



E-Learning y agentes de IA en la educación superior: un marco teórico para la adopción, el valor educativo y la implementación responsable.

Dr. Luis F. Toro Dupouy

Investigador de OBS Business School

Septiembre, 2026

Partners académicos:



UNIVERSITAT DE
BARCELONA

unie* Universidad



Autor

Dr. Luis F. Toro Dupouy

Investigador de
OBS Business School

Director de programas académicos de OBS Business School (*School of Business Administration & Leadership*). Ha desarrollado y dirigido programas máster y de grado en gestión empresarial, dirección de marketing y gestión comercial, investigación de mercados y análisis de clientes, *business intelligence*, *international business*, Executive MBA y Global MBA. Actualmente es profesor del Departamento de Administración de Empresas y del Departamento de Investigación de Mercados y Métodos Cuantitativos de ESIC University. También colabora como docente en otras universidades e instituciones entre las que destacan EAE Business School, Universitat Carlemany, Fundación Ortega y Gasset Gregorio Marañón y el Institute for American Universities. Ha sido director académico de EAE Business School (campus Madrid) y profesor visitante en diversas universidades y *business schools* de Europa y de América. Conferenciante internacional que ha impartido seminarios, conferencias y programas de formación *in company* en diversos países de Europa y de Latinoamérica. Experto en *e-learning* y *business simulation*. Autor de diversas publicaciones sobre estrategia empresarial, gestión de la diversidad, gestión intercultural, marketing, *e-learning* y tecnología educativa.

Doctor en Ciencias Económicas y Empresariales por la Universidad de Granada (España). Master of Business Administration (MBA) por The University of Saint Thomas (Houston, Texas). Máster en Dirección de Marketing y Gestión Comercial por ESIC Business & Marketing School (Madrid, España). Licenciado en Psicología por la Universidad Central de Venezuela (Caracas, Venezuela).



ÍNDICE DE CONTENIDOS



Introducción.

Resumen ejecutivo

PÁG. 5

01

Capítulo 1.

Introducción y justificación

PÁG. 7

02

Capítulo 2.

Metodología de la revisión de fuentes secundarias

PÁG. 14

03

Capítulo 3.

Fundamentos conceptuales

PÁG. 18

3.1. El aprendizaje en línea como ecosistema educativo

PÁG. 19

3.2. De la inteligencia artificial a la IA generativa

PÁG. 20

3.3. Agentes de IA e IA Agentica

PÁG. 21

04

Capítulo 4.

Aplicaciones y creación de valor en el aprendizaje electrónico

PÁG. 25

4.1. Aplicaciones dirigidas al estudiante: Personalización y agencia

PÁG. 26

4.2. Aplicaciones orientadas al profesorado: productividad
y aumento pedagógico

PÁG. 27

4.3. Aplicaciones orientadas a la institución: eficiencia
y diferenciación institucional

PÁG. 29

4.4. Evaluación del valor a través del riesgo y la supervisión

PÁG. 30



ÍNDICE DE CONTENIDOS

05

Capítulo 5.

Perspectivas teóricas

PÁG. 32

- 5.1. Determinantes de la aceptación y uso individual PÁG. 33
- 5.2. Alineación de tareas y rendimiento del sistema PÁG. 34
- 5.3. La perspectiva pedagógica: procesos de aprendizaje y preparación del profesorado PÁG. 36
- 5.4. Contexto institucional y de gobernanza PÁG. 37
- 5.5. Argumento integrativo y relaciones multidimensionales PÁG. 37

06

Capítulo 6.

Beneficios, riesgos y condiciones de contorno

PÁG. 39

- 6.1. Beneficios potenciales y evidencia emergente PÁG. 40
- 6.2. Riesgos principales y decisiones de alta consecuencia PÁG. 41
- 6.3. Condiciones de frontera y tensiones contextuales PÁG. 43

07

Capítulo 7.

Marco Teórico Integrador Propuesto y Agenda de Investigación Futura

PÁG. 46

- 7.1. Componentes del Marco Integrador PÁG. 48
- 7.2. Moderadores clave y condiciones límite PÁG. 50
- 7.3. Agenda de Investigación Futura PÁG. 50

08

Capítulo 8.

Implicaciones para la gestión de las escuelas de negocios

PÁG. 52

9

Capítulo 10.

Conclusiones

PÁG. 56

Introducción.

Resumen ejecutivo



RESUMEN

E-Learning y agentes de IA en la educación superior: una hoja de ruta para la adopción estratégica

Fuente: Elaboración propia

CIMIENTOS Y ANÁLISIS

1 Introducción y justificación



Explica por qué los agentes de IA importan en el e-learning y define el problema central.

2 Metodología de la revisión de fuentes secundarias



Describe cómo se identificaron y sintetizaron la literatura y la evidencia institucional.

3 Cimientos conceptuales



Establece definiciones claras para el e-learning, la IA, la IA generativa y los agentes de IA autónomos.

CREACIÓN DE VALOR Y TEORÍA

4 Aplicaciones y creación de valor



Identifica aplicaciones orientadas a estudiantes, profesores e instituciones y sus propuestas de valor.

5 Perspectivas teóricas



Presenta las teorías primarias utilizadas para analizar la adopción, los patrones de uso y los resultados educativos.

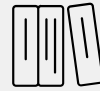
6 Beneficios, riesgos y condiciones límite



Destaca oportunidades y riesgos, identificando las condiciones específicas donde los agentes de IA generan valor real.

MARCO E IMPLEMENTACIÓN ESTRATÉGICA

7 Marco teórico integrador propuesto



Integra hallazgos en un modelo conceptual con una agenda de investigación para estudios futuros.

8 Implicaciones gerenciales para escuelas de negocios



Traduce los hallazgos de la investigación en recomendaciones operativas y estratégicas para el liderazgo institucional.

9 Conclusiones



Refuerza la necesidad de un enfoque basado en la evidencia y responsable para la adopción de la IA.



MENSAJE CENTRAL

Los agentes de IA no mejoran automáticamente el e-learning. Su éxito depende estrictamente de la alineación pedagógica, la idoneidad de la tarea, las capacidades del usuario, la confianza, la preparación institucional, la supervisión humana y la gobernanza responsable.

01

Capítulo 1.

Introducción y justificación



El panorama de la educación superior está experimentando un profundo cambio estructural, ya que el e-learning pasa de ser una modalidad periférica para convertirse en una prioridad estratégica central. En este contexto, es fundamental distinguir entre el aprendizaje en línea diseñado intencionalmente, que utiliza la tecnología para transformar metodologías y ecosistemas, y la enseñanza remota de emergencia, que sirvió como una medida provisional necesaria durante las disrupciones globales, pero que a menudo simplemente conservó las prácticas educativas existentes (OECD, 2021; OECD, 2023). Mientras que la primera representaba una transformación digital fundamental destinada a mejorar la personalización y el apoyo al alumnado, la segunda replicaba con frecuencia los métodos tradicionales de pizarra en formato digital, lo que a menudo resultaba en fatiga visual y estrés por la adaptación (OECD, 2021; OECD, 2023). Para las escuelas de negocios y universidades, el imperativo estratégico ahora radica en integrar la tecnología no solo como herramienta de imposición, sino también como catalizador de una experiencia educativa más eficiente y basada en datos (OECD, 2023). Como se muestra en la Figura 1, el aumento exponencial de la participación en la IA refleja un cambio estructural en las prioridades globales de investigación y educación.

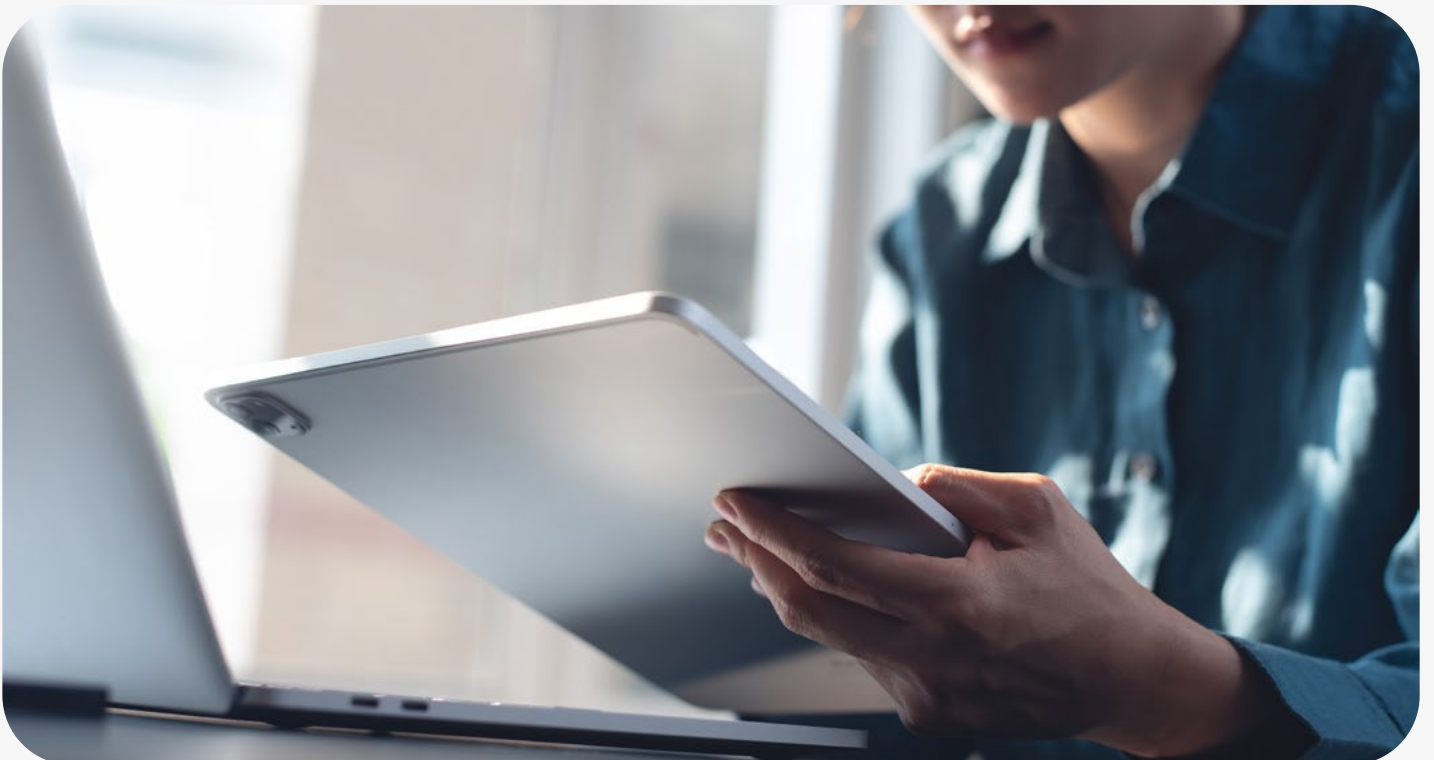
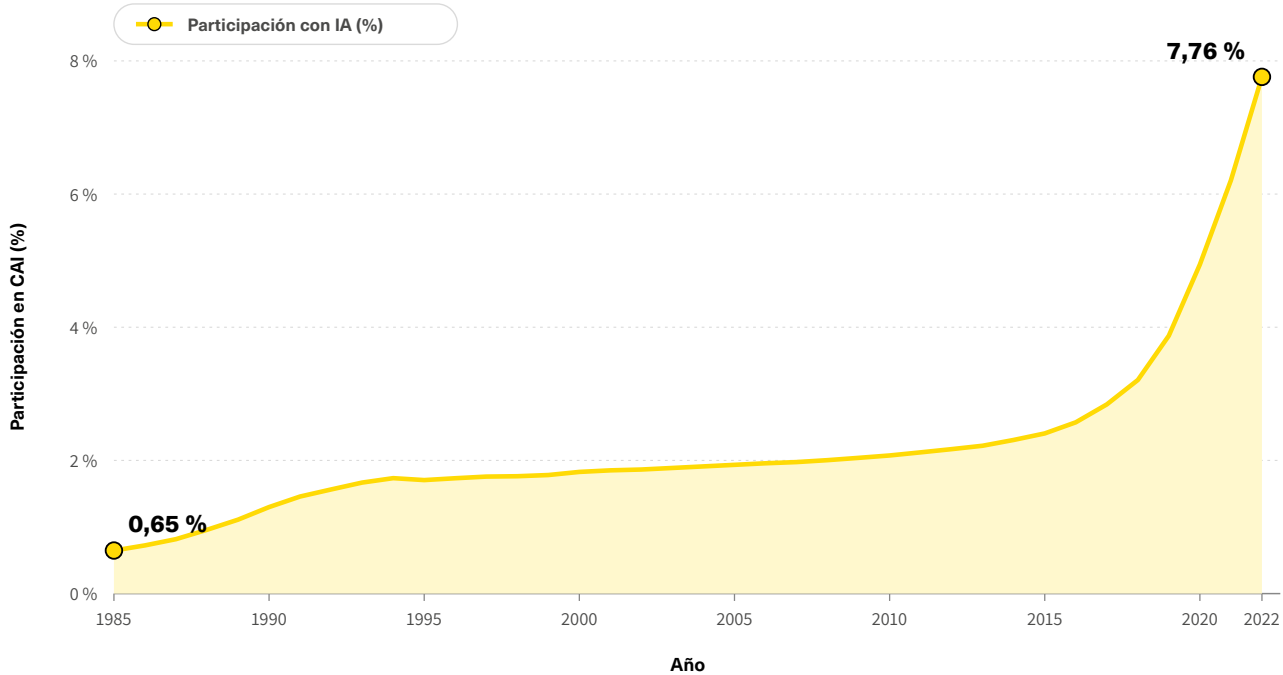


