



Microcredențial de IA etică

CARTEA

CU6 | Etica IA, o abordare practică

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. Înțelegerea implicațiilor etice ale IA

03. Identificarea și diminuarea riscurilor de încălcare a eticii practici în domeniul IA

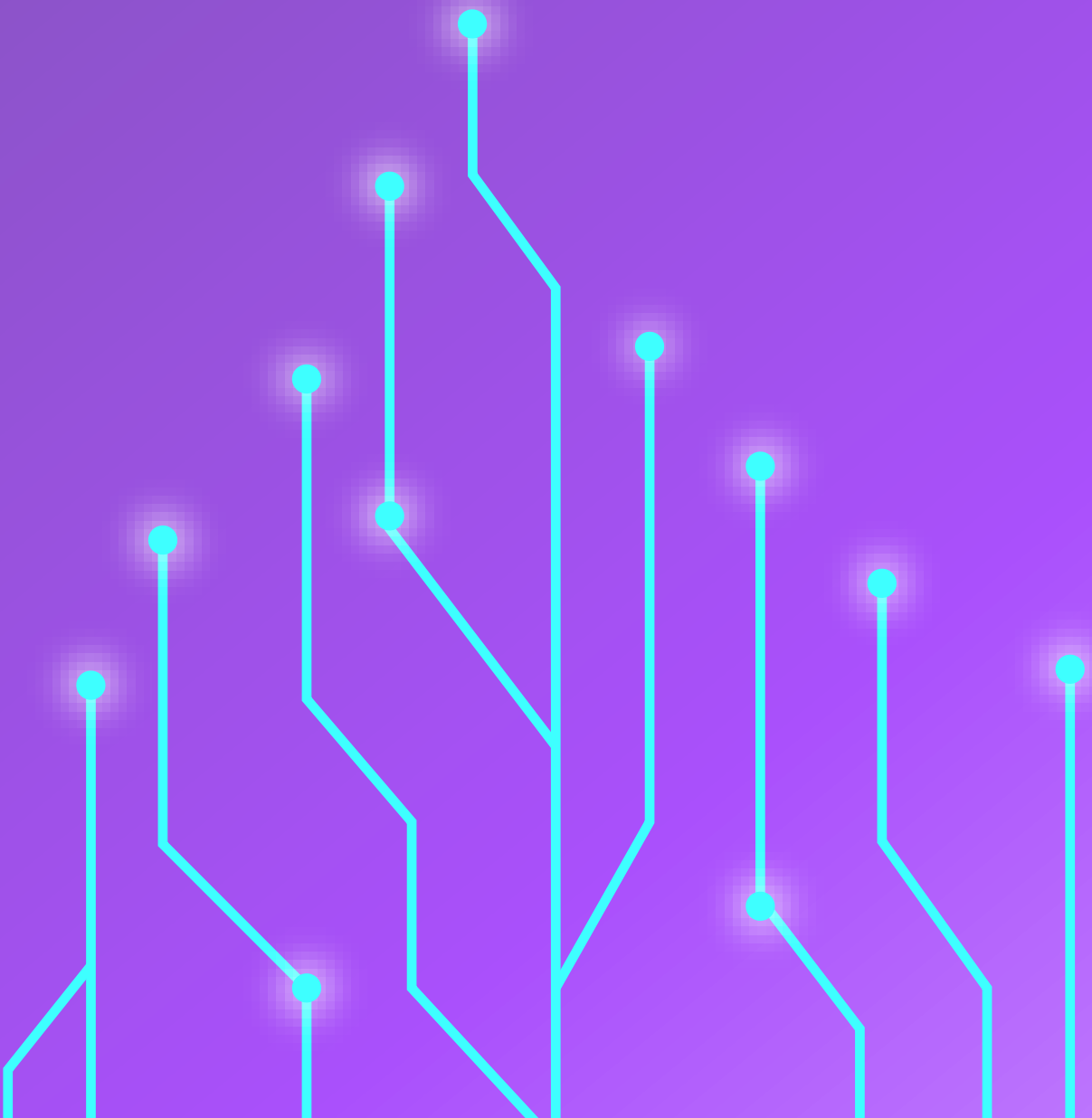
04. Aplicarea practică a orientărilor etice în IA

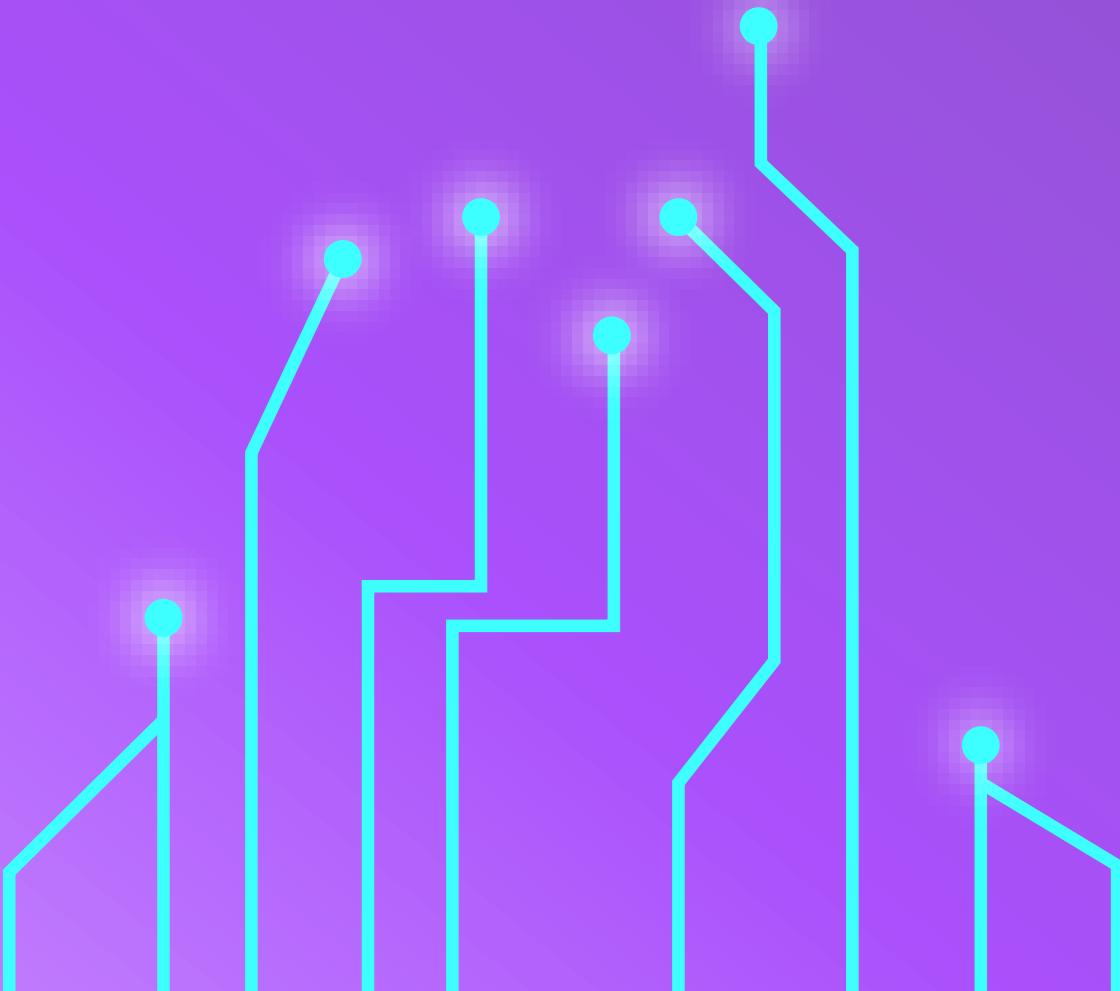
05. Promovarea dezvoltării responsabile a IA și implementare

06. Concluzie

01. Introducere

CU6 | Etica IA, o abordare practică





01. Introducere

Această unitate de competență este concepută pentru ca studenții să aprofundeze aspectele practice ale eticii IA, încorporând principiile dezvoltării și utilizării responsabile a IA în scenarii din lumea reală.

Studenții sunt încurajați să fie proactivi în punerea în aplicare a orientărilor etice în mediile IA, pornind de la cunoștințele teoretice și transpunându-le în perspective concrete.

Rezultatele acestei unități de competență cuprind:

- **Înțelegerea implicațiilor etice ale IA:** studenții vor începe cu o bază solidă în diferitele dimensiuni etice pe care le cuprinde IA. Urmând perspectivele oferite de Floridi et al. (2018) cu privire la preocupările etice din jurul IA, studenții vor fi introduși în implicațiile potențiale ale IA asupra societății, indivizilor și comunității globale.
- **Identificarea și atenuarea practicilor lipsite de etică în IA:** abordează nevoia de recunoaștere și atenuare a potențialelor practici lipsite de etică în IA. Pornind de la reflecțiile lui Bostrom (2014), care a explorat viitorul IA și alinierea sa la valorile umane, studenții vor învăța cum să identifice și să prevină practicile neetice în dezvoltarea și implementarea IA.

