașelului se distinge cuvântul *cunt* (engl.), „(trivial, peiorativ și ofensator) vagin”. În limbajul curent a fost introdusă expresia *Scunthorpe problem*, care desemnează o dificultate ce apare atunci când filtrele de internet blochează nume legitime pentru că acestea conțin anumite combinații de litere, care, scoase din context, pot fi considerate obscene (n. tr.).

În spatele fiecăreia dintre aceste invenții stau algoritmi. Fragmente invizibile de cod care alcătuiesc mecanismele și roțile mașinilor epocii moderne, algoritmi au oferit lumii totul, de la fluxurile de știri ale mass-media și până la motoarele de căutare, de la navigarea prin satelit și până la programele de recomandare a muzicii, făcând tot atât de mult parte din infrastructura noastră modernă pe cât au făcut, cândva, și podurile, clădirile sau fabricile. Există algoritmi în spitale, în sălile de judecată și în autovehiculele pe care le conducem. Tot ei sunt folosiți de forțele de poliție, de către lucrătorii din supermarketuri, dar și de către angajații studiourilor cinematografice. Algoritmi au învățat ce ne place și ce ne displace; ei ne spun ce să urmărim, ce să citim și cu cine să avem o întâlnire romantică. Iar, în tot acest timp, au ajuns să dețină puterea ascunsă de a schimba, încet-încet și în mod subtil, regulile în privința a ceea ce înseamnă să fii om.

În această carte vom descoperi marea varietate de algoritmi pe care ne bazăm din ce în ce mai mult, poate chiar fără să ne dăm seama. Ne vom îndrepta atenția asupra a ceea ce pretind ei că fac, le vom analiza puterea nedeclarată și vom încerca să ne confruntăm cu întrebările rămase fără răspuns pe care le ridică. Vom face cunoștință cu algoritmi folosiți de către polițiști pentru a decide cine să fie arestat, pe aceia care ne fac să alegem dacă să le protejăm pe victimele infracționalității sau să credem în inocența celor care sunt puși sub acuzare. Vom afla care sunt algoritmi pe care îi folosesc judecătorii pentru a stabili sentințele date infractorilor care au fost găsiți vinovați, adică acelea care ne somează, practic, să decidem cum ar trebui să fie sistemul nostru judiciar. Vom descoperi că există algoritmi pe care medicii îi folosesc pentru a-și reexamina sau chiar pentru a-și infirma propriile diagnostice, algoritmi în interiorul autoturismelor fără șofer care tot insistă să ne reevaluăm propria moralitate, dar și unii care cântăresc modul în care ne exprimăm emoțiile, precum și alții care dețin puterea de a ne submina democrațiile.

Nu susțin, nici pe departe, că algoritmiile în sine ar fi nocivi. Așa cum veți vedea în paginile următoare, există nenumărate motive pentru care merită să rămânem pozitivi și să privim cu optimism spre ceea ce ne așteaptă în viitor. Tot ce contează este modul în care sunt utilizați algoritmiile. GPS-ul a fost inventat pentru lansarea rachetelor nucleare, iar acum ajută la livrarea pizzei, câtă vreme s-a descoperit că repetarea la nesfârșit a unei melodii pop poate fi folosită ca metodă de tortură. Totodată, să știți că, oricât de frumos va fi fost împletită o ghirlandă de flori, dacă-și dori neapărat acest lucru, chiar aș putea folosi una pentru a strangula pe cineva! Să îți formezi o opinie în privința unui algoritm înseamnă să înțelegi relația care există între om și mașină. Fiecare algoritm este, în mod obligatoriu, legat de persoanele care l-au conceput și îl utilizează.

Ceea ce înseamnă că, în esență, aceasta este o carte despre oameni. Despre felul în care suntem, despre direcția în care ne îndreptăm, despre ceea ce este important pentru noi și despre felul în care toate acestea sunt transformate prin intermediul tehnologiei. Este o carte despre relația pe care o avem cu algoritmiile care există deja, cu aceia care lucrează împreună cu noi, ne potențează abilitățile, ne corectează greșelile, ne rezolvă problemele, deși ne mai creează și altele noi pe parcurs.

