Cum protejează Regulamentul UE privind IA drepturile fundamentale și cum ar trebui să funcționeze?

1. Introducere

Sistemele de inteligența artificială IA poate părea o tehnologie disruptivă și imposibil de fi reglementată de om. Zilnic, vedem titluri în presă despre cum inteligența artificială va înlocui o anumită meserie, va schimba fața educației și cum poate pune în pericol munca unor creatori. În această optică, inteligența artificială pare o tehnologie - sau un set de tehnologii - care nu pot fi cuprinse într-un cadru legal care să le gestioneze și reglementeze, find prea complexă, prea diversă și mereu în schimbare.

De fapt, sistemele IA sunt deja reglementate. Și în UE, și în SUA sau în China ori India. Pentru a vorbi de lucrurile pe care le știm și ne afectează direct în Uniunea Europeană sau în România, există legislație care se aplică deja. Sigur, s-ar putea să fie norme diferite, dacă ar fi cazul despre prelucrarea datelor personale (caz în care se va aplica GDPR), fie despre aspecte care implică drepturile de autor ale operelor folosite în instruirea AI (caz în care se va aplica legislația națională care transpune directivele UE) sau dacă ar fi o situație despre folosire a

sistemelor de inteligență artificială în scop anticoncurențial (caz în care se va aplica cadrul legal național privind concurența).

În acest context, noul Regulamentul UE privind Inteligența Artificială (AI ACT sau mai pe scurt AIA) își propune să creeze un cadru general care să suplimenteze aceste reglementări deja existente, cu noi obligații specifice doar pentru acest tip de tehnologii complexe, diverse și mereu în schimbare.

Regulamentul a fost propus de Comisia Europeană pe data de 21 aprilie 2021, a fost adoptat de Parlamentul European la 13 martie 2024 și apoi aprobat în unanimitate de Consiliul UE pe data de 21 mai 2024, deci publicarea lui în Jurnalul Oficial al UE este iminentă

Aplicarea AIA se va face etapizat, practic începând cu anul 2025 cu o aplicare generală abia de la jumătatea anului 2026. AIA creează un Consiliu European privind Inteligență Artificială pentru a promova cooperarea națională și a asigura respectarea regulamentului. Obligațiile AIA se aplică și furnizorilor din afara UE dacă aceștia direcționează activitatea către utilizatori din UE, similar cu prevederi din Regulamentul GDPR, de exemplu.

Implementarea Regulamentului privind Inteligența Artificială la nivelul întregii UE este efectuată de European AI Office. Printre scopurile AI Office, se numără, de exemplu "promovarea de acțiuni și politici pentru a profita de beneficiile sociale și economice ale IA în întreaga UE" și facilitarea dezvoltării ecosistemelor de IA de încredere la nivel European.

AIA este o așa-numită „legislație orizontală", care este numită astfel deoarece se aplică în aproape toate sectoarele de activitate umană, industrială și economică. Această lege urmărește îmbunătățirea pieței europene de tehnologie prin promovarea utilizării sigure a inteligenței artificiale (IA), respectând "drepturile omului și protejând sănătatea, siguranța și mediul". Accentul cade pe respectarea drepturilor omului în acest Regulament european, având în vedere diversele moduri în care sistemele de inteligență artificială ne periclitează și vulnerabilizează drepturile noastre fundamentale.

În următoarele capitole, detaliem modul de funcționare al regulamentului, cum ne protejează drepturile noastre fundamentale, care-i sunt limitările și ce ne spune asta despre viitorul tehnologiilor noastre care folosesc sisteme IA.

2. Cum ar funcționa regulamentul privind inteligența artificială?

Pentru a înțelege ce se reglementează, regulamentul privind inteligența artificială are nevoie de o definiție specifică și destul de solidă conceptual pentru sistemele IA. Mai mulți observatori, printre care și cercetătoarea în domeniul IA Kate Crawford, au punctat faptul că definiția dată sistemelor IA poate să includă unele tehnologii care folosesc sisteme IA în mod concret, iar pe altele să le excludă. Așadar, definițiile sunt esențiale în abordarea legală a sistemelor IA.

