



Microcredențial de IA etică

CARTEA

CU5 | Drepturile omului și echitatea

Numărul proiectului:
2022-1-ES01-KA220-HED-000085257



Cum să utilizați acest Flipbook?

Acest document este interactiv. De-a lungul documentului, veți găsi linkuri către informații suplimentare.



Buton care vă duce la începutul documentului. Această pictogramă apare în colțul din dreapta sus al paginilor.



Ori de câte ori vedeți această săgeată, înseamnă că aveți un **text color interactiv** pe care trebuie să faceți clic, care are asociat un link extern.

DECLINARE DE RESPONSABILITATE: Vă rugăm să rețineți că nu putem garanta disponibilitatea continuă a conținutului extern, cum ar fi videoclipurile, deoarece acestea pot fi modificate sau eliminate de către autorii sau platformele gazdă.

Index

Faceți clic pe meniu

01. Introducere

**02. Relevanța drepturilor omului și a echității
în sistemele AI**

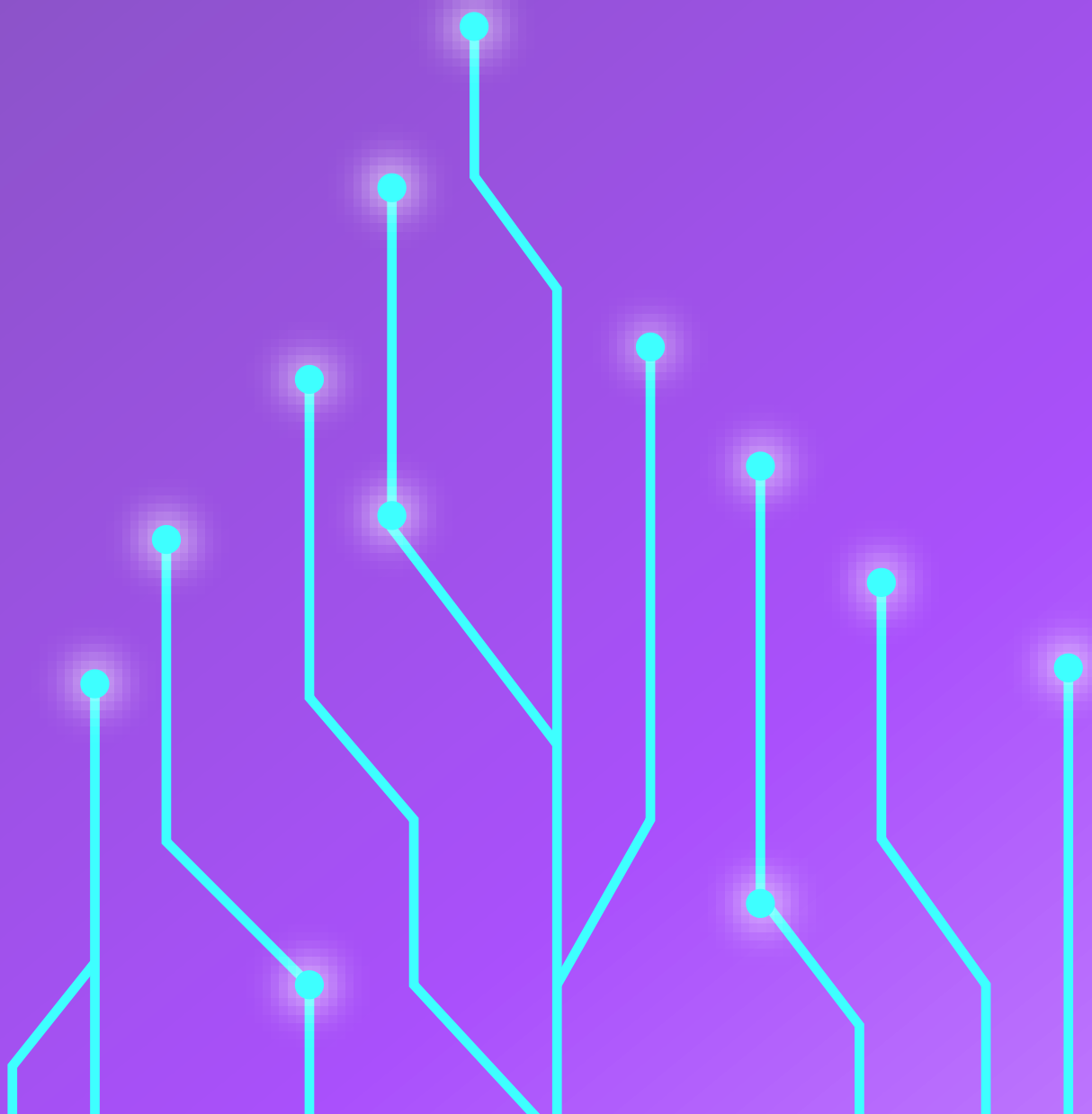
**03. Intersecția drepturilor omului
și echitatea algoritmică**

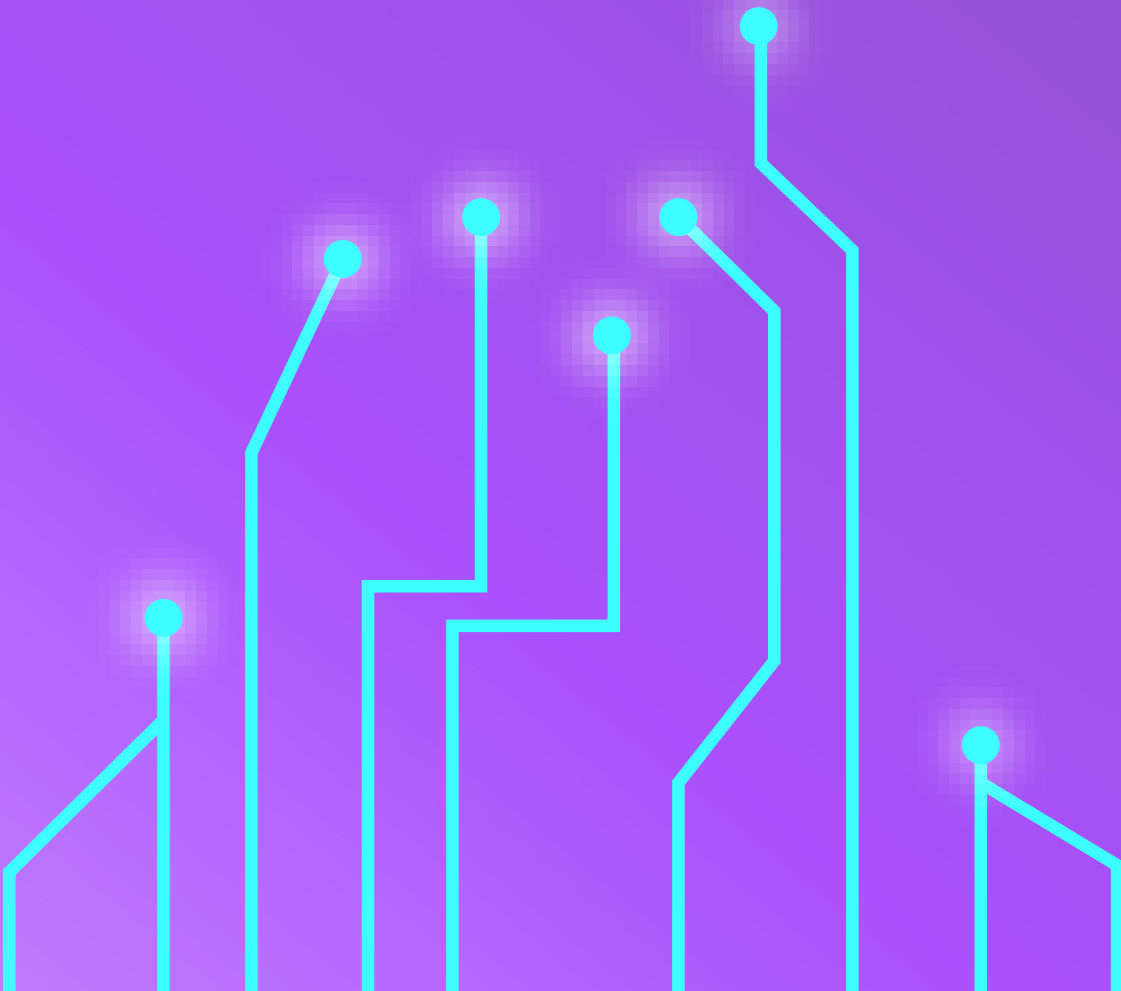
04. Principii de echitate în sistemele AI

