

ERA DIGITALĂ ȘI JUSTIȚIA (I). OBIECTIVE, METODE DE PROIECTARE A ALGORITMILOR ȘI CONSECINȚE ALE FOLOSIRII IA ÎN JUSTIȚIE



Conf. univ. dr., av. Andreea CIUREA
Facultatea de Drept,
Universitatea „Transilvania” din Brașov

Abstract

After the development of several sectors of human activity, it is the turn of the judiciary to experience a digital revolution, which should put an end to many of the problems of traditional justice. At the same time, these technologies, which give machines tremendous power, can be worrisome. How can we ensure that technology remains an aid in the service of man, respecting the powers of the judge, of trusted third parties (lawyers, notaries, bailiffs) and the interests of litigants? What will be the role of the State in the accountability and verification of algorithmic code designers? During several studies, we aim to shed light on the issue, making it possible to deepen some reflections already initiated by law professionals in Romania and abroad.

Keywords: *algorithm, artificial intelligence, justice, consequences.*

1. Introducere: despre algoritmi și folosirea lor în Justiție

Utilizarea Inteligenței Artificiale (IA) în Justiție pare a fi subiectul momentului în mediul judiciar din România¹. După dezvoltarea mai multor sectoare ale

¹ L.M. Stănilă, *Inteligența Artificială. Dreptul penal și sistemul de justiție penală, Amintiri despre viitor*, Editura Universul Juridic, 2020. Apoi, a se vedea, spre exemplu, pe universuljuridic.ro: B.C. Trandafirescu, *Raportul juridic în mediul digital – o abordare din perspectiva teoriei generale a dreptului* (21.12.2021); G. Manu, *Impactul digitalizării asupra acțiunii administrative și dreptului administrative* (11.11.2021); E. Hurubă, *Digitalizarea executării silite* (07.11.2021), G. Mihai, *Provocările arbitrajului în era digitală* (27.10.2021); M. Voicu, *Cyberjustiția. Unele considerații asupra digitalizării procedurilor judiciare* (09.10.2020); M. Dumitrescu, C. Donciu, *Suntem pregătiți să permitem tehnologiei să ia decizii în locul nostru?* (04.10.2021); X. Moldovan, *Către o nouă etică digitală* (22.09.2021); R. Matefi, *Impactul Inteligenței Artificiale asupra drepturilor la egalitate și nediscriminare în lumina Propunerii de Regulament al Parlamentului European și al Consiliului de stabilire a unor norme armonizate privind Inteligența Artificială (Legea privind Inteligența Artificială) și de modificare a anumitor acte legislative ale Uniunii* (16.07.2021);

activității umane, este rândul sistemului judiciar să experimenteze o revoluție digitală, care ar trebui să pună capăt multor probleme ale Justiției tradiționale: aglomerarea instanțelor, durata mare a procedurilor sau accesul inegal la Justiție... Dar putem ști că Justiția algoritmică va fi mai bună decât Justiția umană?

Codurile algoritmice influențează deja funcționarea mediului judiciar, într-un proces ireversibil. În același timp, Justiția (din România și nu numai) se simte vulnerabilă în fața unor progrese tehnologice. Juriștii sunt refractari la folosirea IA, pentru că văd în IA antiteza filozofiei care ghidează, în mod tradițional, activitatea jurisdicțională: Justiția trebuie să ofere fiecărui caz un răspuns individualizat luând în considerare particularitățile speței, pe când IA își propune să raționalizeze și să „masifice” soluțiile.

Încă din anii '60 ai secolului trecut, în Statele Unite ale Americii, oamenii de știință au început să construiască modele matematice bazate pe probabilități sau corelații menite să ajute la anticiparea hotărârilor judecătorești². În 1963, avocatul Reed Lawlor, care a pus bazele procesării computerizate a datelor unui dosar pentru a anticipa asupra admisibilității acestuia, scria: „Nu există nicio modalitate prin care dreptul să poată evita controlul științei. Dacă avocații și judecătorii nu participă la această lucrare, totul va fi făcut de alții”³. Apoi, în 2012, s-a dezvoltat un model matematic prin care verdictele Curții Supreme a Statelor Unite pot fi determinate cu o rată de fiabilitate de până la 75%⁴.

Cuvântul *algoritm* a invadat limbajul obișnuit, dar aceasta popularizare a schimbat oarecum sensul tehnic al cuvântului. În general, dicționarele definesc

S. Uscov, M. Morar, *Scurte considerații privind răspunderea Inteligenței Artificiale în România sub unghiul mort al AI* (26.08.2020); N.R. Dominte, *Inteligența artificială versus inteligența umană* (19.05.2020) etc.

A se vedea, pe juridice.ro: M. Duțu, „Justiția digitală” la momentul autodefinirii (15.11.2021); D. Mocanu, *Inteligența artificială: nici inteligentă, nici artificială*. Atlas of AI de Kate Crawford (26.10.2021); L. Tec, „Mașinării ca mine”: despre Drept și Inteligență Artificială (14.10.2022); C. Drăgușin, *Revoluția inteligenței artificiale și domeniul juridic* (21.11.2019); L. Tec, *Viața noastră în noua eră a tehnologiei 5G și a inteligenței artificiale* (27.08.2019); M.A. Hotca, *Conștiința, conștiința, inteligența artificială și inteligența mixtă* (19.07.2019), etc.

² F. Kort, *Predicting Supreme Court Decisions Mathematically: a Quantitative Analysis of the “Right to Counsel” Cases*, American Political Science Review 1957, n° 51, p. 1 s.; J. A. Segal, *Predicting Supreme Court Cases Probabilistically: the Search and Seizure Cases (1962-1981)*, American Political Science Review 1984, p. 891 s.; S. S. Nagel, *Applying correlation analysis to case prediction*, Texas Law Review 1963, p. 1006 s., citați de B. Barraud, *Un algorithme capable de prédire les décisions des juges: vers une robotisation de la justice?*, Les cahiers de la justice 2017 p. 121-139.