Aici ne vom pune întrebări precum: aduc sau nu algoritmiile beneficii certe societății? Când ar trebui să ai mai multă încredere într-o mașină decât în propria putere de judecată și când ar trebui să rezizi tentației de a lăsa computerele să comande? Vom descompune algoritmiile și le vom descoperi limitările, iar, în cele din urmă, ne vom cerceta cu atenție pe noi înșine și le vom descoperi pe ale noastre. În această carte vom încerca să separăm ce e rău de ce e bine și să hotărâm în ce fel de lume dorim să trăim.

Pentru că viitorul nu e ceva ce ni se întâmplă, pur și simplu. Noi suntem aceia care avem puterea de a-l crea.

Capitolul 1

PUTEREA

Garry Kasparov știa exact cum să-și intimideze adversarii. La 34 de ani, era cel mai mare jucător de șah care existase vreodată, având o reputație suficient de redutabilă, încât să îl facă pe orice rival al său să fie cuprins de neliniște. Chiar și așa, apela la un șiretlic enervant pe care adversarii săi ajunseseră să-l urască. În timp ce erau așezați, asudând să câștige ceea ce era, probabil, cea mai grea partidă din viața lor, rusul își lua, nonșalant, ceasul de mână de unde fusese pus, lângă tabla de șah, și și-l fixa din nou la încheietură. Era semnalul pe care îl recunoștea toată lumea, acela care însemna că Garry Kasparov se plictisise să se joace cu adversarul său. Ridicarea ceasului semnala că era timpul ca rivalul său să abandoneze partida. Acesta putea să refuze, însă Kasparov avea să-l învingă oricum în curând, iar acest lucru era inevitabil¹.

Dar atunci când Deep Blue al lui IBM* l-a înfruntat pe Kasparov, în meciul celebru din mai 1997, mașina era imună la asemenea tactici. Rezultatul partidei este bine-cunoscut, însă povestea modului în care Deep Blue și-a asigurat victoria este

* În mai 1997, Gari (sau Garry) Kasparov a jucat la New York un meci de șase partide cu super-computerul IBM, Deep Blue. Evenimentul a fost urmărit de milioane de oameni, în timp real, pe site-ul IBM. Dintre cele șase partide desfășurate pe parcursul a șapte zile, Kasparov nu a reușit să învingă calculatorul decât în prima partidă (n. tr.).

mai puțin știută. Acea triumfare simbolică a computerului asupra ființei umane, care a marcat, în multe privințe, începutul erei algoritmilor, nu s-a datorat doar simplei puteri de calcul superioare. Pentru a-l învinge pe Kasparov, Deep Blue fusese nevoit să-l înțeleagă nu doar ca pe un procesor extrem de eficient pentru calcularea unor mutări de șah senzaționale, ci și ca pe o ființă umană.

Pentru început, inginerii de la IBM au luat geniala decizie de a-l proiecta pe Deep Blue astfel încât să pară mult mai nesigur decât era în realitate. Pe durata celebrului meci format din șase partide, mașina s-a abținut, din când în când, să anunțe ce mutarea dorea să efectueze după ce își terminase deja calculele, uneori chiar și pentru câteva minute bune. Din perspectiva lui Kasparov, întârzierile păreau să lase de înțeles că mașina se căznea, chinuindu-se să efectueze din ce în ce mai multe calcule. Asta părea să confirme ceea ce Kasparov credea că știa: și anume, că reușise să aducă jocul într-o situație în care numărul posibilităților era atât de năucitor de mare, încât Deep Blue nu putea să ia o hotărâre judicioasă². În realitate, acesta stătea în repaus, știind exact ce piesă să mute și lăsând, pur și simplu, timpul să se scurgă. Era un șiretlic plin de viclenie, dar a funcționat. Chiar din prima partidă a meciului, Kasparov a început să se lase din ce în ce mai distras, tot încercând să aprecieze cât de priceput la șah putea fi Deep Blue³.

Deși Kasparov a câștigat prima partidă, de la a doua, Deep Blue a reușit să i se vâre pe sub piele, deslușindu-i modul de a gândi. Kasparov a încercat să-l ademenească într-o capcană, ispitindu-l să captureze câteva piese, în timp ce el însuși se pregătea – cu câteva mișcări în avans – să-și elibereze regina și să lanseze un atac⁴. Toți șahiștii care urmăreau partida, ca și Kasparov însuși, se așteptau ca Deep Blue să muște momeala și să cadă în capcană. Numai că, nu se știe cum, computerul a mirosit el ceva. Spre uimirea lui Kasparov, Deep Blue și-a dat seama ce pune la cale marele maestru și i-a

blocat acestuia regina, distrugându-i, astfel, orice șansă de a obține o victorie umană⁵.

