

# DREPTURILE OMULUI ÎN ERA INTELIGENȚEI ARTIFICIALE

ANDREEA MOROIANU\*

## Abstract:

The use of artificial intelligence has an impact on the fulfilment of fundamental human rights and freedoms. The regulation of AI systems and algorithms is one of the main priorities of international and regional human rights bodies. Given the increasing use of these technologies, it is essential to consider issues such as the right to privacy, confidentiality and responsibility in case of human rights violations. This paper refers to a series of documents issued by the UN, EU and the Council of Europe in order to establish a framework for the development and implementation of AI systems. A human rights based approach allows all stakeholders to participate in the creation and definition of principles representing the foundation of AI regulations.

**Keywords:** right to privacy, artificial intelligence, confidentiality, data processing, non-discrimination

## Résumé:

L'utilisation de l'intelligence artificielle a des effets sur la réalisation des droits et libertés fondamentales de l'homme. La réglementation des systèmes et des algorithmes fondés sur l'intelligence artificielle représente l'une des priorités des organismes internationales et régionales des droits de l'homme. Etant donné l'utilisation de plus en plus large de ces technologies, il est essentiel de prendre en compte certains aspects tels que le respect du droit à la vie privée, la confidentialité et la responsabilité dans le cas de possibles violations des droits de l'homme. Cet article se propose de présenter quelques documents élaborés au niveau de l'ONU et du Conseil de l'Europe qui visent d'établir le cadre général nécessaire au développement et à la mise en œuvre des systèmes fondés sur l'intelligence artificielle. Une approche fondée sur les droits de l'homme représente une modalité par laquelle les règlements en matière d'intelligence artificielle permettent la participation de toutes les parties intéressées dans l'élaboration et la définition des principes en la matière.

**Mots-clés:** droit à la vie privée, intelligence artificielle, confidentialité, traitement de données, non discrimination

Tehnologiile bazate pe inteligență artificială (IA)<sup>1</sup> devin treptat o parte integrantă din viața fiecărei persoane. Fie că ne referim la motoare de căutare,

---

\* E-mail: andreea.moroianu@irdo.ro

<sup>1</sup> În contextul lucrării de față, inteligența artificială se referă inclusiv la tehnologii precum învățare automată, robotică, rețele neurale, procesare lingvistică.

la electrocasnice inteligente, aplicații de rețele sociale sau mașini autonome, toate acestea au un impact asupra dezvoltării societății și, implicit, asupra vieților oamenilor. Tehnologiile IA sunt utilizate de companiile de rețele sociale pentru a modera conținutul de pe paginile lor, recomandă pagini care ar putea fi de interes, locuri de vizitat sau de petrecere a timpului liber. De asemenea, sistemele IA joacă un rol important în sistemul de sănătate, putând fi utile pentru depistarea unor afecțiuni precum cancerul. Mai mult decât atât, pot fi folosite pentru a îmbunătăți sistemul de educație sau pentru a determina candidatul potrivit pentru un anumit loc de muncă.

Având în vedere domeniile de utilizare și efectele pe care IA le poate avea este necesar să se adopte reglementări menite să asigure respectarea drepturilor și libertăților fundamentale ale omului. Astfel, statele, companiile și societatea civilă trebuie să se asigure că tehnologiile bazate pe IA ajută la realizarea drepturilor omului și nu aduc atingere acestora.

În ultimii ani, organismele internaționale au început să acorde o atenție din ce în ce mai mare implicațiilor IA asupra drepturilor și libertăților fundamentale și au fost elaborate o serie de recomandări, studii și rapoarte menite să ofere orientări statelor în vederea asigurării unei reglementări optime în materie.

### **Conceptul de Inteligență Artificială**

Termenul „intelență artificială” a apărut prima dată în anul 1956 când John McCarthy a ținut prima conferință academică pe această temă<sup>2</sup>. Cu toate acestea, nu a fost primul care a abordat subiectul sistemelor care au capacitatea de a gândi în mod autonom. În eseu *As we may think*<sup>3</sup>, publicat în anul 1945, Vannevar Bush vorbește despre dispozitive mecanizate care pot fi utilizate pentru a stoca informații și care pot fi consultate rapid, iar în anul 1950, Alan Turing scrie o lucrare pe tema sistemelor capabile să simuleze gândirea ființelor umane.

Inteligența artificială se bazează pe algoritmi, respectiv coduri informatice ce reprezintă instrucțiuni prin care anumite date pot fi transformate în concluzii, informații sau rezultate. Cantitatea de date existentă în momentul actual și capacitatea de prelucrare și analizare a acestora reprezintă motorul IA. Cu cât volumul și calitatea datelor sunt mai mari, cu atât algoritmiile produc rezultate mai precise. IA înglobează o serie de subdomenii, respectiv: învățare automată (*eng. machine learning*), învățare profundă (*eng. deep learning*), știința datelor (*eng. data science*) și robotică.

