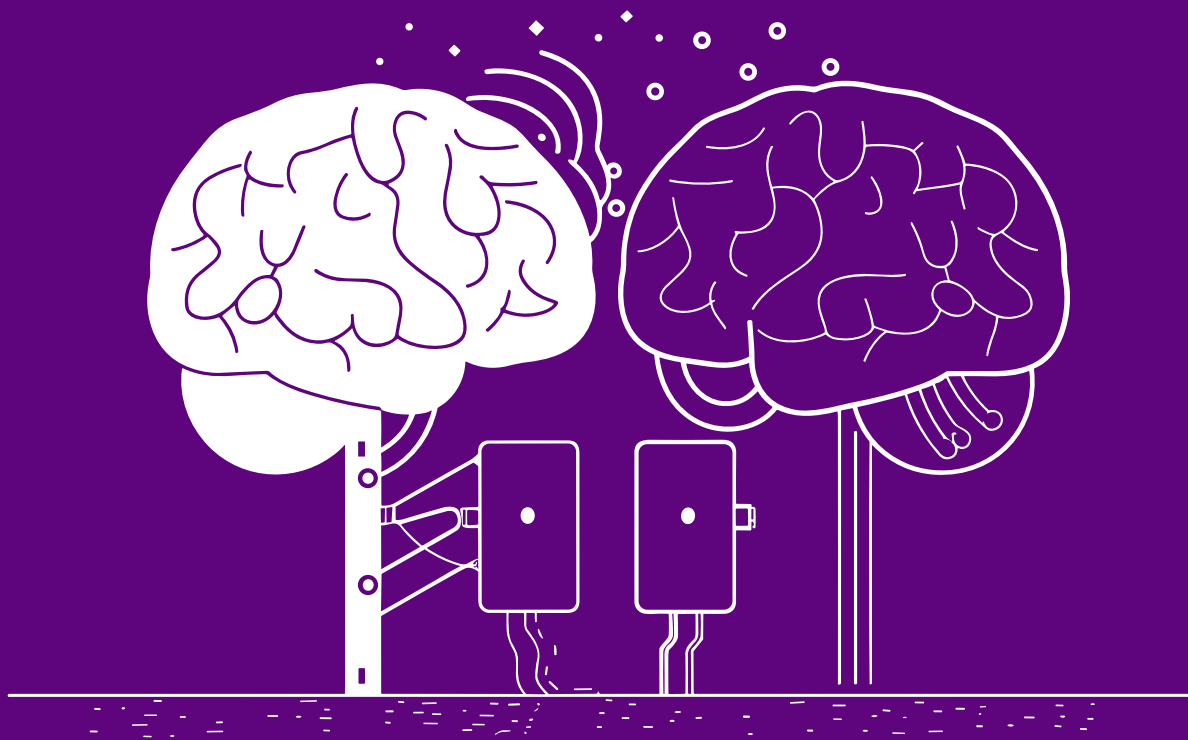


Ghid Etic pentru Inteligența Artificială în Sectoarele de Comunicare și Informare

Cum să punem întrebări morale și cum să decidem ce este corect în era IA



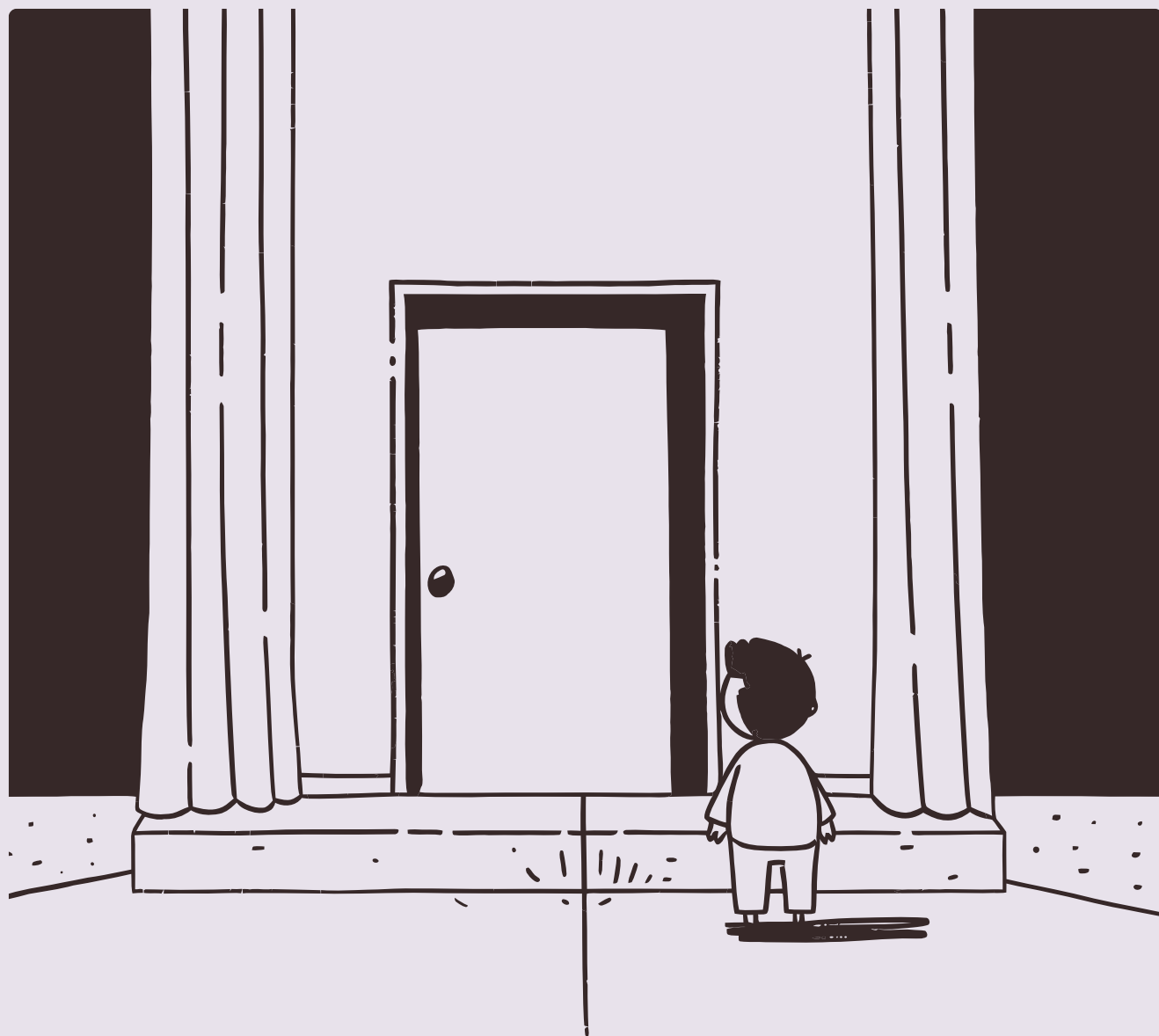
Co-funded by
the European Union

Sinteză: Ghid pentru Etica IA

CAPITOLUL I: Introducere	3
1.1 Scopul Liniilor Directoare	4
1.2 Domeniu și Destinatari	5
1.3 Fundamentele Liniilor Directoare	6
Capitolul II: Principii Etice Fundamentale	7
2.1 Introducere	8
2.2 Transparența în IA	10
2.3 Responsabilitate	12
2.4 Acuratețe	15
2.5 Confidențialitate	18
2.6 Proprietate Intelectuală	20
Capitolul III: Aplicații Practice	23
3.1 Utilizarea IA în Crearea de Conținut	24
3.2 IA pentru Jurnalism și Media	28
3.3 IA pentru Analiza Datelor și Cercetare	30
Capitolul IV: Listă de Verificare pentru Utilizarea Etică a IA	32
4.2. Scenarii de Exemplu	34
Capitolul V: Implementare și Monitorizare	38
5.1 Pași pentru Adoptarea Liniilor Directoare	39
5.2 Monitorizare și Conformitate	41
Anexe	43
A1 – Termeni Cheie și Definiții	44
A2 – Referințe Legale și Politice	46

CAPITOLUL I

Introducere



1.1 Scopul Ghidului

De ce este importantă utilizarea etică a IA în sectoarele de comunicare și informare?

Scopul acestui ghid este de a oferi un cadru pentru integrarea etică a Inteligenței Artificiale (IA) în sectoarele de comunicare și informare, asigurând alinierea cu principiile definite în "Recomandarea privind Etica Inteligenței Artificiale" a UNESCO (2021). Dezvoltat ca parte a proiectului Erasmus+ "Me, Myself, and AI", acest ghid își propune să ofere profesioniștilor cunoștințele și instrumentele necesare pentru a aborda în mod responsabil provocările etice complexe prezentate de tehnologiile IA.

Abordând domeniile recomandate precum transparența, responsabilitatea, confidențialitatea și incluziunea, ghidul urmărește să:

Capacitate profesională

Să integreze IA în mod etic în fluxurile de lucru, înțelegând când utilizarea acesteia este benefică și când ar trebui evitată.

Promoveze încrederea și responsabilitatea

Să se străduiască să susțină adevărul într-o societate caracterizată de o neîncredere crescândă în practicile bazate pe IA.

Protejeze demnitatea umană

Să evite practicile discutabile care afectează viețile umane, protejând drepturile omului și sustenabilitatea mediului.

Informată de abordarea UNESCO privind etica IA, acest ghid reflectă importanța:

- Recunoașterii impacturilor sociale profunde ale IA, atât pozitive, cât și negative.
- Asigurării că IA favorizează prosperitatea umană minimizând riscuri precum discriminarea, inegalitatea și dezinformarea.
- Încurajării reflecției etice pe tot parcursul ciclului de viață al IA, de la dezvoltare la implementare și mai departe.

După cum subliniază UNESCO, abordarea preocupărilor etice nu inhibă inovația, ci mai degrabă oferă un cadru pentru a asigura că tehnologiile IA contribuie la justiție, incluziune și dezvoltare durabilă. Aderând la aceste principii, proiectul Erasmus+ "Me, Myself, and AI" aspiră să ofere un exemplu în promovarea utilizării responsabile a IA în sectoarele de comunicare și informare.

"Comunicarea și informarea, întrucât tehnologiile IA joacă un rol din ce în ce mai important în procesarea, structurarea și furnizarea informațiilor; problemele de jurnalism automatizat și furnizare algoritmică de știri și moderarea și curățarea conținutului pe rețelele sociale și motoarele de căutare sunt doar câteva exemple care ridică probleme legate de accesul la informații, dezinformare, informare greșită, discurs de ură, apariția unor noi forme de narațiuni sociale, discriminare, libertate de exprimare, confidențialitate și alfabetizare media și informațională, printre altele." (Recomandarea privind Etica Inteligenței Artificiale, 2021, p.11)

1.2 Domeniu de aplicare și destinatari

Aceste ghiduri etice sunt concepute pentru a sprijini profesioniștii din sectoarele de comunicare și informare în înțelegerea și aplicarea responsabilă a tehnologiilor IA. Ca parte a proiectului Erasmus+ "Me, Myself, and AI", documentul se adresează principalilor actori ale căror roluri se intersectează cu instrumentele bazate pe IA, asigurându-se că aceștia sunt echipați pentru a lua decizii informate și etice în practicile lor.

Ghidurile se concentrează pe:

- Integrarea principiilor etice în aplicațiile IA pentru crearea de conținut, jurnalism, design grafic și educație media.
- Abordarea provocărilor legate de transparență, responsabilitate, confidențialitate, proprietate intelectuală și acuratețe în utilizarea IA.
- Furnizarea unui cadru practic și accesibil pentru evaluarea și îmbunătățirea practicilor etice ale IA.

Ghidurile sunt destinate pentru:



Profesioniști în Comunicare și Informare

Jurnaliști, creatori de conținut, designeri grafici și fotografi care utilizează în mod regulat IA pentru a-și îmbunătăți fluxurile de lucru și rezultatele.



Educatori și Formatori pentru Adulți

Profesioniști care predau alfabetizarea digitală și etica IA și care pot încorpora aceste ghiduri în curriculumul lor pentru a forma următoarea generație de utilizatori etici ai IA.



Legiuitori și Factori de Decizie

Persoane responsabile pentru crearea de politici și reglementări privind utilizarea IA în media și comunicare.



Instituții Educaționale și ONG-uri

Școli, universități și organizații care se concentrează pe alfabetizarea media și tehnologia etică pot utiliza aceste ghiduri pentru a-și îmbogăți programele educaționale.



1.3 Fundamentele Liniilor Ghid

Liniile Ghid Etice pentru Inteligența Artificială în sectoarele Comunicării și Informației se bazează pe principiile și obiectivele proiectului Erasmus+ „Me, Myself, and AI”. Aceste fundamente reflectă un angajament de a alinia progresul tehnologic la considerațiile etice, pentru a promova un ecosistem digital resoinabil și incluziv.

Principiile Etice ale UNESCO

Liniile ghid se bazează pe „Recomandarea privind Etica Inteligenței Artificiale” a UNESCO, care oferă un cadru etic universal pentru utilizarea IA, subliniind demnitatea umană, echitatea și sustenabilitatea.

Obiectivele Cheie ale Proiectului

Aceste linii ghid sprijină obiectivele cheie ale proiectului:

- Creșterea gradului de conștientizare asupra tehnologiilor IA în rândul profesioniștilor din comunicare.
- Furnizarea de instrumente practice pentru integrarea etică a IA.
- Îmbunătățirea alfabetizării digitale și a rezilienței etice a profesioniștilor din domeniu.

Abordare Incluzivă și Accesibilă

Liniile ghid urmăresc să garanteze că tehnologiile IA sunt incluzive și accesibile, răspunzând nevoilor diferitelor categorii de utilizatori și părți interesate. O atenție deosebită a fost acordată protejării grupurilor marginalizate și asigurării unui acces echitabil la oportunitățile oferite de IA.

Provocări Etice Specifice Sectorului

Liniile ghid sunt adaptate pentru a aborda dilemele etice unice cu care trebuie să se confrunte profesioniștii din comunicare și informare, precum combaterea dezinformării, garantarea confidențialității în jurnalism și respectarea proprietății intelectuale în crearea de conținut.

Dezvoltare Colaborativă

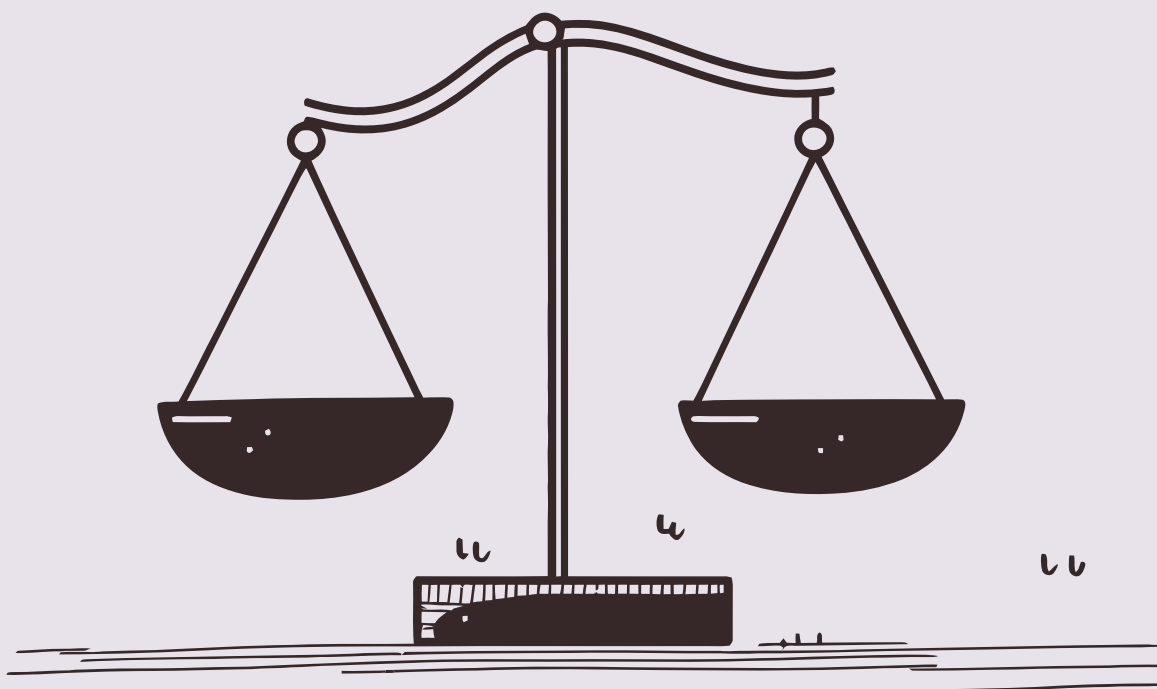
Elaborate prin consultări și contribuții ale experților în cadrul meselor rotunde europene ale proiectului, aceste linii ghid reflectă înțelepciunea colectivă a profesioniștilor, experților în etică, educatorilor și dezvoltatorilor de IA.