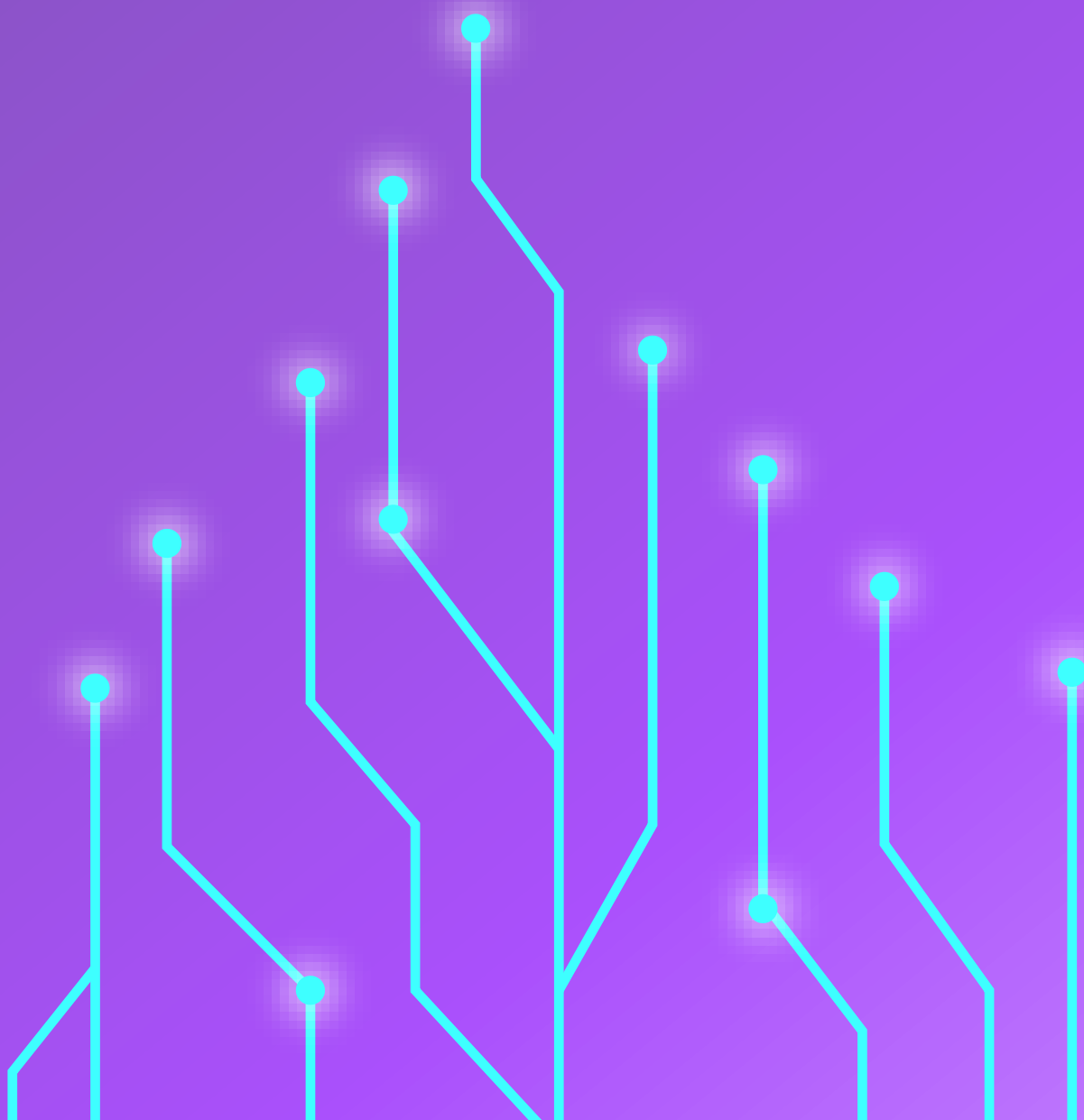


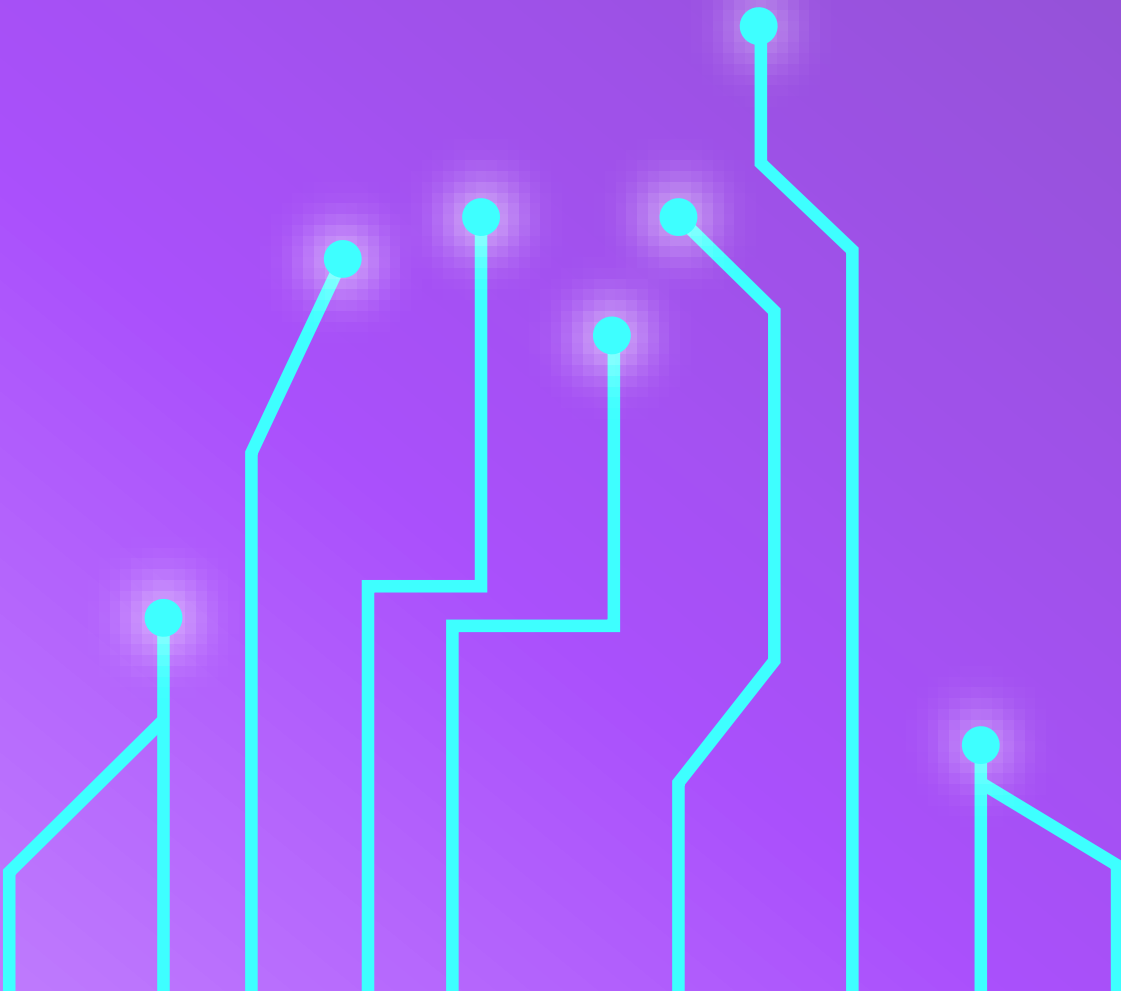
- **Aplicarea practică a orientărilor etice în IA:** studenții din acest modul sunt încurajați să transpună înțelegerea teoretică în acțiuni practice. Pornind de la lucrările lui Ryan & Stahl (2020), care se concentrează pe formularea de orientări etice, studenții se vor angaja în activități care promovează aplicarea practică a acestor orientări în scenariile IA.
- **Promovarea dezvoltării și implementării IA responsabile:** se va concentra pe promovarea unui mediu care încurajează dezvoltarea și implementarea IA responsabile. Pe baza cadrului propus de Jobin et al. (2019), care a discutat perspectivele globale privind etica IA, studenții vor explora metode de promovare a colaborării globale și a creării unei IA responsabile.



02. Înțelegerea implicațiilor etice ale IA

CU6 | Etica IA, o abordare practică





02. Înțelegerea implicațiilor etice ale IA

Pe parcursul acestui curs, am făcut o incursiune profundă în lumea inteligenței artificiale, explorând capacitățile acesteia și riscurile asociate. Până acum, probabil că sunteți conștienți de puterea de transformare a IA și de prezența sa tot mai mare în diverse sectoare. Cu toate acestea, o astfel de putere imensă vine la pachet cu o responsabilitate etică profundă.

Pe măsură ce navigăm în peisajul complex al inteligenței artificiale etice, diverse școli de gândire etică oferă îndrumări valoroase. Utilitarismul, de exemplu, pune accentul pe maximizarea bunăstării generale, determinându-ne să analizăm modul în care IA poate fi concepută pentru a aduce beneficii unui număr cât mai mare de persoane. Etica deontologică, pe de altă parte, se concentrează pe corectitudinea sau incorectitudinea inerentă a acțiunilor, îndemnându-ne să ne asigurăm că sistemele IA funcționează în cadrul etic stabilit. În sfârșit, etica virtuții subliniază importanța caracterului și a dezvoltării morale, încurajându-ne să promovăm o cultură a dezvoltării responsabile a IA, care să acorde prioritate valorilor umane.

În această ultimă unitate, vom explora aceste școli de gândire de bază, transpunându-le în considerații practice pentru construirea unei IA etice.





Vom examina aspecte legate de corectitudine, transparență și responsabilitate și vom discuta despre cum să dezvoltăm cadre solide care să asigure că IA servește umanitatea într-un mod responsabil și benefic. Prin evaluarea critică a implicațiilor etice ale IA, putem deschide calea către un viitor în care aceste tehnologii puternice să contribuie la o societate mai dreaptă și mai echitabilă.

> **Prezentare generală a teoriilor etice și aplicarea lor în AI**

Utilitarismul - o teorie etică proeminentă, care pune accentul pe maximizarea bunăstării generale sau a "utilității" pentru cel mai mare număr de persoane. Aplicat la inteligența artificială, acest lucru înseamnă prioritizarea rezultatelor care aduc beneficii societății în ansamblu, minimizând în același timp potențialele prejudicii.

- **Impact individual:** Imaginați-vă un sistem de camere de securitate alimentat cu inteligență artificială într-un cartier. Deși ar putea spori siguranța publică pentru locuitori, ar putea duce, de asemenea, la o supraveghere sporită și la sentimentul de a fi monitorizat în permanență. O perspectivă utilitaristă ar pune în balanță dreptul individului la viață privată cu potențiala reducere a criminalității pentru întreaga comunitate.
- **Impactul asupra comunității:** Algoritmii alimentați cu inteligență artificială sunt din ce în ce mai utilizați în platformele de socializare pentru a optimiza implicarea utilizatorilor. Cu toate acestea, acești algoritmi pot contribui la răspândirea dezinformării, afectând discursul public și putând dăuna comunităților. Utilitarismul ar pleda pentru măsuri de protecție care să minimizeze răspândirea dezinformării, menținând în același timp implicarea utilizatorilor pe platformă.



- **Impactul societal:** Dezvoltarea sistemelor de arme autonome ridică probleme etice complexe. În timp ce IA ar putea minimiza pierderile umane în războaie, aceasta introduce, de asemenea, întrebări cu privire la responsabilitate și la moralitatea delegării deciziilor de viață sau de moarte către mașini. Utilitarismul ar îndemna la o analiză atentă a implicațiilor societale și a potențialelor consecințe pe termen lung ale implementării unor astfel de tehnologii.

Etica deontologică - o altă teorie influentă, prioritizează corectitudinea sau incorectitudinea inerentă a acțiunilor, independent de rezultatele acestora. Această abordare pune accentul pe reguli și principii morale stabilite, cum ar fi demnitatea umană, consimțământul utilizatorului, protecția vieții private și standardele etice. Aceste principii ar trebui să ghideze dezvoltatorii în proiectarea și implementarea sistemelor AI.

