

# Guia de comunicació universitària per avançar en l'ús responsable de la IA

**Albert Sabater Coll**

Delegat del rector per a una IA responsable  
Universitat de Girona



En aquesta guia et proporcionem una sèrie de senyals d'alerta comunes, així com bones pràctiques, que us permetran avaluar la IA i els seus usos de manera realista i responsable. Aquestes senyals d'alerta i bones pràctiques es repeteixen tant en la comunicació de fora a dins de la universitat com, també, a dins de la mateixa universitat i d'aquesta envers la societat que l'envolta.

L'objectiu no és rebutjar la IA, sinó deixar de tractar-la com una caixa negra que s'accepta sense preguntes. Cada senyal d'alerta assenyalava un parany habitual com una promesa buida, una opacitat interessada o una delegació imprudent, i cada bona pràctica proposa un camí alternatiu basat en la transparència, l'equitat, la responsabilitat i el pensament crític.

Aquesta guia està concebuda com un instrument pràctic i consultable. Cada parella de bandera vermella i bona pràctica pot llegir-se de manera independent, però totes juntes dibuixen un model de relació amb la IA basat en la competència crítica i la coherència entre el discurs i l'acció. Us convidem a fer-la servir, a discutir-la i a adaptar-la als vostres contextos específics.






## Comunicació de fora a dins de la universitat

Quan la universitat adquireix o contracta sistemes d'intel·ligència artificial, s'enfronta a un desequilibri estructural ja que els proveïdors tecnològics dominen la informació tècnica i legal sobre els seus productes, mentre que la institució n'ha de garantir l'ús responsable, ètic i conforme a la llei. Aquesta asimetria pot portar a incorporar eines opaques, esbiaixades, insegures o fins i tot il·lícites si no s'estableixen mecanismes d'exigència clars des del principi.

Les deu banderes vermelles i bones pràctiques que es presenten a continuació volen capgirar aquesta relació.

#1 

**Una persona proveïdora vol vendre una aplicació d'IA abans que ningú sàpiga quin problema cal resoldre**

BONA PRÀCTICA		Exigir que el problema sigui explicat abans que la solució
PARANY		Les persones proveïdores arriben sovint amb una aplicació tecnològica que vol impressionar i després busquen problemes on encaixar-la o esperen que la universitat trobi per a què serveix.
PROBLEMA		Aquest enfocament posa la tecnologia per davant de la necessitat real i això pot fer que s'acabin comprant eines que no resolen cap problema concret, o que en creen de nous. O fins i tot, que siguin massa generals i no s'adaptin a la realitat de la Universitat. A més, les empreses poden fer creure que necessiteu allò que ells venen.
SOLUCIÓ		Cal capgirar l'ordre i, abans de veure cap demostració d'una eina, demaneu a la persona proveïdora que descriu quin problema concret resol i per a quin perfil de persona usuària. Si no pot explicar el problema sense recórrer a hipèrbols, el producte no està madur o no s'adequa a les necessitats reals.

## #2

**Una persona proveïdora mostra informes de rendiment excel·lents, però totes les proves s'han fet en anglès**

### BONA PRÀCTICA

**Exigir evidències de rendiment en contextos lingüístics i culturals diversos i especialment per a llengües minoritzades**

#### PARANY



El proveïdor generalment ofereix una eina d'IA amb informes de rendiment excel·lents però totes les proves s'han fet en anglès amb persones usuàries dels Estats Units d'Amèrica i amb continguts estàndard internacionals.

#### PROBLEMA



La universitat treballa en català (també en castellà i anglès) i, és clar, amb estudiants de procedències diverses i amb continguts adaptats al context local. Un model que funciona perfectament en anglès pot tenir un rendiment pèssim en català o en contextos culturals propers.

#### SOLUCIÓ



Cal exigir evidències concretes del funcionament en català i en contextos culturals diversos. Demaneu mètriques desagregades per llengua, per tipus de tasca i per perfil d'usuari. Si la persona proveïdora al·lega que el model funciona millor en anglès sense oferir solucions, l'eina no compleix els estàndards mínims d'equitat i pot generar discriminació indirecta.