05. Concluzii

01. Introducere

CU5 | Drepturile omului și echitatea





01. Introducere

În această unitate de competență, cursanții vor dobândi cunoștințe cuprinzătoare privind rolul vital al drepturilor omului și al echității în domeniul sistemelor AI. Această explorare va aprofunda conceptele fundamentale, subliniind intersecția dintre părtinirea algoritmică și drepturile omului, precum și principiile de echitate inerente sistemelor AI. Prin înțelegerea acestor concepte, cursanții vor aprecia implicațiile din lumea reală și valoarea principiilor corectitudinii, echității și justiției în atenuarea prejudecăților algoritmice și încurajarea unor rezultate mai echitabile, cu atenția cuvenită pentru interesele generațiilor viitoare.

Rezultatele acestui curs cuprind:

- **Relevanța drepturilor omului și a corectitudinii în sistemele de inteligență artificială** pentru a promova rezultate echitabile și prevenirea daunelor în cadrul dezvoltării și implementării inteligenței artificiale. Vom prezenta beneficiile și riscurile potențiale asociate cu tehnologiile IA, recunoscând necesitatea de a aborda aspecte precum disparitatea socio-economică, încălcarea vieții private și erodarea autonomiei.



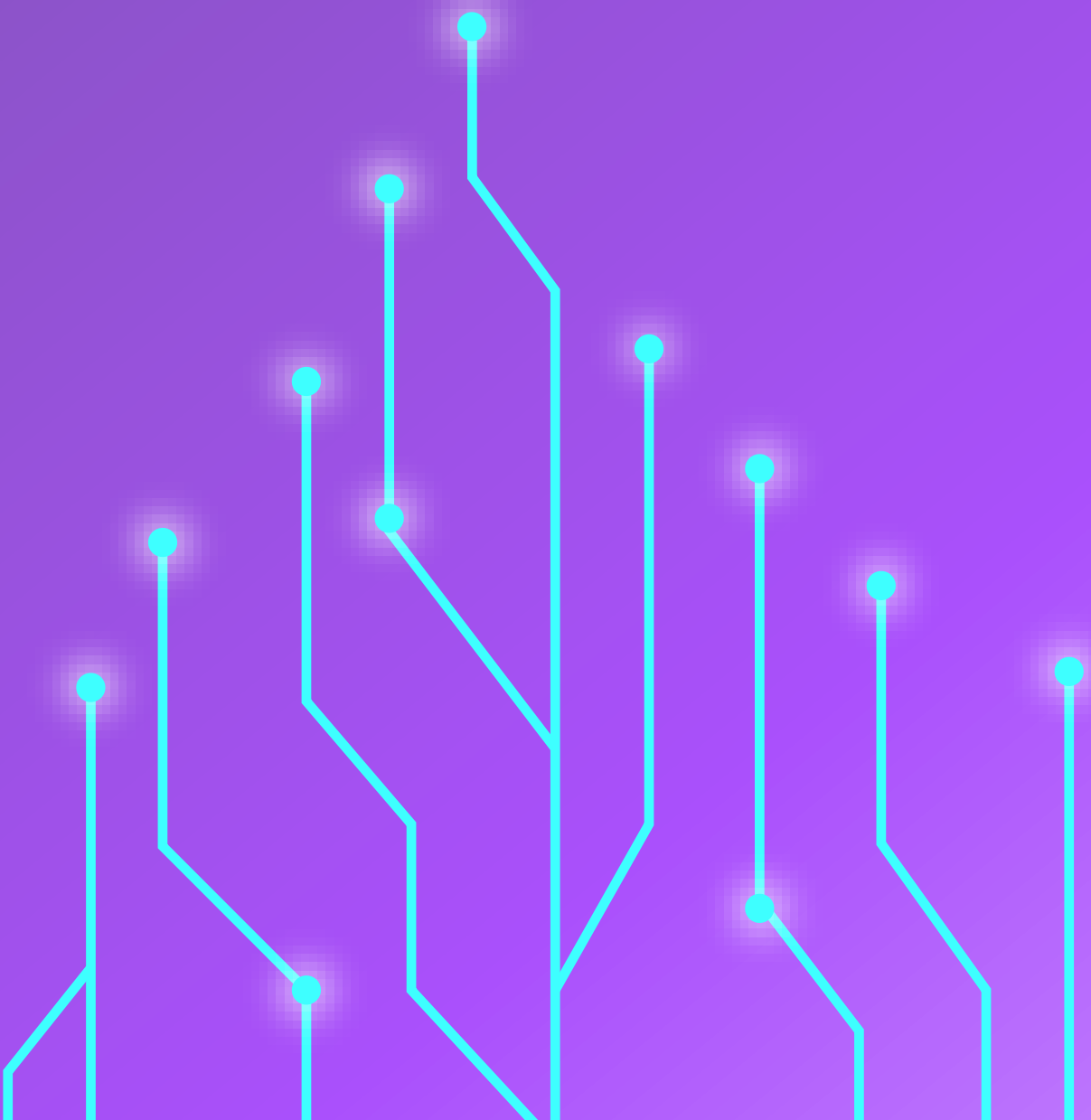


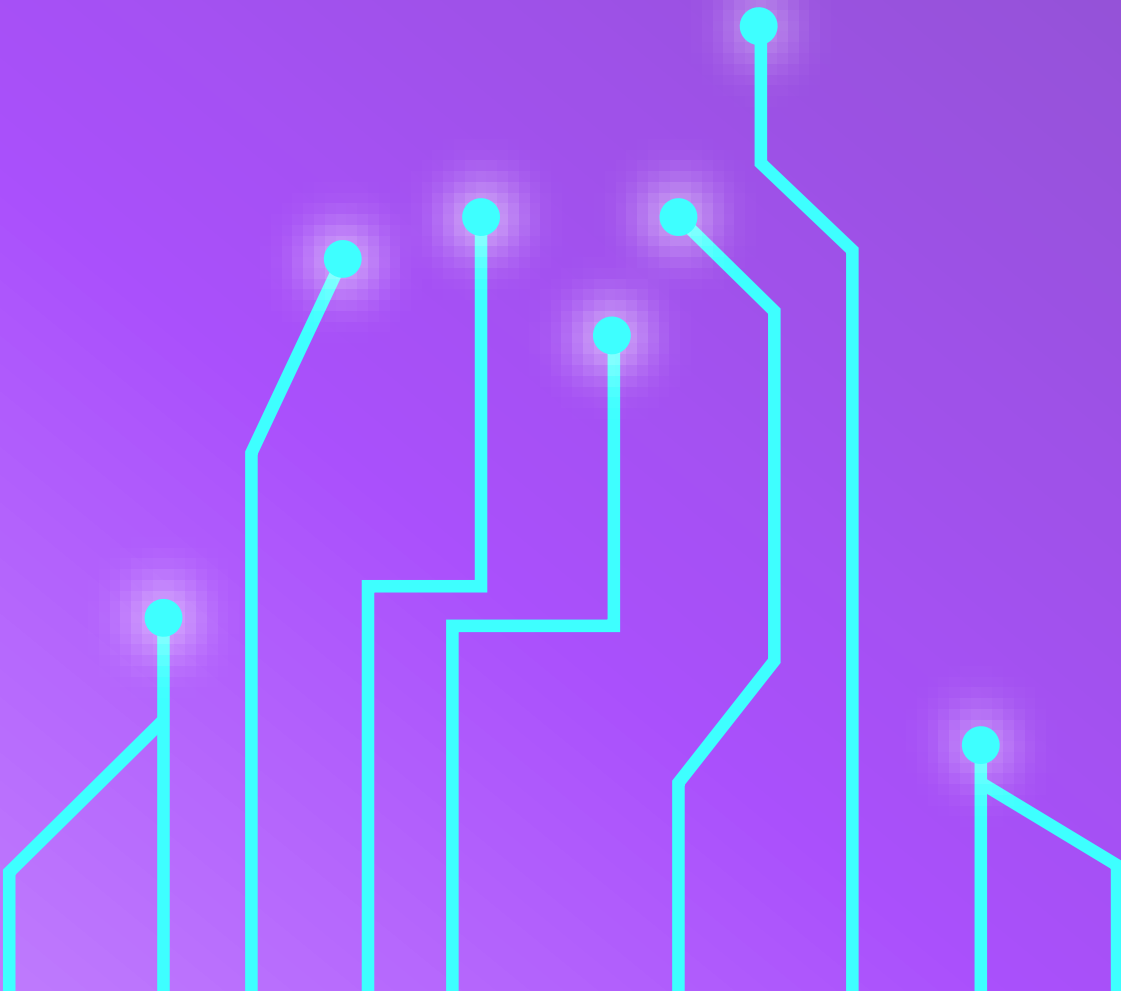
- **Intersecția dintre drepturile omului și echitatea algoritmică:**
Această unitate fundamentală va introduce cursanții la intersecția critică dintre drepturile omului și echitatea algoritmică, bazându-se pe cadre teoretice pentru a analiza implicațiile etice ale tehnologiei. Prin studierea evoluției istorice și a dezbaterilor contemporane, cursanții vor înțelege modul în care procesele algoritmice afectează drepturile omului și vor explora studii de caz care ilustrează conflictele dintre tehnologiile algoritmice și drepturile omului.
- **Principiile corectitudinii în sistemele AI**, inclusiv egalitatea de șanse, nediscriminarea, corectitudinea procedurală, echitatea și justiția. Recunoscând importanța acestor principii în modelarea unor rezultate echitabile pentru generațiile viitoare, cursanții vor fi echipați să facă față provocărilor etice și să pledeze pentru dezvoltarea și implementarea responsabilă a IA.

În ultimii ani, integrarea sistemelor de inteligență artificială (AI) în diferite aspecte ale societății a generat preocupări etice semnificative privind protecția drepturilor omului și promovarea echității. Această unitate își propune să aprofundeze semnificația critică a drepturilor omului și a echității în sistemele de inteligență artificială și rolul lor esențial în promovarea unor rezultate echitabile, prevenind în același timp daunele.

02. Relevanța drepturilor omului și a echității în sistemele AI

CU5 | Drepturile omului și echitatea





02. Relevanța drepturilor omului și a echității în sistemele AI

Importanța drepturilor omului și a corectitudinii în sistemele AI nu poate fi supraestimată.

