

AI@MEDIATORS

GUÍA DE PRÁCTICAS ÉTICAS PARA EL USO DE LA IA EN LA EDUCACIÓN EN EUROPA



DISEÑO GRÁFICO REALIZADO POR
IDEC S.A.

[AI.MEDIATORS](https://www.aimediatorsproject.eu)



[AI@MEDIATORS](https://www.aimediatorsproject.eu)



[AI.MEDIATORS](https://www.aimediatorsproject.eu)



[WWW.AIMEDIATORSPROJECT.EU](https://www.aimediatorsproject.eu)





ABOUT OUR PROJECT

La inteligencia artificial (ia) está transformando diversos aspectos de nuestras vidas, y su impacto en la educación es particularmente significativo, sobre todo para los adolescentes, que son vulnerables a sus riesgos y beneficios. La ia ofrece ventajas como el acceso universal a la información, los sistemas de tutoría inteligentes, los facilitadores virtuales y la automatización de las tareas administrativas. Estas ventajas pueden mejorar la educación, pero es fundamental evaluar y supervisar cuidadosamente los cambios que introduce la ia para evitar consecuencias negativas en los comportamientos y la forma de pensar de la población.

Esta guía se centra en proporcionar a los profesores, alumnos y educadores las herramientas necesarias para comprender las principales cuestiones relacionadas con la ia. Destaca la importancia de desarrollar y utilizar la ia de manera ética y responsable, en conformidad con los valores democráticos y los derechos humanos, así como de tomar medidas para maximizar sus beneficios y minimizar sus riesgos.

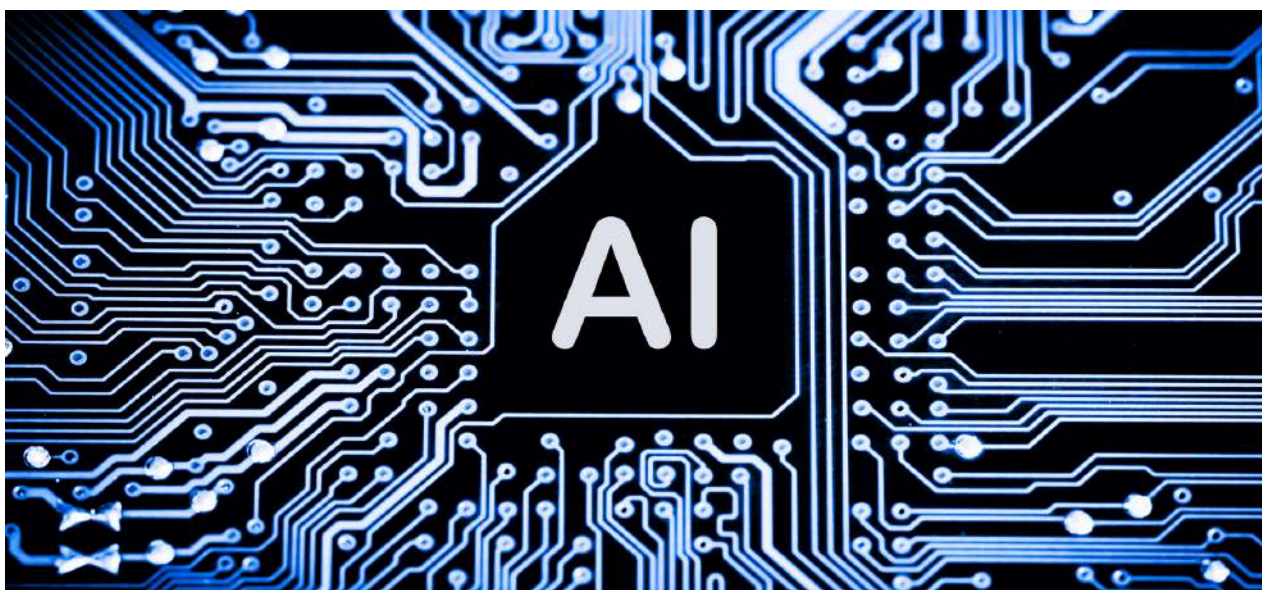
A lo largo de las últimas dos décadas, la interacción social entre los jóvenes se ha visto influida por el uso de internet, en particular a través de las redes sociales, debido a la profunda transformación de las telecomunicaciones en el siglo xxi. El acceso a internet y a plataformas como facebook, instagram y tiktok ha creado nuevas formas de socialización, que pueden reproducir comportamientos violentos tradicionales, como el acoso escolar y el sexismo, o generar nuevos comportamientos nocivos, como la adicción a internet. Estas dinámicas están mediadas por nuevas formas de comunicación y por la estructura de las redes sociales. El ciberacoso en las redes sociales se define como un comportamiento deliberado y hostil, en el que una persona o un grupo utiliza las redes sociales para comunicar repetidamente contenidos agresivos con la intención de causar daño o malestar a una persona objeto del acoso. Este tipo de acoso suele estar relacionado con un uso problemático de internet. Las víctimas del acoso en internet pueden presentar síntomas como depresión, insomnio, ansiedad general y social, baja autoestima, sentimientos de soledad, insatisfacción con la vida, absentismo escolar, falta de concentración, sentimientos de vergüenza y culpa, ideas suicidas y comportamientos antisociales, entre otros.