³ R.C. Lawlor, *What computers can do: analysis and prediction of judicial decisions*, American Bar Association Journal, 1963, 49, p. 337, citat de B. Dondero, *Justice prédictive: la fin de l'aléa judiciaire?*, Recueil Dalloz 2017, p. 532.

⁴ A. Aft, J. Blackman, C. M. Carpenter, *FantasySCOTUS: Crowdsourcing a Prediction Market for the Supreme Court*, Northwestern Journal of Technology & Intellectual Property 2012, n° 10, p. 125 s.; J. Blackman, M. J. Bommarito, D. M. Katz, *Predicting the Behavior of the Supreme Court of the United States: A General Approach*, SSRN Electronic Journal, 21.07.2014.

algoritmul ca fiind o secvență finită și fără echivoc de operații sau instrucțiuni permițând rezolvarea unei clase de probleme complexe. Funcționează din „intrări” (datele originale) și duce la „ieșiri” (rezultatele) urmând pași diferiți care necesită decizii logice sau comparații.

În materie juridică, regulile de drept ar trebui să facă obiectul instrucțiunilor din robot, să poate fi exprimate prin relația IF... THEN. De exemplu: *DACĂ* țara cumpărătorului = non UE, *ATUNCI TVA* = nu și expeditor = da. *DACĂ* plata = numerar, *ATENȚIE garanții* = niciuna”.

Codurile algoritmice sunt traduse în limbaje de programare care permit execuția automatizată a operațiilor care le compun, cu ajutorul unui computer. Cele mai cunoscute sunt cele ale motoarelor de căutare (Algoritmul Google, de exemplu) care permit, în funcție de cuvintele cheie introduse de utilizator, să ajungă la cele mai relevante pagini web. Astfel, extragerea dintr-o mare bază de date a unor exemple, care respectă o serie de criterii, necesită un algoritm. În acest caz, activitatea algoritmică vizează, în principal, să asigure că răspunsul furnizat este corect și că va fi redat în timp scurt, chiar și atunci când baza de date este foarte mare.

Calculatorul efectuează mecanic operațiile care alcătuiesc algoritmul și permite algoritmului să *auto-învețe*, adică să învețe de la sine datele pe care le procesează. Astfel, *codurile algoritmice* mai sofisticate simulează atribute asociate cu inteligența umană, cum ar fi raționamentul și învățarea. Acest tip de cod este *inteligența artificială* (IA).

Având în vedere puterea de calcul pe care o mobilizează și cantitatea de date pe care o pot folosi, algoritmiile pot primi funcții multiple: de a facilita căutarea informațiilor, de a oferi statistici, de a automatiza redactarea, clasificarea și traducerea documentelor, de a recomanda produse sau servicii, de a ajusta prețurile în funcție de piață, de a selecta atribuind scoruri (pentru a acorda un împrumut, a recruta un angajat). Ca urmare, IA este folosită cu succes în transporturi, medicina predictivă, finanțe etc.

În prezent, și în domeniul Dreptului, algoritmiile au devenit *furnizori de servicii juridice*. Dar ar putea deveni și *producători de drept*? Este una dintre întrebările generate de folosirea IA în drept și asupra căreia vom reveni pe parcursul studiului nostru, căci implementarea algoritmilor în sectorul juridic ne obligă la o inversare a perspectivei pentru a înțelege, nu modul în care legea poate controla algoritmiile, ci modul în care algoritmiile pun stăpânire pe lege. Este vorba de a stabili măsura crescândă a intervenției algoritmilor în procesele de producție și de aplicare a legii⁵.

Apoi, sunt dezbaterile despre *Justiția predictivă*, despre algoritmiile care ar permite computerului să pronunțe decizii judecătorești așa cum fac magistrații

⁵ S. Chassignard-Pinet, *Les usages des algorithmes en droit: prédire ou dire le droit?*, Dalloz IP/IT 2017, p. 495.

umani. „Justiție predictivă”, „Justiție algoritmică”, „Justiție actuarială”, „Justiție Buton” – sunt exprimări tot mai frecvente, care creează imaginea unui sistem în care oamenii ar fi înlocuiți de roboți-judecători sau roboți-avocați.

Este vorba de algoritmi sofisticăți care dezvăluie diverse corelații și care se bazează pe trecut pentru a modela prezentul și a imagina viitorul. Scopul unui algoritm predictiv este acela de a calcula o probabilitate pe baza informațiilor pe care proiectantul le oferă.

Așadar, *datele* sunt „inima” dezvoltării algoritmului: funcționarea unui algoritm depinde de corelația dintre foarte multe date, alese în prealabil, ceea ce implică o mare putere de decizie. Care sunt „datele bune”? Căci această alegere va avea consecințe juridice asupra persoanei în cauză. De exemplu, dacă, din punct de vedere istoric, în sistemul de Justiție al unor comunități sociale s-au comis discriminări rasiale sau economice, atunci utilizarea în scop predictiv a datelor colectate va conduce în mod necesar la simpla reproducere a greșelilor care au fost făcute în trecut; datele originale fiind părtinitoare, nu ne putem aștepta ca predicția realizată prin algoritm să fie la polul opus.

Pentru a cunoaște efectele acestei *cyber-justiții*, trebuie să ajungem la o înțelegere și o definiție a acestor concepte. Vom descoperi definiții care variază de la o persoană la alta, în funcție de profesie sau de mișcarea filosofică sau tehnologică de care aparține. De exemplu, într-o lucrare de căpătâi în domeniu, cercetătorii francezi Antoine Garapon și Jean Lassègue definesc Justiția digitală ca fiind „capacitatea dată mașinilor de a mobiliza rapid, în limbaj natural, regula de drept relevantă pentru a trata un caz, de a o pune în context propriile caracteristici (locul, personalitatea judecătorilor, firme de avocatură etc.) și de a anticipa probabilitatea oricărei decizii care ar putea fi luate”⁶.