Drepturile omului, astfel cum sunt consacrate în diverse declarații și convenții internaționale, reprezintă piatra de temelie a guvernancei etice și a bunăstării societății. În contextul sistemelor AI, aplicarea principiilor drepturilor omului asigură respectarea și protejarea demnității, autonomiei și vieții private a persoanelor. În plus, corectitudinea sistemelor de inteligență artificială este esențială pentru promovarea unor rezultate echitabile și atenuarea perpetuării prejudecăților și inegalităților existente în cadrul societății.

Sistemele de inteligență artificială au potențialul de a afecta diferite aspecte ale vieții umane, de la oportunitățile de angajare și accesul la servicii până la administrarea justiției și protecția libertăților civile. Prin urmare, protejarea drepturilor omului și promovarea echității în dezvoltarea și implementarea AI sunt imperative pentru prevenirea daunelor și respectarea standardelor etice. Prin integrarea drepturilor omului și a principiilor echității în sistemele de inteligență artificială, dezvoltatorii pot reduce riscul practicilor discriminatorii, al încălcării vieții private și al rezultatelor nedrepte, promovând astfel încrederea și responsabilitatea în rândul părților interesate.



Aplicarea principiilor drepturilor omului și echității în contextul dezvoltării și implementării inteligenței artificiale **necesită o abordare multidisciplinară care să ia în considerare implicațiile juridice, etice și sociale**. Dezvoltatorii trebuie să adere la cadrele consacrate ale drepturilor omului, cum ar fi Declarația Universală a Drepturilor Omului și Pactul internațional cu privire la drepturile civile și politice, pentru a se asigura că sistemele de IA respectă drepturile și libertățile fundamentale.

Un aspect esențial al aplicării drepturilor omului și a principiilor echității în dezvoltarea inteligenței artificiale este asigurarea transparenței și a responsabilității pe tot parcursul ciclului de viață al inteligenței artificiale. Transparența permite părților interesate să înțeleagă modul în care funcționează sistemele AI, datele pe care le utilizează și procesele decizionale implicate. Prin furnizarea de documente și explicații clare, dezvoltatorii pot permite utilizatorilor să evalueze corectitudinea și fiabilitatea sistemelor de inteligență artificială și să îi tragă la răspundere pe cei responsabili de dezvoltarea și implementarea acestora.



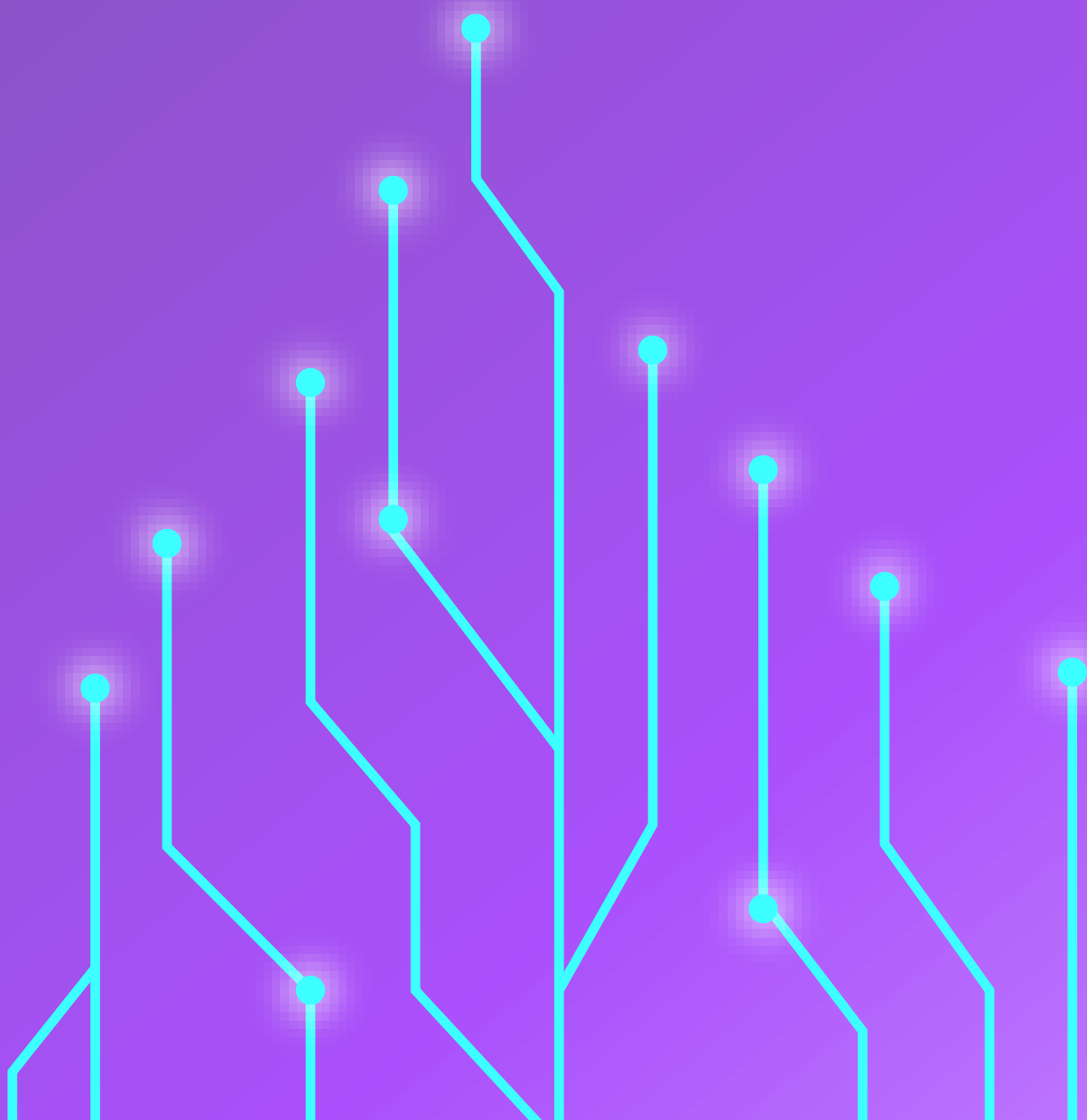
În plus, echitatea în dezvoltarea inteligenței artificiale necesită o analiză atentă a strategiilor de atenuare a prejudecăților și a mecanismelor algoritmice de responsabilizare. Dezvoltatorii trebuie să identifice și să abordeze în mod proactiv prejudecățile din datele de formare, proiectarea algoritmilor și procesele decizionale pentru a preveni rezultatele discriminatorii. În plus, punerea în aplicare a mecanismelor de supraveghere și despăgubire garantează că persoanele afectate de sistemele de inteligență artificială au posibilitatea de a recurge în caz de tratament nedrept sau de prejudiciu.