- **Impactul individual:** Tehnologia de recunoaștere facială utilizată la locul de muncă poate spori securitatea, dar ridică, de asemenea, preocupări cu privire la consimțământul utilizatorului și la potențialele prejudecăți ale algoritmilor. Etica deontologică ar impune asigurarea unui consimțământ clar al utilizatorului pentru recunoașterea facială și monitorizarea continuă a potențialelor prejudecăți din cadrul sistemului.

- **Impactul asupra comunității:** Algoritmii predictivi de poliție care urmăresc reducerea criminalității prin identificarea zonelor cu risc ridicat ridică probleme legate de corectitudine și discriminare potențială. Etica deontologică ar pleda pentru asigurarea faptului că acești algoritmi sunt imparțiali și nu vizează în mod disproporționat comunitățile marginalizate.
- **Impactul societal:** Utilizarea tot mai frecventă a inteligenței artificiale în sistemele de arme autonome prezintă o dilemă morală. Etica deontologică ar ridica probabil întrebări cu privire la implicațiile etice ale delegării deciziilor de viață sau de moarte către mașini și la importanța menținerii responsabilității umane în război.

Etica virtuților - inspirată de Aristotel, pune accentul pe cultivarea caracterului moral și a virtuților precum onestitatea, integritatea și empatia. Aplicată la dezvoltarea inteligenței artificiale, această teorie încurajează promovarea comportamentului etic și a procesului decizional în acest domeniu. Aceasta înseamnă construirea de sisteme AI care să acorde prioritate unor trăsături precum empatia, transparența și integritatea.

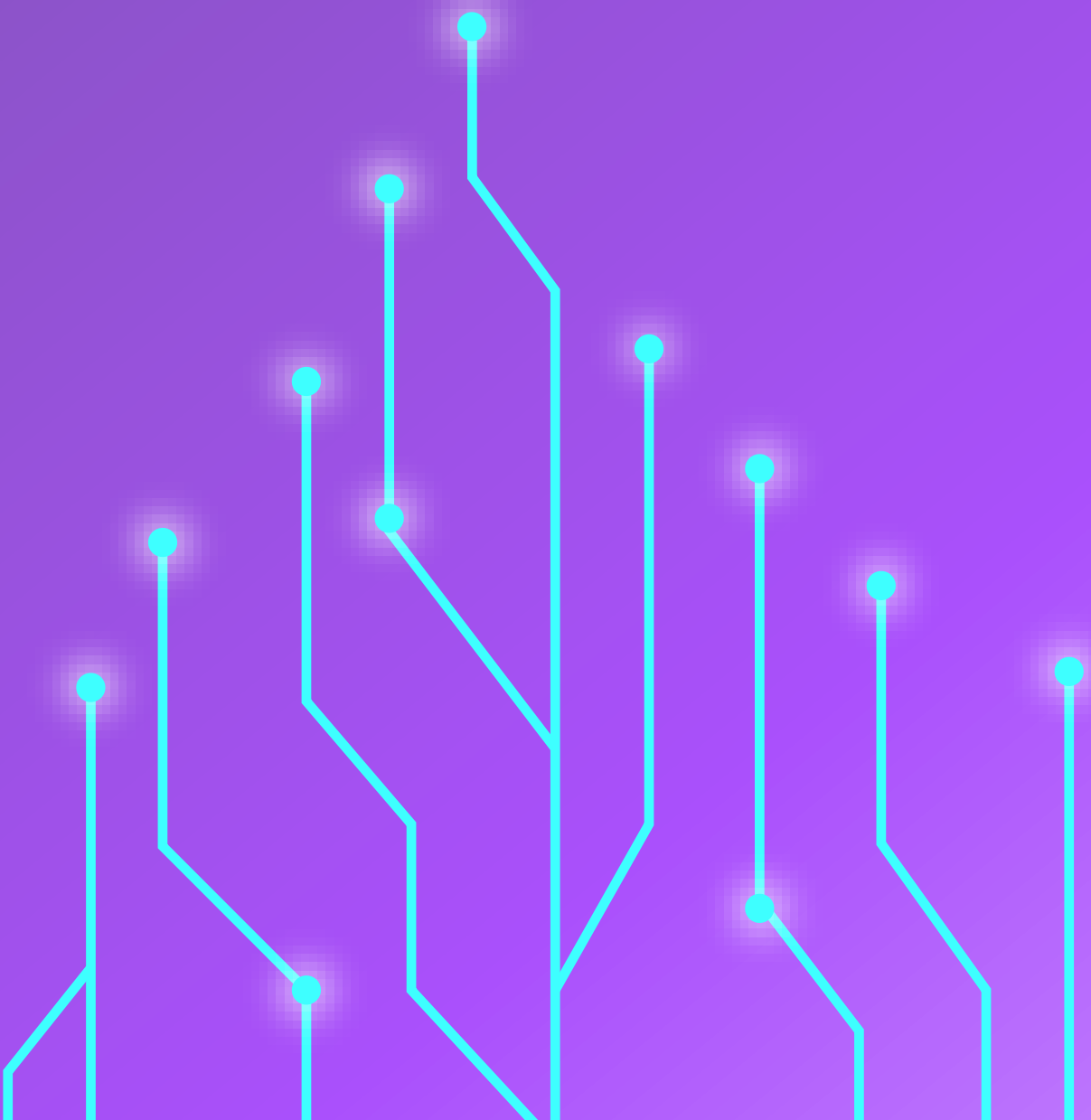
- **Impact individual:** Dezvoltatorii de platforme de social media acordă adesea prioritate implicării utilizatorilor pentru a genera venituri. Cu toate acestea, o abordare bazată pe etica virtuții le-ar cere să ia în considerare impactul societal mai larg al platformelor lor, cum ar fi potențialul de dependență sau răspândirea dezinformării.

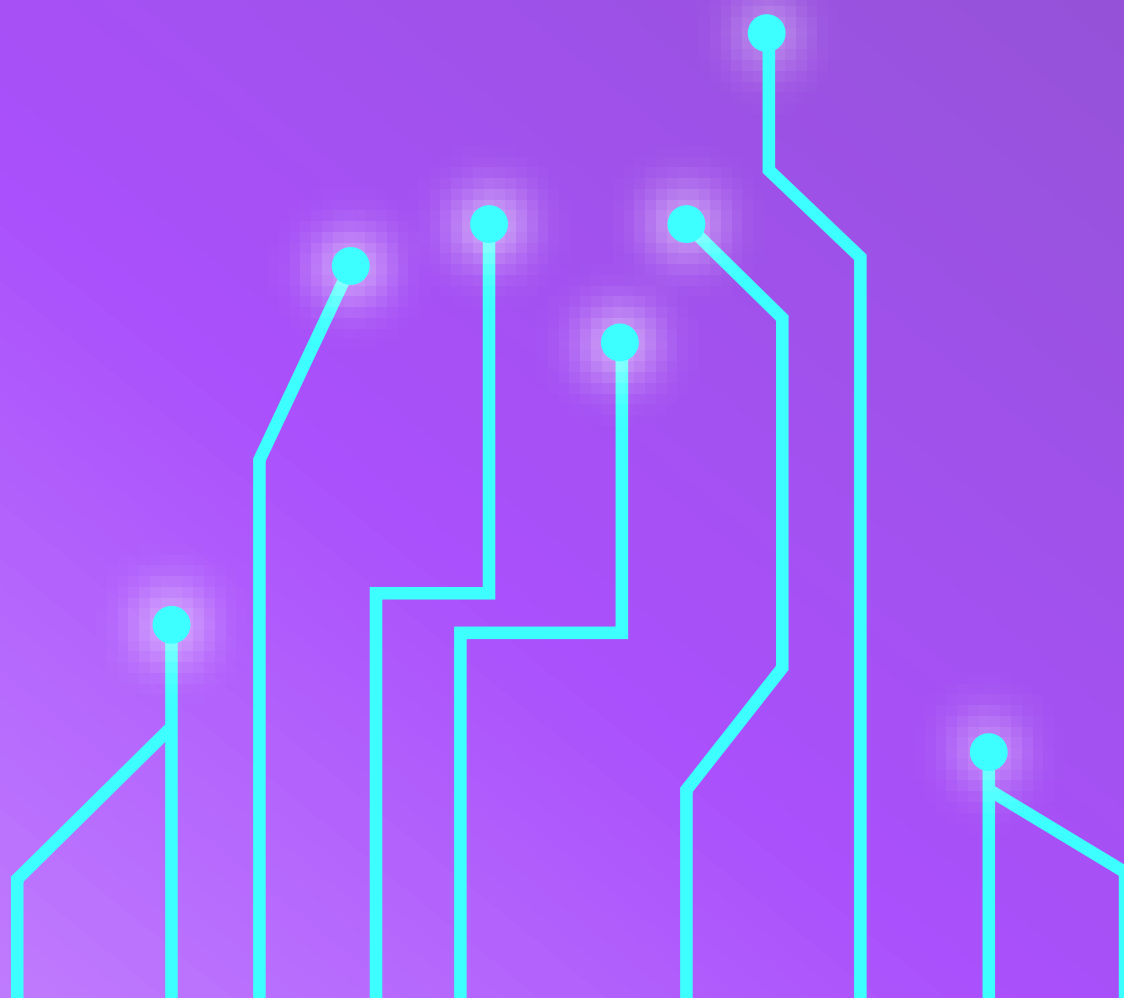


- **Impactul asupra comunității:** Tehnologiile de automatizare a muncii bazate pe inteligența artificială pot crește productivitatea și eficiența. Cu toate acestea, ele suscită, de asemenea, îngrijorări cu privire la dislocarea locurilor de muncă, ceea ce duce la șomaj și inegalitate economică. O perspectivă a eticii virtuților ar încuraja dezvoltatorii să ia în considerare aceste consecințe potențiale și să exploreze modalități de atenuare a impactului negativ asupra comunităților.
- **Impactul societal:** IA este utilizată din ce în ce mai mult în domeniul asistenței medicale pentru a îmbunătăți diagnosticarea și planurile de tratament. Cu toate acestea, îngrijirea centrată pe pacient se bazează în mare măsură pe empatie și compasiune. Etica virtuții ar pleda pentru asigurarea faptului că sistemele de inteligență artificială din domeniul asistenței medicale completează judecata umană și acordă prioritate considerentelor etice, alături de eficiență și precizie.

