

Elena Lazăr

Dreptul inteligenței artificiale

O scurtă introducere

Editura
Stamangiu
2024

3

Răspunderea inteligenței artificiale

Caracterul adecvat și complet al regimurilor de răspundere în fața provocărilor tehnologice este de o importanță crucială pentru societate. În cazul în care sistemul este inadecvat, are lacune sau prezintă deficiențe în abordarea daunelor cauzate de tehnologiile digitale emergente, victimele pot sfârși prin a fi total sau parțial necompensate. Impactul social al unei potențiale inadecvări a regimurilor juridice existente, în ceea ce privește noile riscuri create de tehnologiile digitale emergente, ar putea compromite beneficiile așteptate.

Funcționarea sistemelor de inteligență artificială depinde deseori de date și de alte intrări colectate de senzorii proprii ai sistemului sau adăugate de surse externe. Nu numai că astfel de date pot fi eronate în sine, dar și procesarea unor date care altfel ar fi corecte poate fi imperfectă. Aceasta din urmă se poate datora unor defecte inițiale de proiectare a manipulării datelor sau poate fi consecința unor distorsiuni ale capacității de autoînvățare a sistemului din cauza volumului mare de date colectate, al căror caracter aleatoriu poate determina sistemul de inteligență artificială în cauză să perceapă și să clasifice

greșit datele de intrare ulterioare. Toate aceste aspecte fac foarte dificilă aplicarea normelor de răspundere la sistemele IA^[1].

O altă problemă specifică sistemelor IA este dată de existența unor surse potențiale multiple de prejudicii. Spre exemplu, este posibil ca un robot chirurgical sau o dronă să fi fost proiectată defectuos (fie că este vorba de hardware, de softul preinstalat sau de ambele), dar este posibil și ca sistemul să fi citit greșit datele corecte sau să fi primit date incorecte, sau ca o actualizare a software-ului efectuată de producătorul inițial sau de o terță parte să fi fost defectuoasă, ori ca utilizatorul să nu fi reușit să instaleze o actualizare care ar fi prevenit un anumit incident (intervenție medicală eșuată)^[2].

Prezentul capitol va aborda răspunderea IA pe două planuri, respectiv răspunderea individuală (delictuală, contractuală și penală, fără însă a intra în detalii la cea de natură contractuală și penală) și răspunderea statală (răspunderea statelor pentru folosirea tehnologiei IA).

Înainte de a analiza cadrul legal în materie, înțelegem să facem următoarele precizări privind posibilitatea acordării de personalitate juridică sistemelor IA.

^[1] CoE, A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, 2019.

^[2] Many hands doctrine, A. GIANNINI, (2024), "Artificial Intelligence, Criminal Liability for", în M. SELLERS, S. KIRSTE (eds.), *Encyclopedia of the Philosophy of Law and Social Philosophy*, Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6730-0_1094-1.

De-a lungul anilor, s-au vehiculat propuneri de extindere a unui anumit tip de personalitate juridică la tehnologiile digitale emergente^[1]. În raportul intitulat „Normele de drept civil privind robotica”, Parlamentul European (PE) a solicitat Comisiei să creeze un instrument legislativ pentru a aborda răspunderea pentru prejudiciile cauzate de roboți. De asemenea, acesta a solicitat Comisiei să ia în considerare „un statut juridic specific pentru roboți” și, „eventual, aplicarea personalității electronice”, ca o soluție de răspundere. Totuși, considerăm aceste propuneri ca fiind controversate și ne-practice din mai multe motive.

Personalitatea juridică îmbracă mai multe forme, chiar și în cazul persoanelor fizice, cum ar fi copiii (cărora le incumbă o răspundere limitată), care pot fi tratați diferit de adulți. Acordarea unei personalități juridice roboților sau IA nu ar necesita includerea tuturor drepturilor de care beneficiază persoanele fizice sau chiar societățile comerciale. Cu toate acestea, o astfel de soluție nu ar fi utilă din punct de vedere practic, deoarece răspunderea civilă este o răspundere patrimonială, care necesită ca titularul să dețină bunuri.

