1.1 Potencial y Riesgos

> *Si solo recordáis una frase de esta página:* la IA ha llegado a las escuelas como un medicamento que viene sin prospecto. Como docente, no sabéis cuándo usar la IA, os faltan las instrucciones de dosificación y no hay información sobre los efectos secundarios. Este capítulo sirve como el prospecto del Playbook.

¿Debería la IA venir con un prospecto, como un medicamento?

¿Conocéis ese papelito que va dentro de cada caja de medicamentos? Ese que probablemente descartáis sin mirarlo siquiera. Puede parecer aburrido, pero ese pequeño papel es un triunfo silencioso del siglo pasado. Ese prospecto detalla cuatro elementos esenciales: qué trata el medicamento, cuánto tomar, las sorpresas que puede traer y quién debe evitarlo.

Detrás de cada línea de ese prospecto hay toda una maquinaria: ensayos clínicos, reguladores y una responsabilidad de revelar los errores. Nadie cree que el prospecto en sí cure. Su verdadero poder es hacer que el medicamento sea seguro para cualquiera. Esa es la magia que nos falta cuando dejamos que la IA enseñe sin orientación.
Las herramientas de IA que ahora se instalan silenciosamente en vuestra escuela heredan un problema más antiguo que la propia IA: la tecnología educativa rara vez ha venido con instrucciones. En el Reino Unido, una revisión encontró que solo 7 de cada 100 empresas de tecnología educativa habían sometido sus productos a un ensayo aleatorizado, y solo 12 de cada 100 habían buscado una certificación externa. Las escuelas rara vez piden más. En una encuesta realizada en 17 estados estadounidenses, solo 11 de cada 100 docentes y miembros de equipos directivos habían solicitado pruebas revisadas por pares antes de adoptar una nueva herramienta. El resultado, en muchos casos, es dinero desperdiciado. En EE. UU., por término medio, dos tercios de las licencias de software educativo quedaron sin usar, y 98 de cada 100 nunca se utilizaron de forma intensiva. El EdTech Genome Project, también en EE. UU., analizó unas 7000 herramientas educativas con un valor de 13 000 millones de dólares y descubrió que 85 de cada 100 o bien no eran adecuadas o bien se habían implantado de forma incorrecta.
Este capítulo puede parecer la sección árida y necesaria, pero revela las historias ocultas detrás de las normas. Cuenta lo que realmente ocurrió, a quién afectó, y arroja luz sobre los efectos secundarios, incluso aquellos que llegaron silenciosamente a niños que nunca dieron su consentimiento.

Potencial de la IA

La mayoría de los estudios sobre el potencial de la IA se centran en estudiantes universitarios. El alumnado escolar ha aparecido hasta ahora relativamente poco en estos estudios. Y la mayoría de los estudios se refieren a intervenciones a corto plazo, que duran desde unas horas hasta unas semanas. En consecuencia, el llamado efecto novedad puede haber influido positivamente en los resultados. Una amplia revisión de 68 experimentos encontró exactamente este patrón: cuanto más corto era el estudio y menor el número de participantes, mayor era el beneficio de la IA del que se informaba.

(1) Dejad que la IA explique un tema nuevo y reservad la práctica para vuestra clase

Imaginad esto. Mañana, cuando un alumno o alumna se pierda un tema o necesite que le expliquen de nuevo ese concepto complicado, dejad que un tutor de IA se encargue de la primera explicación. Reservad la práctica para vuestra clase, donde cobra vida el aprendizaje real.

Por qué es importante. En un ensayo bien controlado, un tutor de IA enseñó contenidos nuevos al menos tan bien como una buena clase. En un estudio de Harvard, 194 estudiantes de física aprendieron dos temas. Para uno de los temas, usaron un tutor de IA desarrollado para este fin. Para el otro tema, aprendieron en una clase de aprendizaje activo bien llevada, con debate y resolución de problemas por parejas, no una clase magistral. Cada estudiante completó ambos enfoques, de modo que nadie puede alegar que el grupo más fuerte recibió el mejor tratamiento. En una prueba realizada inmediatamente después, los estudiantes enseñados por la IA obtuvieron resultados significativamente más altos. Los investigadores registraron un tiempo mediano de 49 minutos para completar la tarea, frente a los 60 minutos asignados para la clase en el horario. Los autores son cautelosos al interpretar estos resultados: el material era nuevo para los estudiantes, y gran parte de la tarea consistía simplemente en explicarlo bien. Se niegan expresamente a extender su resultado a tareas que exigen relacionar varias ideas. Y se trataba de estudiantes universitarios, no de alumnado de 2.º de la ESO.

