

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 llegaron sin instrucciones. En el Reino Unido, una revisión encontró que solo 7 de cada 100 empresas de tecnología educativa habían sometido alguna vez sus productos a un ensayo controlado, y solo 12 de cada 100 habían buscado una certificación externa. El aula cuenta la misma historia. Solo 11 de cada 100 directores y docentes de 17 estados estadounidenses habían solicitado alguna vez pruebas revisadas por pares antes de adoptar una nueva herramienta.

El resultado, en muchos casos, es dinero desperdiciado. En EE. UU., dos tercios de las licencias de software escolar acumulan polvo, y el 98 % apenas se utiliza. El EdTech Genome Project analizó unas 7000 herramientas educativas con un valor de 13 000 millones de dólares y descubrió que el 85 % estaban mal adaptadas o mal utilizadas.

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

Dejemos esto claro desde el principio. 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. 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. Por último, pero no menos importante, la prueba utilizada para evaluar la eficacia a menudo fue diseñada por las mismas personas que impartieron las clases. Esto también puede haber distorsionado los datos.

(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. La IA puede apoyar la adquisición de conocimiento factual. 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 tradicional. 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. Rechazan expresamente la idea de que lo mismo se aplicaría a tareas que requieren integrar múltiples ideas. Además, se trataba de estudiantes universitarios, no de alumnado de secundaria.

Una amplia revisión que combinó los resultados de 228 estudios llegó a una conclusión similar. Los chatbots basados en IA generativa produjeron un efecto mayor sobre el conocimiento (y la comprensión) que los sistemas de tutoría inteligente y el software de práctica adaptativa.

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 estudio británico investigó en qué medida el uso específico de la IA para planificar clases puede ayudar a reducir la

carga de trabajo docente. Los investigadores asignaron aleatoriamente a 259 profesores de ciencias de 68 escuelas a un grupo que usaba un chatbot junto con breves instrucciones escritas o a un grupo que continuaba planificando las clases como de costumbre. El tiempo semanal de planificación para las clases de ciencias de 1.º y 2.º de la ESO se redujo de unos 81 minutos a unos 56. 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, lo que significa que el tiempo ahorrado no se produce automáticamente ni se mantiene por sí solo. (2) El ensayo midió el tiempo que el profesorado dedicaba a planificar las clases, no los resultados de aprendizaje del alumnado.

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 señalar la siguiente aplicación. Los chatbots son la aplicación más antigua y mejor documentada de la «IA» en educación, y esto es precisamente lo que a menudo 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: una pequeña máquina física situada sobre una mesa, a menudo operada por un investigador desde otra sala. 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). 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. Imagen creada con NanoBanana 2, creada el 20 de agosto de 2026. Prompt de un solo intento: a doctor, a nurse (un médico, una enfermera)

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 sistemática, Midjourney representaba casi siempre a los médicos como hombres blancos, mientras que Adobe Firefly hacía intentos visibles, aunque imperfectos, de diversificar.

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: la imagen no es un espejo, y no es casualidad. ¡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.

Esta lección puede empezar pronto. Cuando 209 estudiantes finlandeses de 12 clases de 4.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.

Para profundizar. Vartiainen, H., et al. (2025). Enhancing children's understanding of algorithmic biases in and with text-to-image generative AI. *New Media & Society*, 27(9). Gratuito, DOI 10.1177/14614448241252820

Weinmann, H., et al. (2026). Gender bias in text-to-image generative artificial intelligence: Neglect and stereotypical presentations across three popular platforms. *New Media & Society*, Online First

¿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 beneficio significativo 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 externos habían identificado incoherencias en la síntesis de los estudios. 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 que merece mencionarse: cuanto más se acerca la evidencia al contexto escolar, al menos por el momento, menor es el efecto.

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). *HSSC*, 13, 684. Gratuito, DOI 10.1057/s41599-026-07019-z. La cifra referida solo al ámbito escolar: Yi, L., Liu, D., Jiang, T., y Xian, Y. (2025). *International Journal of Science and Mathematics Education*, 23(4), 1105-1126. DOI 10.1007/s10763-024-10499-7.

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.

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 comerciales analizaron 91 redacciones escritas por hablantes no nativos que realizaban un examen de inglés, esas redacciones fueron señaladas incorrectamente como generadas por máquina en alrededor de seis de cada diez casos. En cambio, las redacciones escritas por estadounidenses de 14 años 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](https://www.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. A personas adultas que ya habían terminado su formación se les proporcionó un asistente de IA. La brecha de rendimiento entre quienes tenían niveles más altos y más bajos de formación reglada se redujo significativamente. Cuando los investigadores retiraron después el asistente, la brecha reapareció parcialmente. Los investigadores concluyeron que el asistente de IA había realizado el trabajo, pero no había ayudado a los participantes a desarrollar ninguna competencia.

Una encuesta a directores de escuelas estadounidenses reveló el siguiente panorama: cuanto más desfavorecido era el alumnado de una escuela, menor era el impacto de la IA en la enseñanza.

Si combinamos los hallazgos de ambos estudios, podemos afirmar lo siguiente: una buena herramienta puede cerrar una brecha mientras esté en manos del alumno. Sin embargo, normalmente no llega a quienes más la necesitan, y no deja ninguna competencia 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 y el 55 % de las menos acomodadas veían la IA como una oportunidad. Y a ese reconocimiento le sigue la acción: el 70 % del alumnado de bachillerato general usaba la IA para los deberes, frente al 58 % en la educación secundaria general.

¿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 sobre IA. Esto se compara con el 21 % en escuelas públicas. Dentro del sistema público, las cifras oscilaban entre el 26 % en las escuelas más ricas y el 18 % en las más pobres.

(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 usar la nueva herramienta.

Por qué es importante. Una amplia revisión analizó cuatro aspectos distintos del aprendizaje (conocer y comprender, trabajar con IA, la implicación, la motivación y la confianza, y la capacidad de planificar, supervisar y corregir el propio trabajo). Este último no mejoró, mientras que los demás sí.

- En cuanto a conocer y comprender, la IA generativa tuvo efectos mayores que los sistemas de tutoría inteligente y el software de práctica adaptativa ya utilizados en las escuelas.
- En cuanto a aplicar el conocimiento, y en motivación e implicación, la IA no obtuvo mejores resultados que el software ya existente.
- En cuanto a la autorregulación, no disponemos de suficientes estudios para determinar ningún efecto.

Una breve nota al margen: los mayores impactos se encuentran en las artes y las humanidades, más que en las matemáticas o las ciencias naturales.

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 #3

Created 2026-09-01 09:14:59 UTC by Alexander

Updated 2026-09-01 14:19:30 UTC by Alexander