Marile așteptări ale susținătorilor Justiției predictive vizează accesul la o Justiție mai echitabilă, descongestionarea instanțelor, neutralitatea și transparența. Însă, în starea actuală a cunoașterii și tehnologiei, s-a remarcat că aceste așteptări tind să aibă un efect advers⁷. De exemplu, transparența deciziilor este afectată de motive economice, de conflictele de drepturi privind secretul comercial și proprietatea intelectuală, iar metodologiile algoritmilor rămân opace și inaccesibile, lăsând „inima reflecției” ascunsă, așa cum explică un autor în cartea sa⁸.

De asemenea, este ferventă dezbaterea asupra fezabilității modelării raționamentului juridic. Pe de o parte, sunt cei care cred că Justiția umană este deja mecanică, astfel încât Justiția predictivă ar reproduce același tipar. Atât algoritmul,

⁶ A. Garapon, J. Lassegue, *Justice digitale, accepteriez-vous d'être jugés par des algorithmes?*, PUF, 2018.

⁷ A. Pembellot, *Justice prédictive, solution ou simple reproduction du passé?*, pe <https://www.cyberjustice.ca>, 2019.

⁸ F. Pasquale, *Black Box Society. Les algorithmes secrets qui contrôlent l'économie et l'information*, Fyp Edition, citat de A. Pembellot pe <https://www.cyberjustice.ca/en/2019/07/18/justice-predictive-solution-ou-simple-reproduction-du-passe/>.

cât și silogismul juridic ar fi comparabile cu rețetele de gătit: prin amestecarea diferitelor ingrediente identificate cu precizie, orice bucătar va obține același fel de mâncare; orice judecător va ajunge la aceeași soluție⁹.

Pe de altă parte, sunt cei care consideră că umanitatea din centrul Justiției este imposibil de reprodus. De altfel, unii oameni de știință a datelor (*data scientist*) atrag atenția asupra limitelor existente în încercarea de a descrie fenomenele sociale folosind ecuațiile. Printre aceștia, matematiciana Catherine O'Neil susține, în cartea sa devenită celebră - *Weapons of Math Destruction*, că matematicile pot deveni o nouă „armă de distrugere în masă” deoarece, utilizând calcule algoritmice foarte dense, „ar fi dificil, dacă nu imposibil, să facă transparente alegerile deseori arbitrare ale creatorilor lor”. În ciuda neutralității lor aparente, algoritmi ar putea să consolideze inegalitățile dintre oameni într-un mod fără precedent, utilizând încrederea acestora în matematică pentru a-i împiedica să pună întrebări¹⁰.

Cum putem impune, pentru Proiectantul algoritmilor, datoria de a se asigura că tehnologia rămâne un ajutor în serviciul omului, respectând atribuțiile judecătorului, ale terților de încredere (avocați, notari, executori judecătorești) și interesele justițiabililor? Care va fi rolul Statului în responsabilizarea și verificarea acestor proiectanți?

În mod cert, este vremea unor alegeri decisive cu privire la modul de a introduce AI în Justiție, la nivel național și european, astfel încât trebuie să supunem dezbaterii chestiuni susceptibile de a determina aceste alegeri. Este momentul să ne întrebăm ce vrem în legătură cu utilizarea IA, unde ne poate conduce această utilizare, care sunt riscurile și limitele unei digitalizări *ex-abrupto* și ce măsuri de precauție ar trebui respectate.

Acest studiu își propune să arunce o lumină asupra problematicii, făcând posibilă aprofundarea unor reflecții deja inițiate de profesioniștii dreptului din România.

Efectuând cercetări, am descoperit o doctrină europeană - și franceză, mai ales - foarte consistentă, bazată pe cercetări științifice numeroase și interdisciplinare, la care ne vom raporta cu preponderență în prezentul studiu. Această alegere este motivată, printre altele, de faptul că Franța se numără printre primele patru țări din lume la producția de articole despre IA, alături de China, SUA și Marea Britanie; de exemplu, în 38 dintre universitățile din Franța se predau peste 135 de cursuri legate de IA, există peste 260 de echipe de cercetători în domeniul IA¹¹ etc.

De asemenea, în Franța, este concepută o strategie națională de cercetare în materie de IA, pentru a cărei dezvoltare s-a preconizat, din partea statului francez,

⁹ B. Barraud, *La croisée des savoirs - Un algorithme capable de prédire les décisions des juges: vers une robotisation de la justice?*, Les cahiers de la justice 2017 p. 121.

¹⁰ C. O'Neil, *Weapons of Math Destruction*, Crown, 2016 - <https://www.theguardian.com/books/2016/oct/27/cathy-oneil-weapons-of-math-destruction-algorithms-big-data>.

¹¹ Pentru informații detaliate, a se vedea pe <https://www.enseignementsup-recherche.gouv.fr>.

o finanțare de 665 milioane de euro în perioada 2018-2022. În cadrul acesteia a fost elaborat un amplu Raport¹², coordonat de matematicianul Cédric Villani, care identifică patru sectoare prioritare în care Franța trebuie să își concentreze eforturile pentru a dezvolta inteligența artificială: sănătatea, transportul, mediul și apărarea; în plus, sunt abordate aspecte etice, deoarece aceste tehnologii, care oferă mașinilor puteri gigantice de analiză, pot avea aspecte îngrijorătoare. Raportul Villani subliniază că „este nevoie de un organism pentru a emite avize, pentru a da hotărâri în deplină independență, care poate fi sesizat de guvern și de cetățeni, și care ne spune ce este acceptabil sau inacceptabil”. În acest context, remarcăm că Franța a adoptat în anul 2016, Legea pentru o republică digitală¹³ care vizează, printre altele, și luarea de decizii automatizate.

