ETICA ALGORITMICĂ ȘI STANDARDELE EUROPENE: ÎNTRE AUTO-REGLEMENTARE ȘI INTERVENȚIE NORMATIVĂ

ALGORITHMIC ETHICS AND EUROPEAN STANDARDS: BETWEEN SELF-REGULATION AND NORMATIVE INTERVENTION

CZU: 004.89:17:006.35(4)CEN **https://doi.org/10.59295/spd2025s.16**

Mariana TACU
ORCID: 0000-0003-4930-2395
E-mail: mariana.tacu@usm.md
Universitatea de Stat din Moldova, Chișinău, Republica Moldova

Summary. *In the context of the rise of artificial intelligence and the widespread use of decision-making algorithms in critical fields such as healthcare, justice, security, and digital communication, the ethical dimension of these technologies is becoming an indispensable component of academic inquiry and public policy development. Within the European Union, the debate over the regulation of algorithmic technologies unfolds between two distinct yet interconnected normative paradigms: voluntary self-regulation, promoted by industry actors, and legislative intervention, undertaken by public institutions. The former is grounded in the principles of flexibility and the private sector's capacity to respond swiftly to technological innovation, but it carries the risk of fragmented approaches and a diffuse sense of ethical responsibility. The latter seeks to ensure the systematic protection of fundamental rights and to establish clear safeguards concerning transparency, fairness, and accountability, although it introduces requirements that, if not properly calibrated, may reduce the responsiveness and adaptability of the technological ecosystem to the rapid pace of digital transformation.*

Keywords: *artificial intelligence, algorithmic governance, ethical responsibility, self-regulation, fundamental rights, transparency, technological innovation, digital policies.*

Europa se află astăzi la confluența a două direcții de acțiune aparent opuse, dar în realitate complementare, în ceea ce privește configurarea normativă a eticii algoritmice: pe de o parte, autoreglarea voluntară a industriei tehnologice și, pe de altă parte, intervenția normativă asumată de autoritățile publice. În ultimii ani, dinamica accelerată a dezvoltării inteligenței artificiale, în special în domeniile aplicative de mare impact social, precum cel mediatic comunicațional, a generat preocupări tot mai evidente legate de legitimitatea deciziilor automatizate, de posibilitatea perpetuării unor inegalități sistemice prin intermediul algoritmilor și de dificultatea identificării responsabilității în cadrul unor procese din ce în ce mai autonome și obscure.

Deși, teoretic, aceste provocări ar fi putut fi abordate prin autoreglare instituțională și prin angajamente etice voluntare asumate de actori privați, amploarea consecințelor asupra drepturilor individuale și structurilor sociale impune o reflecție mai profundă

asupra eficienței și legitimității acestui model autoreferențial. În absența unui cadru normativ clar și validat la nivel european, dar și internațional, există riscul ca discursul etic promovat de marii actori tehnologici – precum Google, Microsoft sau OpenAI – să rămână unul declarativ, cu o funcție mai degrabă simbolică decât operațională, consolidând aparența unei responsabilități asumate, dar fără garanții privind aplicarea sa efectivă. În acest context, fenomenul cunoscut sub denumirea de „greenwashing" etic (afișarea unei preocupări etice în scopuri de reputație, se referă la acțiunea de a face afirmații false sau înșelătoare cu privire la impactul pozitiv asupra mediului pe care o companie, un produs sau un serviciu l-ar avea) [1] – devine un obstacol semnificativ în realizarea unei guvernanțe algoritmice cu adevărat orientate spre interesul public.

Etica algoritmică între potențial transformator și riscuri sistemice

Algoritmii au devenit un element esențial care stă la baza celor mai importante infrastructuri și servicii ale societăților informaționale. Indivizii interacționează zilnic cu sisteme de recomandare – sisteme algoritmice care sugerează utilizatorului ce i-ar putea plăcea, fie că este vorba de alegerea unei melodii, a unui film, a unui produs sau chiar a unui prieten. În același timp, școlile și spitalele**,** instituțiile financiare**,** instanțele de judecată**,** autoritățile publice locale și cele centrale se bazează din ce în ce mai mult pe algoritmi pentru a lua decizii importante.