Una amplia revisión que combinó los resultados de 228 estudios apunta en la misma dirección, con un límite importante. La IA generativa produjo mayores avances en conocer y comprender que formas más antiguas de IA, como los sistemas de tutoría inteligente y el software de práctica adaptativa. Pero este fue el único de los cuatro resultados medidos por la revisión en el que la IA generativa obtuvo ventaja.

Para profundizar. Kestin, G., Miller, K., Klales, A., Milbourne, T., y Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports, 15, artículo 17458. Acceso libre, DOI 10.1038/s41598-025-97652-6. El artículo describe cómo se diseñó y controló el tutor. Esto resulta útil para quien desee desarrollar su propio tutor de IA.

(2) Aprovechar mejor el tiempo de planificación de las clases

Imaginad esto. Ahorrad parte del tiempo que necesitáis para planificar las clases utilizando la IA para redactar un primer borrador de vuestra planificación y luego revisarlo a fondo.

Por qué es importante. Nadie discute que la carga de trabajo del profesorado es alta. Por ello, uno de los posibles beneficios de la IA es aprovechar mejor las horas de trabajo. Un ensayo realizado en Inglaterra investigó si el uso específico de la IA para planificar clases puede reducir la carga de trabajo docente. Los investigadores asignaron aleatoriamente a 259 profesores de ciencias de 68 escuelas a un grupo que usaba ChatGPT junto con una guía en línea sobre cómo usarlo bien, o a un grupo al que se pidió que no usara IA generativa en absoluto. **El profesorado del grupo de ChatGPT dedicó unos 56 minutos semanales a planificar las clases de ciencias de los dos primeros cursos de secundaria (alumnado de 11 a 13 años), frente a unos 81 minutos en el grupo de comparación.** Un panel de especialistas en la materia evaluó la calidad de las planificaciones, sin saber si estas se habían creado con o sin ayuda de la IA. Los especialistas no encontraron pruebas de que la calidad de las planificaciones difiriera. Sin embargo, deben destacarse dos puntos. (1) **El profesorado utilizó la herramienta cada vez menos a medida que pasaban las semanas.** Los evaluadores lo interpretan como el paso de una fase de aprendizaje a la rutina; pero también puede significar que el ahorro de tiempo no se mantiene por sí solo sin apoyo. (2) El ensayo midió el tiempo que el profesorado dedicaba a planificar las clases, no los resultados de aprendizaje del alumnado.

«Todo se reduce al tiempo. Incluso el profesorado que ve el valor de la IA rara vez tiene tiempo para aprender a usarla».

Para profundizar. Roy, P., Poet, H., Staunton, R., Aston, K., y Thomas, D. (2024, 12 de diciembre). ChatGPT in lesson preparation: A Teacher Choices Trial. NFER, encargado por la Education Endowment Foundation y la Hg Foundation. Gratuito en educationendowmentfoundation.org.uk.

(3) Ofrecer mejor apoyo al alumnado que no se beneficia plenamente de las clases estándar

Imaginad esto. Vuestra escuela decide cómo quiere gastar sus fondos. Quiere comprar un chatbot de IA. ¿Qué probabilidad creéis que hay de que los proveedores del chatbot os aporten pruebas de los beneficios del bot para las distintas materias?

Por qué es importante. El proveedor puede remitirse a la investigación sobre alumnado con discapacidad, uno de los campos de la IA en educación estudiados desde hace más tiempo y que con facilidad se pasa por alto al asignar fondos. En un análisis de 29 estudios con 41 grupos de alumnado con discapacidad, los investigadores encontraron beneficios moderados en una amplia variedad de contextos. Pero leed la letra pequeña. En siete de cada diez de estos estudios, «IA» significaba un robot, normalmente humanoide, a menudo manejado por un investigador fuera de la vista del niño o la niña. Solo uno de cada cinco estudios utilizó software. Por tanto, esto no es prueba de que un chatbot vaya a ayudar a un alumno autista en vuestra clase. Tampoco pudieron

los revisores determinar qué hace que estas intervenciones sean eficaces. A partir de su investigación, no pudieron identificar por qué los estudios que funcionaron diferían de los que no lo hicieron.