2. Clarificarea limbajului

Legat de inteligența artificială, înainte de orice încercare de analiză, trebuie depus un efort de definire, deoarece fără un vocabular clar nu se poate construi o gândire clară, după cum remarca cercetătorul francez Jean Lassègue¹⁴. Expresia „Inteligența artificială” și-a schimbat sensul, de când a apărut – în anii '50 ai secolului trecut, însă utilizarea omniprezentă a acronimului „IA” este o indicație a aspectului său ermetic, părând o materie rezervată doar inițiaților.

În cazul IA, cuvântul *algoritm* desemnează atât procedeul care permite furnizarea unui răspuns la o interogare, prin intermediul unei formule de calcul, cât și a procesului cunoscut sub numele de „învățare automată” care a făcut posibilă stabilirea precisă a acelei formule.

Într-un amplu studiu¹⁵, J. Lassègue pune în discuție diferite definiții ale inteligenței artificiale, pornind de la cea a dicționarului Larousse: „Toate teoriile și tehnicile implementate în vederea construirii de mașini capabile să efectueze sarcini rezervate anterior oamenilor”.

Autorul observă că această definiție comportă **trei probleme distincte** care dezvăluie trei ezitări:

- **o ezitare între știință și tehnică**: o știință are ca scop înțelegerea unui câmp de cunoștințe deschis la nesfârșit, în timp ce o tehnică vizează atingerea obiectivelor practice definite în avans. În cazul Inteligenței artificiale, care este relația pe care o întreține între știință și tehnologie?

¹² Rapp. Villani, *Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne*, Doc. fr., mars 2018, disponibil pe <https://www.enseignementsup-recherche.gouv.fr/cid128577/rapport-de-cedric-villani-donner-un-sens-a-l-intelligence-artificielle-ia.html>.

¹³ Loi n° 2016-1321 du 7 oct. 2016 pour une République numérique, JO n° 0235 du 8 oct.

¹⁴ Filosof-epistemolog, cercetător la CNRS (Le Centre national de la recherche scientifique), Franța.

¹⁵ J. Lassègue, *L'Intelligence artificielle, technologie de la vision numérique du monde*, Les cahiers de la justice, 2019 p. 205 și urm.

- o ezitare asupra noțiunii de *mașină*. Trebuie să distingem astăzi între *computer*, ca mașină materială și *mașina* care întruchiează o structură logică: o mașină materială supusă pericolelor oricărei componente fizice și *mașina* reprezentând noțiunea universală de *calcul*. Prin urmare, conceptul de mașină se bazează acum pe o separare strictă între ceea ce numim *software-ul* și *hardware-ul*. Astăzi, prin „mașină” se înțelege o structură logică universală capabilă să execute într-o manieră uniformă modele de repetiție exprimabile sub formă de numere și scrise într-un format dat în avans, numite „programe”; dacă descoperim un model repetitiv într-o activitate, o „mașină de calcul” poate descrie acest tipar sub forma unui program care, ulterior, poate fi executat prin intermediul unui *computer*. Prin urmare, inteligența artificială pare să se extindă de la aspectul cel mai abstract al noțiunii de *mașină* la efectuarea etapelor unui program într-un *computer hardware*.

- ezitări cu privire la **sfera sarcinilor rezervate sau nu oamenilor**. Sarcinile îndeplinite de oameni devin tot mai complexe în timp și nu avem o delimitare clară a sarcinilor pe care le vor putea îndeplini oamenii în viitor. Se pune întrebarea dacă mașinile trebuie să îndeplinească unele sau toate sarcinile pe care oamenii, astăzi și în viitor, le pot îndeplini. Este posibil să suprapunem absolut sarcinile efectuate de oameni și mașini, sau ar trebui, dimpotrivă, să le separăm radical? sau să găsim o variantă hibrid?

Apoi, descoperim o definiție mai actuală și mai tehnică a IA, făcută într-o lucrare de referință a genului - *Artificial Intelligence: A Modern Approach*¹⁶, publicată în 1995 de doi informaticieni renumiți: „Tema principală unificatoare este cea a agentului inteligent. Definim inteligența artificială ca studiul agenților care primesc percepții din mediul înconjurător și care realizează acțiuni până la finalizare. Fiecare dintre acești agenți materializează o funcție care raportează secvențe de percepții la acțiuni [...]”

Făcând o comparație, J. Lassègue observă că, de la o definiție la alta, noțiunea de *ființă umană* a dispărut: inteligența, în acest citat, nu mai este apanajul oamenilor și noțiunea de IA pare să reproducă stricta distincție între mașina software și computerul hardware.

Putem vedea inteligența ca un concept *obiectiv* care poate fi aplicat altor comportamente decât cele ale oamenilor? În esență, este vorba despre relația *agentului inteligent* cu lumea, văzută ca o relație matematică de tip funcțional: aceasta implică asocierea „informațiilor” colectate în „input” și procesarea lor, astfel încât să producă alte „informații” în „Iesire”. Apar aici câteva întrebări: de ce relația *agentului inteligent* cu lumea ar trebui să fie proiectată pe modelul matematic al unei funcții? Și cum explicăm faptul că nu acesta este modul în care oamenii își concep relația cu lumea și de ce ar trebui, așadar, să o schimbăm? trebuie ca mașinile să se adapteze

¹⁶ S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, disponibilă pe <http://aima.cs.berkeley.edu/>.

oamenilor sau oamenii să se adapteze la mașini? Pornind de la aceste întrebări, J. Lassègue identifică anumite *zone de tensiune* având ca obiect noțiunea de IA. Poate că cea mai importantă este tensiunea dintre cele două concepții ale scopului IA: inteligența artificială este destinată să *înlocuiască oamenii* (așa-numita IA puternică) sau să *îmite oamenii* (IA slabă)?