<sup>2</sup> Curs, *Istoria inteligenței artificiale*, Universitatea Washington, disponibil la: <https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf>

<sup>3</sup> <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>

Grupul de experți la nivel înalt privind IA al Uniunii Europene definește conceptul astfel: „Sistemele de inteligență artificială (IA) sunt sisteme software (și, eventual, hardware) proiectate de oameni, care, dacă li se dă un obiectiv complex, acționează în dimensiunea fizică sau digitală, percepând mediul prin intermediul preluării datelor, prin interpretarea datelor structurate sau nestructurate colectate, prin raționarea cu privire la cunoștințe sau prin prelucrarea informațiilor obținute din aceste date și prin deciderea celei/ celor mai bune acțiuni care trebuie întreprinse pentru a realiza obiectivul dat. Sistemele IA pot să utilizeze reguli simbolice sau să învețe un model numeric și, de asemenea, își pot adapta comportamentul analizând modul în care mediul este afectat de acțiunile lor anterioare.”<sup>4</sup>

Potrivit raportului special al ONU<sup>5</sup>, există 3 caracteristici ale IA care trebuie monitorizate din punct de vedere al drepturilor omului, respectiv: automatizarea, analiza de date și adaptabilitatea. Prin automatizare, factorul uman este eliminat din anumite etape ale procesului de luare a deciziilor, fiind înlocuit cu instrumente de calcul. Automatizarea permite prelucrarea unei cantități mari de date la viteze și la o scară care depășesc capacitatea umană, însă anumite seturi de date pot fi influențate de părtinire (*eng. bias*)<sup>6</sup> și pot avea efecte discriminatorii. Analiza de date stă la baza majorității aplicațiilor de IA, fie că este vorba de istoricul căutărilor pe internet, de datele privind traficul sau alte date personale. O problemă care apare din acest punct de vedere este reprezentată de drepturile asupra acestor date, de capacitatea sistemelor IA de a „deanonimiza date anonimizate”, de elementele părtinitoare cuprinse în seturile de date sau care sunt „întipărite prin formarea umană”<sup>7</sup>. În sfârșit, adaptabilitatea la care face referire raportul se referă la adaptabilitatea învățării automate, întrucât algoritmi aflați la baza sistemelor IA pot identifica, progresiv, noi probleme și genera noi rezolvări. În funcție de gradul de monitorizare, sistemele pot identifica anumite modele și ajunge la anumite concluzii neprevăzute de programatori sau de persoanele care au dat anumite

<sup>4</sup> Pentru mai multe informații se poate consulta documentul „O definiție a inteligenței artificiale (IA): principalele capacități și discipline științifice” elaborat de Grupul de experți la nivel înalt privind IA la UE, disponibil la: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=60665](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60665)

<sup>5</sup> Raport privind impactul IA asupra libertății de opinie și exprimare, A/73/348, disponibil la: <https://undocs.org/A/73/348>

<sup>6</sup> În limba engleză conceptul *bias* este definit ca acțiunea de a sprijini sau a se opune unei persoane sau unui lucru într-o manieră incorectă, din cauza părerilor personale care influențează opinia unei persoane. (<https://dictionary.cambridge.org/dictionary/english/bias>). În limba română termenul a fost tradus prin părtinire sau subiectivitate, însă se folosește termenul din engleză *bias*, inclusiv în traduceri realizate la nivelul Uniunii Europene

<sup>7</sup> Raportul ONU privind impactul IA asupra libertății de opinie și exprimare, A/73/348, p. 6.

sarcini. În acest fel, dacă factorul uman este exclus din definirea rezultatelor generate de IA, transparența și responsabilitatea devin o provocare, la fel ca și accesul la un remediu efectiv.

## **Drepturile omului în contextul sistemelor bazate pe inteligență artificială**

O abordare bazată pe drepturile omului<sup>8</sup> este esențială în dezvoltarea de politici, strategii și în reglementarea tehnologiilor bazate pe inteligență artificială. De asemenea, statele trebuie să țină cont de *Orientările ONU privind întreprinderile și drepturile omului*, având în vedere că majoritatea tehnologiilor bazate pe inteligență artificială sunt dezvoltate de companii private. În plus, instituțiile naționale de drepturile omului joacă un rol important în furnizarea de asistență statelor, pentru a identifica măsura în care elaborarea și implementarea normelor privind inteligența artificială respectă obligațiile în materie de drepturile omului, dar și în furnizarea de orientări companiilor și întreprinderilor private și altor actori non-statali.<sup>9</sup>

Așa cum am menționat anterior, organismele internaționale au început să acorde o atenție din ce în ce mai mare implicațiilor IA asupra drepturilor și libertăților fundamentale și au fost elaborate o serie de recomandări, studii și rapoarte la nivelul Consiliului Europei, iar la nivelul Uniunii Europene, printre altele, a fost elaborată Cartea albă privind „Inteligența artificială – O abordare europeană axată pe excelență și încredere”<sup>10</sup>. Într-un studiu al Consiliului Europei – „Algoritmi și drepturile omului”<sup>11</sup> se analizează impactul tehnologiilor moderne asupra unor drepturi, printre care se numără dreptul la un proces echitabil, confidențialitatea și protecția datelor, libertatea de exprimare și dreptul la un remediu efectiv. De asemenea, în anul 2018, la nivelul Oficiului Înaltului Comisariat al Națiunilor Unite pentru Drepturile Omului, raportorul special a înaintat Adunării Generale un raport privind

