

# LA INTELIGENCIA ARTIFICIAL EN LA EDUCACIÓN: OPORTUNIDADES, RETOS Y EQUIDAD EN UN NUEVO PARADIGMA DE APRENDIZAJE

**Almudena Sevilla**

*London School of Economics*

**Pilar Cuevas-Ruiz**

*London School of Economics y Universidad de Sevilla*

**Luz Rello**

*IE*

**Ismael Sanz**

*Funcas, URJC y LSE*

## Resumen

Este artículo analiza el impacto de la inteligencia artificial generativa en el ámbito educativo, comparando el rendimiento de sistemas como GPT-4 con el de estudiantes en pruebas PISA y PIAAC. Se identifican oportunidades, como el potencial de la IA para apoyar la enseñanza en lectura y ciencias, y retos como la equidad digital y las brechas socioeconómicas en el acceso a estas tecnologías. A través de evidencia empírica y datos de Google Trends, se muestra cómo el uso de ChatGPT en educación se ha expandido en España de manera desigual, aunque en el último año se ha producido una tendencia hacia la convergencia territorial.

*Palabras clave:* inteligencia artificial, educación, brecha digital, equidad, PISA, PIAAC, ChatGPT, aprendizaje.

## Abstract

This paper examines the impact of generative artificial intelligence on education, comparing the performance of models like GPT-4 with students in PISA and PIAAC assessments. It identifies opportunities—such as AI's potential to support reading and science education—and challenges, including digital equity and socioeconomic disparities in access. Using empirical evidence and Google Trends data, the article shows that ChatGPT usage in Spain's educational sector initially expanded unevenly but is now showing signs of regional convergence.

*Keywords:* artificial intelligence, education, digital divide, equity, PISA, PIAAC, ChatGPT, learning.

*JEL classification:* C83, I21, J24, O33.

## I. INTRODUCCIÓN

En los últimos años, la inteligencia artificial (IA) ha pasado de ser una herramienta integrada en múltiples aspectos de la vida cotidiana. En especial, la irrupción de los modelos generativos de lenguaje natural, como ChatGPT, ha transformado radicalmente la manera en que las personas acceden a la información, interactúan con contenidos digitales y desarrollan procesos de aprendi-

zaje. Esta revolución tecnológica ha desencadenado una reflexión profunda sobre el papel que debe desempeñar la educación en un mundo donde el conocimiento está al alcance de un clic y donde las capacidades cognitivas de las máquinas avanzan a una velocidad sin precedentes.

Este artículo se plantea una pregunta clave: ¿cómo puede la irrupción de la inteligencia artificial (IA), y en particular de los modelos generativos

de lenguaje como ChatGPT, transformar el aprendizaje en los sistemas educativos sin agravar las desigualdades existentes? Esta cuestión resulta cada vez más urgente, ya que la IA ha dejado de ser una tecnología de futuro para convertirse en una realidad cotidiana que redefine la forma en que las personas acceden a la información, interactúan con contenidos digitales y desarrollan procesos de aprendizaje.

La importancia de responder a esta pregunta radica en que la incorporación de estas tecnologías, si no se hace de forma reflexiva y equitativa, puede acentuar desigualdades educativas preexistentes. La velocidad a la que avanzan las capacidades de la IA amenaza con desbordar los marcos institucionales y pedagógicos tradicionales, generando una brecha entre quienes pueden aprovechar sus beneficios y quienes quedan rezagados. Si los sistemas educativos no se adaptan a tiempo, corren el riesgo de perpetuar nuevas formas de exclusión digital y limitar las oportunidades de aprendizaje precisamente en los contextos que más lo necesitan.

En este contexto, los sistemas educativos enfrentan una doble tarea. Por un lado, deben preparar a los estudiantes para convivir con tecnologías que replican —e incluso superan— ciertas competencias humanas. Por otro, deben garantizar que esta transición no profundice las desigualdades, sino que contribuya a una educación más equitativa e inclusiva. La IA no debe entenderse únicamente como un instrumento para automatizar tareas o aumentar la eficiencia. Implica también promover una nueva cultura del aprendizaje basada en la creatividad, la autonomía, el pensamiento crítico y la responsabilidad ética. Un enfoque integral debe combinar innovación tecnológica con inclusión educativa, evitando que las brechas digitales existentes se amplíen con la expansión de estas herramientas. Así, el uso de la IA en educación debe orientarse hacia la equidad, asegurando que todos los estudiantes, independientemente de su origen socioeconómico, puedan beneficiarse de sus ventajas.

Además, la rápida evolución de las capacidades de la IA en tareas cognitivas fundamentales plantea preguntas sobre los desafíos que supondrá para el empleo y sobre cómo deberían adaptarse

los sistemas educativos. Una posible vía de actuación es que la educación y la formación ayuden a las personas a mantenerse al ritmo del progreso de la IA mediante la mejora de sus competencias básicas. Fortalecer las habilidades en lectura, matemáticas y ciencias no solo es crucial para seguir siendo competitivos en entornos laborales digitalizados, sino que también constituye la base para adquirir nuevas competencias y acceder a aprendizajes futuros.

Para abordar estas cuestiones, el artículo se estructura en cinco secciones. Tras esta introducción, la segunda parte presenta los resultados más recientes sobre las capacidades de la inteligencia artificial generativa comparadas con las de los estudiantes, utilizando las pruebas PISA y PIAAC. La tercera sección examina los retos de equidad y accesibilidad digital que plantea la integración de la IA en los sistemas educativos. La cuarta sección aborda un estudio de caso sobre el uso real de ChatGPT en España a partir de datos de Google Trends. Por último, el quinto apartado recoge las principales conclusiones y propone recomendaciones para una adopción justa y efectiva de estas tecnologías en educación.

## II. OPORTUNIDADES: CAPACIDADES DE LA IA Y NUEVAS COMPETENCIAS EDUCATIVAS

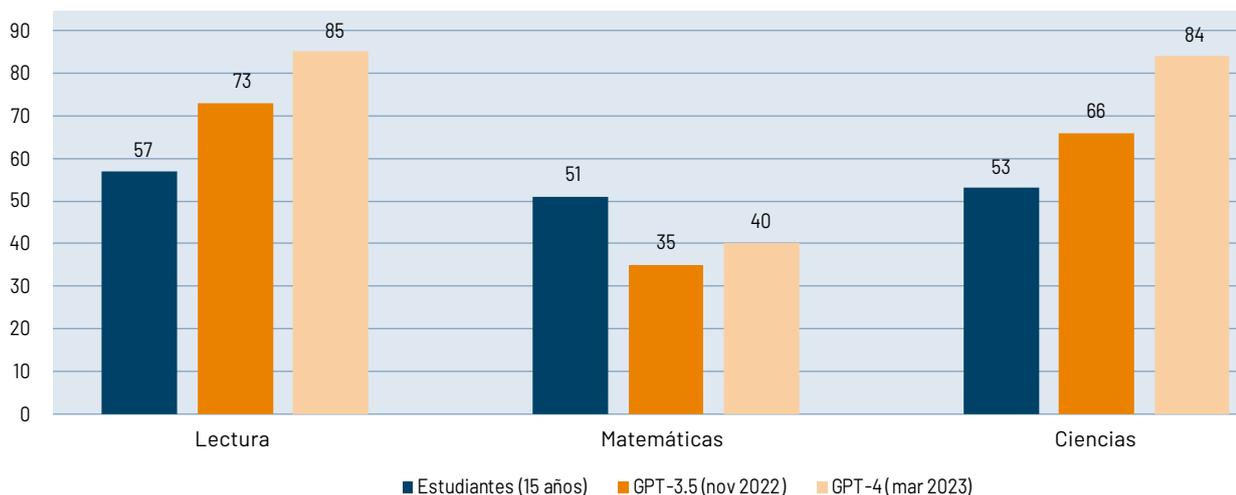
Los desarrollos recientes en inteligencia artificial (IA), particularmente en el ámbito del procesamiento del lenguaje natural, están redefiniendo el alcance de las habilidades cognitivas que pueden ser replicadas por sistemas automatizados. Esta transformación plantea un desafío fundamental: identificar hasta qué punto las habilidades humanas siguen siendo insustituibles y en qué medida pueden ser complementadas —o incluso superadas— por sistemas de IA. Con este objetivo, la OCDE (Programa para la Evaluación Internacional de las Competencias de la Población Adulta) (2023) llevó a cabo una evaluación comparativa del rendimiento de modelos de lenguaje generativo de gran tamaño (LLMs, por sus siglas en inglés) frente a estudiantes de 15 años de los países que componen la OCDE en tareas de lectura, matemáticas y ciencias del programa PISA.

La OCDE publica cada tres años los resultados de PISA, la prueba que evalúa las competencias de los alumnos de 15 años en lectura, matemáticas y ciencias. El promedio de la puntuación de la OCDE en esta evaluación se fijó en 500 puntos en edición inicial de cada competencia (2.000 lectura, 2.003 matemáticas y 2.006 ciencias). Cada edición se centra en una de las tres competencias, lo que significa que cerca de dos tercios de las pruebas se centran en esa área y la estimación de sus resultados es más precisa que las de las otras dos materias secundarias. En 2018, la competencia principal fue lectura, en 2022 matemáticas y en 2025 está siendo ciencias. PISA 2022 incluirá también además pensamiento creativo y PISA 2025 las competencias en un idioma extranjero. La comparación se realizó utilizando ítems liberados oficialmente por la OCDE, lo que permite una evaluación directa y estandarizada de las capacidades de los modelos GPT-3.5 (noviembre 2022) y GPT-4 (marzo 2023). El gráfico 1 muestra el porcentaje de acier-

tos obtenidos por estudiantes de 15 años, GPT-3.5 y GPT-4 en cada una de las tres áreas de la prueba. El análisis revela una asimetría significativa en el rendimiento relativo de los modelos de IA según el dominio evaluado. En lectura, GPT-4 obtiene un 85 por 100 de aciertos, seguido de GPT-3.5 con un 73 por 100, frente al 57 por 100 del alumnado. Esta diferencia de 28 puntos entre GPT-4 y los estudiantes indica que las habilidades de comprensión textual, inferencia y análisis lingüístico son replicables –e incluso mejorables– mediante sistemas entrenados masivamente con corpus textuales. El rendimiento superior de GPT sugiere una capacidad no solo para decodificar información, sino para estructurar respuestas coherentes en función del contexto, una competencia clave en los entornos laborales actuales. En Ciencias, el tipo de razonamiento requerido es más aplicado y contextual que en lectura. Aun así, los modelos también superan con claridad a los estudiantes: GPT-4 alcanza un 84 por 100, frente al 66 por 100 de GPT-3.5 y al 53

**GRÁFICO 1**  
**GPT Y RENDIMIENTO DE LOS ESTUDIANTES EN PISA**

Porcentaje de preguntas correctamente respondidas por estudiantes, GPT-3.5 y GPT-4 en ítems liberados de las pruebas PISA



*Nota:* Las barras reflejan el porcentaje de preguntas respondidas correctamente por estudiantes de 15 años, GPT-3.5 y GPT-4 en ítems liberados de lectura, matemáticas y ciencias. En el caso de los estudiantes, los porcentajes se calcularon a partir de la información disponible públicamente sobre su rendimiento, utilizando el banco de ítems liberados de PISA. Para los sistemas GPT, los porcentajes se calcularon promediando las puntuaciones obtenidas por los modelos en todos los ítems. El número de ítems de PISA utilizados en este estudio fue de 44 para lectura, 42 para matemáticas y 34 para ciencias. Estos ítems liberados se utilizaron en las encuestas principales e incluyen información sobre la dificultad de cada pregunta, el porcentaje de estudiantes que la respondieron correctamente y las competencias evaluadas en cada dominio.

*Fuente:* OCDE (2023). *Putting AI to the text*. [https://www.oecd.org/en/publications/putting-ai-to-the-test\\_2c297e0b-en.html](https://www.oecd.org/en/publications/putting-ai-to-the-test_2c297e0b-en.html)

por 100 de los alumnos. Es decir que los modelos no solo interpretan textos científicos, sino que también infieren hipótesis, reconocen relaciones causales y operan con conceptos abstractos. Esta capacidad tiene implicaciones profundas para la enseñanza de ciencias, donde la transmisión de conocimientos básicos puede ser fácilmente automatizada, dando la oportunidad a que la enseñanza se centre en habilidades experimentales y críticas.