03. Identificare și atenuarea problemelor imorale practici în domeniul IA

CU6 | Etica IA, o abordare practică





03. Identificarea și diminuarea riscurilor de încălcare a eticii practici în domeniul IA

După cum știți deja, practicile lipsite de etică în inteligența artificială se pot manifesta sub diverse forme, inclusiv părtinire, încălcări ale vieții private și lipsă de transparență. În această secțiune vom trece rapid în revistă unele dintre principalele practici lipsite de etică abordate până acum, modul de identificare a acestora și apoi vom demonstra cum ar putea funcționa o potențială strategie de atenuare pentru fiecare dintre ele.

> Prejudecăți și discriminare

- **Indicatori:** Impact disproporționat asupra anumitor grupuri demografice. Erori sistematice sau disparități în rezultatele procesului decizional. Lipsa diversității în datele de formare sau subreprezentarea comunităților marginalizate.
- **Exemplu:** Algoritmi de recunoaștere facială care prezintă rate de eroare mai mari pentru persoanele de culoare, ceea ce conduce la rezultate discriminatorii în aplicarea legii și în supraveghere.
- **Soluție:** Seturi de date diverse și parametri de evaluare mai corecți



- **Explicație:** Modelele AI învață din datele pe care sunt antrenate. Dacă datele sunt distorsionate, modelul va perpetua aceste distorsiuni. Pentru a atenua acest aspect, dezvoltatorii se concentrează din ce în ce mai mult pe utilizarea unor seturi de date diverse care reprezintă populația lumii reale. În plus, se dezvoltă metrice de evaluare mai echitabile pentru a evalua dacă modelele AI sunt părtinitoare și pentru a se asigura că acestea nu dezavantajează pe nedrept anumite grupuri.



> Lipsa de transparență

- **Indicatori:** Algoritmi de tip black-box cu procese decizionale opace. Acces limitat la informații privind sursele de date, arhitectura modelului și criteriile de decizie. Absența explicațiilor sau a justificărilor pentru rezultatele generate de AI.
- **Exemplu:** Algoritmi financiari care resping cererile de împrumut fără a oferi motive sau criterii clare de respingere, exacerband neîncrederea și frustrarea în rândul solicitanților.
- **Soluție:** Inițiative privind inteligența artificială explicabilă (XAI)
- **Explicație:** Inteligența artificială explicabilă (XAI) este un domeniu în creștere care se concentrează pe dezvoltarea de modele de inteligență artificială care sunt mai transparente și mai interpretabile. Acest lucru permite oamenilor să înțeleagă modul în care modelul ajunge la deciziile sale, promovând încrederea și permițându-ne să identificăm și să abordăm potențialele prejudecăți. De exemplu, un sistem XAI ar putea evidenția factorii dintr-un set de date privind o cerere de împrumut care au influențat cel mai mult decizia modelului.





> Încălțări ale dreptului la viață privată

- **Indicatori:** Accesul neautorizat sau utilizarea abuzivă a datelor cu caracter personal. Garanții inadecvate pentru protejarea informațiilor sensibile. Neobținerea consimțământului informat pentru colectarea și prelucrarea datelor.
- **Exemplu:** Platformele de social media care utilizează algoritmi de inteligență artificială pentru a analiza comportamentul și preferințele utilizatorilor fără mecanisme transparente de dezvăluire sau excludere voluntară, compromițând viața privată și autonomia utilizatorilor.
- **Soluție:** Reglementări privind confidențialitatea datelor și controlul utilizatorului
- **Explicație:** Reglementările privind confidențialitatea datelor, precum GDPR (General Data Protection Regulation - Regulamentul general privind protecția datelor) și CCPA (California Consumer Privacy Act - Legea privind confidențialitatea consumatorilor din California) oferă utilizatorilor mai mult control asupra informațiilor lor personale. Aceste reglementări impun companiilor să fie transparente cu privire la practicile de colectare a datelor, să obțină consimțământul în cunoștință de cauză și să ofere utilizatorilor posibilitatea de a accesa, corecta sau șterge datele lor.

> Lipsa de responsabilitate

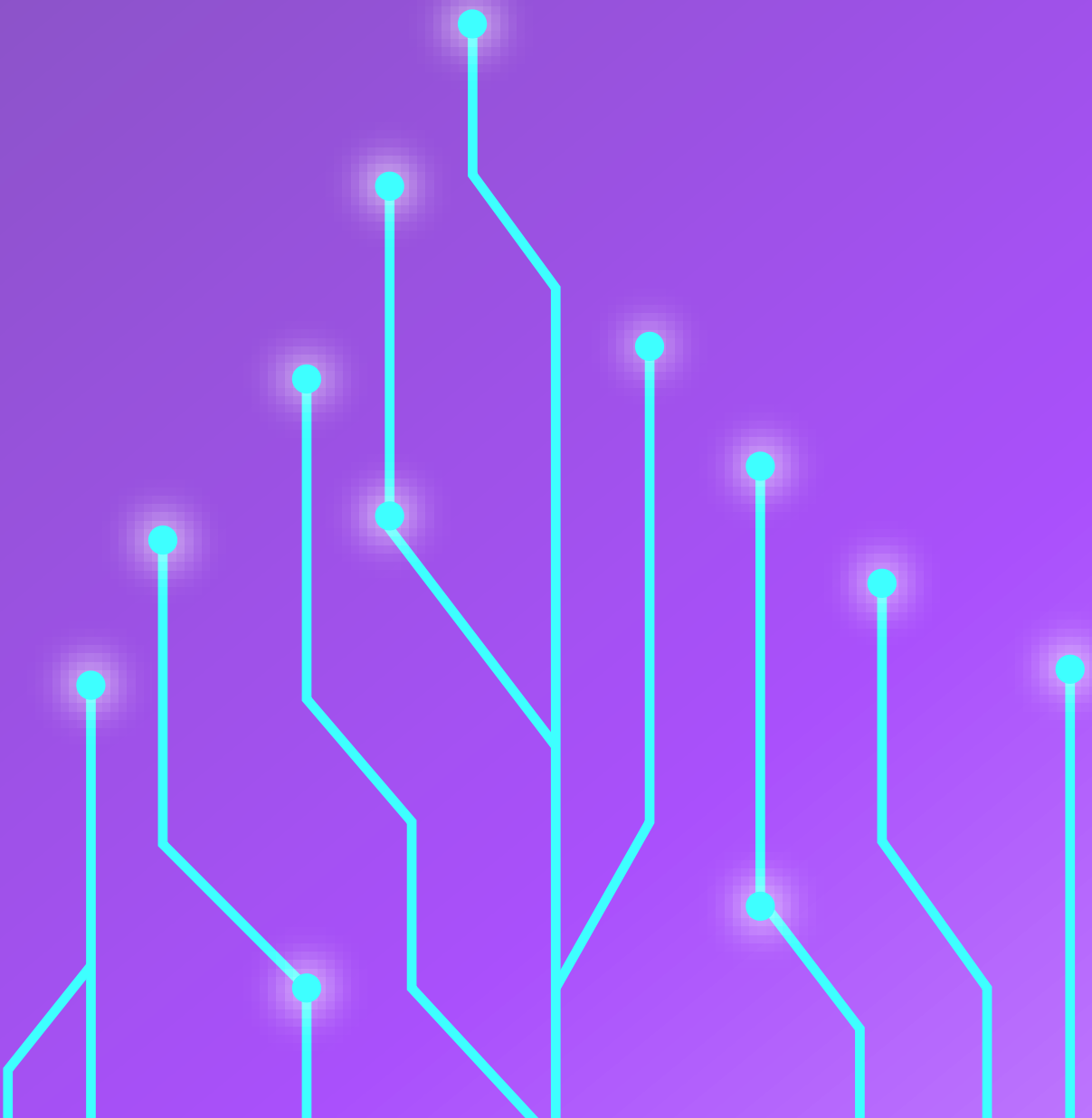
- **Indicatori:** Absența mecanismelor de identificare a părților responsabile sau de tragere la răspundere a acestora pentru prejudiciile cauzate de IA. Căi de atac sau remedii limitate pentru persoanele afectate de comportamentul necorespunzător al IA. Neefectuarea de evaluări ale impactului sau de audituri ale sistemelor AI.
- **Exemplu:** Vehicule autonome implicate în accidente datorate unor defecțiuni ale inteligenței artificiale, în care producătorii evită răspunderea sau responsabilitatea pentru daune, ceea ce ridică întrebări cu privire la responsabilitatea juridică și etică în cazul tehnologiilor bazate pe inteligența artificială.
- **Soluție:** Supravegherea umană și cadrele juridice
- **Explicație:** Pentru aplicațiile IA cu risc ridicat, supravegherea umană poate fi esențială. Aceasta înseamnă implicarea oamenilor în procesul de luare a deciziilor pentru a asigura utilizarea responsabilă a IA și atenuarea riscurilor potențiale. În plus, sunt în curs de elaborare cadre juridice pentru a stabili linii clare de responsabilitate pentru prejudiciile legate de inteligența artificială. Acest lucru va contribui la asigurarea faptului că dezvoltatorii și companiile sunt responsabili pentru implicațiile etice ale sistemelor lor de IA.