FIGURA 1

Cambio en la participación de la IA en todos los campos científicos

Fuente: OECD Digital Education Outlook 2026.



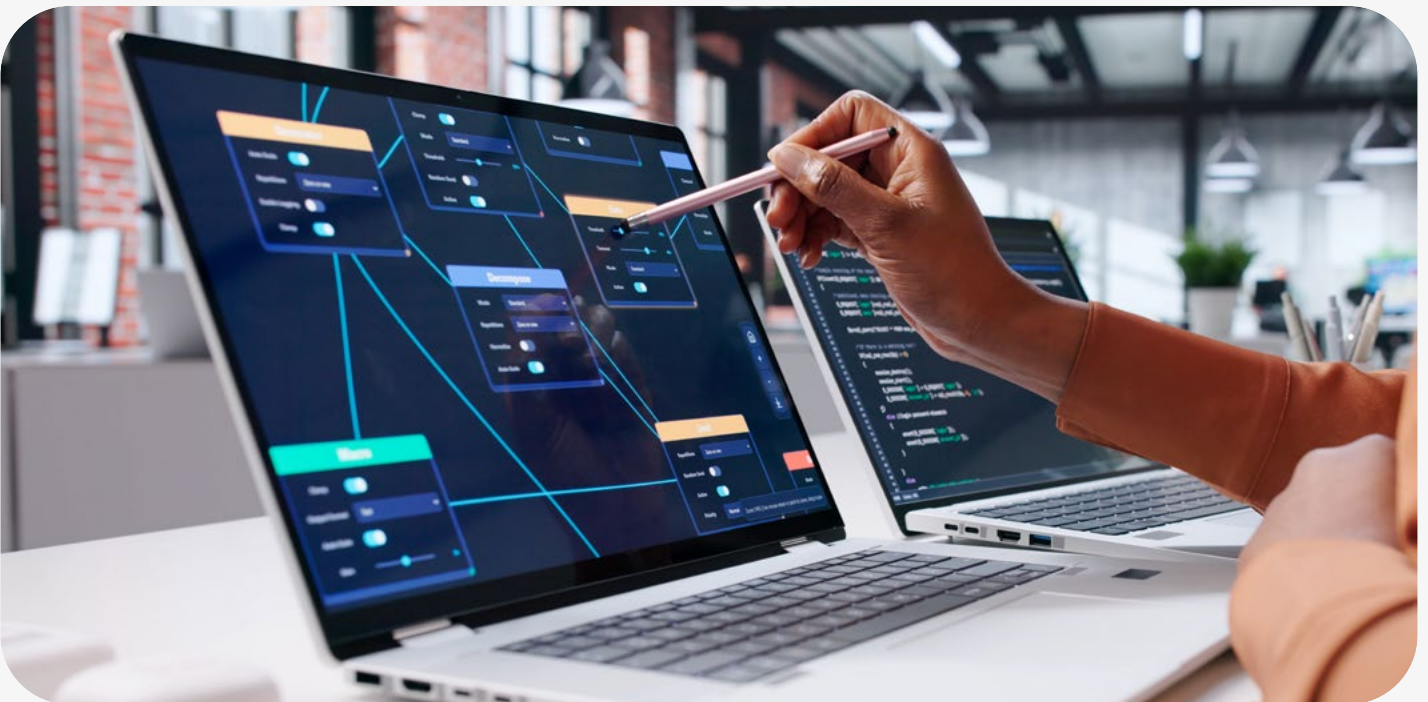
La integración de la inteligencia artificial en la educación (AIEd) representa la última fase de esta evolución y se basa en décadas de investigación en análisis del aprendizaje, sistemas inteligentes de tutoría (ITS), aprendizaje adaptativo y retroalimentación automatizada (Ouyang et al., 2022; OECD, 2021; Zawacki-Richter et al., 2019). El ITS clásico, que surgió en los años 70, utilizaba sistemas expertos basados en reglas para simular tutorías individuales, proporcionando apoyo paso a paso para la resolución de problemas basado en el conocimiento experto específico de cada área (Holmes et al., 2019; OECD, 2026; Zawacki-Richter et al., 2019). Los avances posteriores en el análisis de aprendizaje han permitido a las instituciones utilizar datos a gran escala para comprender mejor las trayectorias de los estudiantes, proporcionando paneles que apoyan la toma de decisiones del profesorado e identifican a los estudiantes en riesgo de deserción.

(OECD, 2021; Zawacki-Richter et al., 2019). Los sistemas de aprendizaje adaptativo perfeccionaron esto al diagnosticar las brechas de conocimiento de los estudiantes y al ajustar automáticamente la dificultad o la secuencia de tareas para optimizar las rutas de aprendizaje individuales (OECD, 2021; OECD, 2026).

Ahora se está produciendo una transición tecnológica significativa desde la IA generativa, que genera automáticamente contenido nuevo, como textos, imágenes o código, en respuesta a indicaciones en lenguaje natural, hacia sistemas de "agentes" más sofisticados (Parasuraman et al., 2000; Miao & Holmes, 2023; OECD, 2026). Mientras que la IA tradicional identifica patrones y la IA generativa produce representaciones simbólicas del pensamiento humano, los agentes de IA se definen como sistemas autónomos capaces de planificar, recuperar información, utilizar herramientas digitales, ejecutar tareas de varios pasos y monitorizar sus propios resultados dentro de límites predefinidos (OECD, 2026; UNESCO, 2021). Estos agentes funcionan como mucho más que simples interfaces conversacionales; pueden actuar como "conductores" que pilotan múltiples modelos y herramientas de IA hacia un objetivo determinado, simulando eficazmente el comportamiento de un equipo de investigación especializado o de un socio instruccional multifacético (OECD, 2026).

Para la educación superior y las escuelas de negocios, la relevancia de los agentes de IA abarca todo el ciclo de vida de los estudiantes y de los profesores. En roles docentes, agentes como "JeepyTA" apoyan al profesorado proporcionando retroalimentación formativa 24/7 sobre artefactos complejos de los estudiantes, respondiendo a consultas logísticas y facilitando discusiones en profundidad en los foros del curso (OECD, 2026). Más allá del aula, los sistemas agentivos están transformando la orientación académica y los servicios al estudiante al ofrecer cursos personalizados y recomendaciones de especialidad basadas en patrones complejos de matrícula y en objetivos profesionales (OECD, 2026). Administrativamente, los agentes agilizan los flujos de trabajo, mapean equivalencias de cursos a través de los límites institucionales, automatizan la evaluación de expedientes académicos y ayudan a crear ítems de evaluación estandarizados, anclados en los planes de estudio existentes (OECD, 2026).

A pesar de estas capacidades, el problema central de investigación identificado en este informe técnico es que los agentes de IA no mejoran de forma intrínseca el aprendizaje en línea; su valor depende de la calidad tecnológica, la alineación pedagógica, las capacidades de los usuarios, la confianza, la preparación institucional, la supervisión humana y la gobernanza responsable (Ouyang et al., 2022; Miao & Holmes, 2023; OECD, 2026). La evidencia sugiere que los buenos resultados educativos no se logran simplemente desplegando tecnologías avanzadas; más bien, la IA debe ser "pedagogía primero", basada en teorías de aprendizaje establecidas, como el constructivismo social o el andamiaje vygotskiano (Ouyang et al., 2022; OECD, 2026). Además, el impacto real está mediado por el "ajuste tarea-tecnología", en el que las capacidades específicas de la IA deben ajustarse a las demandas cognitivas y metacognitivas de la tarea educativa (Ogunleye et al., 2024; OECD, 2026; Long and Magerko, 2020). Sin un diseño cuidadoso que implique a las personas en el bucle, los agentes de IA corren el riesgo de provocar "pereza metacognitiva" o descarga cognitiva, en la que estudiantes y profesores dependen en exceso de resultados automatizados a costa del aprendizaje profundo y del razonamiento crítico (OECD, 2026).



Persisten carencias significativas en la literatura actual, que a menudo está fragmentada y se centra en exceso en las intenciones de adopción a corto plazo, más que en el impacto educativo a largo plazo (Zawacki-Richter et al., 2019; Ogunleye et al., 2024). La mayoría de las implementaciones empíricas de IA en la educación son pilotos a corto plazo, con una persistente falta de evidencia longitudinal sobre cómo estas herramientas afectan el conocimiento y la disposición de los estudiantes a lo largo del tiempo (Zawacki-Richter et al., 2019; Ogunleye et al., 2024; Roe & Perkins, 2024). Además, en la literatura no existe una distinción clara entre *chatbots* simples y agentes autónomos, y hay una notable falta de reflexión crítica sobre la preparación institucional, el sesgo algorítmico y la erosión de la agencia humana (Zawacki-Richter et al., 2019). Las cuestiones relacionadas con la privacidad de los datos, la responsabilidad por errores generados por la IA y la dependencia a largo plazo de los alumnos del soporte automatizado siguen, en gran medida, sin resolverse (Zawacki-Richter et al., 2019; Park et al., 2025; Martin et al., 2020; Miao & Cukurova, 2024).