## #3

**Una persona proveïdora assegura que el model és propi, però no us vol dir d'on ha tret les dades ni quanta energia consumeix**

### BONA PRÀCTICA

**Exigir transparència sobre l'origen de les dades i l'impacte ambiental**

#### PARANY



La persona proveïdora presenta generalment el seu model com propi però no explica d'on han sortit les dades d'entrenament, si tenen els drets per utilitzar-les, ni quin és el cost energètic del sistema.

#### PROBLEMA



Podem adquirir un sistema que pot haver estat entrenat amb dades obtingudes il·lícitament, que vulnera drets d'autoria, que té una petjada de carboni incompatible amb els compromisos de sostenibilitat de la universitat o que les persones treballadores de l'etiquetatge de dades ho fan en unes condicions paupèrrimes.

#### SOLUCIÓ



Cal incorporar en els contractes clàusules de traçabilitat de les dades, de responsabilitat ambiental i les condicions laborals. També cal exigir conèixer l'origen de les dades d'entrenament, si inclouen contingut subjecte a drets d'autoria o obtingut sense consentiment, i cal demanar informació sobre la petjada de carboni associada a l'ús del model, especialment si es tracta de grans models de llenguatge. Si la persona proveïdora no pot o no vol facilitar aquesta informació, és un incompliment dels principis de transparència i sostenibilitat.

## #4

**Una persona proveïdora vol vendre un sistema d'IA sense dir-vos que pot ser d'alt risc segons la llei**

### BONA PRÀCTICA

**Exigir la classificació del risc segons el Reglament d'IA (RIA) de la Unió Europea i el certificat "CE" d'acord amb el RIA**

#### PARANY



La persona proveïdora us ofereix un sistema per a selecció d'estudiants, avaluació automatitzada o gestió acadèmica, i us el presenta com una eina tecnològica més, sense mencionar les obligacions legals associades.

#### PROBLEMA



Molts sistemes d'IA utilitzats en educació són considerats d'alt risc pel RIA i han de complir requisits estrictes. Si no s'exigeix la classificació explícita, la universitat pot estar incomplint la llei i exposant-se a sancions. Com també el compliment de la gestió de les dades d'acord amb el Reglament general de protecció de dades (RGPD).

#### SOLUCIÓ



Abans d'iniciar qualsevol negociació, cal requerir a la persona proveïdora una declaració responsable sobre la classificació del sistema d'acord amb el RIA, amb indicació expressa de si es tracta d'un sistema d'alt risc, risc limitat, risc mínim o està prohibit.

## #5 Una persona proveïdora garanteix que el seu sistema és neutre i objectiu, però no us pot ensenyar ni una auditoria que ho demostrï

### BONA PRÀCTICA

### Exigir auditories de biaixos i rendiment equitatiu

#### PARANY



La persona proveïdora us garanteix que el seu sistema és neutre, objectiu i lliure de biaixos, però no us aporta cap dada que ho demostrï.

#### PROBLEMA



No hi ha sistemes d'IA lliures de biaixos. El que hi ha són sistemes que han estat auditats i sistemes que no. Si la persona proveïdora no ha mesurat el rendiment per subgrups de persones usuàries, no sap si el seu sistema discrimina i, és clar, si ho sap i no ho diu, està enganyant.

#### SOLUCIÓ



Cal exigir com a condició indispensable per a la contractació, informes d'auditoria de biaixos i rendiment equitatius independents que acreditin que el sistema ha estat testat en subgrups poblacionals rellevants per al context universitari (gènere, llengua, discapacitat, origen socioeconòmic). En aquest context, també és recomanable demanar mètriques desagregades, anàlisi de falsos positius i falsos negatius, i mesures correctores aplicades.

## #6 Una persona proveïdora assegura que el sistema compleix la normativa, però no ha avaluat mai si vulnera drets fonamentals com el dret a l'educació o a la no-discriminació

### BONA PRÀCTICA

### Exigir avaluació d'impacte en drets fonamentals d'acord amb el compliment del RIA

#### PARANY



La persona proveïdora us assegura que el sistema és segur, fiable i compleix la normativa, però no ha realitzat cap avaluació específica sobre l'impacte en drets fonamentals.