Instrumente pentru Aplicarea Practică

Pentru a garanta relevanța și aplicabilitatea, liniile ghid includ instrumente practice, precum o listă de verificare pentru utilizarea etică a IA și un videoclip animat pentru a simplifica și a disemina principiile cheie.

CAPITOLUL II

Principii Etice Fundamentale



2.1 Introducere

Avansarea rapidă a Inteligenței Artificiale (IA) prezintă atât oportunități, cât și provocări, în special în sectoarele comunicării și informării. În timp ce IA poate îmbunătăți eficiența, creativitatea și accesibilitatea, ridică și probleme etice complexe care ar trebui abordate pentru a asigura o **utilizare responsabilă**.

De ce principiile de Transparență, Responsabilitate, Acuratețe, Confidențialitate și Proprietate Intelectuală **sunt importante**?

Consideră următoarele exemple și explorează punându-ți întrebările de sub fiecare exemplu. Ține cont că nu există răspunsuri corecte sau greșite, doar răspunsuri etice.

1

Transparență

O organizație jurnalistică folosește IA pentru a genera articole, dar fără a dezvălui acest lucru cititorilor, riscă să erodeze încrederea. Fiind transparenți cu privire la implicarea IA, ar putea susține standardele etice și promova credibilitatea.

- + Care sunt implicațiile erodării încrederii ca organizație jurnalistică?
- + Care ar putea fi impactul asupra organizației dacă ar fi transparentă?
- + Ce ai face dacă ai fi un manager al unei astfel de organizații?

2

Responsabilitate

Un designer grafic utilizează un instrument de IA pentru a crea elemente vizuale pentru un client. Dacă IA produce rezultate nepotrivite sau părtinitoare, designerul ar trebui să își asume responsabilitatea de a aborda aceste erori, mai degrabă decât să le atribuie exclusiv tehnologiei.

- + Care sunt riscurile creării unor rezultate nepotrivite sau părtinitoare?
- + În opinia ta, care este cel mai bun mod de a utiliza IA într-un proiect creativ?

3

Acuratețe

Un jurnalist utilizează IA pentru verificarea faptelor, dar nu verifică manual rezultatele. Acest lucru ar putea duce la răspândirea dezinformării dacă rezultatele IA sunt defectuoase. Susținerea acurateței asigură că IA îmbunătățește, mai degrabă decât compromite, calitatea conținutului.

- + Care sunt consecințele răspândirii dezinformării?
- + În opinia ta, care este cel mai bun mod de a asigura calitatea conținutului?

4

Confidențialitate

Instrumentele de IA care analizează comportamentul publicului pentru recomandări de conținut trebuie să protejeze datele utilizatorilor. Nerespectarea poate duce la încălcări ale încrederii și ale drepturilor la confidențialitate.

- + Care sunt consecințele unei încălcări a datelor?
- + În opinia ta, care este cel mai bun mod de a gestiona problemele legate de drepturile de confidențialitate?

5

Proprietate Intelectuală

Un creator de conținut folosește IA pentru a remixa opere existente fără o atribuire adecvată, ridicând preocupări etice și legale. Respectarea proprietății intelectuale asigură utilizarea corectă și recunoașterea pentru creatorii originali.

- + Cum sunt afectați creatorii de conținut de atribuirea corectă a muncii lor?
- + De ce este important să respecti creația altcuiva?

Etica ca alegere personală:

Adoptarea IA nu este doar o provocare tehnologică, ci una profund personală și etică. Modul în care fiecare profesionist alege să utilizeze IA - fie ca instrument pentru a-și spori competențele, fie ca înlocuitor al creativității și judecății umane - reflectă valorile și prioritățile sale individuale.

De exemplu:



- Un fotograf ar putea folosi IA pentru a îmbunătăți imaginile, asigurându-se că deciziile artistice finale rămân ale sale, tratând IA ca un **instrument de suport**.
- În schimb, un alt profesionist s-ar putea baza în întregime pe IA pentru a automatiza crearea de conținut, riscând pierderea responsabilității personale și a aportului creativ, poziționând astfel IA ca un **înlocuitor**.

Poziția etică a fiecărui profesionist determină cum să echilibreze beneficiile IA cu potențialele sale riscuri. Ghidurile din acest capitol au ca scop să capaciteze indivizii să ia decizii informate și chibzuite care să se alinieze cu valorile lor personale și profesionale.

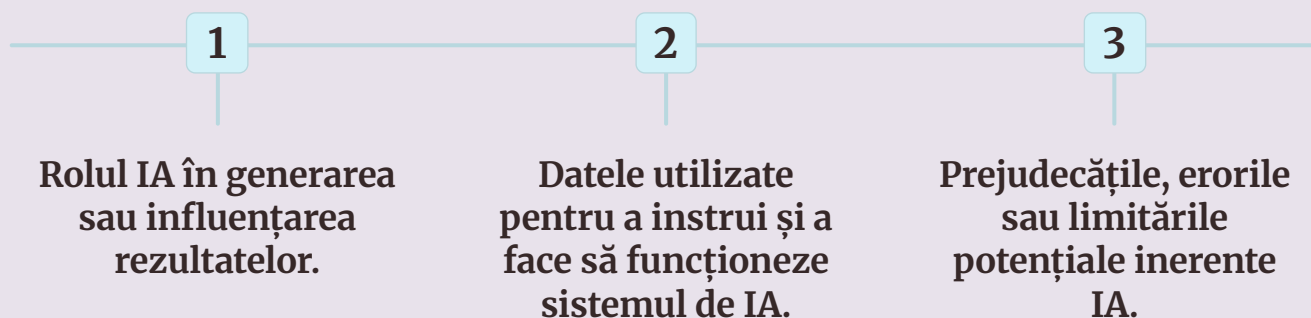
Fiecare principiu este explorat în profunzime, oferind îndrumare practică și strategii pentru abordarea provocărilor unice prezentate de IA. De la asigurarea unei comunicări clare privind rolul IA în crearea de conținut până la protejarea datelor personale și respectarea proprietății intelectuale, acest capitol oferă un cadru cuprinzător pentru luarea deciziilor etice.

2.2 Transparența în IA

Transparența se referă la dezvăluirea clară și deschisă a modului în care sunt utilizate tehnologiile de IA, care sunt limitele lor și procesele care stau la baza rezultatelor lor. Într-o lume în care IA influențează din ce în ce mai mult procesul decizional și crearea de conținut, menținerea transparenței asigură încredere, responsabilitate și o implicare informată cu sistemele de IA.

Ce este Transparența în IA?

Transparența în IA implică furnizarea către părțile interesate—fie că sunt cititori, utilizatori sau clienți—informații clare și accesibile despre:



De ce este importantă Transparența?

Construirea Încrederii

Când părțile interesate sunt conștiente de modul în care este utilizată IA, este mai probabil să aibă încredere în rezultate.

De exemplu, un organ de presă care declară în mod clar când sunt utilizate instrumente de IA pentru a crea titluri sau a genera articole își construiește credibilitatea în fața publicului său.

Promovarea Responsabilității

Transparența permite organizațiilor să își asume responsabilitatea pentru rezultatele IA, asigurând că supravegherea umană rămâne o parte integrantă.

Dacă un instrument de IA creează conținut distorsionat sau inexact, un proces transparent face mai ușoară identificarea și corectarea problemei.

Încurajarea Practicilor Etice

Transparența încurajează utilizarea etică a IA, solicitând organizațiilor să abordeze și să dezvăluie prejudecăți, surse de date și riscuri potențiale.

Capacitarea Părților Interesate:

Utilizatorii pot lua decizii mai informate când înțeleg cum IA le influențează interacțiunile cu conținutul

Provocări în Realizarea Transparenței

1 Complexitatea Sistemelor IA

Multe sisteme de IA funcționează ca „cutii negre”, făcând dificilă explicarea modului în care ajung la anumite rezultate. Simplificarea acestor informații fără a le trivializa reprezintă o provocare semnificativă.

2 Tehnologii Proprietare

Organizațiile pot fi reticente în a divulga procesele IA din cauza preocupărilor legate de proprietatea intelectuală, limitând potențial transparența.

3 Prejudecăți în Datele de Antrenament

Transparența poate evidenția prejudecățile prezente în datele de antrenament sau în algoritmi, care ar putea reflecta negativ asupra organizației.

Cum să Implementăm Transparența în Practică



Declarați Implicarea IA

Etichetați clar conținutul generat de IA. De exemplu, dacă un articol, o grafică sau un videoclip a fost creat parțial sau complet de IA, acest lucru ar trebui declarat în mod explicit.



Explicați Procesele IA în Termeni Simpli

Utilizați un limbaj simplu pentru a descrie modul în care funcționează instrumentele IA și nivelul de supraveghere umană implicat. Evitați jargonul tehnic care ar putea aliena sau confuza utilizatorii.



Evidențiați Limitările

Fiți sinceri cu privire la potențialele slăbiciuni sau prejudecăți ale sistemelor IA. De exemplu, declarați dacă IA are o înțelegere limitată a anumitor subiecte sau se bazează pe date învechite.



Activați Mecanisme de Feedback

Permiteți utilizatorilor să raporteze erori sau prejudecăți pe care le observă în rezultatele generate de IA, încurajând o abordare colaborativă pentru îmbunătățirea transparenței.

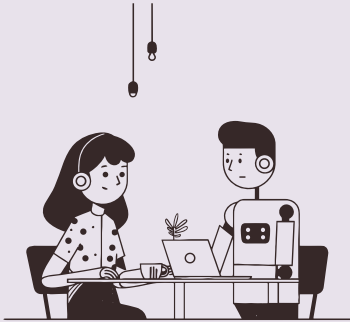


Verificați și Documentați Sistemele IA:

Mențineți o documentație internă despre modul în care sistemele IA sunt implementate și monitorizate, asigurând un lanț clar de responsabilitate.

Studiu de caz: IA în jurnalism

O companie media implementează un instrument de IA pentru a rezuma articole lungi de știri pentru consultare rapidă. Pentru a asigura transparența:



- Fiecare rezumat include o declarație: "Acest rezumat a fost generat folosind IA cu supraveghere editorială."
- Compania oferă o pagină web care explică modul în care funcționează instrumentul de IA, inclusiv datele sale de antrenament și limitările sale.
- Linkuri pentru feedback sunt incluse cu rezumatele, permițând cititorilor să raporteze inexactități sau să solicite clarificări.

Această abordare nu numai că informează publicul, dar demonstrează și angajamentul organizației față de utilizarea etică a IA.

Echilibrarea transparenței și a caracterului practic

Deși transparența este esențială, aceasta trebuie echilibrată cu considerații practice, cum ar fi menținerea cunoștințelor proprietare și gestionarea supraîncărcării de informații pentru utilizatori. Obiectivul este de a oferi o transparență semnificativă fără a supraîncărca părțile interesate cu detalii tehnice excesive.

Privind spre viitor

Transparența nu este doar o linie directoare, ci o practică dinamică care evoluează odată cu avansarea tehnologiilor IA. Organizațiile și profesioniștii trebuie să evalueze și să actualizeze continuu măsurile lor de transparență pentru a reflecta noile evoluții în IA.

Acordând prioritate transparenței, profesioniștii din sectoarele de comunicare și informare pot promova încrederea, susține standardele etice și ghida integrarea responsabilă a IA.

2.3 Responsabilitatea

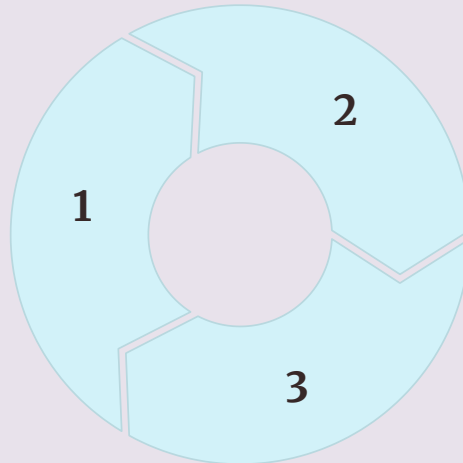
Responsabilitatea este un principiu fundamental pentru asigurarea utilizării etice a Inteligenței Artificiale (IA) în sectoarele de comunicare și informare. Se concentrează pe atribuirea clară a responsabilității pentru sistemele de IA și rezultatele acestora, asigurând că supravegherea umană rămâne o parte integrantă pe tot parcursul ciclului de viață al IA.

Ce este Responsabilitatea în IA?

Responsabilitatea în IA implică atribuirea clară și precisă a obligațiilor și îndatoririlor către persoane, echipe sau organizații specifice pentru:

Proiectarea și Implementarea IA

Dezvoltarea sistemelor de IA care respectă principiile etice, oferă transparență și sunt responsabile față de utilizatori și societate.



Monitorizarea și Gestionarea Impacturilor

Supravegherea cu atenție a apariției erorilor, prejudecăților sau consecințelor negative - intervenind prompt și eficient atunci când este necesar.

Menținerea Standardelor Etice

Păstrarea și consolidarea principiilor de responsabilitate pe tot parcursul ciclului de viață al IA, de la concepție până la retragere.

De ce este Importantă Responsabilitatea

Asigurarea că sistemele de IA funcționează cu transparență, sunt de încredere și rămân aliniate cu valorile umane este esențială. Responsabilitatea constituie fundamentul indispensabil pentru dezvoltarea și implementarea etică a IA.

1. Garantează Supravegherea Umană:

Sistemele de IA, în ciuda sofisticării lor, nu pot înlocui judecata umană în domenii critice, în special în deciziile care afectează drepturile fundamentale sau bunăstarea oamenilor.

2. Promovează Încrederea:

Utilizatorii tind să aibă mai multă încredere în conținutul generat sau în deciziile susținute de IA atunci când au certitudinea cu privire la cine supraveghează acuratețea, echitatea și integritatea acestora.

3. Promovează Practici Etice:

Responsabilitatea asigură respectarea unor standarde etice riguroase, în special în abordarea și prevenirea consecințelor neintenționate, cum ar fi prejudecățile algoritmice, răspândirea dezinformării sau încălcarea confidențialității utilizatorilor.

4. Atenuază Riscurile:

O responsabilitate clar definită și atribuită garantează intervenții prompte pentru corectarea erorilor sau problemelor, minimizând astfel potențialele daune rezultate din output-uri inadecvate sau defecțiuni ale sistemelor de IA.