<sup>8</sup> Specificul abordării bazate pe drepturile omului constă în dezvoltarea de politici și de planuri de acțiune prin alocarea unei poziții centrale drepturilor omului și obligațiilor corelative ale statelor, având două obiective majore: sprijinirea titularilor de drepturi în exercitarea prerogativelor pe care le dețin și consolidarea capacității actorilor statali și non-statali în asigurarea protecției, promovării și protejării drepturilor omului.

<sup>9</sup> Orientările ONU privind întreprinderile și drepturile omului, p. 6, material disponibil la adresa: [https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr\\_en.pdf](https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf)

<sup>10</sup> COM(2020) 65 final.

<sup>11</sup> Comitetul de experți privind intermediarii (serviciilor) de internet (MSI-NET), Algorithms and human rights – Study on the human rights dimensions of automated data processing techniques and possible regulatory implications (Algoritmi și drepturile omului – Studiu privind dimensiunea drepturilor omului a tehnicilor de prelucrare automate de date și posibile implicații de reglementare), Consiliul Europei, 2017, disponibil la adresa <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

efectele IA asupra libertății de opinie și exprimare<sup>12</sup>, care examinează impactul inteligenței artificiale asupra mediului informațional, propunând un cadru bazat pe drepturile omului pentru dezvoltarea și utilizarea tehnologiilor bazate pe IA, de către state și actori privați.

## Uniunea Europeană

În anul 2018, Comisia Europeană a înaintat o comunicare Parlamentului European, Consiliului European, Comitetului Economic și Social European și Comitetului Regiunilor, intitulată „Inteligența artificială pentru Europa”<sup>13</sup>. Comunicarea stabilea o inițiativă europeană privind IA cu accent pe „creșterea capacității tehnologice și industriale a UE și utilizarea IA în întreaga economie, atât în sectorul privat, cât și în cel public<sup>14</sup>, pregătirea pentru schimbările socio-economice pe care IA le aduce, și asigurarea unui cadru etic și legal adecvat, bazat pe valorile Uniunii și în conformitate cu Carta drepturilor fundamentale a UE.”

În ceea ce privește asigurarea unui cadru etic, Comisia a constituit Grupul de experți la nivel înalt privind inteligența artificială (AI HLEG) care, în 2019, a publicat *Orientările în materie de etică pentru o inteligență artificială fiabilă*<sup>15</sup>. Potrivit acestui document, o inteligență artificială de încredere presupune respectarea a trei componente, respectiv: „(a) IA ar trebui să fie legală, respectând toate legile și reglementările aplicabile; (b) ar trebui să fie etică, asigurând respectarea principiilor și a valorilor etice și (c) ar trebui să fie solidă, atât din perspectivă tehnică, cât și socială, deoarece, chiar dacă au intenții bune, sistemele IA pot provoca daune neintenționate.”

De asemenea, grupul de experți la nivel înalt a prevăzut și o listă de evaluare cu 7 cerințe cheie pentru a verifica respectarea componentelor menționate mai sus, respectiv: a) implicare și supraveghere umană; b) robustețe tehnică și siguranță; c) respectarea vieții private și governanța datelor; d) transparență; e) diversitate, nediscriminare și echitate; f) bunăstare societală și de mediu; g) responsabilitate. Aceste șapte cerințe și cele trei componente menționate

<sup>12</sup> Disponibil la adresa: <https://undocs.org/A/73/348>

<sup>13</sup> Comisia Europeană, COM (2018) 237 final.

<sup>14</sup> IA poate îmbunătăți în mod semnificativ serviciile publice și poate contribui la îndeplinirea obiectivelor stabilite în Declarația ministerială privind guvernarea electronică – Declarația de la Tallinn (octombrie 2017, <https://ec.europa.eu/digital-single-market/en/news/ministerial-declaration-egovernment-tallinn-declaration>). De exemplu, Comisia va studia potențialul IA de a analiza volume mari de date și de a contribui la verificarea modului în care se aplică regulile pieței unice.

<sup>15</sup> Titlul original al documentului este *Guidelines is to promote Trustworthy AI*. În limba română „trustworthy” a fost tradus cu „fiabil” în documentul la care se face referire. Documentul este disponibil în limbile oficiale UE la adresa: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

anterior au făcut obiectul Comunicării Comisiei către Parlamentul European, Consiliu, Comitetul Economic și Social european și Comitetul Regiunilor – Cum construim încrederea cetățenilor într-o inteligență artificială centrată pe factorul uman, din anul 2019<sup>16</sup>.