En contraste con los dos dominios anteriores, en Matemáticas la ventaja se revierte. Los estudiantes resuelven correctamente el 51 por 100 de los ítems, superando tanto a GPT-3.5 (35 por 100) como a GPT-4 (40 por 100). Si bien se observa una mejora entre versiones (GPT mejora su rendimiento en 5 puntos porcentuales entre noviembre de 2022 y marzo de 2023), el rendimiento de los modelos sigue siendo sustancialmente inferior al de los estudiantes de 15 años de los países de la OCDE. Las matemáticas, especialmente las tareas que implican razonamiento abstracto, manipulación sim-

bólica o resolución paso a paso, continúan siendo un área en la que los LLM encuentran dificultades. Esta diferencia evidencia que el razonamiento lógico secuencial aún no es una competencia plenamente internalizada por la IA, posiblemente porque se basa en estructuras cognitivas más robustas que las estadísticas de lenguaje. Las matemáticas se consideran, en general, una tarea difícil para la inteligencia artificial, de modo que la IA todavía no ha logrado dominar las pruebas de razonamiento cuantitativo (Choi, 2021; Hendrycks *et al.*, 2021).

En conjunto, el gráfico 1 pone de relieve dos conclusiones clave. En primer lugar, que la IA ha superado ya a los estudiantes en tareas complejas de procesamiento textual y razonamiento científico. En segundo lugar, que aún existen nichos cognitivos —como el razonamiento matemático formal— donde la ventaja humana se mantiene. Sin embargo, la rapidez en la mejora del rendimiento entre GPT-3.5 y GPT-4 plantea la posibilidad de que esa ventaja sea transitoria.

#### RECUADRO 1

##### EJEMPLO DE PREGUNTA DE PISA, PUNTUACIÓN Y VALORACIÓN CON CHATGPT (OCDE, 2023A)

El Centro de la OCDE para la Investigación e Innovación Educativa (CERI) trabajó con expertos en ciencias de la computación para poner a prueba a GPT-3.5 y GPT-4 utilizando preguntas del programa PISA. La evaluación se centró en tres competencias: lectura (44 ítems), matemáticas (42 ítems) y ciencias (34 ítems). Todos los ejercicios fueron seleccionados de las versiones de PISA que han sido publicadas abiertamente correspondientes a las ediciones de 2000, 2003 y 2006. Los modelos GPT fueron evaluados con materiales accesibles públicamente, por lo que no se puede descartar que sus respuestas estuvieran condicionadas por haber sido entrenados con contenidos iguales o similares disponibles en línea.

La metodología empleada para valorar el desempeño de los modelos consistió en varias etapas. En primer lugar, se extrajeron manualmente las preguntas en formato de texto plano a partir de los documentos PDF originales, excluyendo cualquier contenido no textual —como gráficos o esquemas— que no puede ser

procesado por modelos de lenguaje que solo aceptan texto. La evaluación de las respuestas fue realizada manualmente por el equipo investigador utilizando como referencia las soluciones oficiales de PISA. La puntuación final se calculó en función del porcentaje de respuestas correctas, otorgándose 0,5 puntos a las respuestas parcialmente correctas.

#### A) *Pregunta original de PISA:*

El Instituto Nacional de Evaluación Educativa (INEE) del Ministerio de Educación, Formación Profesional y Deportes tiene en su web las preguntas de PISA liberadas por la OCDE <https://www.educacionfpydeportes.gob.es/inee/publicaciones/items-liberados/pisa-ocde.html>

Un ejemplo de una pregunta de PISA matemáticas es la referida a un monopatín:

Marcos es un gran fan del monopatín. Entra en una tienda denominada PATINADORES para mirar algunos

RECUADRO 1 (continuación)

**EJEMPLO DE PREGUNTA DE PISA, PUNTUACIÓN Y VALORACIÓN CON CHATGPT (OCDE, 2023A)**

precios. En esta tienda puedes comprar un monopatín completo. Pero también puedes comprar una tabla, un juego de cuatro ruedas, un juego de dos ejes y un conjunto de piezas para ensamblar los tres componentes anteriores y montar tu propio monopatín. Los precios de los productos de la tienda son:

PRODUCTO	PRECIO EN ZEDS
Monopatín completo	82 o 82
Tabla	40, 60 o 65
Un juego de cuatro ruedas	14 o 36
Un juego de dos ejes	16
Un juego de piezas para montar (cojinetes, almohadillas de goma, tornillos y tuercas)	10 o 20

**Pregunta 1**

Marcos quiere montar su propio monopatín. ¿Cuál es el precio mínimo y el precio máximo de los monopatines montados por uno mismo en esta tienda?

- a) Precio máximo: ..... zeds
- b) Precio mínimo: ..... zeds

**Pregunta 2**

La tienda ofrece tres tablas diferentes, dos juegos diferentes de ruedas y dos conjuntos diferentes de piezas para montar. Solo hay un juego de ejes para elegir. ¿Cuántos monopatines distintos puede construir Marcos?

- a) 6
- b) 8
- c) 10
- d) 12

**Pregunta 3**

Marcos tiene 120 zeds para gastar y quiere comprar el monopatín más caro que pueda. ¿Cuánto dinero puede gastar Marcos en cada uno de los cuatro componentes? Escribe tu respuesta en la tabla de abajo.

COMPONENTE	CANTIDAD (ZEDS)
Tabla	
Ruedas	
Ejes	
Piezas para ensamblar	

**B) Pregunta de PISA extraída para GPT:**

Marcos es un gran fan del monopatín. Entra en una tienda denominada PATINADORES para mirar algunos precios. En esta tienda puedes comprar un monopatín completo. Pero también puedes comprar una tabla, un juego de cuatro ruedas, un juego de dos ejes y un conjunto de piezas para ensamblar los tres componentes anteriores y montar tu propio monopatín. Los precios de los productos de la tienda son:

PRODUCTO	PRECIO EN ZEDS
Monopatín completo	82 o 82
Tabla	40, 60 o 65
Un juego de cuatro ruedas	14 o 36
Un juego de dos ejes	16
Un juego de piezas para montar (cojinetes, almohadillas de goma, tornillos y tuercas)	10 o 20

**C) Criterios de corrección**

**Máxima puntuación:**

Código 21: Tanto el mínimo (80) como el máximo (137) correctos.

**Puntuación parcial:**

Código 11: Solo el mínimo (80) correcto.  
Código 12: Solo el máximo (137) correcto.

**Sin puntuación:**

Código 00: Otras respuestas.  
Código 99: Sin respuesta.

**Dificultad:**

- Puntuación 2: 496 (nivel 3)
- Puntuación 1: 464 (nivel 2)

**Porcentaje de aciertos:**

**Puntuación 2**

OCDE: ..... 66,7 por 100  
España: ..... 66,6 por 100

**Puntuación 1**

OCDE: ..... 10,6 por 100  
España: ..... 10,1 por 100

Más allá de comparar directamente el rendimiento de modelos de lenguaje con el de estudiantes, la OCDE (2023b) ha desarrollado un segundo enfoque metodológico para valorar las capacidades de la IA. En lugar de observar el comportamiento de los sistemas frente a las pruebas, este segundo método se basa en el juicio experto para estimar si la IA, utilizando las tecnologías actuales, sería capaz de resolver las tareas del *Survey of Adult Skills* del programa PIAAC. Esta estrategia, recogida en el capítulo 3 del informe *Is Education Losing the Race with Technology?* (OCDE, 2023b), consiste en presentar a un grupo de expertos en inteligencia artificial las preguntas de lectura (alfabetización) y matemáticas (numeración) del Programa para la Evaluación Internacional de Competencias de Adultos (PIAAC, por sus siglas en inglés), con el fin de que valoren hasta qué punto un sistema podría resolverlas con acierto. PIAAC es una iniciativa de la OCDE que se considera una versión para adultos del programa PISA. Evalúa competencias clave en comprensión lectora, matemáticas y resolución de problemas en entornos tecnológicos (denominada “resolución de problemas adaptativos”) entre personas de 16 a 65 años. A través de una escala de 0 a 500 puntos, el rendimiento en lectura y matemáticas se organiza en seis niveles de dificultad, desde “por debajo del nivel 1” hasta el nivel 5, mientras que en resolución de problemas el nivel 5 no está definido. Las tareas están diseñadas para reflejar situaciones reales en los ámbitos del trabajo, la vida cotidiana, la comunidad y la educación. PIAAC permite comparar la preparación de la población adulta para afrontar las exigencias cognitivas del mundo contemporáneo, ofreciendo datos sobre el grado de alfabetización funcional de los adultos, su capacidad de razonamiento cuantitativo y su competencia digital. En este contexto, la OCDE ha desarrollado una metodología innovadora para estimar si los sistemas de inteligencia artificial serían capaces de superar estas pruebas, lo cual se convierte en un indicador revelador del potencial y los límites actuales de la IA en dominios cognitivos complejos.

Las tareas del PIAAC miden competencias funcionales en contextos reales —como interpretar gráficos, gestionar información escrita o realizar cálculos básicos—, y están diseñadas para reflejar las demandas cognitivas de la vida diaria y el entor-

no laboral. El panel estuvo compuesto por quince especialistas internacionales en áreas como el razonamiento automatizado, la comprensión de lenguaje, la representación del conocimiento o el razonamiento matemático. Siguiendo una adaptación del método Delphi, los expertos evaluaron 113 ítems (57 de lectura y 56 de matemáticas), asignando a cada uno una probabilidad —de 0 a 100 por 100— de que la IA pudiera resolverlo. Además, se les pidió que justificaran sus valoraciones, identificaran las partes de la tarea que podrían ser abordadas por la IA y, al finalizar, proyectaran su evolución de cara a 2026.

A diferencia del estudio piloto de 2016 (Elliott, 2017), la nueva edición introdujo mejoras metodológicas sustanciales: mayor interacción entre los participantes, revisión iterativa de respuestas y recogida sistemática de comentarios cualitativos. Esta aproximación permitió no solo obtener una medida agregada de la capacidad actual de la IA en tareas del PIAAC, sino también captar el grado de consenso (o desacuerdo) entre los expertos, especialmente en el dominio de matemáticas, donde las opiniones resultaron más divididas.

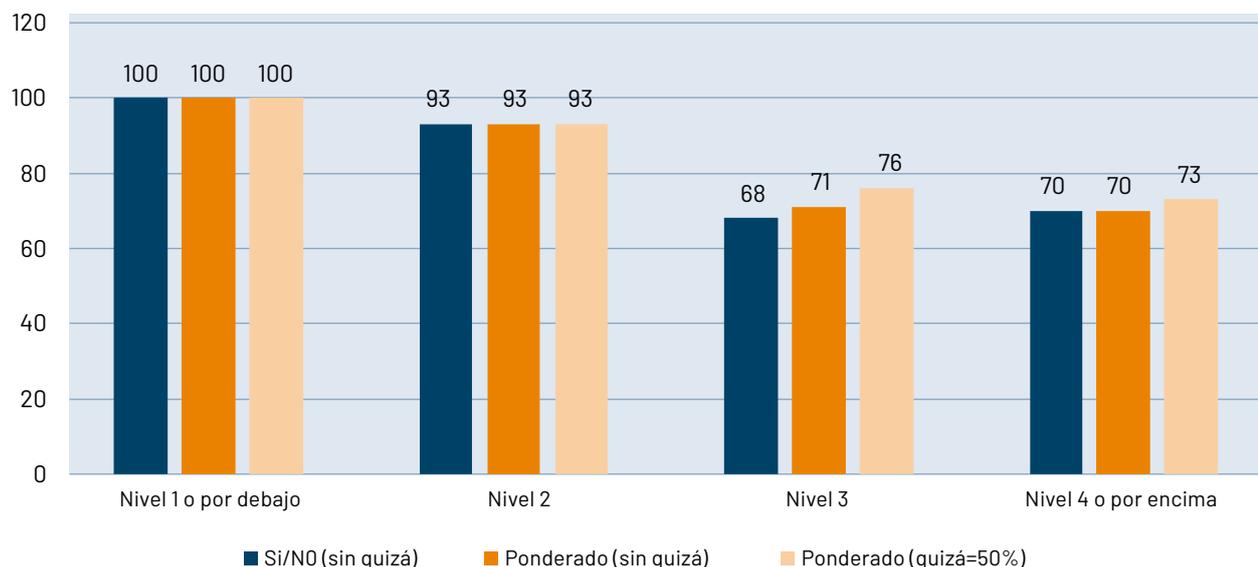
Este segundo enfoque —desarrollado también por la OCDE (2023b)— resulta útil para explorar los límites y posibilidades de la IA sin depender de un único sistema o versión concreta. En lugar de observar el rendimiento observado, capta el potencial de la tecnología en función del conocimiento experto acumulado. Los resultados permiten construir indicadores ajustados por nivel de dificultad, nivel de confianza y grado de incertidumbre, ofreciendo así una imagen más matizada del estado actual y futuro de la IA en relación con las competencias humanas fundamentales.