Orice răspuns posibil trebuie să aibă componente în plan social, tehnic, cognitiv, etic etc., deoarece delegarea puterii la mașini nu poate fi făcută fără amenințarea retrogradării umane; în prezent, mulți oameni trăiesc cu teama de a-și pierde locul de muncă și de a fi înlocuiți de mașini care se presupune că au eficiență intelectuală sau practică superioară oamenilor. Cu toate acestea, înlocuirea unui om cu o mașină este o *decizie eminamente umană* care implică o structură de comandă într-o ierarhie socială, astfel că realitatea nu ar trebui, indiferent de performanța pe care mașinile o au, să scape persoanelor.

Analizând distincția dintre oameni și mașini, J. Lassègue subliniază că *reflexivitatea și originalitatea gândirii* au o specificitate umană. *Sensul uman* este esențial plurivoc, construit în interacțiune și perceput ca atare de către oameni. Este vorba despre o *maleabilitate* permanentă a sensului rezultată din interacțiuni de moment și care necesită evaluări continue din partea vorbitorilor¹⁷. Interacțiunile umane, în diverse împrejurări concrete, fac posibilă precizarea sensului unui cuvânt, a unui comportament, a unei vestimentații – toate acestea fiind supuse unor procese de evaluare și aprobare constante, în orice comunitate umană.

Pe de altă parte, în cazul procesării informațiilor într-un sistem de intrare-ieșire, *sensul este definit în avans și rămâne fixat în timpul interacțiunii*. Este adevărat că mașina poate memora toate sensurile deja întâlnite (într-un dicționar, de exemplu); însă, dicționarul nu oferă informații despre modul în care sensurile cuvintelor vor evolua în viitor.

Deși s-au creat mii de limbaje de programare, din punctul de vedere al expresivității lor matematice, acestea sunt identice (se spune că sunt funcții recursive). Proliferarea lor se explică nu prin diversitatea expresiei matematice, ci prin necesitatea de a răspunde nevoilor unor comunități specifice: comunitatea statisticienilor va avea nevoie de un limbaj de programare în care să fie ușor să ia medii, dar un astfel de limbaj nu va interesa un designer web care are nevoie de alte funcții.

Prin urmare, este foarte complicat ca oamenii să integreze propria construcție a sensului uman, atât de maleabilă, în procesele digitale, după cum este complicat să înțelegem natura relației om/mașină.

De asemenea, se pot identifica tensiuni cu privire la *relația dintre mediu și modelele de acțiune care implică alegerile făcute în mod automat* și tensiuni cu privire la *natura cunoștințelor angajate de Inteligența artificială*: IA este o descriere a realității

¹⁷ J. Lassègue, *idem*, Les cahiers de la justice, 2019, p.205 și urm.

efective a proceselor din creier? sau vizează atingerea unor obiective, indiferent de procesele implicate și care nu trebuie să semene cu inteligența umană?

Toate aceste întrebări rămân încă deschise... Cert este însă că informatica s-a generalizat treptat la toate practicile și cunoștințele umane, tinzând să proiecteze o viziune alfanumerică asupra lumii, sugerând că acesta este fundalul realității în sine. Totuși, această proiecție nu ar trebui luată ca valoare nominală, deoarece „calculul” nu are sfera globală care i se atribuie: existența proceselor incalculabile este un fapt dovedit în științele exacte (în logică) și în științele naturii (în fizică)¹⁸.

3. Obiectivele folosirii IA în Justiție

În Justiție sunt numeroși actori (judecători, procurori, avocați și consilieri juridici, grefieri, executor judecătoresc, experți, justițiabili, Ministerul Justiției) care solicită – mai mult sau mai puțin intens – introducerea și dezvoltarea IA în mediul judiciar; însă aceștia speră la rezultate foarte variate și nu întotdeauna convergente. Prin urmare, ar fi util să decidem **care așteptări merită să aibă prioritate**, pentru a cunoaște ce instrumente trebuie dezvoltate pentru a servi în mod optim obiectivele.

Astfel, doi autori francezi (un jurist și un matematician) au identificat câteva axe principale ale **așteptărilor** legate de introducerea IA în Justiție¹⁹, corelate cu **sarcinile** care ar putea fi încredințate IA pentru munca juridică:

a) securitatea juridică și previziunea răspunsului judiciar, adică speranța unei mai bune uniformizări a hotărârilor judecătorești, ceea ce ar determina o mai bună securitate juridică.

Hotărârile judecătorești sunt supuse unor variații, care provin din necesitatea ca judecătorul să se adapteze la fiecare situație, dar și din motive mai puțin conștiente sau intenționate, legate de persoana judecătorului (temperamentul, educația sa, credințele sale). Astfel, se naște așa-numitul *risc judiciar* al pronunțării unor hotărâri diferite, în situații asemănătoare sau aproape identice. Pentru diminuarea acestuia, există mai multe garanții: completuri colegiale în faza controlului judiciar, instrumentele procesuale de asigurare a unei jurisprudențe unitare etc.

În această privință, însă, IA ar putea face posibilă rafinarea funcției unificatoare, mulțumindu-i pe cei care consideră că este esențial să ai o Justiție care să garanteze identitatea soluției, într-un design de fiabilitate bazat pe egalitate și certitudine: orice situație A va da naștere la răspunsul B²⁰.

¹⁸ F. Bailly, G. Longo, *Mathematiques et sciences de la nature. La singularite physique du vivant*, Hermann, Paris, 2006, cap. 2, citați de J. Lassègue, *op. cit.*

¹⁹ L. Pécaut-Rivolier, S. Robin, *Justice et intelligence artificielle, préparer demain* – ep. I, publicat pe Dalloz Actualité, la 14.04.2020.

²⁰ L. Pécaut-Rivolier, S. Robin, *idem*, Dalloz Actualité, la 14.04.2020.

De asemenea, des exprimată este dorința de face hotărârile judecătorești mai previzibile, dorința de a putea prezice, în funcție de diferitele circumstanțe, ce răspuns ar trebui dat unei situații particulare: „încercarea de a prezice, cu cea mai puțină incertitudine posibilă, care va fi răspunsul jurisdicției X atunci când se confruntă cu cazul Y”²¹. Previzibilitatea ar permite justițiabilului să facă alegeri mai informate, să măsoare mai bine conținutul angajamentelor și concesiunilor sale într-o negociere, de exemplu.