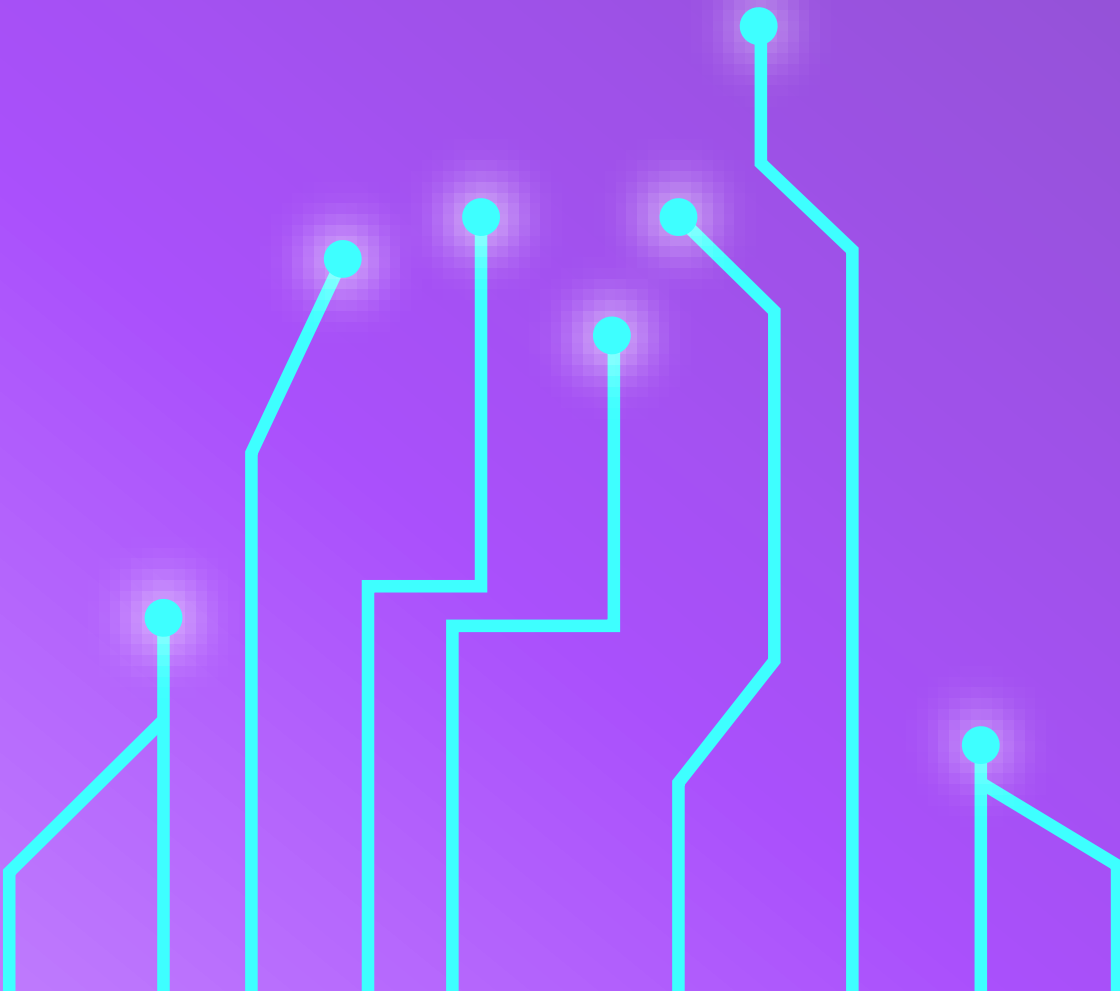


0101

04. Aplicarea practică a orientărilor etice în IA

CU6 | Etica IA, o abordare practică





04. Aplicarea practică a orientărilor etice în IA

➤ **Construirea setului de instrumente etice pentru IA - Un cadru de acțiune**

Orientările etice oferă o foaie de parcurs pentru luarea deciziilor etice în dezvoltarea și implementarea IA. Pe parcursul acestei unități, am explorat rolul esențial al considerațiilor etice în dezvoltarea și implementarea IA. Am examinat diverse școli de gândire etică, fiecare oferind cadre valoroase pentru navigarea în peisajul etic complex al IA. Dar cum putem transpune aceste principii abstracte în acțiuni practice?

Această secțiune este un cadru simplu în patru pași care vă va ghida în elaborarea propriilor orientări etice privind IA.

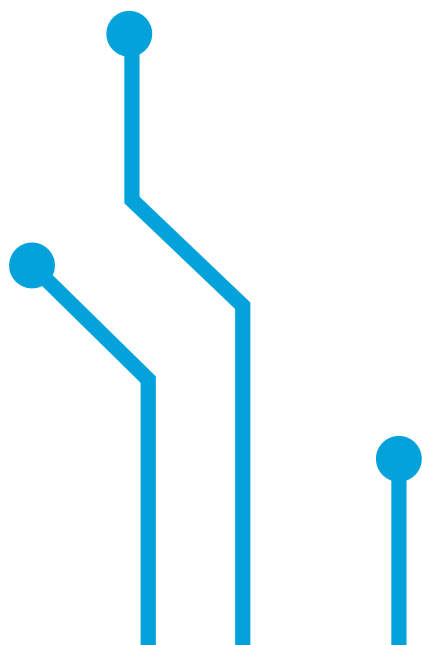
1. Pasul 1: Delimitarea scopului sistemului AI

Primul pas implică definirea clară a cazului de utilizare preconizat pentru sistemul de inteligență artificială. Ce sarcină specifică este proiectată să îndeplinească AI? Luați în considerare exemple precum software-ul de recunoaștere facială în scopuri de securitate, un instrument de angajare bazat pe inteligență artificială pentru a automatiza verificarea CV-urilor sau un asistent de diagnosticare medicală pentru a sprijini profesioniștii din domeniul sănătății. Înțelegerea scopului sistemului este esențială pentru identificarea părților interesate implicate și a aspectelor etice care vor apărea.



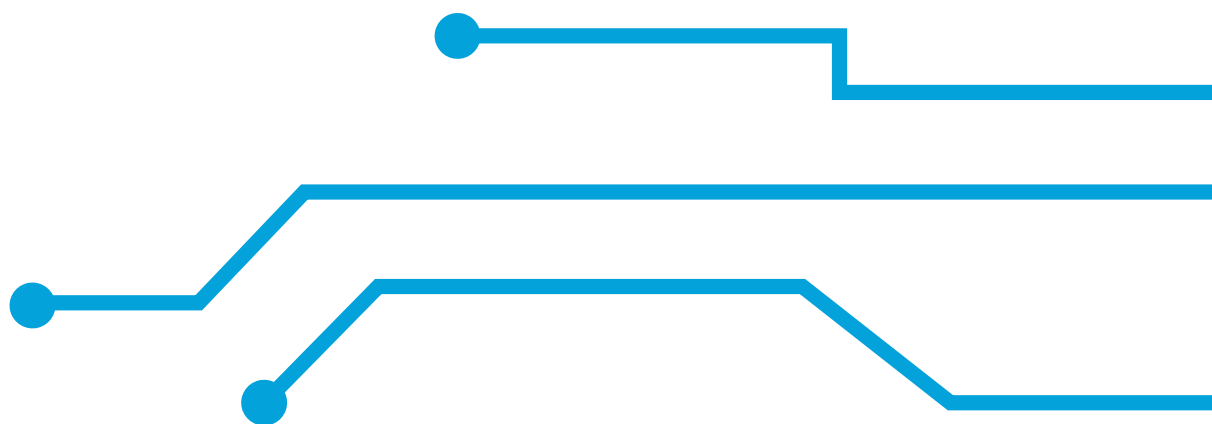
2. Etapa 2: Identificarea părților interesate și a valorilor acestora

Odată stabilit cazul de utilizare, a doua etapă constă în identificarea exactă a diferitelor părți interesate care vor fi afectate de sistemul de inteligență artificială. Părțile interesate pot cuprinde o gamă largă de persoane și grupuri, inclusiv utilizatorii sistemului, dezvoltatorii, cei afectați de deciziile acestuia (de exemplu, solicitanții de împrumuturi pentru o platformă de creditare bazată pe IA) și societatea în general. Fiecare grup de părți interesate va avea propriul set de valori și priorități care trebuie luate în considerare la elaborarea orientărilor etice. De exemplu, utilizatorii ar putea acorda prioritate confidențialității și corectitudinii în ceea ce privește modul în care sunt gestionate datele lor, în timp ce dezvoltatorii s-ar putea concentra asupra eficienței și preciziei în funcționarea sistemului. Este important să recunoaștem aceste valori diverse și să găsim modalități de a le echilibra atunci când formulăm principii etice.



3. Etapa 3: Enunțarea principiilor etice fundamentale

Pe baza înțelegerii cazului de utilizare și a părților interesate identificate, a treia etapă implică definirea principiilor etice de bază care vor ghida dezvoltarea și implementarea sistemului de IA. Aceste principii servesc drept bază pentru asigurarea utilizării responsabile și etice a tehnologiei IA. Printre principiile frecvent luate în considerare se numără echitatea, transparența, responsabilitatea, confidențialitatea și siguranța. De exemplu, în cazul unui instrument de angajare bazat pe inteligență artificială, echitatea ar necesita asigurarea faptului că sistemul nu discriminează anumite categorii demografice pe baza unor date pătinoitoare. Transparența ar necesita o comunicare clară cu privire la modul în care funcționează instrumentul și la factorii pe care îi ia în considerare la evaluarea candidaților.





4. Etapa 4: Transpunerea principiilor în măsuri aplicabile

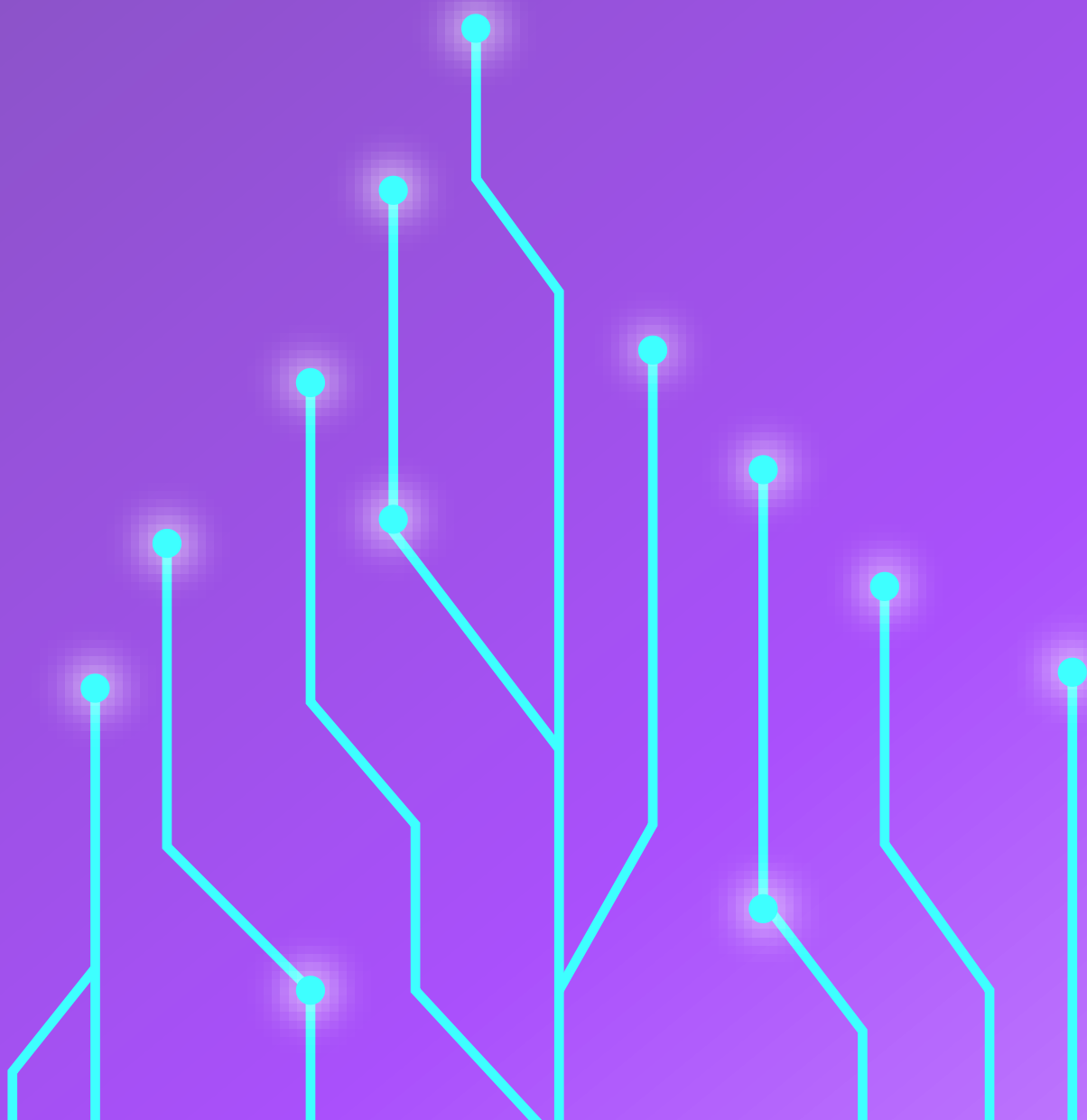
Etapa finală transpune principiile etice abstracte în acțiuni concrete care pot fi puse în aplicare pentru a se asigura că sistemul AI aderă la aceste principii. Aceasta ar putea implica utilizarea unor seturi de date diverse pentru a atenua părtinirea în procesul decizional al sistemului. Pentru transparență, pot fi încorporate tehnici de inteligență artificială explicabilă (XAI) pentru a permite utilizatorilor să înțeleagă modul în care inteligența artificială ajunge la rezultatele sale. În plus, este esențial să se stabilească linii clare de responsabilitate pentru eventualele prejudicii cauzate de sistemul AI. Aceasta ar putea implica atribuirea responsabilității dezvoltatorilor sau crearea de mecanisme de despăgubire pentru cei afectați negativ de deciziile AI.