Potențialul algoritmilor de a îmbunătăți bunăstarea individuală și socială este însoțit, însă, și de riscuri etice majore, printre acestea:

- reproducerea și amplificarea inegalităților sociale;
- eroziunea transparenței și a explicabilității decizionale;
- fragmentarea responsabilității morale și juridice;
- profilarea invazivă și restrângerea autonomiei personale;
- supravegherea generalizată și intensificarea tendințelor autoritare;
- consolidarea asimetriilor de putere informațională și inechitatea în accesul la tehnologie.

În acest sens, cercetările privind etica algoritmilor s-au intensificat considerabil după 2016, când guvernele naționale, organizațiile neguvernamentale și companiile private au început să joace un rol central în dezbaterea despre o Inteligență Artificială (IA) „echitabilă" și „etică" (Sandvig et al. 2016; Binns 2018; Selbst et al. 2019; Wong 2019; Ochigame 2019) [2].

Tot în anul 2016, Mittelstadt și colaboratorii săi au propus o hartă conceptuală care continuă să reprezinte un reper teoretic solid în analiza critică a problematicii etice asociate algoritmilor. Această structură conceptuală oferă un cadru coerent și sistematic de interpretare a dinamicii dintre tehnologia algoritmică și implicațiile sale asupra valorilor morale, facilitând atât identificarea principalelor tipuri de îngrijorări etice, epistemice și normative, cât și formularea unor răspunsuri etice pertinente. Relevanța sa se menține în actualul context tehnologic marcat de proliferarea sistemelor automatizate și de necesitatea regândirii relației dintre responsabilitatea umană și decizia algoritmică.

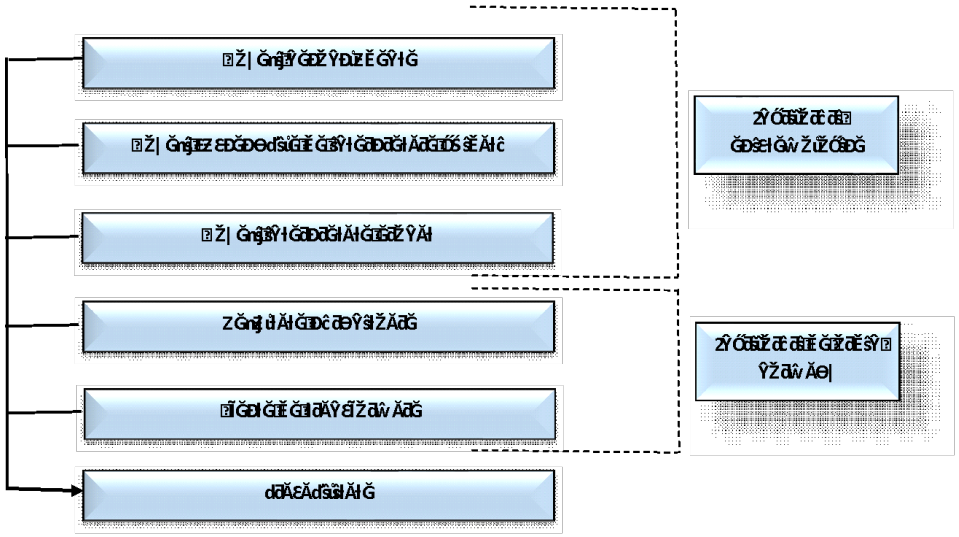


Figura 1. Șase tipuri de preocupări etice generate de algoritmi

Sursa: Tsamados, A., Aggarwal, N., Cowls, J. The ethics of algorithms: key problems and solutions. Disponibil: https://link.springer.com/article/10.1007/s00146-021-01154-8

Raportate la mass-media, aceste riscuri pot fi prezentate și contextualizate astfel:

1. Devin vizibile sub forma **dovezilor neconcludente** (*inconclusive evidence*), adică a acelor informații care, deși preluate și difuzate în spațiul public, nu oferă suficientă claritate pentru a fundamenta o concluzie coerentă. În jurnalismul bazat pe algoritmi (supranumit astăzi și jurnalism computațional: de la chatboți și sisteme de recomandare, până la inteligență artificială și automatizare în timp real) [3, p. 23], astfel de dovezi pot fi promovate pe criterii de atractivitate sau de probabilitate de viralizare, contribuind la consolidarea incertitudinii și la erodarea încrederii în mijloacele de informare în masă.

