



Bruxelles, 8.4.2019
COM(2019) 168 final

**COMUNICARE A 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**

1. INTRODUCERE — STRATEGIA EUROPEANĂ ÎN DOMENIUL INTELIGENȚEI ARTIFICIALE (IA)

Inteligența artificială (IA) are potențialul de a ne transforma lumea în bine: poate îmbunătăți asistența medicală, poate reduce consumul de energie, poate spori siguranța traficului rutier și le poate permite fermierilor să folosească mai eficient apa și resursele naturale. IA poate fi folosită pentru a prevedea schimbările de mediu și climatice, pentru a îmbunătăți gestionarea riscurilor financiare și oferă instrumentele necesare pentru a fabrica, cu mai puține deșeuri, produse adaptate nevoilor noastre. De asemenea, IA poate contribui la detectarea fraudelor și a amenințărilor pentru securitatea cibernetică și permite autorităților de aplicare a legii să combată infracțiunile într-un mod mai eficient.

IA poate fi utilă întregii societăți și economiei. Este o tehnologie strategică care se dezvoltă și se folosește într-un ritm din ce în ce mai rapid în întreaga lume. Însă IA aduce cu sine noi provocări pentru viitorul muncii și ridică probleme de natură juridică și etică.

Pentru a face față acestor provocări și pentru a valorifica la maximum oportunitățile oferite de IA, Comisia a publicat o strategie europeană¹ în aprilie 2018. Această strategie plasează oamenii în centrul dezvoltării IA — **IA centrată pe factorul uman**. Strategia este formată din trei axe, al căror scop este acela de a îmbunătăți capacitatea tehnologică și industrială a UE și implementarea IA în economie, de a asigura pregătirea pentru schimbările socioeconomice și de a asigura un cadru etic și juridic adecvat.

Pentru a asigura implementarea strategiei privind IA **Comisia a elaborat, împreună cu statele membre, un plan coordonat în ceea ce privește IA²**, pe care l-a prezentat în decembrie 2018, pentru a se crea sinergii, a se pune în comun date - materiile prime pentru numeroase aplicații IA - și a se spori investițiile comune. Scopul este acela de a stimula cooperarea transfrontalieră și de a mobiliza toți jucătorii să sporească investițiile publice și private la **cel puțin 20 de miliarde EUR** anual pe parcursul următorului deceniu³. Comisia și-a dublat investițiile în IA în cadrul programului Orizont 2020 și intenționează să investească anual 1 miliard EUR din programele Orizont Europa și Europa digitală, sprijinind în special spațiile comune ale datelor în domeniul sănătății, al transporturilor și al producției, dar și marile unități de experimentare, precum spitalele inteligente și infrastructura pentru vehicule automate, și o agendă strategică în materie de cercetare.

Pentru a pune în aplicare o astfel de agendă strategică comună în materie de cercetare, inovare și implementare, Comisia și-a identificat **dialogul cu toate părțile interesate relevante** de la nivelul industriei, al institutelor de cercetare și al autorităților publice. Noul program Europa digitală va juca, de asemenea, un rol esențial pentru a pune IA la dispoziția întreprinderilor mici și mijlocii din toate statele membre, prin centre de inovare digitală, instalații consolidate de testare și experimentare, spații ale datelor și programe de formare.

Bazându-se pe reputația sa în ceea ce privește produsele sigure și de calitate superioară, abordarea etică a Europei privind IA consolidează încrederea cetățenilor în dezvoltarea digitală și vizează crearea unui avantaj competitiv pentru întreprinderile europene din domeniul IA. Scopul prezentei comunicări este să lanseze o etapă-pilot cuprinzătoare, care implică părțile interesate pe o scară cât mai largă, pentru a testa punerea practică în aplicare a orientărilor etice pentru dezvoltarea și utilizarea IA.

¹ COM(2018) 237.

² COM(2018) 795.

³ Pentru a contribui la atingerea acestui obiectiv, Comisia a propus, în cadrul următoarei perioade de programare 2021-2027, ca Uniunea să aloce o finanțare de cel puțin 1 miliard EUR pe an din programele Orizont Europa și Europa digitală pentru investiții în IA.

2. CONSOLIDAREA ÎNCREDERII ÎN IA CENTRATĂ PE FACTORUL UMAN

Strategia europeană și planul coordonat privind IA clarifică faptul că **încrederea este o condiție prealabilă pentru asigurarea unei abordări centrate pe oameni în materie de IA**. IA nu reprezintă un scop în sine, aceasta este un instrument care trebuie să deservească populația, având drept obiectiv final creșterea bunăstării umane. În acest scop, **ar trebui asigurată încrederea în IA**. Valorile pe care sunt bazate societățile noastre trebuie să fie integrate pe deplin în modul în care se dezvoltă IA.