Provocări în Realizarea Responsabilității

1. Ambiguitate în Roluri:

În proiectele colaborative de IA care implică mai multe părți interesate, identificarea responsabililor poate fi dificilă.

2. Complexitatea Sistemelor de IA:

Sistemele de IA pot funcționa ca „cutii negre”, făcând dificilă identificarea sursei erorilor sau prejudecăților.

3. Avansul Tehnologic Rapid:

Pe măsură ce IA evoluează, cadrele etice și legale ar putea rămâne în urmă, creând lacune în mecanismele de responsabilitate.

Cum să Implementăm Responsabilitatea în Practică

1. Definirea Rolurilor și Responsabilităților	Delimitarea clară a responsabililor pentru fiecare etapă a ciclului de viață al IA, de la dezvoltare până la monitorizarea post-implementare.
2. Stabilirea unor Cadre Clare de Governanță	Crearea de politici și protocoale pentru utilizarea etică a IA, asigurând că toate părțile interesate sunt conștiente de responsabilitatea lor.
3. Încorporarea Supravegherii Umane	Solicitarea revizuirii umane a conținutului generat de IA, în special în domeniile sensibile precum reportajul jurnalistic, design-ul grafic sau procesul decizional.
4. Efectuarea de Audituri Regulate	Realizarea de evaluări de rutină a sistemelor de IA pentru a asigura conformitatea cu standardele etice și pentru a identifica riscurile potențiale.
5. Furnizarea de Mecanisme Transparente de Feedback	Permiterea utilizatorilor să raporteze erori sau prejudecăți și asigurarea unor răspunsuri prompte pentru a aborda preocupările acestora.
6. Conformitate Legală și Etică	Asigurarea aderenței la legile internaționale (vezi Anexa A2) și la reglementările naționale care guvernează utilizarea IA.

Studiu de caz: IA în Generarea Automatizată de Știri

Revizuire Editorială

Compania desemnează o echipă editorială pentru a revizui și aproba tot conținutul generat de IA înainte de publicare.

Gestionarea Erorilor

Erorile din conținutul generat de IA sunt înregistrate, iar o echipă desemnată este responsabilă pentru rezolvarea acestora.

Politica de Transparență

Compania implementează o politică transparentă care afirmă: "Toate articolele generate de IA sunt revizuite și editate de echipa noastră editorială pentru a asigura acuratețea și echitatea."

Echilibrarea Responsabilității și Inovației

Deși responsabilitatea este fundamentală, aceasta nu ar trebui să împiedice inovația. Cadrele flexibile care se adaptează tehnologiilor emergente și permit spațiu pentru experimentare sunt esențiale. Obiectivul este găsirea unui echilibru între promovarea inovației și asigurarea responsabilității etice.

Privind spre Viitor

Responsabilitatea trebuie să rămână centrală în orice proces ghidat de IA. Prin definirea clară a rolurilor, menținerea supravegherii umane și promovarea transparenței, organizațiile din sectoarele de comunicare și informare pot asigura că IA servește ca instrument pentru progres, menținând în același timp standardele etice.

2.4 Acuratețe

Acuratețea este un principiu etic fundamental în utilizarea Inteligenței Artificiale (IA), în special în sectoarele de comunicare și informare, unde difuzarea informațiilor fiabile și factuale este de importanță primordială. Asigurarea acurateței înseamnă că sistemele de IA produc în mod constant rezultate care sunt precise, imparțiale și aliniate cu scopurile lor prevăzute. Acest principiu stă la baza încrederii, minimizează dezinformarea și asigură că instrumentele de IA contribuie pozitiv la societate.

Ce este Acuratețea în IA?

Acuratețea se referă la capacitatea sistemelor de IA de a produce rezultate—fie că sunt decizii, previziuni sau conținut—care sunt corecte, relevante și aliniate cu realitatea. Aceasta necesită atât precizie tehnică (de ex., algoritmi corecți și date fiabile), cât și considerații etice (de ex., evitarea daunelor cauzate de inexactități).

De ce este importantă acuratețea?



Păstrarea credibilității

Rezultatele precise ale IA consolidează credibilitatea organizațiilor și a persoanelor care se bazează pe aceste instrumente.



Prevenirea daunelor

Rezultatele inexacte ale IA pot duce la dezinformare, daune reputaționale sau chiar vătămări fizice, în special în sectoare precum sănătatea sau jurnalismul.



Promovarea standardelor etice

Asigurarea acurateței se aliază cu angajamente etice mai largi față de adevăr și echitate.



Cultivarea încrederii în sistemele IA

Încrederea în sistemele IA se bazează pe capacitatea lor de a oferi rezultate coerente și corecte.

Provocări în realizarea acurateței

1. Date distorsionate sau incomplete: Sistemele IA antrenate pe seturi de date distorsionate sau incomplete pot produce rezultate inexacte sau părtinitoare.
2. Complexitatea sistemelor IA: Unele modele IA, în special sistemele de învățare profundă, funcționează ca "cutii negre", făcând dificilă identificarea și corectarea inexactităților.
3. Medii dinamice din lumea reală: Sistemele IA implementate în medii în continuă schimbare pot avea dificultăți în menținerea acurateței în timp fără actualizări frecvente.
4. Obiective nealiniate: Dacă un sistem IA este optimizat pentru viteză sau cost, mai degrabă decât pentru acuratețe, rezultatele sale ar putea să nu îndeplinească standardele etice.

Cum să asigurăm acuratețea în practicile IA

Calitatea și validarea datelor

Utilizarea unor seturi de date diverse și de înaltă calitate pentru antrenarea modelelor IA. Validarea continuă a datelor pentru a evita inexactitățile cauzate de informații învechite sau distorsionate.

Transparența și testarea algoritmilor

Implementarea unor protocoale riguroase de testare pentru evaluarea și îmbunătățirea acurateței algoritmilor IA. Verificarea regulată a sistemelor IA pentru erori sau prejudecăți.

Supravegherea umană

Implicarea profesioniștilor umani în verificarea rezultatelor generate de IA, în special în aplicații sensibile precum jurnalismul, crearea de conținut sau luarea deciziilor.

Mecanisme de Feedback

Stabilirea unor canale care să permită utilizatorilor să raporteze inexactități în rezultatele IA, asigurând corecții prompte.

Monitorizarea Performanței

Monitorizarea continuă a acurateții sistemelor de IA în contexte reale, adaptându-le după necesitate pentru a menține performanța.

Comunicarea Clară a Limitelor

Informarea utilizatorilor despre rata de acuratețe a sistemului IA, posibile distorsiuni și zone de incertitudine pentru a gestiona așteptările

Studiu de caz: IA în Verificarea Faptelor

O organizație de verificare a faptelor folosește un sistem de IA pentru a identifica și evalua potențiale afirmații false online. Pentru a asigura acuratețea:

- Modelul IA este antrenat pe un set de date diversificat de afirmații verificate și dezinformare.
- Rezultatele sunt examinate de verificatori umani înainte de publicare.
- Sistemul include un ciclu de feedback prin care utilizatorii pot raporta erori, care sunt apoi utilizate pentru a perfecționa modelul IA.
- Un avertisment însoțește toate rezultatele asistate de IA: "Acest rezultat a fost generat cu IA și verificat de echipa noastră editorială."

Echilibrarea Acurateții cu Alte Principii

Deși acuratețea este esențială, uneori trebuie echilibrată cu alte principii precum confidențialitatea sau explicabilitatea. De exemplu, creșterea transparenței prin explicarea deciziilor unui sistem IA ar putea expune date sensibile, așadar compromisurile trebuie gestionate cu atenție pentru a evita probleme etice.

Privind spre Viitor

Acuratețea este un principiu dinamic care necesită atenție continuă pe măsură ce tehnologiile IA și aplicațiile lor evoluează. Angajându-se pentru acuratețe, organizațiile din sectoarele de comunicare și informare pot asigura că sistemele lor IA susțin practici etice, promovează încrederea publică și contribuie la dezvoltarea responsabilă a IA.

2.5 Confidențialitate

Confidențialitatea este un pilon fundamental al utilizării etice a Inteligenței Artificiale (IA), în special în sectoarele de comunicare și informare unde datele sunt centrale pentru operațiuni. Asigurarea confidențialității implică protejarea informațiilor personale și sensibile ale indivizilor împotriva accesului neautorizat, utilizării necorespunzătoare sau exploatării. Acest principiu este esențial pentru menținerea încrederii, apărarea demnității umane și respectarea standardelor legale și etice.

Ce este Confidențialitatea în IA?

Confidențialitatea în IA se referă la gestionarea responsabilă a datelor pe parcursul întregului ciclu de viață al IA, de la colectare și procesare până la stocare și eliminare. Aceasta include:



- Protejarea informațiilor personale și sensibile.
- Asigurarea că utilizarea datelor este în conformitate cu consimțământul informat.
- Minimizarea colectării și păstrării datelor care nu sunt necesare.

De exemplu, un sistem de recomandare bazat pe IA pentru conținut trebuie să analizeze doar datele pe care utilizatorii au acceptat în mod explicit să le împărtășească și să se asigure că acestea sunt anonimizate acolo unde este cazul.

De ce este importantă Confidențialitatea?

1. Susține Încrederea	Utilizatorii sunt mai predispuși să folosească servicii bazate pe IA dacă simt că datele lor sunt gestionate în mod responsabil.
2. Protejează Demnitatea Umană	Protejarea confidențialității previne utilizarea necorespunzătoare a informațiilor personale care ar putea dăuna indivizilor sau grupurilor.
3. Se Aliniază Obligațiilor Legale	Multe regiuni au legi stricte privind confidențialitatea (de exemplu, GDPR în UE), făcând conformitatea esențială pentru operațiuni legale și etice.
4. Reduce Riscul de Prejudicii	Protejarea datelor minimizează riscul de furt de identitate, supraveghere sau alte rezultate dăunătoare.

Provocări în Asigurarea Confidențialității

1. **Dependența de Date în IA:** Multe sisteme de IA se bazează pe seturi mari de date, crescând riscul de încălcare a confidențialității dacă datele sunt gestionate necorespunzător.
2. **Prejudecăți în Colectarea Datelor:** Colectarea doar a datelor parțiale poate exclude sau reprezenta greșit anumite grupuri, afectându-le confidențialitatea.
3. **Evoluția Tehnologică Rapidă:** Tehnologiile emergente pot depăși legile și cadrele de reglementare existente privind confidențialitatea, creând lacune de reglementare.
4. **Fluxuri Globale de Date:** Diferențele în reglementările internaționale privind confidențialitatea complică conformitatea pentru organizațiile care operează peste granițe.

Cum să Asigurăm Confidențialitatea în Practicile de IA

Adoptarea Confidențialității prin Design

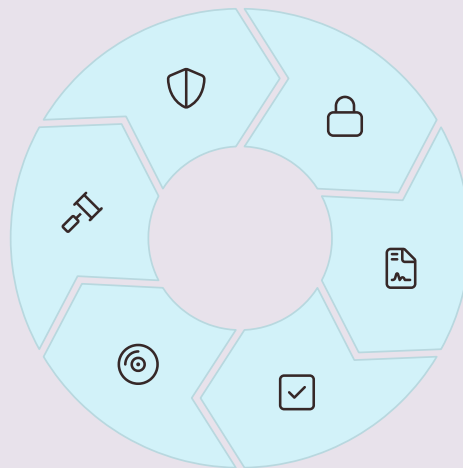
Integrarea măsurilor de confidențialitate în sistemele de IA de la început, cum ar fi anonimizarea datelor și limitarea colectării de date inutile.

Asigurarea Conformității cu Reglementările

Alinierea practicilor privind datele cu legile aplicabile, cum ar fi GDPR, și actualizarea proceselor pe măsură ce reglementările evoluează.

Efectuarea Evaluărilor de Impact asupra Confidențialității (PIAs)

Evaluarea regulată a riscurilor de confidențialitate ale sistemelor de IA și implementarea măsurilor pentru a le atenua.



Stocarea Securizată a Datelor

Utilizarea criptării și a altor măsuri de securitate pentru a proteja datele în repaus și în tranzit.

Implementarea Politicilor Clare privind Datele

Stabilirea liniilor directoare pentru colectarea, procesarea, păstrarea și eliminarea datelor.

Obținerea Consimțământului Informat

Asigurarea că utilizatorii sunt conștienți de modul în care vor fi utilizate datele lor și au opțiunea de a refuza.

Studiu de Caz: IA în Publicitatea Țintită

O companie folosește IA pentru a analiza comportamentul utilizatorilor și a furniza reclame personalizate. Pentru a asigura confidențialitatea:

- Datele sunt anonimizate și criptate înainte de analiză.

- Utilizatorii trebuie să își dea acordul explicit pentru colectarea datelor, cu o explicație clară despre modul în care vor fi utilizate datele lor.
- O evaluare a impactului asupra confidențialității este efectuată anual pentru a identifica și aborda potențialele riscuri.
- Utilizatorii sunt dotați cu instrumente pentru a-și gestiona preferințele privind datele, inclusiv posibilitatea de a-și șterge propriile date.

Echilibrarea confidențialității cu inovația

Deși IA se bazează puternic pe date, măsurile de confidențialitate nu ar trebui să împiedice inovația. Găsirea unui echilibru implică dezvoltarea unor mecanisme robuste pentru protejarea confidențialității, permițând în același timp sistemelor de IA să funcționeze eficient. De exemplu, învățarea federată permite modelelor de IA să se antreneze pe date descentralizate fără a compromite confidențialitatea individuală.

Privind spre viitor

Confidențialitatea va rămâne un principiu dinamic și în evoluție pe măsură ce tehnologiile de IA avansează. Organizațiile trebuie să abordeze proactiv provocările emergente privind confidențialitatea pentru a promova încrederea și a-și demonstra angajamentul față de practicile etice în sectoarele de comunicare și informare.

2.6 Proprietatea Intelectuală

Proprietatea intelectuală (PI) în contextul Inteligenței Artificiale (IA) cuprinde drepturile și responsabilitățile legate de crearea, proprietatea și utilizarea conținutului și instrumentelor generate de IA. Pe măsură ce IA devine un instrument din ce în ce mai puternic pentru crearea de conținut în sectoarele de comunicare și informare, asigurarea respectării drepturilor de PI este crucială pentru promovarea inovației, echității și practicilor etice.

Ce este Proprietatea Intelectuală în IA?

Proprietatea intelectuală în IA se referă la:



- Proprietatea asupra operelor generate de IA, cum ar fi articole, grafică, muzică și videoclipuri.
- Asigurarea că instrumentele de IA nu încalcă drepturile de autor sau mărcile înregistrate ale operelor existente.
- Recunoașterea și atribuirea contribuțiilor creatorilor umani împreună cu rezultatele IA.

De ce este importantă Proprietatea Intelectuală?

1. Protejează Creativitatea și Inovația: Respectarea drepturilor de proprietate intelectuală asigură că creatorii și inovatorii sunt compensați echitabil și motivați să-și continue activitatea.
2. Previne Problemele Juridice: Organizațiile pot întâmpina procese sau sancțiuni pentru încălcarea drepturilor de proprietate intelectuală, făcând conformitatea esențială.
3. Menține Standardele Etice: Practicile etice în IA includ respectarea drepturilor altora, inclusiv cele legate de opere creative și invenții.
4. Promovează Încrederea: O abordare transparentă și respectuoasă a proprietății intelectuale construiește încredere cu părțile interesate, inclusiv creatori, clienți și publicul.

