



Codi sobre l'Ètica de la Intel·ligència Artificial d'Andorra



Govern d'Andorra

Índex

I. Introducció.....	3
II. Glossari de termes	9
III. Àmbit d'aplicació.....	12
IV. Fins i objectius	14
V. Valors i principis	17
V.1. Valors	18
V.2. Principis.....	20
VI. Marc per a una intel·ligència artificial fiable	23
VI.1. Fonaments i requisits per a una IA fiable.....	24
VI.2. Mètodes i tècniques per al foment de la IA fiable	26
VI.3. “Human Rights By Design (HRbD)” a l’era de la IA	28
VII. Intel·ligència artificial i criteris ESG.....	29
VIII. Drets digitals i neurodrets	31
IX. IA Generativa (GenAI)	33
X. Àmbits tractors d’actuació	35
X.1. Àmbit d’actuació 1: sector públic.....	36
X.2. Àmbit d’actuació 2: sector privat.....	39
X.3. Àmbit d’actuació 3: ciutadania	42
XI. Impuls i promoció d’aquest Codi	44
XII. Annex. Referències (bibliografia i webgrafia).....	46

I. Introducció

La intel·ligència artificial (IA) està transformant la societat, l'economia i la forma en què opera tant el sector públic i el Govern¹ en l'exercici de les seves funcions, com el sector privat, i planteja, en general, nous desafiaments i reptes per als drets dels ciutadans, així com oportunitats d'avenç lligades a la innovació, la competitivitat i el bé comú.

Tots aquests aspectes han de ser considerats per traçar estratègies coherents i equilibrades, de manera que es permeti **desenvolupar i extreure tot el potencial inherent a aquesta tecnologia emergent sense que minvin les garanties que han de ser considerades per a la millor protecció dels drets i dels interessos legítims** que es puguin veure afectats.

Això requereix desplegar un enfocament holístic i responsable que fomenti la **confiança en l'aplicació i l'ús d'aquesta tecnologia (enfocament IA confiable²)** i que, per tant, n'afavoreixi el desplegament adequat de forma responsable **posant les persones al centre d'aquest desenvolupament (human-centered AI³)**.

En aquest sentit, la comunitat internacional està implementant, en l'àmbit comparat i regional, diverses iniciatives que promouen l'ús legal i ètic de la intel·ligència artificial. Resulten destacables, entre d'altres, i als efectes d'aquest Codi sobre l'ètica de la intel·ligència artificial d'Andorra (d'ara endavant, "el Codi"), les accions següents:

- a) **L'Organització de les Nacions Unides per a l'educació, la ciència i la cultura (UNESCO)⁴:** aquesta entitat compta amb una línia d'acció ben definida al voltant de la intel·ligència artificial, i ha establert el marc internacional més complet del món per promoure un ús ètic de la IA. Així, la Recomanació sobre l'ètica de la intel·ligència artificial⁵ va ser adoptada per 193 estats membres a la Conferència general de la UNESCO el mes de novembre de l'any 2021 i constitueix, actualment, el marc mundial més acceptat en la matèria.

Així mateix, entre altres iniciatives llançades per la UNESCO destaquen, en primer lloc, el Model d'avaluació d'impacte ètic (EIA),⁶ que és part d'un pla d'implementació més ampli de la recomanació anterior i que, alhora, complementa una altra eina elaborada per aquesta organització, com és la Metodologia d'avaluació de la preparació (RAM).⁷ La RAM ajuda els governs a avaluar com són de sòlides i àgils les seves lleis, polítiques i institucions per abordar els riscos de la IA.⁸ En segon lloc, també són ressenyables les iniciatives al voltant d'aquesta intel·ligència en l'àmbit educatiu i de la investigació,⁹ les línies d'acció en l'àmbit de la protecció dels neurodrets¹⁰ i les actuacions en l'àmbit de l'ètica digital i la dona.¹¹

- b) **L'Organització de les Nacions Unides (ONU):¹²** Andorra forma part de les Nacions Unides com a membre de ple dret des de l'any 1993. Per la seva banda, el mes de març del 2024, l'Assemblea General de l'ONU va adoptar una resolució històrica sobre la promoció de sistemes d'intel·ligència artificial (IA) "segurs i fiables" que beneficiïn també el desenvolupament sostenible per a tothom. El text va tenir el suport de més de 120 estats membres. L'Assemblea General també va reconèixer el potencial dels sistemes d'IA per accelerar i permetre el progrés cap a la consecució dels 17 Objectius de desenvolupament sostenible.¹³

Aquesta resolució anima els estats membres i convida múltiples interessats, inclosos el sector privat, la societat civil, els cercles acadèmics i les institucions de recerca i comunitats i persones de l'àmbit tècnic, entre altres entitats, a elaborar, dins de les seves funcions i responsabilitats respectives, enfocaments i marcs normatius i de governança referents a sistemes segurs, protegits i fiables d'intel·ligència artificial, i a donar-hi suport. Aquests enfocaments i marcs han de crear un ecosistema propici en tots els nivells, entre altres coses per a la innovació, l'emprenedoria i la difusió de coneixements i tecnologies en condicions acordades mútuament, reconeixent que les

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

aliances i la cooperació eficaces entre els governs i múltiples interessats són necessàries per elaborar aquests enfocaments i marcs.

Precisament, aquest Codi pretén impulsar aquesta recomanació i oferir a les parts interessades orientació sobre certes eines de suport existents en l'àmbit internacional per desenvolupar aquests enfocaments i marcs fiables. De la mateixa manera, és molt important conèixer els principis rectors sobre les empreses i els drets humans de l'ONU i poder-s'hi atènyer (posada en pràctica del marc de les Nacions Unides per "protegir, respectar i posar remei").¹⁴

- c) **L'Organització per a la Cooperació i el Desenvolupament Econòmics (OCDE):**¹⁵ aquesta organització internacional desenvolupa, des de fa anys, diverses estratègies al voltant de la intel·ligència artificial per assegurar que aquesta tecnologia aportí el màxim benefici possible per a la societat en el seu conjunt, i intenta mitigar en paral·lel i al màxim possible els riscos per a les persones i els seus drets, així com altres possibles riscos associats. Amb aquesta finalitat, l'OCDE dona suport als governs i a altres entitats amb informació, metodologies, eines¹⁶ i altres recursos disponibles per generar estratègies i polítiques nacionals al voltant d'una IA confiable. En aquest àmbit, l'Observatori de Polítiques d'IA de l'OCDE (OECD.AI)¹⁷ combina recursos de tota l'OCDE i dels seus socis, i facilita el debat i la reflexió compartida proporcionant dades d'impacte d'una rellevància enorme.¹⁸

Sobre aquest tema, destaca la Recomanació del Consell sobre intel·ligència artificial,¹⁹ que conté una **definició del que s'entén per sistemes d'IA**,²⁰ definició que s'ha incorporat en regulacions sobre IA com l'europea,²¹ així com els **principis**²² **per a una gestió responsable d'una IA confiable**.²³ Hi ha altres posicions recents importants, per exemple en matèria d'IA generativa (GenAI),²⁴ que s'esmentaran al llarg d'aquest Codi.

Juntament amb aquests principis, s'estableixen certes recomanacions per als responsables de la formulació de polítiques, entre les quals destaca la inversió en R+D de la IA; el foment d'un ecosistema digital per a la IA; propiciar un entorn polític adequat per a la IA; el desenvolupament de la capacitat humana, la preparació per a la transformació del mercat laboral²⁵ i la cooperació internacional per a una IA confiable.²⁶

En general, l'OCDE treballa per fomentar la cooperació internacional per a una IA confiable (una de les recomanacions principals per als responsables de la formulació de polítiques al voltant dels principis anteriors) i propicia:

- L'organització de fòrums globals i regionals per impulsar l'intercanvi de coneixements sobre IA i, amb això, de bones pràctiques, i la generació d'estàndards tècnics globals impulsats per consens i de múltiples parts interessades per a una IA interoperable i responsable.
- L'establiment, a través dels governs, de mètriques comparables internacionalment per mesurar la recerca, el desenvolupament i la implementació de la IA, i amb això poder avaluar el progrés en la implantació dels principis esmentats.

Resulta útil el Catàleg d'eines i mètriques per a una IA confiable que posa a disposició aquesta organització, en què aporta recursos pràctics de cara a establir un marc de governança de la IA.²⁷ De la mateixa manera, destaca el Marc de control d'incidents de la IA (AIM), que ajuda a mostrar patrons i a establir una comprensió col·lectiva dels incidents de la IA i de la seva naturalesa, cosa que serveix, com diu la mateixa OCDE, com a eina fonamental per aconseguir una IA confiable.²⁸

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

- d) **El Consell d'Europa (COE):** el COE també ha dut a terme un ampli treball entorn de la intel·ligència artificial i els drets humans,²⁹ i ha creat el Comitè sobre Intel·ligència Artificial (CAI),³⁰ que pren el relleu de la tasca prèvia efectuada per un altre comitè anterior (CAHAI)³¹ i que treballa en l'elaboració d'un conveni marc sobre el desenvolupament, el disseny i l'aplicació de la intel·ligència artificial,³² entre altres iniciatives relacionades amb la regulació de la IA.³³ Andorra treballa de manera activa en els grups de treball associats al desenvolupament d'aquest conveni marc, i espera poder adherir-s'hi quan formalment sigui possible fer-ho.

De la mateixa manera, el Comitè Europeu de Cooperació Jurídica (CDCJ) d'aquesta organització internacional està desenvolupant diverses iniciatives per afrontar els reptes nous i emergents que sorgeixen en l'àmbit del dret públic,³⁴ el dret privat i l'Administració de justícia, entre moltes altres línies de treball i estudis d'interès.³⁵

Sens dubte, el treball en curs del Consell d'Europa sobre aquests temes és prolix³⁶ i interessant per la gran transversalitat de temes que s'aborden (privacitat, justícia i Administració pública, salut i biomedicina, no-discriminació i igualtat de gènere,³⁷ drets socials, drets dels nens, joventut, educació, llibertat d'expressió, cultura, etc.).

- e) **Unió Europea (UE):** sense perjudici de les iniciatives anteriors, la UE, com a regió, també compta amb una estratègia definida en intel·ligència artificial.³⁸ L'any 2018, la Comissió Europea va emetre unes directrius ètiques per a una IA fiable,³⁹ i a l'abril de l'any 2021 la Comissió també va presentar el seu paquet d'IA, que inclou una comunicació sobre el foment d'un enfocament europeu de la IA, una revisió del Pla coordinat d'intel·ligència artificial (amb els estats membres de la UE) i la proposta del marc regulador sobre intel·ligència artificial. Aquesta proposta regulatòria (Reglament europeu d'intel·ligència artificial (AI Act)) finalment es convertirà en el primer marc regulador de tipus transversal del món, després que s'hagi assolit un acord sobre el seu contingut el mes de desembre del 2023⁴⁰ i que s'hagi aprovat posteriorment per part del Parlament Europeu.⁴¹

Tot això, a més, està en consonància amb l'estratègia continguda en la denominada "Brúixola digital europea (Dècada Digital)"⁴² i en la Declaració sobre drets i principis digitals europeus (Declaració sobre els drets i principis digitals per a la Dècada Digital), que promou una transició digital cisellada pels valors europeus.⁴³

Així mateix, recentment s'ha creat l'Oficina Europea d'IA⁴⁴ i el paquet d'innovació en IA per donar suport a les empreses emergents i les pimes en matèria d'intel·ligència artificial.⁴⁵

En tot cas, la UE continua treballant en fòrums com els del G7, l'OCDE, el Consell d'Europa, el G20⁴⁶ i les Nacions Unides per promoure normes consensuades sobre una IA fiable en l'àmbit internacional. Per exemple, ha donat suport a l'acord dels líders del G7 en el marc del procés d'IA d'Hiroshima sobre principis rectoris internacionals i a un codi de conducta voluntari per a sistemes avançats d'IA.⁴⁷

- f) **Altres iniciatives internacionals d'interès:** al marge de les línies de treball anteriors, n'hi ha moltes altres en l'àmbit internacional⁴⁸ que resulten de gran interès.

Entre altres, destaquen els treballs duts a terme per l'Associació Global sobre Intel·ligència Artificial (GPAI),⁴⁹ que és una iniciativa de múltiples parts interessades que té com a objectiu avançar cap a la IA. De la mateixa manera, els treballs efectuats pel Fòrum Econòmic Mundial o Fòrum de Davos (WEF, per les sigles en anglès)⁵⁰ pel que fa a la IA⁵¹ són ressenyables.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

Així mateix, també són interessants els avenços del Regne Unit en relació amb la intel·ligència artificial,⁵² amb una atenció especial als aspectes reguladors amb un enfocament a la innovació,⁵³ entre altres documents polítics i directrius.⁵⁴

També són rellevants les iniciatives i els esforços desenvolupats pels Estats Units⁵⁵ en aquest àmbit, i en destaca la recent Ordre executiva sobre el desenvolupament i ús segur i fiable de la intel·ligència artificial i els posteriors actes d'execució d'aquesta ordre.⁵⁶ En aquest àmbit, també destaca com a marc de referència de control de riscos en intel·ligència artificial el marc proposat pel Centre de Recursos d'IA Fiable i Responsable del NIST, en col·laboració amb els sectors públic i privat nord-americans.⁵⁷

En aquest context mundial, Andorra compta amb el **Programa de transformació digital (PdTDA)**, emmarcat en el Pla estratègic de transformació digital 2030, dissenyat pel Govern d'Andorra, els objectius del qual convergeixen alhora amb els objectius 2030 de l'Agenda Digital Europea.