Kasparov a fost, în mod vizibil, oripilat. Faptul că se înșelase în legătură cu ceea ce era computerul capabil să facă l-a terminat, pur și simplu. Câteva zile mai târziu, într-un interviu luat după meci, el a apreciat că Deep Blue jucase „dintr-odată, pentru o clipă, ca un zeu⁶”. Mulți ani mai târziu, reflectând la cum se simțise atunci, Kasparov a scris că „făcuse greșea să creadă că mutările pe care era surprinzător ca un calculator să le facă reprezentau, totodată, și niște mutări grozave, imbatabile, în mod obiectiv”.⁷ În orice caz, geniul algoritmului triumfase. Felul în care reușise să înțeleagă mintea omenească și propensiunea acesteia spre greșală îl ajutase să îl atace și să îl învingă pe geniul cel mult prea uman.

Demoralizat, Kasparov a preferat să abandoneze a doua partidă, în loc să lupte pentru a obține o remiză. Din acel moment, încrederea lui în sine a început să se destrame.

A treia, a patra și a cincea partidă s-au soldat cu remize. Când s-a ajuns la a șasea, Kasparov era deja terminat. Scorul final al meciului: Deep Blue, 3½, Kasparov, 2½.

A fost o înfrângere ciudată. Kasparov era mai mult decât capabil să iasă din toate pozițiile dificile în care îi ajunseseră piesele de pe tabla de șah, însă el subestimase capacitatea algoritmului, iar apoi se lăsase intimidat de acesta. Reflectând la acel meci, el avea să scrie în 2017:

Am fost atât de impresionat de felul în care juca Deep Blue! Am devenit atât de preocupat de ceea ce era capabil să facă, încât nu am mai izbutit să-mi dau seama că toate problemele mele se datorau mai mult faptului că eu jucam prost, decât faptului că el juca prea bine.⁸

Așa cum vom vedea în repetate rânduri în această carte, așteptările, speranțele noastre, sunt importante. Povestea înfrângerii unui mare maestru de șah de către Deep Blue arată că puterea unui algoritm nu se limitează neapărat la ceea ce

conțin liniile sale de cod. Înțelegerea propriilor defecte și slăbiciuni – precum și pe acelea ale computerului – reprezintă cheia pentru deținerea controlului.

Dar dacă un geniu precum Kasparov nu a înțeles acest lucru, ce speranțe mai avem noi, toți ceilalți? În aceste pagini vom vedea cum algoritmi s-au strecurat, practic, în absolut toate aspectele vieții moderne – de la serviciile medicale și lupta împotriva infracționalității, până la transport și politică. Am reușit, cumva, în decursul timpului, și să-i ignorăm, și să ne lăsăm intimidati de ei, dar și să ne minunăm de capacitățile lor, toate în același timp. Rezultatul este că nu prea avem habar cât de multă putere le conferim și nici dacă nu cumva am îngăduit deja ca lucrurile să meargă prea departe.

ÎNAPOI LA ELEMENTELE FUNDAMENTALE

Înainte de a ajunge să discutăm despre toate acestea însă, ar fi bine, poate, să ne oprim puțin și să ne întrebăm ce înseamnă, mai exact, un *algoritm*. Este o noțiune care, deși frecvent utilizată, nu reușește, cel mai adesea, să transmită prea multe. Asta se întâmplă, în parte, deoarece termenul în sine este destul de neclar. Oficial, el este definit după cum urmează⁹:

algoritm (s.m.), o procedură pas cu pas de soluționare a unei probleme sau de îndeplinire a unei sarcini, în principal de către un computer.

Asta e tot! Un algoritm reprezintă, pur și simplu, o serie de instrucțiuni logice care arată, de la început și până la sfârșit, cum să se îndeplinească o anumită sarcină. Conform acestei definiții generale, o rețetă pentru prăjituri poate fi considerată un algoritm. La fel și o succesiune de îndrumări pe care i le-ai oferi unui străin care s-a rătăcit. Manualele de la IKEA, videoclipurile despre cum să repara diverse chestii de pe YouTube, până și cărțile de tip *self-help* – teoretic, orice listă de instrucțiuni

autonome cu ajutorul cărora se poate atinge un obiectiv specific și bine definit ar putea fi considerată un algoritm.