2. Se manifestă prin **dovezile susceptibile de interpretare ghidată** (*instructable evidence*), acele informații care, deși aparent obiective, sunt integrate într-un cadru narativ prestabilit, modelat de preferințele algoritmice, de interesele editoriale sau de logica de funcționare a platformelor digitale. Într-un astfel de cadru, algoritmii nu doar selectează conținutul, ci influențează interpretarea acestuia, ceea ce poate conduce la o formă subtilă, dar persistentă, de manipulare mediatică.

3. Diseminarea **dovezilor distorsionate sau eronat interpretate** (*misguided evidence*). În lipsa unui control editorial riguros, datele pot fi scoase din context, combinate necorespunzător sau prezentate în mod tendențios, contribuind la dezinformare. Acest fenomen este amplificat de algoritmii care privilegiază conținutul emoțional, controversat sau polarizant, în detrimentul celui echilibrat, dar mai puțin spectaculos.

În acest context, putem vorbi și de conceptul de post-adevăr, care reflectă o transformare profundă a imaginarului colectiv, fenomen analizat de filosoful și sociologul francez al teoriei informațiilor Edgar Morin, care subliniază că, odată cu expansiunea

mass-mediei, imaginarul a devenit o componentă structurală a vieții cotidiene. În era rețelelor sociale, acesta se fragmentează în bule informaționale, unde circulă narațiuni și convingeri nefiltrate critic. Absența mecanismelor editoriale tradiționale și democratizarea producției de conținut favorizează proliferarea informațiilor distorsionate, determinând o cultură a adevărului personalizat, în care indivizii își construiesc realitatea pe baza selecției subiective a surselor informaționale [4].

4. Impactul acestor disfuncții se traduce adesea în **rezultate părtinitoare sau inechitabile** (*unfair outcomes*), care afectează reprezentarea corectă a anumitor categorii sociale, marginalizează vocile minoritare sau consolidează stereotipuri deja existente. Într-un peisaj mediatic contemporan dominat de logici algoritmice, distribuția inegală a vizibilității devine o problemă etică fundamentală, în mod special atunci când algoritmii operează în mod lipsit de transparență și fără supraveghere democratică.

5. Un rol important în acest context este jucat de **efectele de transformare profundă** (*transformative effects*) pe care aceste procese le exercită asupra relației dintre mass-media și public. Algoritmii nu doar intermediază accesul la informație, ci modelează preferințele, comportamentele și chiar modul în care utilizatorii înțeleg realitatea. Mass-media algoritmică influențează astfel nu doar ce află publicul, ci și cum învață, cum reacționează și cum se poziționează față de evenimentele sociale, culturale sau politice.

Efectele de transformare pot fi cercetate și prin prisma tendinței de personalizare algoritmică a conținutului, care conduce la distribuții adaptative și fluxuri de știri filtrate, ceea ce marchează o tranziție de la accentul tradițional al jurnalismului pe importanța comună, spre experiențe profund individualizate. Această schimbare are implicații semnificative asupra formării publicului și a efectelor mediatice. În mod particular, proliferarea algoritmilor a fost asociată cu apariția „bulelor de filtrare" și cu dezvoltarea unei realități algoritmice, care poate fi manipulată pentru a răspândi dezinformarea și informațiile false [5].

6. **Trasabilitatea informațională** (*traceability*) reprezintă o provocare tot mai vizibilă în peisajul media actual. Capacitatea de a identifica sursa originară a unei informații, a unei imagini de a urmări modificările suferite pe parcursul circulației digitale și de a atribui responsabilitatea pentru eventualele distorsionări este frecvent limitată de dinamica algoritmică. Lipsa trasabilității subminează mecanismele clasice de verificare jurnalistică, făcând mai dificilă deconstrucția critică a fluxurilor informaționale și creând un spațiu fertil pentru manipulare și ambiguitate.

Aceste riscuri etice, contextualizate în sfera mass-mediei, relevă complexitatea relației dintre algoritmi, cunoaștere și responsabilitate. Ele indică necesitatea unor noi cadre de reflecție etică și a unor mecanisme de reglementare capabile să răspundă provocărilor emergente într-un peisaj informațional dinamic, intermediat tot mai mult de tehnologii autoreproductive.