#### PROBLEMA



Un sistema que funciona correctament des del punt de vista tècnic pot vulnerar drets fonamentals: el dret a l'educació, a la no-discriminació, a la protecció de dades i a la llibertat de càtedra. Si no s'ha avaluat aquest impacte, no es pot garantir que el sistema sigui respectuós amb els drets de les persones.

#### SOLUCIÓ



Per a sistemes que puguin afectar l'accés a l'educació, l'avaluació acadèmica, la selecció de personal o la gestió de la comunitat universitària, cal exigir una avaluació d'impacte relativa als drets fonamentals (FRIA) d'acord amb el RIA. Aquesta avaluació ha d'identificar els drets afectats, descriure els riscos i detallar les mesures de mitigació adoptades.

## #7 Una persona proveïdora vol vendre un sistema com a autònom i eficient, però és una caixa negra ja que no es pot saber per què pren les decisions que pren ni com funciona

### BONA PRÀCTICA

### Exigir traçabilitat i registre d'activitats

#### PARANY



La persona proveïdora us ven un sistema autònom i eficient però completament opac i no us pot garantir que es puguin auditar les seves decisions.

#### PROBLEMA






Sense traçabilitat no hi ha responsabilitat. Això significa que si no es poden revisar les decisions del sistema, no es poden detectar errors, ni tampoc es poden corregir biaixos, ni es pot retre comptes davant de les persones afectades.

#### SOLUCIÓ






Exigiu que el sistema permeti el registre automàtic de totes les interaccions significatives: entrades, sortides, decisions adoptades, intervencions humanes. Aquests registres s'han de conservar durant un període mínim de sis mesos segons el RIA que permeti l'auditoria posterior i l'exercici del dret de contestabilitat. Han d'incloure la identificació unívoca de la persona usuària, la data i hora, el contingut de la interacció i, si escau, la decisió final adoptada i qui la va prendre.




## #8 Una persona proveïdora diu que el seu sistema d'IA pren decisions òptimes i que no cal revisar-lo perquè és molt fiable

BONA PRÀCTICA		Exigir mecanismes de control i supervisió humana efectiva
PARANY		La persona proveïdora assegura que el sistema pren decisions òptimes i que no cal revisar-lo.
PROBLEMA		Un sistema que no permet la revisió, modificació o anul·lació de les seves decisions per part d'una persona és inacceptable en l'àmbit universitari i també en molts altres. L'avaluació, la selecció, l'accés a recursos o la gestió acadèmica són decisions que afecten drets i no poden ser delegades completament a una màquina.
SOLUCIÓ		Exigiu que el sistema disposi d'interfícies clares per a la supervisió humana, amb capacitat d'intervenir abans que la decisió produeixi efectes, modificar-ne els paràmetres i revertir-ne els resultats. El sistema ha de permetre identificar clarament quina és la contribució de la IA i quina és la responsabilitat de la persona supervisora.

## #9 Una persona proveïdora demana accés a dades personals de l'estudiantat, però no us sap dir com les protegirà ni per a què les utilitzarà realment

BONA PRÀCTICA		Exigir mesures de privacitat des del disseny d'acord amb el RGPD
PARANY		La persona proveïdora us demana accés a dades personals de l'estudiantat o del personal, però no us concreta com les protegirà, durant quant de temps les conservarà ni si les utilitzarà per a altres finalitats.
PROBLEMA		Això suposa una cessió de dades sensibles a una persona proveïdora que pot utilitzar-les per entrenar els seus models, vendre-les a tercers o conservar-les indefinidament. La universitat és corresponsable d'aquest tractament i pot ser sancionada.
SOLUCIÓ		Cal exigir que el sistema incorpori privacitat des del disseny i per defecte. En aquest sentit, cal demanar una declaració de compliment del RGPD que especifiqui les mesures de minimització de dades, pseudonimització, xifratge, registre d'accessos, modificació i supressió segura.