În ceea ce privește supravegherea umană, se urmărește ca sistemele de IA să nu submineze autonomia umană și să nu provoace efecte adverse. Potrivit Comunicării Comisiei – Cum construim încrederea cetățenilor într-o inteligență artificială centrată pe factorul uman, este necesar să se asigure măsuri de control, iar supravegherea se poate realiza prin „abordări care implică factorul uman, de exemplu „human-in-the-loop (omul care este implicat în operațiune), human-on-the-loop (omul care monitorizează operațiunea) sau human-in-command (omul care conduce operațiunea).”<sup>17</sup> Robustetea tehnică și siguranța reprezintă un mijloc prin care să se asigure pe de o parte că erorile sistemului IA vor putea fi soluționate, iar pe de cealaltă parte vizează rezistența acestuia la atacuri cibernetice sau la încercări de manipulare a datelor sau algoritmilor utilizați.

Respectarea vieții private și guvernanta datelor reprezintă un element esențial al unei AI de încredere. Întrucât sistemele IA se bazează pe o cantitate mare de date personale din care pot fi deduse preferințe, vârstă, opinii religioase și politice, orice persoană ar trebui să poată controla modul de prelucrare al datelor și ar trebui să i se garanteze că datele care le vizează nu vor fi utilizate în scopuri nelegale sau discriminatorii. În acest sens, este important de menționat că, la nivelul Uniunii, Regulamentul General privind Protecția Datelor<sup>18</sup> permite informarea persoanelor vizate asupra identității operatorului de date și scopurile prelucrării, precum și oferirea de informații suplimentare, pentru a asigura o prelucrare echitabilă și transparentă în ceea ce privește persoanele fizice vizate și dreptul acestora de a li se confirma și comunica datele cu caracter personal care le privesc.

Transparența sistemelor IA presupune ca întregul proces care stă la baza deciziilor luate de aceste sisteme să fie cunoscut, furnizându-se explicații cu privire la proiectarea sistemului și a raționamentelor pe care se bazează aplicarea acestuia. Totodată, utilizatorii ar trebui să fie înștiințați cu privire

<sup>16</sup> COM(2019) 168 final, disponibil la <https://eur-lex.europa.eu/legal-content/RO/TXT/?uri=CELEX:52019DC0168>

<sup>17</sup> Cum construim încrederea cetățenilor într-o inteligență artificială centrată pe factorul uman, Comunicare a Comisiei către Parlamentul European, Consiliu, Comitetul Economic și Social european și Comitetul Regiunilor, COM(2019) 168 final, p. 5.

<sup>18</sup> Regulamentul (UE) 2016/679 al Parlamentului European și al Consiliului din 27 aprilie 2016 privind protecția persoanelor fizice în ceea ce privește prelucrarea datelor cu caracter personal și privind libera circulație a acestor date și de abrogare a Directivei 95/46/CE (Regulamentul general privind protecția datelor), JO L 119, 4.5.2016.

la faptul că interacționează cu un astfel de sistem, și să cunoască cine este persoana responsabilă.

Diversitatea, nediscriminarea și echitatea trebuie să stea la baza sistemelor IA, având în vedere că acestea se bazează pe seturi de date care sunt supuse părtinirii sau subiectivității – fie că este vorba despre subiectivități pre-existente în seturile de date sau care apar ca urmare a modului în care este scris codul de programare al unui algoritm. Astfel, Comisia consideră că este necesar ca echipele de proiectare să fie diverse și să existe mecanisme care asigură participarea la dezvoltarea IA, în special a cetățenilor.

În ceea ce privește bunăstarea societală și de mediu, se are în vedere impactul pe care astfel de sisteme le au asupra mediului înconjurător, specificându-se nevoia de a încuraja durabilitatea și responsabilitate ecologică a sistemelor IA, având în vedere și obiectivele de dezvoltare durabilă ale ONU. Întrucât sistemele IA afectează societatea în ansamblu, utilizarea acestora ar trebui să respecte principiile democratice, inclusiv formarea de opinii sau alte decizii politice.

Un element care ridică de multe ori probleme în ceea ce privește utilizarea sistemelor de IA se referă la responsabilitate și la asumarea răspunderii pentru aceste sisteme și a rezultatelor acestora. În comunicarea Comisiei se subliniază nevoia ca astfel de sisteme să fie supuse auditului, atât intern, cât și extern, mai ales aplicațiile care afectează drepturile fundamentale ale omului. Având în vedere că sistemele IA pot avea un impact negativ, se impune asigurarea unei protecții pentru cei care sunt afectați în mod „(in)direct”.

Se observă că utilizarea sistemelor IA poate avea efecte negative asupra drepturilor fundamentale ale omului, dacă nu se instituie o serie de principii pe care să se bazeze dezvoltarea acestora. Ținând cont că seturile de date pe care se bazează algoritmiile pot fi supuse părtinirii, este posibil să se ajungă la situații discriminatorii. Așadar, asigurarea transparenței procesului de luare a deciziilor de sistemele IA și determinarea persoanei responsabile în cazul unor prejudicii sunt esențiale în vederea asigurării unei IA de încredere.