Provocări în Abordarea Proprietății Intelectuale cu IA

- Ambiguitatea Proprietății: Determinarea proprietății operelor generate de IA poate fi complexă, mai ales când mai multe părți contribuie la proces.
- Probleme cu Datele de Antrenament: Sistemele de IA antrenate pe material protejat de drepturi de autor pot produce involuntar rezultate care încalcă drepturile de proprietate intelectuală.
- Lipsa Cadrelor Juridice: Legile actuale privind proprietatea intelectuală din multe regiuni nu abordează complet conținutul generat de IA, creând incertitudine.
- Încălcarea Involuntară: Sistemele de IA pot replica sau imita involuntar conținut protejat de drepturi de autor.

Cum să Abordăm Proprietatea Intelectuală în Practicile de IA

- Definirea Clară a Proprietății: Stabilirea drepturilor de proprietate pentru conținutul generat de IA în contracte și politici.
- Asigurarea Utilizării Etice a Datelor de Antrenament: Utilizarea datelor din domeniul public, licențiate sau autorizate explicit pentru antrenarea IA.
- Atribuirea Contribuțiilor: Recunoașterea atât a contribuțiilor umane, cât și a celor ale IA în operele de colaborare.
- Efectuarea Evaluărilor de Risc privind Proprietatea Intelectuală: Evaluarea regulată a sistemelor de IA pentru potențiale încălcări ale proprietății intelectuale și implementarea măsurilor pentru atenuarea riscurilor.
- Alinierea cu Standardele Juridice: Menținerea la curent cu evoluția legilor privind proprietatea intelectuală și adaptarea politicilor pentru a rămâne conform.
- Implementarea Acordurilor de Licență: Utilizarea acordurilor de licență pentru a clarifica modul în care conținutul sau instrumentele generate de IA pot fi utilizate și partajate.

Studiu de Caz: IA în Compoziția Muzicală

O companie de producție muzicală folosește IA pentru a genera coloane sonore. Pentru a aborda preocupările legate de proprietatea intelectuală:

- Sistemul de IA este antrenat exclusiv pe muzică din domeniul public sau cu licență de utilizare.
- Contractele specifică faptul că firma deține drepturile asupra compozițiilor generate de IA.
- Compozitorii umani care colaborează cu IA primesc credit pentru contribuțiile lor.
- Compania efectuează audituri periodice pentru a asigura conformitatea cu reglementările privind proprietatea intelectuală.

Echilibrarea inovației și protecția PI

IA poate conduce la inovații incredibile, dar acest lucru nu ar trebui să se întâmple în detrimentul drepturilor de PI. Găsirea unui echilibru implică crearea unor cadre care să permită explorarea creativă, respectând în același timp drepturile creatorilor și aliniindu-se cu standardele legale și etice.

Privind spre viitor

Pe măsură ce tehnologiile de IA evoluează, proprietatea intelectuală va rămâne un domeniu cheie de atenție etică. Organizațiile trebuie să abordeze proactiv aceste provocări, contribuind la un mediu echitabil și inovator în sectoarele de comunicare și informare.

CAPITOLUL III

Aplicații Practice



Această secțiune a ghidurilor etice își propune să ofere perspective practice pentru profesioniștii din sectoarele de comunicare și informare cu privire la modul de integrare responsabilă a IA în fluxurile lor de lucru. Această secțiune a ghidurilor etice este concepută pentru a oferi perspective practice și orientări concrete profesioniștilor care lucrează în sectoarele de comunicare și informare. Își propune să doteze acești profesioniști cu cunoștințele și instrumentele necesare pentru a integra responsabil și etic tehnologiile de Inteligență Artificială în fluxurile lor de lucru zilnice și în strategiile operaționale mai ample.

3.1 Utilizarea IA în Crearea de Conținut

Integrarea IA în domeniul creării de conținut, care include generarea de text, producția video și designul grafic, introduce o serie de considerații etice care merită o examinare atentă. În timp ce IA oferă eficiență și posibilități creative fără precedent, prezintă și provocări care necesită definirea unor linii directoare clare și a celor mai bune practici.



Proprietate Intelectuală și Titularitate: Utilizarea IA în generarea de conținut ridică întrebări despre titularitate și drepturile de proprietate intelectuală. Determinarea cui îi aparțin drepturile asupra conținutului generat de IA — utilizatorului, dezvoltatorului IA sau o combinație — necesită o considerare atentă și cadre de reglementare adecvate.



Dezinformare și Deepfake: Capacitatea IA de a genera conținut realist și persuasiv, inclusiv deepfake-uri, reprezintă o amenințare semnificativă în termeni de dezinformare și manipulare. Asigurarea autenticității și integrității conținutului generat de IA este fundamentală pentru menținerea încrederii și combaterea răspândirii informațiilor false.



Prejudecăți și Discriminare: Modelele IA pot perpetua în mod neintenționat prejudecăți prezente în datele pe care au fost instruite, ducând la rezultate discriminatorii în conținutul generat. Abordarea prejudecăților în algoritmi IA și asigurarea echității și incluziunii în conținutul generat de IA este crucială pentru a evita întărirea prejudecăților sociale.



Transparență și Divulgare: Cu răspândirea crescândă a conținutului generat de IA, transparența și divulgarea devin esențiale. Identificarea clară a conținutului care a fost generat sau manipulat de IA ajută la menținerea încrederii și permite utilizatorilor să ia decizii informate.



Responsabilitate și Răspundere: Stabilirea unor linii clare de responsabilitate și răspundere pentru conținutul generat de IA este crucială. Determinarea cine este responsabil pentru potențialele daune cauzate de conținutul generat de IA — utilizatorul, dezvoltatorul IA sau platforma — necesită o reflecție atentă și cadre de reglementare adecvate.

Soluții Potențiale și Bune Practici

Dezvoltarea Etică a IA:	Implementarea ghidurilor și principiilor etice în dezvoltarea și implementarea modelelor de IA pentru crearea de conținut este esențială. Aceasta include asigurarea transparenței, echității și responsabilității în algoritmi de IA.
Educație și Conștientizare:	Promovarea educației și conștientizării cu privire la capacitățile și limitele conținutului generat de IA ajută utilizatorii să evalueze critic și să interpreteze acest conținut.
Cadre Juridice Robuste:	Dezvoltarea unor cadre juridice clare care să abordeze proprietatea intelectuală, responsabilitatea și preocupările etice legate de conținutul generat de IA este crucială pentru navigarea în peisajul în evoluție.
Colaborare și Standarde Industriale:	Încurajarea colaborării între părțile interesate, inclusiv dezvoltatorii de IA, creatorii de conținut și factorii de decizie politică, pentru a stabili standarde industriale și bune practici pentru conținutul etic generat de IA este esențială pentru o inovare responsabilă.
Monitorizare și Evaluare Continuă:	Monitorizarea și evaluarea continuă a conținutului generat de IA pentru potențiale încălcări etice și prejudecăți ajută la asigurarea faptului că IA este utilizată în mod responsabil și în conformitate cu valorile societății.

Prin abordarea proactivă a acestor provocări etice și implementarea practicilor responsabile, sectoarele de comunicare și informare pot valorifica potențialul IA pentru crearea de conținut, menținând în același timp standarde etice și protejând bunăstarea societății.

Cele mai bune practici

Generarea textului

Instrumentele de IA sunt valoroase pentru redactarea conținutului, sintetizarea textelor lungi și brainstorming-ul ideilor. Cu toate acestea, ar trebui să fie utilizate ca ajutor, nu ca înlocuitor pentru creativitatea umană și judecata editorială. Profesioniștii umani trebuie să revizuiască și să modifice textul generat de IA pentru a asigura acuratețea, claritatea și tonul adecvat. Transparența este menținută prin divulgarea utilizării IA în crearea conținutului.

Crearea video

IA poate îmbunătăți calitatea video prin creșterea rezoluției, reducerea zgomotului și stabilizarea filmărilor instabile. De asemenea, poate ajuta la scrierea scenariilor, planificarea pre-producției și sarcinile de post-producție, cum ar fi etichetarea scenelor și subtitrarea. Cea mai bună practică implică utilizarea IA pentru a îmbunătăți procesul de creare video, în timp ce videoziariștii umani mențin controlul asupra narațiunii și considerațiilor etice. Instrumentele de detectare deepfake bazate pe IA pot fi utilizate pentru a verifica autenticitatea conținutului video.

Design grafic

Instrumentele bazate pe IA pot genera rapid opțiuni de design și pot automatiza sarcini repetitive de design, îmbunătățind eficiența și explorând posibilități creative. Designerii umani ar trebui să mențină controlul artistic și să se asigure că designurile sunt valide din punct de vedere etic și conforme din punct de vedere legal în ceea ce privește proprietatea intelectuală.

Provocări etice

- **Transparență:** Nedivulgarea utilizării IA în crearea conținutului poate eroda încrederea publicului. Transparența este un element cheie al cadrelor de reglementare europene privind IA.
- **Responsabilitate:** Determinarea persoanelor responsabile pentru erori, prejudecăți sau conținut necorespunzător generat de IA reprezintă o provocare. Responsabilitatea necesită roluri clar definite și supraveghere umană.
- **Proprietate intelectuală:** Utilizarea instrumentelor de IA antrenate pe material protejat de drepturi de autor ridică îngrijorări cu privire la proprietate și posibile încălcări. Utilizarea etică a datelor de antrenament și atribuirea corectă a contribuțiilor, atât umane cât și IA, sunt cruciale.
- **Prejudecată:** Modelele de IA pot perpetua prejudecăți prezente în datele lor de antrenament, ducând la conținut distorsionat sau discriminatoriu. Selecția atentă și validarea datelor de antrenament, împreună cu revizuirea umană, pot atenua acest risc.
- **Impact asupra creativității umane și ocupării forței de muncă:** O dependență excesivă de IA pentru crearea de conținut ar putea diminua oportunitățile pentru creatorii umani și ar putea duce la un declin al originalității și gândirii critice. Echilibrarea utilizării IA cu valoarea competențelor umane și adaptabilitatea este esențială.

Soluții



Divulgare

Etichetarea clară a conținutului generat de IA pentru a informa publicul despre nivelul de implicare al IA.



Supraveghere Umană

Implementarea fluxurilor de lucru care necesită revizuire umană și modificarea conținutului generat de IA înainte de publicare sau difuzare.



Due Diligence pentru PI

Înțelegerea termenilor de licență ai instrumentelor de IA și originea datelor lor de instruire. Evitarea utilizării instrumentelor care se bazează pe material protejat de drepturi de autor dobândit în mod necorespunzător.



Detectarea și Atenuarea Prejudecăților

Evaluarea continuă a modelelor de IA pentru potențiale prejudecăți și utilizarea datelor de instruire diverse și reprezentative.



Accent pe Augmentare

Încadrarea IA ca un instrument pentru a spori abilitățile și creativitatea umană, mai degrabă decât un înlocuitor. Sublinierea capacității unice a ființelor umane de a se adapta la schimbările culturale.



Mecanisme de Feedback

Stabilirea canalelor pentru ca utilizatorii să raporteze erori sau prejudecăți în conținutul generat de IA.



Educație și Formare

Investiția în programe de formare și actualizare a competențelor pentru creatorii de conținut pentru a se adapta la peisajul în evoluție și pentru a utiliza eficient IA.



Colaborare

Încurajarea colaborării între dezvoltatorii de IA, creatorii de conținut și experții în etică pentru a asigura o dezvoltare și implementare responsabilă a IA.



Audituri Regulate

Efectuarea de audituri regulate ale sistemelor de IA pentru a identifica și aborda potențiale prejudecăți, erori și preocupări etice.



Ghiduri și Standarde Etice

Dezvoltarea și implementarea unor ghiduri etice clare și standarde pentru utilizarea IA în crearea de conținut.

3.2 IA pentru Jurnalism și Media

IA pentru Verificarea Faptelor

Instrumentele de verificare a faptelor bazate pe IA pot analiza cantități enorme de date, inclusiv articole de ziar, postări pe rețelele sociale și înregistrări istorice, pentru a identifica inconsecvențe, erori logice și potențiale dezinformări. Aceste instrumente pot verifica afirmațiile prin comparație cu surse de încredere, pot semnaliza modele suspecte și pot oferi jurnaliștilor dovezi pentru a susține sau infirma afirmații. Cu toate acestea, instrumentele de verificare a faptelor bazate pe IA nu sunt infailibile și pot produce rezultate fals pozitive sau negative. Jurnaliștii trebuie să evalueze critic rezultatele furnizate de IA și să efectueze cercetări suplimentare pentru a verifica acuratețea informațiilor.

Detectarea Deepfake

Deepfake-urile reprezintă o amenințare semnificativă pentru integritatea jurnalismului și încrederea publicului în mass-media. Instrumentele de detectare a deepfake bazate pe IA pot analiza înregistrări video și audio pentru a identifica inconsecvențe subtile, cum ar fi expresii faciale nenaturale, mișcări ale buzelor sau modele vocale, care pot indica manipulare. Aceste instrumente pot utiliza, de asemenea, tehnici de informatică forensică pentru a urmări originea fișierelor media și a identifica potențiale surse de deepfake. Jurnaliștii ar trebui să utilizeze instrumente de detectare a deepfake împreună cu alte metode de verificare pentru a asigura autenticitatea conținutului media.

Verificarea Surselor

Creșterea conținutului generat de IA și proliferarea informațiilor online au făcut din ce în ce mai dificilă distincția între surse credibile și nesigure. IA poate asista jurnaliștii în verificarea surselor prin analizarea reputației site-urilor web, a conturilor de social media și a autorilor individuali. IA poate, de asemenea, să urmărească răspândirea informațiilor online și să identifice

potențiale surse de dezinformare. Cu toate acestea, IA nu poate înlocui judecata umană când vine vorba de evaluarea credibilității surselor. Jurnaliștii trebuie să ia în considerare contextul, expertiza și potențialele prejudecăți ale surselor atunci când evaluează fiabilitatea informațiilor. Aprofundare a considerațiilor etice în raportarea automatizată.

Transparența Autoratului IA

Când IA este utilizată pentru a genera articole de știri sau rapoarte, este esențial să se comunice acest lucru în mod clar publicului. Acest lucru poate fi realizat prin includerea unei declarații la începutul sau sfârșitul articolului care indică faptul că a fost generat cu asistența IA. Transparența ajută la menținerea încrederii publicului în jurnalism și asigură că cititorii pot lua decizii informate cu privire la informațiile pe care le consumă.

Responsabilitatea pentru Erori

Când rapoartele de știri generate de IA conțin erori, este fundamental să se stabilească cine este responsabil pentru corectarea acestora. Organizațiile media ar trebui să aibă politici și proceduri clare pentru revizuirea conținutului generat de IA și abordarea oricărui inexactități. Aceasta poate implica desemnarea unor editori umani pentru a supraveghea rapoartele generate de IA sau implementarea unor sisteme automatizate pentru a semnaliza potențiale erori.

Potențial de Prejudecată

Algoritmii IA pot perpetua prejudecăți prezente în datele lor de antrenament, ducând la o acoperire jurnalistică distorsionată sau nedreaptă. Pentru a atenua acest risc, organizațiile media ar trebui să utilizeze seturi de date diversificate și reprezentative pentru a antrena modelele IA și să implementeze supraveghere umană pentru a identifica și corecta potențiale prejudecăți.

Impactul asupra Rolurilor Jurnalistice

Utilizarea tot mai mare a IA în jurnalism transformă rolurile jurnaliștilor umani. În timp ce IA poate automatiza anumite sarcini, cum ar fi colectarea și analiza datelor, jurnaliștii umani rămân esențiali pentru a oferi context, interpretare și judecată etică. Jurnaliștii trebuie să se adapteze la aceste roluri în evoluție dezvoltând noi competențe, cum ar fi alfabetizarea în domeniul IA și jurnalismul de date, și concentrându-se pe domenii în care expertiza umană rămâne indispensabilă.