Uniunea se întemeiază pe **valorile respectării demnității umane, libertății, democrației, egalității, și statului de drept, precum și pe respectarea drepturilor omului**, inclusiv a drepturilor persoanelor care aparțin minorităților⁴. Aceste valori sunt comune societăților tuturor statelor membre, în care prevalează pluralismul, nediscriminarea, toleranța, justiția, solidaritatea și egalitatea. În plus, **Carta drepturilor fundamentale a UE** reunește într-un singur text toate drepturile personale, civice, politice, economice și sociale de care beneficiază persoanele în cadrul UE.

UE are un **cadru solid de reglementare** care va stabili standardul global pentru IA centrată pe factorul uman. Regulamentul general privind protecția datelor asigură un standard ridicat de protecție a datelor cu caracter personal și impune punerea în aplicare de măsuri pentru asigurarea protecției datelor din faza de proiectare și în mod implicit⁵. Regulamentul privind libera circulație a datelor fără caracter personal în Uniunea Europeană elimină barierele din calea liberei circulații a datelor fără caracter personal și asigură prelucrarea tuturor categoriilor de date, peste tot în Europa. Legea privind securitatea cibernetică, adoptată recent, va contribui la consolidarea încrederii în lumea online, iar propunerea de regulament privind respectarea vieții private în comunicațiile electronice⁶ vizează același obiectiv.

Cu toate acestea, IA aduce noi provocări pentru că le permite echipamentelor să „învețe”, să ia și să pună în aplicare decizii fără intervenția umană. În curând, acest tip de funcționalitate va deveni standard pentru numeroase produse și servicii, de la telefoane inteligente la mașini automate, roboți și aplicații online. Cu toate acestea, deciziile luate pe baza unor algoritmi ar putea rezulta din date care sunt incomplete și, prin urmare, care nu sunt fiabile, și ar putea fi manipulate prin atacuri cibernetice, ar putea fi subiective sau pur și simplu ar putea fi greșite. Aplicarea fără discernământ a tehnologiei pe măsură ce aceasta se dezvoltă ar conduce, așadar, la rezultate problematice, precum și la reticența cetățenilor de a o accepta sau a o folosi.

Tehnologia IA ar trebui dezvoltată însă într-o manieră care să fie centrată pe factorul uman și care, astfel, să merite încrederea populației. Aceasta înseamnă că aplicațiile IA nu ar trebui să fie doar conforme cu legea, ci și să respecte principiile etice și să garanteze că punerea acestora în aplicare nu provoacă daune nedorite. Diversitatea din punctul de vedere al genului, al originii rasiale sau etnice, al religiei sau al convingerilor, al handicapului și al vârstei ar

⁴ Totodată, UE este parte a Convenției Organizației Națiunilor Unite privind drepturile persoanelor cu handicap.

⁵ Regulamentul (UE) 2016/679. Regulamentul general privind protecția datelor (RGPD) garantează libera circulație a datelor cu caracter personal în interiorul Uniunii. Acesta conține dispoziții privind luarea deciziilor bazate exclusiv pe prelucrarea automată, inclusiv realizarea de profiluri. Persoanele vizate au dreptul de a fi informate cu privire la existența unui proces decizional automat și de a primi informații pertinente referitoare la logica procesului decizional automat și la semnificația și la consecințele preconizate ale prelucrării pentru acestea. De asemenea, ele au dreptul, în astfel de cazuri, să obțină intervenția umană, să își exprime punctul de vedere și să conteste decizia.

⁶ COM(2017) 10.

trebui asigurată în fiecare fază a dezvoltării IA. Aplicațiile IA ar trebui să ofere cetățenilor mijloace pentru a acționa și ar trebui să respecte drepturile fundamentale ale acestora. Scopul lor ar trebui să fie acela de a spori abilitățile persoanelor, nu de a le înlocui, și de a permite totodată accesul persoanelor cu handicap.

Prin urmare, este nevoie de **orientări în materie de etică**, care să se bazeze pe cadrul de reglementare existent și care ar trebui să fie aplicate de dezvoltatorii, furnizorii și utilizatorii de IA pe piața internă, care să stabilească condiții etice pentru desfășurarea activității în toate statele membre. Din acest motiv, Comisia a instituit un **grup de experți la nivel înalt privind IA**⁷, care reprezintă o gamă largă de părți interesate, și a încredințat acestui grup misiunea de a elabora orientări în materie de etică în ceea ce privește IA, precum și de a pregăti un set de recomandări pentru o politică mai amplă referitoare la IA. Totodată, a fost înființată **Alianța europeană în domeniul inteligenței artificiale**⁸, o platformă multilaterală deschisă, cu peste 2 700 de membri, care să ofere o gamă mai variată de observații pentru activitatea grupului de experți la nivel înalt privind IA.

