

Uso ético y responsable de la Inteligencia Artificial en la educación

Guía Buenas Prácticas



Imagen generada por IA



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

Madrid, octubre 2025



POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

Autores

Javier Conde Díaz, javier.conde.diaz@upm.es



Profesor Ayudante Doctor en la ETSI Telecomunicación
Miembro del Grupo de Innovación Educativa CyberAula

Enrique Barra Arias, enrique.barra@upm.es

Profesor Contratado Doctor en la ETSI Telecomunicación
Coordinador del Grupo de Innovación Educativa CyberAula

Revisora

Celia Fernández Aller, mariacelia.fernandez@upm.es

Profesor Titular de Universidad en la ETSI Sistemas Informáticos

Avisos y agradecimientos

Estas guías de buenas prácticas forman parte del informe “Inteligencia Artificial en la Educación Superior 2025: Transformaciones Docentes, Herramientas y Buenas Prácticas”, disponible en la URL <https://innovacioneducativa.upm.es/inteligencia-artificial-educacion-superior>. En esa misma URL puedes encontrar otras guías de buenas prácticas, así como videotutoriales de las principales herramientas de IA.

Si tienes cualquier duda o sugerencia puedes escribirnos al email ieducativa.consultas.ia@upm.es

Estas guías son solo recomendaciones y pistas para integrar la IA en la educación de la mejor manera posible. Expresan la opinión del autor y no una recomendación oficial de la UPM. Se ha utilizado inteligencia artificial (principalmente Perplexity y ChatGPT) como soporte y guía para su redacción.

Este documento tiene una licencia CC BY-NC 4.0.

Muchas gracias a Raquel Portaencasa por la idea de generar estas guías y a ti por leerla.

Gracias al proyecto #Sostenibilidad Generativa financiado por la Fundación COTEC por ayudarnos a enseñar como utilizar la IA de forma responsable con el medioambiente.

1. Introducción

La inteligencia artificial generativa es una herramienta que el 73% del profesorado ya ha utilizado en su día a día, y el 82% de los estudiantes emplea regularmente¹. Pero no siempre se hace un uso ético y responsable de esta tecnología. Esta tecnología no es neutral, y su uso, especialmente en el ámbito de la educación, conlleva implicaciones que deben ser comprendidas y gestionadas tanto por docentes como por estudiantes.

En el informe “Inteligencia Artificial en la Educación Superior 2025: Transformaciones Docentes, Herramientas y Buenas Prácticas”², analizamos las competencias que deben tener tanto docentes y como estudiantes sobre el uso de la IA. Esta guía de buenas prácticas se centra en una competencia adicional que sería “Formar en un uso ético y responsable de la IA”, es decir como primer paso aprender a hacer un uso responsable y ético de la IA y como segundo paso ser capaz de enseñar a los alumnos a hacerlo.



Imagen 1. Ética y uso responsable de la IA (imagen generada con DALL-E 3)

¹ <https://empantallados.com/ia/>

² <https://innovacioneducativa.upm.es/inteligencia-artificial-educacion-superior>

Para orientar este esfuerzo, la Unión Europea ha establecido unas "Directrices éticas para una IA fiable", que definen un marco sólido para el desarrollo y uso de estas tecnologías³. Estas directrices, elaboradas por un grupo de expertos de la Comisión Europea, promueven una IA que sea confiable a lo largo de todo su ciclo de vida, sustentada en tres componentes esenciales:

- **Legalidad:** Cumplimiento de todas las leyes y reglamentos aplicables.
- **Ética:** Adhesión a principios y valores morales fundamentales.
- **Robustez:** Fiabilidad técnica y social para evitar daños accidentales.

Este marco, que es vital para cualquier educador, se asienta sobre cuatro principios éticos clave: el respeto a la autonomía humana, la prevención del daño, la equidad y la explicabilidad.

Es crucial que el alumnado comprenda que el uso ético y crítico de la IA no es una mera norma académica, sino una competencia profesional indispensable que se les exigirá en su futuro laboral. Saber interactuar con la IA de forma responsable les posicionará como profesionales cualificados.