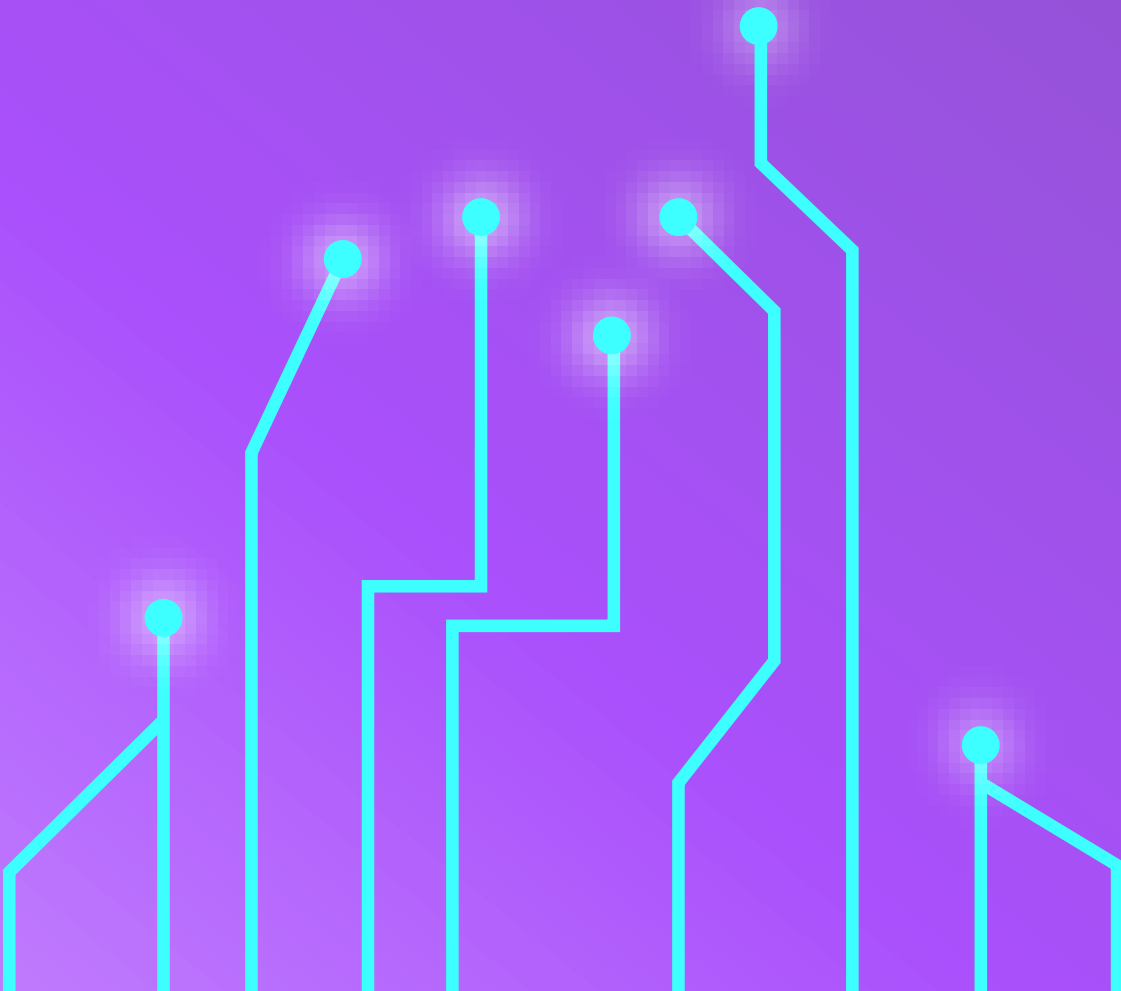




03. Intersecția dintre drepturile omului și echitate algoritmică

CU5 | Drepturile omului și echitatea





03. Intersecția dintre drepturile omului și echitate algoritmică

Relația dintre părtinirea algoritmică și drepturile omului este o problemă cu multiple fațete.

Relația dintre prejudecățile algoritmice și drepturile omului este o problemă complexă și multifacetată care necesită o examinare atentă. În această unitate, cursanții vor pătrunde în complexitatea acestei relații, recunoscând impactul potențial asupra populațiilor vulnerabile și perpetuarea inegalităților. Prin înțelegerea dinamicii în joc, cursanții pot identifica exemple din lumea reală de sisteme AI părtinitoare care afectează drepturile omului și pot elabora strategii pentru a aborda aceste probleme în mod eficient.

Prejudecata algoritmică se referă la preferințele sau prejudecățile sistematice și nedrepte care pot fi prezente în datele, algoritmiile sau procesele decizionale ale sistemelor AI. Atunci când nu sunt controlate, prejudecățile algoritmice pot avea implicații profunde asupra drepturilor omului, în special pentru populațiile vulnerabile precum minoritățile rasiale, femeile, persoanele în vârstă și persoanele cu handicap.

Prin perpetuarea inegalităților existente și consolidarea practicilor discriminatorii, **sistemele AI părtinitoare pot submina principiile fundamentale ale drepturilor omului, precum egalitatea, nediscriminarea și dreptul la viață privată.**



Impactul prejudecăților algoritmice asupra drepturilor omului se poate manifesta în diverse moduri în diferite sectoare și contexte. De exemplu, algoritmiile părtinitori utilizați în procesele de angajare pot duce la practici discriminatorii care refuză persoanelor șanse egale la angajare pe baza unor factori irelevanți precum rasa, sexul sau vârsta. În mod similar, în sistemul de justiție penală, algoritmiile predictivi părtinitori ai poliției pot viza în mod disproporționat comunitățile marginalizate, ducând la arestări nejustificate și la încălcări ale dreptului la un proces echitabil și la un proces corect.

