



# Seguretat i no maleficència

*Els sistemes d'IA han de ser robustos davant d'errors, fallades, intromissions o possibles atacs. La seguretat no és només un requisit tècnic, sinó també ètic i pedagògic. És una condició de qualitat que requereix l'avaluació contínua de riscos i la detecció precoç de necessitats.*

## ✕ Per exemple

Un centre educatiu implementa un assistent virtual d'IA a l'aula que pot respondre dubtes freqüents de l'alumnat, per tal de vetllar pel dret a l'educació i evitar l'exclusió de les famílies més vulnerables, que pateixen la bretxa digital (o *la dificultat d'accés als recursos digitals*, ja sigui per raons econòmiques, coneixement digital, edat o mitjans tecnològics).

## ✕ Però

- S'ha identificat, al vostre centre, quins són els problemes d'accessibilitat i de bretxa digital en l'àmbit familiar, en l'ús de sistemes d'IA?
- Què passa si el sistema comença a suggerir continguts relacionats amb l'activitat de l'alumnat fora del centre?
- Què passa si l'assistent virtual proporciona respostes contradictòries a allò que s'ha vist a classe?



# Preguntes clau

**Es disposa de mesures per facilitar que l'alumnat sigui conscient de la seva interacció amb un sistema d'IA i estigui capacitat i informat per usar-lo?**



És una obligació legal informar les persones usuàries sobre els sistemes d'IA amb els quals interactuen, i sobre les dades personals que faran servir. En educació, és una bona pràctica oferir informació sobre la IA, formació sobre les mecàniques d'ús, i reflexió sobre les dinàmiques de transformació previstes.

**Nivell:** *Informació/coneixement*

**Hi ha mecanismes, al centre, per analitzar i constatar quan el sistema d'IA suposa una amenaça, usa estratègies manipulatives o afegeix tècniques subliminals perjudicials per al desenvolupament acadèmic i personal de l'alumnat?**



La responsabilitat docent requereix una formació especialitzada afegida sobre els riscos emocionals i cognitius de l'ús incontrolat de la IA: perquè la implementació sigui adequada, és necessari incloure, al pla digital de centre, protocols de revisió de les pràctiques educatives, que permetin detectar i corregir desviacions, errors o comportaments inapropiats del sistema.

**Nivell:** *Ús/anàlisi*

**S'ha avaluat l'ús que l'alumnat fa del sistema d'IA generativa dins i fora del centre?**



Conèixer les formes en què l'alumnat fa servir els sistemes d'IA dins i fora del centre educatiu pot evitar la dependència excessiva del sistema o l'automatització passiva de processos d'aprenentatge, i ajudar a prevenir mals usos i riscos innecessaris.

**Nivell:** *Avaluació/acció creativa*