Mai ales în litigiile patrimoniale, poate fi folosită IA pentru a prezice cuantumul unor despăgubiri, fondat pe o bază de date a hotărârilor pronunțate de instanțe în litigii cu același obiect. Se pune întrebarea: Magistratul poate păstra controlul asupra acestei operațiuni definind el însuși criteriile pe care le consideră relevante pentru a extrage situații similare? Sau „formula de predicție” ar fi opacă pentru utilizator, presupunându-se că a respectat toate hotărârile judecătorești folosite pentru a o determina?

b) suport în procesul de deliberare

Judecătorul nu caută același răspuns pentru toți, ci dimpotrivă un răspuns care să țină cont de particularitățile speței. De altfel, judecătorilor le este interzis să stabilească dispoziții general obligatorii prin hotărârile pe care le pronunță, acestea producând efecte doar pentru părți și succesorii lor (art.5, art.435 NCPC). Dificultatea muncii judecătorului constă în a stabili corect starea de fapt, în a determina cadrul legal aplicabil și a stabili consecințele specifice pentru fiecare speță dedusă judecării.

În acest context, IA poate tinde să ajute la pronunțarea unei soluții adaptate fiecărui caz, considerat unic, **fără a lua decizia în locul utilizatorului**. IA furnizează judecătorului baze de date cuprinzătoare pentru o cercetare diversă și eficientă, pe măsura complexității misiunii pe care o are acesta.

Folosind baze de date jurisprudențiale și instrumente informatice interactive, judecătorul poate să extragă un set de cazuri similare cu cel care îl interesează. Această extracție poate fi lăsată în întregime **la inițiativa utilizatorului**, pentru a determina criteriile care caracterizează cel mai bine cazul său (texte aplicate, circumstanțe, cuvinte cheie etc.); sau operațiunea **ar putea fi încredințată unui algoritm derivat din IA** al cărui obiectiv ar fi să definească o tipologie a cazurilor înregistrate în baza de date, furnizând „automat” judecătorului un set de spețe apropiate de cea care îl preocupă, fără a fi necesar să specifice judecătorul criteriile de asemănare.

Se pune, evident, aceeași întrebare: Magistratul poate păstra controlul asupra acestei operațiuni definind el însuși criteriile relevante pentru a extrage situațiile similare? Dacă algoritmul este astfel proiectat încât să lase utilizatorul liber să

²¹ B. Dondero, *La justice prédictive*, pe blogul profesorului Bruno Dondero: www.brunodondero.com.

evalueze rezultatele pe care le consideră cele mai relevante, se creează condițiile pentru un dialog om-mașină, ceea ce conduce la o creștere a încrederii în tehnologiile algoritmice și în răspunsurile pe care le oferă.

c) îmbunătățirea procesării cantitative a dosarelor

Încrederea în Justiție se bazează pe respectarea procedurilor, care sunt adesea greoaie și cronofage, dar acceptăm deja că o anumită lentoare este inerentă Justiției.

Însă, durata mare procedurilor este greu suportabilă pentru cei a căror viață este suspendată în așteptarea unei hotărâri judecătorești (pentru un divorț sau un conflict de vecinătate). Altfel, durata proceselor rămâne o preocupare majoră, în multe state democratice. În acest în cadru, utilizarea IA accelerează activitatea judiciară; aceasta implică fluidizarea procedurilor în amonte și soluționări accelerate în cazurile considerate „de masă” (neplata creditelor de nevoi personale sau încălcarea codului rutier). Este un scop urmărit mai ales de către persoanele responsabile de conducerea instanțelor.

d) analiza a posteriori a hotărârilor judecătorești

Aceasta este o operațiune importantă pentru a asigura coerența jurisprudenței și pentru îmbunătățirea concepțiilor judiciare, prin eliminarea prejudecăților pe care le conțin uneori. Analizele se pot baza pe experiența operatorului uman sau pot fi făcute „automat”, prin folosirea IA. Mai concret, omul poate păstra inițiativa pentru a face o anumită comparație mai degrabă decât alta, criteriile de analiză rămânând explicite, sau analiza poate fi încredințată unui algoritm având sarcina de a constitui o tipologie de decizii, instanțe, judecători.

Ca și în cazul predicției, există pericolul ca regula pentru clasificarea datelor să rămână opacă, nefiind proiectată de utilizatorul-jurist.

4. Alegerea metodelor de proiectare a algoritmilor folosiți în Justiție

Știm că orice prelucrare automată necesită *introducerea prealabilă de date* disponibile într-o bază computerizată. Această organizare a datelor într-o formă exploatabilă **nu este neutră**, deoarece necesită o codificare a informațiilor, de exemplu prin intermediul cuvintelor cheie, astfel încât acestea să poată fi apoi compilate, analizate.

Este important să determinăm *modul în care pot fi proiectați astfel de algoritmi*, deoarece caracterul aparent automat al funcționării acestor sisteme se bazează, de fapt, pe o serie de decizii perfect umane, pe care este important să le stăpânim.

De exemplu, stabilirea unei liste de cazuri de divorț similare sau evaluarea unei indemnizații de concediere reprezintă sarcini care pot fi delegate, cel puțin parțial, algoritmilor. Prin urmare, IA va avea ca scop să ajungă la un rezultat comparabil cu cel pe care l-ar produce un om.

Există de obicei două modalități de abordare a unei astfel de probleme, pe care ni le ne explică L. Pécaut-Rivolier (jurist) și S. Robin (matematician), în studiul lor²²:

- **metoda mecanicistă:** constă în codificarea (sau modelarea) procesului intelectual al omului care ar fi trebui să îndeplinească sarcina.