## Consiliul European

La nivelul Consiliului European au fost elaborate o serie de documente cu privire la sistemele IA printre care se numără un studiu cu privire la responsabilitate și IA<sup>19</sup>, recomandarea „10 pași pentru protejarea drepturilor omului”<sup>20</sup> și studiul „Algoritmi și drepturile omului”.

<sup>19</sup> Comitetul de experți privind dimensiunea drepturilor omului a prelucrării automate de date și alte forme de inteligență artificială, *Responsibility and AI (Responsabilitate și IA)*, Consiliul European, 2019, disponibil la: <https://rm.coe.int/responsability-and-ai-en/168097d9c5>

<sup>20</sup> Comisarul pentru Drepturile Omului al Consiliului European, *Unboxing Artificial Intelligence: 10 steps to protect human rights (Dezvăluirea Inteligenței Artificiale: 10 pași pentru protecția drepturilor omului)*, Consiliul European, 2019.

Recomandarea „10 pași pentru protejarea drepturilor omului”<sup>21</sup> este destinată statelor membre, principiile sale fiind însă valabile pentru orice persoană care influențează – direct sau indirect – dezvoltarea, implementarea sau efectele unui sistem bazat pe IA, respectiv: (1) Evaluarea impactului asupra drepturilor omului; (2) Consultare publică; (3) Obligația statelor membre de a facilita implementarea standardelor de drepturile omului în sectorul privat; (4) Informare și transparență; (5) Monitorizare independentă; (6) Egalitate și nediscriminare; (7) Protecția datelor și confidențialitate; (8) Libertatea de exprimare, de întrunire și asociere și dreptul la muncă; (9) Remedii; (10) Promovarea „alfabetizării IA”. Cele 10 principii au în vedere asigurarea unui cadru legal orientat pe drepturile omului, prevăzându-se implicarea părților interesate și a structurilor pentru drepturile omului în elaborarea și monitorizarea sistemelor IA. De asemenea, se recomandă ca statele membre să se asigure că legislația existentă nu creează bariere în ceea ce privește răspunderea și remediile pentru încălcările de drepturile omului „realizate” de sisteme IA.

În recomandarea Consiliului Europei se remarcă principiul dreptului la exprimare, libertatea de întrunire și asociere și dreptul la muncă. Astfel, se are în vedere ca statele membre să ia în considerare toate standardele internaționale de drepturile omului care pot fi implicate în utilizarea tehnologiilor bazate pe IA. În ceea ce privește libertatea de exprimare se specifică responsabilitatea statelor de a dezvolta un mediu informațional divers și pluralist și de a analiza impactul negativ pe care IA îl poate avea asupra dreptului la exprimare, accesului la informații, având în vedere că există situații în care astfel de tehnologii pot limita fluxul liber de informații.

De asemenea, se pune accent pe libertatea de asociere și întrunire care poate fi afectată de utilizarea pe scară largă a tehnologiilor de recunoaștere facială de către state. Potrivit recomandării Consiliului Europei, asemenea tehnologii trebuie reglementate strict de către statele membre, inclusiv din punct de vedere al limitării utilizării acestora și al transparenței necesare pentru a proteja exercitarea dreptului la liberă întrunire.

În ceea ce privește dreptul la muncă, se intenționează efectuarea de evaluări privind numărul de locuri de muncă create sau dispărute ca urmare a automatizării. În urma unor asemenea evaluări pot fi avute în vedere măsuri de adaptare a curriculumului școlar, pentru a asigura accesul la locuri de muncă ce presupun noi competențe specifice, în contextul sistemelor IA.

La fel ca și în cazul recomandărilor Comisiei Europene, Consiliul Europei se arată preocupat de responsabilitatea și tragerea la răspundere în cazul sistemelor IA, întrucât pot apărea situații de încălcare a drepturilor omului. În

---

<sup>21</sup> Ibidem.



această situație, orice persoană ale cărei drepturi au fost încălcate ca urmare a utilizării IA ar trebui să aibă acces la un remediu efectiv.

În acest sens, poate fi reamintit studiul cu privire la responsabilitate și IA, care face distincție între două aspecte ale responsabilității.<sup>22</sup> Pe de o parte, responsabilitate retrospectivă, care se referă la comportamente sau evenimente care au avut loc în trecut și, pe de altă parte, responsabilitate potențială, care stabilește obligații și îndatoriri asociate rolurilor și sarcinilor din viitor, care este orientată spre obținerea de rezultate bune. Responsabilitatea potențială are funcție de îndrumare, furnizând orientări privind drepturi și obligații față de ceilalți. Astfel, se subliniază că, în vederea stabilirii responsabilității în cazul unor consecințe negative ale utilizării tehnologiilor IA, este important să se acorde atenție celor două tipuri de responsabilitate. În plus, se are în vedere diferențierea conceptului de responsabilitate morală și responsabilitate legală, întrucât în cazul utilizării sistemelor IA există două tipuri de efecte adverse, respectiv încălcarea unor drepturi și prejudicii aduse sănătății umane, proprietății sau mediului. Studiul Consiliului Europei analizează și „controlul” IA, întrucât multe dintre sistemele IA funcționează fără implicarea directă a factorului uman, iar în acest caz se poate vorbi despre existența unui vid de responsabilitate întrucât deciziile ajung să fie luate de sistemele IA fără implicarea oamenilor.