Para profundizar. Zhang, L., Carter, R. A., Jr., Liu, Y., y Peng, P. (2026). Let's CHAT about artificial intelligence for students with disabilities: A systematic literature review and meta-analysis. *Review of Educational Research, 96*(1), 215–257. Gratuito, DOI 10.3102/00346543241293424. La verdadera joya es la tabla de estudios incluidos: revela de inmediato lo poco que este campo se centra realmente en las herramientas que hoy se comercializan para las escuelas.

(4) Enseñar al alumnado cómo falla la IA

Imaginad esto. Pedid a vuestra clase que cree diez imágenes de «un médico» y diez de «una enfermera». Contad quién aparece en cada conjunto y luego iniciad un debate: ¿de dónde vienen estos patrones y quién tomó esas decisiones?



Fig. 1: «Prompt de un solo intento: a doctor, a nurse (un médico, una enfermera)» (imagen generada con IA, NanoBanana 2)

Por qué es importante. La versión clásica de esta lección puede parecer anticuada, y precisamente por eso merece atención. Durante años, la demostración era predecible: pedir un médico y aparecía un hombre. En una prueba de la organización sin ánimo de lucro AlgorithmWatch,

Midjourney representaba casi siempre a los médicos como hombres blancos, mientras que Adobe Firefly hacía intentos visibles, aunque imperfectos, de diversificar. Un estudio revisado por pares de 2400 imágenes de cirujanos generadas con IA encontró la misma división entre productos: dos de tres generadores populares casi no mostraban mujeres, mientras que el tercero se acercaba a la proporción real de mujeres en la cirugía.

Probad hoy el mismo prompt y puede que aparezca una médica o un enfermero (Fig. 1). Pero no os dejéis engañar pensando que el sesgo ha desaparecido. Un análisis de 2026 de 1344 imágenes de tres generadores populares reveló lo contrario: al pedir una «persona competente», un sistema no mostró ninguna mujer.

Lo que ha cambiado no es el sesgo, sino la capa en la que vive, ¡y esa es la lección importante! Detrás de todas estas imágenes hay decisiones humanas: con qué se entrenó el sistema y qué decidieron después sus creadores que debía mostrar. Esta segunda decisión es ajustable, invisible y la toma una empresa.

La primera capa de decisiones humanas es lo que el sistema ha absorbido. Millones de imágenes, creadas y etiquetadas por personas, cada una reflejando el mundo tal como ya era. Esta es la capa que revela la lección clásica sobre el sesgo, y el alumnado la comprende con facilidad: la máquina se convirtió en un reflejo de lo que se le dio de comer.
La segunda capa está determinada por lo que una empresa decide mostraros a continuación, aunque esta influencia no deja rastro alguno en la propia imagen. Aquí ocurren tres cosas, cada una descrita por sus creadores en sus propias palabras.

- **Vuestras palabras se reescriben antes de que el modelo las reciba.** Según la propia documentación para desarrolladores de OpenAI: «usamos GPT-4 para optimizar todos vuestros prompts antes de pasarlos a DALL-E» — y «esta función no se puede desactivar por el momento». Lo que escribió vuestro alumno no es lo que se le pidió a la máquina.
- **El aspecto se completa cuando el prompt lo deja abierto.** OpenAI, describiendo su técnica de diversidad: «Esta técnica se aplica a nivel de sistema cuando se le da a DALL·E un prompt que describe a una persona sin especificar raza o género, como “bombero”». Después de aplicarla, los usuarios tenían doce veces más probabilidades de decir que las imágenes mostraban personas de orígenes diversos. «Un médico» es exactamente ese tipo de prompt.
- **Y ese ajuste puede fallar.** Google, tras pausar la generación de imágenes de Gemini en 2024: «Nuestro ajuste para garantizar que Gemini mostrara una variedad de personas no tuvo en cuenta los casos en los que claramente no debía mostrarse esa variedad».

Así que cuando un alumno o alumna escribe «un médico» y aparece una mujer, la imagen no ofrece ninguna pista sobre qué capa la hizo aparecer. Esa incertidumbre no es un defecto de la lección. Es la lección misma.
La vieja narrativa afirmaba que la máquina simplemente absorbía los prejuicios de la sociedad. La nueva narrativa va más al fondo y perdura más: ¡una empresa tomó una decisión! No podéis detectar esa decisión en el resultado, y puede cambiar un martes cualquiera, en silencio, sin previo aviso. Esa verdad perdura a través de cada nuevo modelo. La lección del estereotipo se desvanece.