Grupul de experți la nivel înalt privind IA a publicat o primă versiune a orientărilor în materie de etică în decembrie 2018. În urma unei **consultări cu părțile interesate**⁹ și a **întâlnirilor cu reprezentanții statelor membre**¹⁰, grupul de experți privind IA a transmis Comisiei un document revizuit în martie 2019. Până în prezent, în feedback-ul lor, părțile interesate au salutat, per ansamblu, caracterul practic al orientărilor și îndrumările practice pe care le oferă dezvoltatorilor, furnizorilor și utilizatorilor de IA în ceea ce privește asigurarea încrederii.

2.1. Orientări privind o IA de încredere, elaborate de grupul de experți la nivel înalt privind IA

Orientările elaborate de grupul de experți la nivel înalt privind IA la care se referă prezenta comunicare¹¹ valorifică în special activitatea Grupului european pentru etică în domeniul științei și al noilor tehnologii și a Agenției pentru Drepturi Fundamentale a Uniunii Europene.

Orientările prevăd că, pentru o „IA de încredere”, sunt necesare trei componente: (1) aceasta ar trebui să respecte legea; (2) ar trebui să respecte principiile etice și (3) ar trebui să fie solidă.

Pe baza acestor trei componente și a valorilor europene prevăzute în secțiunea 2, orientările identifică șapte cerințe-cheie pe care aplicațiile IA ar trebui să le respecte pentru a fi considerate de încredere. Orientările includ și o listă de evaluare, care poate fi folosită pentru a verifica îndeplinirea acestor cerințe.

Cele șapte cerințe-cheie sunt:

⁷ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

⁸ <https://ec.europa.eu/digital-single-market/en/european-ai-alliance>

⁹ Această consultare a condus la transmiterea de observații din partea a 511 organizații, asociații, societăți, institute de cercetare, persoane fizice și alte entități. O sinteză a feedback-ului primit este disponibilă la următoarea adresă:

https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Activitatea grupului de experți a primit aprecieri pozitive din partea statelor membre, iar la 18 februarie 2019 au fost adoptate concluziile Consiliului în care s-a consemnat, printre altele, viitoarea publicare a orientărilor în materie de etică și s-a exprimat susținerea pentru efortul Comisiei de a introduce o abordare etică a UE pe scena mondială: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/ro/pdf>

¹¹ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

- implicare și supraveghere umană;
- robustețe tehnică și siguranță;
- respectarea vieții private și guvernanta datelor;
- transparență;
- diversitate, nediscriminare și echitate;
- bunăstare societală și de mediu;
- responsabilitate.

Deși aceste cerințe sunt concepute pentru a se aplica tuturor sistemelor de IA în contexte și industrii diferite, pentru punerea lor în aplicare într-o manieră concretă și proporțională ar trebui să se țină cont de contextul specific de aplicare, adoptându-se o abordare ce are în vedere impactul. De exemplu, o aplicație IA care sugerează, drept recomandare de lectură, o carte inadecvată este mult mai puțin periculoasă decât stabilirea în mod eronat a unui diagnostic de cancer și, prin urmare, ar putea face obiectul unei supravegheri mai puțin stringente.

Orientările elaborate de grupul de experți la nivel înalt privind IA nu au un caracter obligatoriu și, prin urmare, nu creează nicio nouă obligație juridică. Totuși, numeroase dispoziții existente (și deseori cu utilizare specifică sau specifice unui anumit domeniu) ale dreptului Uniunii reflectă deja, desigur, una sau mai multe dintre aceste cerințe-cheie, de exemplu normele privind siguranța, protecția datelor cu caracter personal, respectarea vieții private sau protecția mediului.

Comisia salută activitatea grupului de experți la nivel înalt privind IA și o consideră valoroasă pentru procesul său de elaborare a politicilor.

2.2. Cerințe-cheie pentru o IA de încredere

Comisia sprijină următoarele cerințe-cheie pentru o IA de încredere, care sunt bazate pe valori europene. Aceasta încurajează părțile interesate să aplice cerințele și să testeze lista de evaluare care le operaționalizează, pentru a crea un mediu adecvat de încredere pentru dezvoltarea și utilizarea cu succes a IA. Comisia salută feedback-ul primit de la părțile interesate pentru a evalua în ce măsură această listă de evaluare prevăzută în orientări necesită ajustări suplimentare.