În studiul Consiliului Europei cu privire la algoritmi și drepturile omului se remarcă faptul că dezbateră privind dimensiunea drepturilor omului în contextul algoritmilor tinde să rămână în urma dezvoltărilor tehnologice, fiind necesar ca drepturile și interesele oamenilor fie asigurate cât mai rapid, în concordanță cu prevederile Convenției Europene a Drepturilor Omului și alte tratate internaționale.

### **Raportul raportorului special al ONU privind impactul IA asupra libertății de opinie și exprimare**

Raportul propune o serie de recomandări adresate atât statelor, cât și companiilor, cu privire la utilizarea tehnologiilor IA. În ceea ce privește statele, se prevede obligația acestora de a respecta principiile de drepturile omului, de a se asigura că drepturile și libertățile fundamentale se află în centrul dezvoltării, utilizării și implementării sistemelor de inteligență artificială, inclusiv prin actualizarea reglementărilor în vigoare, în special cele privind protecția datelor, pentru a fi aplicate în materie de IA. Statele ar trebui să prevadă reglementări astfel încât companiile dezvoltatoare să întreprindă

<sup>22</sup>Comitetul de experți privind dimensiunea drepturilor omului a prelucrării automate de date și alte forme de inteligență artificială, *Responsibility and AI (Responsabilitate și IA)*, Consiliul Europei, 2019, disponibil la: <https://rm.coe.int/responsability-and-ai-en/168097d9c5>

evaluări de impact asupra drepturilor omului, asigurând, de asemenea, mecanisme externe pentru asigurarea responsabilității, în cazul încălcării drepturilor. În plus, companiile ar trebui să aibă în vedere consultări publice în ceea ce privește utilizarea și aplicarea sistemelor de IA.

Se are în vedere faptul că IA are efecte asupra mediului informațional din trei puncte de vedere. În primul rând, se pune accent pe afișarea de conținut și personalizare. Practic, în rețelele sociale se acordă atenție deosebită utilizării datelor de navigare pe internet și istoricului căutărilor, pentru a afișa unui utilizator un anumit tip de conținut, determinat de un algoritm, în baza datelor prelucrate. În acest fel, utilizatorii au acces limitat sau nu au deloc acces la anumite informații. De asemenea, ținând cont că afișarea unor rezultate se bazează pe gradul de relevanță al acestora, este posibil ca anumite informații să fie excluse din rezultatele căutărilor.

În al doilea rând, se analizează efectele IA în ceea ce privește crearea de profiluri, reclame și țintirea acestora. Pentru a avea acces la aplicații, conținut și servicii gratuite, utilizatorii pun la dispoziția dezvoltatorilor acestora date personale. Având în vedere cantitatea mare de date personale disponibile, devine facil pentru companii să creeze profiluri și să țintească reclame direct utilizatorilor de pe urma cărora pot profita. Însă, această țintire poate reprezenta o practică discriminatorie. De exemplu, Facebook permite întreprinderilor să utilizeze reclame țintite către anumite grupuri pe baza unor criterii precum vârsta, sexul, locația geografică, interese și alte caracteristici<sup>23</sup>. Astfel, deși reclamele ajung să se adreseze categoriei relevante, există posibilitatea ca unele persoane să fie excluse. Un studiu<sup>24</sup> efectuat de cercetători de la Universitatea Northeastern și Universitatea Southern California arată că, în cazul unui grup de persoane care ar putea fi eligibile să vadă o reclamă, Facebook va alege persoanele în baza unor algoritmi care să maximizeze profitul companiei, uneori afișând reclamele unor audiențe influențate de rasă și gen. De exemplu, Facebook a afișat reclame din industria prelucrătoare a lemnului unei audiențe formate din aproximativ 70% persoane albe, și 90% bărbați, și reclame pentru funcția de casier unei audiențe formate din 85 % femei. Similar, alte studii arată că unele reclame limitează vizibilitatea pentru persoane dintr-o anumită categorie de vârstă, și astfel lucrătorii mai în vârstă nu le pot vedea.<sup>25</sup>

<sup>23</sup> Facebook, Țintirea reclamelor, <https://ro-ro.facebook.com/business/ads/ad-targeting>

<sup>24</sup> Ali M. et al., *Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes (Discriminare prin optimizare: Modalități prin care afișarea de reclame Facebook duce la rezultate distorsionate)*, 2019, disponibil la: <https://arxiv.org/pdf/1904.02095.pdf>