Esta es la lección que merece la atención del alumnado, y perdurará más allá de la próxima actualización del modelo: la imagen no es un simple reflejo de la realidad, y no ocurrió por casualidad. Alguien tomó una decisión.

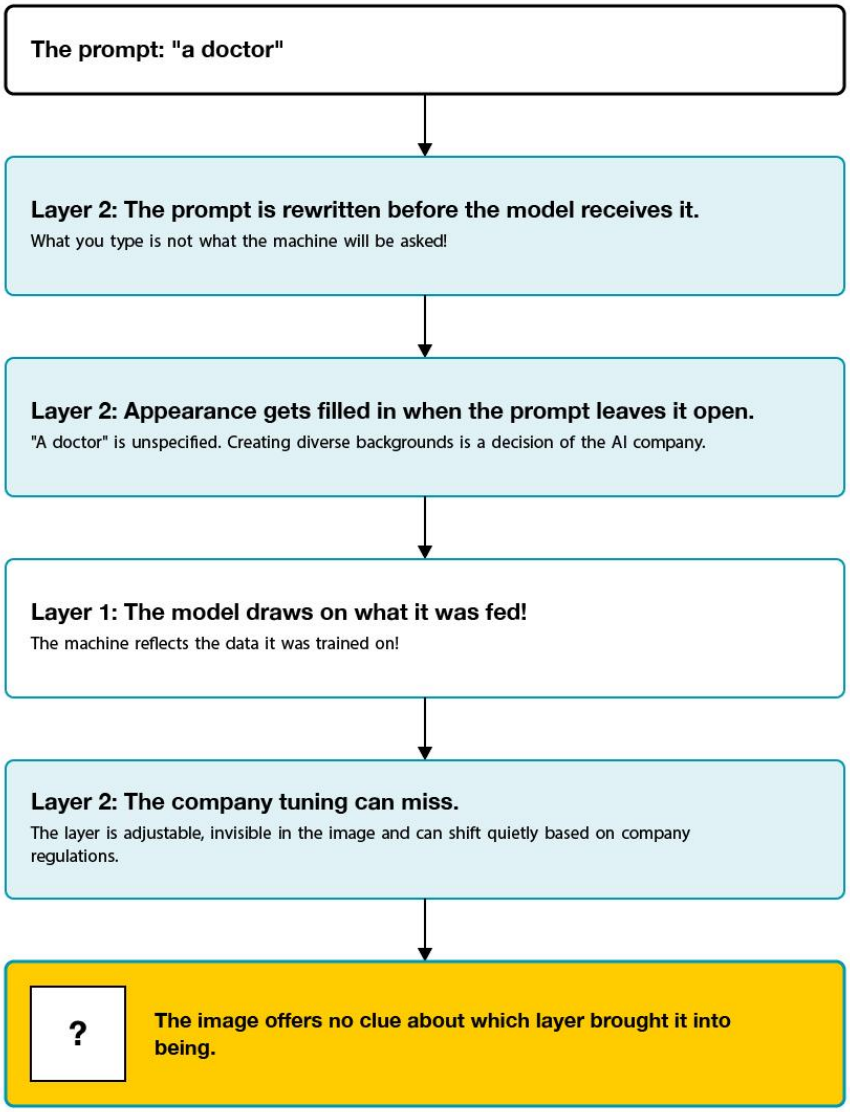


Fig. 2: «Por qué la imagen generada a partir de un prompt como “dibuja un médico” tiene el aspecto que tiene y por qué la propia imagen no os lo dirá» (imagen generada con IA, NanoBanana 2)

Esta lección puede empezar pronto. Cuando 209 estudiantes finlandeses de 12 clases de 6.º y 7.º curso exploraron este tema, el porcentaje que podía explicar el sesgo usando datos pasó de unos 7 de cada 100 a 44 de cada 100. La reserva honesta: el estudio midió lo que los niños y niñas podían explicar, no lo que harían más adelante.

«El sesgo no es exclusivo de la IA. Incluso los libros de texto pueden contenerlo. Lo importante es saber sobre el azar, sobre los datos que hay detrás de una respuesta y sobre las posibles distorsiones».