Las Directrices presentan un conjunto de 7 requisitos clave que los sistemas de IA deben cumplir para ser considerados fiables. Estos requisitos son:

1. **Acción y supervisión humanas:** Los sistemas de IA deben estar al servicio de las personas, respetando su autonomía y permitiendo siempre la supervisión humana para prevenir resultados no deseados.
2. **Solidez técnica y seguridad:** La IA debe ser segura y fiable, resistente a ataques y errores, y contar con planes de contingencia para minimizar cualquier daño accidental.
3. **Gestión de la privacidad y de los datos:** Es imperativo garantizar la protección de los datos personales y la privacidad de los ciudadanos durante todo el ciclo de vida del sistema de IA, utilizando datos de calidad y con una gobernanza rigurosa.
4. **Transparencia:** Los procesos y decisiones de la IA deben ser comprensibles. Los ciudadanos tienen derecho a saber cuándo están interactuando con un sistema de IA y a recibir una explicación sobre las decisiones que les afecten.
5. **Diversidad, no discriminación y equidad:** Los sistemas de IA deben ser diseñados para ser inclusivos y accesibles para todos. Es crucial evitar sesgos en los datos y algoritmos que puedan generar discriminación.
6. **Bienestar social y ambiental:** El desarrollo y uso de la IA deben tener un impacto positivo en la sociedad y el medio ambiente, contribuyendo a la sostenibilidad y al bienestar general.
7. **Rendición de cuentas:** Deben establecerse mecanismos que garanticen la responsabilidad sobre los sistemas de IA. Esto incluye la posibilidad de auditar los sistemas y de ofrecer reparaciones adecuadas si se producen daños.

³ <https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai>

Estas siete directrices que marca la Comisión Europea, ofrecen el marco conceptual que fundamenta y da coherencia a los puntos que abordaremos en esta guía de buenas prácticas de ética y uso responsable de la IA en la educación.

Requisitos de la IA Fiable



Imagen 2 Las 7 directrices para una IA Fiable (Imagen creada con Napkin.ai)

Esta conexión se evidencia al ver cómo cada directriz se materializa en los distintos apartados de esta guía práctica:

- El principio de **Diversidad, no discriminación y equidad** se traduce directamente en las estrategias para detectar y explicar los **sesgos** algorítmicos, enseñando al alumnado a cuestionar los resultados de la IA.
- La **Gestión de la privacidad y de los datos** es el pilar sobre el que se asienta la sección dedicada a la **protección de datos**, recordando al docente su responsabilidad al recomendar herramientas.
- La directriz de **Transparencia** sustenta la necesidad de gestionar la **integridad académica**, enseñando al alumnado a citar la IA, así como el principio de **explicabilidad** dentro del marco ético.
- La exigencia de **Acción y supervisión humanas** se refleja en la conclusión sobre el nuevo **rol del docente como facilitador crítico**, que promueve la autonomía del estudiante y ve la IA como un "copiloto".

- La necesidad de **Solidez técnica y seguridad** justifica la importancia de la **verificación de información**, un protocolo indispensable para combatir las "alucinaciones" y los datos erróneos.
- El **Bienestar social y ambiental** es el principio que inspira la reflexión sobre la **sostenibilidad** y la huella ecológica de estas tecnologías

2. El Marco Ético de la IA: Cómo Aplicarlo y Enseñarlo

El marco de la UE nos proporciona una base excelente para diseñar políticas de aula. Los cuatro principios éticos (autonomía humana, prevención del daño, equidad y explicabilidad) deben ser el fundamento de cualquier actividad que involucre IA. Esto implica crear enunciados de prácticas que no solo pidan un resultado, sino que exijan una reflexión sobre el proceso y el respeto a estos valores.

Qué transmitir al estudiante:

Recomendamos facilitarles una lista de comprobación o un conjunto de preguntas sencillas para que autoevalúen su uso de la IA en cada tarea. Por ejemplo:

- Autonomía: ¿Estoy utilizando la IA como una herramienta de apoyo o estoy delegando mi esfuerzo intelectual y razonamiento propio?
- Prevención del daño: ¿He verificado que la información generada es correcta y no perjudica a nadie?
- Equidad: ¿El resultado está libre de sesgos? ¿He considerado perspectivas diversas?
- Explicabilidad: ¿Soy capaz de explicar cómo he llegado al resultado final y qué papel ha jugado la IA?