Los autores de acoso cibernético también presentan niveles significativamente más altos de insatisfacción con la vida, baja autoestima, depresión, paranoia, ansiedad, sensibilidad interpersonal, hostilidad, ideas suicidas, consumo de drogas y actividad delictiva, en comparación con aquellos que no participan en el ciberacoso. Los comportamientos de acoso cibernético pueden manifestarse de muchas formas diferentes. De manera directa, adopta la forma de insultos, amenazas, chantaje, acoso sexual o lynchamiento público cuando implica a un gran número de personas. De forma indirecta, el acoso cibernético puede incluir la difusión de rumores, la divulgación de contenido sensible de terceros sin su consentimiento o la privación de la socialización de una persona (marginación), entre otras formas.

En este contexto, se identifican tres figuras principales en el ciberacoso: la víctima, que es la persona a la que se dirige el comportamiento violento; el agresor, que es quien lleva a cabo el comportamiento violento; y los observadores, que son aquellos que son testigos del ciberacoso sin participar directamente, actuando como espectadores. Estos últimos pueden dividirse a su vez en participantes (que ayudan al agresor o a la víctima) o no participantes. Es fundamental tener en cuenta que estos papeles no son estáticos, ya que la mayoría de las personas implicadas en situaciones de ciberacoso pueden desempeñar los tres papeles en momentos diferentes. Independientemente del papel que se asuma, es posible abordar el ciberacoso desde diferentes perspectivas con el fin de prevenirlo y mitigarlo.

En el contexto del acoso cibernético, la ia puede ser una herramienta poderosa para combatir este problema. Los algoritmos de ia pueden analizar grandes volúmenes de datos en tiempo real para identificar patrones de comportamiento abusivo, lo que permite intervenciones tempranas. Además, los sistemas de ia pueden ofrecer apoyo a las víctimas a través de chatbots que proporcionan orientación y recursos, así como detectar señales de malestar emocional en los mensajes y publicaciones. Estas capacidades pueden ayudar a crear un entorno escolar más seguro y protector, contribuyendo a la lucha contra el ciberacoso y mejorando el bienestar de los alumnos.

A continuación se presenta un conjunto de principios con preguntas orientativas que pueden ayudarte a identificar las deficiencias y los puntos fuertes en caso de que necesites utilizar la inteligencia artificial en tu centro educativo. Recuerde que estas preguntas **no son definitivas**, sino una ayuda para una buena práctica ética que debe adaptar a las circunstancias de su centro escolar.



PREVENCIÓN DE DAÑOS

El desarrollo y el uso de la inteligencia artificial para combatir el ciberacoso no debe causar daño (psicológico, físico, social, económico, etc.) A los alumnos, padres, profesores o cualquier otra persona que utilice el sistema. Por eso, deben eliminarse o minimizarse los posibles riesgos de su uso

- ¿han debatido los profesores qué daños a corto plazo puede causar el uso de la IA? ¿y a largo plazo?
- ¿interfiere su uso en alguna actividad importante para el desarrollo del alumno?
- ¿puede poner en peligro la salud física o mental de la persona y sus relaciones interpersonales?
- ¿sabemos en qué grado puede la aplicación cometer errores o producir alucinaciones?

¿el uso del sistema de ia hace que los adolescentes se sientan desmotivados en su aprendizaje, en el uso de herramientas digitales o menos implicados en las relaciones sociales?