În contexte de criză, precum pandemia de COVID-19, războiul din Ucraina, campania electorală pentru alegerile prezidențiale din România, din noiembrie 2024 - riscurile etice asociate algoritmilor se amplifică semnificativ, devenind nu doar mai vizibile, ci și mai influente asupra formării opiniei publice. Presiunea informativă, nevoia de reacție

rapidă și creșterea consumului digital determină o favorizare a conținutului emoțional, fragmentar și ușor de viralizat, în detrimentul informației verificate și contextualizate. Astfel, dovezile neconcludente și cele distorsionate sunt diseminate rapid, generând confuzie și afectând capacitatea publicului de a distinge între fapte și opinii. Algoritmii personalizează fluxurile informaționale în funcție de istoricul utilizatorului, consolidând narațiuni ideologizate și facilitând polarizarea discursului public. În același timp, vocile minoritare sau criticile legitime devin adesea invizibile, iar mecanismele de trasabilitate informațională sunt grav afectate de volumul și viteza circulației datelor. În asemenea circumstanțe, algoritmii nu mai sunt simpli mediatori, ci actori activi ai realității sociale, modelând percepțiile, comportamentele și valorile. Aceste fenomene evidențiază limitele autoreglării și impun dezvoltarea unor mecanisme normative proactive, capabile să protejeze integritatea spațiului public în situații de vulnerabilitate colectivă.

Guvernanța algoritmică europeană: între flexibilitate și rigoare normativă

Uniunea Europeană (UE) a devenit la nivel global un reper al reglementării stricte în domeniul IA, prin instrumente legislative precum: **General Data Protection Regulation (GDPR)** - *Regulamentul General privind Protecția Datelor*. Regulamentul UE 2016/679 al Parlamentului European și al Consiliului, adoptat la 27 aprilie 2016 este aplicabil din 25 mai 2018 în toate statele membre ale Uniunii Europene. GDPR stabilește cadrul legal pentru protecția datelor cu caracter personal ale persoanelor fizice și se aplică tuturor organizațiilor care prelucrează astfel de date, inclusiv în afara UE, dacă vizează cetățeni europeni. Tehnologia a transformat atât economia, cât și viața socială și ar trebui să faciliteze în continuare libera circulație a datelor cu caracter personal în cadrul Uniunii, precum și transferul acestora către țări terțe și organizații internaționale, asigurând totodată un nivel ridicat de protecție a datelor personale [6].

Digital Services Act (DSA) - *Actul legislativ privind serviciile digitale* a fost adoptat la 4 octombrie 2022. DSA formează un set unic de norme aplicabile în întreaga Uniune Europeană. Acestea urmăresc două obiective principale: 1. Crearea unui spațiu digital mai sigur, în care drepturile fundamentale ale tuturor utilizatorilor de servicii digitale să fie protejate; 2. Stabilirea unor condiții echitabile de concurență care să stimuleze inovația, creșterea economică și competitivitatea, atât pe piața unică europeană, cât și la nivel global [7].

Artificial Intelligence Act (AI Act) - *Legea privind inteligența artificială*. Este prima lege cu aplicare transnațională dedicată explicit inteligenței artificiale, care a intrat în vigoare la 1 august 2024 (Regulamentul 2024/1689) și prevede o implementare etapizată până în anul 2027. Potrivit acestei legi sistemele de IA sunt clasificate în patru categorii de risc: 1. risc inacceptabil (interzise începând cu 2 februarie 2025); 2. risc înalt; 3. risc limitat și 4. risc minim. În plus, legea introduce o categorie distinctă pentru inteligența artificială generativă și modelele de uz general (GPAI), cărora li se aplică obligații suplimentare de transparență, documentare tehnică și responsabilitate juridică, în conformitate cu cerințele reglementării europene [8].

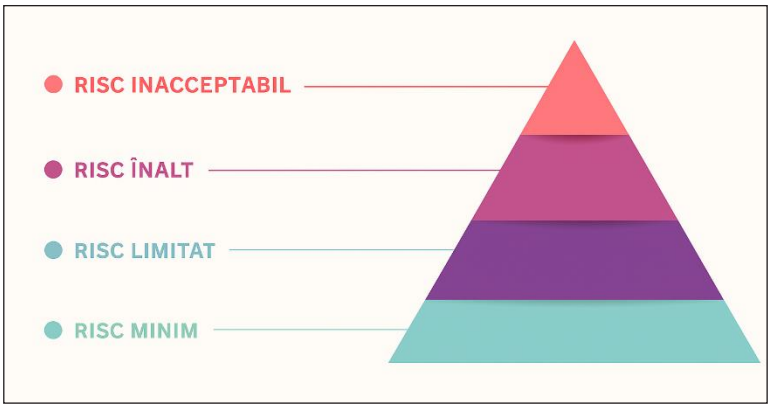


Figura 2. AI Act - 4 niveluri de risc pentru sistemele de inteligență artificială
Sursa: European Commission. AI Act. Disponibil: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