Para abordar estos desafíos, este informe técnico justifica la necesidad de un marco teórico integrador que reúna múltiples perspectivas analíticas. Dicho marco debe incorporar el Modelo de Aceptación de Tecnología (TAM) y la Teoría Unificada de Aceptación y Uso de la Tecnología (UTAUT) para comprender la percepción de utilidad y de facilidad de uso por parte del usuario (Ogunleye et al., 2024; Park et al., 2025; Batista et al., 2024). Además, debe integrar los modelos de Ajuste Tarea–Tecnología (TTF) y de Éxito de Sistemas de Información para evaluar cómo la calidad técnica y la alineación de tareas impulsan el rendimiento (Ogunleye et al., 2024; Parasuraman et al., 2000; Al-Fraihat et al., 2020). Desde un punto de vista pedagógico, el marco se basa en el Aprendizaje Autorregulado (SRL) y en la Comunidad de Investigación (Col) para preservar la agencia del alumno y su presencia social (OECD, 2026; Ogunleye et al., 2024; Park et al., 2025). Además, los modelos de Conocimiento Pedagógico Tecnológico de Contenidos (TPACK) y Tecnología–Organización–Entorno (TOE) son esenciales para comprender la integración entre el profesorado y el contexto institucional (Ogunleye et al., 2024; Ng et al., 2021; OECD, 2026; Mishra et al., 2006). Por último, son necesarias perspectivas sobre la confianza

en la automatización y en la IA responsable para garantizar que los sistemas de agentes se implementen con la debida confianza y con salvaguardas éticas (Parasuraman et al., 2000; Ogunleye et al., 2024; Glikson y Woolley, 2020).

Este informe utiliza el análisis de fuentes secundarias para sintetizar diversas perspectivas y desarrollar un marco conceptual que sirva de hoja de ruta para futuras investigaciones primarias. Al articular las condiciones necesarias para el éxito, desde el diseño de interfaces hasta las estructuras de gobernanza, este marco proporciona a los líderes universitarios y gestores de tecnología educativa la orientación basada en la evidencia necesaria para la implementación responsable de agentes de IA. Las siguientes secciones detallarán las metodologías específicas empleadas para analizar la literatura académica y los informes institucionales que fundamentan este marco.



02

Capítulo 2.

Metodología de la revisión de fuentes secundarias



Este informe está diseñado como una revisión integradora estructurada de la literatura, basada exclusivamente en fuentes secundarias (Chen et al., 2020; Long & Magerko, 2020). No constituye una revisión sistemática completa ni un metaanálisis formal, que requeriría presentar tamaños de efecto específicos a partir de un conjunto fijo de datos empíricos (Ogunleye et al., 2024; OECD, 2026). En cambio, proporciona una síntesis transparente de la literatura académica, institucional y de políticas relevantes para construir un marco teórico integral para la adopción e implementación de agentes de IA en la educación superior (Long & Magerko, 2020; Toro-Dupouy & Erdmann, 2026; Zawacki-Richter et al., 2019).

La revisión se basa en una amplia variedad de fuentes de alta calidad, incluyendo artículos revisados por pares, revisiones sistemáticas previas, trabajos teóricos fundamentales y estudios empíricos publicados (Panadero, 2017; Ogunleye et al., 2024; Ng et al., 2021). Para proporcionar contexto institucional y político, la metodología incorpora informes y directrices de organismos internacionales acreditados, como la UNESCO y la OCDE (Miao & Holmes, 2023; OECD, 2023; Zawacki-Richter et al., 2019). La búsqueda se realizó en las principales bases de datos y repositorios académicos, incluyendo Scopus, Web of Science, ERIC, ScienceDirect y Google Scholar, así como en sitios web institucionales de documentación de políticas (Ogunleye et al., 2024; Park et al., 2025; Ouyang et al., 2022; Roe & Perkins, 2024).

Se seleccionaron términos de búsqueda representativos para captar la intersección entre tecnología y pedagogía: "e-learning", "aprendizaje en línea", "inteligencia artificial en la educación", "IA generativa", "agentes de IA", "IA agente", "adopción de IA", "alfabetización en IA", "confianza en la automatización", "IA responsable" y "educación superior" (Ogunleye et al., 2024; Park et al., 2025; Ouyang et al., 2022; Roe & Perkins, 2024). El informe mantiene un enfoque destacado en las publicaciones de 2015 a 2026, con especial énfasis en el período transformador posterior a la publicación, a finales de 2022, de grandes modelos de lenguaje (Park et al., 2025; Ogunleye et al., 2024; Batista et al., 2024; OECD, 2026). Simultáneamente, la revisión incluye trabajos fundamentales antiguos para aportar rigor teórico, específicamente aquellos que establecen el Modelo de Aceptación de Tecnología (TAM), la Teoría Unificada de

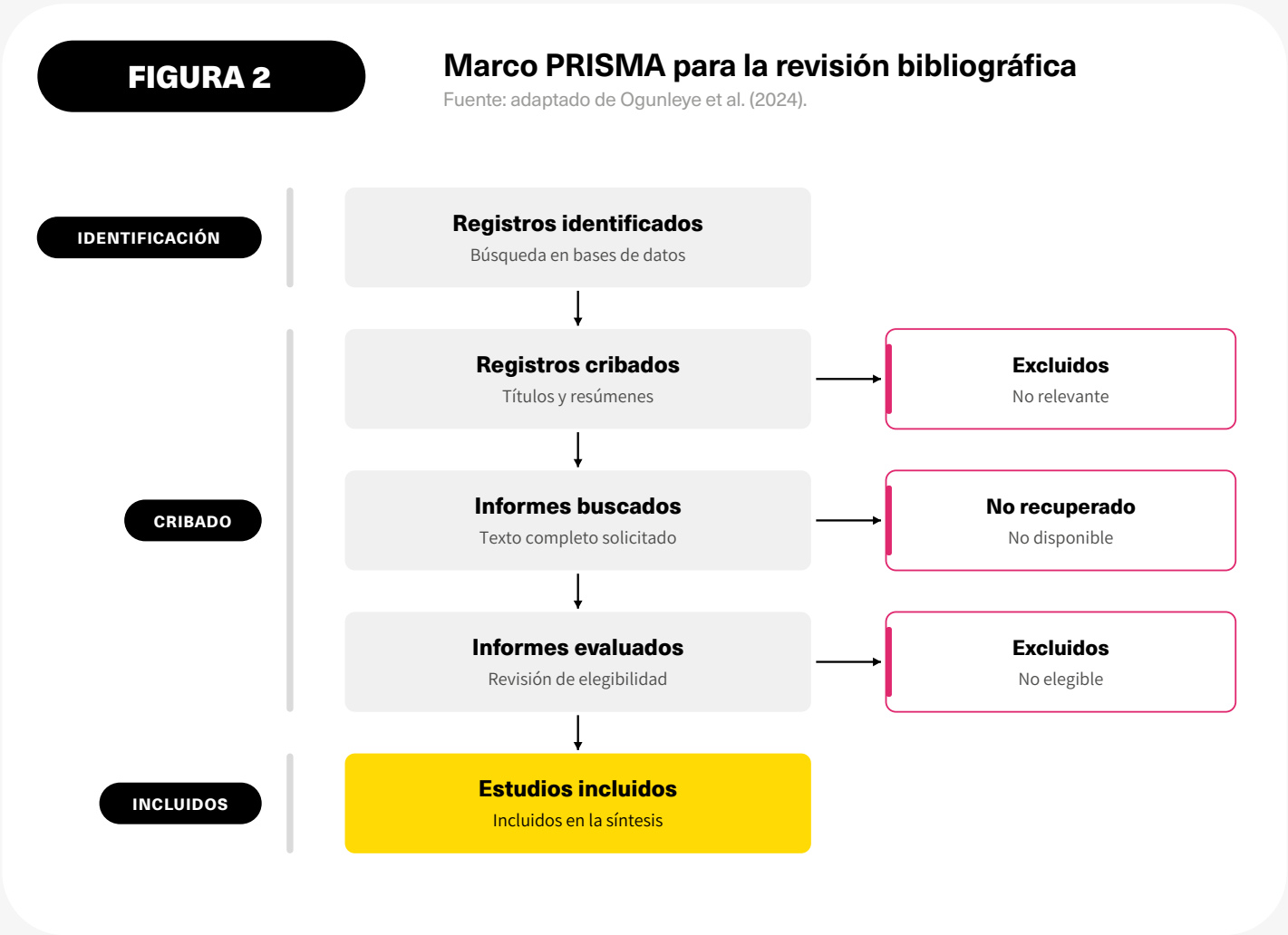
Aceptación y Uso de la Tecnología (UTAUT), el Ajuste Tarea–Tecnología (TTF), el Conocimiento Tecnológico de Contenidos Pedagógicos (TPACK), el marco de Comunidad de Investigación (Col), el Aprendizaje Autorregulado (SRL) y el marco Tecnología–Organización–Entorno (TOE) (Panadero, 2017; Toro-Dupouy & Erdmann, 2026; Mishra & Koehler, 2006).

Los criterios de inclusión se definieron estrictamente para garantizar la relevancia en la educación superior y en los ecosistemas de e-learning, centrándose en aplicaciones de IA, impulsores de la adopción, resultados de aprendizaje, confianza en la automatización, preparación institucional, gobernanza e interacción humano–IA (Ouyang et al., 2022; Zawacki-Richter et al., 2019). Por el contrario, la revisión excluyó material de poca relevancia para el sector educativo, un enfoque puramente técnico o algorítmico, contenido promocional o comercial, y obras que carecían de suficiente credibilidad académica o de estatus de revisión por pares (Park et al., 2025; Ouyang et al., 2022; Lo, 2023).

La literatura se organizó mediante una síntesis temática que categoriza la base de evidencia en áreas centrales: la evolución del e-learning, la transición de la IA generativa a la agentiva, los determinantes de adopción, la integración pedagógica, la confianza de los usuarios, la alfabetización en IA, la preparación institucional, los riesgos y la gobernanza de los sistemas automatizados (Park et al., 2025; Batista et al., 2024).

Hay que reconocer varias limitaciones. El campo se caracteriza por un rápido cambio tecnológico, lo que provoca una escasez de evidencia empírica longitudinal, en particular sobre los agentes de IA (Zawacki-Richter et al., 2019). Además, la terminología es inconsistente en la literatura actual y predomina el uso de estudios exploratorios o piloto en lugar de despliegues a gran escala (OECD, 2026; Roe & Perkins, 2024; Batista et al., 2024). Finalmente, como revisión de fuentes secundarias, este estudio carece de datos primarios y se basa en la validez de las fuentes subyacentes. A pesar de estas limitaciones, el enfoque elegido proporciona la base de evidencia rigurosa necesaria para desarrollar un marco teórico integrador y formular propuestas significativas para futuras investigaciones primarias (Ogunleye et al., 2024; OECD, 2026).