Així, aquesta estratègia de digitalització 2020-2030 per a Andorra s'operativitza actualment a través dels quatre eixos estratègics que vertebreren el PdTDA en el període 2024-27: 1) Avenç en la digitalització de l'Administració pública; 2) Impuls a la transformació digital del teixit empresarial privat; 3) Consolidació i evolució de les bases tecnològiques actuals i impuls a l'adopció de noves tecnologies, i 4) Donar cobertura als drets i els principis digitals del ciutadà.

En el context d'aquests eixos estratègics, amb la finalitat d'establir des del PdTDA les bases per a una transformació digital responsable i sostenible, és essencial assegurar l'**adopció ordenada de tecnologies emergents clau per al desenvolupament socioeconòmic del país, com és la intel·ligència artificial**. En aquest sentit, Andorra treballa de manera molt activa en grups de treball internacionals tendents al foment de la IA fiable, com per exemple en el si del Consell d'Europa o la UNESCO.

Així mateix, la futura **Agència d'Intel·ligència de la Dada d'Andorra** serà l'entitat encarregada de promoure la governança i la supervisió de la IA a Andorra, fet que suposa un pas fonamental en l'avenç cap a la consecució de models responsables i fiables en relació amb la intel·ligència artificial, d'acord amb el PdTDA.

En conseqüència, pel que fa als objectius del PdTDA, i de cara a conformar el futur marc regulador a Andorra d'intel·ligència artificial, **cal disposar d'un codi ètic de la IA que ajudi en l'aportació de certes directrius, recomanacions i orientacions sobre el desplegament, l'ús i l'aplicació ordenada de la IA al sector públic, al sector privat i entre la ciutadania**, en general. Dins aquests sectors i parts s'hi inclou la referència a l'àmbit educatiu i de recerca.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

En definitiva, a través d'aquest Codi es pretén:

1. **Servir de guia**, a través de l'aportació de certes directrius i recomanacions, per **a les parts interessades i els destinataris del mateix Codi**, amb l'únic objectiu d'oferir-los el suport o els recursos bàsics necessaris per al desplegament i l'aplicació pràctica de models i marcs tendents a l'impuls de la IA confiable en tot el cicle de vida dels sistemes IA que correspongui.⁵⁸
2. **Establir certes bases de cara al possible proper desenvolupament regulador de la IA a Andorra**, la qual cosa pivotaria sobre dos eixos fonamentals, que són: a) la promoció de la IA legal, és a dir, que els sistemes d'IA compleixin la normativa que resulti aplicable a aquest efecte (una vegada es desenvolupi i promulgui d'acord amb la llei), i b) el foment de la IA ètica, en relació amb els principis i valors associats a l'ètica digital, atès el ple respecte dels drets fonamentals de les persones.

Per tant, **aquest Codi pretén merament oferir orientacions sobre el foment i la garantia d'una IA ètica (enfocament basat en els drets fonamentals).**

En resum: Andorra, en línia amb altres iniciatives internacionals i nacionals similars, que succeeixen en l'àmbit global, per al foment de la IA de confiança, així com amb els objectius i eixos estratègics del Programa de transformació digital (PdTDA) del país, advoca per donar suport a models i ecosistemes digitals responsables, apostant per la innovació sense cap minvament dels drets i les llibertats fonamentals de les persones garantits per la Constitució i la legislació andorrana. El Codi d'ètica de la intel·ligència artificial d'Andorra busca secundar, guiar i orientar el sector públic, el sector privat i la ciutadania (parts interessades i destinataris del Codi), en el necessari camí cap a la plena consecució i foment de la IA confiable en els diferents àmbits d'actuació de què es tracti.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

II.

Glossari de termes

A continuació, es defineixen i desenvolupen certs termes d'aquest Codi perquè s'entenguin millor:

- **Sistema d'intel·ligència artificial (IA):** un sistema d'IA és un sistema basat en una màquina que, per a objectius explícits o implícits, infereix, a partir de l'entrada que rep, com generar resultats com ara prediccions, contingut, recomanacions o decisions que poden influir en entorns físics o virtuals. Els diferents sistemes d'IA varien en els nivells d'autonomia i adaptabilitat després del desplegament.⁵⁹
- **Cicle de vida del sistema d'IA:** es refereix, amb caràcter general, a les fases següents: i) disseny, dades i models, que és una seqüència que depèn del context que abasta la planificació i el disseny, la recopilació i el processament de dades, així com la construcció de models; ii) verificació i validació, iii) desplegament i iv) operació i seguiment. Aquestes fases solen tenir lloc de manera iterativa i no necessàriament són seqüencials.
- **Parts interessades:** es refereix a totes les organitzacions, públiques o privades, i individus o ciutadans involucrats o afectats pels sistemes d'IA, directament o indirectament, a Andorra. Dins d'aquestes organitzacions, ja siguin públiques o privades, s'hi inclouen el sector educatiu i el de la recerca. Els actors de la IA, s'integrin en el sector públic o en el sector privat andorrà, són un subconjunt de parts interessades.
- **Actors de la IA:** són els que exerceixen un paper actiu en el cicle de vida del sistema d'IA a Andorra, s'integrin en el sector públic, en el Govern o en el sector privat, incloses les organitzacions i les persones que implementen o operen la IA en aquest país o amb efecte en aquest país.
- **IA confiable:** es refereix a la IA que respecta els principis basats en valors d'acord amb el que descriu aquest Codi. La fiabilitat és un requisit previ perquè les entitats, les organitzacions, les persones i les societats desenvolupin, despleguin i utilitzin sistemes d'IA. Si aquests sistemes –i les persones que es troben al darrere– no demostren ser mereixedors de confiança, es poden produir conseqüències no desitjades que n'obstaculitzin l'adopció, i que impedeixin que s'assoleixin els enormes beneficis econòmics i socials que poden comportar els sistemes d'IA. La fiabilitat de la intel·ligència artificial (IA) es recolza en tres components que s'han de poder complir al llarg de tot el cicle de vida del sistema:
 - **La IA ha de ser lícita**, de manera que es garanteixi el respecte a la normativa aplicable.
 - **La IA ha de ser ètica**; és a dir, ha d'assegurar que es compleixin els principis i valors ètics que siguin aplicables, i ha de poder servir d'orientació i guia a les parts interessades d'aquest Codi.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

- **La IA ha de ser robusta** i ha d'assegurar un control adequat i mitigar els riscos que puguin concórrer, garantir la seguretat i evitar danys, i és que les persones i la societat han de poder confiar que els sistemes d'IA no els provocaran un dany involuntari i que no afectaran els seus drets i llibertats fonamentals. En aquest context, la robustesa significa la capacitat de resistir o superar condicions adverses, inclosos els riscos de seguretat digital. Els sistemes d'IA no han de plantejar riscos de seguretat no raonables –conforme a l'estat de la tècnica en cada moment–, inclosa la seguretat física, en condicions d'ús o mal ús normal o previsible durant tot el cicle de vida.

Garantir la confiança de la IA és una responsabilitat individual i col·lectiva compartida per totes les parts interessades, amb una atenció especial als actors de la IA. Els actors de la IA poden fer servir un enfocament de gestió de riscos per identificar-los, avaluar-los, prioritzar-los i mitigar-los.

III.

Àmbit d'aplicació

El Codi d'ètica de la intel·ligència artificial d'Andorra aborda qüestions ètiques relacionades amb l'àmbit de la intel·ligència artificial, i s'alinea amb la Recomanació sobre l'ètica de la intel·ligència artificial de la UNESCO; els treballs en curs del Conveni marc sobre el disseny, el desenvolupament i l'aplicació de sistemes d'intel·ligència artificial basats en les normes del Consell d'Europa en matèria de drets humans, democràcia i estat de dret, i propicis per a la innovació; la Recomanació del Consell de l'OCDE sobre intel·ligència artificial, i les directrius ètiques per a una IA fiable de la Comissió Europea, entre altres textos, i documents internacionals rellevants recollits en l'apartat I d'aquest Codi.

En aquest sentit, el Codi ofereix una base i un punt de partida per al necessari foment de la de IA confiable en totes les fases del cicle de vida dels sistemes d'IA (disseny, desenvolupament, desplegament, aplicació o utilització d'IA, etc.), d'acord amb les diverses iniciatives internacionals en aquest mateix sentit, i aporta, per a això, certes directrius i recomanacions amb aquest propòsit i que poden evolucionar en el temps, amb un focus especial en tres àmbits principals d'actuació:

- a) El sector públic.
- b) El sector privat.
- c) La ciutadania, en general.

Per consegüent, aquest Codi ofereix una orientació als seus destinataris (parts interessades), és a dir, a les entitats associades al sector públic i al Govern, al sector privat i als ciutadans. L'objectiu primordial és que aquestes parts comptin amb informació, recursos i elements d'interès per identificar i afrontar les principals implicacions, desafiaments i oportunitats associats al desplegament i a l'ús ètic de la IA en les seves àrees d'actuació respectives, i que, en tot cas, puguin actuar de manera respectuosa amb els drets i interessos legítims concernits. També persegueix propiciar la recerca i la innovació responsables.

Les parts interessades, i en concret els actors d'IA compromesos amb una IA fiable, poden optar voluntàriament per utilitzar aquest Codi com a forma de materialitzar el seu compromís. No obstant això, **el sector públic pot requerir l'adhesió al mateix Codi per part de certes entitats com a part dels requisits d'accés o de participació en determinats processos de contractació pública,** o en una altra mena de processos públics, sempre que la normativa aplicable a Andorra ho permeti.

Finalment, les directrius i recomanacions contingudes en aquest Codi **no pretenen reemplaçar cap política o normativa actual o futura a Andorra, ni impedir-ne el desenvolupament o l'aprovació futura.** A més, aquest Codi s'ha de considerar com un **document viu i en evolució contínua, que caldrà revisar i actualitzar quan sigui necessari o s'estimi convenient,** precisament per garantir-ne la coherència amb la normativa que pugui resultar aplicable, amb l'evolució de la tecnologia, l'entorn social i econòmic o el desenvolupament de la cooperació internacional i els acords existents en aquest àmbit.

En resum: aquest Codi ofereix una orientació als seus destinataris (parts interessades), és a dir, a les entitats associades al sector públic, al sector privat i als ciutadans. L'objectiu primordial és que aquestes parts comptin amb informació, recursos i elements d'interès per identificar i afrontar les principals implicacions, desafiaments i oportunitats associats al desplegament i l'ús ètic de la IA en les seves àrees d'actuació respectives, i que, en tot cas, puguin actuar de manera respectuosa amb els drets i interessos legítims concernits. També persegueix propiciar la recerca i la innovació responsables.

IV. Fins i objectius

Aquest Codi pren com a referència bàsica, i sense ànim exhaustiu, els aspectes següents:

- Promoure la transformació digital i la tecnologia responsable i confiable, sota paràmetres de disseny legal i ètic de tota mena d'iniciatives, projectes, serveis i productes en l'entorn digital.
- Fomentar el benestar comú i l'interès general.
- Protegir i defensar la dignitat i l'autonomia humanes, i els drets bàsics de les persones, posant aquests conceptes al centre.
- Donar suport als drets digitals, fent esment especialment de la transparència algorítmica i el risc d'exclusió i de bretxa digital, així com promoure la igualtat i evitar situacions derivades del biaix algorítmic.
- Atendre les necessitats específiques de persones i grups de persones vulnerables, com els menors d'edat, els joves, les persones grans i les persones amb discapacitat, entre altres possibles col·lectius.
- Considerar situacions caracteritzades per asimetries de poder o en la relació.
- Prevenir de forma general els riscos, les amenaces i els danys derivats de sistemes d'IA, promovent-ne la robustesa i la seguretat.
- Impulsar la conscienciació, la formació i l'impuls de les capacitats digitals al voltant de la IA.
- Assegurar que el possible desplegament de l'actuació pública a través de sistemes d'IA respongui, en tot cas, a criteris transparents i ètics.
- Potenciar la llibertat d'empresa i l'ètica digital al sector privat.
- Donar suport a les empreses emergents i a les pimes en matèria d'intel·ligència artificial.
- Propiciar la innovació i la recerca responsables.
- En general, procurar un enfocament divers, inclusiu i sostenible.

Per tant, aquest Codi té com a finalitats i objectius primordials els següents:

Objectius i finalitats
Situat les persones al centre del desenvolupament, desplegament i ús de sistemes d'IA.
Promoure i respectar els drets humans i les llibertats fonamentals.
Ajudar el sector privat en el desplegament de polítiques d'IA fiable d'acord amb criteris d'ètica digital.
Garantir, per part del sector públic, la plena promoció i aplicació de paràmetres responsables, transparents i ètics a la IA.
Donar suport a les empreses emergents i a les pimes en matèria d'intel·ligència artificial.
Procurar que hi hagi més conscienciació, informació i educació sobre els riscos, els desafiaments i les oportunitats de la IA.
Impulsar la innovació i la recerca responsables en IA.
Fomentar la protecció del medi ambient, la sostenibilitat i la consecució dels indicadors ESG.
Estimular la utilització dels sistemes d'IA amb fins pacífics.
Garantir el desenvolupament socioeconòmic i el bé comú amb suport en sistemes d'IA fiables.