Ejemplo práctico: para combatir el ciberacoso, se pueden utilizar programas que emplean chats grupales en los que los alumnos intercambian insultos. Los profesores deben tener esto en cuenta de antemano para evaluar cómo abordar la situación. También debería observarse, a corto y largo plazo, cómo modifica el comportamiento de las personas, ya que existen aplicaciones móviles, con y sin ia, que generan adicción y, por lo tanto, no serían recomendables para la lucha contra el acoso digital.

PRINCIPIO DEL BIENESTAR

La aplicación de una herramienta de ia no solo debe significar la ausencia de daño, sino también la mejora de la condición humana, ya sea desarrollando el aspecto físico, mental, social y/o emocional. La lucha contra el ciberacoso en sí misma justifica este principio, pero debemos ser conscientes de si se están alcanzando los objetivos propuestos:

- ¿se centra la aplicación de ia en algún aspecto del bienestar de los alumnos? ¿lo está consiguiendo?
- ¿mejora la aplicación las relaciones interpersonales del alumno?
- En el caso de una evaluación de costes y beneficios, ¿merece la pena utilizar la aplicación para el bienestar del alumno, y ¿supone el coste un perjuicio significativo para la integridad de la persona?
- En cuanto al tiempo y el esfuerzo que supone utilizar el programa de ia, ¿es lo más eficiente posible en cuanto a sus beneficios para el individuo y para la sociedad?
- ¿fomenta el sistema de ia una comunicación clara y no violenta entre los alumnos y con los adultos (profesores, personal, padres)?

Ejemplo práctico: el centro educativo podría implementar un chatbot basado en ia que ofrezca apoyo emocional y estrategias de afrontamiento a los alumnos. En colaboración con la ayuda psicológica que prestan los psicólogos del centro.



RESPECTO POR LA AUTONOMÍA HUMANA

Se refiere a mantener la capacidad de las personas para actuar de forma incondicional en función de sus propias acciones, sin presiones externas, y para ser responsables de su comportamiento.

- ¿tiene la aplicación algún elemento adictivo?
- ¿se ha obtenido el consentimiento informado de las personas implicadas?
- ¿es obligatorio para el alumno utilizar la aplicación de IA? ¿existe alguna alternativa?
- ¿sienten los alumnos que pueden interactuar libremente con la aplicación o se sienten cohibidos (por ejemplo, ¿se sienten observados)?
- ¿disponen los alumnos de una persona de contacto a la que acudir cuando tienen dudas sobre el uso de la aplicación de IA?

Ejemplo práctico: antes de implementar un sistema de IA para combatir el ciberacoso, el colegio debería obtener el consentimiento informado de los alumnos y los padres, explicándoles su finalidad y su uso. La ia debe estar diseñada para que no resulte intrusiva, garantizando que los alumnos puedan darse de baja si se sienten incómodos, respetando así su autonomía.

NO DISCRIMINACIÓN

It Es una obligación moral garantizar que los sistemas de ia destinados a combatir el ciberacoso no discriminen a los menores creando una situación de desventaja por motivos (entre otros) de: raza, nacionalidad, género, orientación sexual, religión o convicciones políticas. También debe prestarse especial atención a las necesidades especiales que puedan tener algunos alumnos para poder participar.

- ¿pueden utilizarlo todos los alumnos sin prácticamente ninguna barrera tecnológica?
- ¿se puede adaptar para estudiantes con discapacidades o necesidades especiales?
- ¿son equitativos los beneficios del uso de la IA para todos los estudiantes que la utilizan?
- ¿se ha entrenado la ia de forma imparcial, sin riesgo de que discrimine a algún grupo?
- ¿existen mecanismos para detectar posibles sesgos y discriminación en el uso o los resultados de la IA?

EJEMPLO PRÁCTICO: LA APLICACIÓN DEBE TRADUCIRSE AL MAYOR NÚMERO POSIBLE DE IDIOMAS PARA GARANTIZAR LA INCLUSIÓN Y EL ACCESO A LA MISMA PARA TODOS LOS ESTUDIANTES, SIN DISTINCIÓN ALGUNA



TRANSPARENCIA

Greater transparency contributes to more peaceful, just, democratic and inclusive societies by enabling public scrutiny and reducing corruption and discrimination. It is based on the fact that users should be informed in a clear and understandable way about the use of AI and all that it entails.