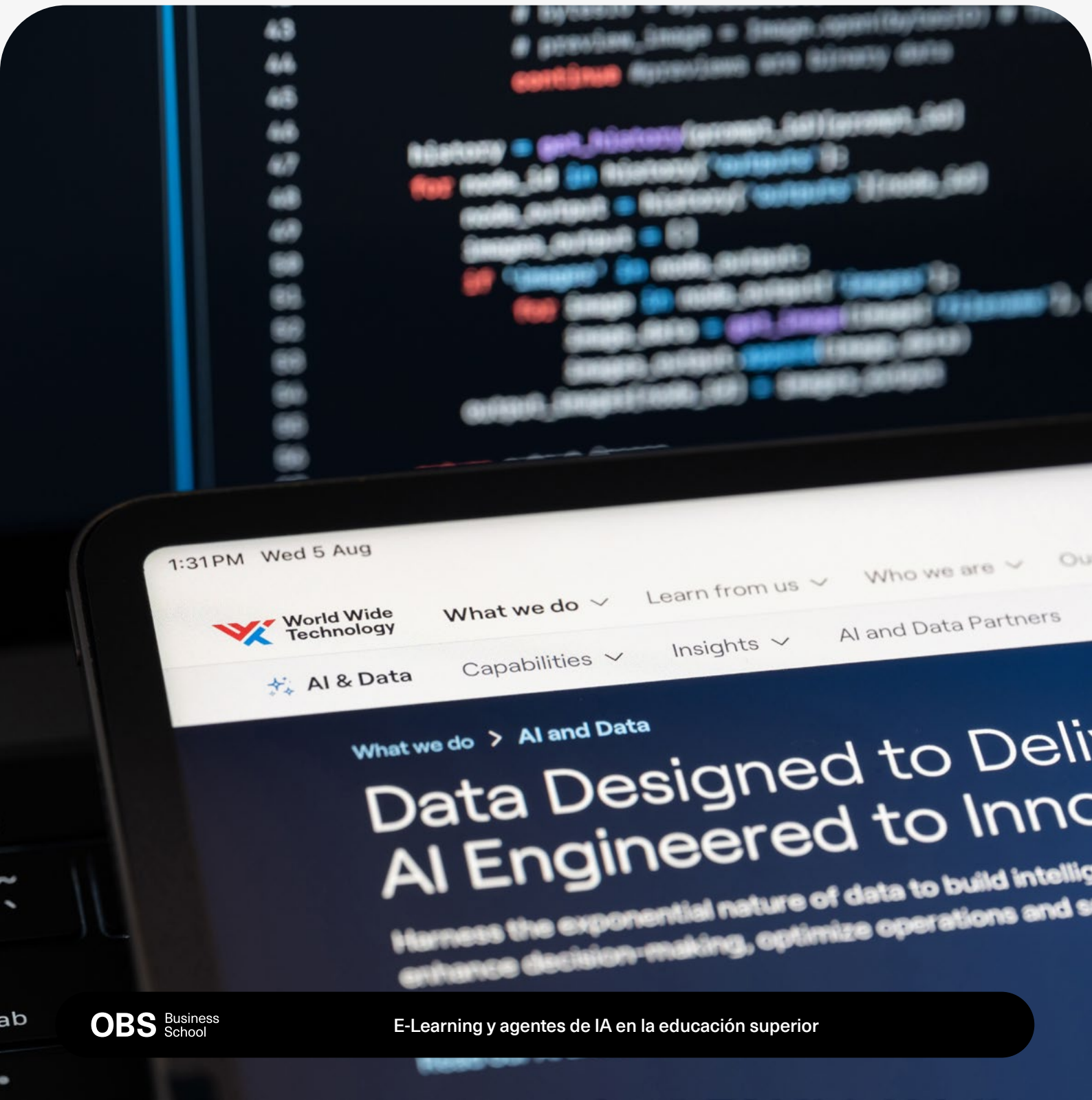
La síntesis siguió un proceso de selección riguroso, basado en el marco PRISMA que se muestra en la Figura 2. Este avanza a través de tres etapas —identificación de registros mediante búsqueda en bases de datos, cribado de títulos/resúmenes e informes de texto completo (con exclusiones registradas en cada punto de control), e inclusión final de los estudios en la síntesis— reflejando la estructura utilizada en la revisión sistemática de Ogunleye et al. (2024) sobre la IA generativa en la enseñanza y el aprendizaje en la educación superior.



03

Capítulo 3.

Fundamentos conceptuales



3.1

La integración de la inteligencia artificial (IA) en la educación superior no ocurre en el vacío, sino que constituye la última fase de la evolución del aprendizaje digital. Para evaluar el impacto potencial de los agentes de IA, es necesario establecer una base conceptual clara sobre la naturaleza del e-learning, la evolución de la IA predictiva a la generativa y las características funcionales específicas de los sistemas de agentes.

El aprendizaje en línea como ecosistema educativo

El e-learning es más que un mecanismo de imposición; se define como un ecosistema educativo complejo que integra contenido digital, interacción síncrona y asincrónica, sistemas de gestión del aprendizaje (LMS) y mecanismos sofisticados de evaluación, retroalimentación y análisis (Al-Fraihat et al., 2020; Salloum et al., 2019). En esencia, el e-learning aprovecha los medios electrónicos y los dispositivos para mejorar la disponibilidad de la formación y de la comunicación, facilitando formas novedosas de adquirir y comprender conocimientos (Salloum et al., 2019). El éxito en este ecosistema es multidimensional y se determina por la sinergia entre la calidad del sistema (incluidas las dimensiones técnicas, educativas y de soporte), la calidad de la información y las características tanto de los alumnos como de los instructores (Al-Fraihat et al., 2020).

Debe mantenerse una distinción conceptual crítica entre el aprendizaje electrónico diseñado intencionalmente y la enseñanza remota de emergencia (ERT). Aunque el aprendizaje online, diseñado intencionalmente, busca transformar las metodologías docentes y el apoyo al alumnado mediante una transformación digital fundamental, la ERT sirvió como una medida temporal durante las disrupciones globales (Toro-Dupouy & Erdmann, 2026; OECD, 2021; OECD, 2023). La ERT a menudo simplemente conservaba las prácticas educativas existentes, replicando métodos tradicionales en formato digital, mientras que una verdadera transformación digital reinventaba el ecosistema educativo para priorizar la personalización y el compromiso (OECD, 2021; OECD, 2023). La eficacia educativa de estos ecosistemas

3.2

depende no solo de las herramientas en sí, sino también del diseño instruccional: de cómo se utiliza la tecnología para crear un entorno que fomente la presencia cognitiva, social y docente (Martin et al., 2000).

De la inteligencia artificial a la IA generativa

La inteligencia artificial en la educación (AIEd) ha evolucionado a través de varios paradigmas. Los primeros sistemas se basaban principalmente en reglas (o "IA de la vieja usanza"), utilizando lógica experta predefinida para simular comportamientos inteligentes (UNESCO, 2021; Zawacki-Richter et al., 2019). Estas iniciativas pioneras apoyaron el desarrollo de Sistemas de Tutoría Inteligente (ITS), que buscan simular la tutoría individual humana y proporcionar apoyo paso a paso para la resolución de problemas, basado en modelos específicos de cada dominio (Holmes et al., 2019; Zawacki-Richter et al., 2019).

El campo avanzó significativamente con la llegada del aprendizaje automático, un subcampo de la IA capaz de descubrir patrones en los datos para realizar predicciones y clasificaciones sin estar explícitamente programado para cada escenario (Zawacki-Richter et al., 2019). Esto llevó al auge de la analítica del aprendizaje y de los sistemas de aprendizaje adaptativo, que utilizan datos a gran escala para diagnosticar las brechas de conocimiento de los estudiantes, proporcionar retroalimentación automática y recomendar vías de aprendizaje personalizadas (Chen et al., 2020; OECD, 2021). Hasta hace poco, la mayoría de la IA en la educación era predictiva y producía clasificaciones o valoraciones, como identificar a estudiantes en riesgo de abandono mediante el análisis de relaciones en grandes volúmenes de datos históricos (OECD, 2026).

El lanzamiento de los Grandes Modelos de Lenguaje (LLM) y de la IA Generativa (GenAI) supuso un cambio de paradigma. A diferencia de la IA predictiva, las herramientas GenAI producen automáticamente nuevo contenido, como texto, código o imágenes, en respuesta a indicaciones en lenguaje natural (Miao & Holmes, 2023; OECD, 2026). Estos sistemas utilizan arquitecturas *transformer* para aprender relaciones estadísticas

3.3

en conjuntos de datos, lo que les permite resumir, explicar y adaptar contenidos complejos con un nivel de fluidez conversacional antes inalcanzable (Ogunleye et al., 2024; Tlili et al., 2023; Kasneci et al., 2023). Mientras que la IA predictiva se centra en identificar "lo que es" o "lo que probablemente será", la IA generativa se centra en "crear lo que podría ser" a partir de patrones aprendidos (OECD, 2026).

Agentes de IA e IA Agentica

Aunque la mayoría de los usuarios han experimentado con GenAI a través de *chatbots* simples, que principalmente imitan conversaciones humanas y responden a indicaciones específicas, la tecnología está evolucionando hacia agentes de IA (OECD, 2026; Tlili et al., 2023). Un agente de IA se define como un sistema autónomo que puede perseguir objetivos específicos mediante un ciclo sofisticado de recuperación de información, planificación, descomposición de tareas, gestión de la memoria y uso de herramientas (OECD, 2026; Miao et al., 2024). A diferencia de los *chatbots* simples, los agentes pueden ejecutar acciones de varios pasos, monitorizar sus propios resultados y adaptar su comportamiento dentro de límites predefinidos para lograr un objetivo determinado (OECD, 2026).

Para analizar su papel en la educación superior, es esencial distinguir entre diferentes arquitecturas de sistemas:

- **Sistemas de IA tradicionales:** centrados en tareas específicas y concretas, como la clasificación o la predicción, sin generar contenido (Zawacki-Richter et al., 2019; Long & Magerko, 2020).
- **Herramientas de IA generativa:** interfaces que producen contenido, pero que normalmente carecen de la capacidad de planificar o de utilizar herramientas externas de forma autónoma (Miao & Holmes, 2023).
- **Chatbots:** principalmente interfaces conversacionales que responden a indicaciones inmediatas (Tlili et al., 2023).

- **Copilotos:** sistemas diseñados para trabajar junto a un usuario, potenciando la acción humana, pero que normalmente dependen de transacciones iniciadas por humanos (OECD, 2026).
- **Agentes de IA semiautónomos y autónomos:** sistemas que pueden realizar acciones automáticamente y gestionar secuencias completas de operaciones, actuando esencialmente como "conductores" que pilotan múltiples modelos de IA hacia un objetivo (OECD, 2026).

La Tabla 1 distingue entre los modelos predictivos no generativos y la clase emergente de herramientas educativas de IA generativa.

TABLA 1

Ejemplos de diferentes categorías de IA

Fuente: adaptado de OECD (2026), OECD Digital Education Outlook 2026, Capítulo 1, "Exploring effective uses of generative artificial intelligence in education: An overview", Tabla 1.1, p. 14.