<sup>25</sup> Angwin J., Scheiber N., Tobin A., *Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads*, ProPublica Dec. 20, 2017, <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>

În al treilea rând se analizează moderarea și eliminarea (ștergerea) conținutului. Tehnologiile bazate pe IA ajută companiile de rețele sociale să modereze conținutul afișat pe platformele lor, în baza reglementărilor și standardelor stabilite de acestea, printre care se pot număra: postările cu conținut violent, cele care aduc atingere anumitor drepturi ale omului, mesaje instigatoare la ură, abuz sexual și altele. Însă, moderarea conținutului prin intermediul tehnologiilor IA prezintă anumite riscuri având în vedere că există limitări privind evaluarea contextului sau a particularităților lingvistice sau culturale. Acest lucru presupune că poate fi eliminat conținut sau pot fi suspendate conturi de utilizator care nu sunt problematice.

Raportorul Special atrage atenția că instrumentele IA trebuie proiectate, dezvoltate și implementate astfel încât să respecte obligațiile statelor și responsabilitățile actorilor privați aferente dreptului internațional al drepturilor omului. În ceea ce privește dreptul la opinie<sup>26</sup> s-a remarcat faptul că noile tehnologii pot influența dreptul la opinie. Având în vedere că se promovează un anumit tip de conținut, capacitatea unei persoane de a-și forma opinii informate poate scădea. În același timp, dreptul la liberă exprimare presupune și obligația de a promova diversitatea mass-media, independența și protecția accesului la informație. Având în vedere că sistemele IA nu pot detecta întotdeauna elemente de ironie sau sarcasm, se poate elimina conținut de acest tip, ceea ce poate fi interpretat ca o formă de cenzură.

Raportul privind impactul IA asupra libertății de opinie și exprimare abordează și dreptul la viață privată. Orice ingerință la adresa vieții private trebuie să respecte standardele internaționale în materie de legalitate, necesitate și proporționalitate<sup>27</sup>. Având în vedere că tehnologiile IA se bazează pe seturi de date și creează altele noi, se reduce capacitatea indivizilor de a cunoaște, înțelege și controla modul de utilizare a datelor lor. Raportorul Special atrage atenția că, dacă anumite seturi de date sunt scoase din contextul original, crește riscul ca acele date să devină nereale, iar persoanele nu pot avea posibilitatea de a le rectifica sau șterge.

<sup>26</sup> Art. 19, Pactul internațional cu privire la drepturile politice și civile, „1. Nimeni nu trebuie să aibă de suferit din cauza opiniilor sale. 2. Orice persoană are dreptul la libertatea de exprimare; acest drept cuprinde libertatea de a căuta, de a primi și de a răspândi informații și idei de orice fel, indiferent de frontiere, sub formă orală, scrisă, tipărită ori artistică sau prin orice alt mijloc, la alegerea sa. 3. Exercitarea libertăților prevăzute la paragraful 2 al prezentului articol comportă obligații și răspunderi speciale. În consecință, ea poate fi supusă anumitor limitări care trebuie însă stabilite în mod expres prin lege și care sunt necesare; a) respectării drepturilor sau reputației altora; b) apărării securității naționale, ordinii publice, sănătății sau moralității publice.”

<sup>27</sup> Oficiul Înaltului Comisar pentru Drepturile Omului, A/HRC/27/37, para. 23 și rezoluția Comitetului pentru drepturile omului 34/7, para. 2.

În același timp, se subliniază importanța abordării bazate pe drepturile omului în domeniul inteligenței artificiale. Atât companiile, cât și statele, trebuie să se asigure că respectarea drepturilor și libertăților fundamentale stă la baza dezvoltării și implementării tehnologiilor de IA. Se recomandă ca statele să ia în considerare efectuarea de consultări publice înainte de a implementa orice tehnologii IA, în vederea evaluării impactului acestora asupra titularilor de drepturi, cu accent pe minorități, activiști politici. Obligația de a evalua impactul tehnologiilor IA revine și companiilor care ar trebui să aibă în vedere organizarea de consultări publice încă din momentul proiectării sistemelor IA.

## Concluzii

Sistemele de inteligență artificială au un potențial ridicat în sprijinirea drepturilor și libertăților fundamentale, însă prezintă o serie de riscuri care trebuie reglementate în vederea asigurării dreptului la viață privată și la nediscriminare.

Tehnologiile bazate pe inteligență artificială pot avea efecte benefice asupra unor aspecte ale vieții, precum educație, muncă, asistență socială, sănătate și aplicarea legii, însă pot avea și un impact negativ asupra grupurilor vulnerabile, precum refugiați, solicitanți de azil și persoane vârstnice. În mod special, prin utilizarea sistemelor bazate pe IA, apar riscuri cu privire la realizarea dreptului la egalitate și nediscriminare, precum și a dreptului la viață privată.