En resum: aquest Codi promou, d'acord amb altres posicions a escala internacional, determinats objectius i finalitats, entre els quals destaquen situar les persones al centre del desenvolupament, el desplegament i l'ús de sistemes d'IA, i protegir-ne els drets i les llibertats fonamentals, sense que això suposi un fre per a la innovació i la competitivitat empresarials. És fonamental atendre els drets dels grups de persones més vulnerables, com són els menors, els joves, les persones grans i les persones amb discapacitat.

Promoure, entre altres aspectes, l'alfabetització en intel·ligència artificial i la reducció de la bretxa digital, i fomentar la sostenibilitat també són objectius prioritaris a Andorra.

V. Valors i principis

V.1. Valors

V.2. Principis

Aquest Codi parteix de certs valors i aborda determinats principis importants basats en els principis per promoure estratègies d'IA fiable entre els actors de la IA, pertanyin al sector públic o al sector privat.

V.1. Valors

La intel·ligència artificial ha d'actuar com una **eina al servei de les persones i de la societat i ha de promoure, en tot cas, el benestar humà i la sostenibilitat**. Per consegüent, és molt important partir de certs valors en relació amb els sistemes d'IA. Aquests valors són, com no pot ser d'altra forma, consubstancials als mateixos valors democràtics i a l'ordenament jurídic vigent a Andorra, i es concreten en els punts següents:

- a) **Desenvolupar una estratègia de transformació digital centrada en les persones:** és crucial que els sistemes d'IA se centrin en les persones, de manera que les parts interessades assumeixin el compromís d'utilitzar aquesta tecnologia al seu servei i al servei de la humanitat, la consecució de les seves aspiracions, la solidaritat i, en general, el bé comú, sense deixar ningú enrere. Això implica tractar de maximitzar els beneficis i les oportunitats dels sistemes d'IA en benefici de la societat i les persones i de la seva plena inclusió digital, i fomentar un nivell elevat de protecció per a la seguretat i la garantia dels seus drets fonamentals.
- b) **Promoure el disseny legal i ètic de projectes, iniciatives, productes i serveis basats en IA:** s'ha d'atendre un criteri que tendeixi, amb caràcter general, al disseny legal i ètic des de l'inici, i durant tot el desenvolupament, en qualsevol projecte, iniciativa, producte i servei basat en IA, tingui caràcter públic o privat, prenent en consideració el principi de *Human Rights by Design (HRbD)*.
- c) **Adoptar una estratègia i actuació responsables i diligents:** s'advoca per fomentar i garantir l'estratègia i acció responsables i diligents per part de totes les parts interessades i actors digitals, públics i privats, durant tot el cicle de vida dels sistemes d'IA.
- d) **Respectar, protegir i promoure la dignitat, els drets humans i les llibertats fonamentals:** el respecte, la protecció i la promoció de la dignitat humana i dels drets i de les llibertats fonamentals reconeguts internacionalment⁶⁰ s'han de considerar al llarg del cicle de vida dels sistemes d'IA, amb una atenció especial als drets dels col·lectius o persones especialment vulnerables. La dignitat humana fa referència al reconeixement del valor intrínsec i igual de cada ésser humà, al marge de qualsevol factor personal.⁶¹ De la mateixa manera, els sistemes d'IA han de poder assistir i ajudar l'ésser humà en tot moment, empoderar-lo i amplificar les seves capacitats, i respectar la seva autonomia de manera que les persones que interactuïn amb sistemes d'IA puguin mantenir sempre una autonomia plena i efectiva sobre si mateixes.
- e) **Garantir la diversitat i la inclusió:** cal assegurar la protecció i la promoció de la diversitat i la inclusió al llarg del cicle de vida dels sistemes d'IA, de conformitat amb el dret aplicable en l'àmbit internacional i nacional. Per a aquest fi, s'hauria de promoure la participació activa de totes les persones o grups, amb independència de qualsevol condició personal, i atendre de manera especial la inclusió de persones o grups especialment vulnerables, com són les persones amb discapacitat.
- f) **Assegurar la conscienciació, el dret a l'educació i l'impuls de les capacitats digitals:**⁶² tot ésser humà té dret a la informació i a l'educació, en particular quan és crucial per assegurar el ple desenvolupament de la seva dignitat i personalitat, o quan sigui important per reforçar el respecte dels seus drets i llibertats fonamentals. Per això, les parts interessades destinatàries d'aquest Codi

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

han de fomentar aquests drets⁶³ al màxim possible. És important prestar una atenció especial a l'alfabetització digital en IA dels infants i joves, i a altres col·lectius vulnerables.

- g) Evitar qualsevol tipus de dany, manipulació o sotmetiment, ja sigui de caràcter físic, mental, econòmic, social, polític o cultural, a un ésser humà:** tot individu té dret a la vida, a la llibertat, incloent-hi la llibertat de pensament, consciència i d'opinió i expressió, i a la seguretat de la seva persona.⁶⁴ Cap ésser humà ni cap comunitat humana no han de patir aquest tipus de danys, manipulació o sotmetiment al llarg del cicle de vida d'un sistema d'IA. Al contrari, els sistemes d'IA han de poder protegir i millorar la qualitat de vida dels éssers humans, la qual cosa implica dissenyar, desenvolupar i desplegar aquesta tecnologia de forma ètica i responsable des de l'inici, prevenint, evitant i controlant, en tot moment, possibles riscos associats als drets fonamentals; per exemple, a través de la realització prèvia de les anàlisis i avaluacions d'impacte pertinents respecte a aquests drets abans de desenvolupar, desplegar o utilitzar sistemes d'IA.
- h) Facilitar i potenciar a través de la IA nous mitjans i eines per poder defensar i exercir els drets humans, i no per vulnerar-los:** és possible fer servir la IA per promoure la dignitat i la igualtat, entre altres drets i llibertats fonamentals, utilitzant sistemes d'IA per millorar el grau de compliment d'aquests drets i la seva observança en diferents nivells.
- i) Promoure l'aplicació i l'ús pacífic i just de la IA:** els sistemes d'IA han de propiciar societats pacífiques i justes, en benefici de tothom, compatibles amb els drets humans i les llibertats fonamentals i, en general, el benestar social i el bé comú. Els sistemes d'IA no han de dividir i enfrontar entre si les persones ni les comunitats o els grups en què s'integren, i han d'evitar d'amenaçar la coexistència entre els éssers humans i la seva interconnexió amb el medi natural i la resta dels éssers vius. Per contra, han de poder fomentar l'harmonia i la pau mundial.
- j) Advocar per la sostenibilitat i la protecció del medi ambient:** s'ha de promoure la sostenibilitat i protegir el medi ambient i els ecosistemes al llarg del cicle de vida dels sistemes d'IA i fer, per exemple, les corresponents avaluacions d'impacte ambiental amb aquest objectiu. Les parts interessades que participin en el cicle de vida dels sistemes d'IA han de respectar el dret internacional i la normativa nacional aplicable, així com les bones pràctiques reconegudes a escala mundial, com són els 17 Objectius de desenvolupament sostenible (ODS) de l'ONU⁶⁵ o els criteris o indicadors ESG del Fòrum de Davos.⁶⁶

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

V.2. Principis

La transformació digital i la tecnologia emergent com la intel·ligència artificial ofereixen grans oportunitats per a la societat, les persones i les empreses; però, perquè siguin coherents amb certs valors assentats a Andorra, conforme es recull en la seva legislació actual, han de basar-se també en principis ètics sòlids per a la millor protecció de les persones i els seus drets bàsics.

La IA ha de promoure una millor qualitat de vida, el creixement econòmic i la sostenibilitat. En definitiva, la IA ha de respectar i promoure els drets humans i les llibertats fonamentals, així com els objectius socials, econòmics i ambientals d'Andorra.

Per això, la IA ha de ser, en tot moment, **justa, transparent, explicable, responsable i no discriminatòria, i ha de garantir la dignitat, l'autonomia i aquests drets i llibertats fonamentals, així com la seguretat de les persones. La IA també ha de ser accessible, inclusiva i diversa, i ha d'empoderar les persones i la societat civil.**

És per això que a Andorra es promouen els principis següents en relació amb el foment de la IA confiable:

- a) Desenvolupament inclusiu i sostenible en favor del benestar de les persones i la societat:** durant tot el cicle de vida dels sistemes d'IA, aquests sistemes han de ser capaços d'assistir i ajudar els éssers humans empoderant-los i afavorint, en tot moment, el desenvolupament inclusiu i sostenible en favor de les persones i la societat. Així mateix, com indica la UNESCO, l'avaluació continuada dels efectes humans, socials, culturals, econòmics i ambientals de les tecnologies de la IA s'hauria de dur a terme amb ple coneixement de les repercussions d'aquestes tecnologies en la sostenibilitat com a conjunt de metes en evolució constant en tota una sèrie de dimensions, com les que es defineixen actualment en els Objectius de desenvolupament sostenible (ODS) de les Nacions Unides.
- b) IA centrada en l'ésser humà i l'equitat:** cal respectar l'estat de dret, els drets humans i els valors democràtics durant tot el cicle de vida del sistema d'IA, situar les persones al centre en tot moment, i promoure un ús just i adequat d'aquests sistemes. Els destinataris d'aquest Codi, en la mesura que sigui aplicable, haurien de fer tot el que fos raonablement possible per reduir al mínim resultats discriminatoris o esbiaixats (biaix algorítmic), i per evitar reforçar-los o perpetuar-los al llarg del cicle de vida dels sistemes d'IA, per tal de garantir l'equitat dels sistemes esmentats. També s'haurà de prestar una atenció particular a les situacions en què els sistemes d'IA puguin provocar efectes adversos (o puguin agreujar els ja existents) a causa d'asimetries de poder o d'informació en certes relacions o situacions.
- c) Principi de disseny legal i ètic des de l'inici d'iniciatives, projectes, serveis i productes d'IA (HRbD):** aquest Codi promou l'aplicació i el desenvolupament d'aquest principi, reconegut internacionalment pel Consell d'Europa, entre altres organitzacions internacionals i estats. Aquest principi pretén garantir des del mateix disseny i projecció de projectes, serveis i productes basats en IA que es consideri la protecció plena dels drets humans des d'aquest moment.

d) Proporcionalitat i innocuïtat de la IA: la IA ha de ser proporcional i no ha d'anar més enllà del que sigui necessari per assolir els fins o objectius legítims que persegueix, i ha de ser adequada al context. La IA ha d'evitar o prevenir qualsevol dany per a les persones, els seus drets, la seguretat o el medi ambient, i ha d'aplicar eines d'avaluació de riscos i les mesures de salvaguarda pertinents.

Per tant, la decisió d'utilitzar sistemes d'IA ha de poder acompanyar-se de l'anàlisi documentada corresponent de necessitat i proporcionalitat, i ha d'acreditar que el sistema d'IA en qüestió és una tècnica o un mètode adequat i proporcional per assolir un o diversos objectius legítims prèviament identificats i definits, i sense que això suposi vulnerar drets o llibertats fonamentals de les persones. La necessitat és una condició prèvia a la proporcionalitat i l'esmentada anàlisi de necessitat i proporcionalitat s'ha de poder mantenir en el temps durant tot el cicle de vida del sistema d'IA. Una avaluació de la proporcionalitat d'un sistema d'IA implica avaluar quines mesures i salvaguardes resulten igualment necessàries en aquests casos.

e) Robustesa, seguretat i prevenció del dany: la IA ha de ser fiable i robusta, i ha de complir els estàndards de qualitat, seguretat i eficàcia durant tot el cicle de vida del sistema d'IA de què es tracti. Així mateix, la IA ha d'estar subjecta a una supervisió humana adequada i permetre la intervenció i la rectificació humanes quan sigui procedent. Així mateix, la IA ha de prevenir i mitigar els riscos de danys o errors, i ha de ser resistent davant les amenaces cibernètiques i els atacs maliciosos.

f) Transparència i explicabilitat: més transparència contribueix a crear societats més pacífiques, justes, democràtiques i inclusives, per la qual cosa cal fomentar una comprensió general dels sistemes d'IA, i conscienciar i informar les parts interessades sobre les seves interaccions amb els sistemes d'IA, per permetre que els afectats per un sistema d'IA entenguin que estan interactuant amb aquesta mena de sistemes i el funcionament de la lògica aplicada en aquests sistemes, i, en definitiva, perquè puguin comprendre'n el resultat i puguin qüestionar, d'acord amb la llei aplicable, certes decisions automatitzades que els afectin o que els puguin afectar de manera significativa o negativa i/o oposar-s'hi. Així mateix, la transparència i l'explicabilitat estan estretament relacionades amb la responsabilitat i la rendició de comptes, així com amb la confiabilitat dels sistemes d'IA.

g) Traçabilitat i gestió dels riscos de la IA: s'ha de garantir la traçabilitat i la gestió dels riscos dels sistemes d'IA, en relació amb els conjunts de dades i la qualitat d'aquests conjunts, els processos i les decisions preses durant tot el cicle de vida del sistema d'IA, per prevenir danys i per permetre l'anàlisi dels resultats i les decisions adoptades a cada moment, perquè es puguin avaluar i abordar els possibles impactes negatius de la IA en les persones, la societat i el medi ambient.

h) Dret a la intimitat i protecció de dades: per les seves implicacions especials, cal prestar atenció als riscos associats a la intimitat i a la protecció de les dades personals,⁶⁷ ja que s'erigeixen en drets essencials per protegir la dignitat, l'autonomia i la capacitat d'actuar dels éssers humans. És important que les dades personals tractades pels sistemes d'IA es recopilin, s'utilitzin, es comparteixin, s'arxivin i se suprimeixin de manera coherent amb el dret aplicable i, si escau, les directrius i recomanacions que publiquin les autoritats de control competents en la interpretació d'aquest dret. En molts casos pot resultar obligatori o recomanable fer avaluacions d'impacte en matèria de protecció de dades en relació amb els sistemes d'IA que pretenguin desenvolupar-se o desplegar-se, segons correspongui.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

- i) **Supervisió i decisió humanes:** això implica garantir la supervisió i el control humans sobre els processos de treball dels sistemes d'IA. Un sistema d'IA mai no podrà reemplaçar la responsabilitat final dels éssers humans i la seva obligació de retre comptes, per la qual cosa s'haurà de vetllar sempre perquè sigui possible atribuir la responsabilitat ètica o legal associada a un sistema d'IA a un ésser humà, en qualsevol etapa del cicle de vida dels sistemes d'IA.
- j) **Responsabilitat:** s'ha de garantir el funcionament correcte dels sistemes d'IA i el respecte dels principis que conté aquest Codi, en funció del context de què es tracti i d'acord amb l'estat de l'art en cada moment. Per això, les parts interessades i els agents destinataris d'aquest Codi han de poder demostrar la seva responsabilitat mitjançant accions i processos transparents i auditables.