Numai că termenul nu este utilizat chiar în acest sens. De regulă, când spunem *algoritmi* ne referim la ceva special: încă mai reprezintă, în esență, niște liste de instrucțiuni care trebuie urmate pas cu pas, dar acești algoritmi sunt, aproape întotdeauna, niște obiecte matematice. Ei preiau secvențe de operațiuni matematice și (folosind ecuații, calcule aritmetice sau algebrice, ori chiar analiza matematică, logica și probabilitățile) le traduc în cod-program. Algoritmii sunt „alimentați” cu informații din lumea reală, li se atribuie un obiectiv și sunt puși la lucru, să calculeze și să tot calculeze până își ating scopul. Algoritmii sunt aceia care fac din informatică o adevărată știință, iar, în decursul timpului, au facilitat multe dintre cele mai extraordinare creații moderne realizate de diverse mașini.

Există un număr aproape incalculabil de algoritmi diferiți. Fiecare cu propriile obiective, propriile caracteristici, ingenioasele sale ciudățenii și deficiențe, așa că nu s-a ajuns încă la un consens cu privire la cel mai bun mod de a-i clasifica. Dar, în mod general, poate fi util să ne gândim la sarcinile din lumea reală pe care le îndeplinesc algoritmii, grupându-le în patru mari categorii¹⁰:

Stabilirea priorităților: conceperea unei liste ordonate

Google Search anticipează ce site cauți, deoarece ierarhizează, stabilește prioritatea unui site pe care s-ar putea să dorești să-l accesezi în funcție de frecvența cu care l-ai utilizat în trecut, comparând rezultatele și clasificându-le. În mod similar procedează Netflix când îți sugerează filme pe care s-ar putea să-ți placă să le urmărești sau TomTom* când îți alege ruta cea mai rapidă. Toate folosesc un proces matematic pentru a ordona marea varietate de alegeri posibile. Deep Blue a fost, în principal, tot un algoritm de stabilire

* Companie fondată în 1991, cu sediul central în Amsterdam, Olanda, specializată în sisteme de navigare auto și hărți (n. tr.).

de priorităţi, unul care trecea în revistă toate mutările posibile de pe tabla de şah şi calcula care dintre ele ar putea oferi şansa cea mai bună de a învinge.

Clasificarea: alegerea unei categorii

Când m-am apropiat de vârsta de 30 de ani, am fost bombardată pe Facebook de reclame la inele de logodnă, împodobite cu diamante. Iar, de îndată ce m-am căsătorit, reclamele pentru testele de sarcină mă urmăreau peste tot pe internet. Pentru aceste mici săcăieli trebuie să le mulţumesc algoritmilor de clasificare. Aceştia, care sunt atât de îndrăgii de cei care lucrează în agenţiile de publicitate, rulează în spatele scenei şi te clasifică drept o persoană interesată de lucrurile respective pe baza caracteristicilor personale. (Şi, cu toate că s-ar putea să aibă dreptate, tot este supărător când reclamele la teste de fertilitate îţi tot răsar pe laptop în mijlocul unei şedinţe!)

Există algoritmi care pot să clasifice şi să îndepărteze automat conţinutul inadecvat de pe YouTube, alţii care pot eticheta pozele de vacanţă în locul tău şi algoritmi care îţi pot scana scrisul de mână şi identifica orice semn de pe o pagină drept o literă din alfabet.

Asocierea: găsirea legăturilor

Asocierea se referă la găsirea şi evidenţierea relaţiilor dintre lucruri. Algoritmii pentru site-urile de anunţuri matrimoniale, precum OKCupid, se bazează pe asociere, ei căutând legături între membri şi sugerând potriviri în funcţie de rezultatele găsite. Motorul de recomandări al companiei Amazon foloseşte o idee similară, făcând asocieri între ceea ce te interesează şi ceea ce i-a interesat pe foştii clienţi. Aceasta a dus la uimitoarea sugestie de cumpărături care i s-a făcut utilizatorului *Kerobotat* de pe site-ul Reddit, după ce cumpăraseră o bătă de baseball pe Amazon: „Poate vă interesează să cumpăraţi această cagulă?”¹¹

Filtrarea: izolarea a ceea ce este important

Algoritmii trebuie adesea să îndeparteze o parte din informații pentru a se concentra asupra a ceea ce este important, să separe, adică, semnalul de zgomot. Uneori fac asta *ad litteram*: astfel, algoritmii de recunoaștere a vorbirii, cum sunt cei care rulează în interiorul Siri, Alexa și Cortana, trebuie, mai întâi, să-ți filtreze vocea, pentru a o deosebi de zgomotele din fundal înainte de a începe să lucreze la descifrarea spuselor tale. Uneori o fac în sens figurat: Facebook și Twitter filtrează, în funcție de activitatea ta anterioară, postările care au legătură cu ceea ce te interesează, pentru a concepe un flux de știri (*feed*) personalizat, care ți se adresează numai și numai ție.