- **metoda empirică:** propune să producă rezultate asemănătoare cu cele oferite de un om fără a căuta să-i reconstituie abordarea.

În ultimii ani, algoritmii IA folosesc mai ales metoda empirică: nu se modelează raționamentul (legal) al unui om, ci se imită deciziile sale. Cu alte cuvinte, acești algoritmi se bazează pe procesul de învățare prin care mașina își determină propriile reguli de decizie bazate pe multe exemple, care țin loc de experiență și pe care ea va încerca să le imite cât mai bine.

Această pre-judecată este pragmatică deoarece pune accent pe **rezultat**. Dacă este vorba de a produce rezultate apropiate de ceea ce ar produce un om, mai degrabă decât de a încerca reconstruirea raționamentului uman, se poate încerca determinarea unei formule matematice care să reducă la minimum diferențele dintre rezultatele produse inițial de oameni (și stocate într-o bază de date) și cele furnizate de aparat. Problema învățării se transformă astfel într-o problemă de *optimizare*: dat fiind un criteriu care măsoară fidelitatea predicțiilor față de observații, algoritmul de învățare vizează direct optimizarea acestui criteriu.

Autorii mai sus menționați atrag atenția că IA tinde în mod natural să-și impună criteriile matematice (în acest caz, apropierea maximă de deciziile deja luate) în detrimentul altora, mai puțin matematizabile, cum ar fi nevoia de a justifica o decizie.

Apoi, cu privire la sarcinile care pot fi încredințate algoritmilor, autorii²³ observă o distincție între învățarea „nesupravegheată” și cea „supravegheată”, prima având un obiectiv descriptiv, iar cealaltă un scop predictiv:

- **Învățarea nesupravegheată:** desemnează problemele pentru care, în faza de învățare, algoritmul nu are informații exogene care să îi permită să evalueze calitatea rezultatelor pe care le produce. Algoritmul determină o tipologie de caz, pornind de la o mare bază de date, care reunește, de exemplu, litigiile privind creditele de consum similare cu un caz dat; tipologia expusă de algoritm este bazată exclusiv pe contemplarea datelor, fără nici o grilă specială de analiză. Rezultatul obținut se bazează pe o reprezentare „matematică” a datelor (texte) și pe un criteriu matematic măsurând asemănarea dintre două texte. Deși inițial părea greu de imaginat că putem matematiza comparația a două texte, acest lucru este posibil: reprezentăm un text ca un simplu „sac de cuvinte”, uitând de orice structură sintactică și măsurând asemănarea dintre doi „saci” prin compararea frecvenței de utilizare a diferitelor cuvinte. Apelăm apoi la un algoritm care are ca

²² L. Pécaut-Rivolier, S. Robin, *idem*, Dalloz Actualité, 14.04.2020.

²³ L. Pécaut-Rivolier, S. Robin, *idem*, Dalloz Actualité, 14.04.2020.

scop definirea unei categorii „optime” a textelor, adică determinarea unor grupuri de texte asemănătoare.

- **Învățare supravegheată:** problemele supravegheate sunt cele pentru care este disponibilă o validare a predicțiilor din timpul fazei de învățare. Acesta este cazul evaluării indemnizației de concediere pentru care am avea o bază de date cuprinzând atât descrierea cazurilor, cât și suma alocată la finalizarea procedurii judiciare. Este vorba de o funcție de predicție care va asocia o sumă cu un caz, care funcție ar trebui să imite, cât mai bine, deciziile înregistrate în baza de date. IA poate aduce în joc zeci de mii de coeficienți a căror combinație asigură predicția; astfel, învățarea supravegheată poate face o predicție nu doar ca valoare punctuală (o sumă), ci și sub forma unei distribuții de valori (interval) sau chiar a unei distribuții de probabilități. Însă, în contrapartidă cu performanțele acestor metode, este caracterul opac al formulelor de predicție rezultate, pe care nu le putem înțelege intuitiv, în general.

5. Consecințele folosirii IA în diverse domenii socio-umane

Cum este cazul oricărei tehnici, realizările în domeniul IA **pot avea consecințe pozitive sau negative** în funcție de obiectivele atribuite acestora. Astfel, există domenii de activitate în care IA a avut cert **un impact pozitiv:** îmbunătățirea accesului la cunoștințe, ajustarea reciprocă a mișcărilor în spațiu, precizie în diagnosticare.

IA permite clasificarea dinamică datelor și accesul la aceste date indiferent de loc, timp și limbă. Astfel, asistența la traducere a înregistrat mari progrese, dar rămâne imperfectă, reprezentând un ajutor și *nu o înlocuire a competenței umane*, deoarece construcția site-urilor bilingve în cauză este rezultatul competenței și al muncii umane. Munca umană este esențială pentru a stabili corespondența corectă între fraze dintr-o limbă în alta, iar utilizatorul uman judecă relevanța în contextul de traducere care îi este propriu. În mod similar, s-a dezvoltat și *asistența în domeniul juridic:* constituirea unor baze de date vaste care enumeră nu numai legile, ci și hotărârile judecătorești grupate după diferite criterii, asistență la redactarea contractelor multilingve deja formate parțial, proiecția rezultatelor posibile ale unei dispute în termeni financiari etc.

În Drept, putem observa un dublu efect benefic al folosirii IA: stimulatив și de constrângere. *Efectul stimulatив* se manifestă fie prin îndeplinirea voluntară a unei obligații contractuale, fie prin decizia de a încheia o tranzacție, ca urmare a acceptării estimării pe care o face un algoritm pentru cheltuielile de judecată, de exemplu. Debitorul își îndeplinește în mod deliberat obligațiile datorită efectului descurajant al amenințării de a suporta sancțiunea prezisă. De asemenea, *contractul inteligent* („smart contract”) dă speranța unei eficiențe mai mari și *tranzacția părților*

într-un litigiu este, la rândul ei, stimulată de sentimentul de obiectivitate și imparțialitate atașat metodelor statistice sau probabiliste folosite de IA.