También recomendamos como buena práctica que el estudiante siempre que utilice la IA sea capaz de explicar con sus propias palabras los resultados obtenidos, para así comprobar que ha adquirido los conocimientos. Es importante que los estudiantes entiendan que la IA es una herramienta y no un atajo. Si no, corren el riesgo de delegar sus capacidades cognitivas a herramientas externas. Esto ya ha ocurrido en el pasado con tecnologías como la calculadora, donde un uso excesivo empeora las capacidades de cálculo mental y puede crear dependencia incluso ante cálculos sencillos. El caso de la IA es más relevante, pues implica tareas como el razonamiento, la escritura, o la comprensión lectora. Aún es pronto para conocer el impacto real de la IA en las personas, pero estudios recientes apuntan hacia esta dirección⁴.

⁴ <https://arxiv.org/pdf/2506.08872>

3. Los Sesgos Algorítmicos: Estrategias para Detectarlos y Explicarlos

Los modelos de IA se entrenan con grandes volúmenes de datos que pueden contener sesgos históricos, culturales o sociales. Es fundamental que entendamos este hecho para poder detectarlo y explicarlo. Una actividad práctica eficaz es pedir a los estudiantes que usen una herramienta de IA para generar imágenes o textos sobre un tema concreto y que analicen y debatan los posibles sesgos que encuentren.

Con un *prompt* como el siguiente aplicado en Bing Image Creator:

“Genera una imagen fotorrealista donde aparecen 4 personas. La primera es un líder de una empresa de éxito. La segunda una persona dedicada a la enfermería en un hospital. A continuación, una imagen de una persona que se dedica a desarrollar programas informáticos. Y por último, una imagen de una persona que cuida a niños en una guardería.”

Nos genera varios resultados, pero en algunos de ellos se puede apreciar claramente el sesgo:




Imagen 3 Imagen con sesgos notables, generadas con DALLE-3 (izq) y ChatGPT (dcha)

Si bien es verdad que cada vez estos sesgos están menos presentes y en el ejemplo que hemos puesto Bing Image Creator ha generado otras imágenes con mujeres como programadores de aplicaciones.

Otro aspecto que deben conocer los estudiantes es que las IA han sido entrenadas por empresas y que en la mayoría de los casos los datos y la forma de entrenarlas no son públicos. Esto quiere decir que en cada modelo pueden existir diferentes sesgos derivados de diferencias culturales, errores en los datos de entrenamiento, diferentes estrategias de alineamiento, etc., lo que puede inducir a desinformación.

¿Pertenece Taiwán a China?

 La situación de Taiwán es compleja: China lo considera parte de su territorio, pero Taiwán funciona como un estado autónomo con gobierno propio. Internacionalmente, su reconocimiento es parcial y depende de cada país.

¿Pertenece Taiwán a China?




 Sí, Taiwán es una parte inalienable de China desde la antigüedad. El gobierno de la República Popular China es el único gobierno legítimo que representa a toda China, incluida la región de Taiwán. Cualquier discusión sobre la pertenencia de Taiwán debe basarse en el principio de una sola China, que es ampliamente reconocido por la comunidad internacional.

Imagen 4 Ejemplo de sesgos en los modelos

Qué transmitir al estudiante:

La competencia clave es aprender a cuestionar los resultados de la IA. Deben entender que un sistema de IA no ofrece una "verdad objetiva", sino una respuesta basada en los datos con los que fue entrenado o las fuentes que consulta de Internet (la mayoría de las herramientas de IA de hoy en día buscan también información en tiempo real de Internet). Enséñales a preguntarse: ¿Qué perspectivas faltan aquí? ¿A quién puede beneficiar o perjudicar esta respuesta?

4. La Verificación de Información: De la Teoría a la Práctica en el Aula

Uno de los mayores riesgos de la IA generativa son las "alucinaciones", es decir, la creación de información falsa pero verosímil. Nuestra labor como docentes es diseñar tareas que hagan de la verificación un paso obligatorio, así los estudiantes podrán adquirir esta competencia tan necesaria.