La valoración de los expertos se organizó en torno a tres grandes ejes: el rendimiento estimado de la IA según el nivel de dificultad de las preguntas, el grado de acuerdo y desacuerdo entre expertos, y la comparación con los niveles de competencia típicos en la población adulta. A ello se suman reflexiones cualitativas sobre el alcance real de la IA en estas competencias fundamentales. Los resultados de la evaluación en lectura indican que la inteligencia artificial actual, especialmente en sus

GRÁFICO 2

**RENDIMIENTO DE LA IA EN COMPRENSIÓN LECTORA EN PIAAC SEGÚN LOS EXPERTOS**

Porcentaje de preguntas de comprensión lectora en PIAAC que la IA puede responder correctamente según la mayoría simple de expertos



Fuente: OCDE (2023b). *Is education losing the race with technology? AI's progress in maths and reading.* [https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology\\_73105f99-en.html](https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology_73105f99-en.html)

aplicaciones de procesamiento del lenguaje natural, puede resolver una proporción significativa de las tareas planteadas en PIAAC.

El gráfico 2 proporciona una representación del rendimiento estimado de la IA en comprensión lectora según el nivel de dificultad de las preguntas del PIAAC. En él se muestran tres formas distintas de computar las valoraciones de los expertos: una versión dicotómica (azul), que interpreta los juicios extremos (0 por 100 y 25 por 100 como “no”, y 75 por 100 y 100 por 100 como “sí”); una versión ponderada (naranja) que asigna pesos diferentes según el grado de confianza, omitiendo las respuestas de “quizás”; y una tercera variante (naranja claro) que incluye estas respuestas intermedias valorándolas como un acierto parcial (50 por 100).

Las tres formas de calcular las respuestas de los expertos conducen a resultados consistentes: los modelos de IA actuales son capaces de resolver correctamente el 100 por 100 de las preguntas de

nivel 1 o inferior, y del 93 por 100 en las de nivel 2. A partir del nivel 3, el rendimiento cae de manera notable, situándose en torno al 70 por 100. Este descenso refleja la mayor complejidad de las tareas, que requieren habilidades de inferencia, interpretación crítica o integración de información procedente de múltiples fuentes textuales, como textos disyuntivos, gráficos o tablas.

Una lectura detallada del gráfico revela que la IA ya ha superado con claridad el umbral de competencia funcional mínima exigida para los niveles más bajos, lo que sugiere que sus aplicaciones actuales podrían asumir con solvencia tareas rutinarias de lectura en contextos laborales y cotidianos. La convergencia de los resultados en los niveles 1 y 2 muestra que existe un consenso sólido entre los expertos en que estas tareas pueden ser resueltas sin necesidad de comprensión profunda, mediante técnicas de procesamiento superficial del lenguaje como el *pattern matching* o la extracción de información literal (OCDE, 2023b).

En cambio, el nivel 3 —considerado el umbral de competencia autónoma para un adulto— se presenta como una frontera más ambigua. Aunque el porcentaje de aciertos estimados ronda el 70 por 100, las diferencias entre métodos de agregación ponen de manifiesto un mayor grado de incertidumbre. Esta ambigüedad refleja que las preguntas de este nivel implican habilidades de evaluación y razonamiento contextual que aún no están plenamente resueltas por la IA. El nivel 3 coincide con el punto de inflexión donde se empieza a exigir una comprensión semántica más profunda, capaz de interpretar ironías, matices argumentativos o contradicciones implícitas en el texto.

Las estimaciones de éxito se mantienen en los mismos niveles en el nivel 4 que en el nivel 3, a pesar del salto cualitativo en la dificultad de las tareas. Aquí las preguntas requieren identificar relaciones complejas entre ideas, evaluar la credibilidad de las fuentes, o integrar información discontinua. Si bien la IA puede abordar parte de estas tareas con técnicas actuales, el rendimiento agregado sugiere que aún existen cuellos de botella cognitivos que no han sido superados.

En conjunto, el gráfico 2 muestra que el rendimiento de la IA en comprensión lectora según los expertos se aproxima al de un adulto con un nivel de competencia 3, lo que equivale a una posición intermedia-alta en la distribución de habilidades en la población adulta de los países de la OCDE. No obstante, el hecho de que el rendimiento estimado descienda con la dificultad pone de manifiesto que el dominio semántico profundo —el que permite navegar textos complejos, ambiguos o intertextuales— sigue siendo terreno difícil para los sistemas actuales.

El gráfico 3 muestra el rendimiento estimado de la IA en tareas de matemáticas de PIAAC, desglosado por nivel de dificultad y según las tres formas de cálculo. Esta evaluación se basa en la valoración de quince expertos que analizaron si un sistema de IA, con las tecnologías actuales, sería capaz de resolver los ítems de matemáticas de PIAAC, que incluyen tareas reales como interpretar gráficos, hacer cálculos con porcentajes o evaluar información cuantitativa presentada en mapas, tablas o textos.

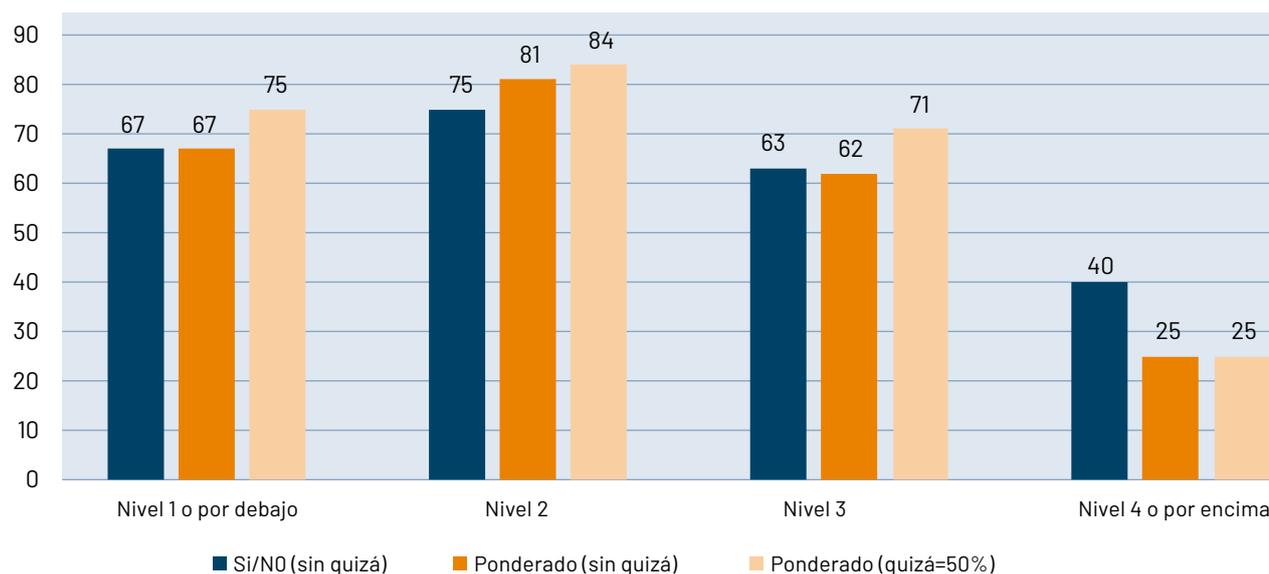
En los niveles más sencillos (nivel 1 o inferior), los expertos estiman que la IA puede resolver correctamente el 67 por 100 de las preguntas según el método dicotómico (azul), que omite respuestas de incertidumbre y computa únicamente votos afirmativos o negativos claros. Esta cifra se eleva al 75 por 100 en nivel 2 y desciende al 63 por 100 en nivel 3, mientras que en los niveles más difíciles (nivel 4 o superior), el rendimiento estimado cae al 40 por 100. Cuando se aplica una ponderación que tiene en cuenta el grado de confianza en la respuesta (método ponderado, naranja), los resultados se mantienen similares para los niveles 1, 2 y 3, pero el rendimiento en el nivel más alto baja al 25 por 100. Por último, la metodología más inclusiva (naranja claro), que considera las respuestas de “quizá” como aciertos parciales (50 por 100), estima que la IA mantiene el mismo rendimiento en niveles intermedios, pero mejora levemente en los extremos. Estos resultados revelan un patrón distinto al observado en lectura. Mientras que en comprensión textual la IA mostraba un descenso progresivo con el aumento de dificultad, en matemáticas el comportamiento es menos lineal. La IA parece manejar mejor las tareas de dificultad media que las más sencillas, lo cual contrasta con el patrón humano típico. Esto puede deberse a que muchas de las tareas de nivel 1 requieren cálculos numéricos básicos, manipulaciones aritméticas simples o interpretación de representaciones visuales (como gráficos o tablas), ámbitos donde los sistemas actuales aún presentan limitaciones, especialmente cuando las instrucciones no están explícitas en formato textual. Otro factor que influye en estos resultados es la falta de inversión y foco en tareas cuantitativas dentro del campo del procesamiento del lenguaje natural. Mientras que la comprensión textual ha sido el centro de muchas aplicaciones de IA, la numeración ha quedado relativamente rezagada, lo que limita el entrenamiento de modelos especializados en este tipo de tareas.

Cuando se comparan los resultados con el rendimiento humano, la IA muestra un desempeño inferior al de los adultos con nivel 2 en las tareas más sencillas (67 por 100 frente a 89 por 100), pero similar o incluso superior en las tareas de dificultad media. Sin embargo, sigue quedando rezagada en los niveles más altos, donde se exige razonamiento

GRÁFICO 3

**RENDIMIENTO DE LA IA EN NUMERACIÓN SEGÚN DISTINTOS MÉTODOS DE CÁLCULO**

Porcentaje de preguntas de numeración que la IA puede responder correctamente según la mayoría simple de expertos



Fuente: OCDE (2023b). *Is education losing the race with technology? AI's progress in maths and reading*. [https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology\\_73105f99-en.html](https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology_73105f99-en.html)

abstracto o integración de múltiples fuentes de información.

Los expertos encontraron más dificultades al evaluar la IA en numeración que en lectura. Las tareas matemáticas del PIAAC son más variadas y requieren razonamiento multietapa, comprensión de tablas, mapas o gráficos, y cálculos con distintas unidades y formatos. Este carácter multimodal complica el desarrollo de un sistema único que las resuelva de forma generalizada. Existe un nivel elevado de disenso entre expertos. Además, el nivel de acuerdo entre los expertos fue limitado: solo se alcanzó mayoría cualificada (dos tercios) en 18 de las 56 preguntas. La mayoría de los ítems reciben tanto valoraciones positivas como negativas, con márgenes muy estrechos entre ambos tipos de juicios, lo que indica que las mayorías simples sobre la capacidad de la IA se construyen sobre un consenso frágil. Esta falta de acuerdo se intensifica en las tareas de mayor dificultad, y en aquellas con componentes visuales, lo que revela que la interpretación de formatos no textuales sigue siendo un obstáculo

significativo para los sistemas actuales. Las diferencias entre expertos también se explican por el enfoque asumido: algunos evaluaron la capacidad de un sistema generalista para resolver todas las tareas, mientras que otros imaginaron soluciones específicas para tipos de preguntas concretos.

En resumen, el gráfico 3 evidencia que, en el dominio de las matemáticas, la IA aún no ha alcanzado un rendimiento comparable al humano en las tareas más básicas ni en las más complejas, pero muestra un desempeño prometedor en niveles intermedios. Esta paradoja —mejor rendimiento en niveles medios que en los extremos— sugiere que la IA actual aún no ha consolidado una comprensión robusta de las tareas más simples ni ha alcanzado la capacidad de razonamiento abstracto necesario para las más complejas.