I. Implicarea și supravegherea umană

Sistemele IA ar trebui să ajute persoanele să ia decizii mai bune și mai bine informate, în conformitate cu propriile obiective. Acestea ar trebui să contribuie la crearea unei societăți înfloritoare și echitabile, sprijinind implicarea umană și **drepturile fundamentale**, fără a reduce, a limita sau a orienta eronat autonomia umană. **Bunăstarea generală a utilizatorului** ar trebui se afle în centrul funcționării sistemului.

Supravegherea umană contribuie la asigurarea faptului că sistemele IA nu subminează autonomia umană și nu provoacă alte efecte adverse. În funcție de sistemul specific bazat pe IA și de domeniul său de aplicare, ar trebui asigurate nivelurile corespunzătoare ale **măsurilor de control**, inclusiv adaptabilitatea, acuratețea și explicitatea sistemelor bazate

pe IA¹². **Supravegherea** se poate realiza prin mecanisme de guvernanta, precum asigurarea unei abordări care implică factorul uman, de tipul *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).¹³ Trebuie să se garanteze faptul că autoritățile publice au capacitatea de a-și exercita atribuțiile de supraveghere în conformitate cu mandatele lor. Dacă toate celelalte condiții rămân neschimbate, cu cât o ființă umană poate exercita mai puțină supraveghere asupra unui sistem IA, cu atât este nevoie de teste mai extinse și de o guvernanta mai strictă.

II. Robustețe tehnică și siguranță

O IA de încredere necesită algoritmi suficient de siguri, de fiabili și de robusti pentru a soluționa erorile sau consecvențele în toate fazele ciclului de viață al sistemului IA și pentru a aborda în mod corespunzător rezultatele eronate. Sistemele IA trebuie să fie **fiabile**, suficient de sigure pentru a fi **reziliente** la atacurile deschise și încercările mai subtile de a manipula datele sau chiar algoritmii și trebuie să asigure un **plan de rezervă** în caz de probleme. Deciziile lor trebuie să fie **exacte** sau cel puțin să reflecte în mod corect nivelul de acuratețe, iar rezultatele lor ar trebui să fie **reproductibile**.

În plus, sistemele IA ar trebui să integreze mecanisme de siguranță și de securitate încă din faza de concepere, astfel încât **siguranța lor să poată fi verificată** în fiecare etapă, punându-se accentul pe siguranța fizică și psihică a tuturor celor implicați. Aceasta include reducerea la minimum și, pe cât posibil, reversibilitatea consecințelor neintenționate sau a erorilor în funcționarea sistemului. Ar trebui instituite procese de clarificare și de evaluare a riscurilor potențiale asociate utilizării sistemelor IA, în diferite domenii de aplicare.

III. Respectarea vieții private și guvernanta datelor

Respectarea vieții private și **protecția datelor** trebuie garantate în **toate fazele** ciclului de viață al sistemului IA. Înregistrările digitale ale comportamentului uman pot permite sistemelor IA să deducă nu numai preferințele, vârsta și sexul persoanelor, ci și orientarea sexuală, opiniile religioase sau politice. Pentru ca persoanele să poată avea încredere în prelucrarea datelor, acestora trebuie să li se garanteze controlul deplin asupra propriilor date și faptul că datele care le vizează nu vor fi folosite pentru a le face rău sau pentru a le discrimina.

Pe lângă protejarea vieții private și a datelor cu caracter personal, trebuie îndeplinite cerințe pentru asigurarea unei calități ridicate a sistemelor IA. Calitatea seturilor de date utilizate este fundamentală pentru performanța sistemelor IA. Atunci când sunt colectate

¹² Regulamentul general privind protecția datelor conferă persoanelor dreptul de a nu face obiectul unei decizii bazate exclusiv pe prelucrarea automată atunci când aceasta produce efecte juridice asupra utilizatorilor sau îi afectează în mod similar într-o măsură semnificativă (articolul 22 din RGPD).

¹³ Abordările de tipul *human-in-the-loop* (HITL) se referă la intervenția umană în fiecare ciclu decizional al sistemului, care, în numeroase situații, nu este nici posibilă, nici de dorit. Abordarea de tipul *human-on-the-loop* (HOTL) se referă la capacitatea de intervenție umană în ciclul de proiectare a sistemului și la monitorizarea funcționării sistemului. Abordarea de tipul *human-in-command* (HIC) se referă la capacitatea de a supraveghea activitatea generală a sistemului IA (inclusiv a impactului său economic, societal, juridic și etic mai amplu) și la capacitatea de a decide când și cum se folosește sistemul în orice situație specifică. Aici poate intra decizia de a nu folosi un sistem IA într-o situație specifică, de a stabili niveluri de marje de apreciere umană pe perioada de utilizare a sistemului sau de a asigura capacitatea de a lua o decizie care să o înlocuiască pe cea luată de sistem.

date, acestea pot reflecta subiectivități determinate social sau pot conține inexactități, erori și greșeli. Aceste nevoi trebuie abordate înainte de programarea unui sistem IA pe baza unui anumit set de date. În plus, trebuie asigurată **integritatea** datelor. Procesele și seturile de date folosite trebuie să fie testate și documentate în fiecare fază, de exemplu, în fazele de planificare, programare, testare și implementare. Acest lucru ar trebui să se aplice și în cazul sistemelor IA care nu au fost dezvoltate intern, ci dobândite din altă parte. În fine, **accesul** la date trebuie să fie gestionat și controlat în mod adecvat.