Regulamentul privind inteligența artificială definește aceste sisteme drept "**un sistem bazat pe o mașină care este conceput pentru a funcționa cu diferite niveluri de autonomie și care poate da dovadă de adaptabilitate după implementare și care, pentru obiective explicite sau implicite, deduce, din datele pe care le primește, cum să genereze rezultate, cum ar fi predicții, conținut, recomandări sau decizii care pot influența mediile fizice sau virtuale**".

Mai departe, regulamentul privind inteligența artificială clasifică sistemele de inteligență artificială pe baza riscului lor. Clasificarea este următoarea:

- **Sisteme IA cu un grad de risc inacceptabil**. Riscul inacceptabil este interzis la nivel de UE (de exemplu, sistemele de credit social, pe modelul chinez, și IA care folosesc "manipularea cognitiv-comportamentală a persoanelor sau a anumitor grupuri vulnerabile" - adică sisteme care în mod explicit caută să înșele și să influențeze negativ comportamentul cetățenilor UE, cum ar fi jucăriile care "încurajează comportamentul periculos la copii").
- **Sisteme IA cu un grad de risc ridicat**. Cea mai mare parte a textului Regulamentului privind IA se referă la sistemele de IA care pot afecta "negativ siguranța sau drepturile fundamentale", sunt reglementate într-un mod strict.
- **Sisteme IA cu un risc limitat.** O secțiune mai mică din regulament se ocupă de sistemele de IA cu risc limitat, care fac obiectul unor obligații de transparență mai ușoare. Dezvoltatorii și implementatorii trebuie să se asigure că utilizatorii finali sunt conștienți de faptul că interacționează cu IA (chatbots și deepfakes), find obligați să-și informeze utilizatorii de acest lucru.

- **Sisteme IA cu un grad de risc minim (sau inexistent)**. Riscul minim nu este reglementat (cele mai multe aplicații cu IA sunt de genul acesta, cum ar fi jocurile video cu IA sau filtrele de spam).

Majoritatea obligațiilor privind respectarea acestor riscuri - sau preîntâmpinarea lor - cad pe furnizorii de sisteme IA, și nu pe utilizatorii lor. Producătorii de sisteme IA trebuie să treacă prin diverse proceduri de verificare și raportare a riscurilor aduse de tehnologiile lor. Verificarea este făcută chiar de compania care dorește să-și împingă produsul pe piața europeană, și nu de o parte terță. Acest lucru a fost criticat dur de unii observatori. Conform European Digital Rights (EDRi), această verificare făcută doar intern de companii, fără observatori externi, riscă să pericliteze și vulnerabilizeze drepturile omului pe care acest regulament vrea să le protejeze.

Riscul este definit, în cadrul acestui regulament drept "combinația dintre probabilitatea de producere a unui prejudiciu și gravitatea acelui prejudiciu" adus la adresa drepturilor fundamentale ale utilizatorului. Drepturile fundamentale ale utilizatorului sunt puse în centrul acestui regulament, iar această clasificare pe bază de risc are ca scop urmărirea (și interzicerea) acelor sisteme care pot afecta sau atinge drepturile fundamentale. Regulamentul instituie o bază de date a UE, accesibilă publicului, pentru a oferi transparență cu privire la sistemele de inteligență artificială care prezintă riscuri mari pentru drepturile sau siguranța persoanelor.

De exemplu, sistemele IA care clasifică automatizat (de exemplu pe bază de stare de sănătate, performanță la muncă, locație, opinii politice, rasă, clasă, interese) persoanele pe baza datelor cu caracter personal colectate și procesate sunt mereu încadrate, conform AIA, ca fiind sisteme cu risc ridicat. Ipotetic, dacă o platformă digitală bazată pe sisteme IA - să spunem Facebook - ar folosi date despre un utilizator ca să infereze din celelalte date în ce crede acest utilizator, care sunt preferințele sale politice etc., acest sistem de la baza platformei ar putea fi clasificat, conform AIA, drept un risc ridicat adus la adresa drepturilor fundamentale ale utilizatorilor.