Para profundizar. Vartiainen, H., Kahila, J., Tedre, M., López-Pernas, S., & Pope, N. (2025). Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI. *New Media & Society, 27*(9), 5342–5368. Gratuito, DOI 10.1177/14614448241252820. El estudio no tenía grupo de comparación: leedlo como un diseño de aula prometedor, no como una prueba.

Weinmann, H., Messingschlager, T. V., & Appel, M. (2026). Gender bias in text-to-image generative artificial intelligence: Neglect and stereotypical presentations across three popular platforms. *New Media & Society*. Online first. Gratuito, DOI 10.1177/14614448261435197.

Ali, R., et al. (2024). Demographic representation in 3 leading artificial intelligence text-to-image generators. *JAMA Surgery, 159*(1), 87–95. DOI 10.1001/jamasurg.2023.5695.

Las declaraciones de los propios fabricantes: OpenAI (2022), *Reducing bias and improving safety in DALL·E 2*; Depue, W. (2023), *What's new with DALL·E 3?*, OpenAI Cookbook; Raghavan, P. (2024), *Gemini image generation got it wrong. We'll do better*, Google.

¿Qué solidez tienen en este momento los resultados de la investigación sobre el potencial de la IA?

Lo más importante. Si alguien os cita un estudio, haced dos preguntas antes de dar por válida su conclusión: «¿Quiénes eran los participantes y cuánto duró el estudio?»

Por qué es importante. En 2025, un metaanálisis de 51 estudios informó de un gran beneficio de ChatGPT en el rendimiento académico del alumnado. El artículo se leyó cientos de miles de veces y sirvió de base para numerosas afirmaciones seguras en las salas de profesores. El 22 de abril de 2026, la revista retiró el artículo. Una *retractación* significa que una revista retira oficialmente algo que ha publicado; el artículo sigue siendo visible, marcado con la nota «RETRACTED» (retractado), para que cualquiera que lo haya citado pueda reconocerlo. Dos investigadores que no habían participado en el estudio habían detectado discrepancias en el análisis. El editor de la revista escribió que los problemas «socavan la confianza del editor en la validez del análisis». Los autores no respondieron a la correspondencia sobre este asunto.

Esto no significa que la IA no tenga ningún impacto. Existen estudios reputados que demuestran beneficios moderados. Este campo de investigación tampoco está aún bien definido, porque es relativamente nuevo y muy dinámico. Significa que la cifra que se citaba con frecuencia era incorrecta, y que hicieron falta un año y dos expertos externos para darse cuenta de ello.

Otra reserva merece mencionarse. Al menos en matemáticas, cuanto más se acerca la evidencia al contexto escolar, menor es el efecto. En el conjunto de las materias el panorama es desigual: algunas revisiones recientes encuentran efectos mayores en las escuelas que en las universidades, pero los estudios escolares siguen siendo pocos y pequeños.

Para profundizar. La propia retractación tiene una página, es gratuita y merece la pena leerla precisamente por ser breve: Wang, J., y Fan, W. (2026). Retraction note: The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking. Humanities and Social Sciences Communications, 13, 528. DOI 10.1057/s41599-026-07310-z.

El original retractado era *HSSC 12,* 621 (2025). Sigue vigente: Wu, X., Zhu, P., Zhang, J., Yin, M., y Wang, Y. (2026). ChatGPT's impact on student learning outcomes: A meta-analysis of 35 experimental studies. *Humanities and Social Sciences Communications, 13*, 684. Gratuito, DOI 10.1057/s41599-026-07019-z.

La cifra referida solo al ámbito escolar en matemáticas: Yi, L., Liu, D., Jiang, T., y Xian, Y. (2025). The effectiveness of AI on K-12 students' mathematics learning: A systematic review and meta-analysis. *International Journal of Science and Mathematics Education, 23*(4), 1105–1126. DOI 10.1007/s10763-024-10499-7.

Un análisis combinado más amplio y más reciente: Han, X., Peng, H., y Liu, M. (2025). The impact of GenAI on learning outcomes: A systematic review and meta-analysis of experimental studies. *Educational Research Review, 48*, 100714. DOI 10.1016/j.edurev.2025.100714.

Riesgos de la IA

(1) Hacer algo no es lo mismo que aprender

Imaginad esto. Antes de poner una tarea que un chatbot pueda completar, pensad cuál es el propósito de esa tarea. ¿Podría ser buena idea cambiarla?