În plus, prejudecățile algoritmice pot exacerba disparitățile existente în ceea ce privește accesul la servicii esențiale precum asistența medicală, locuințele și educația. Sistemele AI părtinitoare utilizate în procesele de acordare a creditelor sau de aprobare a împrumuturilor pot dezavantaja în mod sistematic anumite grupuri demografice, perpetuând inegalitățile economice și împiedicând capacitatea persoanelor de a accesa resurse financiare și oportunități de promovare socio-economică.

Recunoașterea impactului potențial al prejudecăților algoritmice asupra drepturilor omului este esențială pentru protejarea drepturilor și demnității tuturor persoanelor, în special a celor care aparțin comunităților marginalizate sau vulnerabile. Analizând relația dintre prejudecățile algoritmice și drepturile omului, cursanții dobândesc o înțelegere mai profundă a implicațiilor etice ale tehnologiilor AI și a necesității unor măsuri proactive pentru a aborda prejudecățile și a promova corectitudinea și echitatea.

➤ Exemple din lumea reală de sisteme AI părtinitoare care afectează drepturile omului și strategii de abordare a acestor probleme

Exemplele din lumea reală de sisteme AI părtinitoare care afectează drepturile omului abundă, subliniind nevoia urgentă de acțiune pentru abordarea acestor probleme. **De la sistemele discriminatorii de recunoaștere facială la algoritmi predictivi părtinitori utilizați în condamnarea penală, sistemele AI părtinitoare au potențialul de a încălca drepturile persoanelor și de a perpetua inegalitățile sistemice.**

De exemplu, în contextul aplicării legii, s-a constatat că algoritmi predictivi părtinitori ai poliției vizează în mod disproporționat comunitățile minoritare, conducând la o supraveghere sporită, la arestări abuzive și la încălcări ale dreptului persoanelor la viață privată și la libertatea de a nu fi deținute arbitrar.

În mod similar, în domeniul asistenței medicale, algoritmi părtinitori utilizați în diagnosticarea medicală și planificarea tratamentului pot duce la diagnosticarea eronată sau la o îngrijire inadecvată pentru anumite grupuri demografice, exacerband disparitățile în materie de sănătate și subminând dreptul persoanelor la sănătate și bunăstare.





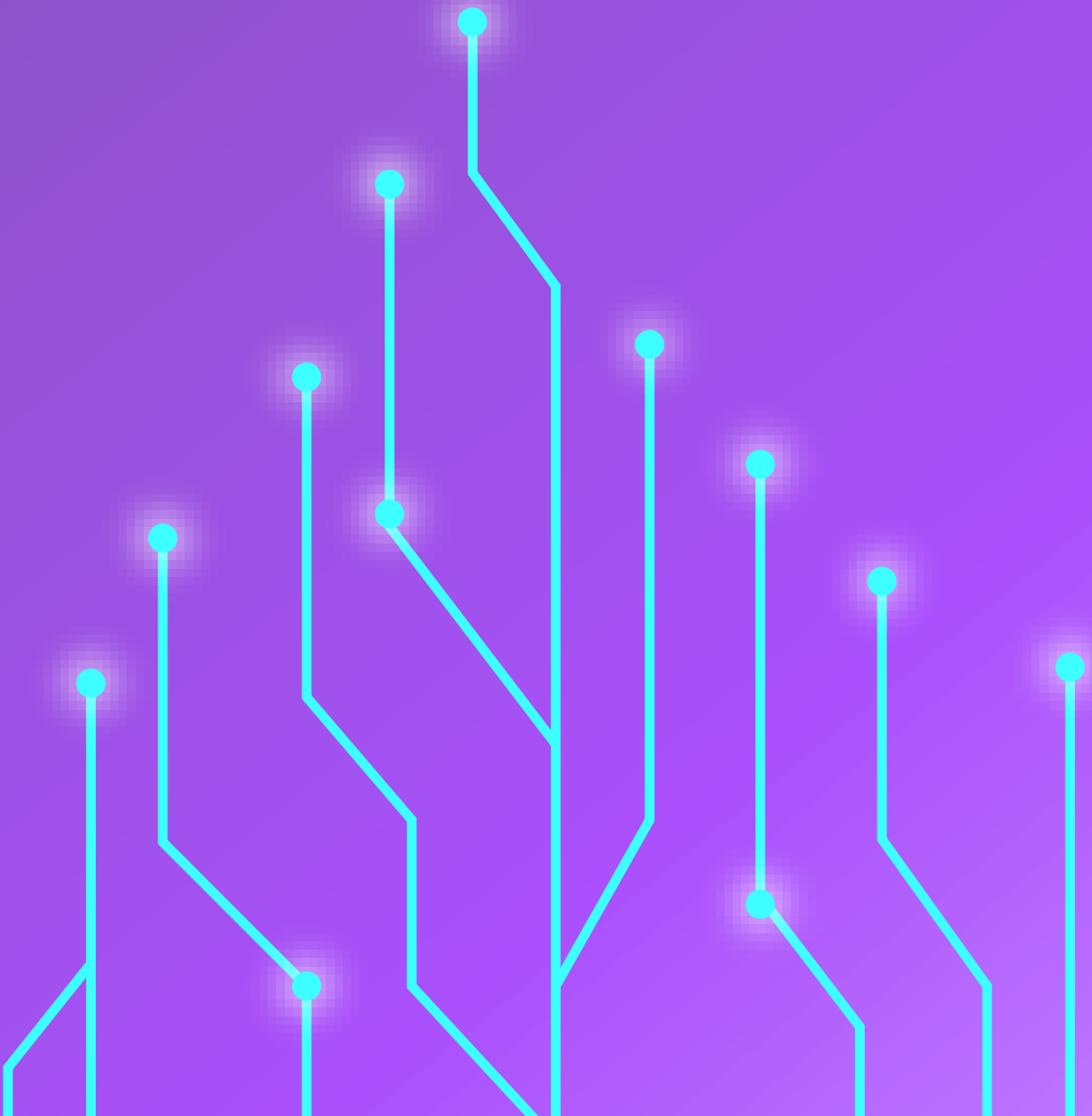
Pentru a aborda aceste probleme, cursanții pot elabora strategii menite să atenueze prejudecățile algoritmice și să promoveze corectitudinea și echitatea în sistemele AI. Acestea pot include punerea în aplicare a tehnicilor de detectare și atenuare a prejudecăților în timpul fazei de dezvoltare, asigurarea unei reprezentări diverse în datele de formare și promovarea transparenței și responsabilității în procesele decizionale algoritmice.