IV. Transparență

Ar trebui asigurată **trasabilitatea** sistemelor IA; este important să fie consemnate și documentate atât deciziile luate de sisteme, cât și întregul proces (inclusiv o descriere a colectării și etichetării datelor și o descriere a algoritmului folosit) care a stat la baza deciziilor. În acest sens, ar trebui asigurată pe cât posibil **explicitatea** procesului decizional algoritmic, adaptat persoanelor implicate. Ar trebui avute în vedere cercetările în curs pentru dezvoltarea unor mecanisme de asigurare a explicității. În plus, ar trebui furnizate explicații privind măsura în care un sistem IA influențează și modelează procesul decizional organizațional, opțiunile în ceea ce privește proiectarea sistemului, precum și raționamentele care stau la baza implementării acestuia (asigurându-se, astfel, nu numai transparența datelor și a sistemului, ci și transparența modelului de afaceri).

În final, este important ca limitările și capacitățile sistemului IA să fie **comunicate** în mod adecvat diferitelor părți interesate implicate, într-o manieră corespunzătoare situației date. În plus, sistemele IA ar trebui să fie identificabile ca atare, astfel încât utilizatorii să știe că interacționează cu un sistem IA și să știe care sunt persoanele responsabile pentru acesta.

V. Diversitate, nediscriminare și echitate

Seturile de date folosite de sistemele IA (atât pentru programare, cât și pentru funcționare) pot include subiectivități istorice neintenționate, date incomplete și modele de guvernare deficitară. Perpetuarea acestor subiectivități ar putea conduce la discriminare (in)directă. De asemenea, exploatarea intenționată a subiectivităților (consumatorilor) sau implicarea în activități de concurență neloială pot avea efecte nedorite. În plus, și modul în care sunt dezvoltate sistemele IA (de exemplu, modul în care este scris codul de programare al unui algoritm) poate fi afectat de subiectivități. Astfel de preocupări ar trebui abordate încă de la începutul dezvoltării sistemului.

Crearea unor **echipe de proiectare diverse** și a unor mecanisme care asigură **participarea** la dezvoltarea IA, în special a cetățenilor, pot contribui la abordarea acestor preocupări. Se recomandă consultarea părților interesate care pot fi afectate în mod direct sau indirect de sistem pe parcursul ciclului său de viață. Sistemele IA ar trebui să ia în considerare întreaga gamă de abilități, competențe și cerințe umane și să asigure accesibilitatea printr-o abordare de proiectare universală, pentru a asigura acces egal pentru persoanele cu handicap.

VI. Bunăstare societală și de mediu

Pentru ca IA să fie de încredere, ar trebui luat în considerare impactul său asupra **mediului înconjurător și asupra altor ființe sensibile**. În mod ideal, toți oamenii, inclusiv generațiile viitoare, ar trebui să beneficieze de biodiversitate și de un mediu

locuibil. Așadar, ar trebui încurajate durabilitatea și **responsabilitatea ecologică** a sistemelor IA. Același lucru este valabil și pentru soluțiile IA care abordează domenii ce reprezintă preocupări globale, cum ar fi, de exemplu, obiectivele de dezvoltare durabilă ale ONU.

În plus, impactul sistemelor IA ar trebui luat în considerare nu numai din perspectivă individuală, ci și din perspectiva **societății în ansamblu**. Utilizarea sistemelor IA ar trebui luată în considerare în mod corespunzător, în special în situații ce au legătură cu procesul democratic, inclusiv cu formarea de opinii, cu deciziile politice sau în contexte electorale. În plus, ar trebui luat în considerare **impactul social** al IA. Deși sistemele IA pot fi utilizate pentru dezvoltarea competențelor sociale, acestea pot contribui și la deteriorarea lor.

VII. Responsabilitate

Ar trebui instituite mecanisme pentru asigurarea responsabilității și a tragerii la răspundere pentru sistemele IA și rezultatele acestora, atât înainte de implementarea lor, cât și ulterior. **Posibilitatea auditării** sistemelor IA este fundamentală în acest sens, întrucât evaluarea sistemelor IA de către auditori interni și externi și disponibilitatea unor astfel de rapoarte de evaluare contribuie în mod semnificativ la încrederea în tehnologie. Posibilitatea auditării externe ar trebui asigurată în special în cazul aplicațiilor care afectează drepturile fundamentale, inclusiv al aplicațiilor critice din punctul de vedere al siguranței.