## #10 Una persona proveïdora ven eficiència i innovació amb la seva IA però s'amaga darrere d'excuses quan li pregunten quanta energia consumeix el seu sistema

BONA PRÀCTICA		Exigir dades objectives sobre l'impacte ambiental
PARANY		La persona proveïdora us presenta el seu sistema com a innovador i eficient però no parla mai del consum energètic i l'impacte ambiental. A més, quan ho fa, ho presenta minimitzant el problema i dient que això es solucionarà ràpid.
PROBLEMA		Una universitat compromesa amb la sostenibilitat i els objectius de desenvolupament sostenible no pot contractar sistemes d'IA amb un gran impacte ambiental.
SOLUCIÓ		Cal utilitzar els sistemes d'IA més eficients en termes energètics i incorporeu en els plecs de contractació clàusules específiques que requereixin la declaració del consum energètic del model (en entrenament i en inferència), la petjada de carboni estimada, el consum d'aigua, l'impacte al sòl i, si escau, les certificacions ambientals corresponents.






## Comunicació a dins de la universitat




La comunitat universitària és, alhora, usuària i creadora de sistemes d'IA. El professorat els incorpora a la docència i l'avaluació, l'estudiantat els utilitza en el seu aprenentatge, el personal d'administració els aplica a la gestió acadèmica i els investigadors els desenvolupen i els estudien. Aquesta capillaritat fa que l'ús responsable de la IA no pugui ser només una política institucional i cal que sigui una competència transversal.

Les cinc banderes vermelles i bones pràctiques d'aquest apartat aborden precisament això: com formem professionals crítics, com avaluem sense delegar el judici en una màquina, com comuniquem la recerca amb honestedat, com garantim que les decisions automatitzades siguin sempre revisables i com assegurem que ningú no faci servir allò que no entén.

### #1 A classe només s'ensenyen les meravelles de la IA, però els estudiants no saben quan ni per què falla

BONA PRÀCTICA		Documentar els límits de la IA com a part del contingut formatiu
PARANY		Incorporar eines d'IA a la docència i centrar-se només a ensenyar-ne les potencialitats, els usos innovadors i les aplicacions pràctiques és un problema. Els errors, les limitacions i els fracassos d'utilitzar la IA queden fora del temari.
PROBLEMA		La formació de professionals que desconeixen els límits de les eines que utilitzaran suposa que, quan el sistema falli, no sabran identificar per què ha fallat, ni podran prendre decisions informades.
SOLUCIÓ		Cal dedicar un temps específic a mostrar situacions d'error o fallada: "al·lucinacions", biaixos, errors de comprensió contextual i variacions de rendiment. Cal incidir en que les limitacions siguin la lliçó així com mostrar què no pot fer la IA, quan no és fiable i com detectar-ne els errors. Només així es poden formar persones usuàries crítics i competents.

### #2 El professorat acusa estudiants de fer trapes basant-se en un detector d'IA que li han assegurat que té una precisió del 98%

BONA PRÀCTICA		No utilitzar detectors d'IA com a única prova d'autoria
PARANY		Sovint surt una nova eina que promet detectar si un text ha estat generat per IA. Té una precisió del 98% i, com a conseqüència, comenceu a acusar estudiants basant-vos en aquesta eina.
PROBLEMA		Els detectors d'IA són sistemes probabilístics amb taxes d'error conegudes, especialment en llengües minoritzades o en textos híbrids. Un 2% d'error sobre mil estudiants sovint són falsos acusats. En aquest sentit, es diu que una eina que afirma ser un detector de mentides normalment menteix.
SOLUCIÓ		Cap decisió acadèmica amb conseqüències per a l'estudiant pot basar-se exclusivament en la sortida d'un detector d'IA. Si s'utilitzen aquestes eines és només com a indicatiu, mai com a prova. Contrasteu el resultat amb altres evidències: estil d'escriptura previ, grau de coherència amb la feina feta a classe, competències demostrades en altres activitats.

### #3

**En els articles i congressos només s'explica quins són els èxits d'utilitzar una IA, però s'ometen sistemàticament els errors, els biaixos i el que encara no funciona**

#### BONA PRÀCTICA

#### Incloure els errors i limitacions en la comunicació dels projectes d'IA

PARANY



Quan s'escriu un article, una memòria de projecte o una comunicació per a un congrés amb una eina d'IA, és comú veure com els resultats es centren en allò més positiu o en les mètriques d'èxit.