	NO GENERATIVA	GENERATIVA
Propósito general	■ Voz a texto, texto a voz ■ Herramientas de toma de notas (imagen a texto) ■ Software de traducción por IA	■ Chatbots (p. ej., ChatGPT, Deepseek, Gemini) ■ Generadores de imágenes, video o sonido
Educativa	■ Sistemas de tutoría inteligente (p. ej., Assistments, Lalilo, PILA) ■ Sistemas de alerta temprana ■ Simulaciones (RA/RV)	■ Tutores de IA generativa (p. ej., Gauth, Khanmigo, Question AI, Socratic Playground) ■ Asistentes de IA para docentes (p. ej., JeepyTA, Coteach, CoTutor)

Los sistemas agentes pueden clasificarse según su nivel de autonomía, que determina la transición de control entre el humano y la tecnología (OECD, 2021; OECD, 2026):

- **Agentes de Asistencia:** proporcionan información, explicaciones o fragmentos para que el usuario los implemente, manteniendo al usuario el control total sobre el diseño instruccional (OECD, 2026).

- **Agentes colaborativos:** trabajan conjuntamente con los usuarios como "socios de co-orquestación", donde las tareas se dividen en función de una conciencia compartida del contexto docente (OECD, 2026).
- **Agentes semi-autónomos:** ejecutan flujos de trabajo aprobados e intervienen en tareas específicas, como señalar errores comunes de los estudiantes para que un profesor los discuta, pero escalan las excepciones al usuario humano (OECD, 2026).
- **Agentes autónomos:** gestionan automáticamente todas las situaciones dentro de un caso de uso definido, generando contenido, retroalimentación y evaluaciones dinámicamente, sin esperar la intervención del profesor (OECD, 2021; OECD, 2026).

El nivel de autonomía concedido a un agente determina la necesidad de supervisión y de una gobernanza con el ser humano en el bucle. Niveles más altos de autonomía requieren una supervisión significativamente más estricta, especialmente cuando el agente está involucrado en decisiones de alto riesgo, como la calificación, el asesoramiento académico, la progresión del estudiante y el apoyo disciplinario o psicológico (OECD, 2023; Miao et al., 2024). En estos contextos, la rendición de cuentas humana es una responsabilidad legal y social que no puede cederse a la IA, ya que el rendimiento algorítmico rara vez es perfecto y carece de la capacidad humana para el juicio contextual y la empatía (Miao et al., 2024; OECD, 2023).

Lo crucial es que la capacidad tecnológica no equivale a la eficacia educativa. Los agentes de IA no mejoran el aprendizaje de forma inherente; más bien, su valor está mediado por la alineación pedagógica, el ajuste tarea-tecnología y la competencia del usuario (Garrison et al., 2000; OECD, 2026). Si los estudiantes dependen en exceso de agentes autónomos para obtener respuestas directas, su compromiso activo y su esfuerzo metacognitivo disminuyen, lo que conduce a un "espejismo de falso dominio" en el que el rendimiento en las tareas mejora, pero no se producen ganancias genuinas en el aprendizaje. Por lo tanto, el marco conceptual para la implementación responsable debe priorizar un enfoque de "pedagogía primero", asegurando que los agentes estén fundamentados en teorías de

aprendizaje establecidas y sometidos a una rigurosa gobernanza institucional (OECD, 2026).

Estas distinciones conceptuales proporcionan el vocabulario y la estructura analítica necesarios para explorar, en las siguientes secciones de este informe, las aplicaciones específicas, los determinantes de la adopción y el valor educativo a largo plazo de los agentes de IA.



04

Capítulo 4.

Aplicaciones y creación de valor en el aprendizaje electrónico



La integración de agentes de IA en ecosistemas de e-learning está impulsada por la promesa de generar valor de manera significativa en las dimensiones estudiantiles, docentes e institucionales. Sin embargo, el valor no es una propiedad inherente de la tecnología; solo se manifiesta cuando las capacidades agenticas están alineadas con objetivos pedagógicos específicos y con contextos organizativos (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; OECD, 2026). Esta sección analiza las principales aplicaciones de los agentes de IA y los mecanismos mediante los cuales generan o potencialmente socavan el valor educativo.

4.1

Aplicaciones dirigidas al estudiante: Personalización y agencia

Las aplicaciones orientadas al estudiante representan la frontera más visible de la IA agente. El principal motor de valor aquí es la personalización a gran escala, que aborda el modelo educativo "industrial" en el que tradicionalmente todos los estudiantes siguen trayectorias idénticas (OECD, 2021). Los agentes de IA actúan como tutores personalizados las 24 horas del día, proporcionando información bajo demanda, explicaciones detalladas y retroalimentación formativa inmediata (Park et al., 2025; OECD, 2026). Sistemas como el *Socratic Playground* demuestran cómo los agentes pueden pasar de dar respuestas simples a preguntas socráticas a guiar a los estudiantes a través del razonamiento paso a paso para fortalecer la argumentación y el pensamiento crítico (OECD, 2026).

Además, los agentes apoyan el aprendizaje autorregulado (SRL) al ayudar a los estudiantes con la planificación de estudios, la programación y el seguimiento del progreso (Park et al., 2025; OECD, 2026). Los sistemas avanzados pueden utilizar análisis en tiempo real para proporcionar andamios metacognitivos que impulsen a los estudiantes a supervisar sus propias estrategias (OECD, 2026). En cuanto a accesibilidad e inclusión, los agentes de IA aportan un valor esencial mediante la conversión de voz a texto, el subtítulo

automático y la traducción de idiomas, que apoyan a estudiantes con discapacidad auditiva o visual y a hablantes no nativos (OECD, 2023; OECD, 2026).

A pesar de estos beneficios, existe una tensión significativa entre la personalización y la dependencia. La dependencia excesiva de agentes autónomos para obtener respuestas directas puede conducir a una "pereza metacognitiva" o a un "espejismo de falso dominio", en los que los estudiantes alcanzan artefactos de alta calidad mediante descargas cognitivas sin interiorizar el conocimiento subyacente (OECD, 2026). Por ello, los agentes orientados al estudiante logran la máxima calidad educativa cuando su grado de autonomía se calibra para proporcionar apoyo en lugar de reemplazo, preservando la agencia del aprendiz (Park et al., 2025; Roe & Perkins, 2024; OECD, 2026).

4.2

Aplicaciones orientadas al profesorado: productividad y aumento pedagógico

Para el profesorado, la creación de valor central radica en la productividad y en el aumento de la competencia docente (Chen et al., 2020; OECD, 2026). Los agentes de IA pueden automatizar tareas "laboriosas" y repetitivas como la redacción del temario, la generación de problemas de práctica y la creación de rúbricas de evaluación (Ogunleye et al., 2024; Chen et al., 2020; Roe & Perkins, 2024; OECD, 2026). Esta automatización permite a los instructores reasignar tiempo a actividades de alto valor centradas en las relaciones, como la mentoría y la facilitación de discusiones críticas en profundidad (Chen et al., 2020; OECD, 2026).

En el ámbito del apoyo a la evaluación, los sistemas agentivos pueden realizar una puntuación inicial alineada con las rúbricas y proporcionar borradores de retroalimentación pulidos estilísticamente para que los profesores los revisen (OECD, 2026). Las investigaciones sugieren que este tipo de colaboración profesor-IA puede reducir el tiempo de corrección hasta en un **44%** mientras mantiene, o incluso mejora, la precisión (OECD, 2026). Además, los agentes mejoran la toma de

decisiones del profesorado mediante la interpretación de los análisis de aprendizaje, identificando malentendidos comunes de los estudiantes o detectando en tiempo real a estudiantes en riesgo, tareas que a menudo resultan imposibles de realizar manualmente a gran escala por los humanos (Ouyang et al., 2022; OECD, 2026).

Un riesgo clave es la desprofesionalización (OECD, 2026). Si los agentes se implementan como un reemplazo en lugar de un complemento, el profesorado puede perder su papel como los principales orquestadores de la experiencia de aprendizaje (OECD, 2026). La creación efectiva de valor requiere un trabajo en equipo "práxico" o "sinérgico", en el que el profesor permanece informado y evalúa los resultados de la IA para determinar su relevancia pedagógica y su veracidad (OECD, 2026).

Como se ilustra en la Figura 3, los agentes de IA generan valor al centrarse en las dimensiones cognitiva, metacognitiva y social del proceso de aprendizaje.



FIGURA 3

Objetivos del apoyo de la IA generativa y ejemplos
Fuente: adaptado de OECD (2026), OECD Digital Education Outlook 2026, Capítulo 4, "Fostering collaborative learning and promoting collaboration skills", Figura 4.3, p. 99

Aspectos (socio) cognitivos	Proceso metacognitivo y retroalimentación entre pares	Aspectos sociales y conductuales
- La IA generativa proporciona información sobre un tema (p. ej., Chu, 2024 y Feng, 2025) - La IA generativa promueve la co-construcción de conocimiento (p. ej., Darmawansah et al., 2025) - La IA generativa impulsa a los miembros del grupo a elaborar la información que aportaron (p. ej., Dang et al., 2025) - La IA generativa realiza tareas por el grupo (p. ej., X. Wei et al., 2025) - La IA generativa facilita explicaciones sobre la solución o los conceptos (Naik et al., 2024b)	- La IA generativa señala episodios de la colaboración para propiciar la reflexión sobre el proceso de colaboración (Ko y Foltz, en Greisel et al., 2025) - La IA generativa proporciona al grupo retroalimentación sobre su solución (dal Ponte et al., 2024; Gong et al., 2024) - La IA generativa proporciona un mensaje de retroalimentación (entre pares) para un texto (Greisel et al., 2025) - La IA generativa proporciona retroalimentación sobre un mensaje de retroalimentación (Greisel et al., 2025) - La IA generativa apoya la reflexión sobre la retroalimentación entre pares (Noroozi et al., en Greisel et al., 2025)	- La IA generativa impulsa el seguimiento de una distribución equitativa de la participación (p. ej., An et al., 2025) - La IA generativa sugiere temas que el grupo podría explorar a continuación (Lin et al., 2024) - La IA generativa impulsa a los miembros del grupo a participar de forma equitativa y a retomar la conversación en el tema (p. ej., Cai et al., 2024; Liu et al., 2024) - La IA generativa orquesta la colaboración, guiando a los grupos a través de la tarea, explicando la tarea y secuenciando el turno de palabra (p. ej., Liu et al., 2024)

4.3

Aplicaciones orientadas a la institución: eficiencia y diferenciación institucional

Las aplicaciones orientadas a instituciones aprovechan la IA a nivel macro para mejorar la capacidad de respuesta del servicio, la eficiencia de costes y la retención (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; OECD, 2026). Los agentes pueden agilizar flujos de trabajo administrativos complejos, como la incorporación, las admisiones y la evaluación de expedientes académicos (Chen et al., 2020; Erdmann & Toro-Dupouy, 2025; OECD, 2026). Un área significativa de innovación es la articulación de cursos y la transferencia de créditos; los modelos de IA basados en incrustaciones vectoriales pueden

identificar similitudes semánticas entre miles de cursos de distintas instituciones, lo que facilita la movilidad del alumnado (OECD, 2026).

