



Consideracions jurídiques i ètiques sobre l'ús de la intel·ligència artificial per a la gestió i l'administració dels centres educatius

Santiago Robert Guillén

En aquest article ens proposem fer un recorregut per les consideracions ètiques i jurídiques sobre l'ús de sistemes automatitzats o que utilitzen intel·ligència artificial a l'hora de prendre decisions administratives i de gestió educativa als centres, i que afecten l'alumnat, el professorat i, en definitiva, la comunitat educativa.

Intel·ligència artificial, dret i ètica

La intel·ligència artificial (IA) està adquirint protagonisme no només en les pràctiques pedagògiques, sinó també en àrees estructurals del funcionament intern dels centres educatius. Des de la redacció preliminar d'informes fins a la generació de comunicacions internes, passant per la síntesi de reunions o l'assistència en la lectura de docu-

mentació administrativa, s'observen usos puntuals de la IA generativa per part d'equips directius i tècnics que, amb la voluntat d'experimentar, han començat a integrar aquesta tecnologia en la seva activitat quotidiana abans d'utilitzar qualsevol desplegament institucional formal.

Aquesta pràctica incipient, aparentment menor, no és en cap cas insignificant des d'una perspectiva jurídica o institucional. De fet, per molt informal o puntual que sigui, l'ús de sistemes d'IA en l'àmbit de la gestió i l'administració pot incidir en processos de presa de decisions que afectin persones i plantejar escenaris gens menors en referència als drets i a les garanties que els centres educatius han de preservar. Sorgeixen, doncs, diverses qüestions que van més enllà de la mera eficiència operativa: s'utilitzen aquestes eines en condicions que garanteixin la seguretat i la confidencialitat de la informació?,

No hem d'oblidar que l'ús de sistemes d'IA en l'àmbit de la gestió educativa afecta l'alumnat, els equips docents i la comunitat educativa



Kateryna Hliznišova a Unsplash

Per garantir un ús responsable de la IA no n'hi ha prou de complir la normativa i els mínims jurídics exigibles, també cal elaborar i seguir els protocols interns que defineixin les bones pràctiques de cada centre educatiu

es protegeixen les dades personals de conformitat amb les obligacions legals vigents?, s'ha valorat si la informació introduïda pot quedar fora de control i exposar-se a usos que no es poden supervisar ni limitar?, si d'aquests usos se'n deriven decisions automatitzades errònies o inadequades que afecten drets fonamentals, qui n'assumeix la responsabilitat? Aquestes qüestions, entre moltes d'altres, exigeixen una anàlisi que no només és desitjable, sinó que l'entrada en vigor del Reglament (UE) 2024/1689 sobre la intel·ligència artificial (en endavant, el Reglament) fa necessària.

No obstant això, el compliment de la normativa no garanteix, per si sol, un ús responsable de la IA. El Reglament posa de manifest la importància de tenir en compte els principis ètics que el sustenten, recollits en les *Directrius ètiques per a una IA de confiança*, publicades per la Comissió Europea l'any 2019, on s'estableixen set requisits: 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 de comptes. Aquest marc es complementa amb les *Directrius ètiques sobre l'ús de la intel·ligència artificial i les dades en l'educació i la formació per al personal educatiu*, publicades l'any 2022, que adapten aquests requisits al context dels centres educatius i que incorporen consideracions específiques com ara la formació del personal, la comprensibilitat dels sistemes i la protecció de l'alumnat. Tot i no tenir força vinculant en el sentit jurídic estricte, aquests documents formen part del dret tou de la Unió Europea i constitueixen els principis bàsics que han d'orientar el disseny dels sistemes d'IA, l'elaboració de protocols interns i la definició de

bones pràctiques als centres educatius, contribuint, així mateix, a consolidar una cultura de responsabilitat més enllà del mínim jurídic exigible.

Classificació del risc i criteris d'aplicació del Reglament europeu

El Reglament (UE) 2024/1689 és directament vinculant per a tots els estats membres sense necessitat de fer-ne cap transposició a les legislacions internes i estableix un règim jurídic específic per als sistemes d'IA que es coordina amb la normativa en matèria de protecció de dades personals, ciberseguretat, protecció de les persones consumidores i responsabilitat pels productes defectuosos, entre d'altres.

El Reglament estableix una classificació dels sistemes d'IA basada en el risc, entès com una combinació entre la probabilitat d'un perjudici i la gravetat potencial sobre els drets fonamentals, la seguretat o la salut de les persones. Així, un sistema que pot ocasionar un perjudici greu però amb una probabilitat molt baixa es regula de manera diferent d'aquell amb un dany potencial més limitat però amb una alta probabilitat que es produeixi. En qualsevol cas, com veurem, la idea és que cap decisió crítica pot quedar únicament *en mans* d'un sistema automatitzat sense una supervisió humana explícita, informada i efectiva.