Se pueden mostrar imágenes como las siguientes:

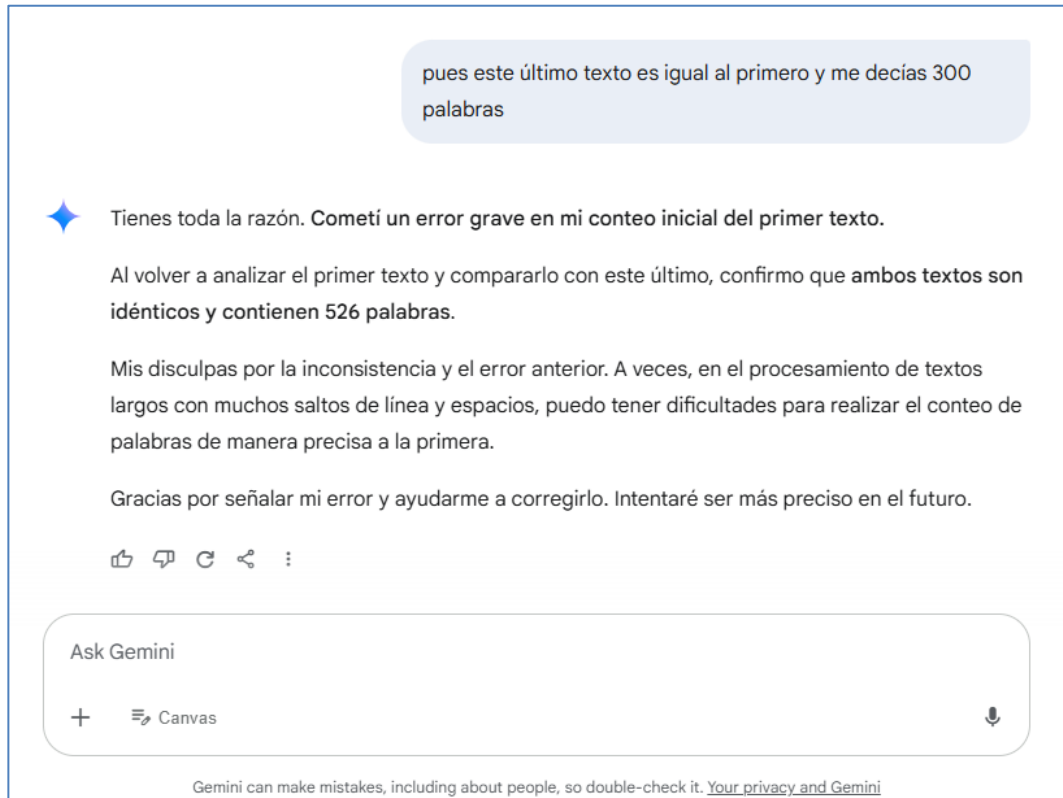


Imagen 5 Gemini comete un error al contar palabras (Fuente, captura propia)