- Does the AI application easily show that it uses AI and how it works?
- Do students, parents and teachers know what data is being collected and for what purpose?
- Does everyone involved know what the risks and limitations of using AI are?
- What mechanisms does the school have in place to know if the AI system is working properly?
- Do teachers and/or student representatives know why the AI is taking these results?

Ejemplo práctico: antes de utilizar la IA, se informa a los estudiantes sobre cómo funciona, sus ventajas, sus riesgos y sus límites. Esta demostración de honestidad ya constituye una forma de educar a los alumnos en valores, lo que contribuye a generar confianza en la herramienta de IA. Cuanta más confianza haya en un programa de IA de este tipo, más se utilizará y más eficaz será a la hora de combatir el ciberacoso

RESPONSABILIDAD PROACTIVA

Esto se refiere al compromiso ético que se exige a los agentes sociales implicados en el desarrollo, distribución, e implementación de estas aplicaciones. Estos agentes sociales deben anticiparse para reducir o eliminar cualquier riesgo potencial que estas tecnologías puedan representar para los adolescentes, ya que una aplicación de IA por sí sola no puede comprender los sesgos o las diferencias de poder que pueda generar, y podría favorecer accidentalmente situaciones de ciberacoso.

- ¿existe un equipo técnico encargado del almacenamiento, la seguridad y el tratamiento de los datos de los alumnos?
- ¿se ha elaborado un formulario de consentimiento informado para padres y alumnos en el que se expliquen los beneficios, los riesgos y las limitaciones de la IA?
- si se utiliza software externo, ¿se conoce la fiabilidad de dicho software?
- ¿se celebran reuniones periódicas para evaluar si la aplicación cumple los requisitos mínimos o si se han detectado probl
- ¿existe una persona legalmente responsable en caso de que se produzca algún incidente?

Ejemplo práctico: Las escuelas organizan reuniones mensuales para analizar el impacto de la herramienta de IA en el comportamiento de los alumnos y en la dinámica del aula. En ellas se debaten los problemas que puedan surgir y se realizan los ajustes necesarios para garantizar el uso ético y eficaz de la herramienta. Este enfoque proactivo contribuye a mantener un entorno de aprendizaje seguro y propicio, al tiempo que se maximizan los beneficios de la IA en la educación.

EXPLICABILIDAD



PARTICIPACIÓN ACTIVA

Users of AI applications should not be seen as mere passive agents, but should be involved (or represented) in the design, development and use in order to avoid problems with the new scenarios posed by AI systems.

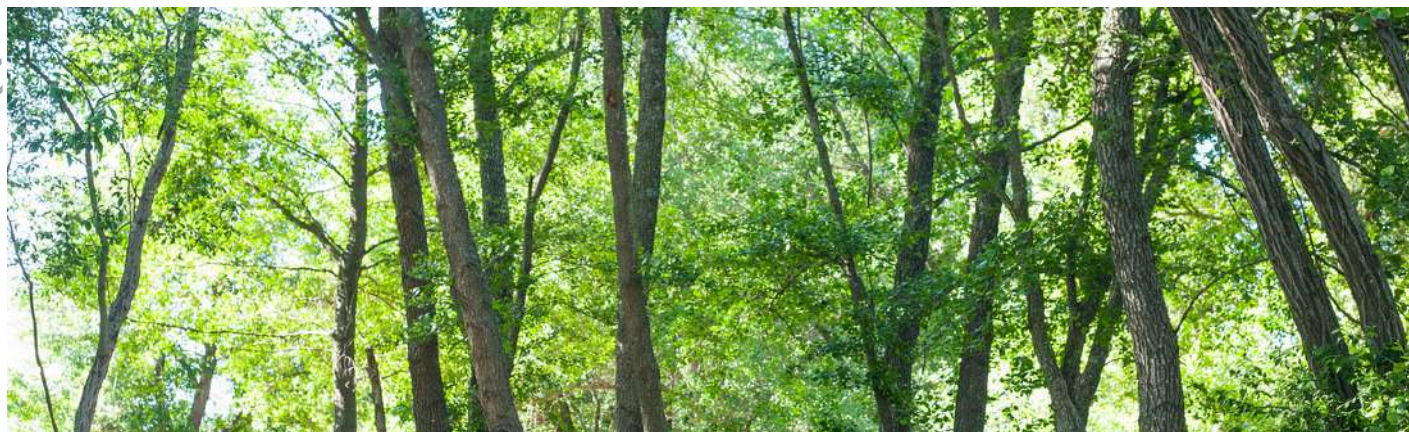
- Has the app been designed/selected by students for their needs?
- Have the students been informed or trained to get the most out of the AI?
- Do teachers know how the app works and what their role is in it?
- Do parents know how the app works? Do they have any role in it?
- Has the app been tested and recommended by an institution representing students' interests?