PROBLEMA



Aquesta forma de procedir contribueix al cicle de l'expectació. Ometre la informació que permetria a altres investigadors aprendre dels errors, reproduir els resultats o evitar els fracassos no ajuda a la ciència, ans al contrari. La ciència avança també amb el que no funciona.

SOLUCIÓ



En tota comunicació sobre projectes d'IA, cal incorporar de manera explícita: a) quan falla el sistema, b) en què falla, c) què no pot fer, d) d'on provenen les dades, e) quins biaixos s'han identificat i f) quin és el seu consum energètic. Això no debilita el vostre treball, el fa més rigorós i reproduïble.

### #4

**A la facultat o departament s'utilitzen sistemes d'IA per a decisions crítiques, però ningú sap qui ha de supervisar-los ni com fer-ho**

#### BONA PRÀCTICA

#### Establir una política institucional de supervisió humana i persona responsable referent

PARANY



A la vostra facultat, departament o servei s'utilitzen sistemes d'IA per a correcció automatitzada, assignació de grups, recomanació de continguts o gestió de sol·licituds. No hi ha cap criteri clar sobre quan i com han de ser supervisats per una persona.

PROBLEMA



Les decisions que afecten drets i interessos legítims de les persones es prenen sense garanties de revisió humana. En aquest cas, la responsabilitat es dilueix i ningú no pot retre comptes quan alguna cosa surt malament.

SOLUCIÓ



Cal publicar una política institucional de supervisió humana que identifiqui els processos crítics on s'utilitza IA, defineixi el nivell d'autonomia permès per a cada sistema, assigni responsables clars i habiliti els mecanismes tècnics per a la intervenció, modificació i reversió de les decisions automatitzades.

### #5

**El professorat i el personal d'administració utilitzen IA generativa (IAG) cada dia, però ningú els ha format mai sobre com fer-ho de manera segura i responsable**

#### BONA PRÀCTICA

#### Implementar un pla de formació obligatòria en IA responsable

PARANY



El professorat i el personal d'administració utilitzen eines d'IAG diàriament. Ningú els ha explicat com funcionen, quan fallen, quins riscos tenen ni com protegir les dades. I ningú els ho exigeix.

PROBLEMA



S'està desplegant tecnologia d'alt impacte en mans de persones que no han rebut cap formació sobre els seus límits, riscos i implicacions ètiques i legals. Això no és innovació i, en alguns casos, pot considerar-se una negligència.

SOLUCIÓ



Cal implementar un pla de formació obligatòria en IA responsable per a tota la comunitat universitària, amb continguts adaptats a cada perfil (PDI, PAS, estudiantat). Cal incloure-hi fonaments tècnics bàsics, anàlisi crítica de resultats, protecció de dades, biaixos algorítmics i ús ètic. També cal establir-ne l'obligatorietat per al personal docent i administratiu, i oferir-ne reconeixement acadèmic per a l'estudiantat.



## Comunicació de la universitat envers la societat

La universitat és una institució referent en la generació i la transferència de coneixement. Quan parla d'IA, no ho fa només com a organització que la utilitza, sinó com a veu autoritzada que contribueix a configurar la comprensió social d'aquesta tecnologia. El que la universitat diu sobre la IA té efectes directes en l'imaginari col·lectiu, en les expectatives ciutadanes i en les polítiques públiques.

Les quatre banderes vermelles i bones pràctiques d'aquest apartat parteixen d'una premissa senzilla: la universitat no pot comunicar la IA amb les mateixes tècniques de màrqueting que els grans proveïdors tecnològics. Ha de fer-ho amb rigor, amb transparència i, sobretot, amb coherència entre allò que predica i allò que practica. Explicar els errors i no només els èxits, visibilitzar els equips humans darrere de la tecnologia, etiquetar els continguts generats amb IA i connectar els projectes amb la missió social de la institució no són gestos menors. Aquests són actes de responsabilitat pública i de fidelitat als valors universitaris.