El estudio de la OCDE (2023b) también arroja luz sobre la incertidumbre existente entre los expertos. En torno al 20 por 100 de las respuestas en el dominio de la lectura y el 12 por 100 en numeración fueron calificadas como "Quizá" o "No lo sé", lo que refleja una

CUADRO N.º 1

**LA MEDIDA EN LA QUE LA ESCASEZ DE RECURSOS EDUCATIVOS Y DIGITALES AFECTA AL APRENDIZAJE EN PISA 2022. RESULTADOS BASADOS EN LOS DIRECTORES DE CENTROS EDUCATIVOS**

PAÍS	ÍNDICE DE ESCASEZ DE MATERIAL		FALTA DE RECURSOS DIGITALES (%)				RECURSOS DIGITALES INADECUADOS O DE BAJA CALIDAD (%)			
	ÍNDICE MEDIO	DESVIACIÓN ESTÁNDAR	NADA	MUY POCO	HASTA CIERTO PUNTO	MUCHO	NADA	MUY POCO	HASTA CIERTO PUNTO	MUCHO
Francia	-0,40 (0,06)	0,88 (0,04)	46,4 (3,8)	30,5 (3,2)	19,7 (2,9)	3,5 (1,1)	46,2 (3,6)	31,2 (3,5)	16,5 (2,8)	6,1 (1,4)
Alemania	-0,07 (0,07)	1,00 (0,04)	27,2 (3,3)	34,5 (3,4)	27,4 (3,5)	10,9 (2,3)	25,2 (3,3)	37,8 (3,4)	26,7 (3,4)	10,3 (2,5)
Italia	-0,21 (0,07)	0,93 (0,05)	50,8 (4,0)	35,6 (3,6)	12,1 (2,4)	1,5 (0,8)	48,6 (4,1)	37,1 (3,9)	12,7 (2,4)	1,7 (1,0)
Portugal	0,24 (0,06)	0,98 (0,05)	30,4 (3,4)	40,4 (3,5)	23,4 (3,3)	5,8 (1,4)	22,7 (2,9)	37,8 (3,1)	29,9 (3,1)	9,6 (1,8)
España	-0,29 (0,04)	1,06 (0,04)	48,3 (2,4)	24,7 (2,1)	20,4 (1,9)	6,5 (1,2)	45,6 (2,3)	30,0 (2,2)	18,2 (1,6)	6,2 (1,1)
Reino Unido	-0,32 (0,06)	0,87 (0,04)	38,3 (4,0)	42,7 (4,4)	14,3 (2,7)	4,7 (1,6)	39,1 (4,0)	39,7 (4,3)	17,5 (3,1)	3,7 (1,4)
Estados Unidos	-0,66 (0,09)	0,92 (0,07)	76,5 (3,9)	16,9 (3,2)	5,1 (2,2)	1,5 (1,1)	73,6 (4,1)	17,0 (3,2)	8,0 (2,7)	1,4 (0,9)
OCDE Promedio	-0,17 (0,01)	0,97 (0,01)	47,2 (0,5)	28,8 (0,5)	16,1 (0,4)	7,8 (0,3)	45,5 (0,5)	29,9 (0,5)	17,1 (0,4)	7,5 (0,3)

Fuente: OCDE (2023). PISA 2022 Results (Volume II): Learning During - and From - Disruption (Tabla II.B1.5.17).

cierta ambigüedad sobre las capacidades reales de la IA. Esta incertidumbre aumenta con el nivel de dificultad de las preguntas, especialmente en las que requieren procesamiento de gráficos o imágenes. Pese a ello, las proyecciones para el futuro inmediato son optimistas. Todos los expertos coincidieron en que, con las tasas actuales de avance, es probable que para 2026 la IA sea capaz de resolver correctamente la totalidad de las preguntas de los dominios de lectura y matemáticas de PIAAC. Este consenso anticipa una transformación radical en los próximos años, especialmente en tareas cognitivas que tradicionalmente se consideraban exclusivamente humanas.

### III. RETOS: BRECHA DIGITAL, EQUITAD Y BARRERAS PARA UNA ADOPCIÓN INCLUSIVA DE LA IA EN EDUCACIÓN

Las herramientas de IA pueden potenciar la adaptación de contenidos, el apoyo a necesidades edu-

cativas especiales o el acceso a recursos digitales avanzados. Sin embargo, la evidencia recogida por la OCDE muestra que la generalización de estos beneficios no está garantizada, y que el acceso desigual a infraestructuras digitales y a recursos de calidad sigue siendo un obstáculo central (OCDE, 2024).

El cuadro n.º 1 presenta el índice global de escasez de materiales educativos reportado por los directores y directoras de los centros educativos en PISA 2022, así como el desglose específico de las carencias en recursos digitales, tanto en cantidad como en calidad. La selección de países incluidos en el cuadro se centra en los principales países de la OCDE más cercanos geográficamente a España, así como en aquellos con especial relevancia comparativa en el contexto europeo y transatlántico, como Francia, Alemania, Italia, Portugal, Reino Unido y Estados Unidos.

El índice de escasez de recursos educativos (EDUSHORT), utilizado en PISA 2022 y en ediciones

CUADRO N.º 2

**LA MEDIDA EN LA QUE LA ESCASEZ DE RECURSOS EDUCATIVOS Y DIGITALES AFECTA AL APRENDIZAJE EN PISA 2022. RESULTADOS BASADOS EN LOS DIRECTORES DE CENTROS EDUCATIVOS (DATOS POR COMUNIDADES AUTÓNOMAS)**

COMUNIDAD AUTÓNOMA	ÍNDICE DE ESCASEZ DE MATERIAL		FALTA DE RECURSOS DIGITALES (%)				RECURSOS DIGITALES INADECUADOS O DE BAJA CALIDAD (%)			
	ÍNDICE MEDIO	DESVIACIÓN ESTÁNDAR	NADA	MUY POCO	HASTA CIERTO PUNTO	MUCHO	NADA	MUY POCO	HASTA CIERTO PUNTO	MUCHO
Andalucía	-0,16 (0,15)	1,19 (0,13)	42,0 (7,0)	32,1 (6,2)	14,00 (5,0)	11,8 (4,9)	39,5 (7,5)	40,4 (6,8)	6,3 (3,7)	13,8 (4,4)
Aragón	-0,14 (0,13)	1,06 (0,15)	35,7 (7,9)	35,6 (8,2)	17,7 (5,2)	11,1 (4,8)	36,2 (7,4)	31,7 (8,3)	25,0 (7,0)	7,0 (4,1)
Asturias	-0,21 (0,13)	0,90 (0,07)	35,3 (6,4)	37,8 (5,4)	20,8 (5,2)	6,1 (3,5)	41,0 (6,7)	32,6 (5,6)	20,3 (6,2)	6,2 (2,8)
Illes Balears	-0,55 (0,21)	1,22 (0,18)	59,1 (7,1)	16,7 (6,5)	16,7 (5,4)	7,6 (4,4)	56,2 (7,6)	19,0 (7,2)	19,6 (4,5)	4,5 (3,2)
País Vasco	-0,26 (0,09)	1,01 (0,09)	62,1 (5,7)	26,3 (4,9)	9,0 (3,1)	2,6 (1,5)	58,5 (5,3)	25,8 (4,4)	12,9 (3,3)	2,8 (1,4)
Canarias	0,36 (0,16)	1,13 (0,10)	30,0 (7,1)	21,6 (5,3)	36,6 (7,9)	11,8 (5,3)	26,1 (7,0)	26,5 (7,1)	38,2 (8,2)	9,2 (4,6)
Cantabria	-0,41 (0,11)	0,84 (0,04)	49,1 (6,4)	28,4 (4,1)	18,8 (5,6)	3,6 (2,5)	46,5 (7,8)	36,8 (7,9)	13,5 (4,6)	3,2 (2,5)
Castilla y León	-0,44 (0,17)	1,12 (0,12)	40,80 (7,80)	32,40 (7,70)	15,10 (5,00)	11,70 (4,70)	36,00 (7,40)	35,50 (7,10)	22,00 (5,60)	6,50 (3,60)
Castilla-La Mancha	0,18 (0,10)	0,88 (0,11)	26,70 (7,10)	26,90 (6,80)	39,60 (8,80)	6,70 (3,60)	26,90 (5,70)	36,40 (6,90)	29,50 (7,30)	7,20 (4,10)
Cataluña	-0,62 (0,16)	1,00 (0,09)	70,0 (7,3)	19,8 (6,7)	10,2 (5,0)	0,0 (-)	66,2 (7,0)	23,1 (7,5)	7,6 (4,3)	3,0 (3,0)
Ceuta	0,04 (0,04)	0,84 (0,02)	16,3 (1,3)	24,1 (2,4)	36,3 (1,8)	23,4 (2,1)	20,5 (1,4)	56,2 (2,5)	0,0 (-)	23,4 (2,1)
Comunidad Valenciana	-0,41 (0,15)	0,94 (0,07)	45,70 (7,70)	15,50 (4,90)	27,30 (5,80)	11,60 (4,90)	38,50 (7,30)	23,10 (4,60)	34,20 (5,90)	4,20 (3,10)
Extremadura	-0,27 (0,17)	1,11 (0,06)	42,80 (6,80)	23,70 (6,30)	26,70 (6,90)	6,80 (3,70)	41,20 (7,00)	24,90 (6,70)	24,00 (6,80)	9,90 (4,50)
Galicia	-0,01 (0,14)	0,98 (0,07)	39,40 (7,10)	20,70 (5,40)	36,80 (6,70)	3,10 (2,30)	36,40 (7,20)	29,00 (6,80)	31,20 (6,00)	3,40 (2,20)
La Rioja	-0,41 (0,01)	0,84 (0,00)	33,20 (0,50)	46,90 (0,50)	13,70 (0,30)	6,20 (0,30)	20,30 (0,40)	56,10 (0,40)	18,50 (0,40)	5,10 (0,30)
Madrid	-0,54 (0,11)	0,91 (0,07)	58,80 (5,80)	22,00 (5,50)	18,50 (5,40)	0,70 (0,90)	56,00 (7,00)	26,00 (5,40)	14,90 (4,90)	3,10 (2,60)
Melilla	0,10 (0,05)	1,18 (0,05)	32,90 (1,90)	32,00 (1,70)	19,30 (1,40)	15,70 (1,70)	37,40 (1,90)	28,30 (1,60)	34,30 (1,80)	0,00 (-)
Murcia	0,00 (0,15)	1,09 (0,09)	28,30 (6,40)	26,50 (6,10)	38,50 (7,70)	6,70 (3,90)	36,80 (7,70)	39,10 (6,70)	22,00 (6,20)	2,10 (2,00)
Navarra	-0,63 (0,09)	0,73 (0,07)	69,50 (5,80)	19,60 (4,40)	8,60 (3,10)	2,30 (2,30)	72,60 (6,40)	15,20 (4,90)	10,40 (4,60)	1,70 (1,80)

anteriores (2015 y 2018), se construye a partir de las respuestas de los directores de centros educativos a la pregunta SC017, que evalúa en qué medida diversos factores obstaculizan la capacidad del centro para impartir instrucción. Las respuestas se recogen en una escala de cuatro niveles: “en absoluto”, “muy poco”, “hasta cierto punto” y “mucho”. El índice EDUSHORT combina cuatro ítems referidos a la cantidad y calidad de materiales educativos y de la infraestructura física del centro y está estandarizado de forma que un valor de 0 representa el promedio de los países de la OCDE. El material educativo incluye libros de texto, equipos TIC (tecnologías de la información y la comunicación), biblioteca, material de laboratorio, etc. La infraestructura física incluye el edificio escolar, los terrenos, los sistemas de calefacción/refrigeración, los sistemas de iluminación y acústica, etc. Valores negativos indican menos escasez de recursos que la media (mejor dotación), mientras que valores positivos señalan una situación peor. En este marco, países como España (-0,29), Francia (-0,40) o Reino Unido (-0,32) muestran una dotación de recursos por encima de la media, mientras que Portugal (0,24) presenta una percepción de mayor escasez. España destaca además por su elevada variabilidad interna (1,06), lo que sugiere grandes diferencias entre centros dentro del propio país.

Un elemento clave para interpretar estos resultados es su evolución temporal. Según el informe de PISA 2022, en aproximadamente la mitad de los sistemas educativos participantes, los equipos directivos de centros educativos en 2022 indicaron que había menos carencias de material educativo que en 2018. Esta mejora fue particularmente significativa en países como Irlanda, Indonesia, Croacia, España o Italia. En cambio, la escasez de personal educativo fue percibida como más acuciante en la mayoría de los países.