Ar trebui identificate, evaluate, documentate și reduse la minimum **efectele negative potențiale** ale sistemelor IA. Utilizarea evaluărilor de impact facilitează acest proces. Evaluările ar trebui să fie proporționale cu amploarea riscurilor pe care le comportă sistemele IA. **Compromisurile** între cerințe – care sunt deseori inevitabile – ar trebui abordate într-o manieră rațională și metodologică și ar trebui justificate. În final, atunci când se produc efecte adverse injuste, ar trebui prevăzute mecanisme accesibile, care să asigure **măsuri reparatorii adecvate**.

2.3. Etapele următoare: o etapă-pilot care implică părțile interesate pe o scară cât mai largă

Obținerea unui consens cu privire la aceste cerințe-cheie pentru sistemele IA reprezintă o primă etapă importantă în vederea elaborării orientărilor privind o IA etică. Într-o etapă următoare, Comisia va garanta că aceste orientări pot fi testate și puse în aplicare în practică.

În acest scop, ea va lansa acum o etapă-pilot specifică, al cărei scop este acela de a se obține feedback structurat de la părțile interesate. Acest exercițiu se va axa în special pe lista de evaluare elaborată de grupul de experți la nivel înalt pentru fiecare dintre cerințele-cheie.

Activitatea se va desfășura pe două paliere: (i) o etapă-pilot pentru orientări, care implică părți interesate ce dezvoltă sau folosesc IA, inclusiv administrațiile publice și (ii) un proces permanent de consultare cu părțile interesate și de sensibilizare a acestora în toate statele membre și la nivelul diferitelor grupuri de părți interesate, inclusiv în sectorul industriei și al serviciilor:

- (i) începând cu iunie 2019, toate părțile interesate și persoanele vor fi invitate să testeze lista de evaluare și să furnizeze feedback privind îmbunătățirea acesteia. În plus, grupul de experți la nivel înalt privind IA va organiza o revizuire aprofundată,

împreună cu părți interesate din sectorul public sau privat, pentru a colecta feedback mai detaliat cu privire la modalitatea de punere în aplicare a orientărilor într-o gamă variată de domenii. Toate feedback-urile privind fezabilitatea și funcționalitatea orientărilor vor fi evaluate până la finalul anului 2019;

- (ii) în paralel, Comisia va organiza activități suplimentare de informare, oferindu-le reprezentanților grupului de experți la nivel înalt privind IA posibilitatea de a prezenta aceste orientări părților interesate relevante din statele membre, inclusiv din sectorul industriei și al serviciilor, precum și posibilitatea suplimentară de a formula observații referitoare la orientările în materie de IA și de a contribui la acestea.

Comisia va ține cont de activitatea grupului de experți în materie de etică pentru mobilitatea conectată și automatizată¹⁴ și va colabora cu proiecte de cercetare finanțate de UE în domeniul IA și cu parteneriate public-private relevante la punerea în aplicare a cerințelor-cheie¹⁵. De exemplu, Comisia va sprijini, în coordonare cu statele membre, dezvoltarea unei baze de date comune cu imagistică de sănătate, dedicată inițial celor mai frecvente tipuri de cancer, astfel încât algoritmi să poată fi programați să diagnosticheze simptome cu o acuratețe foarte mare. În mod similar, cooperarea dintre Comisie și statele membre permite creșterea numărului de coridoare transfrontaliere pentru testarea vehiculelor conectate și automatizate. Orientările ar trebui aplicate și testate în cadrul acestor proiecte, iar rezultatele vor fi folosite în procesul de evaluare.

Etapa-pilot și consultarea părților interesate vor beneficia de contribuția Alianței europene în domeniul inteligenței artificiale și a AI4EU, platforma pentru IA la cerere. Proiectul AI4EU¹⁶, lansat în ianuarie 2019, reunește algoritmi, instrumente, seturi de date și servicii pentru a ajuta organizațiile, mai ales întreprinderile mici și mijlocii, să pună în aplicare soluții IA. Alianța europeană în domeniul IA, împreună cu AI4EU, va mobiliza în continuare ecosistemul IA la nivelul Europei, inclusiv în vederea îndrumării privind aplicarea orientărilor în materie de etică a IA și a promovării respectului pentru IA centrată pe factorul uman.