Riscul Dezinformării la Scară Largă

Capacitatea IA de a genera conținut rapid și la scară largă poate fi exploatată pentru a răspândi dezinformare și propagandă. Pentru a contracara această amenințare, organizațiile media și companiile de tehnologie trebuie să investească în procese solide de verificare, inițiative de alfabetizare media și eforturi de colaborare pentru a identifica și combate campaniile de dezinformare. Aprofundarea Soluțiilor pentru IA Etică în Jurnalism.

Implementarea Politicilor Clare de Divulgare

Organizațiile media ar trebui să adopte politici clare și consecvente pentru a dezvălui utilizarea IA în conținutul de știri. Aceasta include informarea publicului când IA a fost utilizată pentru a genera articole, titluri sau alte elemente ale rapoartelor de știri.

Menținerea Supravegherii Editoriale

Jurnaliștii umani ar trebui să joace un rol activ în revizuirea și verificarea conținutului generat de IA înainte de publicare. Acest lucru ajută la asigurarea acurateței, corectitudinii și standardelor etice.

Prioritizarea Calității și Diversității Datelor

Organizațiile media ar trebui să utilizeze seturi de date de înaltă calitate, diverse și reprezentative pentru a antrena modelele de IA pentru jurnalism.

Acest lucru ajută la minimizarea prejudecăților și asigură că conținutul generat de IA reflectă o gamă largă de perspective.

Investiția în Instrumente de Verificare a Faptelor și Detectare a Deepfake

Organizațiile media ar trebui să echipeze jurnaliștii cu instrumentele și instruirea necesare bazate pe IA pentru a verifica informațiile și a detecta media manipulată. Aceasta include furnizarea accesului la baze de date de verificare a faptelor, software de detectare a deepfake și alte instrumente de verificare.

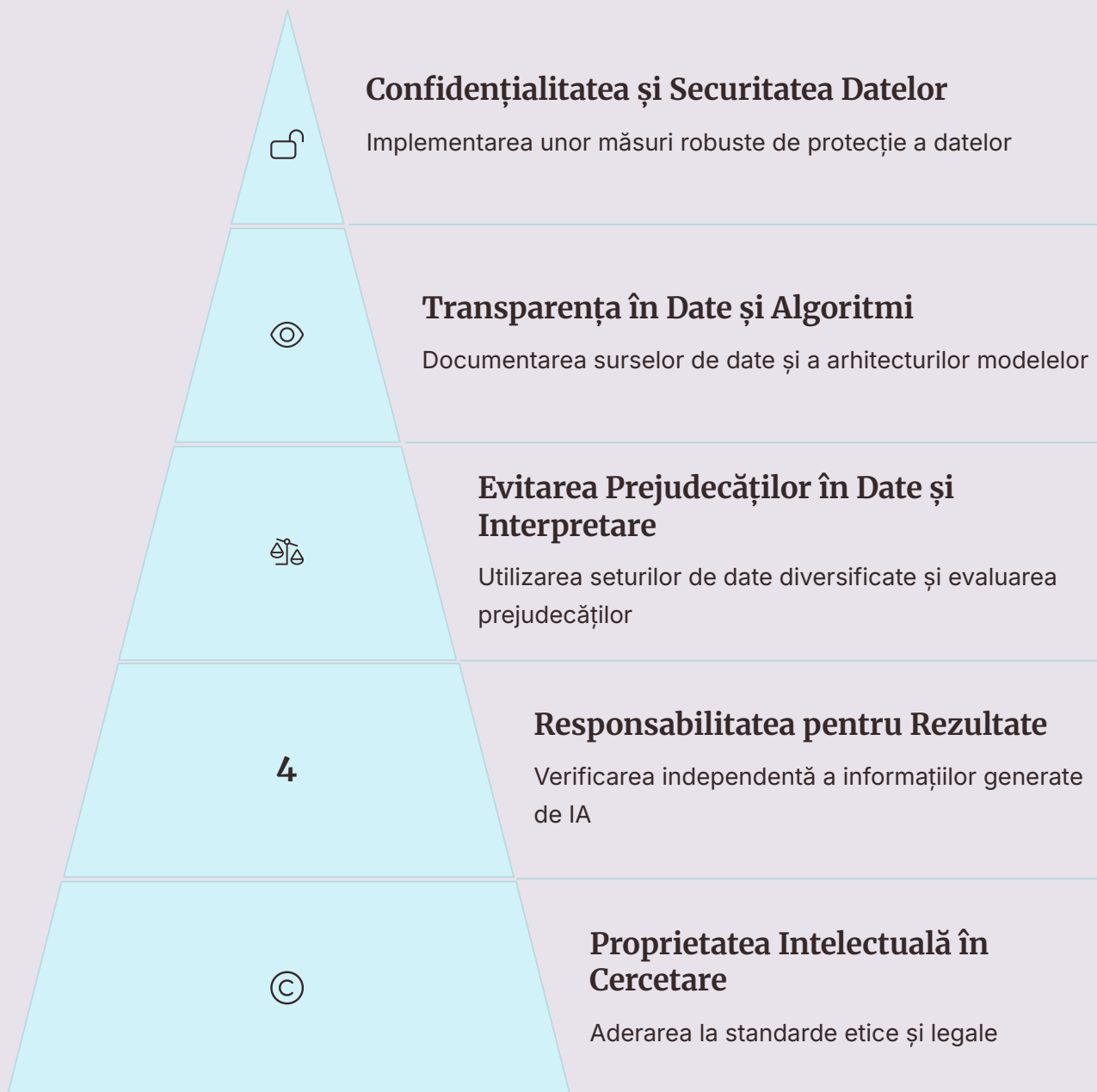
Concentrarea pe Experiența Umană

Organizațiile media ar trebui să recunoască și să valorifice competențele unice și judecata etică a jurnaliștilor umani. Aceasta include investiția în programe de formare și dezvoltare care se concentrează pe gândirea critică, jurnalismul de investigație și luarea deciziilor etice.

Promovarea Alfabetizării Media

Organizațiile media, instituțiile educaționale și agențiile guvernamentale ar trebui să colaboreze pentru a promova alfabetizarea media și a educa publicul cu privire la evaluarea critică a informațiilor, identificarea potențialei dezinformări și navigarea în peisajul media complex.

3.3 IA pentru Analiza Datelor și Cercetare



Confidențialitatea și Securitatea Datelor:

Analiza datelor bazată pe IA implică adesea procesarea unor cantități mari de date personale și sensibile, ceea ce necesită măsuri robuste de protecție a datelor. Aceasta include aderarea la principiile de confidențialitate prin design, asigurarea stocării securizate a datelor și respectarea reglementărilor privind protecția datelor, precum GDPR. În plus, obținerea consimțământului informat pentru colectarea și utilizarea datelor este crucială, așa cum a fost evidențiat de prezentarea lui Mihai Deaconu la masa rotundă despre preocupările legate de confidențialitate în contextul IA. Anonimizarea, criptarea și controalele de acces pot îmbunătăți și mai mult securitatea datelor și pot proteja confidențialitatea persoanelor.

Transparența în Date și Algoritmi:

Cercetătorii ar trebui să fie transparenți cu privire la sursele de date utilizate pentru instruirea modelelor de IA și algoritmi folosiți pentru analiză. O documentație clară a surselor de date, a etapelor de pre-procesare și a arhitecturilor modelelor permite examinarea potențialelor prejudecăți și limitări. Conceptul de "IA Explicabilă", în care procesele decizionale ale IA sunt clarificate, devine din ce în ce mai important în cercetare, în special atunci când IA este utilizată pentru a lua decizii critice. Această transparență promovează încrederea și permite validarea externă a rezultatelor cercetării.

Evitarea Prejudecăților în Date și Interpretare:

Modelele de IA pot amplifica prejudecățile prezente în datele pe care sunt instruite, ducând la interpretări distorsionate și rezultate potențial discriminatorii. Cercetătorii trebuie să abordeze proactiv prejudecățile folosind seturi de date diverse și reprezentative, evaluând modelele de IA pentru a identifica prejudecăți și interpretând critic informațiile generate de IA. Auditurile regulate ale sistemelor de IA pot ajuta la identificarea și atenuarea prejudecăților, asigurând echitatea și imparțialitatea în analiza datelor.

Responsabilitatea pentru Rezultate:

Cercetătorii sunt responsabili pentru integritatea și acuratețea rezultatelor lor, chiar și atunci când utilizează instrumente de IA. Este esențial să verifice informațiile generate de IA folosind metode independente și să se asigure că acestea sunt prezentate în mod responsabil și cu avertismente adecvate privind limitările IA. Studiul lui Susarla et al. (2023) discută necesitatea unei conduite responsabile a activităților academice în era IA generative, subliniind importanța supravegherii umane și a evaluării critice a rezultatelor generate de IA.

Proprietatea Intelectuală în Cercetare:

Probleme de proprietate intelectuală pot apărea în cercetarea ghidată de IA, în special în ceea ce privește proprietatea asupra informațiilor sau modelelor generate de IA. Cercetătorii ar trebui să fie conștienți de aceste probleme și să se asigure că respectă standardele etice și legale privind utilizarea datelor și diseminarea rezultatelor. Acorduri clare și recunoașteri ar trebui să fie în vigoare pentru a aborda drepturile de proprietate intelectuală legate de IA în cercetare, favorizând colaborarea cu respectarea proprietății.

Considerații Suplimentare:

Calitatea și Integritatea Datelor: Asigurarea calității și integrității datelor utilizate pentru analiza cu IA este esențială pentru obținerea unor informații fiabile și semnificative. Procesele de curățare, validare și detectare a erorilor în date ar trebui implementate pentru a minimiza riscul de rezultate inexacte sau înșelătoare.

Colectarea Etică a Datelor: Practicile de colectare a datelor ar trebui să fie etice și să respecte drepturile indivizilor. Aceasta include obținerea consimțământului, evitarea practicilor înșelătoare și minimizarea colectării de date inutile.

Arhivarea și Conservarea pe Termen Lung a Datelor: Ar trebui luate în considerare arhivarea și conservarea pe termen lung a datelor utilizate în cercetarea cu IA, asigurând accesibilitatea acestora pentru cercetări viitoare și reproductibilitatea rezultatelor.

Colaborare și Partajare de Date: Încurajarea colaborării și partajării responsabile a datelor între cercetători poate accelera progresele IA și promova transparența. Cu toate acestea, partajarea datelor ar trebui efectuată în conformitate cu reglementările privind confidențialitatea și cu o anonimizare adecvată pentru a proteja datele persoanelor.

Aderând la aceste bune practici și luând în considerare cu atenție provocările etice și soluțiile prezentate, profesioniștii din sectoarele de comunicare și informare pot valorifica puterea IA în mod responsabil și etic în analiza datelor și cercetare. Informațiile din proiectul "Me, Myself, and AI" și de la masa rotundă subliniază importanța unui dialog continuu și a dezvoltării unor ghiduri practice pentru navigarea în peisajul etic în evoluție al IA.

CAPITOLUL IV

Lista de Verificare pentru Utilizarea etică a IA



Lista de Control cu Referință Rapidă pentru utilizarea etică a IA este o componentă cheie a „Ghidului Etic pentru Inteligența Artificială în Sectoarele de Comunicare și Informare”. Această listă de control este concepută pentru a fi un instrument concis pentru profesioniști pentru a evalua implicațiile etice înainte de a utiliza IA.

Transparență

- Este clar când conținutul a fost generat sau semnificativ asistat de IA? Luați în considerare dacă textul, imaginile, audio sau videoclipurile generate de IA sunt etichetate corespunzător pentru a informa publicul despre originea lor.
- Sunt înțelese și comunicate limitările instrumentului de IA utilizat, acolo unde este relevant? Recunoașteți că IA ar putea produce erori sau conținut distorsionat.
- Pot fi explicate procesele și logica din spatele rezultatului IA în termeni simpli, dacă este necesar? Deși funcționarea internă a unei IA complexe poate fi opacă, scopul general și sursele de date ar trebui să fie ușor de înțeles.
- Există mecanisme de feedback care să permită utilizatorilor să raporteze probleme sau inexactități în conținutul generat de IA? Acest lucru permite îmbunătățirea continuă și identificarea încălcărilor etice.

Responsabilitate

- Sunt definite roluri și responsabilități clare pentru utilizarea și supravegherea instrumentelor de IA? Identificați cine este responsabil pentru implicațiile etice ale rezultatelor IA.
- Există un om în procesul de luare a deciziilor pentru deciziile critice sau generarea de conținut care implică IA? Supravegherea umană este crucială pentru a asigura respectarea considerațiilor etice și corectarea erorilor.
- Există cadre de guvernare și politici interne stabilite privind utilizarea etică a IA în cadrul organizației? Ghidurile formale ajută la asigurarea unor practici etice consecvente.
- Se efectuează audituri regulate ale sistemelor de IA pentru a evalua impactul lor etic și conformitatea cu ghidurile? Această abordare proactivă poate identifica și atenua potențialele riscuri etice.

Acuratețe

- Informațiile procesate sau distribuite de IA au fost verificate pentru corectitudine factuală? Conținutul generat de IA nu ar trebui considerat precis a priori și necesită verificarea faptelor.
- Sunt implementate măsuri pentru a evita răspândirea dezinformării prin instrumente de IA? Fiți conștienți de potențialul IA de a crea și amplifica informații false sau înșelătoare. Instrumente precum InVID Verification Tool și Deepfake Detection with Sensity sunt menționate pentru verificarea autenticității mediilor.
- Sunt sursele de date utilizate de IA de încredere și lipsite de prejudecăți? Datele de antrenament distorsionate pot duce la rezultate distorsionate.
- Este IA utilizată într-un mod care ar putea duce la reprezentări greșite sau distorsiuni ale informațiilor? Asigurați-vă că IA este utilizată pentru a îmbunătăți, nu pentru a submina, integritatea informațiilor.

Confidențialitate

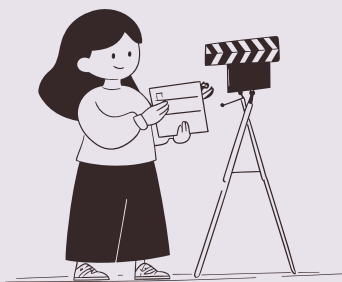
- Datele personale sunt gestionate în conformitate cu legile și reglementările relevante privind protecția datelor (de exemplu, GDPR)? Aplicațiile de IA trebuie să respecte confidențialitatea utilizatorilor.
- Există măsuri adecvate pentru a proteja confidențialitatea persoanelor atunci când se utilizează IA pentru analiza datelor sau generarea de conținut? Luați în considerare tehnici de anonimizare acolo unde este necesar.
- Se obține consimțământul atunci când IA procesează date personale? Asigurați transparența cu privire la modul în care informațiile personale sunt utilizate de sistemele de IA.
- Sunt instrumentele de IA utilizate într-un mod care ar putea duce la colectarea sau divulgarea neautorizată a informațiilor personale? Securitatea și confidențialitatea prin design sunt esențiale.

Proprietate Intelectuală

- Utilizarea IA respectă drepturile de autor și alte drepturi de proprietate intelectuală în datele pe care le utilizează și în conținutul pe care îl generează? Fiți conștienți de implicațiile legale ale utilizării IA pentru a crea conținut bazat pe material protejat de drepturi de autor.
- Drepturile de proprietate intelectuală ale conținutului generat de IA sunt clar definite? Proprietatea asupra operelor create de IA poate fi o chestiune juridică complexă.
- Se acordă credit corespunzător pentru orice conținut creat de om utilizat sau adaptat de IA? Respectarea muncii altora rămâne importantă.
- Este IA utilizată într-un mod care ar putea facilita plagiatul sau încălcarea drepturilor creative? Utilizarea educațională a IA, de exemplu, necesită o atenție deosebită la originalitate.