Más allá del índice global, el cuadro n.º 1 desagrega la información sobre recursos digitales en dos dimensiones específicas: (i) la falta de recursos digitales —como ordenadores, tablets, acceso a Internet o plataformas digitales escolares—, y (ii) la presencia de recursos digitales inadecuados o de baja calidad. En ambos casos, se muestra el porcentaje de estudiantes cuyos directores consi-

deran que la enseñanza se ve obstaculizada “en absoluto”, “muy poco”, “hasta cierto punto” o “mucho”.

En concreto, el cuadro n.º 1 (y n.º 2) se basa en indicadores contruidos a partir de las respuestas recogidas en la pregunta SC017 del Cuestionario para Directores de PISA 2022. Esta pregunta evalúa en qué medida la capacidad del centro para impartir instrucción está obstaculizada por diversos factores relacionados con los recursos escolares digitales. En particular, los ítems SC017Q09JA y SC017Q10JA se centran en:

SC017Q09JA: Falta de recursos digitales (por ejemplo, ordenadores, *tablets* o portátiles, acceso a Internet, sistemas de gestión del aprendizaje o plataformas digitales escolares).

SC017Q10JA: Recursos digitales inadecuados o de baja calidad (por ejemplo, ordenadores, *tablets* o portátiles, acceso a Internet, sistemas de gestión del aprendizaje o plataformas digitales escolares).

En España, un 26,9 por 100 del alumnado está escolarizado en centros donde los equipos directivos perciben que la falta de recursos digitales obstaculiza la enseñanza “hasta cierto punto” o “mucho”, y un 24,4 por 100 en centros donde los recursos existentes son considerados de baja calidad en alguna de esas dos categorías. Aunque estas cifras son similares a la media de la OCDE (23,9 por 100 y 24,6 por 100, respectivamente), reflejan que uno de cada cuatro estudiantes en España asiste a centros con limitaciones digitales que podrían afectar la integración efectiva y equitativa de herramientas de IA en el aprendizaje.

Destacan también casos extremos como Estados Unidos, donde más del 76 por 100 del alumnado estudia en centros sin ninguna falta de recursos digitales (y un 73,6 por 100 sin deficiencias en la calidad), frente a Portugal, donde casi un tercio de los estudiantes asisten a centros con problemas moderados o graves de calidad en los recursos digitales.

Además de las diferencias observadas entre países, la evidencia también apunta a importantes desigualdades dentro de los países, como muestra el cuadro n.º 2, que recoge datos desagregados

por comunidades autónomas en España. Este enfoque territorial permite visibilizar las asimetrías internas en el acceso a materiales educativos y, particularmente, a recursos digitales, que pueden condicionar de forma significativa la capacidad de los centros para incorporar innovaciones tecnológicas como la inteligencia artificial.

El índice de escasez de material educativo, elaborado a partir de las respuestas de los directores en PISA 2022, revela una brecha territorial clara. Comunidades como Cataluña (-0,62), Navarra (-0,63) y Madrid (-0,54) se sitúan entre las mejor dotadas, con valores muy por debajo de la media de la OCDE, lo que sugiere una percepción generalizada de suficiencia de recursos en sus centros educativos. En el caso de Cataluña, por ejemplo, el 70 por 100 del alumnado asiste a centros donde los directores indican que no existe ninguna carencia digital, y ningún centro declara sufrir obstáculos graves. Navarra muestra una situación similar, con más del 69 por 100 del alumnado en centros sin limitaciones digitales, y una desviación interna muy baja, lo que indica una mayor equidad territorial.

En contraste, comunidades como Canarias (0,36), Castilla-La Mancha (0,18), Murcia (0,00) o Melilla (0,10) presentan valores positivos en el índice, lo que indica una dotación de recursos inferior a la media de la OCDE. En estas regiones, entre un 35 por 100 y un 47 por 100 del alumnado estudia en centros donde los directores reportan escasez moderada o alta de recursos digitales. En Canarias, por ejemplo, el 38,2 por 100 de los estudiantes están en centros donde la enseñanza se ve obstaculizada “hasta cierto punto” por la baja calidad de los recursos digitales, y un 9,2 por 100 en centros donde esta limitación es grave. Estas cifras duplican las de comunidades como el País Vasco o La Rioja, y sitúan a Canarias como una de las regiones con mayores barreras para una implementación equitativa de tecnologías emergentes como la IA.

La desigualdad interna entre centros de una misma comunidad es también un factor preocupante. Comunidades como Andalucía, Castilla y León, Islas Baleares, Aragón o Extremadura presentan desviaciones estándar elevadas en el índice de escasez (superiores a 1), lo que indica una fuerte

dispersión en la dotación de recursos entre centros. Esta dispersión sugiere que, incluso dentro de regiones con dotaciones promedio aceptables, existen focos significativos de exclusión digital concentrados en ciertos centros, habitualmente en zonas rurales o en barrios con menor renta. Estos datos muestran que la brecha digital en España no es solo una cuestión internacional o entre sectores público y privado, sino también una cuestión territorial, que atraviesa la geografía educativa del país.

En resumen, los datos por comunidades autónomas confirman que, si bien en términos agregados España muestra un desempeño relativamente favorable respecto a la media de la OCDE, la realidad interna es mucho más heterogénea. Los estudiantes que residen en determinadas regiones o que asisten a centros con menor dotación de recursos digitales pueden ver limitada su posibilidad de acceder a las oportunidades de aprendizaje que ofrecen las nuevas herramientas basadas en IA.

El despliegue de la IA en educación puede amplificar, en lugar de reducir, las brechas existentes entre centros, estudiantes y comunidades. Factores como la disponibilidad de dispositivos, la calidad de la conectividad, la formación del profesorado y la capacidad de las instituciones para integrar tecnologías emergentes condicionan la posibilidad de que todos los estudiantes se beneficien por igual. Como advierte la OCDE, la adopción “no regulada” de herramientas de IA puede acelerar la polarización: los centros con más recursos acceden antes y mejor a las innovaciones, mientras que los más desfavorecidos quedan rezagados tanto en oportunidades como en resultados. A estos retos materiales se suman otras cuestiones críticas: la presencia de sesgos algorítmicos, los riesgos para la privacidad y la protección de datos, la necesidad de reforzar la integridad educativa frente a la presión comercial y la importancia de dotar a los docentes de competencias para un uso responsable de la IA. Además, es fundamental que el desarrollo y la implantación de soluciones tecnológicas sean culturalmente sensibles y capaces de responder a la diversidad de los contextos educativos (Holmes, 2023).

El desafío, por tanto, no reside únicamente en incorporar IA a las aulas, sino en asegurar que su

adopción no agrave las desigualdades ya existentes. Un entorno educativo equitativo requiere no solo acceso físico a tecnología, sino también apoyo institucional, formación continua del profesorado y un marco ético claro que vele por los intereses de los más vulnerables. Uno de los principales retos para una adopción inclusiva de la IA en la educación reside en la persistencia de desigualdades materiales en el acceso a recursos digitales básicos. Estas desigualdades, lejos de haberse mitigado en los últimos años, siguen afectando de forma desproporcionada a los estudiantes más vulnerables, como se evidencia en el análisis de PISA 2022. El cuadro n.º 2 muestra el porcentaje de estudiantes cuyos directores de centro señalan carencias de recursos digitales —como ordenadores, conectividad a Internet o plataformas de gestión del aprendizaje—, desagregado por nivel socioeconómico del centro educativo y por tipo de titularidad (pública o priva-

da). El índice socioeconómico de PISA, denominado ESCS (Economic, Social and Cultural Status), es una medida compuesta que resume el nivel socioeconómico y cultural del entorno familiar del estudiante. Este índice se calcula a partir de tres dimensiones fundamentales. En primer lugar, el nivel educativo más alto alcanzado por alguno de los padres, codificado según la clasificación internacional ISCED-2011 y transformado en años de escolarización acumulada mediante una conversión internacional (PA-REDINT). En segundo lugar, se incorpora el estatus ocupacional más alto de los padres (HISEI), derivado de la codificación de sus ocupaciones según el sistema ISCO-08 y asignado al índice internacional de estatus socioeconómico ocupacional (ISEI). Este componente refleja la posición social asociada a la ocupación laboral de los padres, más allá del ingreso económico directo. El tercer componente es el índice HOMEPOS, que recoge la presencia de bienes

CUADRO N.º 3

**PORCENTAJE DE ESTUDIANTES EN CENTROS CON CARENCIAS DE RECURSOS DIGITALES SEGÚN NIVEL SOCIOECONÓMICO DEL CENTRO Y TITULARIDAD**

PAÍS	TOTAL ESTUDIANTES (1)	CENTROS DESFAVORECIDOS (2)	CENTROS DE NIVEL MEDIO (3)	CENTROS FAVORECIDOS (4)	DIFERENCIA FAVORABLE DESFAVORABLE (5)=(2)-(4)	CENTROS PÚBLICOS (6)	CENTROS PRIVADOS (7)	DIFERENCIA PRIVADOS - PÚBLICOS (6)-(7)
Francia	23,2 (3,0)	22,0 (5,6)	28,4 (4,8)	13,8 (4,5)	-8,2 (6,9)	21,8 (3,3)	27,9 (6,7)	6,1 (7,4)
Alemania	38,3 (3,6)	39,8 (7,3)	41,3 (5,0)	31,6 (7,1)	-8,2 (10,7)	39,3 (3,6)	9,5 (10,3)	<b>-29,8*</b> <b>(10,7)</b>
Italia	13,6 (2,5)	14,3 (5,2)	13,6 (3,3)	13,0 (5,6)	-1,3 (7,6)	13,3 (2,6)	21,3 (10,7)	8,0 (10,9)
Portugal	29,2 (3,2)	26,7 (5,8)	34,4 (4,8)	21,5 (6,1)	-5,2 (8,3)	32,0 (3,7)	13,6 (5,7)	<b>-18,3*</b> <b>(7,4)</b>
España	27,0 (2,0)	31,6 (4,8)	30,3 (2,8)	16,1 (3,3)	<b>-15,5*</b> <b>(5,9)</b>	29,5 (2,4)	21,6 (3,0)	<b>-7,9*</b> <b>(3,7)</b>
Reino Unido	19,0 (3,1)	26,6 (7,5)	19,7 (3,9)	12,7 (5,1)	-13,9 (8,8)	26,8 (4,8)	15,1 (3,9)	-11,7 (6,3)
Estados Unidos	6,6 (2,4)	8,0 (6,4)	5,6 (2,6)	5,8 (3,9)	-2,1 (7,4)	5,2 (1,8)	—	—
OCDE Promedio	23,9 (0,4)	27,8 (0,9)	22,4 (0,6)	18,3 (0,9)	<b>-9,5*</b> <b>(1,3)</b>	26,0 (0,6)	13,4 (1,0)	<b>-13,5*</b> <b>(1,2)</b>

Notas: El perfil socioeconómico de los centros se define según el índice ESCS de PISA: los centros desfavorecidos corresponden al cuartil inferior y los favorecidos al cuartil superior del índice dentro de cada país. En negrita se destacan las diferencias estadísticamente significativas.

Fuente: OCDE (2023). PISA 2022 Results (Volume II): Learning During – and From – Disruption (Tabla II.B1.5.19). <https://doi.org/10.1787/a97db61c-en>.

educativos, tecnológicos y culturales en el hogar, como número de libros, ordenador para uso escolar, conexión a Internet, escritorio propio, calculadora, libros de literatura, libros de referencia (como diccionarios o enciclopedias), lavadora, lavavajillas, y otros objetos que varían según el país. Las variables se agrupan en un índice mediante análisis de componentes principales y son adaptadas culturalmente en cada país. El valor final del índice socioeconómico estandariza, de modo que 0 representa el nivel medio de estatus socioeconómico de la OCDE y una desviación estándar es igual a 1. Así, un valor positivo indica un entorno más favorecido que la media de la OCDE, mientras que un valor negativo señala un contexto más desfavorecido.