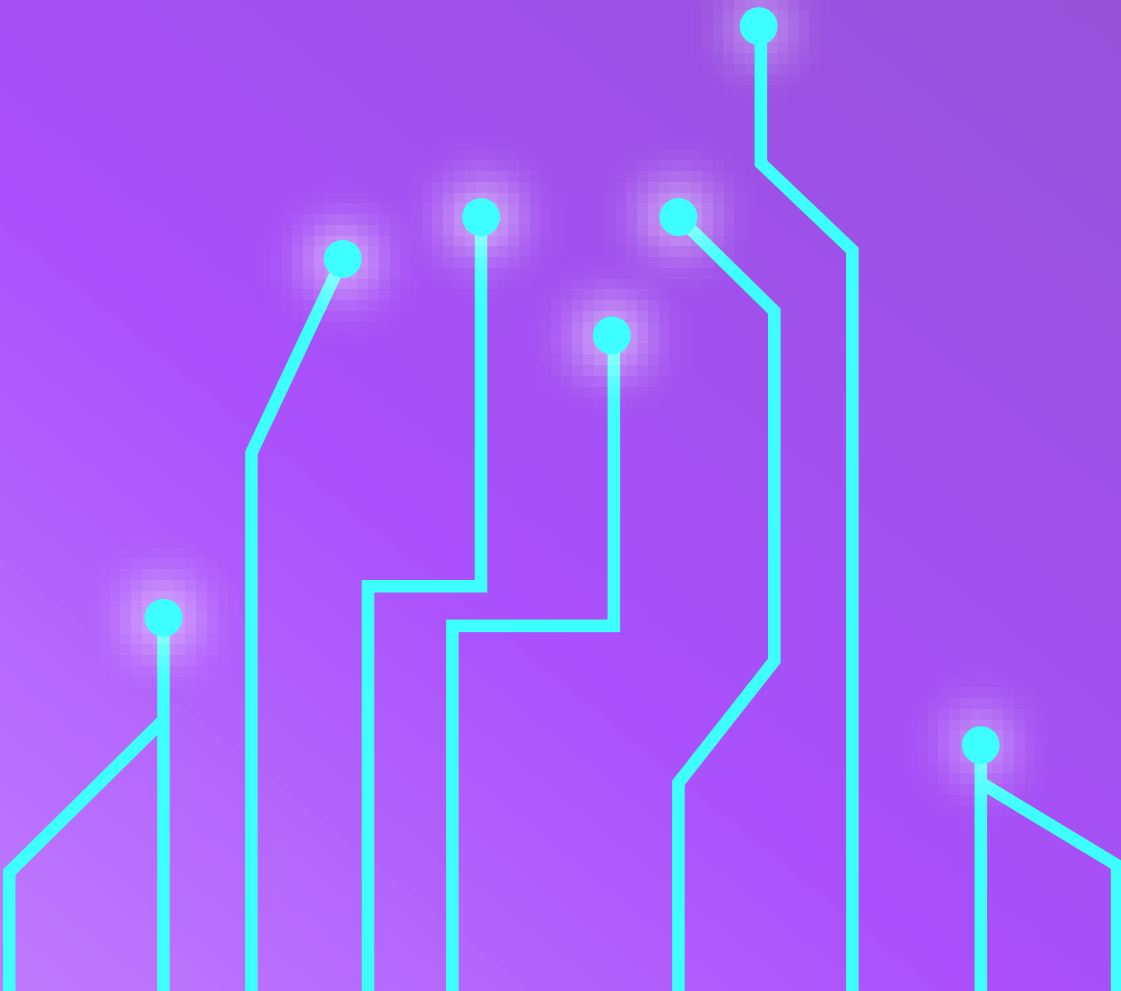
Urmând acești pași, veți fi în măsură să evaluați critic sistemele AI și să contribuiți la dezvoltarea și utilizarea responsabilă a acestor tehnologii puternice.



05. Promovarea dezvoltării responsabile a IA și implementare

CU6 | Etica IA, o abordare practică





05. Promovarea dezvoltării responsabile a IA și implementare

> Construim împreună un viitor mai bun - Colaborare globală pentru IA responsabilă

Am explorat importanța dezvoltării responsabile a IA pe parcursul acestui curs. Dar asigurarea unei IA etice nu este doar o provocare locală - necesită un efort global. Aici vom examina inițiativele internaționale și strategiile de colaborare care lucrează pentru a promova dezvoltarea și implementarea inteligenței artificiale responsabile la nivel mondial.

- **Parteneriatul global privind inteligența artificială (GPAI):**
Aceasta este o inițiativă multilaterală care implică guverne, lideri din industrie și organizații ale societății civile. GPAI își propune să elaboreze cele mai bune practici și recomandări pentru dezvoltarea și implementarea responsabilă a IA în diverse sectoare. <https://gpai.ai/>
- **Observatorul OCDE privind politica în domeniul IA:**
Această inițiativă a Organizației pentru Cooperare și Dezvoltare Economică (OCDE) oferă o platformă pentru schimbul de informații și de bune practici privind elaborarea de politici în domeniul IA între țările membre. Scopul său este de a promova dialogul și colaborarea internațională privind guvernarea etică a IA. <https://oecd.ai/>





- **Recomandarea UNESCO privind etica inteligenței artificiale:**

Acesta este un instrument internațional neobligatoriu adoptat de UNESCO în 2021. Acesta subliniază principiile etice cheie pentru dezvoltarea și utilizarea inteligenței artificiale, promovând drepturile omului, corectitudinea, transparența și responsabilitatea.

<https://unesdoc.unesco.org/ark:/48223/pf0000380455#:~:text=AI%20actors%20and%20Member%20States,law%2C%20in%20particular%20Member%20States>



➤ **Actul UE privind inteligența artificială**

Actul UE privind IA este prima încercare de a stabili un cadru cuprinzător pentru reglementarea aplicațiilor IA în cadrul Uniunii Europene. Acesta a fost propus pentru prima dată de Comisia Europeană în 2021, fiind în cele din urmă aprobat de Parlamentul și Consiliul UE în mai 2024.

Legea clasifică sistemele AI în patru categorii de risc - inacceptabil, ridicat, limitat și minim - plus o categorie suplimentară pentru AI cu scop general. Aplicațiile AI considerate a avea riscuri inacceptabile, cum ar fi cele care manipulează comportamentul uman sau exploatează vulnerabilități, sunt interzise categoric.

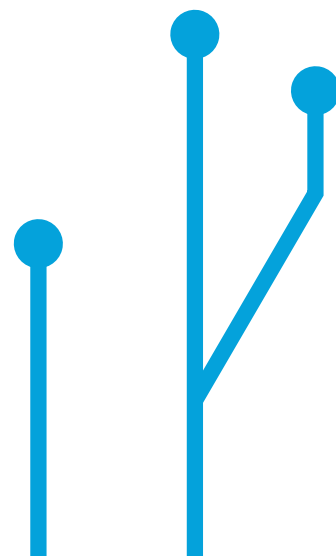
Aplicațiile cu risc ridicat trebuie să respecte cerințe stricte de securitate, transparență și calitate și să fie supuse unor evaluări ale conformității pentru a asigura conformitatea. Aplicațiile cu risc limitat sunt supuse obligațiilor de transparență, în timp ce aplicațiile cu risc minim nu sunt reglementate. IA de uz general, în special cele cu capacități ridicate, trebuie să îndeplinească cerințe de transparență și pot fi supuse unor evaluări suplimentare.