Majoritatea algoritmilor sunt astfel concepuți, încât să execute o combinație a sarcinilor de mai sus. Spre exemplu, UberPool potrivește, în mod automat, pasagerii potențiali cu alții, care se deplasează în aceeași direcție. Odată ce i se dau punctul de plecare și destinația, aplicația trebuie să selecteze posibilele rute care te-ar duce acasă, să te asocieze cu alți utilizatori care se îndreaptă în aceeași direcție și să te aloce unui grup pe care l-ai ales – toate acestea, în timp ce selectează, cu prioritate, rutele care vor obliga șoferul să execute cele mai puține schimbări de direcție, pentru a face cursa cât mai eficientă cu putință.¹²

Prin urmare, asta pot face algoritmii. Acum, să vedem cum reușesc să ducă la îndeplinire toate acestea. Din nou, deși posibilitățile sunt, practic, infinite, există un mod de a sintetiza lucrurile. Putem considera că abordările utilizate de algoritmi se încadrează, în linii mari, în două paradigme principale pe care le vom întâlni în această carte.

Algoritmi care se bazează pe reguli

Primul tip de algoritmi se bazează pe reguli. Instrucțiunile lor sunt scrise de un operator uman, sunt directe și lipsite de ambiguitate. Ne putem imagina că algoritmii aceștia urmează o schemă logică precum rețeta unei prăjituri. Primul pas:

fă asta. Al doilea pas: dacă faci asta, atunci rezultă cealaltă. Ceea ce nu înseamnă deloc că acești algoritmi ar fi simpli – există suficient spațiu pentru construirea unor programe puternice în cadrul amintitei paradigme.

Algoritmi de învățare automată

Al doilea tip de algoritmi s-a inspirat din modul în care învață ființele vii. Pentru a oferi o analogie, să ne gândim cum am putea să-l învățăm pe un câine să ne dea laba. Nu este nevoie să întocmim o listă exactă de instrucțiuni pe care să i le comunicăm. Dresorul nu trebuie decât să își stabilească bine în minte un obiectiv (ce dorește să facă acel câine), precum și o modalitate oarecare de a-l răsplăti atunci când duce la îndeplinire în mod corect sarcina respectivă. Totul se rezumă, cât se poate de simplu, la consolidarea comportamentului dezirabil și la ignorarea celui nedorit, și, desigur, la exersare, astfel încât câinele să ajungă să-și dea singur seama ce are de făcut. Când vorbim despre algoritmi, echivalențul celor de mai sus este cunoscut sub denumirea de „algoritm de învățare automată”, care se înscrie în categoria mai largă a „intelenței artificiale”, sau IA*. Îi furnizezi mașinii informații, un obiectiv și feed-back atunci când se află pe calea cea bună – și o lași să descopere singură modalitatea cea mai bună de a-și atinge scopul.

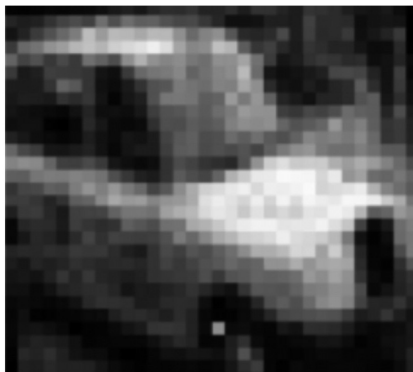
Ambele tipuri de algoritmi au părți pozitive și negative. Din moment ce algoritmi care se bazează pe reguli au instrucțiuni scrise de ființe umane, ei sunt ușor de înțeles. Teoretic, oricine îi poate descompune și urmări raționamentul a ceea ce se petrece înăuntrul lor.¹³ Însă această binecuvântare este, pentru ei, și un blestem. Algoritmii care se bazează pe reguli vor funcționa numai pentru rezolvarea problemelor pentru care operatorii umani știu cum să scrie instrucțiunile.