Los datos del cuadro n.º 3 reflejan un patrón claro: en todos los países analizados, los centros con alumnado socioeconómicamente desfavorecido indican una mayor frecuencia de carencias digitales que aquellos con alumnado más favorecido. Se calcula el porcentaje de estudiantes que asisten a centros donde los directores indican que la instrucción se ve afectada “hasta cierto punto” o “mucho” por la escasez (cuadro n.º 3) o la baja calidad de los recursos digitales, en caso de haberlos (cuadro n.º 4). Ambos indicadores permiten evaluar con precisión hasta qué punto la infraestructura digital de los centros educativos condiciona la equidad en el acceso a las oportunidades de aprendizaje, especialmente ante el despliegue de herramientas de la IA en la educación. En promedio, el 27,8 por 100 del alumnado de centros desfavorecidos en países de la OCDE estudia en centros con limitaciones en el acceso a tecnologías digitales, frente al 18,3 por 100 en los centros favorecidos. Esta diferencia de 9,5 puntos porcentuales —estadísticamente significativa— pone de manifiesto una desigualdad estructural que condiciona la capacidad de los centros para incorporar herramientas de IA en condiciones de equidad. España es uno de los países donde esta brecha es especialmente acusada y también significativa: el 31,6 por 100 del alumnado en centros desfavorecidos experimenta limitaciones digitales, frente al 16,1 por 100 en los favorecidos, lo que representa una diferencia de 15,5 puntos, por encima de los 9,5 puntos de la OCDE.

El problema no se limita únicamente al nivel socioeconómico de los centros, sino también a su

titularidad. En promedio en la OCDE, el 26 por 100 del alumnado en centros públicos está expuesto a carencias digitales, frente al 13,4 por 100 en centros privados, lo que se traduce en una brecha significativa de 13,5 puntos. Esta desigualdad es aún mayor en Alemania (29,8 puntos) y Portugal (18,3 puntos), mientras que en España también alcanza niveles relevantes (7,9 puntos), siendo significativa en todos estos casos. Estas cifras ilustran cómo los centros privados, en general con mayor capacidad de inversión tecnológica, están en mejores condiciones para integrar herramientas educativas avanzadas, lo que puede reforzar la segmentación del sistema educativo y limitar la capacidad redistributiva de las políticas públicas. Como advierte la OCDE (2024), sin una estrategia que garantice un acceso universal a dispositivos, conectividad y recursos digitales de calidad, el uso de IA en educación puede amplificar las brechas existentes entre centros y estudiantes.

El cuadro n.º 4 profundiza en las desigualdades en el acceso a recursos digitales adecuados, de haberlos, diferenciando por nivel socioeconómico del centro y titularidad pública o privada. Los datos confirman que los estudiantes que asisten a centros desfavorecidos o públicos se enfrentan con mayor frecuencia a limitaciones en la calidad de los recursos digitales que los de centros favorecidos o privados. En promedio, el 28 por 100 del alumnado de centros desfavorecidos en los países de la OCDE está en institutos o colegios con recursos digitales inadecuados o de baja calidad, frente al 19,5 por 100 de los centros favorecidos. Esta diferencia de 8,5 puntos porcentuales, estadísticamente significativa, refleja una desventaja estructural en el acceso a tecnologías educativas básicas. En países como España, esta brecha es aún más acusada: el 26,6 por 100 de los estudiantes en centros desfavorecidos sufre esta carencia, frente al 12,7 por 100 en centros favorecidos, lo que supone una diferencia de 13,9 puntos.

En lo que respecta a la titularidad de los centros, la tendencia también es clara. En promedio, un 26,2 por 100 del alumnado de centros públicos tiene acceso limitado o de baja calidad a recursos digitales, frente a un 15 por 100 en centros privados, lo que representa una diferencia de 12 puntos

CUADRO N.º 4

**PORCENTAJE DE ESTUDIANTES EN CENTROS CON RECURSOS DIGITALES INADECUADOS O DE BAJA CALIDAD SEGÚN NIVEL SOCIOECONÓMICO DEL CENTRO Y TITULARIDAD**

PAÍS	TOTAL ESTUDIANTES (1)	CENTROS DESFAVORECIDOS (2)	CENTROS DE NIVEL MEDIO (3)	CENTROS FAVORECIDOS (4)	DIFERENCIA FAVORABLE – DESFAVORABLE (2)-(4)	CENTROS PÚBLICOS (6)	CENTROS PRIVADOS (7)	DIFERENCIA PRIVADOS – PÚBLICOS (6)-(7)
Francia	22,6 (3,0)	26,7 (6,1)	25,8 (4,4)	11,8 (4,8)	<b>-14,9</b> (7,2)	23,1 (3,4)	20,8 (5,9)	-2,3 (6,7)
Alemania	37,0 (3,3)	40,9 (7,5)	38,4 (5,0)	31,8 (7,3)	-9,1 (11,0)	38,4 (3,3)	0,0 (c)	<b>-38,4</b> (3,3)
Italia	14,3 (2,6)	15,1 (6,0)	14,5 (3,5)	13,3 (4,9)	-1,8 (7,6)	14,3 (2,7)	14,2 (9,7)	-0,1 (9,9)
Portugal	39,5 (3,4)	33,9 (6,6)	44,0 (4,5)	36,0 (6,7)	2,1 (9,5)	43,9 (3,7)	14,2 (5,8)	<b>-29,7</b> (7,1)
España	24,4 (1,8)	26,6 (4,5)	29,3 (2,7)	12,7 (2,4)	<b>-13,9</b> (5,1)	28,9 (2,5)	14,9 (2,4)	<b>-14,0</b> (3,4)
Reino Unido	21,2 (3,2)	27,9 (8,2)	23,9 (4,6)	12,5 (4,8)	-15,4 (9,2)	27,9 (5,5)	18,2 (3,9)	-9,6 (6,8)
Estados Unidos	9,4 (2,9)	9,2 (6,5)	11,7 (4,2)	3,5 (3,2)	-5,8 (7,2)	8,1 (2,5)	-	-
OCDE Promedio	24,6	28,0	23,1	19,5	<b>-8,5</b>	26,2	15,0	<b>-12,0</b>

Notas: El perfil socioeconómico de los centros se define según el índice ESCS de PISA: los centros desfavorecidos corresponden al cuartil inferior y los favorecidos al cuartil superior del índice dentro de cada país. En negrita se destacan las diferencias estadísticamente significativas.

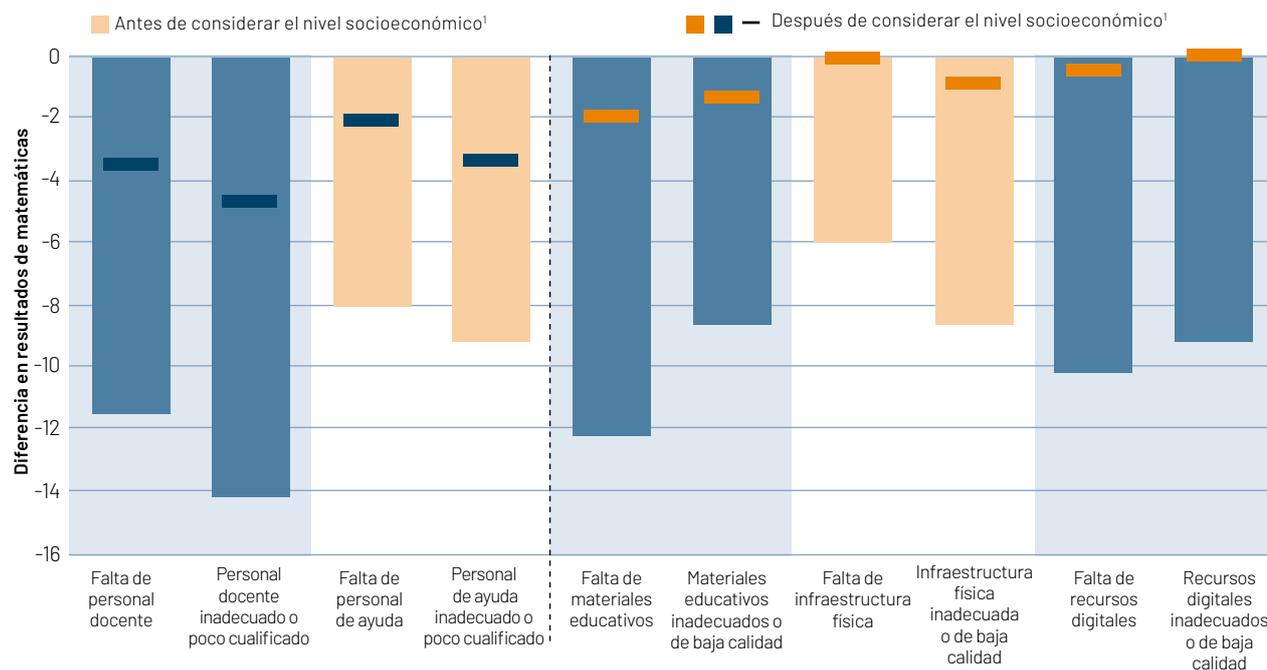
Fuente: OCDE (2023). PISA 2022 Results (Volume II): Learning During – and From – Disruption (Tabla II.B1.5.19). <https://doi.org/10.1787/a97db61c-en>.

porcentuales. En el caso español, esta brecha alcanza los 14 puntos, situando a los centros públicos en una posición de desventaja en términos de equidad digital.

Cabe destacar también casos extremos, como el de Alemania, donde la diferencia entre centros públicos y privados es de 38,4 puntos, o el de Portugal, con una diferencia de 29,7 puntos, ambos en detrimento de los centros públicos. Estas cifras revelan que, más allá del número de dispositivos disponibles, existen grandes disparidades en su calidad y adecuación pedagógica, lo cual puede impactar directamente en las oportunidades de aprendizaje del alumnado más vulnerable. En resumen, el cuadro confirma que la calidad de los recursos digitales en los centros escolares no está distribuida de manera equitativa.

Los cuadros n.º 3 y 4 han mostrado que los centros con alumnado más desfavorecido y de titularidad pública disponen de menos recursos digitales y de peor calidad. Cabe preguntarse si estas limitaciones materiales pueden estar influyendo también en el aprendizaje del alumnado. Esta cuestión resulta especialmente relevante en el contexto actual, en el que la incorporación de herramientas de IA en la educación requiere una infraestructura digital mínima para poder ser efectiva y equitativa. A este respecto, el gráfico 4 muestra que, en promedio, los estudiantes que asisten a centros donde los directores señalan escasez de materiales educativos –incluidos los digitales– obtienen peores resultados en matemáticas, el área central de PISA 2022, en los países de la OCDE. En el gráfico del informe PISA 2022 se muestra la relación entre la escasez de recursos en los centros educativos y el

**GRÁFICO 4**  
**ESCASEZ DE PERSONAL EDUCATIVO Y RECURSOS MATERIALES, Y RENDIMIENTO EN MATEMÁTICAS**



Notas: Cambio en el rendimiento en matemáticas asociado con directores que informan que la capacidad de la escuela para impartir instrucción se ve obstaculizada en cierta medida o considerablemente por los siguientes factores: personal docente, personal de ayuda, materiales educativos, infraestructura física o recursos digitales. Las diferencias en las puntuaciones que son estadísticamente significativas se muestran en un tono más oscuro. Todas las diferencias en las puntuaciones son estadísticamente significativas antes de considerar el perfil socioeconómico de los estudiantes y los centros educativos.

Fuente: OCDE, PISA 2022 Database, Annex B1, Capítulo 5.

rendimiento de los estudiantes en matemáticas. En concreto, representa la diferencia media en puntuación entre los estudiantes que asisten a centros donde los directores indican que la capacidad de impartir instrucción está obstaculizada —“hasta cierto punto” o “mucho” — por la falta o baja calidad de recursos (materiales o humanos), y aquellos que asisten a centros sin esas limitaciones (“muy poco” o “en absoluto”). Las diferencias se expresan en puntos PISA, donde 30 puntos equivalen a un curso académico de aprendizaje. A partir de este análisis más general sobre la escasez de recursos, nos centraremos a continuación en la dimensión específica de interés en este artículo: los recursos digitales. Su disponibilidad y calidad no solo son fundamentales para el funcionamiento ordinario de los centros, sino que también constituyen la base imprescindible para que las herramientas basadas

en IA puedan ser desplegadas con eficacia y equidad en el ámbito educativo.