4.2 Scenarii de Exemplu

Exemple practice de aplicare a listei de verificare în situații reale.



Această parte este destinată să ofere exemple practice despre cum principiile etice prezentate în document pot fi aplicate în situații reale. Aceste scenarii urmăresc să facă principiile abstracte mai tangibile și să ofere îndrumare profesioniștilor în navigarea complexităților etice ale utilizării IA în munca lor. Bazându-se pe principiile etice fundamentale de Transparență, Responsabilitate, Acuratețe, Confidențialitate și Proprietate Intelectuală, și ținând cont de discuțiile de la masa rotundă despre IA și Etică, următoarele sunt câteva scenarii exemplificative:

Scenariul 1: Generarea de articole de știri asistată de IA

Un mic ziar online se confruntă cu o presiune crescândă de a produce mai mult conținut cu resurse limitate. Ei decid să implementeze un instrument de IA care poate genera schițe inițiale ale articolelor de știri bazate pe date și surse furnizate.

Considerații etice: Acest scenariu ridică probleme legate de Transparență, Responsabilitate și Acuratețe.

Aplicarea principiilor etice:

Transparență:

Ar trebui ziarul să dezvăluie cititorilor săi că IA a fost utilizată în crearea articolului? Principiul transparenței sugerează că o astfel de dezvăluire este crucială pentru construirea și menținerea încrederii. O etichetă clară precum „Acest articol a fost generat cu asistența IA și revizuit de editori umani” ar putea fi utilizată.

Responsabilitate:

Cine este responsabil dacă articolul generat de IA conține erori factuale sau prezintă o perspectivă distorsionată? Stabilirea unor roluri și responsabilități clare, cu jurnaliști umani care supraveghează rezultatul IA și verifică informațiile, este esențială pentru responsabilitate.

Acuratețe:

Cum poate ziarul să garanteze că instrumentul de IA utilizează date fiabile și evită perpetuarea dezinformării sau a prejudecăților prezente în datele sale de antrenament? Auditurile regulate ale sistemului de IA și ale surselor pe care le utilizează, împreună cu verificarea umană a faptelor, sunt necesare pentru menținerea acurateței.

Scenariul 2: IA în designul grafic pentru materiale de marketing

O agenție de marketing utilizează instrumente bazate pe IA pentru a genera rapid diverse opțiuni de design pentru campania de social media a unui client. Instrumentul de IA utilizează o bibliotecă vastă de imagini stock și elemente de design în procesul său de generare.

Considerații etice: Acest scenariu se referă în principal la Proprietatea Intelectuală și Transparență.

Proprietate Intelectuală

Are agenția dreptul de a utiliza comercial designurile generate de IA? Materialele sursă utilizate de IA sunt licențiate adecvat și există riscul de a încălca drepturi de autor existente? Agenția trebuie să se asigure că înțelege termenii de licență ai instrumentului de IA și originea datelor sale de antrenament.

Transparență

Ar trebui agenția să informeze clientul că IA a fost implicată semnificativ în crearea opțiunilor de design? Deși nu este întotdeauna cerut legal, transparența cu privire la procesul creativ poate construi încredere cu clientul.

Scenariul 3: Moderarea Conținutului Ghidată de IA pe Rețelele Sociale

O platformă de social media folosește algoritmi de IA pentru a identifica și a elimina automat conținutul care încalcă liniile directoare ale comunității, inclusiv discursul instigator la ură și dezinformarea.

Considerații Etice: Acest scenariu ridică probleme legate de Acuratețe, Responsabilitate și potențiale prejudecăți care afectează diferite grupuri.

Aplicarea Principiilor Etice:

Acuratețe:

Cât de precisă este IA în identificarea conținutului dăunător? Există cazuri de rezultate fals pozitive (conținut legitim care este eliminat) sau fals negative (conținut dăunător care nu este detectat)? Platforma trebuie să evalueze și să perfecționeze continuu instrumentele de moderare bazate pe IA pentru a îmbunătăți acuratețea și a minimiza erorile.

Responsabilitate:

Dacă conținutul legitim este eliminat din greșală, ce mecanisme sunt implementate pentru a permite utilizatorilor să facă apel și să caute remediere? Cadrele clare de responsabilitate și procesele transparente de contestare sunt esențiale.

Prejudecăți:

Algoritmul IA este antrenat pe date care conțin prejudecăți sociale, conducând la o moderare disproporționată a conținutului provenit de la anumite grupuri demografice? Asigurarea incluzivității și abordarea potențialelor prejudecăți în datele de antrenare ale IA este crucială pentru o moderare etică a conținutului.

Scenariul 4: IA pentru Învățare Personalizată în Educația Media

O instituție educațională folosește o platformă de IA pentru a personaliza materialele de învățare și a recomanda resurse studenților în programul lor de studii media, bazat pe progresul lor individual și stilurile de învățare.

Considerații Etice: Acest scenariu implică Confidențialitate și Transparență.

Aplicarea Principiilor Etice:

Confidențialitate

Ce date sunt colectate despre studenți, cum sunt stocate și sunt utilizate exclusiv în scopul personalizării experienței lor de învățare? Instituția trebuie să asigure conformitatea cu reglementările privind protecția datelor și să obțină consimțământul informat de la studenți cu privire la colectarea și utilizarea datelor lor.

Transparență

Sunt studenții informați despre cum funcționează platforma de IA, ce date colectează și cum le influențează parcursul de învățare? Furnizarea de informații clare și accesibile despre sistemul de IA promovează încrederea și permite studenților să înțeleagă și să interacționeze cu tehnologia.

Scenariul 5: IA în Jurnalismul Automatizat pentru Rapoartele Financiare

O agenție de știri financiare utilizează IA pentru a genera automat rapoarte despre profiturile companiilor bazate pe date financiare disponibile public. IA poate produce aceste rapoarte mult mai rapid decât jurnaliștii umani.

Considerații Etice: Acest scenariu implică Acuratețea și Responsabilitatea. Aplicarea Principiilor Etice:

Acuratețe

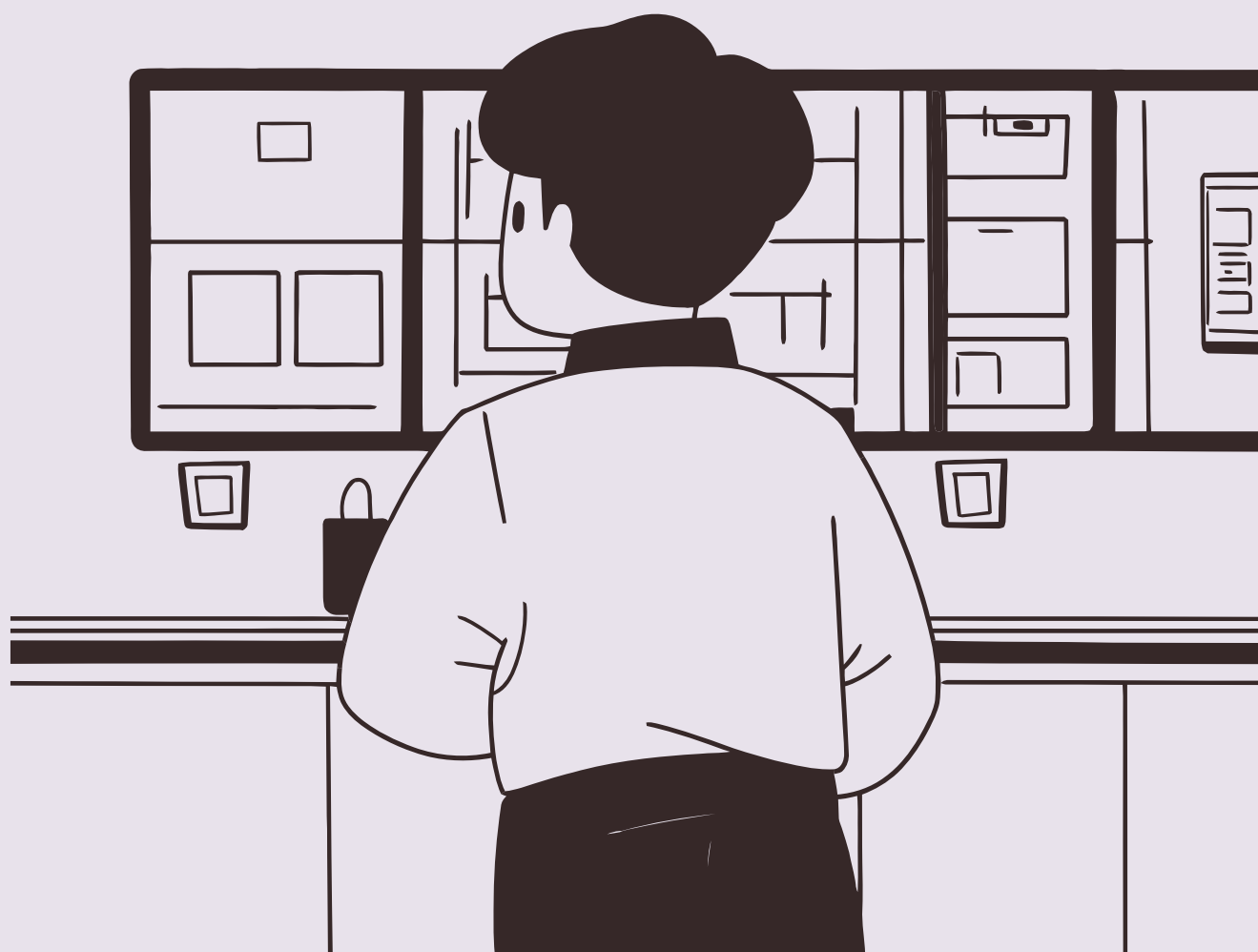
Deși sursa datelor poate fi de încredere, cum interpretează și prezintă IA aceste informații? Există măsuri de siguranță pentru a preveni interpretările greșite sau omiterea contextelor cruciale pe care un jurnalist uman le-ar putea include? Sunt necesare teste robuste și validarea capacităților de raportare ale IA.

Responsabilitate

Dacă un raport financiar generat de IA conține inexactități care duc la pierderi financiare pentru cititori, cine este responsabil? Agenția de știri trebuie să aibă protocoale clare pentru verificarea rezultatelor IA și să își asume responsabilitatea pentru eventualele erori.

CAPITOLUL V

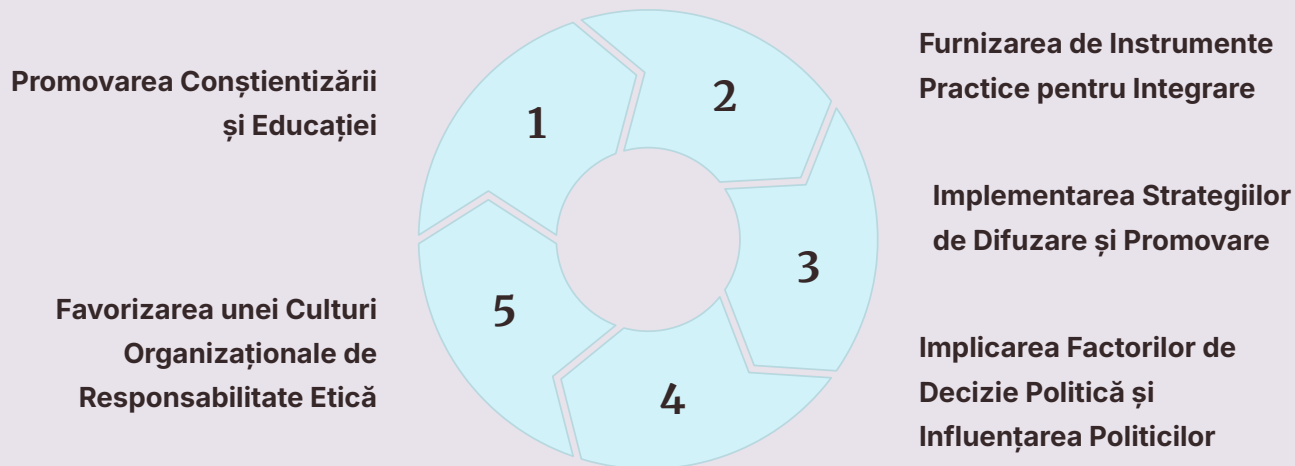
Implementare și Monitorizare



Această secțiune prezintă strategiile pentru integrarea principiilor etice ale IA în politicile operaționale ale organizațiilor din sectoarele de comunicare și informare și discută mecanismele pentru monitorizarea și asigurarea aderenței la aceste linii directoare.

5.1 Pași pentru Adoptarea Liniilor Directoare

„Liniile Directoare Etice pentru Inteligența Artificială în Sectoarele de Comunicare și Informare”, dezvoltate ca rezultat cheie al proiectului Erasmus+ „Me, Myself, and AI”, sunt destinate să fie o resursă practică pentru profesioniștii care navighează prin complexitățile etice ale tehnologiilor de IA. Designul și obiectivele proiectului sugerează puternic o abordare multifacțată pentru a facilita integrarea acestor principii în politicile organizaționale.



Promovarea Conștientizării și Educației: Un pas fundamental în adoptarea oricărui nou set de principii este asigurarea că părțile interesate sunt conștiente de existența lor și înțeleg importanța acestora. Proiectul „Me, Myself, and AI” prioritizează creșterea conștientizării cu privire la tehnologiile IA și implicațiile lor etice în rândul profesioniștilor din comunicare. Crearea documentului de linii directoare etice în sine, împreună cu lista de verificare practică care îl însoțește și videoclipul animat care rezumă liniile directoare, sunt toate inițiative strategice menite să educe profesioniștii și să facă principiile etice accesibile și ușor de înțeles. În plus, proiectul se adresează în mod specific educatorilor și formatorilor pentru adulți, încurajându-i să încorporeze aceste linii directoare în programele lor de studiu, integrând astfel considerații etice în formarea viitorilor profesioniști din domeniu. Proiectul își propune să atingă un public larg și diversificat, asigurându-se că o cultură a utilizării responsabile a IA este cultivată prin diverse roluri și niveluri în cadrul organizațiilor. Masa rotundă transnațională planificată în Italia servește, de asemenea, ca platformă pentru catalizarea conștientizării în rândul unei porțiuni semnificative (90%) a grupului țintă cu privire la cele mai recente tehnologii de IA.

Furnizarea de Instrumente Practice pentru Integrare: Pentru ca principiile etice să fie adoptate în mod eficient, acestea trebuie traduse în acțiuni concrete și integrate în fluxurile de lucru zilnice. Proiectul „Me, Myself, and AI” recunoaște acest lucru prin dezvoltarea de instrumente practice concepute pentru a facilita integrarea etică a IA.

„Lista de Control pentru Utilizarea Etică a IA” este un exemplu primar, servind ca instrument de referință rapidă pe care profesioniștii îl pot utiliza pentru a evalua proactiv implicațiile etice ale instrumentelor de IA înainte de implementare. Această abordare pragmatică își propune să meargă dincolo de principiile abstracte și să ofere ghidare tangibilă pentru procesul decizional privind utilizarea IA în contexte organizaționale. Toolkit-ul online în curs de dezvoltare de către proiect va explica în continuare „de ce” și „cum” să se utilizeze IA în mod eficient și etic, oferind profesioniștilor exemple concrete și studii de caz pentru a-i ghida în integrarea responsabilă a IA.