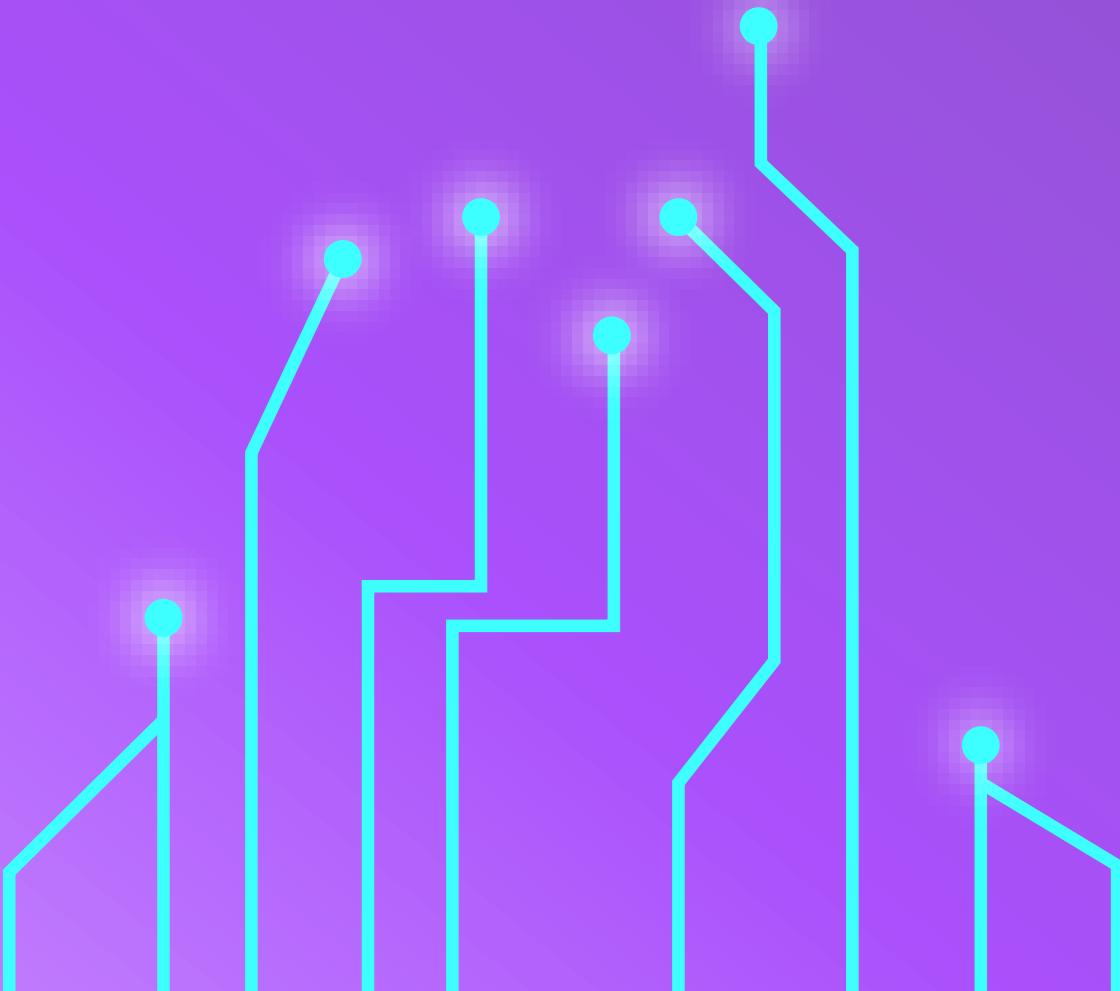
În plus, părțile interesate pot pleda pentru elaborarea și punerea în aplicare a unor cadre de reglementare și orientări care să acorde prioritate echității, nediscriminării și respectării drepturilor omului în dezvoltarea și implementarea IA.



04. Principii de echitate în sistemele AI

CU5 | Drepturile omului și echitatea





04. Principii de echitate în sistemele AI

Echitatea în sistemele AI cuprinde o serie de principii.

Echitatea în sistemele AI cuprinde o serie de principii care sunt fundamentale pentru asigurarea unui tratament și a unor rezultate echitabile pentru toate persoanele:

- Unul dintre principiile de bază este **egalitatea de șanse**, care presupune oferirea acelorași șanse și oportunități indivizilor, indiferent de originea sau caracteristicile acestora. Acest principiu urmărește să uniformizeze condițiile de concurență și să prevină discriminarea bazată pe factori precum rasa, sexul sau statutul socioeconomic.
- **Nediscriminarea** este un alt principiu esențial care interzice utilizarea de criterii arbitrare sau irelevante în procesele decizionale. Sistemele de inteligență artificială nu trebuie să discrimineze persoane sau grupuri pe baza unor caracteristici protejate, cum ar fi rasa, sexul, religia sau handicapul. În schimb, acestea trebuie să trateze toate persoanele în mod corect și imparțial, indiferent de atributele lor personale.





- **Echitatea procedurală** se referă la echitatea proceselor utilizate pentru luarea deciziilor în cadrul sistemelor AI. Aceasta cuprinde transparența, responsabilitatea și dreptul de a face apel sau de a contesta deciziile. Echitatea procedurală garantează faptul că persoanele au un cuvânt de spus în procesele decizionale și că deciziile sunt luate într-un mod transparent și responsabil.
- **Echitatea** este un principiu care se concentrează pe realizarea echității și justiției prin abordarea inegalităților sistemice și prin furnizarea de resurse și oportunități celor care au cea mai mare nevoie de acestea. În contextul sistemelor de inteligență artificială, echitatea presupune conceperea unor algoritmi și politici care să acorde prioritate nevoilor grupurilor marginalizate sau dezavantajate și să urmărească reducerea disparităților în ceea ce privește accesul la oportunități și resurse.
- **Justiția** este un concept mai larg care cuprinde corectitudinea, echitatea și protecția drepturilor omului. Ea urmărește să se asigure că persoanele primesc un tratament corect și că drepturile și demnitatea lor sunt respectate. Justiția în sistemele AI necesită aderarea la standarde etice, reglementări juridice și norme societale care promovează egalitatea, corectitudinea și respectarea drepturilor omului.

Distincția între aceste principii este esențială pentru înțelegerea considerentelor etice nuanțate implicate în dezvoltarea și implementarea inteligenței artificiale. În timp ce egalitatea de șanse se concentrează pe asigurarea șanselor egale pentru toate persoanele, nediscriminarea interzice tratamentul inechitabil bazat pe caracteristici personale. Echitatea procedurală pune accentul pe transparență și responsabilitate în procesele decizionale, în timp ce echitatea vizează abordarea inegalităților sistemice și promovarea echității pentru grupurile marginalizate. În cele din urmă, justiția cuprinde considerații etice și juridice mai ample legate de drepturile omului și bunăstarea societății.

Aplicarea principiilor echității în dezvoltarea IA necesită o abordare proactivă și multidimensională care să ia în considerare implicațiile etice, sociale și juridice ale tehnologiilor IA. Aceste principii pot fi integrate în sistemele IA pentru a promova rezultate echitabile și a atenua prejudecățile algoritmice, punând astfel bazele unui viitor mai corect și mai echitabil pentru generațiile viitoare.

Un aspect esențial al aplicării principiilor echității în dezvoltarea inteligenței artificiale este asigurarea faptului că algoritmi sunt concepuți și instruiți folosind seturi de date diverse și reprezentative. Prin încorporarea diverselor perspective și experiențe în procesul de dezvoltare, dezvoltatorii pot reduce riscul unor rezultate părtinitoare și se pot asigura că sistemele AI sunt echitabile și incluzive.