Această abordare bazată pe riscuri urmărește să atenueze potențialele prejudicii, promovând în același timp inovarea și asigurând alinierea dezvoltării IA la valorile și principiile europene.

https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_EN.pdf



<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>



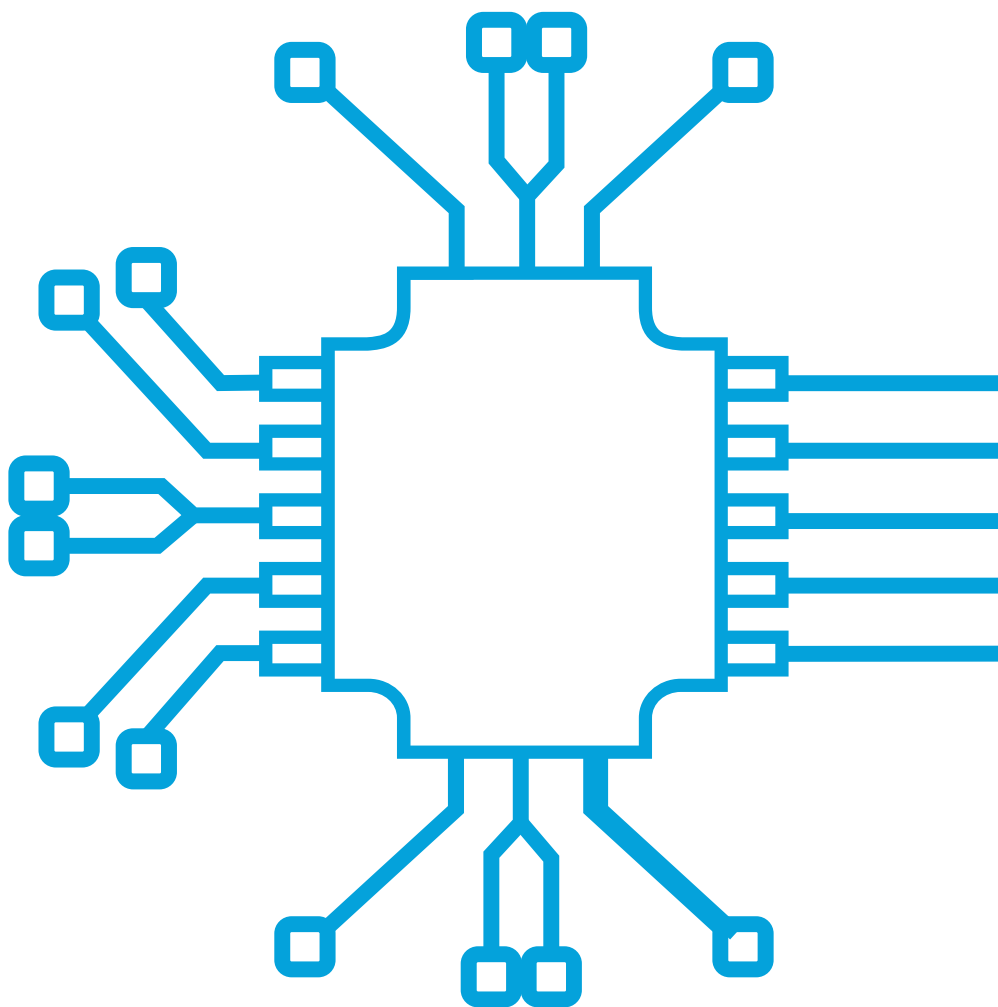


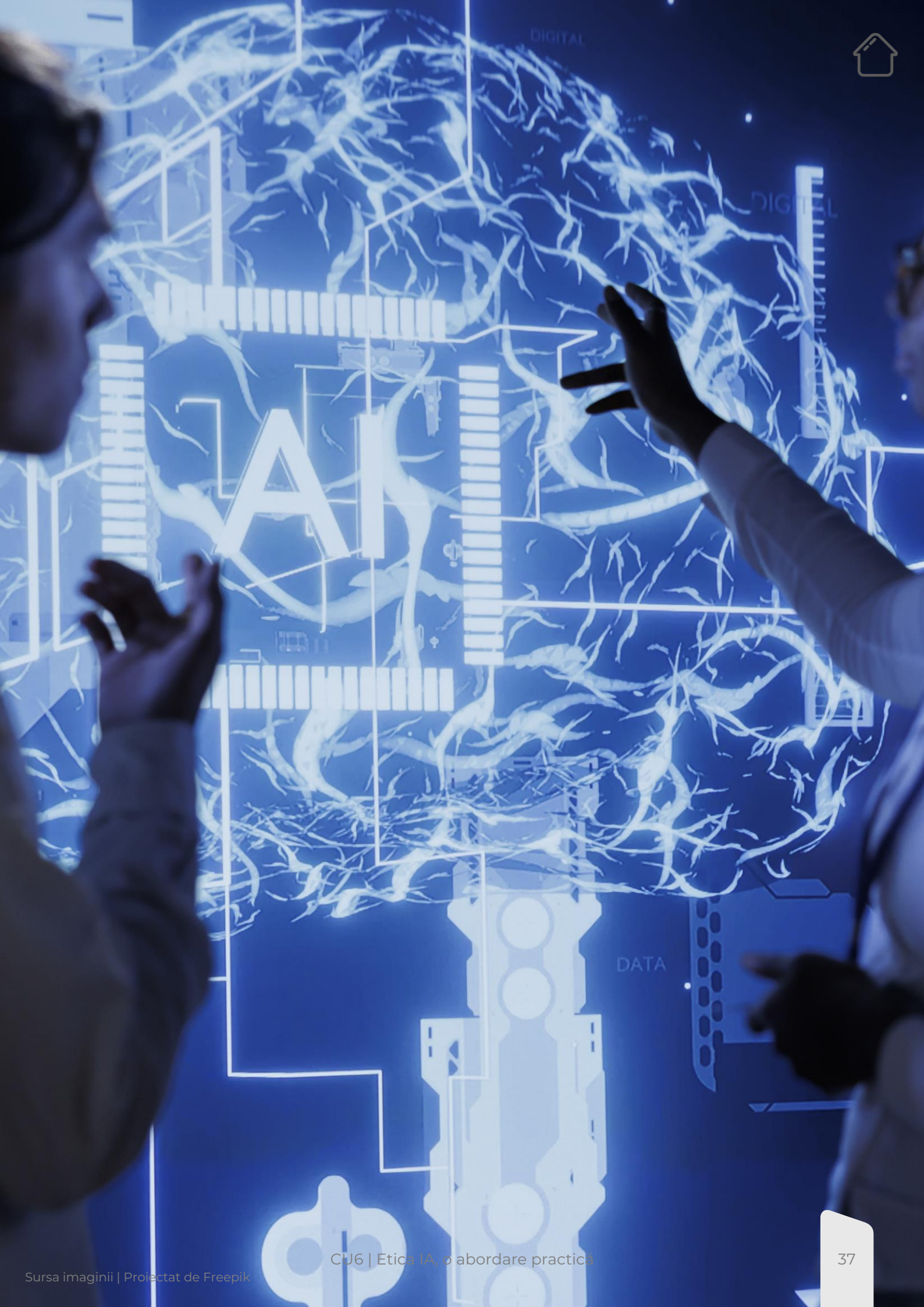
> **Puteți face diferența - Strategii pentru construirea unui viitor responsabil al IA**

Ați dobândit acum cunoștințe valoroase despre etica IA și importanța dezvoltării responsabile a IA. Dar această călătorie nu se încheie aici. Viitorul IA se conturează chiar acum, iar vocea dumneavoastră contează. Mai jos sunt prezentate câteva sugestii pentru a vă oferi strategii practice prin care să vă implicați și să contribuiți la un viitor mai responsabil al IA în comunitatea dumneavoastră și nu numai.

- **Creșteți gradul de conștientizare:** Alegeți o problemă etică specifică în domeniul inteligenței artificiale care vă interesează (de exemplu, prejudecăți în recunoașterea facială, preocupări legate de confidențialitate în cazul asistenților alimentați cu inteligență artificială). Cercetați problema și pregătiți o scurtă prezentare sau un atelier pentru clubul școlar, centrul comunitar sau chiar platformele online. Adaptați-vă prezentarea la publicul dumneavoastră și faceți-o atractivă (utilizați elemente vizuale, interactive).
- **Militați pentru schimbare:** Identificați inițiativele politice existente sau organizațiile care lucrează în domeniul eticii IA (de exemplu, reprezentanții guvernului local, grupurile de advocacy). Cercetați activitatea acestora și identificați domeniile în care vă puteți face auzită vocea. Luați în considerare posibilitatea de a scrie o scrisoare reprezentantului dvs. local în care să vă exprimați îngrijorarea cu privire la o anumită aplicație a IA sau să pledați pentru măsuri de transparență și responsabilitate.

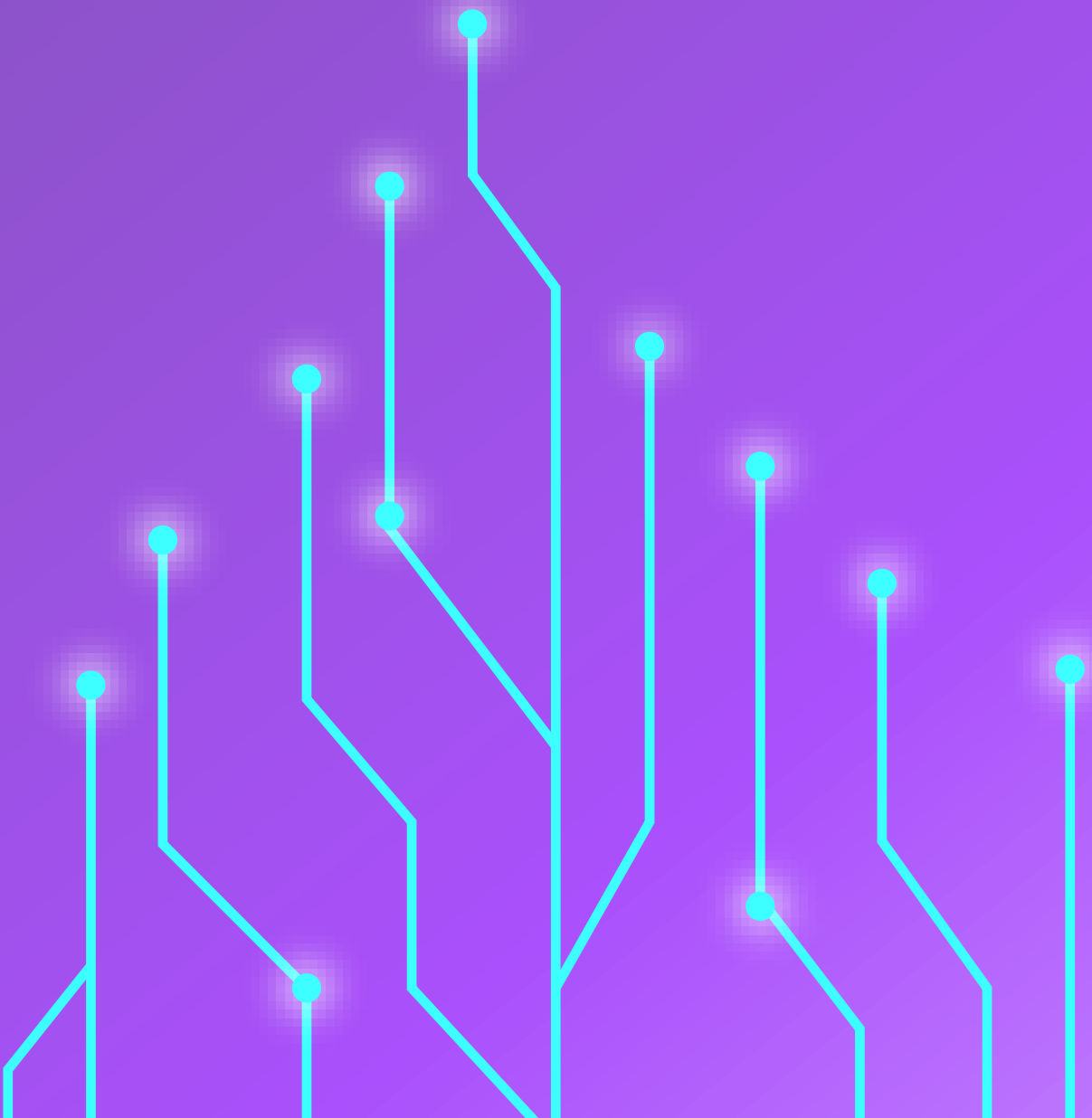
- **Participați la conversație:** Explorați forumuri online, discuții și grupuri de social media axate pe etica IA. Participați la discuții respectuoase, împărtășiți-vă învățămintele de la acest curs și învățați din perspectivele altora. Luați în considerare posibilitatea de a vă crea propriul blog sau propria platformă online pentru a vă împărtăși gândurile și pentru a crește gradul de conștientizare cu privire la etica IA.