La începutul anului 2020, valorificând evaluarea feedback-ului primit în etapa-pilot, grupul de experți la nivel înalt privind IA va revizui și va actualiza orientările. Pe baza revizuirii și a experienței dobândite, **Comisia va evalua rezultatul și va propune eventualele etape următoare.**

IA etică reprezintă o propunere reciproc avantajoasă. Garantarea respectului pentru valorile și drepturile fundamentale nu este numai o misiune esențială în sine, dar aceasta facilitează, de asemenea, acceptarea de către public și sporește avantajul competitiv al întreprinderilor europene în domeniul IA, creând o marcă a inteligenței artificiale de încredere și centrate pe factorul uman, cunoscută pentru produse etice și sigure. Aceasta îmbunătățește, per ansamblu, reputația întreprinderilor europene în ceea ce privește furnizarea de produse sigure, de înaltă calitate. Etapa-pilot va contribui la asigurarea faptului că produsele IA respectă această promisiune.

¹⁴ A se vedea Comunicarea Comisiei privind mobilitatea conectată și automatizată, COM(2018) 283.

¹⁵ În cadrul Fondului european de apărare, Comisia va elabora și îndrumări etice specifice pentru evaluarea propunerilor de proiecte în domeniul IA pentru apărare.

¹⁶ <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-ai4eu-project-launches-1-january-2019>

2.4. Către orientări internaționale în materie de etică în domeniul IA

Discuțiile internaționale privind etica în domeniul IA s-au intensificat după ce Președinția G-7 deținută de Japonia a plasat acest subiect ca prioritate pe agenda 2016. Având în vedere interdependențele internaționale ale dezvoltării IA în ceea ce privește circulația datelor, dezvoltarea algoritmilor și investițiile în cercetare, **Comisia își va continua eforturile de promovare a abordării Uniunii pe scena mondială și de obținere a unui consens în ceea ce privește o IA centrată pe factorul uman**¹⁷.

Activitatea desfășurată de grupul de experți la nivel înalt privind IA și, mai exact, lista de cerințe și procesul de implicare a părților interesate, furnizează Comisiei observații suplimentare valoroase pentru contribuția la discuțiile internaționale. Uniunea Europeană poate juca un rol de lider în dezvoltarea de orientări internaționale în materie de IA și, dacă este posibil, a unui mecanism de evaluare aferent.

Prin urmare, Comisia:

Va consolida cooperarea cu parteneri care au aceeași viziune:

- explorând măsura în care se poate realiza convergența cu proiectele de orientări în materie de etică ale țărilor terțe (de exemplu, Japonia, Canada, Singapore) și, valorificând existența acestui grup de țări cu o viziune similară, pregătind discuții mai ample, sprijinite de acțiuni de punere în aplicare a instrumentului de parteneriat pentru cooperarea cu țările terțe¹⁸; și
- explorând în ce fel pot contribui întreprinderile din țări din afara UE și organizațiile internaționale la „etapa-pilot” a orientărilor, prin testare și validare.

Va juca în continuare un rol activ în cadrul discuțiilor și inițiativelor la nivel internațional:

- contribuind la foruri multilaterale precum G-7 și G-20;
- implicându-se în dialoguri cu țări din afara UE și organizând întâlniri bilaterale și multilaterale pentru a obține un consens cu privire la o IA centrată pe factorul uman;
- contribuind la activități relevante de standardizare în cadrul organizațiilor internaționale de elaborare de standarde, pentru promovarea acestei viziuni; și
- consolidând colectarea și diseminarea de opinii privind politicile publice, prin colaborare cu organizații internaționale relevante.

¹⁷ Înaltul Reprezentant al Uniunii pentru afaceri externe și politica de securitate, cu sprijinul Comisiei, se va baza pe consultările din cadrul Organizației Națiunilor Unite, al Global Tech Panel și al altor foruri multilaterale și va coordona, în special, propunerile care vizează abordarea provocărilor complexe în materie de securitate pe care le implică IA.

¹⁸ Regulamentul (UE) nr. 234/2014 al Parlamentului European și al Consiliului din 11 martie 2014 de instituire a unui instrument de parteneriat pentru cooperarea cu țările terțe (JO L 77, 15.3.2014, p. 77). De exemplu, proiectul planificat privind „O alianță internațională pentru o abordare centrată pe factorul uman a inteligenței artificiale” va facilita inițiativele comune cu parteneri cu o viziune similară, pentru a promova orientări în materie de etică și a adopta principii comune și concluzii operaționale. Acesta va permite ca UE și țările cu o viziune similară să discute concluziile operaționale ce rezultă din orientările în materie de etică în domeniul IA propuse de grupul de experți, pentru a ajunge la o abordare comună. În plus, va permite monitorizarea adoptării tehnologiei IA la nivel global. În final, proiectul prevede organizarea de activități de diplomatie publică, în marja unor evenimente internaționale, desfășurate, de exemplu, de către G-7, G-20 și Organizația pentru Cooperare și Dezvoltare Economică.