În această ordine de idei, abordarea bazată pe riscuri a fost analizată critic și de cercetători, precum Jonas Schuett, subliniind dificultățile normative inerente în evaluarea și definirea riscurilor legate de demnitatea umană, autodeterminare informațională, justiție și bine comun – elemente care nu se pretează ușor statisticii și clasificărilor binare [9].

Din perspectivă teoretică, absența unor mecanisme clare și eficiente de responsabilizare juridică, inclusiv în ceea ce privește răspunderea civilă sau penală pentru prejudiciile generate de sistemele de inteligență artificială, continuă să mențină un decalaj semnificativ între principiile normative declarate și aplicabilitatea lor concretă în practică. Pentru a reduce acest dezechilibru, legislația europeană, prin AI Act, prevede o serie de măsuri aplicative care presupun instituirea unor coduri de conduită obligatorii, elaborate în scopul detalierii obligațiilor de transparență, efectuării evaluărilor de impact asupra drepturilor fundamentale și implementării unor mecanisme de monitorizare publică.

În contextul mijloacelor de informare în masă tradiționale transpuse în spațiul virtual și a noilor media, adoptarea DSA și a AI Act marchează o reconfigurare substanțială a cadrului normativ în domeniu. Aceste acte reglementează modul în care este gestionat conținutul online, impun standarde de transparență algoritmică și introduc cerințe privind responsabilitatea platformelor în raport cu diseminarea informației, combaterea dezinformării și protejarea pluralismului. Pentru media digitală, în special, utilizarea IA în generarea sau moderarea conținutului devine condiționată de evaluări de risc, audituri și conformitate etică. În ansamblu, acest cadru legislativ favorizează o guvernanță echilibrată, care integrează libertatea de exprimare cu imperativele reglementării democratice într-un ecosistem media digital tot mai complex.

Acest cadru de guvernanță, caracterizat printr-o arhitectură normativă hibridă, ce îmbină transferul de responsabilități către actori privați cu supravegherea publică și impunerea de obligații cu caracter coercitiv, este implementat și prin intermediul instituțiilor

specializate la nivel european. Printre acestea, se numără **European AI Office** *(Oficiul European pentru Inteligența Artificială)* - https://digital-strategy.ec.europa.eu/en/policies/ai-office, **European Artificial Intelligence Board** *(Comitetul European pentru Inteligență Artificială)* - https://digital-strategy.ec.europa.eu/en/policies/ai-board.

În plus, inițiative precum **European Centre for Algorithmic Transparency (ECAT)** - *Centrul European pentru Transparența Algoritmică*, contribuie la consolidarea cadrului de guvernanță digitală prin furnizarea de expertiză științifică și tehnică esențială pentru implementarea DSA. Activitatea acestui centru vizează analiza sistemică a algoritmilor utilizați de platformele online de foarte mari dimensiuni, precum și formularea de recomandări privind bunele practici în materie de transparență algoritmică [10]. Astfel, se configurează un cadru de reglementare în care autoreglarea nu mai reprezintă o practică opțională sau unilaterală a actorilor privați, ci devine un mecanism integrat într-un cadru instituționalizat, caracterizat prin transparență, evaluare independentă și corelare cu mecanismele de supraveghere publică.