Apreciem așadar că prejudiciile cauzate chiar și de tehnologiile complet autonome sunt, în general, reductibile la riscurile imputabile persoanelor fizice sau categoriilor existente de persoane juridice. Orice personalitate suplimentară ar trebui să meargă mână în mână cu fondurile alocate acestor entități IA, astfel încât să poată fi formulate în mod eficient cereri de despăgubire împotriva lor. Acest lucru ar echivala cu plafonarea răspunderii și cu încercările ulterioare de a eluda aceste restricții prin urmărirea cererilor de despăgubire împotriva persoanelor fizice sau juridice cărora le pot fi atribuite persoanele electronice. În plus, pentru a da o dimensiune reală

^[1] Civil rules on robotics, P8_TA(2017)0051, disponibil la https://www.euro-parl.europa.eu/doceo/document/TA-8-2017-0051_EN.html.

răspunderii, sistemele IA ar trebui să fie capabile să dobândească active pe cont propriu, or, acest lucru nu este posibil^[1].

Pentru a demonstra lipsa de pragmatism a acordării personalității juridice sistemelor IA, să ne imaginăm următorul exemplu. Un robot complet autonom realizează o intervenție chirurgicală în cadrul unui spital, iar răspunderea ar reveni acestuia, și nu medicului. În cazul unei disfuncțiuni, victima operației ar primi despăgubiri doar dacă s-ar încheia o asigurare pentru sistemul IA care a performat operația? Și cine ar plăti primele sau cine ar pune la dispoziția robotului active din care ar putea fi plătite daunele? În cazul în care astfel de active nu ar fi suficiente pentru a despăgubi pe deplin victimele unui incident, acestea ar fi puternic încurajate să ceară despăgubiri de la persoana care beneficiază de pe urma exploatării sistemului (posibil clinica/spitalul care se folosește de robot).

Pentru toate aceste motive, considerăm că nu este oportună acordarea personalității juridice, fie ea și limitată, sistemelor IA.

Înainte de a trece la analiza cadrului legal european privind răspunderea delictuală, ținem să facem referire și la cele 4 modele de răspundere anglo-americane: (1) modelele bazate pe intenție/culpabilitate; (2) modelele bazate pe risc/neglijență; (3) răspunderea strictă și (4) sistemele de asigurare obligatorie^[2].

Modelul (1)^[3], bazat pe intenție/culpabilitate, ce constituie modelul de bază al responsabilității care stă la baza dreptului penal, se

^[1] CoE, A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, 2019.

^[2] J. OBERDIEK, (2017), *Imposing risk: a normative framework*, Oxford: Oxford University Press.

^[3] CoE, A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework, 2019.

concentrează în primul rând pe caracterul voluntar al comportamentului agentului. Acesta poate fi interpretat ca necesitând îndeplinirea a două condiții: în primul rând, condiția de „control”, care demonstrează că agentul a fost responsabil din punct de vedere cauzal pentru comportamentul interzis de lege, în măsura în care agentul a avut o alegere liberă și voluntară cu privire la faptul de a acționa sau nu, și, în al doilea rând, „condiția epistemică”, care necesită dovedirea „culpei”, înțeleasă în sens larg ca necesitând ca agentul să fi avut cunoștință și conștientizare reală a faptelor specifice ce au legătură cu consecințele dăunătoare ale comportamentului său, iar acțiunea agentului poate fi înțeleasă ca fiind bazată pe aceste fapte.

Cel puțin deocamdată, deoarece agenții computaționali nu au capacitatea de cunoaștere subiectivă, conștientizare și intenție, aceste modele de responsabilitate nu pot fi aplicate cu ușurință agenților computaționali în sine, întrucât nu pot satisface condiția epistemică necesară. Modelele bazate pe intenție/culpabilitate pot fi, totuși, aplicate dezvoltatorilor sau utilizatorilor umani ai unor astfel de agenți computaționali. Comportamentul persoanelor care dezvoltă sau implementează însă în mod intenționat tehnologii de inteligență artificială în scopuri periculoase sau rău intenționate (de exemplu, pentru a comite fraude sau pentru a-și însuși bunuri) ar satisface în mod clar cerințele pentru stabilirea responsabilității în cadrul unui model bazat pe intenție/culpabilitate^[1].