Practical example: There should be communication between teachers and students to know how cyberbullying happens, its consequences and to know how to choose the best AI application to fight cyberbullying.

is essential for meaningful informed consent. Users should receive clear, accessible information about how the AI system works, what data it collects, and its potential benefits and limitations. In cyberbullying prevention, students need to understand how the AI application supports them and why its recommendations can be trusted.

- Have teachers or technical staff been trained in understanding how the AI application works?
- Can the AI system's decisions and processes be easily understood by non-experts?
- Are the benefits, risks and limitations explained in a simple and understandable way?
- Does the AI system have a 'black box' algorithm that does not allow understanding of your decision?
- Will there be any kind of online support system that will be active 24/7 to assist users with any queries, problems or technical issues?

Ejemplo práctico: Las escuelas organizan reuniones mensuales para analizar el impacto de la herramienta de IA en el comportamiento de los alumnos y en la dinámica del aula. En ellas se debaten los posibles problemas que surjan y se realizan los ajustes necesarios para garantizar que la herramienta se utilice de forma ética y eficaz. Este enfoque proactivo contribuye a mantener un entorno de aprendizaje seguro y propicio, a la vez que maximiza los beneficios de la IA en la educación.



PRIVACIDAD

Es la protección de los datos personales de las personas, en especial de los niños y adolescentes, frente a la recopilación, el uso y el tratamiento no autorizados. Las herramientas de IA no solo deben justificar que los datos que recopilan son legítimos y que se destinan a fines específicos consentidos por el niño y su tutor legal, sino que también deben garantizar que no recopilan datos que no sean los necesarios para combatir el ciberacoso.

- ¿cumple la IA con la normativa de protección de datos y otras disposiciones similares?
 - ¿existen políticas claras sobre el almacenamiento, el acceso y el intercambio de datos?
 - ¿cuenta la institución con profesionales técnicos que puedan asumir la responsabilidad de prevenir infracciones de seguridad?
 - ¿saben los padres, los profesores y los alumnos qué datos se almacenan y cuáles no?
- ¿es posible no recopilar los datos personales de los alumnos o anonimizarlos?

Ejemplo práctico: los datos se almacenan en un servidor protegido por la dirección del centro, que se mantiene actualizado contra posibles ciberataques. La filtración de información confidencial es peligrosa, ya que puede dar lugar a discriminación y acoso en las redes sociales, fomentando así el ciberacoso.

CUIDADO DEL MEDIO AMBIENTE

El uso de la inteligencia artificial implica un consumo energético considerable, por lo que los desarrolladores y distribuidores deben actuar de forma responsable y utilizar la IA de la manera más sostenible y respetuosa con el medio ambiente, minimizando su impacto medioambiental y maximizando su rentabilidad.

- ¿conoce la institución cuál es el coste energético de la implantación de este sistema de IA, y puede este compensarse con ahorros en otros ámbitos?
- ¿existe alguna forma de proporcionar a los equipos que utilizan IA una vida útil más larga o más eficiente?
- ¿contribuyen las consecuencias del uso de la IA a mejorar la sociedad?
- ¿demuestra la empresa proveedora de la herramienta que esta y su uso son respetuosos con el medio ambiente?

Ejemplo práctico: los equipos que utilizan sistemas de IA emplean herramientas de ahorro energético. Fomentar valores más empáticos, solidarios y colectivistas mejora no solo la lucha contra el ciberacoso, sino también la concienciación medioambiental.



Co-funded by
the European Union

GRACIAS

Financiado por la Unión Europea. Las opiniones y puntos de vista expresados en este sitio web son responsabilidad exclusiva de su(s) autor(es) y no reflejan necesariamente los de la Unión Europea ni los del Servicio Español para la Internacionalización de la Educación (SEPIE). Ni la Unión Europea ni la agencia nacional del SEPIE pueden ser consideradas responsables de las mismas.

N.º de proyecto: 2023-1-ES01-KA220-SCH-000160237.

SOCIOS



Universidad de Alicante



AI@MEDIATORS