El gráfico n.º 4 muestra que los estudiantes que asisten a centros con menos carencias de recursos digitales obtuvieron mejores resultados en matemáticas, en promedio, en los países de la OCDE y en España. Aunque el efecto negativo de esta carencia desaparece al controlar por el perfil socioeconómico de los estudiantes y de los propios centros, el dato pone de relieve que los contextos más vulnerables no solo tienen menos acceso a recursos, sino que además estos déficits están asociados con rendimientos académicos más bajos. Esta conclusión es coherente con los hallazgos del artículo de Marcenaro y López-Agudo (2025), incluido también en este número, que identifica una correlación negativa entre la falta de dotación

digital en los centros y el rendimiento medio del alumnado en matemáticas y lectura, especialmente en comunidades con menor nivel de “normas” para el uso pedagógico de estos recursos. Este patrón también ha sido documentado en estudios internacionales como el de Guan *et al.* (2025), quienes, en un análisis comparado entre China y países occidentales, muestran que un acceso adecuado a recursos digitales se asocia a mejores resultados de aprendizaje, mientras que un uso excesivo o no regulado puede tener efectos contraproducentes. Sus resultados refuerzan la importancia de diseñar políticas que no solo expandan la disponibilidad tecnológica, sino que también promuevan su uso pedagógico equilibrado y eficaz, especialmente en entornos socioeconómicamente vulnerables.

El impacto negativo de la escasez de recursos materiales y, particularmente, de recursos digitales no es homogéneo, sino que afecta más intensamente a quienes ya parten de una posición de desventaja. Además, aunque los directores en 2022 mostraron una menor preocupación por la escasez de materiales educativos en comparación con 2018, la variabilidad interna entre centros educativos de un mismo país, y más aún en España, sigue siendo elevada.

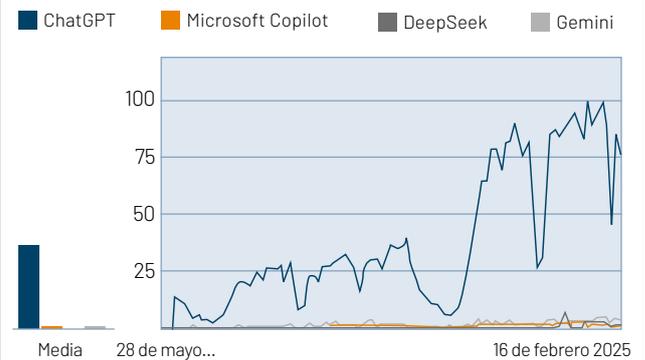
Los sistemas educativos donde los estudiantes asistieron a centros con menos carencias de recursos digitales, o con recursos digitales adecuados y de alta calidad, obtuvieron también mejores resultados en matemáticas, en los países de la OCDE antes de tener en cuenta los perfiles socioeconómicos; pero esta relación deja de ser significativa al tener en cuenta los perfiles socioeconómicos similares.

#### IV. EL ESTUDIO DE LA EQUIDAD EN EL ACCESO A CHATGPT EN ESPAÑA: UN ANÁLISIS DE GOOGLE TRENDS

En esta sección analizamos empíricamente la equidad en el uso de inteligencia artificial en el ámbito formativo. Con este propósito se ha llevado a cabo una estimación econométrica que permite identificar diferencias relevantes en su intensidad de uso a lo largo del tiempo y entre territorios en

GRÁFICO 5  
EVOLUCIÓN DEL INTERÉS DE BÚSQUEDA EN HERRAMIENTAS DE IA GENERATIVA EN EDUCACIÓN (ESPAÑA, 2023-2025)

Interés a largo tiempo



Fuente: Google Trends.

España. Bacher-Hicks *et al.* (2021) utilizan los datos de búsqueda en Internet para estudiar en tiempo real cómo buscaron los hogares estadounidenses recursos de aprendizaje en línea cuando las escuelas cerraron debido a la pandemia de la COVID-19. Estos autores llegaban a la conclusión de que la intensidad de búsqueda a nivel nacional de recursos de aprendizaje en línea se duplicó con respecto a los niveles previos. Sin embargo, las zonas del país con mayores ingresos, mejor acceso a internet y menos escuelas rurales experimentaron aumentos significativamente mayores en la intensidad de búsqueda. Bacher-Hicks *et al.* (2021) concluían que es probable que la pandemia amplíe las brechas de rendimiento académico dado que las distintas escuelas y familias interaccionan de forma diferente con los recursos en línea para compensar el tiempo de aprendizaje perdido en la escuela.

El gráfico n.º 5 muestra la evolución del interés en búsquedas web sobre cuatro herramientas de inteligencia artificial generativa –ChatGPT, Microsoft Copilot, DeepSeek y Gemini– en España, dentro de la categoría de empleo y educación, entre junio de 2023 y mayo de 2025. Los datos provienen de Google Trends y reflejan el nivel de búsquedas relativas en Internet de cada término, normalizado en una escala de 0 a 100. Los resultados ponen de manifiesto una hegemonía de ChatGPT (línea azul) como la herramienta más buscada en este ámbito.

El interés por esta aplicación crece de forma sostenida desde mediados de 2023, experimentando un notable repunte a partir de septiembre de 2024, coincidiendo con el inicio del curso académico. A partir de entonces, las búsquedas se mantienen en valores altos (por encima de 75 puntos), con picos que alcanzan el valor máximo (100) durante algunas semanas del primer trimestre de 2025. El resto de las herramientas analizadas presentan niveles de búsqueda más bajos y relativamente estables. Microsoft Copilot (naranja) y Gemini (gris claro) muestran una ligera presencia en el gráfico, sin variaciones relevantes. DeepSeek (gris oscuro) aparece con un pico visible en febrero de 2025. El gráfico 5 también permite conocer la variable a explicar que se emplea en la regresión. En concreto, se trata de la intensidad de búsqueda definida como el porcentaje de búsquedas de un término en concreto sobre el total de búsquedas en Google. Y además esa información que ofrece Google Trends se encuentra en términos relativos al máximo de intensidad de búsqueda en el periodo analizado, que en el caso del último año en España sería el que alcanzó ChatGPT en la semana del 9 al 15 de marzo de 2025, que es igual a 100. El resto de la información de las demás semanas se expresa en relación con ese máximo.

A continuación, el análisis econométrico se centra en ChatGPT, dado que es la IA más usada en España en empleo y educación. La especificación de la regresión sigue el modelo estimado por Bacher-Hicks *et al.* (2021) y seguido después por Sanz *et al.* (2021):

$$IB_{smac} = \sum_{m=10}^{12-1} \beta_s \text{dummymes}_m + \sum_{a=1}^{3-1} \beta_a \text{dummyaño}_a + \sum_{c=1}^{17-1} \beta_c \text{dummycomunidad}_c + \sum_{a=1}^{3-1} \beta_a \text{PIBpercapita}_c * \text{dummyaño}_a$$

donde  $IB$  es la intensidad de búsqueda de ChatGPT cada semana desde junio de 2023 a mayo de 2025, en mes  $m$  del año  $a$ , en la aplicación  $r$  y en la Comunidad  $c$ . La variable dependiente de intensidad de búsqueda de ChatGPT en relación con el total de búsquedas en Google, proviene de Google Trends y se encuentra en términos relativos al máximo de búsquedas de cada término educativo durante el período. Desde junio de 2023 a mayo de 2025 han transcurrido 101 semanas. El número de observaciones totales de las que se dispone es

de diecisiete comunidades autónomas por 101 semanas, 1.717 observaciones. La variable de interés es el PIB per cápita de la comunidad autónoma  $c$ . El coeficiente asociado al PIB per cápita indicará si en las comunidades autónomas con mayor renta se hacen más búsquedas en Google de ChatGPT que en aquellas con menor renta. Un coeficiente positivo de  $\beta_a$  nos indicará que, como se ha planteado en las secciones anteriores, en las zonas con mayor PIB per cápita se accede con más intensidad a la IA en el ámbito del empleo y la educación. Por el contrario, un coeficiente negativo indicaría que en las comunidades autónomas con menor renta se busca en mayor medida ChatGPT.

Las *dummies* de mes permiten analizar cómo varía el acceso a ChatGPT en España a lo largo del año, en donde el mes omitido será el de diciembre. Las *dummies* para cada uno de los tres años, 2023, 2024 y 2025 muestran la evolución de las búsquedas de ChatGPT en relación con todas las demás búsquedas, en donde 2023 será el año de referencia. Por último, las *dummies* para cada una de las comunidades autónomas señalan si en cada región hay un mayor o menor acceso a esta IA con un objetivo de empleo o educación.

El cuadro n.º 5 presenta los resultados de la intensidad de búsquedas relativas de ChatGPT a lo largo del tiempo y entre comunidades autónomas, tomando como referencia la Comunidad de Andalucía, el mes de diciembre y el año 2023. La interpretación de los coeficientes debe realizarse en comparación con la categoría de referencia: un valor negativo indica una intensidad menor respecto a Andalucía, diciembre o el año 2023, mientras que un valor positivo refleja una mayor intensidad. Se utilizan tres especificaciones que incluyen efectos fijos por comunidad, por mes y por año, incorporando en los modelos 2 y 3 el PIB per cápita y su interacción con los años de referencia. La variable dependiente representa la intensidad de búsquedas en Google sobre ChatGPT en España en términos relativos a cualquier otra búsqueda en Google.

En primer lugar, se observan diferencias territoriales significativas. País Vasco (-25,3), Cantabria (-20,3), Baleares (-17,1), Navarra (-15,2),

## CUADRO N.º 5

**EFFECTO SOBRE LA INTENSIDAD DEL USO DIGITAL POR COMUNIDAD AUTÓNOMA, MES Y AÑO**

Modelo con variables regionales, temporales y PIB per cápita

COMINIDADES	BÚSQUEDAS CHATGPT		
Andalucía (ref.)	–	–	–
Aragón	-8,931*** (4,18)	-8,823*** (1,86)	-8,823*** (1,84)
Asturias	-12,941*** (6,06)	-12,894*** (1,92)	-12,894*** (1,90)
Baleares	-17,198*** (8,06)	-17,093*** (1,86)	-17,093*** (1,84)
Canarias	-2,000 (0,94)	-1,989 (2,07)	-1,989 (2,05)
Cantabria	-20,307*** (9,51)	-20,258*** (1,91)	-20,258*** (1,89)
Castilla-La Mancha	-2,762 (1,29)	-2,649 (1,87)	-2,649 (1,85)
Castilla y León (om.)	–	–	–
Cataluña	-11,891*** (5,57)	-11,887*** (2,11)	-11,887*** (2,09)
Ceuta y Melilla	-2,178 (1,02)	-2,127 (1,90)	-2,127 (1,88)
C. Valenciana	-9,851*** (4,61)	-9,791*** (1,88)	-9,791*** (1,86)
Extremadura (om.)	–	–	–
Galicia	-11,366*** (5,32)	-11,342*** (2,01)	-11,342*** (1,98)
Madrid	12,277*** (5,75)	12,302*** (2,00)	12,302*** (1,98)
Murcia	-15,941*** (7,47)	-15,810*** (1,91)	-15,810*** (1,89)
Navarra	-15,366*** (7,20)	-15,213*** (2,00)	-15,213*** (1,98)
País Vasco	-25,386*** (11,89)	-25,296*** (1,85)	-25,296*** (1,83)
La Rioja	1,287 (0,60)	1,318 (1,98)	1,318 (1,96)
Enero	-19,358*** (10,13)	-19,358*** (1,91)	-19,358*** (1,89)
Febrero	-13,505*** (7,07)	-13,505*** (1,91)	-13,505*** (1,89)
Marzo	-15,039*** (8,26)	-15,039*** (1,82)	-15,039*** (1,80)
Abril	-20,072*** (10,50)	-20,072*** (1,91)	-20,072*** (1,89)
Mayo	-10,989*** (5,40)	-10,989*** (2,03)	-10,989*** (2,01)
Junio	-14,465*** (8,55)	-14,465*** (1,69)	-14,465*** (1,67)
Julio	-17,433*** (10,31)	-17,433*** (1,69)	-17,433*** (1,67)
Agosto	-18,621*** (10,67)	-18,621*** (1,75)	-18,621*** (1,72)
Septiembre	-6,348*** (3,75)	-6,348*** (1,69)	-6,348*** (1,67)
Octubre	6,692*** (3,96)	6,692*** (1,69)	6,692*** (1,67)
Noviembre	14,254*** (8,17)	14,254*** (1,75)	14,254*** (1,72)
Diciembre (ref.)	–	–	–
Año 2023 (ref.)	–	–	–
Año 2024	18,649*** (20,39)	18,649*** (0,91)	32,107*** (4,55)
Año 2025	59,902*** (38,44)	59,902*** (1,56)	98,451*** (6,15)
PIB per cápita		-0,00001 (0,00011)	–
PIB*Año2023	–	–	0,00043*** (0,00015)
PIB*Año2024	–	–	-0,00001 (0,00013)
PIB*Año2025	–	–	-0,00083*** (0,00018)
Constante	19,442*** (10,10)	19,660*** (4,01)	6,243 (4,91)
R <sup>2</sup>	0,66	0,6600	0,6683
N	1,717	1,717	1,717

Murcia (-15,8), Asturias (-12,9) y Cataluña (-11,8) muestran coeficientes negativos estadísticamente significativos, lo que indica una menor importancia en la búsqueda de ChatGPT en Google que en Andalucía. Solo Madrid presenta un valor positivo y significativo (+12,3), lo que sugiere un mayor acceso a ChatGPT con objetivo de empleo o educación en comparación con el resto del país. En segundo lugar, los resultados muestran una clara estacionalidad en las búsquedas de ChatGPT. Los meses del verano y del final del curso escolar —junio, julio y agosto— presentan coeficientes negativos y significativos respecto a diciembre, que es el mes base. Agosto (-18,6), julio (-17,4) y abril (-20) son los meses con menor intensidad de acceso a la principal herramienta de IA, lo que se explica por el calendario académico con los descansos del verano y la Semana Santa. Por el contrario, octubre (+6,6) y noviembre (+14,2) muestran un mayor acceso a ChatGPT, lo que podría estar relacionado con el arranque del curso y la mayor carga de trabajo formativo en el primer trimestre del calendario escolar. En tercer lugar, el modelo revela un fuerte crecimiento en las búsquedas de IA en el ámbito de la educación y el empleo desde 2023.