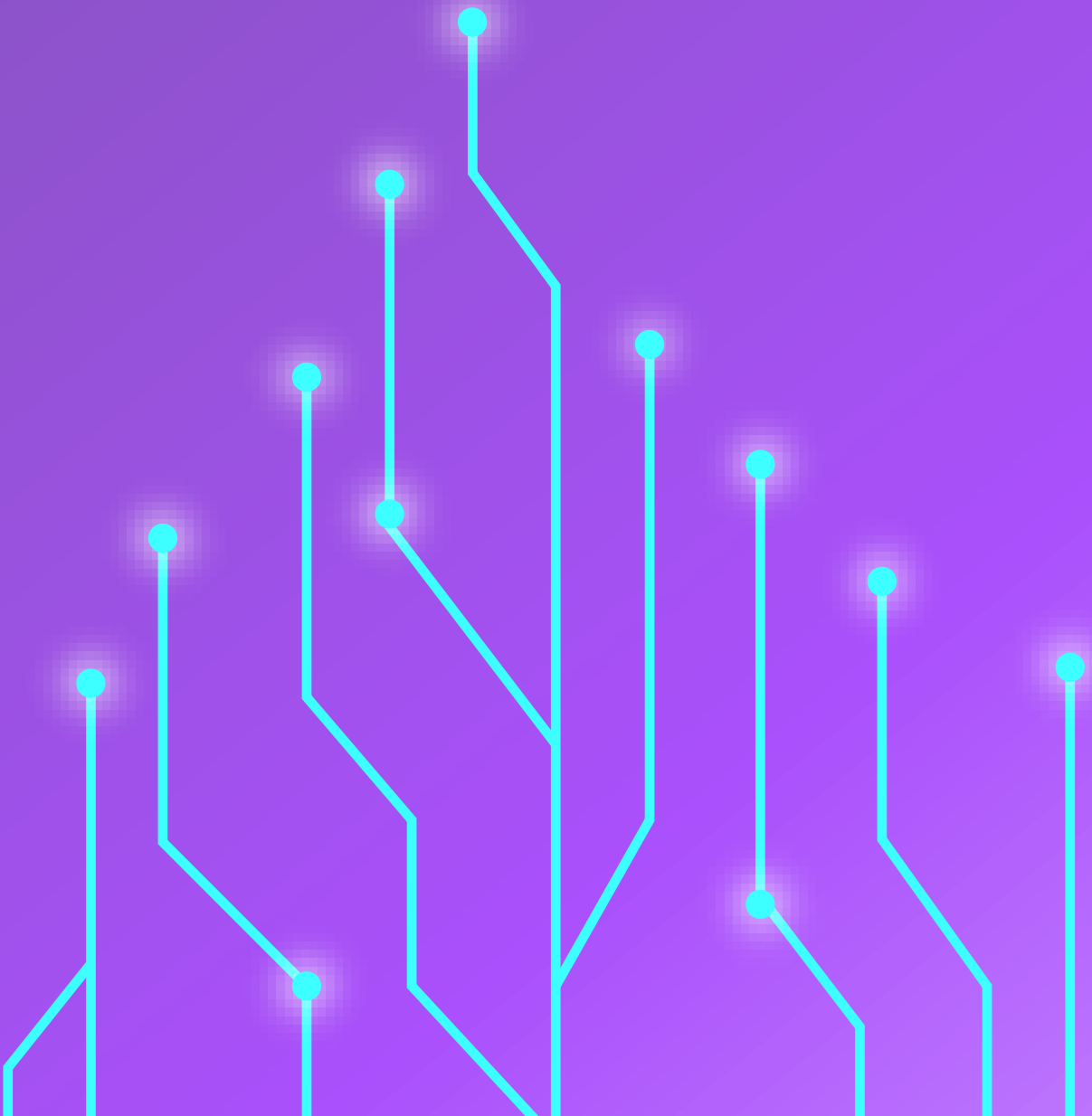
Sistemele de inteligență artificială transparente permit părților interesate să înțeleagă modul în care sunt luate deciziile și să tragă la răspundere dezvoltatorii pentru acțiunile lor. În plus, furnizarea de mecanisme de recurs și de reparație garantează că persoanele afectate de sistemele AI pătinoare au posibilitatea de a căuta dreptate și remediere.

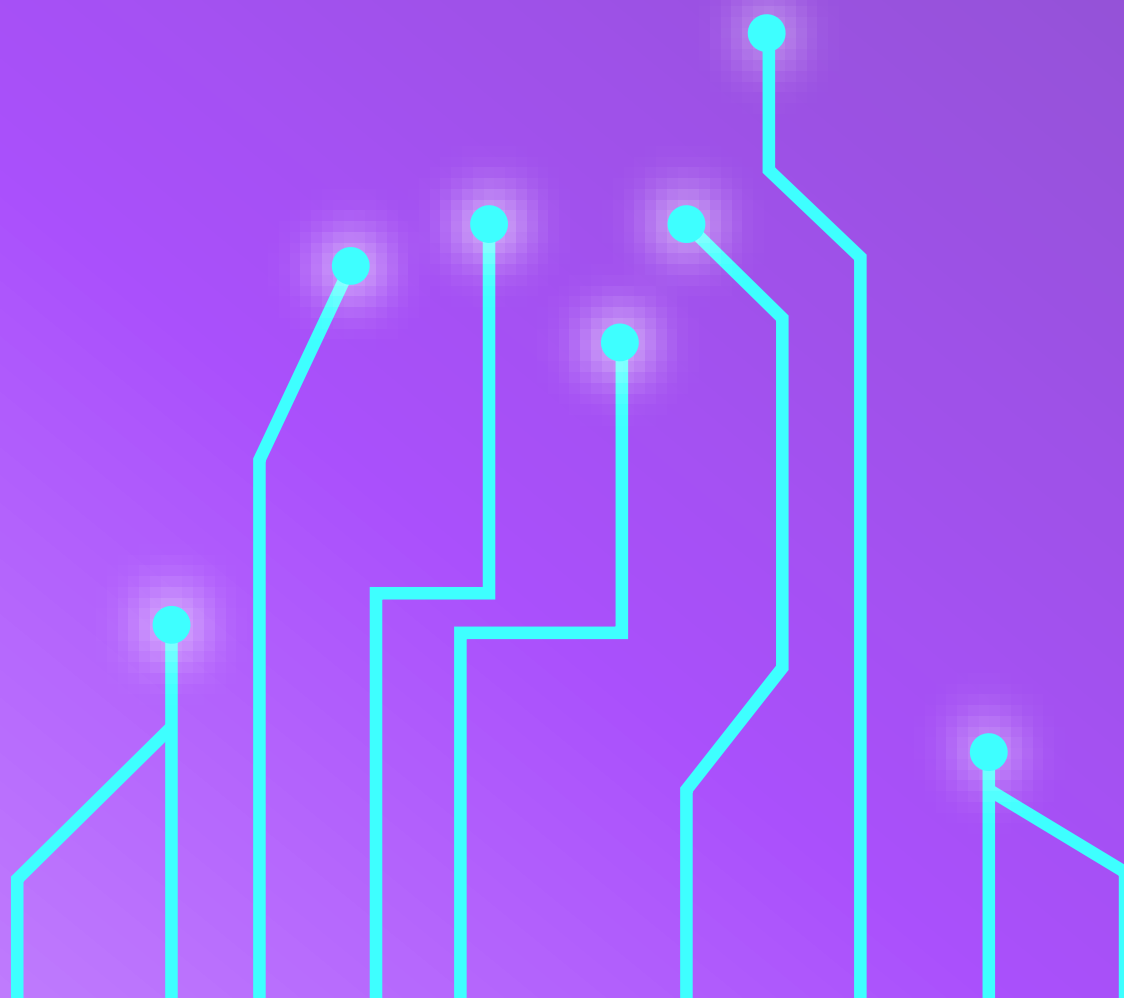
În plus, strategiile de promovare a echității în sistemele de inteligență artificială, cum ar fi proiectarea algoritmilor, acordă prioritate echității și incluziunii și vizează reducerea disparităților în accesul la oportunități și resurse. Prin luarea în considerare a nevoilor și perspectivelor grupurilor marginalizate sau dezavantajate, dezvoltatorii se pot asigura că tehnologiile IA servesc interesele tuturor persoanelor și contribuie la o societate mai echitabilă.