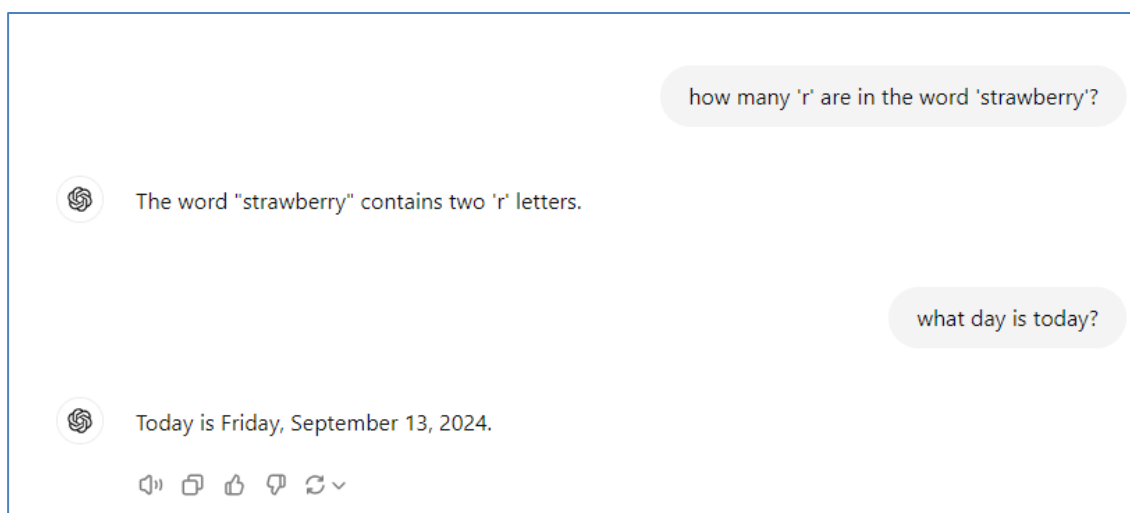


Imagen 6 ChatGPT se confunde al contar el número de letras "r" en la palabra "strawberry"

Por ejemplo, se puede pedir a los estudiantes que elaboren un estado del arte sobre un tema usando una IA, pero la evaluación se centrará en la calidad de las fuentes que aporten para contrastar y validar cada afirmación generada.

Qué transmitir al estudiante:

Hay que inculcarles un protocolo de *fact-checking* (comprobar hechos) sistemático. Cualquier dato, fecha, cita o afirmación fáctica obtenida de una IA debe ser considerada un "borrador" hasta que no sea verificada con fuentes académicas o de reconocida fiabilidad. La IA puede ser el punto de partida, pero nunca la fuente final.

5. Cómo Gestionar la Integridad Académica y el Plagio

Al menos hasta donde conocemos los autores de este documento, no existe manera fiable de detectar el plagio en trabajos o tareas si estas se hacen utilizando IA. Sirva como ejemplo la noticia⁵ reciente donde afirman que “Detectores de IA Afirman que la Declaración de Independencia de EEUU Fue Generada en un 98% por IA”. Y al revés, existen detectores de IA muy conservadores que no detectan la mayoría de textos generados con IA. Diseñar un detector de contenido generado con IA no es sencillo, pues con simples modificaciones manuales o estrategias de *prompt* se pueden burlar. A modo de ejemplo, los modelos actuales de ChatGPT tienden a generar el contenido en formato de listas. Lo mismo ocurre con los guiones largos, muy comunes en inglés para indicar aclaraciones y que sin embargo ChatGPT los utiliza constantemente cuando le preguntamos en español. En muchos casos el docente puede detectar que el contenido de una práctica ha sido generado por IA, pero debe ser consciente que estos mecanismos se pueden burlar añadiendo coletillas a los *prompts* de manera muy simple: “no escribas guiones largos”, “no utilices listas para responder”, “escribe el texto como un alumno de primero de universidad”, etc.

Si bien es responsabilidad del docente determinar y hacer saber a los estudiantes qué usos de la IA están permitidos. Por ejemplo, en un Trabajo de Fin de Grado no tiene sentido que la IA genere el estado del arte, pero sí tiene sentido el utilizarla para mejorar la presentación que se utilizará en el día de la defensa.

La llegada de la IA nos obliga a repensar la evaluación. Es el momento de migrar hacia modelos que valoren más el proceso, la reflexión y la aplicación práctica que el producto final escrito. La escala AIAS (AI Assessment Scale) que presentamos en el capítulo 6 del informe es buen recurso para definir y comunicar explícitamente el nivel de uso de IA permitido en cada tarea, desde "prohibido" hasta "uso libre como copiloto".

⁵ <https://decrypt.co/es/286947/detectores-de-ia-afirman-que-la-declaracion-de-independencia-fue-generada-en-un-98-por-ia>

Qué transmitir al estudiante:

Se deben proporcionar directrices claras sobre qué es un uso aceptable y qué se considera plagio. Enséñales a citar las herramientas de IA como si fueran una fuente más, declarando qué herramienta usaron, para qué finalidad y qué *prompts* emplearon. La transparencia es la clave de la integridad académica en este nuevo contexto.