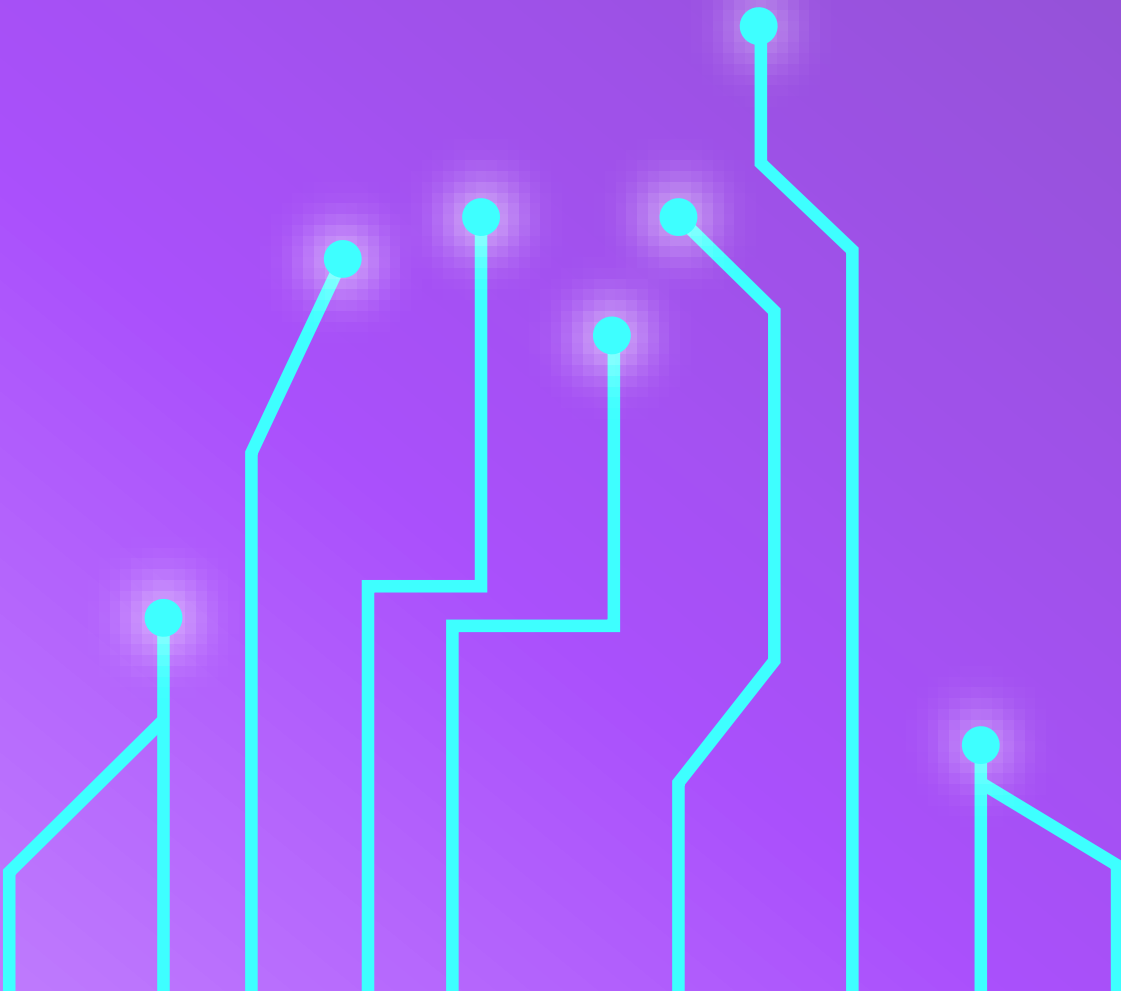




06. Concluzii

CU6 | Etica IA, o abordare practică





06. Concluzie

De-a lungul acestui curs, am explorat puterea de transformare a IA.

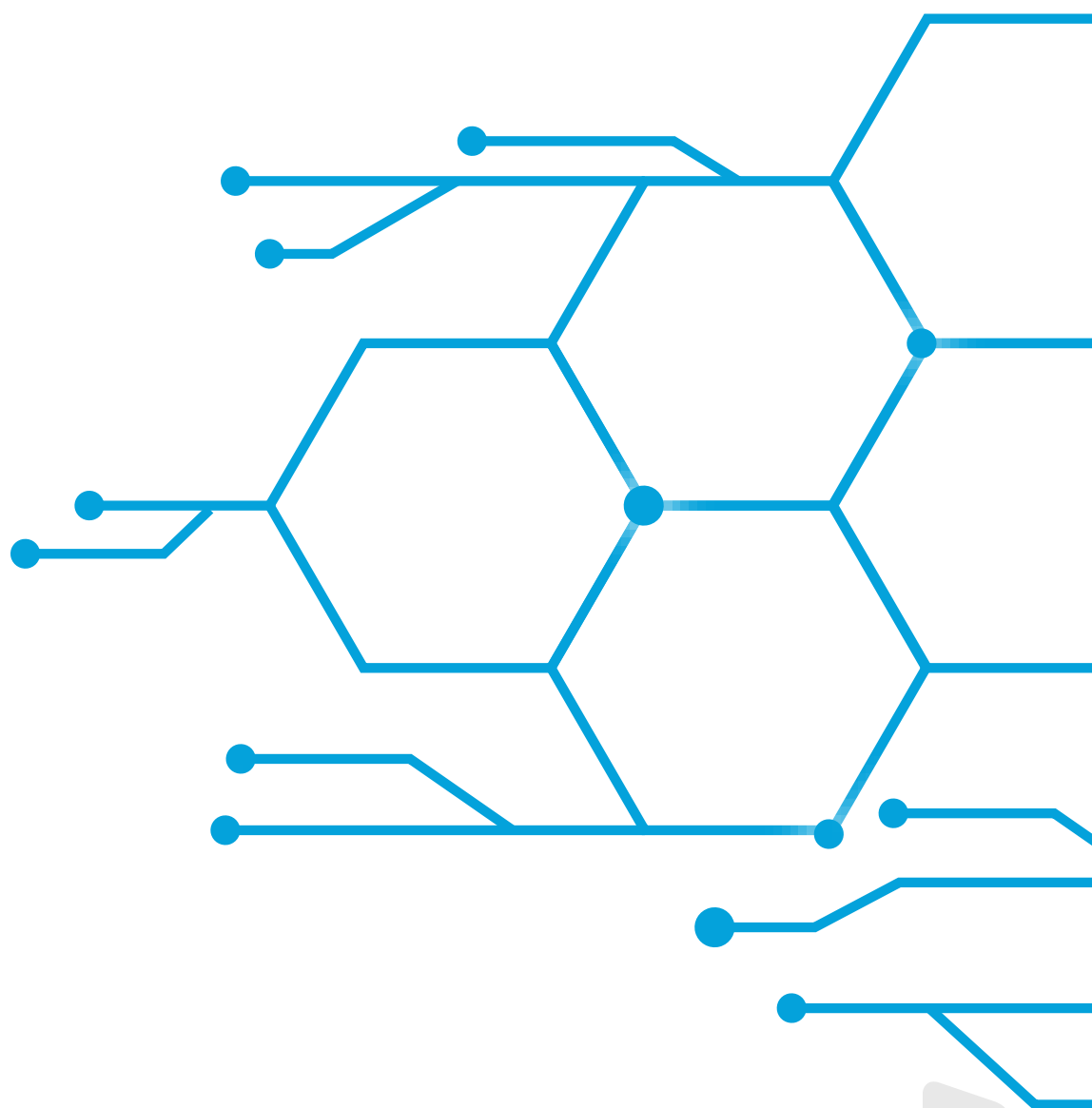
Am început prin a examina natura insidioasă a prejudecăților algoritmice, care este o lentilă critică prin care ar trebui privită întreaga dezvoltare a IA. În unitatea 2, am aprofundat conceptul de non-maleficență, care asigură faptul că sistemele AI nu provoacă daune și sunt concepute având în vedere siguranța. Unitatea 3 a explorat principiul crucial al responsabilității, subliniind necesitatea unor linii clare de responsabilitate în dezvoltarea și implementarea IA. Transparența, care a constituit punctul central al unității 4, asigură faptul că putem înțelege cum funcționează sistemele AI și că putem lua decizii în cunoștință de cauză cu privire la utilizarea acestora. În unitatea 5, am abordat legătura vitală dintre IA și drepturile omului, subliniind importanța echității și a incluziunii în dezvoltarea IA. În cele din urmă, în Unitatea 6, v-am echipat cu un cadru practic pentru elaborarea propriilor orientări etice privind inteligența artificială, permițându-vă să transpuneți principiile în acțiune.

Pe măsură ce ne aflăm în pragul unui viitor condus de inteligența artificială, considerentele etice pe care le-am explorat în acest curs nu sunt simple exerciții academice; ele reprezintă fundamentul dezvoltării responsabile a inteligenței artificiale.



Prin implicarea activă în discuții despre etica inteligenței artificiale, prin susținerea practicilor responsabile și prin tragerea la răspundere a dezvoltatorilor și a factorilor de decizie, ne putem asigura că inteligența artificială servește omenirii într-un mod corect, echitabil și benefic.

Viitorul IA nu este predeterminat; este un viitor pe care îl vom crea împreună. Haideți să ne străduim să construim un viitor în care inteligența artificială devine o forță puternică pentru bine, promovând progresul, bunăstarea și un viitor mai luminos pentru toți.





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



helixconnect



2022-1-ES01-KA220-HED-000085257