En la orientación académica, los agentes ofrecen recomendaciones de carreras y cursos basadas en patrones históricos de matrícula y en objetivos profesionales, lo que aumenta la capacidad de los asesores humanos para gestionar grandes cargas de casos (OECD, 2026). Para asegurar la calidad y la acreditación, los agentes pueden etiquetar y alinear automáticamente el contenido de aprendizaje con los estándares nacionales o profesionales en constante evolución, reduciendo así los costes asociados al rediseño curricular (OECD, 2026).

Los agentes predictivos también apoyan las intervenciones de retención al identificar patrones de desgaste en toda la institución, lo que permite un apoyo proactivo (Ouyang et al., 2022; Erdmann & Toro-Dupouy, 2025; OECD, 2021). Estas capacidades contribuyen a la diferenciación institucional, ya que las universidades que integran con éxito la IA en su infraestructura central pueden ofrecer ecosistemas más flexibles y basados en datos (OECD, 2026).

4.4

Evaluación del valor a través del riesgo y la supervisión

El valor de cualquier aplicación de un agente de IA debe evaluarse en función de las consecuencias del error y de la sensibilidad de los datos (Erdmann & Toro-Dupouy, 2025; Miao & Cukurova, 2024). Las tareas estandarizadas y de bajo riesgo, como consultas logísticas o la generación de planes de estudio, son muy adecuadas para la automatización y para altos niveles de autonomía del agente (Chen et al., 2020; OECD, 2026). En cambio, las decisiones de alto impacto, como la calificación sumativa, la progresión académica, las acciones disciplinarias o el apoyo psicológico, requieren una estricta supervisión humana (Erdmann & Toro-Dupouy, 2025; Miao et al., 2024). Los resultados de la IA en estas áreas de alto riesgo siguen siendo propensos a las alucinaciones y carecen de la capacidad humana para el juicio contextual y la empatía (Miao et al., 2024; OECD, 2026).

Además, la tensión entre el apoyo proactivo y la privacidad debe gestionarse con cuidado (Ogunleye et al., 2024). Aunque la monitorización constante puede mejorar la retención, plantea preocupaciones significativas en torno a la vigilancia y la protección de datos (OECD, 2026; Miao & Cukurova, 2024). Los marcos de gobernanza institucional deben garantizar que la transparencia y la responsabilidad humana constituyan la base legal y social para la implementación de todos los agentes (Miao & Cukurova, 2024; Miao et al., 2024).

Los agentes de IA generan un valor genuino en el aprendizaje electrónico solo cuando sus capacidades funcionales se ajustan a la tarea educativa específica y apoyan objetivos pedagógicos generales (Erdmann & Toro-Dupouy, 2025; Toro-Dupouy & Erdmann, 2026). Ya sea que el objetivo sea mejorar la calidad educativa de los estudiantes, aumentar la productividad del profesorado o impulsar la innovación en la institución, el éxito depende de preservar la agencia humana y de operar dentro de estructuras de gobernanza sólidas. El mero uso frecuente de la IA o la obtención de eficiencia técnica no equivale a una mejora del aprendizaje; más bien, el valor educativo se manifiesta mediante la integración cuidadosa y la "pedagogía primero" de estos sistemas en el tejido centrado en el ser humano de la educación superior (OECD, 2026; Ouyang et al., 2022).

05

Capítulo 5.

Perspectivas teóricas



5.1

La implementación de agentes de IA en la educación superior es un desafío sociotécnico complejo que trasciende el mero despliegue tecnológico. Para entender cómo se adoptan, utilizan y evalúan estos sistemas autónomos, es necesario ir más allá de un enfoque de lente única. Ningún modelo único captura adecuadamente la interacción entre la calidad tecnológica, la eficacia pedagógica, la preparación institucional y la gobernanza responsable (Ouyang et al., 2022; Erdmann & Toro-Dupouy, 2025). Esta sección establece un marco teórico integrador que sintetiza diez modelos establecidos para analizar ecosistemas de agentes de IA a través de cinco dimensiones: aceptación individual, desempeño de tareas, proceso pedagógico, contexto organizativo y gobernanza ética.

Determinantes de la aceptación y uso individual

El Modelo de Aceptación de la Tecnología (TAM) y la Teoría Unificada de la Aceptación y el Uso de la Tecnología (UTAUT/UTAUT2) proporcionan la lente principal para comprender las intenciones del usuario. TAM, según lo establecido por Davis (1989), se centra en dos creencias fundamentales: la utilidad percibida (PU), el grado en que un usuario cree que un sistema mejora el rendimiento, y la facilidad de uso percibida (PEOU), el grado en que el sistema se percibe como libre de esfuerzo (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025). UTAUT los integra con constructos adicionales: expectativa de rendimiento, expectativa de esfuerzo, influencia social y condiciones facilitadoras (Toro-Dupouy & Erdmann, 2026; et al., 2019; Venkatesh et al., 2003).

En el contexto de los agentes de IA, la PU sigue siendo el predictor más fuerte de la intención de adopción tanto para estudiantes como para profesores (Erdmann & Toro-Dupouy, 2025; Venkatesh et al., 2003). Para los estudiantes, los agentes deben proporcionar una utilidad clara en la generación de ideas y en la productividad académica para alcanzar un "umbral" de adopción (Toro-Dupouy & Erdmann, 2026). Para el profesorado, la adopción es más "escalonada" y requiere evidencia

5.2

incremental de que el agente ahorra tiempo en tareas rutinarias sin erosionar los estándares pedagógicos (Toro-Dupouy & Erdmann, 2026). Aunque estos modelos son altamente predictivos de la intención de adopción, su principal limitación radica en la falta de enfoque en los resultados reales de aprendizaje (Ouyang et al., 2022; Al-Fraihat et al., 2020). Para futuras investigaciones empíricas, estas teorías aportan variables clave como la intención conductual, las normas sociales y la formación de hábitos (Toro-Dupouy & Erdmann, 2026; Venkatesh et al., 2003).

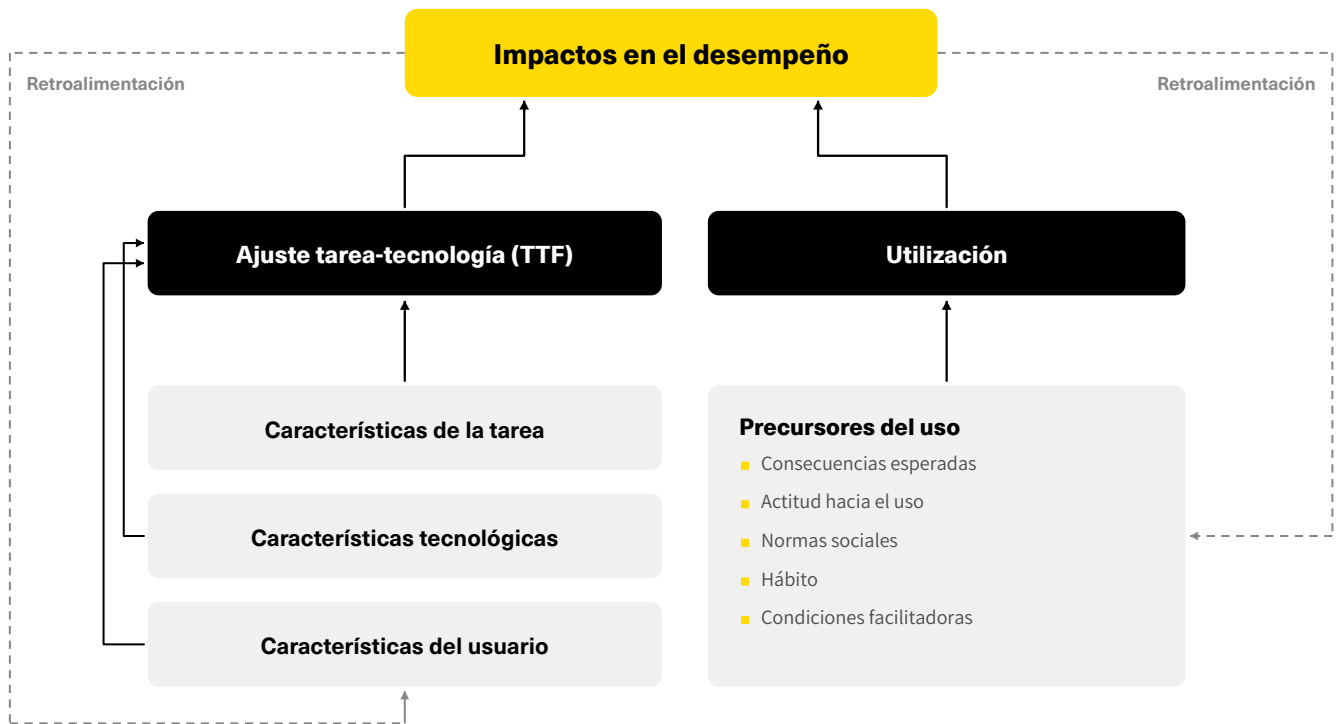
Alineación de tareas y rendimiento del sistema

Mientras que TAM y UTAUT explican por qué un usuario podría *querer* usar un agente de IA, el Ajuste Tarea–Tecnología (TTF) y el Modelo de Éxito de Sistemas de Información de DeLone y McLean (IS) explican si el agente realmente *mejora el rendimiento*. TTF sostiene que los impactos en el rendimiento se deben a la alineación entre las capacidades tecnológicas y los requisitos de la tarea (Toro-Dupouy & Erdmann, 2026; Goodhue & Thompson, 1995). Para los agentes de IA, el ajuste se determina por la capacidad del agente para planificar y ejecutar tareas en varios pasos, como la articulación de recorridos o la generación de retroalimentación compleja, en función de las demandas cognitivas del usuario (OECD, 2026; Goodhue & Thompson, 1995). Un mal ajuste, en el que se utiliza a un agente para tareas que requieren un alto nivel de juicio interpersonal, puede provocar impactos negativos en el rendimiento (Glikson & Woolley, 2020; Goodhue & Thompson, 1995). La Figura 4 ilustra cómo los impactos en el rendimiento dependen de la alineación entre las funcionalidades del agente y los requisitos específicos de la tarea.

FIGURA 4

La cadena Tecnología-Rendimiento (Modelo TTF)

Fuente: adaptado de Technology-to-Performance Chain (TPC) de Goodhue y Thompson (1995).



El Modelo de Éxito de IS complementa esto al evaluar la calidad del sistema (por ejemplo, fiabilidad y capacidad de respuesta), la calidad de la información (por ejemplo, precisión y relevancia) y la calidad del servicio (por ejemplo, el apoyo institucional) (Toro-Dupouy & Erdmann, 2026; Al-Fraihat et al., 2020; DeLone & McLean, 2003). Estos factores de calidad son antecedentes tanto de la satisfacción del usuario como de los beneficios netos (Al-Fraihat et al., 2020; DeLone & McLean, 2003). Existe una relación significativa en la que la alta calidad de la información en los resultados de la IA, como la ausencia de "alucinaciones", influye directamente en la utilidad y la confianza percibidas (Al-Fraihat et al., 2020; OECD, 2026). La fortaleza de estos modelos radica en su capacidad diagnóstica para identificar fallos técnicos, aunque a menudo pasan por alto la dinámica social del aula (Al-Fraihat et al., 2020; Goodhue & Thompson, 1995).