Regulamentul interzice explicit următoarele sisteme IA, clasificate ca fiind cu un grad de risc inacceptabil:

- utilizarea tehnicilor subliminale, manipulatoare sau înșelătoare, exploatarea vulnerabilităților legate de vârstă, handicap sau circumstanțe socio-economice pentru a influența comportamentul și sistemele de clasificare biometrică care deduc date sensibile despre utilizator, cum ar fi rasa sau opiniile politice, care n-au fost exprimate într-un mod explicit.
- De asemenea, se mai interzic în mod explicit tehnologii cu risc inacceptabil: acordarea automatizată de credit social, ca în cazul Chinei, care are ca rezultat un tratament periculos la adresa persoanelor fizice, evaluarea riscului infracțional pe baza exclusivă a clasificării sau categorizării persoanelor sau a trăsăturilor de personalitate (ca în cazul COMPAS publicat de ProPublica despre recidivă),
- compilarea bazelor de date de recunoaștere facială, și deducerea emoțiilor la locul de muncă sau în instituțiile de învățământ - cu excepția cazului în care există motive medicale sau de siguranță.
- În plus, identificarea biometrică de la distanță (RBI) „în timp real” în spații accesibile publicului este limitată la scenarii de ”aplicare a legii”, cum ar fi căutarea persoanelor dispărute, prevenirea amenințărilor la adresa vieții persoanelor fizice sau identificarea suspecților în cazul unor infracțiuni grave, cum ar fi crima sau terorismul.

Modul în care poliția se poate folosi de această excepție în interzicerea identificării biometrice de la distanță, pentru infracțiuni sau terorism, nu este clar și există riscul abuzului unei astfel de excepții pentru himere precum "securitatea națională".

3. Regulamentul IA și drepturile omului: blocarea inovării?

Sistemele de inteligență artificială au stârnit deja anxietăți legitime în cetățenii Europei. Conform unui sondaj Eurobarometer, aproape 9/10 cetățeni europeni consideră faptul că roboții și sistemele IA trebuie gestionate cu grijă.

Știm că liderii Big Tech (cum ar fi Meta, Google, Amazon) din industrie argumentează că acest regulament riscă să stopeze inovarea. Digital Europe - un grup de lobby care reprezintă aproape 45.000 de companii din sectorul digital - au argumentat că regulamentul ar putea sufoca inovarea și competitivitatea. Dar această narațiune falsă și înșelătoare, conform căreia un regulament menit doar să protejeze drepturile cetățenești este un impediment în calea

progresului tehnologic, este o propagandă exact în interesul acestor firme - și foarte des întâlnită în raport cu acest regulament.

În realitate, drepturile fundamentale nu trebuie echilibrate cu inovația în domeniu, ci inovarea în domeniu trebuie să respecte drepturile fundamentale. Președinta Signal Meredith Whittaker observă că **povestea conform căreia UE împiedică inovarea prin intermediul reglementărilor este "atât greșită, cât și suspect de egocentrică atunci când vine din gura giganților din domeniul tehnologiei și a celor care se ocupă cu promovarea acestora"**. Faptul că interesele financiare private ale companiilor Big Tech se ascund sub vălul "inovării" nu este surprinzător, dar astfel de narațiuni tind să ducă la compromisuri dureroase pentru drepturile fundamentale, lucru care deja a fost observat de ONG-uri precum AccesNow, care au criticat ultima versiune a regulamentului pentru concesiile sale în fața industriei de IA și lobby-ului său intens. Criticile AccesNow și a altor ONG-uri (EDRi și Amnesty International) s-au concentrat cel mai mult pe faptul că regulamentul nu reușește să interzică în mod adecvat unele dintre cele mai periculoase utilizări ale inteligenței artificiale, cum ar fi sistemele care permit supravegherea biometrică în masă și sistemele de predicție comportamentală ale poliției (predictive policing).