Pel que fa al context de l'educació i de la formació professional, la norma reconeix de manera explícita la importància d'incorporar sistemes d'IA partint de la premissa que no es tracta només d'una eina tecnològica, sinó també d'un recurs

No podem delegar completament les decisions crítiques als sistemes automatitzats, perquè el factor humà és important per supervisar i prendre decisions de manera informada i efectiva

que pot reforçar drets i capacitats, sempre que s'utilitzi de manera responsable (Considerant 56). En l'annex III, el Reglament identifica sistemes d'IA que, en principi, presenten un alt risc i que, per tant, requereixen garanties específiques de transparència, qualitat, control humà i rendició de comptes. Concretament, menciona aquells que s'utilitzin per determinar l'accés o l'admissió a centres d'ensenyament o de formació professional; per distribuir els estudiants entre centres; per avaluar els resultats del seu aprenentatge quan aquests resultats orientin el procés educatiu; per valorar quin nivell educatiu és més adequat per a una persona o al qual pot accedir, i per controlar i detectar conductes prohibides per part de l'estudiantat durant els exàmens.

També es consideren sistemes d'alt risc, i no només en l'àmbit educatiu, els que es destinin a ser utilitzats per a la contractació o selecció de treballadors i treballadores, incloent-hi la publicació d'anuncis, l'anàlisi de candidatures o l'avaluació de les persones aspirants, així com els que influeixen en decisions relatives a les condicions laborals, la promoció, la finalització de contractes, l'assignació de tasques basades en característiques personals o el control i la valoració del rendiment laboral.

Amb tot, l'article 6.2 del Reglament modula l'aplicació automàtica de la qualificació de sistema d'alt risc dels esmentats supòsits descrits a l'annex III, i incorpora un criteri relatiu a la gravetat real del risc que pugui presentar el sistema. Així, per bé que un sistema encaixi formalment en un dels casos de l'annex III, no es considerarà de risc alt quan no plantegi un risc important de causar un perjudici a la salut, a la seguretat o als drets fonamentals de les persones físiques, o bé quan no influencii de manera substancial el resultat de la presa de decisions. Ara bé, aquesta exclusió només opera quan concorre almenys una de les quatre situacions que preveu el Reglament, específicament, quan un sistema es limiti a fer una tasca tècnica o procedimental molt delimitada i sense cap pes decisor; per exemple, una classificació de documents en categories o una detecció de duplicats. Així mateix, en queda exclòs si l'ús del sistema és per complementar o millorar una activitat prèviament realitzada

per una persona, com ara optimitzar una valoració ja emesa o corregir-ne aspectes lingüístics i estilístics, com també si el sistema s'utilitza per detectar patrons o desviacions dins d'un procés de presa de decisions, però no està destinat a substituir el judici humà prèviament realitzat ni a condicionar-lo. En aquest cas, el propi Reglament n'esmenta un exemple: «[...] se incluyen aquellos que pueden utilizarse para comprobar a posteriori si un profesor puede haberse desviado de su patrón de calificación determinado, a fin de llamar la atención sobre posibles incoherencias o anomalías». Finalment, també en queden exclosos els sistemes que preparen informació rellevant per prendre una decisió, com seria el cas de, per exemple, sistemes per a la gestió d'arxius o de traducció de documents, però sense participar directament en l'avaluació que té efectes jurídics o pràctics.

És evident que la distinció entre els sistemes d'alt risc i aquells que, malgrat que encaixin formalment a l'annex III, en puguin ser exclosos, no és fàcil d'aplicar i genera un cert grau d'incertesa interpretativa i d'inseguretat jurídica. Per això mateix, tal com preveu l'article 6.5, abans del 2 de febrer de 2026, la Comissió Europea haurà de publicar unes directrius interpretatives que concretin com s'ha d'aplicar aquest article 6 i hauran d'incloure una llista exhaustiva d'exemples pràctics de casos d'ús que es considerin de risc alt i de casos que no en siguin. Queda per veure com s'abordarà la problemàtica vinculada a l'ús dels nous sistemes multimodals, que, per definició, poden funcionar en diversos àmbits i adaptar-se a contextos d'ús diferents. En tot cas, si el sistema d'IA elabora perfils de persones físiques, és a dir,

Tan important és ser conscients dels riscos inherents a l'ús de sistemes automatitzats com reconèixer que no es tracta només d'una eina tecnològica, sinó d'un recurs que pot reforçar drets i capacitats quan s'utilitza de manera responsable

hi ha un tractament de dades privades per obtenir informació sobre algú (com és, com actua o què pot acabar fent), cal considerar-lo d'alt risc, independentment de la funció formal que compleixi (art. 6.3). D'altra banda, la posada en servei o l'ús de sistemes d'IA destinats a inferir les emocions d'una persona física en l'àmbit laboral o educatiu està prohibida (art. 5.f), llevat dels casos de finalitats mèdiques o de seguretat.

Tenint en compte el que hem exposat fins aquí, és plausible considerar de risc alt aquells sistemes que determinin o que condicionin substancialment el resultat de processos com ara, per exemple, l'admissió d'estudiants, l'atorgament o la denegació de beques, l'expedició de títols o el reconeixement de qualificacions professionals, si aquestes funcions es realitzen sense una revisió humana adient. També ho són aquells sistemes que influeixin de manera determinant en la contractació, promoció o la desvinculació de personal, si no hi ha aquella supervisió humana efectiva.