En resum: el Codi promou principis importants reconeguts a escala internacional en relació amb la IA confiable, els quals poden resumir-se de la manera següent:

- Desenvolupament inclusiu i sostenible en favor del benestar de les persones i la societat.
- IA centrada en l'ésser humà i l'equitat.
- Principi de disseny legal i ètic des de l'inici d'iniciatives, projectes, serveis i productes d'IA (HRbD).
- Proporcionalitat i innocuïtat de la IA.
- Robustesa, seguretat i prevenció del dany.
- Transparència i explicabilitat.
- Traçabilitat i gestió dels riscos de la IA.
- Dret a la intimitat i protecció de dades.
- Supervisió i decisió humanes.
- Responsabilitat.

VI. Marc per a una intel·ligència artificial confiable

VI.1. Fonaments i requisits per a una IA confiable

VI.2. Mètodes i tècniques per al foment de la IA confiable

VI.3. “Human Rights By Design (HRbD)” a l’era de la IA

Aquest Codi no imposa cap marc de referència per al foment de la IA confiable, però sí que destaca els principis existents en aquest moment a escala internacional al llarg del seu contingut i en les referències finals del mateix document, com a orientació i guia per a les parts interessades en aquest àmbit. En exercici de les actuacions d'impuls i promoció del Codi, el Govern d'Andorra podrà promoure, juntament amb les parts interessades, marcs de referència per al control de riscos de la IA i per al foment de la confiança en aquest àmbit considerant, si escau, els marcs de referència internacional existents.

Les parts interessades podran decidir lliurement els marcs que s'ajusten millor a la seva idiosincràcia i les seves necessitats per fomentar una IA confiable, i podran fer ús dels marcs de referència internacionals esmentats en aquest Codi. Es fa un esment especial de l'estàndard internacional ISO/IEC 23053:2022, Marc per a sistemes d'intel·ligència artificial (IA) que utilitzen l'aprenentatge automàtic (ML),⁶⁸ que pot ser útil per a les parts interessades com a marc comparat de referència per al millor control dels riscos a IA.

Sense perjudici del que s'ha esmentat anteriorment, en aquest apartat es fan una sèrie de recomanacions generals al voltant dels fonaments, els requisits, els mètodes i les tècniques recollits en els documents i les posicions internacionals en aquest sentit.

VI.1. Fonaments i requisits per a una IA confiable

La confiança és un factor clau per garantir que la IA es pugui desenvolupar de manera coherent.

Per generar confiança, i com a **fonaments per a una IA confiable, els sistemes d'IA han de respectar certs valors i principis ètics, i han de complir el dret aplicable.**

A més, **els sistemes d'IA han de ser tècnicament robusts i fiables**, de manera que puguin funcionar de forma segura i previsible sota una sèrie de condicions i escenaris. Així mateix, els sistemes d'IA han de ser **transparents i explicables** (sense que això suposi vulnerar altres possibles drets o interessos legítims, com la llibertat d'empreses o el secret comercial, etc.), de manera que es pugui entendre el seu funcionament, les seves capacitats i les seves limitacions, així com els possibles riscos i impactes associats al seu ús. Finalment, **els sistemes d'IA han de ser fiables i auditables**, de manera que es puguin identificar i atribuir les responsabilitats de totes les parts interessades implicades en el seu cicle de vida i es puguin prendre mesures correctives o compensatòries, si fos necessari.

Aquests **fonaments per a una IA confiable** es basen en **el reconeixement de la dignitat humana, l'autonomia, la igualtat, la justícia, la solidaritat, la democràcia, l'estat de dret i el respecte dels drets humans.**⁶⁹

Així mateix, els **requisits per a una IA fiable es poden agrupar en set categories**, que reflecteixen i inclouen els valors i principis ètics esmentats en aquest Codi. Aquests requisits són els següents:

- **Acció i supervisió humanes:** els sistemes d'IA han de permetre l'acció i la supervisió humanes adequades, de manera que siguin respectuosos amb la dignitat, l'autonomia i l'autocontrol humans. Això implica que els sistemes d'IA han de ser dissenyats, desenvolupats i desplegats de manera que facilitin la interacció humana, la participació, la intervenció i la protecció en tot moment.
- **Solidesa tècnica i seguretat:** els sistemes d'IA han de ser tècnicament robusts i fiables, de manera que puguin funcionar de forma segura i previsible sota una sèrie de condicions i escenaris. Això implica que els sistemes d'IA han de ser resistents als atacs, errors i anomalies, i que han

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

d'incorporar mecanismes de prevenció, detecció i correcció d'errors, així com de recuperació i continuïtat en cas d'emergència.

- **Gestió de la privacitat i de les dades:**⁷⁰ els sistemes d'IA han de garantir la privacitat i la protecció de les dades de les persones i organitzacions involucrades o afectades per aquests sistemes, d'acord amb el dret aplicable. Això implica que els sistemes d'IA han d'assegurar els principis essencials del tractament, la qualitat, la integritat, la legalitat i la traçabilitat de les dades utilitzades. Han d'evitar la identificació indeguda o no autoritzada de les persones físiques, minimitzar l'ús de dades i garantir la concurrència de bases legals suficients per al tractament de la informació personal utilitzada durant tot el seu cicle de vida.⁷¹
- **Transparència:** els sistemes d'IA han de ser transparents i explicables, de manera que es pugui entendre el seu funcionament, les seves capacitats i les seves limitacions, així com els possibles riscos i impactes associats al seu ús. Això implica que els sistemes d'IA han de proporcionar informació clara i accessible sobre els seus objectius, mètodes, components, fonts de dades, resultats, incerteses i biaixos, i que han de facilitar mecanismes d'explicació i comprensió de les seves decisions i accions.
- **Diversitat, no-discriminació i equitat:** els sistemes d'IA han de respectar la diversitat, la no-discriminació i l'equitat, de manera que no perjudiquin ni exclouin cap grup o individu per raons injustes o arbitràries, i que promoguin la inclusió i la participació de tothom. Això implica que els sistemes d'IA han d'evitar i corregir qualsevol biaix, prejudici o estereotip que pugui afectar negativament la igualtat d'oportunitats, el tracte i els resultats, i que han de fomentar la pluralitat, la tolerància i el respecte mutu.
- **Benestar social i ambiental:** els sistemes d'IA han de contribuir al benestar ambiental i social, de manera que millorin la qualitat de vida, el desenvolupament sostenible i el respecte pels ecosistemes i els recursos naturals. Això implica que els sistemes d'IA han de minimitzar el seu impacte ambiental, tant en termes de consum d'energia i de matèries primeres com de generació de residus i d'emissions, i que han de promoure l'ús eficient i responsable dels recursos, la prevenció i la mitigació dels riscos ambientals i la protecció i la restauració de la biodiversitat.
- **Rendició de comptes:** els sistemes d'IA han de ser responsables i auditables, de manera que es puguin identificar i atribuir les responsabilitats de totes les parts interessades implicades en el seu cicle de vida, i es puguin prendre mesures correctives o compensatòries si cal. Això implica que els sistemes d'IA han d'estar subjectes a una supervisió adequada i a un control de qualitat, i que han d'establir mecanismes de reclamació, recurs i reparació per a les persones o les organitzacions que es vegin afectades per les seves decisions o accions.

Aquests requisits no són exhaustius ni immutables, sinó que poden ser adaptats o ampliatos en funció del context, el propòsit i l'impacte de cada sistema d'IA. Així mateix, aquests requisits no són independents ni jeràrquics, sinó que estan interrelacionats i poden entrar en conflicte en situacions determinades. Per tant, cal fer una anàlisi i una avaluació acurades de cada cas concret, tenint en compte els principis ètics, els valors, els drets i els interessos en joc, així com les possibles conseqüències i alternatives disponibles.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

En tot cas, per garantir el compliment d'aquests requisits, és recomanable emprar tant mètodes tècnics com no tècnics que ajudin, conforme a l'estat de l'art i la tècnica en cada moment, a poder avaluar els sistemes d'IA i promoure el compliment en el temps i de manera continuada d'aquests requisits.

En resum: amb el suport dels valors i principis reconeguts en aquest Codi, **és important partir de certs fonaments i requisits per desenvolupar marcs de confiança entorn de la IA. Com a fonaments per a una IA fiable**, els sistemes d'IA han de respectar determinats valors i principis ètics, i han de complir el dret aplicable. També han de ser transparents i explicables, així com responsables i auditables. Per part seva, els principals requisits són els relatius a l'acció i la supervisió humanes; la solidesa tècnica i la seguretat; la gestió de la privacitat i de les dades; la transparència; la diversitat, la no-discriminació i l'equitat; el benestar social i ambiental, i la rendició

VI.2. Mètodes i tècniques per al foment de la IA fiable

Per poder garantir el compliment dels requisits anteriors, és important comptar amb mètodes tècnics i no tècnics apropiats per al foment de la IA fiable.

Entre els mètodes tècnics que poden contribuir a crear sistemes d'IA fiables, hi ha els següents:

- L'ús de dades de qualitat, representatives, diverses i rellevants, que reflecteixin les característiques i necessitats dels usuaris finals i que evitin biaixos o discriminacions indegudes.
- L'aplicació de tècniques d'enginyeria de programari, com ara la revisió de codi, les proves automatitzades, la depuració, el seguiment o l'auditoria, que permetin detectar i corregir errors, vulnerabilitats o anomalies en el funcionament del sistema d'IA de què es tracti.
- La incorporació de mecanismes d'explicació, interpretació o justificació que facilitin la comprensió del funcionament, els resultats o els efectes del sistema d'IA, tant en l'àmbit tècnic com humà (explicabilitat).
- La implementació de mesures de privacitat i protecció de dades, incloent-hi les relacionades amb la seguretat de la informació personal des del disseny del sistema d'IA, de manera que es garanteixin la confidencialitat, la integritat i la disponibilitat de la informació tractada pel sistema d'IA, així com la resta dels principis i obligacions legals aplicables, segons la normativa de protecció de dades personals d'Andorra.⁷²
- La integració de criteris d'accessibilitat, usabilitat i disseny inclusiu, que assegurin que el sistema d'IA sigui fàcil d'utilitzar, que s'adapti a les capacitats i necessitats dels usuaris i que no exclouï o discrimini ningú.
- L'adopció d'estàndards tècnics reconeguts,⁷³ normes o bones pràctiques que orientin el desenvolupament, el desplegament o l'ús del sistema d'IA conforme a criteris de qualitat, eficàcia, eficiència i responsabilitat.

Entre els mètodes no tècnics que poden donar suport al desenvolupament, al desplegament o a l'ús de sistemes d'IA fiables, hi ha els següents:

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

- La realització d'avaluacions d'impacte que analitzin els possibles riscos, els beneficis i les conseqüències de l'aplicació del sistema d'IA per als drets, valors i interessos de les persones i la societat.
- La definició d'objectius, requisits i criteris d'èxit que estableixin les metes, les expectatives i les condicions que ha de complir el sistema d'IA per satisfer les necessitats i resoldre els problemes dels usuaris.
- L'assignació de rols i responsabilitats que determini les funcions, les obligacions i les atribucions dels diferents actors implicats en el cicle de vida del sistema d'IA, així com els mecanismes de rendició de comptes i reparació en cas de danys o perjudicis.
- La participació i implicació de les parts interessades que fomenti la consulta i la col·laboració entre els desenvolupadors, proveïdors, usuaris i afectats pel sistema d'IA, així com la representació de la diversitat i pluralitat de la societat.
- La formació i capacitació que promogui el desenvolupament de competències i habilitats relacionades amb la IA, tant en l'àmbit tècnic com ètic, social i humà, entre els professionals, els usuaris i la ciutadania en general.
- La comunicació i divulgació que faciliti l'intercanvi d'informació, el coneixement i les experiències sobre el sistema d'IA, els seus objectius, el funcionament, els resultats i els efectes, de manera transparent, clara i accessible.
- La realització d'auditories i proves d'acompliment juntament amb mètodes de verificació en el temps capaços de detectar qualsevol anomalia o situació que calgui corregir en relació amb el sistema d'IA en qualsevol fase del cicle de vida.