Implementarea Strategiilor de Diseminare și Promovare: Adoptarea pe scară largă a ghidurilor etice necesită o diseminare și promovare eficientă pentru a asigura că acestea ajung la publicul țintă și sunt recunoscute ca resurse valoroase. Proiectul „Me, Myself, and AI” încorporează un plan cuprinzător de diseminare pentru a împărtăși rezultatele și produsele proiectului, inclusiv ghidurile etice, cu o gamă largă de părți interesate. Acest plan include dezvoltarea materialelor digitale de diseminare, implicarea cu mass-media, participarea la conferințe relevante și utilizarea strategică a platformelor de social media. Secțiunile dedicate pe site-urile web ale organizațiilor partenere și campaniile active pe social media sunt destinate să asigure că ghidurile etice, lista de control și videoclipul animat sunt ușor accesibile profesioniștilor din sectoarele de comunicare și informare. Atelierele de formare naționale și evenimentele de diseminare organizate în ambele țări partenere vor servi, de asemenea, ca platforme cruciale pentru introducerea și promovarea ghidurilor etice, explicând aplicațiile lor practice și încurajând adoptarea lor.

Implicarea Factorilor de Decizie Politică și Influențarea Politicilor: Realizarea unei integrări ample și durabile a principiilor etice necesită adesea implicarea factorilor de decizie politică și potențiala influență asupra politicilor și reglementărilor relevante. Proiectul „Me, Myself, and AI” identifică factorii de decizie politică ca un grup țintă cheie pentru ghidurile etice, cu așteptarea că aceștia vor utiliza aceste resurse pentru a informa dezvoltarea politicilor de guvernare a IA. Natura transnațională a parteneriatului dintre Eu și Lumea și POT Project este menită să consolideze capacitatea proiectului de a influența eforturile de politică și advocacy atât la nivel național, cât și european. Prin prezentarea unei abordări unificate și furnizarea unor ghiduri etice bine cercetate, proiectul își propune să contribuie la discursul mai larg despre etica IA și să informeze potențial viitoarele cadre de reglementare în sectoarele de comunicare și informare.

Promovarea unei Culturi Organizaționale de Responsabilitate Etică: În cele din urmă, integrarea cu succes a ghidurilor etice necesită o schimbare către o cultură organizațională care prioritizează utilizarea responsabilă a IA. Proiectul „Me, Myself, and AI” își propune să contribuie la această schimbare culturală prin încorporarea IA și a eticii în dezvoltarea profesională a forței de muncă din sector. Prin furnizarea de resurse educaționale, instrumente practice și oportunități de discuție și reflecție, proiectul încearcă să permită profesioniștilor să ia decizii etice informate în activitatea lor zilnică. Accentul pe transparență, responsabilitate, confidențialitate și proprietate intelectuală în cadrul ghidurilor oferă un cadru pentru organizații pentru a dezvolta politici și practici interne care se aliniază cu aceste principii etice fundamentale. Viziunea pe termen lung a proiectului include potențialul ca ghidurile etice să influențeze standardele industriei și procesele de certificare, încorporând și mai mult considerațiile etice în structura sectoarelor de comunicare și informare.

5.2 Monitorizare și Conformitate

Contextul mai larg al proiectului "Me, Myself, and AI" și discuțiile aferente oferă idei despre potențiale instrumente și strategii pentru a încuraja aderarea la aceste principii etice. Este important de menționat că în această etapă, acestea sunt în principal abordări propuse mai degrabă decât cadre juridice sau de reglementare stabilite, legate în mod specific de aceste orientări.

Alinierea cu Recomandările UNESCO

Ghidurile etice sunt în mod explicit fondate și alinate cu "Recomandarea privind Etica Inteligenței Artificiale" a UNESCO (2021). Această recomandare a UNESCO servește drept document important de politică internațională, stabilind un cadru universal pentru utilizarea etică a IA. Deși nu este obligatoriu din punct de vedere juridic în sine, aderarea la ghidurile "Me, Myself, and AI" poate fi văzută ca un pas practic spre încorporarea principiilor etice mai largi susținute de UNESCO în sectoarele de comunicare și informare. Monitorizarea aderenței ar putea implica încurajarea organizațiilor să își declare public angajamentul față de aceste principii UNESCO și să demonstreze modul în care practicile lor de IA se aliniază cu acestea, folosind potențial ghidurile "Me, Myself, and AI" ca un cadru concret pentru implementare.

Accent pe Transparență și Responsabilitate

Principiile etice fundamentale prezentate în ghiduri, în special transparența și responsabilitatea, sugerează în mod intrinsec mecanisme pentru monitorizare și conformitate. Pentru transparență, ghidurile recomandă practici precum etichetarea clară a conținutului generat de IA, explicarea proceselor de IA în termeni simpli, evidențierea limitărilor, activarea mecanismelor de feedback și auditarea sistemelor de IA. Monitorizarea aderenței ar putea implica organisme de supraveghere din industrie, organizații profesionale sau chiar utilizatorii înșiși care verifică implementarea acestor practici transparente. Responsabilitatea, așa cum este subliniată în ghiduri, necesită definirea rolurilor, crearea cadrelor de guvernare, încorporarea supravegherii umane, efectuarea de audituri regulate, furnizarea de mecanisme de feedback și asigurarea conformității legale și etice. Conformitatea ar putea fi monitorizată prin audituri organizaționale interne și stabilirea unor linii clare de responsabilitate pentru deciziile și rezultatele legate de IA.

Rolul Organizațiilor Profesionale și al Instituțiilor Educaționale

Proiectul "Me, Myself, and AI" se adresează organizațiilor profesionale, instituțiilor educaționale și ONG-urilor ca părți interesate cheie. Se așteaptă ca aceste entități să joace un rol crucial în promovarea și, potențial, în monitorizarea adoptării practicilor etice de IA. Instituțiile educaționale pot integra ghidurile în programele lor de studiu, asigurându-se că viitorii profesioniști sunt educați cu privire la principiile etice ale IA și înțeleg importanța acestora. Organizațiile profesionale pot aproba ghidurile, le pot încorpora în codurile lor de conduită și pot stabili potențial mecanisme prin care membrii să își demonstreze aderarea. ONG-urile axate pe drepturile digitale și tehnologia etică pot juca, de asemenea, un rol vital în susținerea adoptării acestor ghiduri și în responsabilizarea organizațiilor pentru practicile lor de IA.

Utilizarea Listei de Verificare pentru Utilizarea Etică a IA

„Lista de Verificare de Referință Rapidă” furnizată în cadrul ghidurilor servește ca instrument practic pentru autoevaluare. Organizațiile își pot încuraja profesioniștii să utilizeze această listă de verificare înainte de implementarea instrumentelor de IA, promovând o cultură a considerației etice proactive. Deși este în principal un instrument de autoreglementare, utilizarea consecventă a listei de verificare poate contribui la aderarea generală la principiile etice. Organizațiile ar putea, de asemenea, să ia în considerare încorporarea listei de verificare în procesele lor interne de aprobare pentru proiectele bazate pe IA.

Valorificarea Mecanismelor de Feedback și a Raportărilor Utilizatorilor

Ghidurile susțin stabilirea unor mecanisme transparente de feedback care permit utilizatorilor să raporteze erori sau prejudecăți în rezultatele generate de IA. Acest feedback ghidat de utilizatori poate funcționa ca un sistem informal de monitorizare, evidențiind potențiale zone unde practicile etice nu sunt respectate sau unde sistemele de IA ar putea produce conținut distorsionat sau inexact. Organizațiile ar trebui să aibă procese clare pentru gestionarea acestor feedback-uri și pentru a lua măsuri corective.

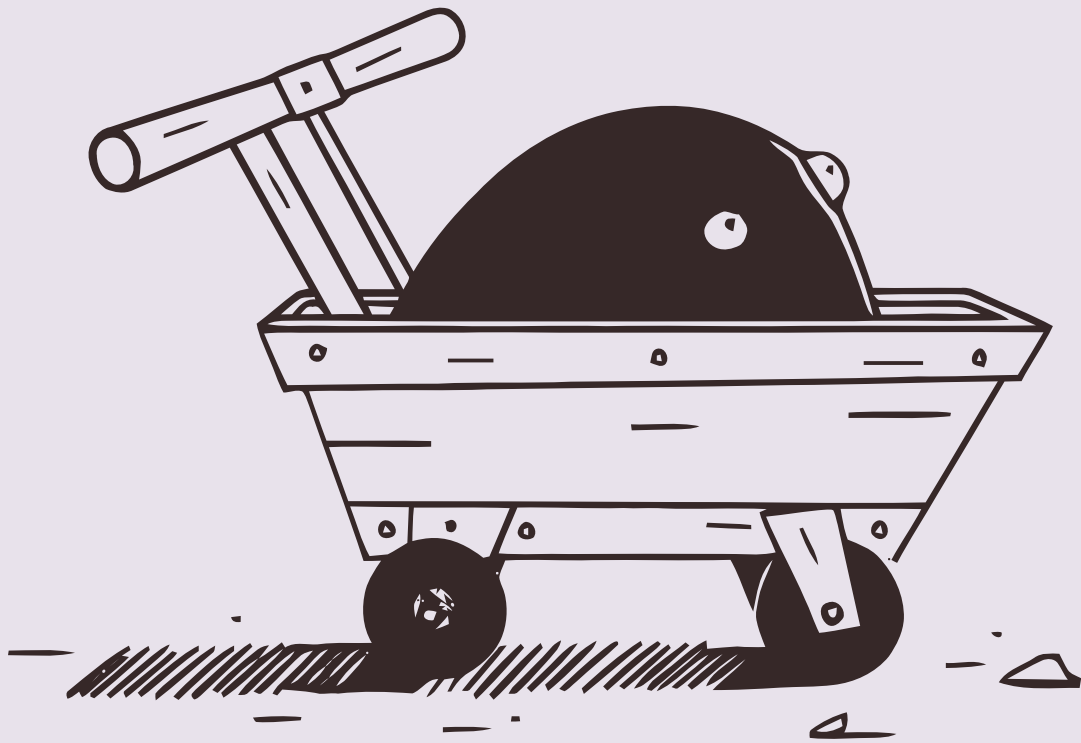
Considerarea Peisajului Legal și de Reglementare Mai Larg

Masa rotundă transnațională planificată ca parte a proiectului „Me, Myself, and AI” include discuții despre „etica IA și cadrele de reglementare europene”, indicând o conștientizare a peisajului legal în evoluție care înconjoară IA. Extrasul din „MEDIA ȘI INTELIGENȚA ARTIFICIALĂ” (Momoc & Surugiu, 2024) menționează, de asemenea, că legislația care reglementează responsabilitățile legate de IA este în curs de dezvoltare în diferite țări, inclusiv România. În timp ce ghidurile etice „Me, Myself, and AI” nu sunt în sine statute legale, aderarea la aceste principii poate ajuta organizațiile să se pregătească și să se alinieze cu reglementările emergente privind IA legate de domenii precum transparența, confidențialitatea datelor (acoperită sub principiul 'Confidențialitate' în ghiduri) și nediscriminarea. Conformitatea cu legile existente privind protecția datelor (de exemplu, GDPR în Europa) ar fi, de asemenea, un aspect relevant al monitorizării utilizării etice a IA în sectoarele de comunicare și informare.

Monitorizarea și Evaluarea Proiectului ca Model

Proiectul „Me, Myself, and AI” însuși are un cadru de monitorizare și evaluare cu Indicatori Cheie de Performanță (KPI) definiți. Deși se concentrează pe evaluarea succesului proiectului în atingerea obiectivelor sale (de exemplu, conștientizare sporită, dezvoltarea competențelor, adoptarea ghidurilor), principiile și metodologiile utilizate în acest cadru ar putea oferi un model pentru modul în care organizațiile ar putea monitoriza adoptarea și impactul ghidurilor etice în cadrul propriilor operațiuni. Aceasta ar putea implica definirea KPI-urilor interne legate de implementarea măsurilor de transparență, efectuarea auditurilor IA sau numărul de personal instruit în practici etice de IA.

Anexe



A1 – Termeni Cheie și Definiții

Iată o listă detaliată a termenilor cheie și definițiilor care se găsesc în cadrul „Ghidului Etic pentru Inteligența Artificială în Sectoarele de Comunicare și Informare”, elaborate cu explicații și citări ale surselor:

<p>Inteligență Artificială (IA)</p>	<p>Acesta este termenul fundamental în jurul căruia se învârtă întregul document, făcând esențială definirea sa. IA se referă probabil la simularea proceselor de inteligență umană de către mașini, în special a sistemelor informatice. Aceste procese includ învățarea (achiziționarea de informații și reguli pentru utilizarea informațiilor), raționamentul (utilizarea regulilor pentru a ajunge la concluzii aproximative sau definitive) și auto-corectarea. În contextul ghidului, definiția ar evidenția probabil integrarea din ce în ce mai mare a IA în sectoarele de comunicare și informare, potențialul său de a îmbunătăți eficiența și creativitatea, și necesitatea rezultată a considerațiilor etice pentru a-i gestiona impactul. Introducerea ghidului menționează în mod explicit „integrarea etică a Inteligenței Artificiale (IA) în sectoarele de comunicare și informare”, subliniind importanța definirii acestui concept fundamental.</p>
<p>Transparență (în IA)</p>	<p>Având în vedere statutul său de principiu etic fundamental discutat pe larg în ghid, „Transparența” ar fi fără îndoială definită. Transparența în IA se referă probabil la dezvăluirea clară, deschisă și comprehensibilă a modului în care tehnologiile IA sunt dezvoltate, implementate și utilizate, inclusiv datele pe care sunt antrenate, algoritmiile pe care îi folosesc, capacitățile și limitările lor, și raționamentul din spatele rezultatelor lor. Acest principiu urmărește să promoveze încrederea între utilizatori și părțile interesate, să permită responsabilitatea pentru rezultatele IA și să favorizeze un proces decizional informat cu privire la utilizarea IA. Ghidul subliniază „transparența în conținutul generat de IA” ca un capitol cheie și evidențiază că nedivulgarea utilizării IA în crearea de conținut poate eroda încrederea publicului.</p>
<p>Responsabilitate (în IA)</p>	<p>Ca un alt principiu etic fundamental evidențiat în tot ghidul, „Responsabilitatea” ar fi un termen crucial de definit. Responsabilitatea în contextul IA se referă probabil la atribuirea clară a responsabilității pentru dezvoltarea, implementarea și consecințele sistemelor IA și a rezultatelor acestora. Aceasta include stabilirea de cadre și procese care să asigure că supravegherea umană rămâne o parte integrantă a ciclului de viață al IA, astfel încât indivizi sau organizații să poată fi trași la răspundere pentru erori, prejudecăți sau impacturi dăunătoare rezultate din utilizarea IA. Ghidul susține pași pentru „stabilirea cadrelor de responsabilitate în organizații” și recunoaște că determinarea cine este responsabil pentru erorile din conținutul generat de IA este o provocare cheie.</p>
<p>Acuratețe (în IA)</p>	<p>Asigurarea fiabilității și imparțialității informațiilor produse de sistemele IA este o preocupare etică centrală în sectoarele de comunicare și informare, făcând „Acuratețea” un termen vital de definit. Acuratețea în acest context înseamnă probabil că sistemele IA generează în mod constant rezultate care sunt corecte din punct de vedere factual, precise, fiabile și lipsite de prejudecăți nejustificate, aliniate cu scopurile lor prevăzute. Acest principiu este fundamental pentru menținerea încrederii în media și informații, minimizarea răspândirii dezinformării și asigurarea că instrumentele IA contribuie pozitiv la societate. Ghidul abordează în mod specific „asigurarea informațiilor fiabile și imparțiale din sistemele IA” și metode pentru „verificarea conținutului generat de IA”.</p>