5.3

La perspectiva pedagógica: procesos de aprendizaje y preparación del profesorado

Para evaluar el valor educativo, el marco integra el Aprendizaje Autoregulado (SRL) y la Comunidad de Indagación (Col). SRL describe el aprendizaje como un proceso cíclico de establecimiento de objetivos, planificación, desempeño y autorreflexión (Panadero, 2017; Toro-Dupouy & Erdmann, 2026; Zimmerman, 2002). Los agentes de IA actúan como "andamios metacognitivos", animando a los estudiantes a monitorizar sus estrategias (OECD, 2021; OECD, 2026). Sin embargo, un riesgo clave identificado en investigaciones recientes es la "pereza metacognitiva" o descarga cognitiva, en la que los estudiantes delegan tareas de pensamiento crítico en agentes autónomos en lugar de interiorizar el conocimiento (OECD, 2026).

El marco de la Col enfatiza que las experiencias educativas valiosas requieren la interacción entre la presencia docente, la presencia cognitiva y la presencia social (Toro-Dupouy & Erdmann, 2026; Garrison et al., 2000). Los agentes de IA pueden aumentar la presencia docente proporcionando retroalimentación diagnóstica y soporte 24/7, así como presencia social mediante interfaces conversacionales integradas (Garrison et al., 2000; OECD, 2026). Una implementación eficaz requiere que el profesorado posea Conocimientos Tecnológicos de Contenido Pedagógico (TPACK), el conocimiento especializado necesario para integrar agentes de IA en un tema específico de manera que apoye, en lugar de sustituir, la instrucción humana (Toro-Dupouy & Erdmann, 2026; OECD, 2023; Mishra & Koehler, 2006). Estas teorías desplazan el enfoque del "uso" al de la "calidad de la interacción", aunque están limitadas por la dificultad de medir la "presencia" y la "metacognición" a gran escala (OECD, 2021; Garrison et al., 2000).

5.4

Contexto institucional y de gobernanza

La adopción institucional se analiza a través del marco Tecnología–Organización–Entorno (TOE). La preparación institucional, que incluye la infraestructura técnica y los recursos humanos, actúa como un "filtro" para la adopción individual (Erdmann & Toro-Dupouy, 2025; Toro-Dupouy & Erdmann, 2026). Factores como la complejidad organizativa, la preparación y la ventaja competitiva son condiciones necesarias para un despliegue exitoso (Erdmann & Toro-Dupouy, 2025).

Finalmente, el marco incorpora la confianza en la automatización y en la IA responsable para abordar la gobernanza. La confianza es un mediador entre la calidad tecnológica y el uso real, influida por la precisión, la fiabilidad y la transparencia del agente (Toro-Dupouy & Erdmann, 2026; Glikson & Woolley, 2020). Los principios responsables de la IA, incluyendo la equidad, la rendición de cuentas, la explicabilidad y la preservación de la agencia humana, son esenciales para gestionar decisiones agenticas de alto riesgo, como la calificación o las acciones disciplinarias (Toro-Dupouy & Erdmann, 2026; UNESCO, 2021; Miao et al., 2024; Miao & Cukurova, 2024). La supervisión humana actúa como moderadora crítica, en la que una mayor autonomía de los agentes requiere una gobernanza más estricta por parte de las personas en el bucle para mitigar los riesgos de error o sesgo (Miao & Cukurova, 2024; OECD, 2026).

5.5

Argumento integrativo y relaciones multidimensionales

Este informe argumenta que estas teorías son funcionalmente interdependientes. La calidad del sistema y de la información (éxito en sistemas de información) es el principal motor de la utilidad y la confianza percibidas (TAM/Confianza) (Toro-Dupouy & Erdmann, 2026; Al-Fraihat et al., 2020). La confianza y la facilidad de uso del usuario están influenciadas de manera significativa por la alfabetización en IA, que es un componente de la capacidad individual y del profesorado en TPACK (Toro-Dupouy & Erdmann, 2026; Ng et al., 2021; Miao & Cukurova,

2024). El valor educativo real creado depende del ajuste tarea–tecnología: el grado en que las capacidades del agente apoyan, en lugar de eludir, los procesos de aprendizaje autorregulados del estudiante (OECD, 2026; OECD, 2021; Goodhue & Thompson, 1995).

Además, la preparación institucional (TOE) determina las condiciones facilitadoras (UTAUT) que posibilitan un uso sostenido (Venkatesh et al., 2003; Erdmann & Toro-Dupouy, 2025). De manera crucial, el riesgo de dependencia del alumno o de desprofesionalización del profesorado se modera mediante la supervisión humana y los marcos de IA responsables, que aseguran que los agentes funcionen como "socios colaborativos" en lugar de como sustitutos autónomos (OECD, 2026; Miao & Cukurova, 2024).

Ninguna perspectiva teórica puede explicar por sí sola el éxito de los agentes de IA en la educación superior. En su lugar, este informe propone un marco integrador multidimensional que combina:

- **Factores tecnológicos:** (IS Éxito, TTF) para garantizar la calidad técnica y la alineación de tareas.
- **Factores individuales:** (TAM, UTAUT, Trust) para entender la aceptación y la confianza del usuario.
- **Factores pedagógicos:** (SRL, CoI, TPACK) para preservar la profundidad de la experiencia educativa y la agencia del profesorado.
- **Factores Organizacionales:** (TOE, Resiliencia Institucional) para gestionar el macrocontexto de la adopción (Toro-Dupouy & Erdmann, 2026).
- **Factores de gobernanza:** (IA responsable) para garantizar la rendición de cuentas y la primacía de la toma de decisiones humanas (Miao & Cukurova, 2024).

Esta síntesis conceptual proporciona la base rigurosa necesaria para analizar, en las secciones posteriores de este documento, las aplicaciones y los resultados específicos de los agentes de IA.

06

Capítulo 6.

Beneficios, riesgos y condiciones de contorno



6.1

El despliegue de agentes de IA en entornos de e-learning representa una doble vía de transformación, con un potencial considerable para mejorar los resultados educativos y, al mismo tiempo, introducir riesgos sistémicos. Lo crucial es que el impacto de estas tecnologías no es inherente ni universal; más bien, es contingente y depende del contexto, el cual está determinado por la interacción entre las capacidades tecnológicas, el diseño pedagógico, las características del usuario, los requisitos de las tareas y la gobernanza institucional (OECD, 2021; OECD, 2026).

Beneficios potenciales y evidencia emergente

Los agentes de IA ofrecen el principal beneficio de la personalización y el soporte adaptativo a gran escala, llevando el e-learning más allá del modelo "industrial" de instrucción única para todos (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; OECD, 2021; OECD, 2026). Estos sistemas pueden funcionar como tutores personalizados 24/7, utilizando preguntas socráticas y el diálogo en tiempo real para proporcionar retroalimentación formativa inmediata, adaptada a las lagunas de conocimiento, los malentendidos y el ritmo de aprendizaje individuales del estudiante (Toro-Dupouy & Erdmann, 2026; Park et al., 2025; OECD, 2026; Zawacki-Richter et al., 2019). Además, los sistemas de agentes apoyan el aprendizaje autorregulado (SRL) actuando como andamios metacognitivos que impulsan a los estudiantes a planificar, monitorizar y reflexionar sobre sus estrategias de aprendizaje (Park et al., 2025; OECD, 2026). En cuanto a accesibilidad e inclusión, los agentes de IA ofrecen apoyo esencial a estudiantes con necesidades especiales y a hablantes no nativos mediante traducción automática, subtítulo y participación multimodal (Chen et al., 2020; OECD, 2021; OECD, 2023; OECD, 2026).

Para el profesorado, la integración de agentes de IA puede aumentar significativamente la productividad al automatizar tareas administrativas "laboriosas" como la redacción del programa, la generación de problemas de práctica y la calificación preliminar

basada en rúbricas (Ogunleye et al., 2024; Park et al., 2025; Chen et al., 2020; Roe & Perkins, 2024; OECD, 2026). Esto permite a los instructores reasignar su tiempo a actividades humanas de alto valor, como la mentoría y las discusiones críticas complejas (OECD, 2026; Zawacki-Richter et al., 2019). A nivel institucional, los agentes mejoran la eficiencia y la capacidad de respuesta mediante servicios estudiantiles automatizados, asesoramiento académico y sistemas de alerta temprana que identifican a estudiantes desconectados o en riesgo mucho antes que los métodos tradicionales (Ogunleye et al., 2024; Ouyang et al., 2022; OECD, 2023; OECD, 2026). Sin embargo, debe mantenerse una distinción crítica entre estos beneficios potenciales y los resultados empíricamente demostrados (Ouyang et al., 2022). Aunque los prototipos muestran un gran potencial, la evidencia longitudinal a gran escala sigue siendo escasa y algunos estudios indican que la mera presencia de IA no conduce automáticamente a ganancias en el aprendizaje (OECD, 2026).

6.2

Riesgos principales y decisiones de alta consecuencia

Los riesgos asociados a los agentes de IA son multifacéticos, comenzando por problemas técnicos como las alucinaciones y los consejos inexactos, en los que los sistemas generan información plausible pero incorrecta sobre hechos que pueden engañar a los estudiantes (Park et al., 2025; Ogunleye et al., 2024; Tlili et al., 2023; OECD, 2026; Miao & Cukurova, 2024). Más profundamente, el uso excesivo puede conducir a sesgos de automatización y a una dependencia excesiva, que los investigadores denominan "pereza metacognitiva" (Toro-Dupouy & Erdmann, 2026; OECD, 2026). En este escenario, los estudiantes pueden experimentar un "espejismo de dominio falso", en el que el rendimiento en la tarea mejora mediante la descarga cognitiva del agente, pero el desarrollo genuino de habilidades y el pensamiento crítico se ven socavados (Park et al., 2025; OECD, 2026; Kasneci et al., 2023). Socialmente, existe un riesgo significativo de deshumanización y de pérdida de la interacción