Por qué es importante. ¿Se aprende más rápido con un chatbot? Cerca de mil alumnos y alumnas de secundaria (13 a 16 años) practicaron matemáticas en tres grupos: uno con un chatbot estándar, otro con una versión programada para no dar la respuesta directamente y hacer una pregunta de seguimiento, y otro sin chatbot. Mientras usaba el chatbot, el grupo con el chatbot estándar completó un 48 % más de ejercicios de práctica que el grupo sin chatbot. Después, todos hicieron el mismo examen sin chatbot. El resultado, que al principio pareció sorprendente, fue que el grupo que usó el chatbot estándar obtuvo un 17 % peor rendimiento que el alumnado que no había usado ningún chatbot. El grupo que había trabajado con el chatbot que no daba la respuesta directamente no mostró ninguna caída de rendimiento en la prueba en comparación con el grupo sin chatbot, y además practicó más. El alumnado que usó el chatbot estándar no fue ni perezoso ni hizo trampas. Trabajaron más y aprendieron menos, porque lo que la herramienta eliminó no fue el esfuerzo, sino el desafío cognitivo que constituye el aprendizaje.

«La IA hace que sea muy fácil obtener una solución sin entenderla».

Para profundizar. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., y Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. PNAS, 122, e2422633122. DOI 10.1073/pnas.2422633122.

(2) Los detectores de IA castigan al alumnado equivocado

Imaginad esto. Recibís la redacción de un estudiante para corregir. Con curiosidad sobre su autenticidad, la introducís en un detector de IA. El resultado aparece: la redacción fue casi con toda seguridad escrita por IA.

Por qué es importante. Usáis el detector de IA para garantizar la equidad, pero para un grupo en concreto, los detectores tienen exactamente el efecto contrario. Un detector no reconoce texto generado por máquina. Mide cuán predecible es la redacción. Y precisamente ahí está el problema, por ejemplo, cuando un alumno escribe en una lengua extranjera, tiene poca práctica escribiendo o dispone de un vocabulario cotidiano limitado. Su escritura es predecible. Cuando siete detectores de uso extendido analizaron 91 redacciones escritas por hablantes no nativos para el examen de inglés TOEFL, recopiladas en un foro chino en línea, las señalaron erróneamente como generadas por máquina en alrededor de seis de cada diez casos. Las redacciones de alumnado estadounidense de octavo curso escritas en su lengua materna, el inglés, se clasificaron casi todas correctamente. Una herramienta que genera falsos positivos exactamente para un grupo no es neutral. Además, ese grupo es el que menos capacidad tiene de defenderse.

Para profundizar. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., y Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7), 100779. Gratuito, DOI 10.1016/j.patter.2023.100779.

(3) ¿Dónde diablos acabarán los datos del alumnado?

Imaginad esto. Usáis una herramienta de IA que ofrece vuestra escuela. Intentad averiguar adónde van los datos del alumnado.

Por qué es importante. ¿Puede vuestra escuela garantizar que los datos no se transmiten a un tercero? Un tercero es una empresa que no es ni vuestra escuela ni el desarrollador de la aplicación. Normalmente son empresas de publicidad o analítica que reciben datos mientras el alumno trabaja. Nadie en la escuela eligió a esa empresa. Cuando los investigadores realizaron un análisis técnico de 163 productos educativos recomendados por gobiernos durante los cierres escolares relacionados con la pandemia, 145 de ellos procesaban los datos de menores de un modo que ponía en peligro o vulneraba sus derechos, al transmitir datos a 196 terceras empresas o concederles acceso, principalmente del sector publicitario. De los 42 gobiernos que desarrollaron sus propios productos en lugar de comprarlos, 39 habían desarrollado productos con el mismo problema. El estudio abarca productos de 2021, y es posible que algunos hayan cambiado desde entonces. Sin embargo, no podemos descartar que ocurra lo mismo con las aplicaciones de IA.

Para profundizar. Human Rights Watch (2022). «How dare they peep into my private life?»: Children's rights violations by governments that endorsed online learning during the Covid-19 pandemic. Gratuito en hrw.org. Los anexos por país permiten consultar qué se recomendó en cada lugar.

(4) La IA llega la última donde más se necesita

Imaginad esto. ¿Cuánta formación habéis recibido ya sobre aprender acerca de la IA o sobre enseñar con IA en vuestra materia?