Dreptul la viață privată<sup>28</sup> este unul dintre cele mai dezbătute subiecte în cazul sistemelor IA. Fiecare persoană are dreptul la respectarea vieții private și nu poate fi supusă unei imixțiuni arbitrare și nelegale în viața particulară, în familia, domiciliul sau corespondența sa, și nu se pot aduce atingeri ilegale onoarei și reputației sale. Cu toate acestea, tehnologiile moderne permit colectarea și prelucrarea unor seturi de date de către state și companii private. Asemenea informații pot ajuta la crearea unor imagini de ansamblu cu privire la interesele și preferințele unei persoane, stilul de viață și legăturile sociale<sup>29</sup>.

Reglementările în materie de IA ar trebui să țină seama de faptul că, de multe ori, tehnologia avansează mai repede decât dezbaterile privind apărarea drepturilor omului, fiind necesar să se stabilească cât mai rapid un cadru general în vederea limitării sau reducerii impactului negativ.

<sup>28</sup> Dreptul la viață privată este prevăzut de art. 12 al Declarației Universale a Drepturilor Omului, de art. 17 al Pactului internațional cu privire la drepturile politice și civile, dar și de Convenția europeană a drepturilor omului la art. 8

<sup>29</sup> McGregor L., Ng V., Shaheed A. et. al, *The Universal Declaration of Human Rights at 70 – Putting Human Rights at the Heart of the Design, Development, and Deployment of Artificial Intelligence*, The Human Rights, Big Data and Technology Project, 2018

## **Bibliografie:**

- Ali M. et al., *Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes (Discriminare prin optimizare: Modalități prin care afișarea de reclame Facebook duce la rezultate distorsionate)*, 2019;
- Angwin J., Scheiber N., Tobin A., *Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads*, ProPublica Dec. 20, 2017, <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>;
- Comisarul pentru Drepturile Omului al Consiliului European, *Unboxing Artificial Intelligence: 10 steps to protect human rights (Dezvăluirea Inteligenței Artificiale: 10 pași pentru protecția drepturilor omului)*, Consiliul European, 2019;
- Comisia Europeană, Cum construim încrederea cetățenilor într-o inteligență artificială centrată pe factorul uman, Comunicare a Comisiei către Parlamentul European, Consiliu, Comitetul Economic și Social european și Comitetul Regiunilor, COM(2019) 168 final;
- Comitetul de experți privind dimensiunea drepturilor omului a prelucrării automate de date și alte forme de inteligență artificială, *Responsibility and AI (Responsabilitate și IA)*, Consiliul European, 2019;
- Comitetul de experți privind intermediarii (serviciilor) de internet (MSI-NET), *Algorithms and human rights – Study on the human rights dimensions of automated data processing techniques and possible regulatory implications (Algoritmi și drepturile omului – Studiu privind dimensiunea drepturilor omului a tehnicilor de prelucrare automate de date și posibile implicații de reglementare)*, Consiliul European, 2017;
- Facebook, Țintirea reclamelor, <https://ro-ro.facebook.com/business/ads/ad-targeting>;
- Grupul de experți la nivel înalt privind IA la UE, *Guidelines is to promote Trustworthy AI (Orientări în materie de etică pentru o inteligență artificială fiabilă)*, 2019;
- Grupul de experți la nivel înalt privind IA la UE, *O definiție a inteligenței artificiale (IA): principalele capacități și discipline științifice*, 2019;
- Istoria inteligenței artificiale*, Universitatea Washington, disponibil la: <https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf>;
- McGregor L., Ng V., Shaheed A. et. al, *The Universal Declaration of Human Rights at 70 – Putting Human Rights at the Heart of the Design, Development, and Deployment of Artificial Intelligence*, The Human Rights, Big Data

- and Technology Project, 2018 [https://ling2s14id7e20wtc8xsceyr-wpengine.netdna-ssl.com/wp-content/uploads/2018/12/UDHR70\\_AI.pdf](https://ling2s14id7e20wtc8xsceyr-wpengine.netdna-ssl.com/wp-content/uploads/2018/12/UDHR70_AI.pdf);
- Carte albă – Inteligența artificială – O abordare europeană axată pe excelență și încredere, Comisia Europeană, COM(2020) 65 final;
- Comunicare a Comisiei Către Parlamentul European, Consiliul European, Consiliu, Comitetul Economic și Social European și Comitetul Regiunilor Inteligența artificială pentru Europa, Comisia Europeană, COM (2018) 237 final;
- Declarația Universală a Drepturilor Omului;
- Orientările ONU privind întreprinderile și drepturile omului, 2011, HR/PUB/11/04;
- Pactul internațional cu privire la drepturile politice și civile;
- Raportul Raportorului Special al ONU privind impactul IA asupra libertății de opinie și exprimare, 2018, A/73/348;
- Regulamentul (UE) 2016/679 al Parlamentului European și al Consiliului din 27 aprilie 2016 privind protecția persoanelor fizice în ceea ce privește prelucrarea datelor cu caracter personal și privind libera circulație a acestor date și de abrogare a Directivei 95/46/CE (Regulamentul general privind protecția datelor), JO L 119, 4.5.2016;