En definitiva, cal apostar per l'avaluació de la fiabilitat de la IA en desenvolupar, desplegar o utilitzar sistemes d'IA i durant tot el cicle de vida. Es tracta d'aplicar un enfocament de risc i identificar i aplicar sobre la base d'aquest enfocament les mesures o els requisits adequats a cada situació i cas, considerant i implicant sempre les parts interessades en el procés.

En resum: per poder garantir el compliment dels requisits anteriors, és important comptar amb mètodes tècnics i no tècnics apropiats per fomentar una IA fiable.

No existeix un catàleg tancat d'aquests mètodes o mecanismes, i es poden aplicar els més apropiats a aquest fi. No obstant això, aquest Codi ofereix un catàleg no exhaustiu dels principals mètodes, tècnics i no tècnics, que es recomana implementar.

En funció de l'evolució de la tècnica i de l'estat de l'art, poden sorgir nous estàndards, bones pràctiques o normes que fomentin i promoguin una IA fiable, per la qual cosa s'aconsella explorar, actualitzar i revisar periòdicament els mètodes i mecanismes que s'apliquin amb aquest propòsit, a fi d'adoptar els que més convinguin segons el sistema d'IA de què es tracti.

VI.3. “Human Rights By Design (HRbD)” a l’era de la IA

Aquest Codi parteix d’un enfocament de l’ètica en la IA basat en els drets fonamentals consagrats tant en la Constitució i legislació d’Andorra⁷⁴ com en el dret internacional de drets humans que resulta aplicable. En aquest sentit, la legislació internacional de drets humans i, en particular, el Conveni europeu de drets humans són de compliment obligat a Andorra.⁷⁵

Conforme a aquest fet, cal considerar que el respecte dels drets fonamentals, en un marc de democràcia i estat de dret com el que es promou a Andorra, proporciona la base més sòlida per identificar els valors i principis ètics que es poden posar en pràctica en el context de la IA. És per això que la identificació d’aquests valors i principis ètics que conté aquest Codi parteix d’aquesta base i hi respon.

En aquest sentit, els drets fonamentals formen part, doncs, del primer component de la IA fiable (la IA lícita), que garanteix el compliment de la legislació aplicable a Andorra.

Per consegüent, com a principi general, cal promoure el principi del disseny legal i ètic de qualsevol tipus d’iniciativa, projecte, producte o servei en l’entorn digital, en particular quan es basen en tecnologies emergents com la intel·ligència artificial.

Aquest principi general es refereix al principi anomenat “Human Rights by Design (HRbD)”,⁷⁶ que reconeixen organitzacions internacionals com ara el Consell d’Europa.

Sobre la base del possible impacte negatiu dels sistemes d’IA en els éssers humans, resulta altament recomanable fer una avaluació d’impacte dels drets fonamentals (EIDH)⁷⁷ en relació amb els sistemes d’IA que es pretenguin desenvolupar, desplegar o aplicar per part de les parts interessades i els destinataris d’aquest Codi. En la realització d’aquesta EIDH cal considerar, en tot cas, el possible impacte que es pot produir durant tot el cicle de vida del sistema d’IA en qüestió.

En resum: per poder atendre de forma coherent els drets i les llibertats fonamentals, cal partir del principi de “Human Rights by Design (HRbD)”, de manera que, com a principi general, s’ha de promoure el principi del disseny legal i ètic de qualsevol tipus d’iniciativa, projecte, producte o servei en l’entorn digital, en particular quan es basen en tecnologies emergents com la intel·ligència artificial.

Sobre la base del possible impacte negatiu dels sistemes d’IA en els éssers humans, resulta altament recomanable fer una avaluació d’impacte dels drets fonamentals (EIDH) en relació amb els sistemes d’IA que es pretenguin desenvolupar, desplegar o aplicar per part de les parts interessades i els destinataris d’aquest Codi.

1-83: Vegeu l’annex I d’aquest document, relatiu a les referències (bibliografia i webgrafia).

VII. Intel·ligència artificial i criteris ESG

La intel·ligència artificial (IA) té el potencial de contribuir a la sostenibilitat ambiental, social i de governabilitat (ESG, “Environmental , Social and Governance”)⁷⁸ de les organitzacions, en oferir solucions innovadores per als desafiaments globals com el canvi climàtic, la pobresa o la desigualtat.

Tot i això, la IA també planteja riscos i dilemes ètics, que han de ser abordats amb responsabilitat i transparència. Per exemple, l'ús intensiu de dades i recursos informàtics pot generar un alt consum d'energia i aigua, així com una petjada de carboni més gran. A més, la IA pot afectar els drets humans, la privacitat, la seguretat, la inclusió i la no-discriminació de les persones, en funció de com es dissenyi, desenvolupi i apliqui.

Per això, cal adoptar un enfocament ètic i basat en els drets humans per al desenvolupament i l'ús de la IA, que tingui en compte els criteris ESG i els principis de confiança, benefici social, justícia i rendició de comptes. Així ho recomanen organismes internacionals com ara l'OCDE, el Consell d'Europa o la Unió Europea, que han elaborat marcs normatius i orientacions pràctiques per fomentar una IA centrada en les persones i alineada amb els valors democràtics.

Aquestes iniciatives busquen garantir que la IA respecti els drets humans, promogui el benestar social i protegeixi el medi ambient, alhora que impulsa la innovació i el creixement econòmic.

Entre els beneficis de la IA per al compliment d'indicadors i mètriques d'ESG, es poden destacar els següents:

- La IA pot ajudar a optimitzar l'ús dels recursos naturals i reduir l'impacte ambiental de les activitats humanes mitjançant l'anàlisi de dades, la predicció d'escenaris, l'automatització de processos i la generació de solucions eficients. Per exemple, la IA pot contribuir a millorar la gestió de l'aigua, l'agricultura de precisió, la producció d'energies renovables, el transport intel·ligent o l'economia circular.
- La IA pot donar suport al desenvolupament social i la inclusió, ja que facilita l'accés a l'educació, la salut, la cultura, la informació i els serveis públics, especialment per a les poblacions vulnerables o marginades. La IA també pot potenciar la participació ciutadana, la transparència i la rendició de comptes, en afavorir la comunicació, la col·laboració i la cocreació entre els actors socials.
- La IA pot impulsar la innovació i la competitivitat, en generar nous productes, serveis, models de negoci i oportunitats d'ocupació que responguin a les necessitats i les expectatives de la societat. La IA també pot millorar la qualitat, la productivitat, l'eficàcia i l'eficiència de les organitzacions, proporcionant eines de suport per a la presa de decisions, l'aprenentatge i la millora contínua.

En resum: la intel·ligència artificial (IA) té el potencial de contribuir a la sostenibilitat ambiental, social i de governabilitat (ESG) de les organitzacions, en oferir solucions innovadores per als desafiaments globals com el canvi climàtic, la pobresa o la desigualtat.

No obstant això, la IA també planteja riscos i dilemes que afecten diversos criteris i indicadors ESG, la qual cosa s'ha de preveure en els projectes i les iniciatives basats en intel·ligència artificial.

Per exemple, l'ús intensiu de dades i recursos informàtics pot generar un alt consum d'energia i aigua, així com una petjada de carboni més gran. A més, la IA pot afectar els drets humans, la privacitat, la seguretat, la inclusió i la no-discriminació de les persones, en funció de com es dissenyi, desenvolupi i apliqui, i pot vulnerar indicadors i mètriques ESG.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

VIII. | Drets digitals i neurodrets

Cada cop és més important protegir i garantir de forma plena els drets i les llibertats fonamentals de les persones a l'entorn digital, i a Andorra l'impuls i la generació d'un marc de protecció integral renovat dels drets digitals és una de les prioritats principals, tal com s'esdevé en aquests moments en altres països i regions del món.

Dins dels drets digitals, cal prestar atenció als drets digitals en l'àmbit de la intel·ligència artificial, amb un focus especial en la protecció del dret a la privacitat mental, entre altres drets que tenen una rellevància especial en aquest àmbit. També és important considerar la protecció dels neurodrets.

Els neurodrets són els drets humans relacionats amb el cervell i la ment, que tenen com a objectiu protegir la integritat, la llibertat, la privacitat i la identitat de les persones davant de possibles amenaces derivades de l'ús de les tecnologies neurocientífiques i d'intel·ligència artificial. Aquests drets es basen en el reconeixement que el cervell és la font de la consciència, la personalitat, el pensament, la memòria, l'emoció i el comportament humà, i que qualsevol alteració o manipulació d'aquestes funcions pot afectar la dignitat i l'autonomia de les persones, així com els seus drets fonamentals.

La UNESCO, que és l'organització de les Nacions Unides encarregada de promoure l'educació, la ciència, la cultura i la comunicació per al desenvolupament humà sostenible, ha mostrat una gran preocupació per les implicacions ètiques, socials i legals de les tecnologies neurocientífiques i d'intel·ligència artificial⁷⁹ que poden tenir un impacte en el cervell humà i les seves capacitats cognitives, afectives i socials.

Per això, la UNESCO ha iniciat un procés de reflexió i diàleg amb experts, representants de la societat civil, governs i organitzacions internacionals, per elaborar una recomanació sobre l'ètica de la intel·ligència artificial que inclogui els aspectes relacionats amb els neurodrets.

Aquesta recomanació serà el primer instrument normatiu internacional que abordarà els principis i els valors ètics que han d'orientar el desenvolupament i l'ús de la intel·ligència artificial, i que servirà de referència per a la cooperació i l'acció global en aquest àmbit.

Les parts interessades i els destinataris d'aquest Codi atendran de manera especial la dimensió digital dels drets i les llibertats fonamentals (drets digitals i neurodrets), en particular pel que fa al desenvolupament, l'ús i l'aplicació de sistemes d'intel·ligència artificial.

En resum: la intel·ligència artificial (IA) té el potencial de contribuir a la sostenibilitat ambiental, social i de governabilitat (ESG) de les organitzacions, en oferir solucions innovadores per als desafiaments globals com el canvi climàtic, la pobresa o la desigualtat.

No obstant això, la IA també planteja riscos i dilemes que afecten diversos criteris i indicadors ESG, la qual cosa s'ha de preveure en els projectes i les iniciatives basats en intel·ligència artificial.

Per exemple, l'ús intensiu de dades i recursos informàtics pot generar un alt consum d'energia i aigua, així com una petjada de carboni més gran. A més, la IA pot afectar els drets humans, la privacitat, la seguretat, la inclusió i la no-discriminació de les persones, en funció de com es dissenyi, desenvolupi i apliqui, i pot vulnerar indicadors i mètriques ESG.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

IX. IA Generativa (GenAI)⁸⁰

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

La IA generativa ofereix grans oportunitats per a la innovació, la creativitat, l'educació i el benestar social, però també planteja desafiaments importants ètics i legals. Entre altres desafiaments i riscos s'hi inclouen el risc de generar continguts falsos, enganyosos o nocius i la possible promoció de la desinformació, així com les implicacions per als drets d'autor, la privacitat, la seguretat i la governabilitat democràtica, entre d'altres.

Per tant, és essencial desenvolupar i aplicar principis i bones pràctiques per garantir que la IA generativa s'utilitzi de manera responsable i respectuosa amb els drets humans, així com amb altres principis i interessos legítims.

A continuació es presenten algunes orientacions, basades en els principis de l'OCDE sobre intel·ligència artificial, que poden ajudar els creadors i usuaris de la IA generativa a maximitzar-ne els beneficis i a minimitzar-ne els riscos:

- Promoure el disseny i el desenvolupament de sistemes d'IA generativa que siguin robustos, segurs, fiables, explicables i auditables, i que respectin els estàndards tècnics i ètics més alts.
- Garantir que els sistemes d'IA generativa se sotmetin a un control de qualitat i a una avaluació d'impacte, en particular sobre els drets fonamentals, abans de desplegar-los, i que se supervisin i actualitzin periòdicament per detectar i corregir qualsevol incidència o risc que calgui pal·liar.
- Implementar mesures adequades per protegir la privacitat, la seguretat i els drets de propietat intel·lectual de totes les parts interessades, garantint la traçabilitat de les dades i les fonts utilitzades en la generació de continguts a partir d'aquests models fundacionals i de tipus generatiu.
- Establir mecanismes de transparència, traçabilitat i rendició de comptes per a la creació i l'ús de continguts generats per la IA, i facilitar la verificació, l'atribució i la correcció de la informació quan sigui necessari.
- Respectar la dignitat, l'autonomia i els drets humans de les persones afectades pels continguts generats per la IA, evitant la difusió de continguts falsos, ofensius, discriminatoris o perjudicials.
- Fomentar l'ús ètic i beneficiós de la IA generativa per a les persones i la societat, donant suport a la diversitat, la inclusió i l'equitat i posant un focus especial en el respecte a la cultura, l'art, l'educació i les activitats científiques.

En resum: la IA generativa ofereix grans oportunitats per a la innovació, la creativitat, l'educació i el benestar social, però també planteja importants desafiaments ètics i legals. Entre altres desafiaments i riscos s'hi inclouen el risc de generar continguts falsos, enganyosos o nocius i la possible promoció de la desinformació, així com les implicacions per als drets d'autor, la privacitat, la seguretat i la governabilitat democràtica, entre altres.

Per tant, és essencial desenvolupar i aplicar principis i bones pràctiques per garantir que la IA generativa s'utilitzi de manera responsable i respectuosa amb els drets humans, així com amb altres principis i interessos legítims.

X. Àmbits tractors d'actuació

**X.1. Àmbit d'actuació 1: sector
públic**

**X.2. Àmbit d'actuació 2: sector
privat**

X.3. Àmbit d'actuació 3: ciutadania

X.1. Àmbit d'actuació 1: sector públic

La intel·ligència artificial (IA) planteja importants desafiaments i oportunitats per al sector públic,⁸¹ i s'han d'aprofitar les potencialitats per millorar les relacions amb la ciutadania i potenciar la qualitat i l'eficiència dels serveis públics, així com per promoure la transparència, la rendició de comptes i la participació ciutadana.

Alhora, el sector públic ha de garantir que el desenvolupament i l'ús de la IA es facin d'acord amb els principis ètics i els valors democràtics, i que respectin els drets i les llibertats fonamentals i protegeixin, en tot moment, l'interès general.

Amb aquesta finalitat, el sector públic ha d'adoptar un enfocament responsable de la IA, basat en la confiança, l'equitat i la sostenibilitat tant quan desenvolupa, aplica o desplega sistemes d'IA *ad intro* en l'exercici de les seves funcions públiques com quan participa en l'establiment de marcs legals i de governança *ad extra* per al desplegament, en general, de la IA confiable.

Sense cap dubte, la confiança és imprescindible atès que hi ha diverses aplicacions d'IA al sector públic que podrien entrar a la categoria d'alt risc; per exemple, els sistemes d'IA destinats a avaluar l'admissibilitat de les persones físiques per beneficiar-se de serveis o prestacions essencials d'assistència pública, inclosos els serveis d'assistència sanitària, així com per concedir, reduir o retirar aquests serveis i prestacions o reclamar-ne la devolució. Un altre exemple podrien ser els sistemes d'IA destinats a avaluar i classificar possibles trucades d'emergència fetes per persones físiques, entre altres sistemes d'IA en l'àmbit públic.

A continuació es fan algunes recomanacions per a les finalitats que pretén aconseguir aquest Codi dirigides al sector públic, de manera orientativa:

BONES PRÀCTIQUES I RECOMANACIONS PER AL SECTOR PÚBLIC:

Objectiu recomanat	Descripció
Establir marcs normatius i de governança de la IA al sector públic	Establir un marc normatiu i una governança adequats per garantir que les aplicacions de la IA al sector públic siguin transparents, justes, segures i responsables.
Alinear l'ús de la IA al sector públic amb els valors democràtics, els drets humans i el bé comú	Alinear l'ús de la IA al sector públic amb els valors democràtics, els drets humans i el bé comú, tenint en compte els possibles impactes socials, econòmics, ambientals i ètics de la IA.
Promoure la participació i la col·laboració dels diferents actors implicats	Promoure la participació i la col·laboració dels diferents actors implicats en el desenvolupament, la implementació i l'avaluació de la IA en el sector públic, incloent-hi, per exemple, els ciutadans, les empreses, les institucions educatives o acadèmiques i, en general, la societat civil.
Fomentar la formació i la capacitació dels empleats públics en matèria d'IA	Fomentar la formació i la capacitació dels empleats públics en matèria d'IA, tant en els

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

Objectiu recomanat	Descripció
(alfabetització en matèria d'IA dels empleats públics)	aspectes tècnics i legals com ètics, així com la creació de perfils professionals especialitzats en IA dins de l'Administració pública.
Afavorir la innovació i l'experimentació responsable amb la IA en el sector públic	Afavorir la innovació i l'experimentació responsable amb la IA en el sector públic, donant suport a projectes pilot, bones pràctiques i plataformes d'intercanvi de coneixement i experiències en aquest àmbit.
Impulsar l'ús de dades obertes, interoperables i de qualitat per alimentar els sistemes d'IA	Impulsar l'ús de dades obertes, interoperables i de qualitat per alimentar els sistemes d'IA, respectant sempre els principis de confidencialitat, privacitat i protecció de dades personals.
Vetllar per la rendició de comptes i la supervisió de les aplicacions de la IA en el sector públic	Vetllar per la rendició de comptes i la supervisió de les aplicacions de la IA en el sector públic, establint mecanismes d'identificació, inventari, classificació, control, auditoria, explicabilitat i revisió humana.
Establir marcs normatius i reguladors adequats i proporcionals per assegurar la protecció dels drets humans i els valors democràtics	Promoure l'elaboració i l'aplicació d'enfocaments i marcs regulatoris i de governança nacionals, d'acord amb les respectives polítiques i prioritats nacionals i amb les obligacions que incumbeixen Andorra, en virtut del dret internacional. Es donarà suport a la innovació i la inversió responsable i inclusiva en intel·ligència artificial per al desenvolupament sostenible, promovent sistemes segurs, protegits i fiables d'IA, i assegurant, en tot cas, la protecció dels drets humans i els valors democràtics existents, la responsabilitat i la rendició de comptes en l'àmbit de la IA.
Promoure sistemes d'intel·ligència artificial que fomentin, protegeixin i preservin la diversitat lingüística i cultural d'Andorra	Promoure sistemes d'intel·ligència artificial que fomentin, protegeixin i preservin la diversitat lingüística i cultural d'Andorra, tenint en compte el multilingüisme en les dades que s'utilitzen per entrenar-les i durant tot el seu cicle de vida.
Promoure la diversitat, la inclusió i l'equitat	Promoure la diversitat, la inclusió i l'equitat en el disseny, el desenvolupament i l'ús de la IA, evitar la discriminació i els biaixos i garantir l'accessibilitat i l'assequibilitat de les solucions d'IA.
Impulsar la capacitat i la formació continuada dels professionals, les empreses i altres parts interessades (població en general)	Sensibilitzar la ciutadania i promoure el coneixement envers les oportunitats, els riscos i, en general, els àmbits d'ús civil adequats dels sistemes d'intel·ligència artificial, per exemple a través de la formació i d'una capacitat. També, promoure la formació continuada sobre

Objectiu recomanat	Descripció
	IA dels professionals, les empreses i altres parts interessades.
Afavorir la creació i l'enfortiment d'ecosistemes d'IA fiables dinàmics i competitius	Afavorir la creació i l'enfortiment d'ecosistemes d'IA fiables dinàmics i competitius, que estimulin la inversió, l'emprenedoria i la col·laboració entre els sectors públic, privat, acadèmic i social.
Donar suport a la recerca i el desenvolupament de solucions d'IA que responguin als desafiaments socials, econòmics i mediambientals	Donar suport a la recerca i el desenvolupament de solucions d'IA que responguin als desafiaments socials, econòmics i mediambientals, i que contribueixin al benestar humà i al progrés democràtic.
Fomentar la innovació oberta i col·laborativa en l'àmbit de la IA	Fomentar la innovació oberta i col·laborativa en l'àmbit de la IA, facilitant l'accés i l'intercanvi de dades, coneixements, recursos i resultats entre els diferents actors públics i privats.
Implicar-se en el diàleg i la cooperació internacional per al desenvolupament i la governança de la IA fiable	Implicar-se en el diàleg i la cooperació internacional per al desenvolupament i la governança de la IA fiable, respectant els principis ètics i els objectius de desenvolupament sostenible, i fomentant la solidaritat i la cooperació entre països i regions.
En general, fomentar la investigació i la cooperació internacionals per comprendre i abordar els possibles beneficis i els riscos associats als sistemes d'intel·ligència artificial, i per poder assolir un equilibri	En general, fomentar la investigació i la cooperació internacionals per comprendre i abordar els possibles beneficis i els riscos associats als sistemes d'intel·ligència artificial, i per poder assolir un equilibri, i evitar així les possibles bretxes digitals i promoure els 17 Objectius de desenvolupament sostenible. ⁸²
Promoure processos de consulta i deliberació pública sobre la regulació i el govern de la IA	Promoure processos de consulta i deliberació pública sobre la regulació i el govern de la IA.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

X.2. Àmbit d'actuació 2: sector privat

La intel·ligència artificial (IA) ofereix oportunitats enormes per al sector privat, ja que permet la innovació, l'eficiència, la competitivitat i el creixement econòmic.

Al mateix temps, la IA planteja importants desafiaments i riscos per a les empreses, els consumidors, els treballadors i la societat en general, que han de ser abordats de manera responsable i ètica.

Per això, és fonamental que el sector empresarial adopti un enfocament basat en els valors i els drets humans per al desenvolupament i l'ús de la IA, i que col·labori amb els altres actors rellevants i parts interessades per garantir una governança efectiva i una rendició de comptes adequada de la IA.

Això implica posar el focus en l'establiment de models responsables quan es desenvolupa, utilitza o aplica la IA, per a la millor protecció dels drets i les llibertats fonamentals i la seguretat. Tot això, a més, per connexió amb l'ètica digital que promouen organitzacions com el Fòrum de Davos en relació amb el degut impuls dels criteris i indicadors ESG.

Sens dubte, cada vegada és més important apostar des del sector privat per la promoció de la innovació responsable, fent ús de la confiabilitat com un element diferenciador i competitiu de primer ordre al mercat.

Pel que fa a això, i d'acord amb la menció explícita continguda al document de "Principis rectors sobre les empreses i els drets humans" de l'ONU, els estats han d'enunciar clarament què s'espera de totes les empreses domiciliades al seu territori i/o jurisdicció de cara al respecte dels drets humans en totes les seves activitats, per preservar de forma plena aquests drets davant de possibles violacions. Aquest Codi respon a aquesta necessitat.

En definitiva, les empreses han d'evitar que les seves activitats provoquin o contribueixin a provocar conseqüències negatives sobre els drets humans, i han de respondre per aquestes conseqüències quan es produeixin.

Així mateix, han de tractar de prevenir o mitigar les conseqüències negatives sobre els drets humans directament relacionades amb operacions, productes o serveis que desenvolupin, o en què participin, aplicant criteris de disseny legal i ètic des de l'inici.

A continuació es fan algunes recomanacions per a les finalitats que pretén aconseguir aquest Codi dirigides al sector privat, de manera orientativa:

BONES PRÀCTIQUES I RECOMANACIONS PER AL SECTOR PRIVAT:⁸³

Objectiu recomanat	Descripció
Confiabilitat de la IA	Garantir que la IA sigui confiable, robusta i segura, adoptant les polítiques, els sistemes o els processos més adequats en cada moment per a la consecució d'aquests propòsits.
Respecte a valors i principis ètics	Promoure el desenvolupament i l'aplicació de sistemes d'IA que respectin, des del disseny, els principis ètics, els valors i els drets humans, així com les normes i els estàndards internacionals rellevants, mitjançant l'adopció de mesures preventives, correctives i de rendició de comptes durant tot el cicle de vida d'aquests sistemes.
Respecte als drets i les llibertats humanes	En particular, respectar la dignitat, l'autonomia, la igualtat, la no-discriminació, la privacitat i la seguretat de les persones afectades pels sistemes d'IA, entre altres drets i llibertats fonamentals, posant-les sempre al centre.
Per complir la seva responsabilitat de respectar els drets humans, les empreses han de comptar amb polítiques i procediments apropiats en funció de la seva dimensió i de les circumstàncies	Per complir la seva responsabilitat de respectar els drets humans, les empreses han de comptar amb polítiques i procediments apropiats en funció de la seva dimensió i de les circumstàncies. Això inclou el següent: 1) El compromís d'assumir la responsabilitat de respectar els drets humans; 2) Processos i funcionament de diligència deguda en matèria de drets humans per identificar, prevenir, mitigar l'impacte sobre els drets humans, i rendir comptes sobre com aborden aquest impacte; 3) Processos que permetin reparar totes les conseqüències negatives sobre els drets humans que hagin provocat o contribuït a provocar.
Identificació i avaluació d'impactes negatius associats a la IA	Identificar i avaluar els impactes negatius, reals o potencials, associats a projectes, iniciatives, productes i serveis de l'empresa en els drets i les llibertats fonamentals de les persones, la seguretat i la sostenibilitat.

1-83: Vegeu l'annex I d'aquest document, relatiu a les referències (bibliografia i webgrafia).

Objectiu recomanat	Descripció
Prevenició d'efectes negatius associats a la IA	Aturar, prevenir i mitigar els efectes negatius, reals o potencials, associats a projectes, iniciatives, productes i serveis de l'empresa en els drets i les llibertats fonamentals de les persones, la seguretat i la sostenibilitat.
Mecanismes de diligència, rendició de comptes i transparència	Preveure mecanismes de diligència deguda, rendició de comptes i transparència en el disseny, el desenvolupament i l'ús de la IA, tenint en compte els impactes i efectes negatius esmentats.
Mecanismes de reclamació i reparació	Establir mecanismes efectius i accessibles de reclamació, recurs i reparació per a les persones perjudicades o amenaçades pels sistemes de IA.
Establir sistemes de gestió i control de riscos de sistemes IA	Establir sistemes de gestió i control de riscos que garanteixin en el temps la seguretat, la fiabilitat i la robustesa dels sistemes d'IA.
Transparència, traçabilitat i explicabilitat dels sistemes d'IA	Garantir la transparència, la traçabilitat i l'explicabilitat dels sistemes d'IA, proporcionant informació clara i accessible sobre els seus objectius, capacitats, limitacions, funcionament i fonts de dades, així com sobre els rols i les responsabilitats dels actors involucrats.
Qualitat de les dades	Assegurar la qualitat, la integritat, la diversitat, la representativitat i la seguretat de les dades utilitzades per a l'entrenament, el desplegament i el funcionament dels sistemes d'IA, evitant biaixos, discriminació, estereotips i prejudicis injustificats o nocius per als drets de les persones.
Capacitació i alfabetització en IA	Fomentar la capacitat digital i l'alfabetització en matèria d'intel·ligència artificial, en particular entre els empleats i les persones assimilades, i entre òrgans directius i d'administració de l'empresa.
Promoció d'una cultura d'innovació responsable	En general, promoure una cultura d'innovació responsable i beneficiosa per a l'empresa i la societat, basada en el diàleg, la col·laboració i la cooperació entre els diferents actors i sectors interessats, incloses les autoritats competents, la societat civil, el món acadèmic i els usuaris finals.
Participar en els processos de consulta i deliberació pública sobre la regulació i el govern de la IA	Participar en els processos de consulta i deliberació pública sobre la regulació i el govern de la IA, i expressar l'opinió i les preocupacions. Donar suport a les iniciatives que promoguin una IA al servei de la innovació, la competitivitat empresarial, les persones i el planeta.

X.3. Àmbit d'actuació 3: ciutadania

La intel·ligència artificial (IA) té el potencial de millorar molts aspectes de les nostres vides, però també planteja desafiaments ètics i riscos per als drets fonamentals i, en relació amb aquests drets, per als drets digitals de les persones.

Per això, és important que els ciutadans coneguin els beneficis i els límits de la IA, així com els principis i les normes que la regeixen. També s'han d'adonar de les grans oportunitats que aquesta tecnologia ofereix i promoure, en tot cas, una informació i formació adequades entorn d'aquesta tecnologia.

Moltes de les recomanacions indicades a continuació atenen els paràmetres esmentats i resulten fonamentals per poder coadjuvar en la generació d'un ecosistema responsable i promoure, alhora, la innovació i el bé comú basat en la IA.

BONES PRÀCTIQUES I RECOMANACIONS ADREÇADES A LA CIUTADANIA:

Objectiu recomanat	Descripció
Conèixer els riscos i les oportunitats associats a la IA	Informar-se sobre què és la IA, com funciona, quines aplicacions té i quines implicacions pot tenir per a la societat i per als seus drets individuals. Buscar fonts fiables i diverses d'informació a aquest efecte, tenint en compte les organitzacions reconegudes.
Conrear una actitud crítica i reflexiva envers la IA	També cal conrear i adoptar una actitud crítica i reflexiva envers la IA, qüestionant-ne la fiabilitat, la precisió, la pertinència i el valor afegit, i posar en relleu això amb els avenços que pot comportar.
Educar i formar-se per a l'era de la IA	Els ciutadans haurien d'adquirir les competències i els coneixements necessaris per comprendre, utilitzar i aprofitar la IA de manera responsable i creativa. També haurien de desenvolupar habilitats transversals, com ara el pensament crític, la comunicació, la col·laboració i la resolució de problemes, que els permetin adaptar-se i prosperar en un món cada cop més digitalitzat i intel·ligent.
Mantenir una actitud diligent en la protecció dels seus drets (i d'aquells que representen)	Comprovar l'ús de la IA en els serveis i productes que es consumeixen o s'utilitzen, i preguntar com es recullen, processen i utilitzen les dades personals, així com quin tipus d'algorismes s'utilitzen i amb quina finalitat.

Objectiu recomanat	Descripció
Exigir que la IA sigui fiable	Exigir que la IA sigui fiable i que, per tant, es desenvolupi i es faci servir d'acord amb els principis de legalitat, necessitat, proporcionalitat, no-discriminació, equitat, transparència, explicabilitat i responsabilitat. Per tant, que els sistemes d'IA responguin a criteris legals i ètics adequats. Reclamar que es facin avaluacions d'impacte i auditories independents sobre els sistemes d'IA i que s'estableixin mecanismes de control i supervisió efectius.
Protecció i defensa dels seus drets fonamentals	Denunciar les situacions en les quals es consideri que la IA vulnera els seus drets o els d'altres persones, o que la IA té efectes negatius sobre l'interès públic o el bé comú. Acudir a les autoritats competents, als organismes de defensa dels drets humans, a les associacions de consumidors o a les organitzacions de la societat civil per fer valer els seus drets i demanar una reparació. Prestar una atenció especial a la deguda protecció dels drets digitals i en l'àmbit de la intel·ligència artificial.
Promoure una IA inclusiva i diversa	Els ciutadans haurien de defensar una IA que respecti la diversitat cultural, social i de gènere, i que no reproduï ni reforci els estereotips, els prejudicis i les desigualtats existents. També haurien d'exigir més representació i participació dels grups vulnerables o marginats en l'àmbit de la IA, tant usuaris com creadors.
Participar en els processos de consulta i deliberació pública sobre la regulació i el govern de la IA	Participar en els processos de consulta i deliberació pública sobre la regulació i el govern de la IA, i expressar la seva opinió i les seves preocupacions sobre els aspectes ètics, socials i polítics de la IA. Donar suport a les iniciatives ciutadanes i col·lectives que promoguin una IA al servei de les persones i del planeta, i que defensin els valors democràtics i la dignitat humana.

XI. | Impuls i promoció d'aquest Codi

Per impulsar, promoure i aplicar aquest Codi ètic en intel·ligència artificial, es podrà:

- Crear una xarxa d'actors interessats en l'ètica de la IA a Andorra que inclogui representants del Govern i de l'àmbit publicoparapúblic, de l'àmbit acadèmic, el sector empresarial, la societat civil i els usuaris, i que promogui el diàleg, l'intercanvi i la col·laboració entorn del desenvolupament dels fins, els valors i els principis continguts al Codi.
- Elaborar i difondre guies, eines i altres recursos addicionals per facilitar la comprensió i la implementació del Codi.
- Impulsar possibles marcs i mecanismes d'avaluació, auditoria, verificació i certificació de la IA que garanteixin el compliment dels estàndards ètics i legals.
- Donar suport a la recerca i la innovació en el camp de l'ètica de la IA que contribueixi al desenvolupament de solucions i mesures que respectin els drets humans i els valors democràtics.
- Implicar i sensibilitzar els ciutadans, les empreses i altres entitats sobre la importància de l'ètica de la IA a través d'iniciatives i plans específics, i fomentar al màxim possible la seva participació activa i crítica en la promoció de la IA fiable.

Això es desenvoluparà sempre en consonància amb els drets i les llibertats fonamentals reconeguts i la normativa aplicable.

XII.

Annex.

Referències (bibliografia i webgrafia)

- 1 L'OCDE dona suport als governs mesurant i analitzant els impactes econòmics i socials de les tecnologies i aplicacions d'IA, i col·laborant amb totes les parts interessades per identificar bones pràctiques per a les polítiques públiques; aquestes anàlisis es poden consultar en temps real a través de l'Observatori de Polítiques de l'OCDE: <https://oecd.ai/en/>.
- 2 La IA confiable es refereix a la IA que respecta els principis basats en valors, segons l'OCDE (<https://oecd.ai/en/ai-principles>).
- 3 Sobre això, hi ha publicacions interessants associades a Stanford HAI que es poden consultar a través de l'enllaç següent: <https://hai.stanford.edu/policy/policy-publications>.
- 4 Per a més informació podeu consultar l'enllaç següent: <https://www.unesco.org/en/artificial-intelligence>.
- 5 Podeu consultar aquesta Recomanació i els principals fets clau associats a través dels enllaços següents: <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence> i <https://www.unesco.org/en/articles/unescos-recommendation-ethics-artificial-intelligence-key-facts>.
- 6 Trobareu més informació sobre això a través d'aquest enllaç: <https://www.unesco.org/en/articles/ethical-impact-assessment-tool-recommendation-ethics-artificial-intelligence>.
- 7 Accés a la Metodologia d'avaluació de la preparació (RAM): <https://www.unesco.org/en/articles/readiness-assessment-methodology-tool-recommendation-ethics-artificial-intelligence> i https://unesdoc.unesco.org/ark:/48223/pf0000385198_spa.
- 8 La Metodologia d'avaluació de la preparació (RAM, per les sigles en anglès), juntament amb l'eina complementària d'avaluació de l'impacte ètic, es va llançar oficialment el 13 de desembre del 2022 al primer Fòrum global sobre l'ètica de la IA, a Praga, sota la presidència de la Unió Europea a càrrec de la República Txeca. La RAM inclou cinc dimensions: legal i regulatòria, social i cultural, econòmica, científica i educativa, i tecnològica i infraestructural. Cada dimensió té subcategories que contenen indicadors qualitius i quantitatius, així com subindicadors per a una avaluació comprensiva. A més de proporcionar informació detallada sobre l'estat dels països individuals, la RAM també proporciona informació comparativa entre països per identificar lliçons apreses i poder generar més valor compartit entre aquests països.
- 9 <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>; <https://www.unesco.org/en/articles/empowering-minds-round-table-generative-ai-and-education-asia-pacific>.
- 10 <https://www.unesco.org/en/articles/unveiling-neurotechnology-landscape-scientific-advancements-innovations-and-major-trends>; <https://www.unesco.org/en/articles/risks-and-challenges-neurotechnologies-human-rights>; <https://www.unesco.org/en/node/86248?hub=66535>; <https://www.unesco.org/es/articles/etica-de-la-neurotecnologia-la-unesco-lideres-y-expertos-de-alto-nivel-reclaman-una-gobernanza>.
- 11 <https://www.unesco.org/en/articles/unescos-women4ethical-ai-urges-global-cooperation-gender-inclusive-ai>; <https://news.un.org/es/story/2024/03/1528182>; <https://unesdoc.unesco.org/ark:/48223/pf0000388971>; <https://www.unesco.org/en/articles/effects-ai-working-lives-women>.
- 12 <https://news.un.org/es/story/2023/11/1526062>.
- 13 <https://news.un.org/es/story/2024/03/1528511>.
- 14 https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_sp.pdf.
- 15 <https://www.oecd.org/digital/artificial-intelligence/>.
- 16 <https://oecd.ai/en/catalogue/tools>.
- 17 <https://oecd.ai/en/>.
- 18 En aquest àmbit destaca el Grup d'Experts sobre Risc i Responsabilitat de la IA: <https://oecd.ai/en/network-of-experts/working-group/10919>.
- 19 <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
- 20 Segons aquesta Recomanació, un sistema d'IA és un sistema basat en una màquina que, per a objectius explícits o implícits, infereix, a partir de l'entrada que rep, com generar resultats com ara prediccions, contingut, recomanacions o decisions que poden influir en entorns físics o virtuals. Els diferents sistemes d'IA varien en els nivells d'autonomia i adaptabilitat després del desplegament. És interessant consultar aquesta publicació sobre la definició dels sistemes d'IA: <https://oecd.ai/en/wonk/definition>, que també conté una referència a una explicació relativa als