Efectul de constrângere al IA este exprimat printr-o execuție forțată *ex ante* sau *ex post*. Executarea forțată *ex ante* constă în aplicarea regulii de drept, când – de exemplu – viteza vehiculului se ajustează automat la limitele prescrise de panourile indicatoare. Executarea forțată *ex post* are loc prin implementarea automată a sancțiunii (blocarea temporară a unui cont de client, perceperea unei penalități asupra fondurilor depuse ca garanție etc.) după o acțiune sau o abținere contrară legii, ceea ce l-ar obliga pe debitorul recalcitrant să se conformeze.

Apoi, IA oferă un suport imens în domeniul *mișcării în spațiu*. Pentru traficul auto, utilizarea telefoanelor mobile a permis ajustarea mișcărilor în timp real, în special prin GPS. Însă interesul pentru utilizarea GPS-ului depășește o simplă ajustare, deoarece permite și automatizarea mișcării în spațiu (în cazul vehiculelor autonome) sau efectuarea unor cercetări, cu ajutorul dronelor, de exemplu (de la monitorizarea recoltelor până la descoperirea siturilor arheologice).

Apoi, *ajutorul în diagnostic, în special în medicină*: recunoașterea automată a modelului face posibilă revizuirea unor mase mari de imagini fără intervenția prealabilă a unui om digitale (raze X, RMN, scanări etc.).

Însă, tot ce a fost descris ca având „consecințe pozitive” poate fi catalogat și la „consecințe negative”. Aceleași instrumente permit supravegherea pe scară largă a indivizilor, organizarea câmpurilor de luptă sau influențarea formării opiniilor, cauzând – în timp – pierderea relației de implicare corporală și pierderea construcției *sensului uman*, care poate fi realizată de om doar în relația sa cu o comunitate vorbitoare.

Astfel, sunt identificate deja numeroase probleme generate de utilizarea IA: problema protecției datelor personale; manipularea opiniilor prin profiluri psihologice la scară largă (afacerea *Cambridge Analytica* sau abuzurile pe Facebook); tendința de a vedea individul doar ca un exemplu statistic de tip general; pierderea controlului piețelor financiare ale căror comenzi de cumpărare și vânzare sunt executate de mașini într-un interval de timp mult prea rapid pentru a fi urmărit sau contestat de oameni.

În Drept, specialiștii în IT și juriștii au identificat anumite riscuri, obstacole și limite inerente folosirii IA în Justiție²⁴, generate de alegerea bazei de învățare, de evoluția algoritmilor, de particularitățile și complexitatea activității judiciare. Printre altele, s-a constatat că lipsa transparenței codurilor algoritmice subminează garanțiile procedurale care protejează prezumția de nevinovăție și dreptul la un proces echitabil. De asemenea, se pun probleme precum: măsura și modul de implicare a Statului (prin politicile sale publice), răspunderea în cazul unor daune

²⁴ L. Pécaut-Rivolier, S. Robin, *Justice et intelligence artificielle, préparer demain* – ep. II, publicat pe Dalloz Actualité, la 15.04.2020.

provocate de o digitalizare prost gândită (căci acceptarea socială a utilizării algoritmilor este strâns legată de atribuirea responsabilității în caz de daune), etica IA, exercitarea ilegală a profesiilor juridice de către algoritmi etc.

Vom detalia analiza asupra acestor probleme pe parcursul unor articole viitoare.

6. Prime concluzii

Suntem în permanență balansați între bine și rău atunci când enumerăm obiectivele și efectele Inteligenței Artificiale, în domeniile socio-umane. În acest context, domeniul juridic este zguduit de apariția noilor tehnologii, iar juriștii sunt îngrijorați de viitorul profesiilor lor. Cei mai sceptici susțin că asistăm la confuzia progresivă a legii persoanelor și a legii bunurilor, ceea ce este fatal pentru umanitate, anunțând dispariția progresivă a statului de drept și instaurarea „post-umanismului”²⁵ sau a „transumanismului”.

Și în România, autorii trag semnale de alarmă asupra consecințelor - unele imprevizibile și ireversibile - ale interacțiunii tot mai complexe dintre om și mașină, asupra tensiunii existente între doi hegemoni: avem, pe de o parte, sistemele robotice care pun în discuție valori fundamentale și standarde tradiționale în materia drepturilor omului, creând pericolul unor forme superioare de inegalitate, iar pe cealaltă parte, „valorile demnității, egalității și solidarității umane, urmărind adaptarea lor la cerințele noilor tehnologii”²⁶. Este cel puțin tulburător acest tablou al omului „augmentat” sau ameliorat prin biotehnologie, „eliberat” de vulnerabilități și de umilitatea creștină, dar și lipsit de „iubire, empatie, răs și plâns”²⁷...

Cert este că dezvoltarea algoritmilor și a inteligenței artificiale ne invită să luăm în considerare modul nostru de a gândi despre Drept, Dreptate și despre Scopul studiului științelor juridice.

Însă, cercetătorul J. Lassègue ne îndeamnă să punem în valoare atuul pe care îl avem: spre deosebire de multe alte profesii, juriștii sunt obișnuiți cu raționamentul formal și, fără îndoială, vor ști, mai bine decât mulți alții, cum să adopte noile tehnologii pentru a continua să producă sentimentul comun al dreptății²⁸.

²⁵ J. X. Labbé, *La fin du monde, la fin du droit ou la transition juridique?*, Recueil Dalloz 2019, p. 78.

²⁶ N. Voiculescu, M.B. Berna, *Eu, Robotul! Actualitatea unei dihotomii: drepturile omului vs. drepturile roboților*, pe www.universuljuridic.ro (18.09.2020).

²⁷ A. Ciucă, *Robinsonada omului augmentat*, pe www.universuljuridic.ro (02.03.2021).

²⁸ J. Lassègue, *idem*, *Les cahiers de la justice* 2019, p. 205.