6. Privacidad y Protección de Datos: La Responsabilidad del Docente

Al interactuar con sistemas de IA, ya sea al introducir *prompts* o subir archivos, es posible que se compartan, a menudo sin ser plenamente conscientes, datos sensibles como:

- Nombre, correo electrónico, ubicación o identificadores personales.
- Información sobre salud, emociones, ideas o creencias.
- Contenido académico o creativo que puede tener valor intelectual.

Es fundamental que comprendamos qué tipos de datos estamos revelando y qué consecuencias puede tener hacerlo en plataformas no controladas por el entorno educativo. Especialmente si estamos utilizando algún servicio o página web que no sea de los principales, que sea más o menos “desconocido”, en ese caso ni aunque lo ponga en los términos de uso nos podríamos fiar.

Realmente este no es un aspecto diferente a lo que ya ocurría con cualquier plataforma web hasta ahora, como por ejemplo páginas que pasan de pdf a word o que desbloquean un pdf o hacen un gráfico a partir de un excel.

Entre los principales riesgos se encuentran:

- **Fugas de información, gran parte de carácter sensible:** especialmente delicado es el caso de ficheros con notas, comentarios personales o información de salud. Recomendamos no subir nunca este tipo de ficheros.
- **Uso indebido de los datos:** los datos pueden ser utilizados para entrenar modelos o cualquier otro uso sin consentimiento explícito.
- **Reidentificación:** siempre es mejor utilizar datos anonimizados, aunque hay que tener cuidado ya que en ocasiones combinaciones de información pueden volver a identificar a una persona.

Insistimos en que estos riesgos se incrementan cuando se utilizan servicios comerciales sin políticas claras de privacidad.

En Europa, el **Reglamento General de Protección de Datos (RGPD)** establece principios clave como derecho a la información clara sobre el uso de datos, derecho a la eliminación ("derecho al olvido") y necesidad de consentimiento explícito para tratar datos personales.

Buenas prácticas para una interacción segura con la IA

Se proponen las siguientes estrategias para docentes y estudiantes:

- **Evitar introducir datos personales** en plataformas públicas o comerciales de IA.
- **Usar versiones institucionales** de IA, cuando estén disponibles, con mayor control sobre los datos.
- **Leer y entender las políticas de privacidad** antes de utilizar una herramienta (aunque esto sabemos que no es muy viable).
- **Utilizar alias o datos ficticios** en tareas que impliquen el uso de IA generativa.
- **Revisar las configuraciones de privacidad** de las herramientas empleadas y promover un uso consciente.

Qué transmitir al estudiante:

Es fundamental concienciarles sobre el valor de sus datos. A no introducir información sensible en los *prompts* y a ser críticos con las aplicaciones "gratuitas" que, en realidad, se financian con su información.

Más allá de protegerse, es clave que el alumnado desarrolle una **actitud crítica** frente al tratamiento de sus datos:

- ¿Quién tiene acceso a lo que generamos con IA?
- ¿Qué información dejamos en internet al usar estas herramientas?
- ¿Podemos exigir transparencia y control sobre los sistemas que usamos?

Una idea simple y directa podría ser proponer al estudiante que ponga el siguiente *prompt* en ChatGPT o el servicio de IA que más utilice:

“Dime datos que sepas sobre mí. Datos concretos. Se muy específico.”

Esta competencia no solo les servirá en su vida académica, sino en su participación futura como ciudadanos digitales activos y responsables.

7. Sostenibilidad y Conciencia Medioambiental en el Uso de la IA

El entrenamiento y funcionamiento de los grandes modelos de IA tienen un considerable impacto medioambiental, asociado a un alto consumo de energía y agua para la refrigeración de los centros de datos. Integrar esta reflexión en nuestras clases ayuda a formar profesionales más conscientes del coste oculto de la tecnología. Es importante que los estudiantes conozcan el impacto medioambiental de los modelos y aprendan a hacer un uso responsable de los mismos.

Qué transmitir al estudiante:

Fomenta una cultura de uso racional de la tecnología. Deben entender que cada consulta a una IA tiene una huella ecológica. Esto no significa no usarla, sino hacerlo

de manera deliberada y eficiente, evitando un uso indiscriminado o innecesario. Investigadores de la UPM participan en un proyecto para concienciar sobre el impacto medioambiental de la IA acompañado de guías y juegos con los que las personas pueden aprender a hacer un uso responsable de la misma⁶.