<p>Confidențialitate (în IA)</p>	<p>Protejarea datelor personale și sensibile în cadrul aplicațiilor de IA reprezintă o piatră de temelie a utilizării etice a IA, în special în sectoarele care procesează cantități enorme de informații, astfel încât "Confidențialitate" ar fi un termen cheie de definit. Definiția ar include probabil protecția informațiilor personale și sensibile ale persoanelor împotriva colectării, accesului, utilizării, divulgării, modificării sau distrugerii neautorizate. Acest principiu este esențial pentru susținerea demnității umane, respectarea standardelor legale și etice precum GDPR și menținerea încrederii în modul în care datele sunt gestionate în sistemele de IA. Ghidul abordează "protejarea datelor personale și sensibile în aplicațiile de IA" și cele mai bune practici pentru "gestionarea datelor și a consimțământului".</p>
<p>Proprietate Intelectuală (PI) (în IA)</p>	<p>Odată cu capacitatea tot mai mare a IA de a crea conținut în diverse medii, considerațiile etice și legale privind "Proprietatea Intelectuală" devin semnificative, făcând probabilă definiția sa. Definiția din ghid include probabil drepturile (precum copyright și brevete) și responsabilitățile legate de crearea, proprietatea și utilizarea conținutului generat de IA, a modelelor de IA și a datelor de antrenament. Aceasta ar aborda preocupări privind proprietatea operelor create de IA, utilizarea etică a materialelor de antrenament protejate de copyright și importanța respectării drepturilor creative.</p>
<p>Crearea de Conținut (IA)</p>	<p>Având în vedere o secțiune dedicată "Utilizării IA în Crearea de Conținut", acest termen ar fi probabil definit. Se referă la utilizarea tehnologiilor IA pentru a genera conținut nou, inclusiv texte, imagini, audio și video. Definiția ar putea aborda, de asemenea, diferitele tipuri de instrumente IA utilizate în acest scop, cum ar fi Modelele Lingvistice Mari (LLM) precum ChatGPT și instrumentele de generare de imagini precum DALL-E și Midjourney. Considerațiile etice specifice pentru conținutul generat de IA, cum ar fi transparența, responsabilitatea și proprietatea intelectuală, ar fi probabil evidențiate în legătură cu această definiție.</p>
<p>Dezinformare</p>	<p>Într-o eră în care neîncrederea este în creștere iar instrumentele IA pot fi utilizate pentru a genera și răspândi informații înșelătoare, definiția "Dezinformării" este crucială. Se referă probabil la informații false sau inexacte care sunt create și răspândite în mod deliberat cu intenția de a înșela sau induce în eroare publicul. Ghidul ar aborda probabil rolul IA atât în facilitarea, cât și în potențiala combatere a dezinformării.</p>
<p>Algoritm</p>	<p>Întrucât sistemele IA funcționează pe baza algoritmilor, o explicație simplificată a acestui termen ar fi utilă pentru un public larg. Un "Algoritm" este probabil definit ca un set de instrucțiuni sau reguli bine definite pe care un computer le urmează pentru a executa o sarcină sau a rezolva o problemă. Înțelegerea conceptului de bază al algoritmilor este importantă pentru a înțelege cum IA procesează datele și generează rezultate, deși funcționarea internă a unor sisteme GenAI este considerată drept "cutii negre".</p>
<p>Prejudecată (în IA)</p>	<p>Posibilitatea ca sistemele IA să manifeste și chiar să amplifice prejudecățile sociale existente este o provocare etică semnificativă. "Prejudecata în IA" se referă probabil la distorsiunea sistematică și nedreaptă a predicțiilor sau rezultatelor modelului IA, adesea derivată din date de antrenament distorsionate sau nereprezentative care reflectă prejudecăți sociale sau presupuneri eronate. Identificarea și atenuarea prejudecăților în IA este crucială pentru asigurarea echității și incluziunii.</p>
<p>Încălcarea Securității Datelor</p>	<p>Având în vedere accentul pus pe confidențialitate și securitatea datelor, conceptul de "Încălcarea Securității Datelor" este probabil relevant. Ar fi probabil definit ca un incident de securitate în care date sensibile, protejate sau confidențiale sunt copiate, transmise, vizualizate, furate sau utilizate de un individ sau o entitate neautorizată să facă acest lucru. Înțelegerea riscurilor și a potențialelor daune asociate încălcărilor securității datelor subliniază importanța practicilor de păstrare a confidențialității în IA.</p>

**Consimțământ
Informat**

În contextul colectării și utilizării datelor în aplicațiile IA, în special pentru instruirea modelelor sau personalizare, "Consimțământul Informat" este un concept fundamental. Se referă probabil la acordul voluntar dat de o persoană pentru a permite colectarea, procesarea și utilizarea datelor sale personale pentru scopuri specifice, bazat pe o înțelegere clară și completă despre ce date sunt colectate, cum vor fi utilizate și cine va avea acces la acestea. Obținerea consimțământului informat este o cerință etică fundamentală pentru respectarea drepturilor la confidențialitate ale persoanelor.

**GenAI (Inteligență
Artificială
Generativă)**

Odată cu ascensiunea sistemelor IA capabile să creeze conținut nou, "GenAI" sau "Inteligența Artificială Generativă" este un termen crucial de definit. Se referă probabil la o categorie de sisteme IA care pot genera conținut original precum text, imagini, audio sau alte forme de media, adesea bazate pe modele învățate din cantități vaste de date de instruire. Exemple precum GPT-4, DALL-E 2, Midjourney și Elevenlabs ar fi probabil menționate în legătură cu acest termen.

A2 – Referințe Legale și Politice

1. "Recomandarea privind Etica Inteligenței Artificiale" a UNESCO (2021)

Aceasta este menționată în mod explicit ca un document fundamental pe care se bazează puternic ghidurile etice. Ca atare, un link direct către documentul oficial UNESCO ar fi o referință primară în această secțiune. Recomandarea oferă un cadru universal pentru utilizarea etică a IA, subliniind principii precum demnitatea umană, echitatea și sustenabilitatea, care sunt reflectate în toate ghidurile "Me, Myself, and AI".

2. Legi și Reglementări privind Protecția Datelor

Având în vedere principiul etic fundamental al "Confidențialității" discutat în ghiduri, referințele la legile și reglementările cheie privind protecția datelor ar fi esențiale. Acestea ar include cel mai probabil: Regulamentul General privind Protecția Datelor (GDPR) (UE): Deoarece proiectul "Me, Myself, and AI" implică parteneri din state membre ale Uniunii Europene (România și Italia) și își propune să promoveze practici etice de IA în sectoarele de comunicare și informare din Europa, GDPR ar fi un cadru legal crucial de citat. GDPR stabilește reguli pentru colectarea și procesarea datelor personale ale persoanelor din UE, abordând aspecte precum consimțământul, drepturile persoanelor vizate și responsabilitățile operatorilor și procesatorilor de date. Relevanța sa pentru aplicațiile IA în gestionarea datelor utilizatorilor, instruirea modelelor IA și procesarea informațiilor personale în contextele media și de comunicare îl face o referință legală fundamentală. Legi Naționale Relevante privind Protecția Datelor: În funcție de contextul național specific al utilizatorilor ghidurilor, ar putea fi incluse și referințe la legile naționale de protecție a datelor care completează sau implementează GDPR. De exemplu, ghidurile ar putea indica legislația specifică privind protecția datelor din România și Italia.

3. Legi și Reglementări privind Proprietatea Intelectuală

Cu accentul important pe "Proprietatea Intelectuală" din ghiduri, această secțiune ar include fără îndoială referințe la cadrele legale cheie care guvernează drepturile de PI: Legi privind Drepturile de Autor: Aceste legi protejează operele originale de autor, precum lucrările literare și artistice. În contextul IA, drepturile de autor sunt relevante pentru proprietatea asupra conținutului generat de IA, utilizarea materialului protejat de drepturi de autor pentru instruirea modelelor IA și posibila încălcare a drepturilor de autor de către sistemele IA. Ar fi probabil incluse referințe la convențiile și legislațiile naționale și internaționale privind drepturile de autor (de exemplu, Convenția de la Berna). Legi privind Brevetele: Deși poate mai puțin direct relevante pentru sectoarele de comunicare și informare comparativ cu drepturile de autor, legile privind brevetele protejează invențiile. Dacă instrumentele sau procesele IA dezvoltate în cadrul acestor sectoare sunt brevetabile, ar putea fi incluse referințe la legislația relevantă privind brevetele. Legi privind Secretele Comerciale: Aceste legi protejează informațiile comerciale confidențiale.

Dacă algoritmi de IA sau datele de antrenament sunt considerate secrete comerciale, ar putea fi citate cadrele juridice relevante. Legi referitoare la colectarea de date (data scraping) și utilizarea neautorizată a conținutului online: Având în vedere considerațiile etice privind utilizarea datelor pentru antrenarea IA, ar fi relevante referințele la legile care abordează legalitatea colectării de date și utilizarea neautorizată a conținutului online, în special a materialelor protejate de drepturi de autor.

Libertatea de exprimare și dreptul mass-media

În contextul jurnalismului și mass-media, unde IA este din ce în ce mai integrată, ar fi relevante referințele la cadrele juridice care protejează libertatea de exprimare și reglementează mass-media. Acestea ar putea include: Garanții constituționale ale libertății de exprimare și ale presei: Referințe la articolele relevante din constituțiile țărilor în care ghidurile sunt destinate a fi utilizate (de exemplu, România și Italia) care garantează libertatea de exprimare și libertatea presei. Legi și reglementări privind mass-media: Legi și reglementări naționale care guvernează conduita, acuratețea și responsabilitatea mass-media. Acestea ar putea fi relevante în contextul știrilor și informațiilor generate de IA pentru a asigura respectarea standardelor etice și legale.

Legislație privind IA și tehnologiile emergente

Uniunea Europeană a stabilit Regulamentul UE privind Inteligența Artificială (Regulamentul (UE) 2024/1689), care a intrat în vigoare la 1 august 2024. Acest regulament stabilește un cadru juridic pentru sistemele de IA din UE, folosind o abordare bazată pe risc cu obligații variabile în funcție de nivelul de risc prezentat de IA. Include interdicții pentru IA cu risc inacceptabil, reguli stricte pentru IA cu risc ridicat și cerințe de transparență pentru anumite sisteme de IA, inclusiv cele de IA pentru scopuri generale. Regulamentul își propune să promoveze inovația abordând în același timp potențialele riscuri ale IA, cu cerințe cheie care vor deveni obligatorii în februarie 2025 și aplicare deplină până în august 2026.

Standarde și ghiduri ale organizațiilor profesionale

Deși nu sunt strict de natură "legală", ar putea fi incluse referințe la coduri etice, ghiduri sau standarde dezvoltate de organizații profesionale în domeniile comunicării, informației și tehnologiei. Diverse organizații profesionale dezvoltă standarde și ghiduri pentru IA pentru a promova o dezvoltare și implementare etică și responsabilă. Iată câteva exemple cheie:

IEEE (Institute of Electrical and Electronics Engineers)

Oferă o gamă de standarde legate de IA și sisteme autonome, concentrându-se pe considerații etice, transparență și responsabilitate. Puteți explora seria lor de standarde SA-P7000. Link: <https://standards.ieee.org/> (Căutați "AI" sau "autonomous systems" în secțiunea de standarde)

ISO/IEC JTC 1/SC 42 (Inteligență artificială)

Acest comitet tehnic comun dezvoltă standarde internaționale pentru inteligența artificială. Munca lor include standarde privind sistemele de management al IA (ISO/IEC 42001) și diverse aspecte ale fiabilității IA. Link: <https://www.iso.org/standard/81230.html>

OCDE (Organizația pentru Cooperare și Dezvoltare Economică)

A dezvoltat Principiile OCDE privind IA, primul standard interguvernamental privind IA, promovând o IA de încredere care respectă drepturile omului și valorile democratice. Link: <https://www.oecd.org/en/topics/artificial-intelligence.html>

IAPP (International Association of Privacy Professionals)

Oferă certificarea Artificial Intelligence Governance Professional (AIGP), concentrându-se pe crearea și implementarea de politici și proceduri pentru o guvernare responsabilă a IA. Link: <https://iapp.org/certify/aigp/>

Convenții și Tratatate Internaționale Relevante

Iată câteva convenții și tratate internaționale relevante, pe lângă Convenția de la Berna, care se referă la protecția datelor, proprietatea intelectuală sau drepturile omului în era digitală, cu exemple și link-uri:

Protecția Datelor:

Convenția pentru Protecția Persoanelor față de Prelucrarea Automată a Datelor cu Caracter Personal (Convenția 108) și Protocolul său Adițional: Această convenție a Consiliului Europei, împreună cu versiunea sa modernizată Convenția 108+, este un instrument internațional cheie pentru protecția datelor. Stabilește principii pentru prelucrarea echitabilă și legitimă a datelor personale, inclusiv garanții pentru drepturile individuale. Exemplu de Relevanță pentru IA: Deoarece sistemele de IA se bazează adesea pe procesarea unor cantități mari de date personale, aceste convenții oferă un cadru pentru a asigura că această prelucrare respectă confidențialitatea și drepturile fundamentale ale persoanelor.

Link către Convenția 108: <https://rm.coe.int/1680078b47>

Link către Convenția 108+: <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/223>

Proprietate Intelectuală:

Tratatul OMPI privind Dreptul de Autor (WCT): Acest tratat, adoptat sub auspiciile Organizației Mondiale a Proprietății Intelectuale (OMPI), actualizează și completează Convenția de la Berna în mediul digital. Abordează protecția operelor și drepturile autorilor în sfera online. Exemplu de Relevanță pentru IA: WCT este relevant pentru protecția drepturilor de autor ale operelor creative generate de IA și utilizarea materialelor protejate de drepturi de autor în instruirea modelelor de IA.

Link: <https://www.wipo.int/treaties/en/ip/wct/>

Tratatul OMPI privind Interpretările și Execuțiile și Fonogramele (WPPT): Acest tratat OMPI se ocupă de drepturile interpreților și producătorilor de fonograme (înregistrări sonore) în era digitală. Exemplu de Relevanță pentru IA: Poate fi relevant în contextul muzicii generate de IA sau al utilizării înregistrărilor sonore existente în aplicațiile de IA.

Link: <https://www.wipo.int/treaties/en/ip/wppt/>

Drepturile Omului în Era Digitală:

Pactul Internațional cu privire la Drepturile Civile și Politice (ICCPR): Acest tratat al Națiunilor Unite protejează o gamă largă de drepturi civile și politice, inclusiv dreptul la libertatea de exprimare, dreptul la viață privată și dreptul la un proces echitabil. Exemplu de Relevanță pentru IA: ICCPR este relevant pentru preocupările legate de tehnologiile de supraveghere îmbunătățite prin IA, potențialul de prejudiciu în procesul decizional al IA care ar putea afecta drepturile fundamentale și impactul IA asupra libertății de exprimare.

Link: <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

Pactul Internațional cu privire la Drepturile Economice, Sociale și Culturale (ICESCR): Acest tratat ONU protejează drepturile economice, sociale și culturale, cum ar fi dreptul la muncă, dreptul la educație și dreptul la sănătate. Exemplu de Relevanță pentru IA: ICESCR este relevant pentru discuțiile despre impactul IA asupra ocupării forței de muncă, accesul la educație și informații, și asistența medicală.

Link: <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights>

Convenția privind Criminalitatea Informatică (Convenția de la Budapesta): Această convenție a Consiliului Europei are ca scop combaterea criminalității informatice și include dispoziții privind infracțiunile împotriva confidențialității, integrității și disponibilității sistemelor și datelor informatice, precum și pornografia infantilă și încălcarea drepturilor de autor. Exemplu de Relevanță pentru IA: Această convenție este relevantă pentru abordarea amenințărilor la adresa securității informatice legate de sistemele de IA și utilizarea necorespunzătoare a IA pentru activități criminale.

Link: <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185>