Por qué es importante. En un experimento, personas adultas con y sin formación universitaria resolvieron una tarea empresarial similar a las del mundo laboral. Con un asistente de IA, desaparecieron aproximadamente tres cuartas partes de la brecha entre ambos grupos. Cuando

los investigadores retiraron después el asistente, una parte considerable de la brecha volvió a aparecer. Los autores concluyen que la IA hizo gran parte del trabajo y que solo una parte de la mejora se mantuvo una vez retirada.

Una encuesta a directores de escuelas estadounidenses añade la segunda mitad del panorama: cuanto más desfavorecido era el alumnado de una escuela, menos había hecho la escuela para integrar la IA (menos normativas, menos formación del profesorado, menos orientación sobre el uso por parte del alumnado). Se trata de una correlación, no de una prueba de causalidad.

En conjunto: una buena herramienta puede reducir una brecha mientras esté en manos del alumno. Pero tiende a llegar a las escuelas que menos la necesitan, y solo una parte de su beneficio permanece cuando se retira.

Las cifras alemanas muestran el mismo patrón. Entre 1590 jóvenes de 14 a 20 años, el 80 % de las familias más acomodadas, pero solo el 55 % de las menos acomodadas, veían la IA como una oportunidad para la escuela. El uso sigue la misma pendiente: entre los jóvenes que usan herramientas de IA, el 70 % del alumnado del Gymnasium (secundaria de orientación académica) las usa para la escuela o los deberes, frente al 58 % en otros centros de secundaria.

¿Y la formación profesional del profesorado? Una encuesta a más de diez mil docentes en Inglaterra reveló que el 45 % del profesorado de escuelas privadas ya había recibido formación reglada sobre IA, frente al 21 % en escuelas públicas. Dentro del sistema público, la cifra oscilaba entre el 26 % en las escuelas con alumnado más acomodado y el 18 % en las de alumnado menos acomodado.

Para profundizar. Cruces, G., Fernández Meijide, D., Galiani, S., Gálvez, R. H., y Lombardi, M. (2026). *Does generative AI narrow education-based productivity gaps? Evidence from a randomized experiment* (NBER Working Paper 34851). Documento de trabajo, aún sin revisión por pares. Campos, C., y Singleton, J. D. (2026). *AI diffusion gaps: Unequal integration of AI across K-12 schools* (NBER Working Paper 35347). Latham, K., y Montacute, R. (2025). *Artificial advantage? AI in the classroom and the inequality gap*. The Sutton Trust. Gratuito en suttontrust.com. Vodafone Stiftung Deutschland (2024), *Pioniere des Wandels*; mpfs (2024), *JIM-Studie 2024*. Ambos gratuitos, en alemán.

(5) Lo que mejora al aprender con IA no es necesariamente lo que se ve

Imaginad esto. Introducís un nuevo chatbot en vuestra docencia y el alumnado se muestra entusiasmado por usarlo.

Por qué es importante. Una amplia revisión de 228 estudios analizó cuatro aspectos distintos que la IA podría mejorar: conocer y comprender; aplicar lo que se sabe; el lado afectivo, como la motivación y la implicación; y la capacidad de planificar, supervisar y corregir el propio aprendizaje. Tres de los cuatro mejoraron. El último no, hasta donde los estudios permitían saber.

- En cuanto a conocer y comprender, la IA generativa obtuvo mejores resultados que formas más antiguas de IA, como los sistemas de tutoría inteligente y el software de práctica adaptativa.
- En cuanto a aplicar el conocimiento y al lado afectivo, la IA generativa no obtuvo mejores resultados que el software más antiguo del que ya disponen muchas escuelas.
- En cuanto a la autorregulación, el efecto combinado fue pequeño y podría deberse al azar; además, había demasiado pocos estudios para preguntarse qué tipo de IA podría ayudar.

Solo un puñado de los resultados procedía de ensayos aleatorizados, y el entusiasmo que veis en las primeras semanas es justo lo que mejor captan los estudios cortos.

Una breve nota al margen: en esta revisión, los mayores avances se encontraron en las artes y las humanidades.

Para profundizar. Yeo, G. H., y Lansford, J. E. (2025). Effects of artificial intelligence on educational functioning: A review and meta-analysis. Educational Psychology Review, 37(4), artículo 110. DOI 10.1007/s10648-025-10085-5.

Revision #5
Created 2026-09-01 09:14:59 UTC by Alexander
Updated 2026-10-20 15:26:53 UTC by Alexander