3. CONCLUZII

UE se întemeiază pe un set de valori fundamentale și, pe baza acestora, și-a construit un cadru de reglementare solid și echilibrat. Pornind de la acest cadru de reglementare existent, este nevoie de orientări în materie de etică pentru dezvoltarea și utilizarea IA, având în vedere caracterul său de noutate și provocările specifice pe care le comportă această tehnologie. IA poate fi considerată de încredere doar dacă este dezvoltată și utilizată într-o manieră ce respectă valori etice împărtășite la nivel larg.

În vederea atingerii acestui obiectiv, Comisia consideră binevenite informațiile furnizate de grupul de experți la nivel înalt privind IA. Pe baza cerințelor-cheie referitoare la încrederea în IA, Comisia va lansa acum o etapă-pilot specifică, pentru a garanta posibilitatea de a pune în aplicare în practică orientările în materie de etică obținute pentru dezvoltarea și utilizarea IA. Comisia va acționa, de asemenea, în vederea atingerii unui consens societal amplu în ceea ce privește IA centrată pe factorul uman, inclusiv cu toate părțile interesate implicate și cu partenerii săi internaționali.

Dimensiunea etică a IA nu reprezintă un element de lux sau o opțiune suplimentară: ea trebuie să facă parte integrantă din dezvoltarea IA. Depunând eforturi pentru o IA centrată pe factorul uman și bazată pe încredere, protejăm respectul pentru valorile noastre societale de bază și creăm o marcă distinctivă pentru Europa și industria sa, în calitate de lider în domeniul unei IA de vârf și de încredere la nivel mondial.

Pentru a asigura dezvoltarea etică a IA în Europa într-un context mai amplu, Comisia adoptă o abordare cuprinzătoare, care include în special următoarele linii de acțiune ce urmează să fie puse în aplicare până în al treilea trimestru al anului 2019:

- va lansa o serie de **rețele de centre de cercetare de excelență în domeniul IA** prin intermediul programului Orizont 2020; va selecta până la patru rețele, axându-se pe provocări științifice sau tehnologice majore, precum explicitatea și interacțiunea avansată om-mașină, ce reprezintă ingrediente-cheie pentru o IA de încredere;
- va începe să creeze **rețele de centre de inovare digitală**¹⁹, punând accentul pe IA în producție și pe volumele mari de date;
- împreună cu statele membre și cu părțile interesate, Comisia va iniția discuții pregătitoare pentru dezvoltarea și punerea în aplicare a **unui model pentru partajarea de date și pentru valorificarea optimă a spațiilor comune ale datelor**, punând accentul în special pe transporturi, asistența medicală și producția industrială²⁰.

În plus, Comisia lucrează la un raport privind provocările asociate IA în ceea ce privește cadrele de siguranță și de răspundere și la un document cu orientări referitoare la punerea în aplicare a Directivei privind răspunderea pentru produsele cu defect²¹. Totodată, întreprinderea comună pentru calculul european de înaltă performanță (EuroHPC)²² va dezvolta următoarea generație de supercalculatoare deoarece capacitatea de calcul este esențială pentru prelucrarea datelor și pentru programarea IA, iar Europa trebuie să controleze întregul lanț valoric digital. Actualul parteneriat cu statele membre și cu industria în ceea ce

¹⁹ <http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs>

²⁰ Resursele necesare vor fi mobilizate prin Orizont 2020 (în cadrul căruia aproape 1,5 miliarde EUR sunt dedicate IA pentru perioada 2018-2020) și din programul prevăzut a-i succede, Orizont Europa, din componenta digitală a Mecanismului pentru interconectarea Europei și, în, special din viitorul program Europa digitală. Proiectele vor folosi și resurse din sectorul privat și din programele statelor membre.

²¹ A se vedea Comunicarea Comisiei „Inteligența artificială pentru Europa”, COM (2018) 237.

²² <https://eurohpc-ju.europa.eu>

privește componentele și sistemele microelectronice (ECSEL)²³ și Inițiativa europeană privind procesoarele²⁴ vor contribui la dezvoltarea tehnologiei pentru procesoarele cu un consum scăzut de energie, pentru a realiza calcule de înaltă performanță, fiabile și sigure.

La fel ca activitatea privind orientările în materie de etică în domeniul IA, toate aceste inițiative au la bază **cooperarea strânsă dintre toate părțile interesate implicate**, statele membre, industrie, actorii societali și cetățeni. Per ansamblu, abordarea Europei în ceea ce privește inteligența artificială ilustrează modul în care competitivitatea economică și încrederea societală trebuie să pornească de la aceleași valori fundamentale și să se consolideze reciproc.

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu