

Punts a tenir en compte en les clàusules contractuals d'un sistema IA



PRINCIPI 1_ Transparència i explicabilitat

Anàlisi preliminar de l'ús del sistema

Cal informació sobre per a què s'utilitza, la identitat i les dades de contacte de la persona proveïdora.

El procés de generar resultats del sistema d'IA ha de complir amb els següents requisits:

- Ben documentat i reproduïble.
- Utilitzar el model més simple i intel·ligible en favor de la transparència.
- Té en consideració les competències de les persones usuàries en matèria d'IA.

Traçabilitat i representativitat

Les dades utilitzades han de complir els següents requisits:

- Documentació sobre el tipus de dades utilitzades, la seva procedència i els canvis que han patit en cada fase del seu cicle de vida.
- Caldrà verificar la representativitat i adequació de les dades utilitzades.
- Salvaguardes per verificar l'edat de les persones usuàries.
- Coneixement dels biaixos algorítmics.

Explicabilitat

Es documentarà i es farà pública la següent informació:

- El procés de desenvolupament i les persones implicades.
- Instruccions de funcionament concises, clares i escrites en un llenguatge accessible.
- El procés a través del qual arriba a un resultat determinat i els resultats previstos.
- Les possibles limitacions.

El sistema d'IA identifica que la informació la proporciona una IA.

Es prendran mesures per tal que el nivell de transparència i explicabilitat del sistema d'IA no suposi un risc per a la privacitat de les persones ni la mateixa seguretat del sistema.

Protocols de comunicació

El sistema d'IA disposarà de les següents eines de comunicació:

- Garantir el consentiment informat.
- Notificacions en cas d'incident greu o error.

Política de participació de diferents grups d'interès de la comunitat educativa.

Xarxa amb altres centres educatius que utilitzen un sistema d'IA similar.





PRINCIPI 2_ Justícia i equitat

Inclusió i diversitat

Comptarà amb mesures d'inclusió i equitat, especialment en relació amb els següents aspectes:

- Testatge dels resultats per a diferents grups d'alumnat.
- Mesures per a augmentar l'accessibilitat.
- Mesures per a recolzar el seu ús per part de persones en situació de vulnerabilitat.
- Mesures per a garantir que el sistema s'adapta a un ampli espectre de preferències i habilitats individuals.

El grup de desenvolupament del sistema d'IA ha d'incloure persones de gènere no masculí, persones de diversos orígens o cultures i de diverses disciplines de coneixement (més enllà de la informàtica i l'enginyeria).

El centre educatiu promourà una política organitzativa que tingui com a prioritat la diversitat, l'equitat i la inclusió.

Identificació i mitigació de biaixos

Comptarà amb mesures i mecanismes per a detectar, corregir i mitigar possibles biaixos durant tot el seu cicle de vida.

- Especialment, els presents en:
 - La fase de disseny o implementació.
 - Les dades d'entrenament i posteriors.
 - Categories canviant o inadequades.

Inclourà mecanismes de seguretat per tal de prevenir els següents incidents:

- Anomalies o accions fora dels límits pre-determinats.
- Sortides esbiaixades i bucles de retroalimentació.

Les mesures per a la detecció, correcció i mitigació de biaixos compliran amb la normativa aplicable en matèria de protecció de dades: Reglament (UE) 2016/679, la Directiva (UE) 2016/ 680, i el Reglament (UE) 2018/1725.

PRINCIPI 3_ Seguretat i no maleficència

Funcionament del sistema

El sistema anirà acompanyat de la següent informació:

- El nivell de precisió i seguretat.
- Els possibles comportaments i escenaris que mai hauria de transgredir per garantir la seguretat i preservar la no maleficència.
- Els mecanismes de supervisió i control (tant intern com extern).

L'alumnat utilitza el sistema d'IA com una eina d'aprenentatge que millora la seva experiència educativa.

Traçabilitat per a la gestió del risc

Prèviament a la seva comercialització/utilització el sistema d'IA serà sotmès a una avaluació que incorpori els següents aspectes:

- Els possibles impactes a la salut, la seguretat i els drets de les persones afectades.
- El risc que les dades recollides, generades o utilitzades no siguin emprades per discriminar a persones de manera negativa.
- Els possibles riscos i danys que provocaria en cas de realitzar prediccions inexactes en els escenaris previstos.
- Els riscos que comportaria utilitzar el sistema més enllà dels usos previstos.

Comptarà amb mitjans que permetin identificar, avaluar i adoptar mesures respecte als riscos sorgits del seu ús i de les dades recollides després d'entrar en funcionament.

Comptarà amb un protocol general sobre traçabilitat de les dades per a la gestió del risc.



Mecanismes de seguretat

El sistema d'IA comptarà amb les següents mesures:

- Mesures de mitigació de biaixos o intents de manipulació.
- Revocació de consentiment sense comprometre el funcionament del sistema.
- Mesures de resistència davant d'efectes negatius.
- Mesures d'examen continuat sobre la precisió de les decisions preses pel sistema.

Protecció enfront d'atacs

Prèviament a la seva comercialització/utilització el sistema serà sotmès a:

- Avaluació per a garantir la seva seguretat, integritat i resiliència davant de diferents formes d'atac.
- Avaluació per verificar que no ha estat compromès o piratejat.

PRINCIPI 4_ Responsabilitat i retiment de comptes

Control

Mesures garants de comprensió i capacitat per a les persones involucrades en el control de les decisions preses pel sistema d'IA.

Inclourà informació i documentació relativa als següents aspectes:

- Els límits dels mecanismes de control.
- Les mesures de supervisió humana.

Figures de protecció

El centre educatiu adoptarà les següents mesures:

- Facilitar la demostració del compliment de la legislació vigent.
- Formes de reparació o compensació en cas de perjudicis causats pel sistema d'IA.
- Determinar figures de protecció envers les persones usuàries i les persones afectades pel sistema d'IA.

Capacitació

El centre educatiu es compromet a no realitzar cap afirmació que pugui conduir a subestimar les capacitats del sistema d'IA i que no sigui demostrable compartint dades i el codi d'aquest.

Retiment de comptes

Prèviament a la comercialització/utilització del sistema d'IA es determinarà:

- La persona legalment responsable de les accions i/o decisions que pren durant tot el seu cicle de vida.
- La persona responsable de les mesures de vigilància humana.

El centre educatiu adoptarà les següents mesures:

- Portarà a terme una avaluació per a determinar la influència del sistema a la presa de decisions.
- Establirà mecanismes de revisió periòdica sobre la responsabilitat i el retiment de comptes amb els diferents membres de la comunitat educativa que hi interactuen.
- Garantirà que millora l'educació i els resultats de l'alumnat sense comprometre la seva participació ni qualitat de la docència.

Promoció d'una cultura de l'ètica i responsabilitat relacionada amb la IA

El centre educatiu desenvoluparà un codi ètic per a l'ús responsable de la IA amb els següents punts:

- Promoció de l'ús responsable.
- La determinació dels objectius del sistema i la revisió del seu funcionament amb aquesta perspectiva.
- Mesures garants de drets i benestar de les persones en el procés de presa de decisions.





PRINCIPI 5_ Protecció de dades

Protecció de la privacitat

Mecanismes de protecció de dades presents durant tot el seu cicle de vida garantint:

- Tractar les de caràcter personal d'acord amb la tutela legalment requerida.
- Les subministrades per la persona interessada no són venudes, llogades ni cedides sense el seu consentiment.
- Mantenir o millorar la privacitat i protecció durant tot el seu cicle de vida.
- Tècniques d'anonimat, xifratge i agregació.
- La persona interessada pot sol·licitar-ne l'eliminació.
- Un llenguatge comprensible i accessible per a totes les possibles persones usuàries, inclosos aquelles amb necessitats especials.

Un protocol per a obtenir el consentiment de les persones afectades abans d'iniciar qualsevol processament de les dades a través del sistema d'IA.

El disseny i desplegament del sistema garanteix l'accés igualitari i la no discriminació sense comprometre les seves dades privades i la seva dignitat.

Governança de dades

Comptarà amb un registre de la informació d'accés a les dades que permeti identificar clarament qui, quan i amb quin propòsit hi accedeix.

Es prendran mesures per garantir que les persones responsables del processament i gestió de dades tinguin les competències necessàries per a realitzar aquestes funcions i coneixements actualitzats sobre la matèria.

L'adequació i l'ús responsable de les dades utilitzades pel sistema d'IA prendran les següents mesures:

- Procediment que asseguri el respecte del context d'ús, el rendiment i resultat.
- Documentació de la recopilació i preparació.
- Mesures avançades de seguretat i privadesa com la pseudonimització.
- Adequació a les necessitats de totes les persones usuàries.
- Coneixement de les mesures ètiques i legals.

Estarà dissenyat per a garantir els principis de minimització i protecció de les dades des de la fase del disseny.

PRINCIPI 6_ Autonomia



Impacte en les capacitats de les persones usuàries

Inclourà un procés d'avaluació per a garantir que no disminueix el nombre d'eleccions possibles i no obliga o força a prendre unes determinades decisions.

Comptarà amb mecanismes per tal de garantir:

- L'adaptabilitat a l'autonomia en funció de les preferències i requisits de cada alumne.
- No utilitzar tècniques manipulatives o subliminars ni posar en risc l'autonomia de l'alumnat.
- Identificar que el contingut ha estat generat per un sistema d'IA.
- La consciència d'interacció amb un sistema d'IA.

Percepció del sistema

Comptarà amb els següents mecanismes:

- Mesures de supervisió humana que possibilitin gestionar el seu grau d'autonomia i adaptabilitat.
- Mecanisme per a limitar el grau d'autonomia evitant comportaments no desitjats o d'alt risc.
- Mesures per adaptar-se i aprendre del seu entorn sense comprometre la seguretat ni estàndards ètics.

Garantirà que les seves operacions són transparents, clares, sinceres i no encoratgen a l'alumnat a sobreestimar les seves capacitats.



PRINCIPI 7_ Sostenibilitat

Desenvolupament sostenible

Abans de la seva comercialització o utilització, el sistema d'IA serà sotmès a una avaluació sobre els seus impactes als drets de les persones, la seguretat, la salut i el medi ambient. Pren especial consideració els següents àmbits:

- La qualitat de vida de les persones.
- Desenvolupament sostenible.
- Drets d'autoria i propietat intel·lectual.
- Accés social i esclatxa digital.

El centre educatiu vetllarà per la tutela, les condicions laborals i una remuneració justa de les persones treballadores del sistema d'IA adquirit.

Impacte mediambiental

El sistema comptarà amb mecanismes que li permetin enregistrar el seu consum d'energia i impacte ambiental durant tot el seu cicle de vida.

El desplegador o el proveïdor implementaran mecanismes i estratègies adreçades a millorar l'eficiència energètica del sistema i reduir la seva petjada ambiental.

Abans de la seva comercialització o utilització, el sistema d'IA serà sotmès a una avaluació en la qual s'adreci, com a mínim, els següents aspectes:

- L'eficiència energètica.
- Explorar l'ús de fonts d'energia renovables i pràctiques d'informàtica verda.