Așadar, țelul principal al AIA este protecția drepturilor noastre fundamentale în fața unei tehnologii noi și bine finanțate, iar acest lucru nu este incompatibil cu inovarea în domeniu. Prin interzicerea unor sisteme specifice de IA (risc inacceptabil) și reglementarea strictă a altora (risc ridicat) inovația nu este afectată în mod negativ, ci canalizată spre alte sisteme care au în design-ul lor, încă de la început, protecția drepturilor fundamentale ca scop. De altfel, conceptul de “etică prin design” poate fi deja găsit în codurile de conduită pe tema inteligenței artificiale de la UE.

4. Limitările AIA

În opinia noastră, Regulamentul privind inteligența artificială are niște limitări și lipsuri concrete. Printre ele se numără, conform unei analize detaliate de la EDRi, unde și ApTI este membru, faptul că regulamentul eșuează în protecția statului de drept și a spațiului civic, privilegiind interesele industriei de IA, serviciilor de informații, și organismelor legii (poliția). De exemplu, excepțiile din regulament privind monitorizarea și identificarea biometrică de la distanță efectuate de către poliție în căutarea teroriștilor lasă loc de abuzuri,

fiindcă securitatea națională (sau terorismul) nu sunt clar menționate, definite și încadrate. Conform aceleași analize EDRi, deși regulamentul obligă dezvoltatorii de IA să mențină ”standarde ridicate pentru dezvoltarea tehnică a sistemelor de IA (de exemplu, în ceea ce privește documentația sau calitatea datelor)”, măsurile menite să protejeze drepturile fundamentale, inclusiv drepturile și libertățile civile ”sunt insuficiente pentru a preveni abuzurile” mai devreme menționate. În final, EDRi adaugă, în raport cu clauzele specifice ale Regulamentului privind Inteligența Artificială de la UE că aceste dintâi clauze sunt ”pline de excepții de mare amploare”, care riscă să reducă protecția cetățenilor europeni ”în special în domeniul aplicării legii și al migrației”.

5. Viitorul nostru comun

Dezvoltarea nestăvilită și neîngrădită a inteligenței artificiale poate să acopere din ce în ce mai multe aspecte ale vieții noastre, ceva numit de activiștii La Quadrature du Net "supra-computerizarea societății". Această supra-computerizare este alimentată de injecțiile mari de capital din industria sistemelor IA. Sistemele de inteligență artificială pot fi noua mină de aur - sau noul mijloc de îmbogățire rapidă și exponențială pentru tehnologiști. Sistemele IA, prin ele însele, nu sunt nici bune nici rele, ci ține de cine le folosește și pentru ce scopuri.

În acest context, regulamentul privind inteligența artificială este o încercare de a înțelege și clarifica viitorul nostru comun digitalizat, supra-computerizat. Sub tutela acestui regulament, viitorul inteligenței artificiale și raportul lor cu drepturile noastre fundamentale ar trebui, teoretic, să devină mai predictibil. Însă adaptabilitatea regulamentului la schimbările tehnologice este incertă și poate lăsa multe lacune neadresate, ca cele mai devreme menționate.

Fiecare regulament sau set de legi poate fi văzut ca un pariu pentru un viitor diferit de trecut, sau pentru preîntâmpinarea unui viitor nedorit. După cum nota faimoasa avocată pe drepturile omului Susie Alegre, scriind despre tehnologiile IA folosite din ce în ce mai mult "Dacă ne permitem să fim dependenți de tehnologie, drepturile noastre vor depinde de funcționarea acelei tehnologii și vom deveni tributari sistemelor, proprietarilor lor umani sau oricui altcuiva care o poate controla sau întrerupe de la sursă". Alegre nu sugerează aici că toate

tehnologiile sunt prin ele însele rele, ci că nu ar trebui să plătim noi - mai ales când vine vorba de drepturile *noastre* fundamentale - doar în numele unui progres tehnologic incert.

Prin urmare, regulamentul privind inteligența artificială - cu unele limitări, carențe și neajunsuri - propune un prim pas spre reglementarea inteligenței artificiale de astăzi și de mâine. Dacă regulamentul va reuși în protecția drepturilor fundamentale ale utilizatorilor de internet, rămâne de văzut. Acest lucru este, pentru moment, un pariu de făcut cu viitorul.