Modelul (2) de răspundere este bazat pe risc/neglijență pentru prejudicii tangibile ce stau la baza unei obligații generale de a depune

^[1] J. DE BRUYNE, E. V. GOOL, T. GILS, “Tort Law and Damage Caused by AI Systems”, în J. DE BRUYNE, C. VANLEENHOVE (eds.), *Artificial Intelligence and the Law*, Centrum voor Verbintenissen-en Goederenrecht. Intersentia; 2021, pp. 359-404.

o atenție rezonabilă pentru a preveni riscurile previzibile de prejudicii. Acest model este aplicat în mod convențional pentru a determina dacă agenții sunt supuși unor obligații legale de reparație față de cei care au suferit un prejudiciu ca urmare a neîndeplinirii de către un agent a acestei obligații generale de diligență. O „condiție de control” similară celei care se aplică modelelor de răspundere bazate pe intenție/culpabilitate se aplică, de asemenea, modelelor bazate pe risc/neglijență, în măsura în care trebuie să se demonstreze că agentul a cauzat prejudiciul sau vătămarea în cauză. Cu toate acestea, condiția epistemică aplicabilă modelelor bazate pe risc/neglijență este considerabil mai puțin exigentă decât cele aplicabile modelelor bazate pe intenție/culpabilitate.

De exemplu, rănirea fatală provocată în anul 2018 de un vehicul Uber^[1] ce a intrat în coliziune cu o femeie care împingea o bicicletă cu pungi de cumpărături agățate de ghidon. Vehiculul a funcționat în modul de autoconducere (selfdriving) timp de 19 minute înainte de a confunda femeia cu o mașină (care, prin urmare, se aștepta să ia măsuri de evitare), recunoscându-și greșeala și redând controlul șoferului uman al vehiculului doar cu câteva secunde înainte de coliziune, pe care șoferul uman nu a putut să o prevină. Pare puțin probabil ca dezvoltatorii mașinii să fi putut prevedea în mod rezonabil că sistemul de detecție IA al vehiculului va crede în mod eronat că o femeie care împinge o bicicletă cu pungi de cumpărături atârând de ghidon este un alt vehicul.

Prin urmare, dezvoltatorii acestor tehnologii ar trebui să fie trași la răspundere pentru că, din neglijență, nu au luat măsuri care ar fi permis evitarea prejudiciilor și a faptelor ilicite rezultate. Este cumva

^[1] EU Commission, “Communication on Liability for Emerging Digital Technologies”, disponibil la <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018SC0137>.

o răspundere prea largă? Aplicăm același standard pe care l-am aplica unui șofer uman? Sau ar trebui să mergem pe modelul (3) – *strict liability*?

Modelul (3) de răspundere juridică aplicabil încălcărilor drepturilor (inclusiv încălcările drepturilor omului și ale libertăților fundamentale) este cel al răspunderii stricte sau „răspunderea juridică strictă”, așa cum este denumită în limbajul juridic anglo-american. Conform acestui model, răspunderea juridică pentru încălcarea drepturilor se atribuie celor care o cauzează (prin raportare la prejudiciu), indiferent dacă agentul responsabil a avut un comportament care a condus la încălcarea unui standard de conduită legal și indiferent dacă acest comportament a fost intenționat sau afectat de o anumită stare mentală.

Modelul (4), Asigurarea obligatorie. Mai degrabă decât să se concentreze pe alocarea responsabilității potențialilor factori care pot fi înțeleși ca fiind cei ce contribuie la daunele și prejudiciile care ar putea apărea în urma funcționării tehnologiilor digitale avansate, o societate ar putea decide, în schimb, să acorde prioritate necesității de a se asigura că toți cei care sunt prejudiciați de funcționarea acestor tehnologii sunt despăgubiți financiar. Acest lucru poate fi realizat prin instituirea unui anumit tip de sistem de asigurare obligatorie (care ar putea fi stabilit pe o bază „fără vină”), creând un fond de asigurare la care ar putea recurge toți cei prejudiciați de funcționarea acestor tehnologii.