Algoritmii de învățare automată, în schimb, s-au dovedit, recent, a fi remarcabil de buni la rezolvarea unor probleme pentru care scrierea unei liste de instrucțiuni nu este posibilă.

* *AI, artificial intelligence* (engl.) (n. tr.).

Ei pot recunoaște obiecte în diverse imagini, ne pot înțelege cuvintele pe măsură ce le rostim și pot să traducă dintr-o limbă în alta – ceva ce a fost întotdeauna dificil pentru algoritmi care se bazează pe reguli. Dezavantajul este că, dacă vei lăsa o mașină să descopere singură soluția, procesul pe care îl va urma pentru a-și îndeplini obiectivul nu va avea sens pentru un observator uman. Căci părțile sale tainice pot fi o enigmă chiar și pentru cei mai pricepuți programatori ai zilelor noastre.

Să luăm, de exemplu, sarcina recunoașterii unei imagini. Un grup de cercetători japonezi a demonstrat de curând cât de bizar i se poate părea unui om modul în care un algoritm „vede” lumea. Poate că ați mai întâlnit iluzia optică în care nu puteți spune cu certitudine dacă vă uitați la imaginea unei vase sau la aceea a două chipuri umane (în caz contrar, aveți un exemplu în notele de la sfârșitul cărții¹⁴). Iată care ar fi echivalentul acesteia pentru un computer: echipa de cercetători a arătat că simpla modificare a unui singur pixel la roata din față a imaginii următoare fusese de ajuns pentru a face ca un algoritm de învățare automată să se răzgândească și să considere că fotografia nu mai reprezenta un automobil, ci un câine.¹⁵



Pentru unii, ideea ca un algoritm să lucreze fără instrucțiuni explicite pare să reprezinte cea mai sigură cale spre dezastru. Cum putem controla ceva ce nu înțelegem? Ce se va întâmpla în cazul în care capacitățile mașinilor conștiente de sine și super-inteligente le vor depăși pe acelea ale creatorilor lor? Cum am putea garanta că o inteligență artificială (IA) pe care nu o înțelegem și pe care nu o putem controla nu completează deja împotriva noastră?

Toate acestea sunt întrebări ipotetice interesante și nu ducem lipsă de cărți despre amenințarea iminentă a unei apocalipse cauzate de IA. Îmi cer scuze dacă la așa ceva vă așteptați, însă cartea de față nu este una dintre ele. Cu toate că IA a făcut progrese rapide în ultima vreme, aceasta rămâne „inteligentă” numai în sensul cel mai restrâns al noțiunii. Ar fi, poate, mai util să ne gândim că am asistat la o revoluție în ceea ce privește statistica informatizată, mai degrabă decât la o revoluție în crearea inteligenței artificiale. Știu că sună mult mai puțin atrăgător (dacă nu cumva *chiar* sunteți pasionați de statistică!), însă este o descriere mult mai corectă a modului în care stau lucrurile în prezent.

Pentru moment, să ne facem griji din cauza unei IA malefice pare, oarecum, ca și cum ne-am face griji în legătură cu... suprapopularea planetei Marte!* Poate că într-o bună zi vom ajunge și în punctul în care inteligența artificială o va depăși pe cea umană, însă, pentru moment, nu suntem nici pe departe acolo. La drept vorbind, suntem încă departe de crearea fie și a unei inteligențe de nivelul celei pe care o are un arici. Până

* Am parafrazat aici un comentariu făcut de Andrew Ng, specialistul în calculatoare și pionierul *deep learning*, într-o conferință pe care a susținut-o în 2015. Vezi Tech Events, GPU Technology Conference 2015, ziua a III-a: „What’s Next in Deep Learning”, discurs postat pe YouTube (<https://www.youtube.com/watch?v=qP9TOX8T-kl>) la 20 noiembrie 2015 (link accesat cel mai recent la 4 martie 2019) (n. a.).

acum, nimeni nu a reușit să depășească nivelul inteligenței unui vierme.*

În plus, toată agitația stârnită în legătură cu IA reprezintă o abatere a atenției de la preocupări mult mai urgente și – cred eu – de la subiecte mult mai interesante. Haideți, așadar, să dăm uitării, pentru moment, omnipotentă inteligență artificială și să ne întoarcem, din viitorul foarte îndepărtat, la timpul prezent, pentru că deja există algoritmi cărora li s-a dat frâu liber să acționeze și să ia, în mod autonom, decizii. Unii care stabilesc perioadele de încarcerare ale infractorilor condamnați, tratamentele pe care trebuie să le urmeze bolnavii de cancer și ce trebuie făcut în cazul unui accident rutier. Pentru că algoritmi iau deja, în numele nostru, hotărâri care ne schimbă viețile.