05. Concluzii

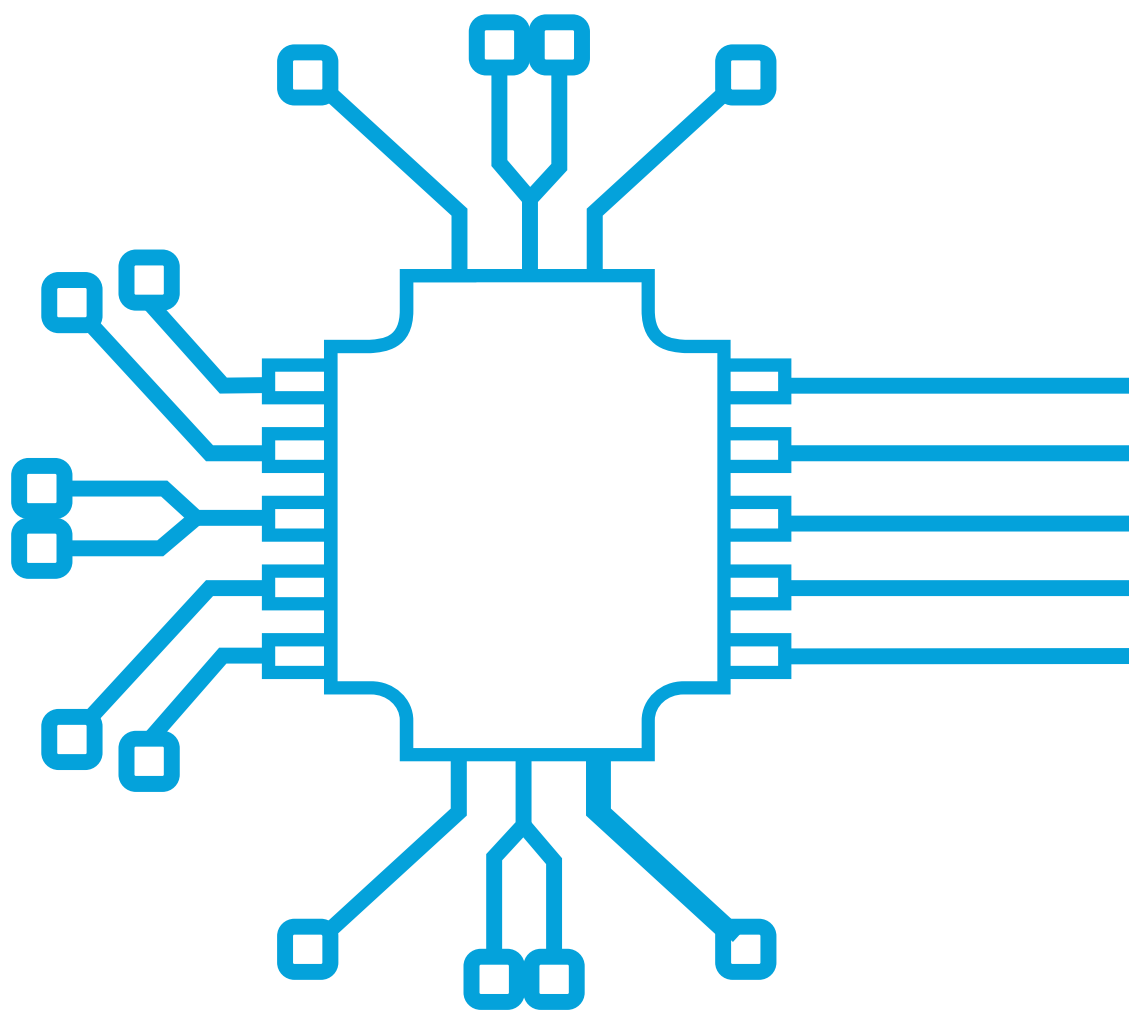
CU5 | Drepturile omului și echitatea

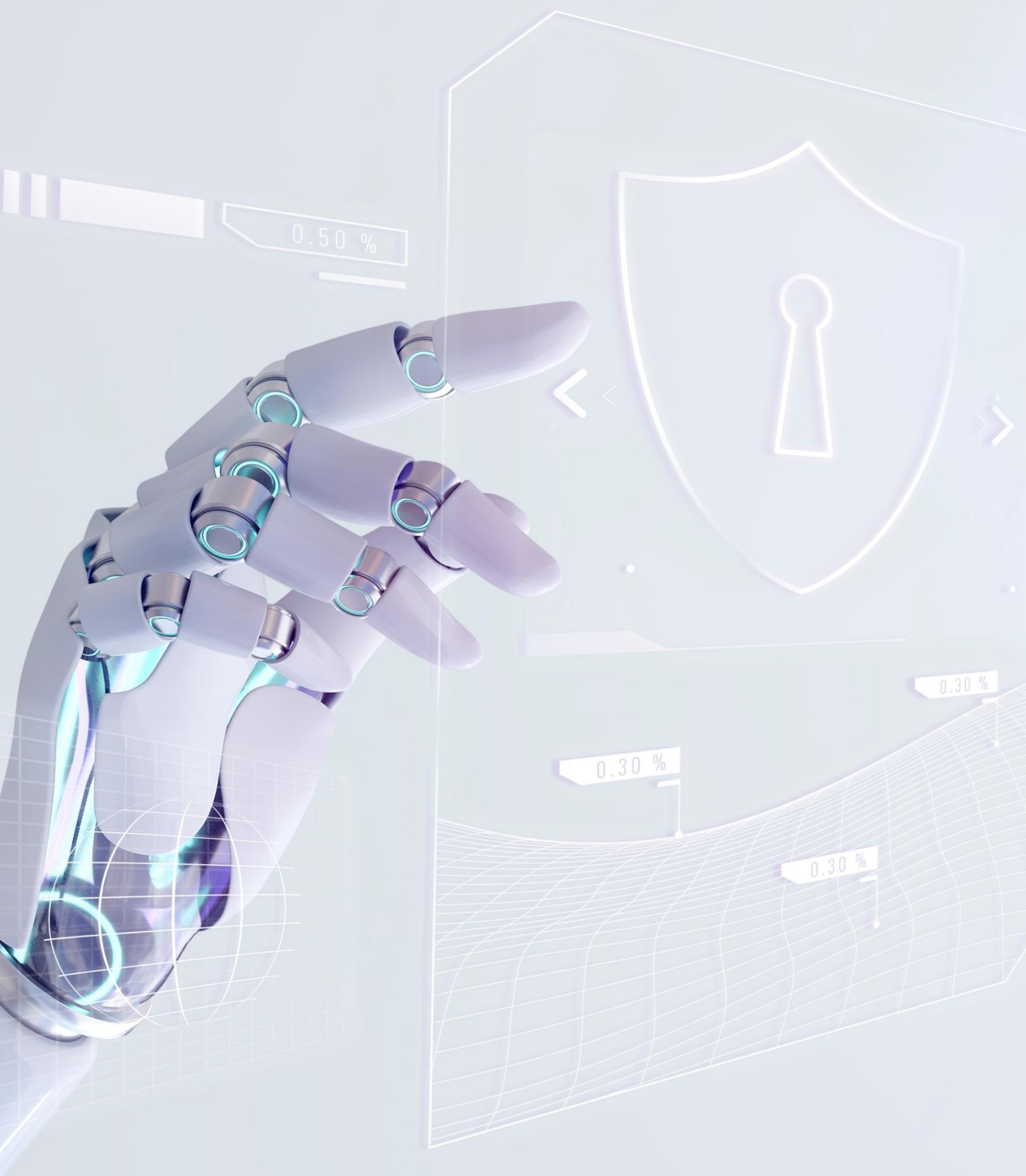




05. Concluzii

În concluzie, înțelegerea și aplicarea principiilor drepturilor omului și ale echității în sistemele de inteligență artificială este esențială pentru promovarea unor rezultate echitabile și protejarea drepturilor fundamentale în era digitală.







Charlie



Cofinanțat de
Uniunea Europeană

Finanțat de Uniunea Europeană. Punctele de vedere și opiniile exprimate aparțin, însă, exclusiv autorului (autorilor) și nu reflectă neapărat punctele de vedere și opiniile Uniunii Europene sau ale Agenției Executive Europene pentru Educație și Cultură (EACEA). Nici Uniunea Europeană și nici EACEA nu pot fi considerate răspunzătoare pentru acestea.



Universitat
de les Illes Balears



ENGAGING PEOPLE



INNOVATION TRAINING CENTER



AARHUS UNIVERSITY



VAMK UNIVERSITY OF APPLIED SCIENCES

helixconnect



2022-1-ES01-KA220-HED-000085257