En la segunda especificación se introduce, además, el PIB per cápita de 2023 de cada comunidad autónoma. Este término no es significativo. Es decir, que no hay evidencia en España de que las búsquedas de ChatGPT en Google entre junio de 2023 y mayo de 2025, en relación con las búsquedas de cualquier otro término, estén relacionadas con el nivel de renta de la región en donde reside la persona que está intentando acceder a la IA.

En la última especificación (modelo 3), se explora además cómo la renta regional interacciona con la evolución temporal. La interacción del PIB per cápita con el año 2023 arroja un coeficiente positivo y significativo (+0,00043), lo que indica que en 2023 las regiones con mayor renta intensificaron más rápidamente el acceso a ChatGPT. Sin embargo, en 2025 el signo de la interacción se revierte (-0,00083), lo que sugiere un proceso de convergencia: las regiones con menor renta parecen haber acelerado su incorporación al uso de la IA en la educación y el empleo. Es decir, que las comunidades autónomas con mayor nivel del

PIB per cápita accedieron con mayor intensidad a ChatGPT, una estela que siguieron las regiones con menor renta casi dos años después. La evidencia que encontramos sugiere que, efectivamente, existe una desigualdad inicial en el acceso a la IA en la educación, que con el tiempo va desapareciendo. Los datos de Google Trends hacen referencia a las búsquedas en Google de términos relacionados con ChatGPT, de modo que una vez que esta herramienta está ya instalada no es necesario que el usuario la busque en internet. Por lo tanto, el cuadro n.º 5 muestra que esos primeros accesos a esta IA en el ámbito de la educación y el empleo se produjeron en buena medida casi dos años antes en las comunidades autónomas de mayor PIB per cápita.

En conjunto, los resultados del cuadro n.º 5 muestran que las búsquedas de ChatGPT en educación y empleo en España ha seguido una trayectoria ascendente, con importantes desigualdades territoriales, variaciones estacionales asociadas al calendario académico y de trabajo, marcada por una caída del uso en los meses de verano y repuntes en otoño, cuestiones estas que reflejan la influencia de los calendarios escolares en la adopción tecnológica.

## V. CONCLUSIONES Y RECOMENDACIONES

La inteligencia artificial generativa ya está alterando de forma tangible el ecosistema educativo, mostrando un rendimiento superior al de los estudiantes en competencias como lectura o ciencias, aunque con notables limitaciones en razonamiento matemático y tareas multimodales. La velocidad de mejora de estos modelos sugiere que muchas de estas limitaciones podrían ser transitorias. Sin embargo, la integración de la IA en la educación no es neutral. Los datos de PISA y PIAAC revelan que su adopción puede amplificar desigualdades si no se acompaña de políticas específicas de equidad digital. La brecha entre centros públicos y privados, entre territorios o entre niveles socioeconómicos condiciona el acceso a las herramientas de IA, tanto por la infraestructura disponible como por las capacidades institucionales para aprovecharlas. Además, la evidencia

obtenida a través de Google Trends confirma que las comunidades autónomas con mayor renta accedieron antes a estas tecnologías, lo que sugiere un patrón de adopción desigual en sus primeras fases. No obstante, se observa una tendencia reciente hacia la convergencia, lo que indica que aún estamos a tiempo de intervenir para garantizar una difusión más equitativa.

En este contexto, resulta preocupante la percepción negativa del uso prolongado o inadecuado de dispositivos, sin distinguir entre un uso pasivo y un uso estructurado, personalizado y orientado al aprendizaje. Como muestra la evidencia del artículo publicado en Sevilla et al. (2024), cuando el aprendizaje asistido por ordenador es impulsado por los propios docentes y contextualizado dentro del centro, puede tener efectos muy positivos en el aprendizaje, especialmente en matemáticas, donde se han documentado mejoras equivalentes a medio curso escolar. Estas herramientas permiten adaptar la enseñanza al nivel de cada estudiante, reforzando sus áreas de mayor déficit, algo difícilmente alcanzable en aulas con alta ratio sin apoyo tecnológico.

Además, se debe poner en valor el enfoque de desarrollo que han seguido algunas herramientas digitales, como la que se analiza en Sevilla et al. (2024) de lectura y escritura en educación Primaria —DyctectiveU—, construidas desde el principio con la participación activa de docentes, alumnado y familias. Esta cocreación mejora la usabilidad, la aceptación en el aula y la eficacia educativa. Por tanto, retirar las herramientas tecnológicas no solo puede limitar oportunidades de aprendizaje, sino también debilitar las soluciones que han demostrado ser eficaces e inclusivas.

Para promover una adopción equitativa de la inteligencia artificial en la educación, es necesario ir más allá de la dotación tecnológica y adoptar un enfoque integral de justicia digital. En este sentido, los desarrolladores del software deberían incluir la participación activa del profesorado y del alumnado en el diseño y evaluación de las herramientas de IA, con el fin de garantizar su adecuación a contextos educativos diversos y evitar posibles sesgos. Asimismo, en el contexto

de España, con cinco lenguas cooficiales, resulta fundamental desarrollar contenidos multilingües y culturalmente relevantes y diversos y evitar posibles sesgos, tan frecuentes en las herramientas actuales. Es decir, la acción debe realizarse desde la creación: desde el punto de vista del diseño y del contenido de la herramienta. Esto requiere fomentar alianzas público-privadas sustentadas en compromisos éticos que no se limiten únicamente a facilitar el acceso económico a herramientas ya desarrolladas, sino que promuevan una inclusión más profunda desde las fases iniciales de diseño y desarrollo. Este enfoque garantiza que las soluciones de IA educativas integren criterios de equidad en su concepción, contenidos y funcionalidades, de modo que respondan a las necesidades de contextos diversos y no reproduzcan sesgos estructurales desde su origen.

Para evitar que la IA se convierta en un nuevo factor de exclusión educativa, proponemos una serie de medidas prioritarias. En primer lugar, es fundamental garantizar una dotación mínima de recursos digitales en todos los centros educativos, especialmente en los más vulnerables. Esto incluye tanto la infraestructura como el acceso a dispositivos y conectividad. En segundo lugar, es importante formar al profesorado en el uso pedagógico, crítico y ético de las herramientas de IA. Solo así será posible integrar estas tecnologías en el aula de forma efectiva, aprovechando su potencial sin renunciar a los valores fundamentales de la educación. Asimismo, deben establecerse marcos normativos que regulen su uso, protejan la privacidad del alumnado y prevengan sesgos algorítmicos.

Finalmente, es necesario fomentar una cultura de aprendizaje centrada en la autonomía, el pensamiento crítico y la creatividad (véase el capítulo de Lastra y Flynn en este mismo número de Papeles), que complemente —en lugar de sustituir— las capacidades humanas con las de la IA. La IA no reemplazará a la educación, pero transformará profundamente cómo, dónde y para qué aprendemos. El reto de los sistemas educativos será, por tanto, no solo adoptar la tecnología, sino hacerlo de manera justa, contextualizada y con una participación activa de toda la comunidad educativa.

## BIBLIOGRAFÍA

- Bacher-Hicks, A., Goodman, J. y Mulhern, C. (2021).** Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, Vol. 193. <https://doi.org/10.1016/j.jpubeco.2020.104345>
- Choi, E. (2021).** The limitations of deep learning in mathematical reasoning. *arXiv preprint arXiv:2111.00331*.
- Education Endowment Foundation (EEF (2024)).** *ChatGPT in lesson preparation: A Teacher Choices trial*. National Foundation for Educational Research (NFER). <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/chatgpt-in-lesson-preparation-a-teacher-choices-trial>
- Elliott, S. (2017).** Computer-based assessment of adult competencies: An overview of the PIAAC approach. *OECD Education Working Papers*, N<sup>o</sup>. 168. OECD Publishing. <https://doi.org/10.1787/5jln3rj01w8s-en>
- Guan, Y., Liang, Z. y Tena, J. (2025).** Uso de internet en la infancia y adolescencia: ¿beneficioso o perjudicial? *Papeles de Economía Española*, 184.
- Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D. y Steinhardt, J. (2021).** Measuring mathematical problem solving with the MATH dataset. *arXiv preprint arXiv:2103.03874*.
- Holmes, W. (2023).** Artificial intelligence and education: Promises and perils. En *The Cambridge Handbook of Education and AI* (33-55). Cambridge University Press.
- Malik, R., Abdi, D., Wang, R. y Demszky, D. (2025).** Scaffolding Middle-School Mathematics Curricula With Large Language Models. *EdWorkingPaper*: 24-1028. Annenberg Institute at Brown University. <https://doi.org/10.26300/b47y-mh41>
- Marcenaro-Gutiérrez, O. D. y López-Agudo, L. A. (2025).** Sobre el uso de recursos digitales y el rendimiento educativo: de las políticas educativas a la práctica. *Papeles de Economía Española*, 184.
- Meyer, K., Page, L. C., Mata, C., Smith, E. N., Walsh, B. T., Fifield, C. L., Tyson, M., Eremionkhale, A., Evans, M., Frost, S. y Jung, E. E. (2024).** Let's Chat: Leveraging Chatbot Outreach for Improved Course Performance. *EdWorkingPaper* No. 22-564. Annenberg Institute at Brown University. <https://doi.org/10.26300/es6b-sm82>
- OCDE (2023a).** *Putting AI to the test: Large Language Models and their relevance for education*. OECD Publishing. [https://www.oecd.org/en/publications/putting-ai-to-the-test\\_2c297e0b-en.html](https://www.oecd.org/en/publications/putting-ai-to-the-test_2c297e0b-en.html)
- OCDE (2023b).** *Is education losing the race with technology? AI's progress in maths and reading*. OECD Publishing. [https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology\\_73105f99-en.html](https://www.oecd.org/en/publications/is-education-losing-the-race-with-technology_73105f99-en.html)
- OCDE (2024).** *PISA 2022 Results (Volume II): Learning During - and From - Disruption*. OECD Publishing. <https://doi.org/10.1787/a97db61c-en>
- OECD (2025).** *Trends Shaping Education 2025*, OECD Publishing, Paris, <https://doi.org/10.1787/ee6587fd-en>
- Petrilli, M. J. (2024).** Next-Gen Classroom Observations, Powered by AI: Let's go to the videotape to improve instruction and classroom practice. *Education Next*, 24(4), 80-82.
- Sanz, I., Cuerdo, M. y Doncel, L. M. (2021).** The use of digital educational resources in times of COVID-19. *Social Media + Society*, Vol. 7/3. SAGE Publications, <https://doi.org/10.1177/20563051211049246>.
- Varsik, S. y Vosberg, L. (2024).** The potential impact of Artificial Intelligence on equity and inclusion in education. *OECD Artificial Intelligence Papers*, No. 23. Paris: OECD Publishing. <https://doi.org/10.1787/15df715b-en>