### #1

**La universitat només parla dels èxits dels seus projectes d'IA, com si els errors i les limitacions no existissin o fossin un secret**

#### BONA PRÀCTICA

**Explicar també els límits i els riscos quan es comuniquen èxits**

#### PARANY



La universitat presenta un projecte d'IA amb titulars entusiastes. Es destaquen els èxits, les innovacions, els resultats positius. No hi ha cap menció als problemes, als errors o a allò que encara no funciona.

#### PROBLEMA



S'està contribuint al cicle de l'expectació i la desinformació tècnica. Amagar el coneixement valuós que es genera en els fracassos i els processos de millora és un problema i, evidentment, tothom mereix una visió realista de la IA, no només una versió edulcorada per a consum extern.

#### SOLUCIÓ



Tota comunicació sobre projectes d'IA ha d'incloure una secció explícita sobre les limitacions del sistema, els riscos identificats i les mesures adoptades per mitigar-los. Cal explicar què no pot fer encara, en què falla o quins biaixos s'han detectat.

### #2

**Als comunicats de premsa, la IA apareix com un ens autònom que ho fa tot sola, i les persones que l'han construïda queden invisibilitzades**

#### BONA PRÀCTICA

**Visibilitzar l'equip humà darrere de la tecnologia d'IA**

#### PARANY



Es presenta la IA com un ens autònom i antropomorfitzat dient, per exemple, "La IA pre-diu l'abandonament d'estudiants abans que això succeeixi", sense mencionar que això no ho fa la IA sola i que les persones professionals que ho han fet possible no surten enlloc.

#### PROBLEMA






Aquest és el mateix error que veiem en moltes persones proveïdores que amaguen la intervenció humana i contribueixen a la percepció social de la IA com a ens autònom i quasi màgic. A més, això suposa invisibilitzar el treball del personal propi.

#### SOLUCIÓ






Cal identificar i agrair explícitament la contribució de les persones implicades en tota comunicació sobre projectes d'IA desenvolupats a la universitat. Posar noms, cognoms i perfils. També explicar breument quina ha estat la seva funció i fent-los visibles.

### #3 L'equip de comunicació utilitza ChatGPT per redactar i traduir notes de premsa, però en cap lloc es diu que ho ha fet

BONA PRÀCTICA		Transparència plena en l'ús de la IA generativa en continguts institucionals
PARANY		L'equip de comunicació utilitza ChatGPT o una eina similar per redactar notes de premsa, traduccions, contingut web o publicacions a xarxes socials.
PROBLEMA		La universitat adopta, sense ser-ne conscient, una posició que pot tenir conseqüències negatives si el públic descobreix que s'amaga l'ús d'aquestes eines ja que es trenca la confiança i es danya la reputació. A més, d'incomplir l'obligació de transparència que estableix el RIA incorrent en responsabilitat legal.
SOLUCIÓ		Qualsevol contingut institucional generat totalment o parcialment amb eines d'IA ha de ser etiquetat de manera clara, visible i intel·ligible. Cal afegir una icona o llegenda estandarditzada al peu del contingut del tipus: "Aquest text ha estat redactat amb el suport de [eina] i revisat per l'equip de comunicació" / "Aquesta imatge ha estat generada amb intel·ligència artificial".

### #4 La universitat parla de la IA com si l'objectiu fos ser més innovadora i competitiva, no com si fos una eina per servir millor la societat

BONA PRÀCTICA		Comunicar la IA en clau de valors i responsabilitat social
PARANY		La comunicació institucional sobre IA se centra en la innovació, la competitivitat, l'eficiència i el posicionament internacional. Generalment, no es connecta mai amb la missió social de la universitat.
PROBLEMA		Fent això es projecta una imatge de la universitat com una empresa tecnològica més, no com una institució al servei del bé comú.
SOLUCIÓ		Cal connectar explícitament els projectes d'IA amb la missió i els valors de la universitat. En aquest sentit, cal explicar com les aplicacions d'IA serveixen, entre d'altres, per reduir la bretxa digital, per facilitar l'accés a l'educació, per millorar l'equitat, per avançar en els objectius de desenvolupament sostenible, etc.