Întrebarea este: dacă le oferim toată această putere, oare, merită ei toată încrederea noastră?

ÎNCREDERE OARBĂ

Duminică, 22 martie 2009, nu a fost o zi bună pentru Robert Jones. Tocmai își vizitase niște prieteni și se întorcea cu mașina prin frumosul oraș Todmorden din West Yorkshire, când a observat că i s-a aprins becul indicatorului de benzină al BMW-ului. Mai avea doar 11 kilometri până să găsească o stație de benzină înainte să i se termine combustibilul, ceea ce era chiar la limită. Din fericire, GPS-ul părea să-i fi găsit o

* Simularea creierului unui vierme constituie chiar obiectivul proiectului științific internațional OpenWorm. Cercetătorii speră să reproducă, pe cale artificială, rețeaua formată din cei 302 neuroni care se găsesc în creierul unui vierme *Caenorhabditis elegans*. Pentru a pune lucrurile în perspectivă, trebuie menționat că noi, oamenii, avem în jur de 100 000 000 000 de neuroni. Pentru mai multe informații, vezi site-ul proiectului OpenWorm: <http://openworm.org/> (link accesat cel mai recent la 4 martie 2019) (n. a.).

scurtătură și l-a trimis pe un drumeag îngust și sinuos care urca pe coasta unui deal.

Robert a urmat indicațiile mașinii, numai că, pe măsură ce conducea, drumul devenea tot mai abrupt și mai îngust, până când, după vreo trei kilometri, s-a transformat într-un adevărat drum de țară pe care abia-abia dac-ar fi încăput niște cai, nicidecum automobile! Dar Robert nu a fost impresionat. Conducea câte 8 000 de kilometri pe săptămână ca să-și câștige pâinea și știa cum să se descurce când era la volan. În plus, s-a gândit el, nu avea „niciun motiv să nu aibă încredere în sistemul său de navigare prin satelit TomTom.”¹⁶

Doar câteva clipe mai târziu, oricine s-ar fi întâmplat să privească în sus din vale ar fi văzut botul BMW-ului lui Robert ivindu-se peste marginea stâncoasă a prăpastiei, fiind salvat de căderea în gol de vreo 30 de metri numai de gardul șubred din lemn de care tocmai se izbise.

A fost nevoie de un tractor și trei ATV-uri ca să-i recupereze mașina din locul în care o abandonase. Ceva mai târziu în același an, când s-a prezentat în instanță sub acuzația de conducere imprudentă, Robert a recunoscut că nu se gândise să ignore indicațiile sistemului său de navigare. „Acesta insista că drumeagul acela era o șosea”, declarase într-un ziar după incident. „Așa că am avut încredere în el. Nu te aștepți să fii ghidat drept într-o prăpastie!”¹⁷

Nu, Robert! Nu cred că se așteaptă cineva la una ca asta!

Istorieara aceasta are, cumva, și o morală. Cu toate că s-a simțit, probabil, puțin ridicol, pentru că nu a ținut cont de ceva ce se afla chiar în fața ochilor săi (cum ar fi, bunăoară, o prăpastie văzută prin parbriz!) și pentru că i-a atribuit unui algoritm o inteligență mai mare decât ar fi meritat, Robert Jones nu era, nici pe departe, singurul om aflat în această situație. În fond, Garry Kasparov însuși căzuse în aceeași capcană cu vreo 12 ani în urmă! Și, în moduri mai puțin

spectaculoase, însă la fel de importante, facem aproape toți această greșeală și, probabil, nici măcar nu ne dăm seama.

În 2015, oamenii de știință au început să studieze în ce fel motoarele de căutare precum Google au puterea de a ne modifica perspectiva asupra lumii.¹⁸ Voiau să afle dacă impunem niște limite sănătoase încrederii pe care o acordăm rezultatelor lor sau dacă nu cumva suntem gata să le urmăm, bucuroși, chiar dacă ne-ar ghida, metaforic vorbind, peste marginea unei prăpastii.