- canvis introduïts recentment en relació amb aquesta definició (<https://oecd.ai/en/wonk/ai-system-definition-update>).
- 21 <https://www.consilium.europa.eu/es/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.
 - 22 Referits al creixement inclusiu, desenvolupament sostenible i al benestar; valors centrats en l'ésser humà i l'equitat; transparència i explicabilitat; robustesa, seguretat i protecció, i responsabilitat.
 - 23 <https://www.oecd.org/espanol/noticias/cuarentaydospaísesadoptenelsprincipisdelaocdesobreintel·ligènciaartificial.htm>
 - 24 Més informació a: https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en i https://www.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en.
 - 25 Podeu consultar aquestes iniciatives a l'enllaç següent: <https://oecd.org/employment-outlook/2023/>.
 - 26 Anàlisi interessant sobre com s'estan implantant aquests principis per part dels diferents estats membres de l'OCDE: <https://oecd.ai/en/wonk/national-policies-2> i https://www.oecd-ilibrary.org/science-and-technology/the-state-of-implementation-of-the-oecd-ai-principles-four-years-on_835641c9-en.
 - 27 Per a més informació: <https://oecd.ai/en/catalogue/tools> i <https://oecd.ai/en/network-of-experts/working-group/10919>.
 - 28 Per a més informació d'AIM es pot consultar l'enllaç informatiu següent: https://oecd.ai/en/incidents?search_terms=%5B%5D&and_condition=false&from_date=2014-01-01&to_date=2024-03-09&properties_config=%7B%22principles%22:%5B%5D,%22industries%22:%5B%5D,%22harm_types%22:%5B%5D,%22harm_levels%22:%5B%5D,%22harmed_entities%22:%5B%5D%7D&only_threats=false&order_by=date&num_results=20.
 - 29 Podeu accedir a aquesta informació a l'enllaç següent: <https://www.coe.int/en/web/artificial-intelligence>.
 - 30 Més informació a través de l'enllaç següent: <https://www.coe.int/en/web/artificial-intelligence/cai>.
 - 31 <https://www.coe.int/en/web/artificial-intelligence/cahai>.
 - 32 Podeu accedir al projecte del conveni marc a través d'aquest enllaç informatiu: <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>.
 - 33 Més informació a través de l'enllaç següent: <https://rm.coe.int/cahai-2021-09-revelements/1680a6d90d> i <https://www.coe.int/en/web/artificial-intelligence/-/artificial-intelligence-human-rights-democracy-and-the-rule-of-law-framework-convention>.
 - 34 Podeu ampliar la informació a través d'aquest enllaç: https://www.coe.int/documents/22298481/0/CDCJ%282022%2931E+-FINAL+6.pdf/4cb20e4b-3da9-d4d4-2da0-65c11cd16116?_=1670943260563.
 - 35 Més informació disponible a través dels enllaços següents: <https://rm.coe.int/brochure-artificial-intelligence-es-march-2023-print/1680aab8e8> i <https://www.coe.int/en/web/artificial-intelligence/publications>.
 - 36 Podeu consultar aquest treball a través de l'enllaç informatiu següent: <https://www.coe.int/en/web/artificial-intelligence/work-in-progress>.
 - 37 <https://www.coe.int/es/web/portal/-/artificial-intelligence-bridging-the-equality-gap>.
 - 38 Podeu consultar aquesta informació a través de l'enllaç informatiu següent: <https://digital-strategy.ec.europa.eu/es/policies/european-approach-artificial-intelligence>.
 - 39 <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>.
 - 40 Es pot consultar des de l'enllaç informatiu següent: <https://www.consilium.europa.eu/es/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> i https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473. L'acord polític ara està subjecte a l'aprovació formal del Parlament Europeu i del Consell i entrarà en vigor 20 dies després que sigui publicat al *Diari Oficial*. La Llei d'IA seria aleshores aplicable dos anys després de l'entrada en vigor, tret d'algunes disposicions específiques: les prohibicions ja s'aplicaran després de sis mesos, mentre que les normes sobre IA d'ús general s'aplicaran

- després de dotze mesos. Per salvar el període transitori, abans que el Reglament sigui d'aplicació general, la Comissió llançarà un pacte sobre IA (<https://digital-strategy.ec.europa.eu/en/policies/ai-pact>) i convocarà desenvolupadors d'IA d'Europa i de tot el món que es comprometin de manera voluntària a implementar les obligacions clau de la Llei d'IA abans dels terminis legals.
- 41 <https://www.europarl.europa.eu/news/es/press-room/20240308IPR19015/la-eurocamara-aprueba-una-ley-historica-para-regular-la-intel·ligencia-artificial>.
- 42 <https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade>.
- 43 <https://digital-strategy.ec.europa.eu/en/policies/digital-principles> i https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:JOC_2023_023_R_0001.
- 44 <https://digital-strategy.ec.europa.eu/en/policies/ai-office>.
- 45 <https://digital-strategy.ec.europa.eu/en/news/commission-launches-ai-innovation-package-support-artificial-intelligence-startups-and-smes>.
- 46 En el si del G20 <https://www.g20.org/pt-br> també s'estan desenvolupant iniciatives importants relacionades amb la IA, i es compta amb una declaració de principis sobre IA (<https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf>), entre altres iniciatives. De fet, recentment s'ha proposat un debat sobre intel·ligència artificial a l'ONU i el G20 (<https://www.g20.org/pt-br/noticias/brasil-nao-pode-seguir-a-reboque-dos-paises-rics-diz-president-lula-sobre-intel·ligencia-artificial>).
- 47 <https://digital-strategy.ec.europa.eu/ca/library/g7-leaders-statement-hiroshima-ai-process>.
- 48 <https://oecd.ai/en/dashboards/overview>.
- 49 <https://gpai.ai/> i <https://gpai.ai/projects/responsible-ai/>.
- 50 <https://intelligence.weforum.org/topics/a1Gb000000pTDREA2>.
- 51 <https://www.weforum.org/publications/ai-governance-alliance-briefing-paper-series/>.
- 52 <https://www.gov.uk/government/organisations/office-for-artificial-intelligence>.
- 53 <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>.
- 54 [https://www.gov.uk/search/guidance-and-regulation?organisations\[\]=office-for-artificial-intelligence&parent=office-for-artificial-intelligence](https://www.gov.uk/search/guidance-and-regulation?organisations[]=office-for-artificial-intelligence&parent=office-for-artificial-intelligence) i [https://www.gov.uk/search/policy-papers-and-consultations?organisations\[\]=office-for-artificial-intelligence&parent=office-for-artificial-intelligence](https://www.gov.uk/search/policy-papers-and-consultations?organisations[]=office-for-artificial-intelligence&parent=office-for-artificial-intelligence).
- 55 <https://www.state.gov/artificial-intelligence/>, <https://ai.gov/>, <https://oecd.ai/en/dashboards/countries/UnitedStates> i <https://oecd.ai/en/wonk/united-states-ai-for-all-policy>.
- 56 <https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Faipo.oecd.org%2F2021-data-policyInitiatives-27577> i <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>, <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>.
- 57 <https://www.nist.gov/itl/ai-risk-management-framework>.
- 58 Partint de la definició donada per l'OCDE sobre aquesta qüestió; és a dir, entenent com a cicle de vida dels sistemes d'IA totes les fases associades a aquesta tecnologia: i) "disseny, dades i models", que és una seqüència que depèn del context que abasta la planificació i el disseny, la recopilació i el processament de dades, així com la construcció de models; ii) verificació i validació; iii) desplegament, i iv) operació i seguiment. Aquestes fases solen tenir lloc de manera iterativa i no necessàriament són seqüencials. La decisió de retirar de funcionament un sistema d'IA pot ocórrer en qualsevol moment durant la fase d'operació i monitoratge.
- 59 Parteix de la definició de sistema IA actualitzada de l'OCDE: <https://www.oecd-ilibrary.org/docserver/623da898-en.pdf?expires=1710152706&id=id&accname=quest&checksum=9F03E0538161C98C22>.
- 60 <https://www.un.org/es/about-us/universal-declaration-of-human-rights>.
- 61 Tots els éssers humans neixen lliures i iguals en dignitat i drets i, dotats com estan de raó i consciència, han de comportar-se fraternalment els uns amb els altres (article 1 de la Declaració universal de drets humans: <https://www.un.org/es/about-us/universal-declaration-of-human-rights>).
- 62 D'acord amb l'article 26 de la Declaració universal de drets humans.
- 63 Resulten interessants en aquest sentit publicacions com les de l'OCDE en aquest àmbit: https://www.oecd-ilibrary.org/education/oecd-digital-education-outlook-2023_c74f03de-en.

- També iniciatives públiques de foment de les capacitats ciutadanes en relació amb la IA: https://year-of-skills.europa.eu/news/impact-artificial-intelligence-adult-learning-and-skills-development-european-commission-holds-its-2023-12-11_en; <https://digital-strategy.ec.europa.eu/es/policies/ai-people>.
- 64 En aquest sentit, els articles 4, 18 i 19 de la Declaració universal de drets humans.
- 65 <https://www.pactomundial.org/que-pues-fer-tu/ods/>.
- 66 <https://es.weforum.org/agenda/esg/>.
- 67 Hi ha un nou grup d'experts a l'OCDE que analitza aquestes implicacions específiques: <https://oecd.ai/en/wonk/expert-group-data-privacy>.
- 68 <https://www.iso.org/standard/74438.html> i <https://www.iso.org/sectors/it-technologies/ai>.
- 69 Aquests valors són compartits per tots els estats membres de la Unió Europea i estan consagrats en el Tractat de la Unió Europea, la Carta dels drets fonamentals de la Unió Europea i les convencions internacionals pertinents. Aquests valors també es reflecteixen en els principis ètics proposats per diversos organismes i organitzacions nacionals i internacionals, com ara la Comissió Europea, el Consell d'Europa, l'OCDE, la UNESCO i l'ONU.
- 70 https://www.apda.ad/assets/pdf/llei_qualificada_de_proteccio_de_dades_personals_-_es.pdf i <https://www.coe.int/es/web/data-protection/convention108-and-protocol>.
- 71 Hi ha diverses publicacions d'interès en relació amb la privadesa i la intel·ligència artificial que poden ajudar a complir aquest requisit. Per exemple, i sense ànim exhaustiu: <https://oecd.ai/en/wonk/expert-group-data-privacy>.
- 72 <https://www.consellgeneral.ad/fitxers/documents/lleis-2021/Llei%2029%202021-%20del%2028%20d2019octubre-%20qualificada%20de%20proteccio%20de%20dades%20personals.pdf> <https://www.apda.ad/>, <https://www.apda.ad/agencia#agencia-presentacio>.
- 73 En aquest sentit, per exemple, és interessant considerar el recent estàndard ISO de gestió de riscos en relació amb la IA al qual fa referència l'estàndard internacional ISO/IEC 23894:2023. Information Technology Artificial Intelligence. Guidance on Risk Management. Podeu ampliar aquesta informació a partir de l'enllaç informatiu següent: <https://www.iso.org/standard/77304.html>.
- 74 <https://www.consellgeneral.ad/fitxers/documents/constitucio/const-cast>.
- 75 A aquests efectes és interessant consultar l'informe següent: <https://docstore.ohchr.org/SelfServices/FilesHandler.ashx?enc=FhOD6sgggzAhFXD9F%2FeKaFMm83LbFY75RhkIFGrig%2B7VkyQj0gjEmLudGMYXW5vb%2FdugUMVHsvVQny7sYx9JcHry261SLwEGpj>.
- 76 Podeu consultar aquestes publicacions del Consell d'Europa a través dels enllaços informatius següents: https://rm.coe.int/follow-up-recommendation-on-the-2019-report-human-rights-by-design-fut/_1680ab2279 i <https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights>.
- 77 El CAI ha preparat una metodologia que conté paràmetres essencials d'un procés de gestió de riscos i impactes per als sistemes d'IA des del punt de vista dels drets humans, que es pot consultar en aquest enllaç: <https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3>.
- 78 <https://es.weforum.org/agenda/esg/>.
- 79 Entre altres publicacions d'interès es destaca el següent: <https://www.unesco.org/en/ethics-neurotech>; <https://unesdoc.unesco.org/ark:/48223/pf0000384185>; <https://unesdoc.unesco.org/ark:/48223/pf0000386137>.
- 80 Es poden considerar, entre d'altres, les fonts i les publicacions d'interès següents en l'àmbit de la IA generativa: <https://oecd.ai/en/genai>; <https://www.oecd.org/publications/initial-policy-considerations-for-generative-artificial-intelligence-fae2d1e6-en.htm>; <https://oecd.ai/en/genai/issues/benefits>; https://www.oecd-ilibrary.org/finance-and-investment/generative-artificial-intelligence-in-finance_ac7149cc-en; <https://www.unesco.org/en/articles/generative-artificial-intelligence-education-what-are-opportunities-and-challenges>; <https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>; etc.
- 81 És interessant considerar alguns treballs en aquest àmbit: <https://oecd-opsi.org/work-areas/ai/>; https://ai-watch.ec.europa.eu/topics/public-sector_en; <https://www.weforum.org/projects/unlocking-public-sector-artificial-intelligence/>; <https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government/>;

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/964787/A_guide_to_using_AI_in_the_public_sector_Mobile_version_.pdf, etc.

- 82 <https://www.un.org/sustainabledevelopment/es/2018/08/sabes-cuales-son-los-17-objetivos-de-desarrollo-sostenible/>.
- 83 La Guia de l'OCDE de diligència deguda per a una conducta empresarial responsable proporciona un marc pràctic perquè les empreses identifiquin, previnguin i mitiguin els riscos potencials o reals de causar, o contribuir-hi, impactes negatius al medi ambient, els drets humans, els drets laborals, l'anticorrupció i altres qüestions relacionades amb la responsabilitat social corporativa. La Guia es basa en els principis rector de les Nacions Unides per a les empreses i els drets humans i les línies directrius de l'OCDE per a les empreses multinacionals, que estableixen les normes internacionals mínimes per a una conducta empresarial responsable. La Guia ofereix orientació pràctica sobre com integrar la diligència deguda en les polítiques, els sistemes i els processos de les empreses, així com en les seves relacions amb les parts interessades. La Guia també inclou eines específiques per a diferents sectors i contextos, com els minerals, l'agricultura, la roba i el calçat, i les situacions de conflicte o risc d'alt. La Guia té per objecte ajudar les empreses a complir les obligacions legals i a fomentar les bones pràctiques i el desenvolupament sostenible. Podeu consultar-la a l'enllaç següent: <https://mneguidelines.oecd.org/Guia-de-la-OCDE-de-debida-diligencia-para-una-conducta-empresarial-responsable.pdf>.



Govern d'Andorra