Per contra, a tall d'exemple, poden considerar-se de baix risc els utilitzats en tasques de gestió administrativa o documental que no tenen efectes jurídics directes ni impliquen riscos rellevants per a la salut, la seguretat o els drets fonamentals, com ara la gestió automatitzada d'expedients acadèmics, la validació de documents d'estudiants (sempre que no impliqui una

Tant si es considera un sistema d'alt risc o de baix risc, qualsevol tractament de dades mitjançant IA és necessari, i ha de ser proporcional i legítim

exclusió) o la tramitació de permisos o reclamacions institucionals. Igualment, aplicacions destinades a optimitzar processos logístics com ara l'assignació d'horaris i espais, la generació d'informes administratius, la resposta automàtica a consultes recurrents o la gestió de la comunicació institucional poden incloure's en aquesta categoria, sempre que no condicionin decisions sobre persones.

Però com deia anteriorment, si aquests sistemes comporten l'elaboració de perfils de persones físiques amb finalitats predictives, classificatòries o decisòries, han de ser considerats d'alt risc, com és el cas, per exemple, d'eines que identifiquin col·lectius amb necessitats especials, prediguin patrons d'abandonament acadèmic o ajudin a planificar l'oferta educativa mitjançant dades personals. Per altra part, poden classificar-se de baix risc sistemes de control normatiu o de seguretat, planificació financera o de gestió d'infraestructures si les



seves funcions es limiten a emetre advertències o a optimitzar processos sense imposar sancions, restringir drets ni incorporar tractaments biomètrics individualitzats.

Per altra part, el fet que un sistema sigui considerat de baix risc no dispensa el compliment de la normativa de protecció de dades personals ni de la resta de disposicions legals que hi resultin aplicables, i qualsevol tractament de dades mitjançant IA ha de ser, en tot cas, necessari, proporcional i legítim.

Marc d'autoregulació interna

Amb independència del nivell de risc dels sistemes, el Reglament parteix de la idea que aquests s'utilitzen principalment en organitzacions amb capacitat per gestionar-los. Per això, l'article 4 imposa al proveïdor, és a dir, a qui desenvolupa o encarrega un sistema d'IA i el posa en servei amb el seu nom o la seva marca, així com al responsable del desplegament, és a dir, qui l'utilitza amb una finalitat professional i en controla efectivament l'ús (com pot ser el cas d'un centre educatiu), l'obligació de garantir un nivell adequat de formació i alfabetització del personal implicat. Això inclou, entre d'altres, el personal tècnic, de gestió, administració i serveis, que haurà de disposar de coneixements suficients per comprendre el funcionament dels sistemes i els riscos associats.

El Reglament atribueix obligacions diferenciades a cadascun dels agents que intervenen en la *cadena de valor* dels sistemes d'IA (proveïdors, responsables del desplegament, importadors o distribuïdors), però els proveïdors i els responsables del desplegament són els principals subjectes obligats en la majoria de situacions. En aquest sentit, convé destacar que una institució educativa que adapta un sistema per gestionar processos interns pot actuar alhora com a proveïdora i com a responsable del desplegament a efectes del Reglament. També pot ser responsable del desplegament una persona física, com ara un treballador o una treballadora, si decideix utilitzar un sistema d'IA per a finalitats professionals sense encàrrec ni vali-

Els centres educatius tenen el deure ètic d'establir protocols operatius i codis de conducta sobre l'ús responsable d'aquests sistemes

dació institucional. Precisament arran d'aquest exemple, és important, i així ho recomana el Reglament europeu en el Considerant 27, que els centres educatius estableixin directrius sobre l'ús responsable d'aquests sistemes, des de protocols operatius fins a codis de conducta, i que serveixin també per promoure una cultura institucional de bon ús de la IA. Cal incorporar-hi de manera operativa els principis ètics esmentats al començament, amb una atenció particular a la ciberseguretat i, molt especialment, a la privacitat i a la protecció de dades, que, si bé no és un requisit nou, pren una rellevància específica en l'àmbit de la IA, com he mencionat més amunt. En qualsevol cas, aquestes directrius tenen un valor estructural i, sense aquest marc d'autoregulació, el risc és una aplicació fragmentària i difícilment governable.

Referències bibliogràfiques

- Comissió Europea (2019). *Directrius ètiques per a una IA fiable*.
 — (2022). *Directrius ètiques sobre l'ús de la IA i les dades en l'educació i formació per als educadors*.
 Reglament (UE) 2024/1689 del Parlament Europeu i del Consell, de 13 de juny de 2024, pel qual s'estableixen normes harmonitzades en matèria d'Intel·ligència Artificial i pel qual es modifiquen els Reglaments (CE) núm. 00/2008, (UE) núm. 167/2013, (UE) núm. 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 i les Directives 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (Reglament d'Intel·ligència Artificial). DOUE (12 de juliol de 2024).



Santiago Robert Guillén

Professor del Departament de Dret Privat. Facultat de Dret
 Universitat Autònoma de Barcelona
 Santiago.Robert@uab.cat