Experimentul s-a concentrat asupra alegerilor care urmau să aibă loc în India. Cercetătorii, sub conducerea psihologului Robert Epstein, au selectat 2 150 de alegători indeciși din întreaga țară și le-au oferit acces la un motor de căutare special conceput în acest scop, numit „Kadoodle”, pentru a-i ajuta să afle mai multe informații despre candidați înainte de a hotărî pe care să-l voteze.

Kadoodle era „măsluit”. Fără ca participanții la studiu să aibă habar, ei fuseseră împărțiți în grupuri, fiecărui grup fiindu-i arătată o versiune ușor diferită a rezultatelor obținute de motorul de căutare, care îl favoriza pe unul sau pe altul dintre candidați. Atunci când membrii unui grup vizitau site-ul, toate link-urile din partea de sus a paginii favorizau un anumit candidat, ceea ce însemna că ei trebuiau să deruleze pagina în jos, link după link, înainte de a găsi fie și o singură trimitere către un site care să îi fie favorabilă oricărui alt candidat. Grupuri diferite au fost încurajate să prefere diverși candidați.

Nu este deloc surprinzător că participanții la experiment și-au petrecut majoritatea timpului citind site-urile semnalate în partea de sus a primei pagini, pentru că, așa cum spune gluma aceea răsuflată, cel mai bun loc unde să ascunzi un cadavru este pe pagina a doua a rezultatelor unei căutări pe Google. Aproape niciunul dintre participanți nu a dat vreo atenție link-urilor care apăreau mult în josul listei. Cu toate acestea, măsura în care ordonarea influențase opiniile

voluntarilor l-a șocat până și pe Epstein. După numai câteva minute de privit rezultatele măsluite ale motorului de căutare, când erau întrebați cu cine aveau să voteze, participanții erau mai dispuși să aleagă candidatul pe care-l favorizase Kadoodle. Cu uluitorul procent de 12%!

În interviul dat revistei *Science* în 2015¹⁹, Epstein a explicat ce se întâmpla:

Ne așteptăm ca motorul de căutare să facă alegeri înțelepte. Participanții la experiment își spun: „Ei bine, da, observ avantajarea unui anumit candidat, iar asta îmi transmite [...] că, de fapt, motorul de căutare își face datoria.

Dar, dacă luăm în considerare cât de multe dintre informațiile noastre le primim de la algoritmi precum motoarele de căutare, este încă și mai amenințător să aflăm în ce măsură voluntarii credeau că își influențaseră propriile opinii: „Când oamenii nu sunt conștienți că sunt manipulați, ei tind să creadă că au adoptat o părere nouă din proprie voință”, scria Epstein în lucrarea sa originală.²⁰

Firește, Kadoodle nu este singurul algoritm acuzat că ar manipula, în mod subtil, opiniile politice ale cetățenilor. Vom insista mai mult asupra acestui fapt în capitolul „Datele personale informatizate”, însă, pentru moment, merită să reținem ce sugerează experimentul acesta că simțim față de algoritmi care au dreptate în majoritatea timpului: sfârșim prin a crede că ei dau întotdeauna dovadă de un raționament superior.²¹ În scurt timp, nu mai suntem nici măcar conștienți de propriul subiectivism față de aceștia.

Pretutindeni în jurul nostru, algoritmi oferă un fel de sursă convenabilă de autoritate. O cale ușoară de a delega responsabilitatea; o scurtătură pe care o întrebuițăm fără să ne mai gândim. Cine, oare, va da clic de fiecare dată pe a doua pagină de rezultate de pe Google și cine, oare, va analiza critic fiecare rezultat? Sau cine va trece pe la fiecare companie

CUPRINS

<i>Cuvânt-înainte</i> de Ciprian I. Tomoiagă	5
<i>Notă despre titlu</i>	11
<i>Introducere</i>	15
Capitolul 1 – <i>Puterea</i>	19
Capitolul 2 – <i>Datele personale informatizate</i>	43
Capitolul 3 – <i>Justiția</i>	75
Capitolul 4 – <i>Medicina</i>	110
Capitolul 5 – <i>Automobilele autonome</i>	150
Capitolul 6 – <i>Infracționalitatea</i>	185
Capitolul 7 – <i>Arta</i>	225
<i>Încheiere</i>	251
<i>Note</i>	259
<i>Mulțumiri</i>	313
<i>Glosar</i>	315
<i>Index</i>	321