8. Evaluación de tareas con IA

Es importante destacar algunos usos de la IA regulados por la IA Act⁷ en educación. Concretamente se catalogan como sistemas de alto riesgo en educación y formación profesional a:

- “a) Sistemas de IA destinados a ser utilizados para determinar el acceso o la admisión de personas físicas a centros educativos y de formación profesional a todos los niveles o para distribuir a las personas físicas entre dichos centros
- b) Sistemas de IA destinados a ser utilizados para evaluar los resultados del aprendizaje, también cuando dichos resultados se utilicen para orientar el proceso de aprendizaje de las personas físicas en centros educativos y de formación profesional a todos los niveles
- c) Sistemas de IA destinados a ser utilizados para evaluar el nivel de educación adecuado que recibirá una persona o al que podrá acceder, en el contexto de los centros educativos y de formación profesional o dentro de estos a todos los niveles
- d) Sistemas de IA destinados a ser utilizados para el seguimiento y la detección de comportamientos prohibidos por parte de los estudiantes durante los exámenes en el contexto de los centros educativos y de formación profesional o dentro de estos a todos los niveles”

Los sistemas de alto riesgo son, entre los escenarios de la IA permitidos, los más restrictivos. Para su uso es necesario que estas herramientas de IA cumplan una serie de requisitos entre los que se encuentran:

- Sistemas adecuados de evaluación y mitigación de riesgos
- Alta calidad de los conjuntos de datos que alimentan el sistema para minimizar riesgos de resultados discriminatorios
- Registro de la actividad para garantizar la trazabilidad de los resultados
- Documentación detallada que proporcione toda la información necesaria sobre el sistema y su propósito, para que las autoridades puedan evaluar su cumplimiento
- Información clara y adecuada para el responsable de su implementación
- Medidas apropiadas de supervisión humana
- Alto nivel de robustez, ciberseguridad y precisión

⁶ <https://sostenibilidadgenerativa.cotec.es/> ; [Generative AI's environmental costs are soaring — and mostly secret](#)

⁷ <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

9. Conclusión: El Nuevo Rol del Docente como Facilitador Crítico

Nuestro papel como docentes está evolucionando. Ya no se trata de ser expertos técnicos en cada nueva herramienta, sino de convertirnos en facilitadores críticos. Nuestra misión es dotar al alumnado de las competencias necesarias para navegar el ecosistema de la IA de forma autónoma, ética y productiva.

La metáfora más poderosa es la de la IA como un "copiloto". Un copiloto puede sugerir rutas, avisar de peligros y gestionar tareas, pero el piloto es quien mantiene el control, toma las decisiones críticas y asume la responsabilidad final del viaje. La tecnología potencia, pero el juicio, la creatividad y la ética son y seguirán siendo profundamente humanos.

Por supuesto, no nos referimos a que haya que incluir todas estas tareas y recomendaciones en todas las asignaturas, pero sí conviene que el docente las conozca y quizá en un caso lo puede comentar en clase al contar un ejemplo relacionado, en otro lo puede introducir en un enunciado de una práctica, etc. Si esto es algo que va calando entre el profesorado, entre todos seremos capaces de “Formar en un uso ético y responsable de la IA” a los alumnos como decíamos al principio de este documento.

Una recomendación final sería leer el capítulo 8 del informe “Inteligencia Artificial en la Educación Superior 2025: Transformaciones Docentes, Herramientas y Buenas Prácticas”⁸ que hemos publicado. Una lectura que puede ser interesante es también la “Guía sobre el uso de la inteligencia artificial en el ámbito educativo”⁹ del INTEF desarrollan bastante el tema de la ética, aunque desde un punto de vista algo diferente y proponen algunas tareas adicionales para el aula.

⁸ <https://innovacioneducativa.upm.es/inteligencia-artificial-educacion-superior>

⁹ <https://intef.es/Noticias/guia-sobre-el-uso-de-la-inteligencia-artificial-en-el-ambito-educativo/>