În 2025, cadrul european privind inteligența artificială se confruntă cu o serie de contradicții semnificative. Pe de o parte, angajamentul față de reglementări stricte, pe de altă parte - influența presiunilor economice și geopolitice. Acestea se manifestă prin intensificarea activităților de lobby ale marilor companii tehnologice americane (precum Google sau OpenAI), prin dezbateri privind limitarea aplicabilității unor dispoziții ale AI Act, dar și prin retragerea Directivei privind răspunderea civilă pentru IA, programată inițial pentru acest an. În timp ce oficialii europeni, precum *Henna Virkkunen*, subliniază necesitatea unui cadru normativ favorabil întreprinderilor mici și mijlocii [11], unele voci din Parlamentul European avertizează că o eventuală relaxare a reglementării modelelor generative ar putea genera riscuri democratice și incertitudine juridică. Pe acest fundal, Uniunea Europeană își reconfigurează strategia prin investiții semnificative în autonomia tehnologică – inclusiv prin programul InvestAI, care prevede 200 de miliarde de euro pentru dezvoltarea infrastructurii digitale, din care 20 de miliarde sunt alocate centrelor de calcul de mare capacitate [12]. Simultan, extinderea capacităților de supraveghere și implementarea de spații experimentale de reglementare („sandbox-uri") până în 2026 reflectă o abordare graduală și echilibrată a guvernanței algoritmice.

Deși nu revendică statutul unei soluții definitive și nu urmărește atingerea unei forme ideale, modelul european propune o traiectorie distinctivă, situată la intersecția dintre libertatea autoreglementării și constrângerile unei intervenții statale excesive. Această cale de mijloc, susținută de principii precum protejarea drepturilor fundamentale, transparența algoritmică și responsabilitatea etică, are potențialul de a deveni un reper global în guvernanța tehnologică.

Pentru ca acest echilibru să fie sustenabil, este esențial ca valorile asociate culturii digitale și eticii în utilizarea inteligenței artificiale să fie asumate la nivel societal. Nu este suficientă conformarea instituțională sau corporativă, este nevoie de alfabetizare digitală extinsă, de acces public la informații privind utilizarea IA și de implicarea activă a cetățenilor în procesele decizionale. Numai printr-o astfel de abordare participativă,

Europa își poate consolida poziția de lider în configurarea unei guvernanțe globale a IA, care să reflecte valorile democratice și umaniste.

REFERINȚE:

1. Gillis, A-S. What is greenwashing? Disponibil: https://www.techtarget.com/whatis/definition/greenwashing [Accesat: 03.06.2025].
2. Tsamados, A., Aggarwal, N., Cowls, J. The ethics of algorithms: key problems and solutions. Disponibil: https://link.springer.com/article/10.1007/s00146-021-01154-8 [Accesat: 03.06.2025].
3. Thurman, N., Lewis, C., Kunert, J. *Algorithms, Automation, and News. New Directions in the Study of Computation and Journalism.* Published by Routledge, 2023. 226 p. ISBN 9780367567545
4. Terra, F. *Fact-checking e DSA: strumenti e strategie contro la disinformazione*. Disponibil: https://www.agendadigitale.eu/cultura-digitale/fact-checking-e-dsa-strumenti-e-strategie-contro-la-disinformazione/?utm_source=chatgpt.com [Accesat: 08.06.2025].
5. Zamith, R. *Algorithms and Journalism.* In: Oxford Research Encyclopedia of Communication Disponibil:https://www.rodrigozamith.com/pubs/Zamith%20%20Algorithms%20and%20Journalism.pdf [Accesat: 08.06.2025].
6. European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council*. Disponibil: https://eur-lex.europa.eu/eli/reg/2016/679/oj [Accesat: 08.06.2025].
7. European Commission. *The Digital Services Act package.* Disponibil: https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package [Accesat: 08.06.2025].
8. The EU Artificial Intelligence Act. *Up-to-date developments and analyses of the EU AI Act*. Disponibil: https://artificialintelligenceact.eu/ [Accesat: 08.06.2025].
9. Schuett, J. *Risk management in the Artificial Intelligence Act*. Disponibil: https://arxiv.org/pdf/2212.03109 [Accesat: 08.06.2025].
10. European Commission. European Centre for Algorithmic Transparency. Disponibil: https://algorithmic-transparency.ec.europa.eu/index_en [Accesat: 10.06.2025].
11. News European Parliament. *Hearing of Executive Vice-President-designate Henna Virkkunen*. Disponibil: https://www.europarl.europa.eu/news/en/press-room/20241029IPR25054/hearing-of-executive-vice-president-designate-henna-virkkunen [Accesat: 10.06.2025].
12. European Commision. *Shaping Europe's leadership in artificial intelligence with the AI continent action plan.* Disponibil: https://commission.europa.eu/topics/eu-competitiveness/ai-continent_en [Accesat: 10.06.2025].