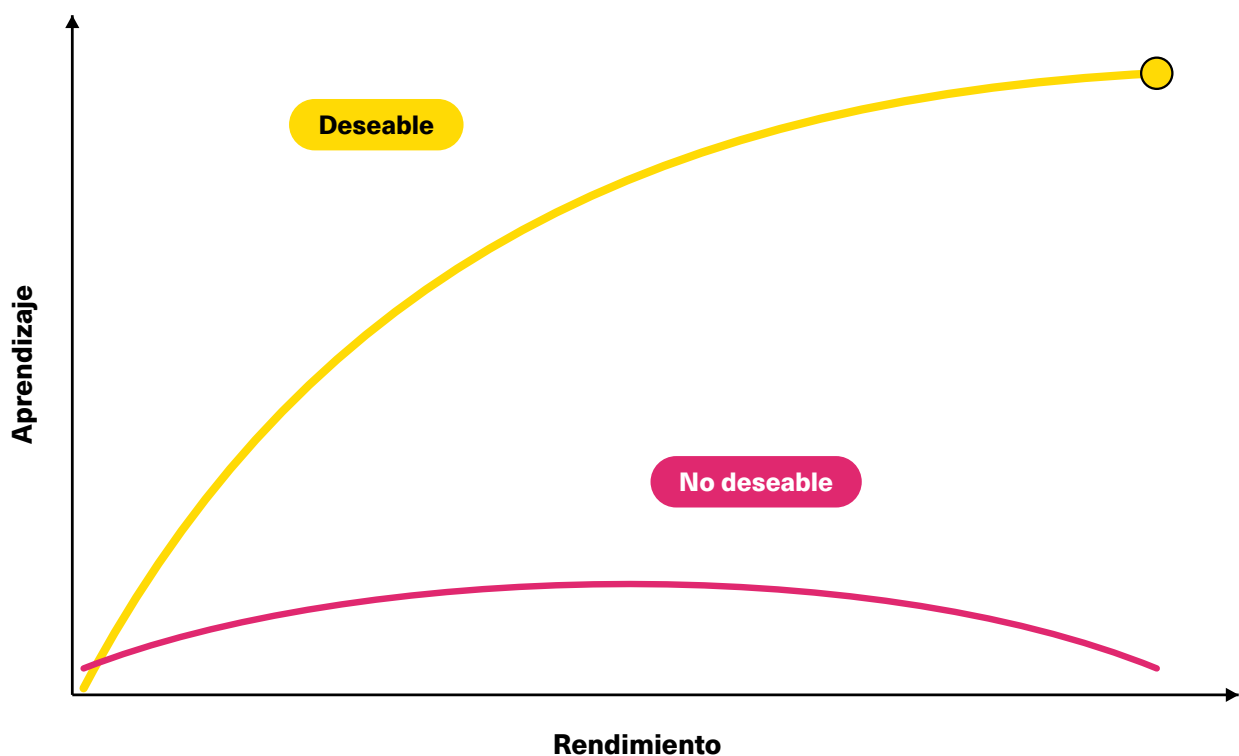
humana esencial entre estudiantes y profesores (Toro-Dupouy & Erdmann, 2026; Miao & Holmes, 2023; Miao & Cukurova, 2024).

Como se muestra en la Figura 5, el principal riesgo de los agentes autónomos es un “espejismo de falso dominio” en el que el rendimiento en la tarea mejora, pero las ganancias reales de aprendizaje se estancan.

FIGURA 5

Equilibrio entre las ganancias de aprendizaje y el rendimiento con la IA generativa

Fuente: adaptado de OECD (2026), OECD Digital Education Outlook 2026, Capítulo 2, "Generative AI for human skill development and assessment", Figura 2.4, p. 51 (numerada aquí como Figura 5).



Los riesgos institucionales incluyen preocupaciones profundas sobre la privacidad, la vigilancia y la seguridad de los datos, ya que los sistemas de agentes suelen requerir la recopilación de datos sensibles de comportamiento y biométricos (Ogunleye et al., 2024; Miao & Cukurova, 2024; Toro-Dupouy & Erdmann, 2026; OECD, 2021; OECD,

2023; UNESCO, 2021). El sesgo algorítmico sigue siendo una amenaza sistémica, donde los datos de entrenamiento que reflejan prejuicios humanos pasados pueden conducir a un trato desigual basado en género, raza o estatus socioeconómico (Park et al., 2025; Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; OECD, 2021; OECD, 2023; OECD, 2026; UNESCO, 2021). Además, la transición hacia agentes autónomos complica la responsabilidad y la integridad académica, planteando cuestiones sin resolver sobre la propiedad del trabajo y la transparencia en la toma de decisiones automatizadas (Ogunleye et al., 2024; Park et al., 2025; Chen et al., 2020; Roe & Perkins, 2024; Batista et al., 2024; Miao & Holmes, 2023; OECD, 2026; Miao et al., 2024; Miao & Cukurova, 2024). Se requiere especial precaución en decisiones educativas de alto impacto, como la calificación sumativa, la progresión académica y el apoyo psicológico, donde el rendimiento algorítmico, rara vez perfecto, nunca debe sustituir el juicio humano ni la empatía contextual (Erdmann & Toro-Dupouy, 2025; OECD, 2021; Miao & Holmes, 2023; OECD, 2023; OECD, 2026; Miao et al., 2024).

6.3

Condiciones de frontera y tensiones contextuales

La realización del valor, o la generación de riesgo, depende de condiciones de contorno específicas. El factor más crítico es la calidad del diseño pedagógico; los agentes de IA deben ser "pedagogía primero", fundamentados en teorías establecidas del aprendizaje en lugar de estar impulsados por la capacidad tecnológica (OECD, 2026; Park et al., 2025; Erdmann & Toro-Dupouy, 2025; Ouyang et al., 2022). El ajuste tarea-tecnología determina si un agente es adecuado para una demanda cognitiva específica, mientras que el grado de autonomía del agente debe calibrarse cuidadosamente para preservar la agencia humana (OECD, 2026; Tlili et al., 2023; OECD, 2021). La implementación efectiva requiere, además, altos niveles de alfabetización en IA de los usuarios y del Conocimiento Tecnológico Pedagógico de Contenidos (TPACK) del profesorado para garantizar que los interesados puedan evaluar y monitorizar críticamente los resultados de la IA (OECD, 2026;

Miao & Cukurova, 2024; Toro-Dupouy & Erdmann, 2026; Ng et al., 2021; OECD, 2023).

Implementar agentes de IA requiere calibrar el control entre humanos y máquinas a lo largo de los seis niveles de automatización que se muestran en la Figura 6.

FIGURA 6

Modelo de seis niveles de automatización

Fuente: adaptado de Anne Hovers e Inge Molenaar, Adaptive Learning Lab, tal como se presenta en OECD (2023), Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem, Figura 13.3.

Control del docente

NIVEL	DOCENTE	TECNOLOGÍA
1 Nivel 1 Solo docente	controla todo el proceso de enseñanza	no interviene
2 Nivel 2 Asistencia al docente	mantiene el control total	ofrece información de apoyo
3 Nivel 3 Automatización parcial	supervisa la tecnología	controla tareas específicas
4 Nivel 4 Automatización condicional	supervisión ocasional; puede retomar el control	señala cuándo se necesita al docente y controla más tareas
5 Nivel 5 Automatización alta	no requiere control ni supervisión en tareas específicas	gestiona casi todas las tareas y solicita ayuda si hace falta
6 Nivel 6 Automatización total	sin intervención	controla todas las tareas de forma automática

Automatización

Los factores organizativos, incluidos la preparación institucional y la madurez de la gobernanza, actúan como filtros del éxito de la adopción (OECD, 2021; Erdmann & Toro-Dupouy, 2025; Miao & Holmes, 2023; OECD, 2026; Miao & Cukurova, 2024). Un enfoque "humano en el bucle" es esencial, ya que proporciona la supervisión y la transparencia necesarias para mitigar riesgos y mantener la confianza (Miao & Holmes,

2023; Kasneci et al., 2023; OECD, 2021; OECD, 2023; OECD, 2026; Miao & Cukurova, 2024). Estas condiciones revelan tensiones fundamentales que los líderes universitarios deben navegar:

- **Personalización frente a dependencia:** Proporcionar apoyo personalizado sin crear un aprendiz incapaz de pensar de forma independiente (OECD, 2026; Park et al., 2025).
- **Automatización frente a agencia humana:** Escalar la eficiencia mientras se evita la desprofesionalización docente y se preserva la primacía de la toma de decisiones humanas (OECD, 2026; Miao & Cukurova, 2024; Miao et al., 2024).
- **Escalabilidad frente a calidad:** Asegurar que las herramientas de IA de gran consumo estén alineadas con rigurosos estándares curriculares y objetivos pedagógicos (OECD, 2026).
- **Intervención basada en datos frente a privacidad:** Utilizar la analítica para la retención de estudiantes evitando la vigilancia intrusiva (OECD, 2021; OECD, 2023; Zawacki-Richter et al., 2019).
- **Autonomía frente a la rendición de cuentas:** Determinar quién es responsable cuando los sistemas semiautónomos cometen errores en contextos de alto riesgo (OECD, 2026; Miao et al., 2024; Miao & Cukurova, 2024; Roe & Perkins, 2024).

Los agentes de IA no mejoran inherentemente el e-learning; su impacto es contingente más que universal. El mismo agente puede crear un valor transformador en un entorno bien gobernado y pedagógicamente sólido y generar riesgos inaceptables en otro. La siguiente sección sintetizará estos beneficios, riesgos y condiciones de frontera en un único marco teórico integrador para la implementación responsable de agentes de IA en la educación superior.

07

Capítulo 7.

Marco Teórico Integrador Propuesto y Agenda de Investigación Futura



La implementación de agentes de IA en la educación superior es un desafío sociotécnico complejo, en el que el valor educativo depende más de lo automático (OECD, 2026). Con base en la literatura sintetizada, este informe propone un marco teórico integrador que posiciona el éxito de los agentes de IA como función de la interacción entre las condiciones tecnológicas, individuales, pedagógicas, organizativas y ambientales (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025).

El marco propuesto (Figura 7) sintetiza los determinantes tecnológicos, individuales y organizativos necesarios para una implementación responsable.

FIGURA 7

Marco integrador para la implementación de agentes de IA

Fuente: Elaboración propia.



7.1

Componentes del Marco Integrador

El marco identifica cinco grupos de antecedentes que impulsan la adopción y eficacia de los sistemas agentes:

- **Factores tecnológicos:** basados en el modelo de éxito de DeLone y McLean (IS) y en el TAM, estos incluyen la calidad del sistema, la precisión de la información y la fiabilidad (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; Salloum et al., 2019). Para los agentes autónomos, la transparencia en la ruta de decisión y la interoperabilidad con los Sistemas de Gestión del Aprendizaje (LMS) existentes son requisitos técnicos críticos para evitar experiencias de usuario fragmentadas (OECD, 2026).
- **Factores individuales:** el éxito depende de la alfabetización en IA y de la competencia digital del usuario (Toro-Dupouy & Erdmann, 2026; Ng et al., 2021; OECD, 2026). Las percepciones individuales están moldeadas por la autoeficacia y la experiencia informática, pero la adopción suele verse limitada por preocupaciones de privacidad y ansiedad por la IA (Toro-Dupouy & Erdmann, 2026; Ng et al., 2021; Al-Fraihat et al., 2020).
- **Factores pedagógicos:** basados en el Aprendizaje Autoregulado (SRL) y la Comunidad de Investigación (Col), estos factores se centran en el diseño instruccional y el ajuste tarea-tecnología (TTF). La alineación pedagógica de un agente determina si proporciona un andamiaje significativo o conduce a una descarga cognitiva (Toro-Dupouy & Erdmann, 2026; Panadero, 2017; OECD, 2026).
- **Factores organizacionales:** a partir del marco TOE, el éxito institucional está impulsado por la adaptabilidad del liderazgo, una estrategia de IA coherente y la disponibilidad de formación y de apoyo técnico (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025). La resiliencia institucional actúa como una metacapacidad que indica estabilidad y fomenta la confianza de los grupos de interés (Toro-Dupouy & Erdmann, 2026).
- **Factores medioambientales:** estos incluyen presiones externas, como el cumplimiento normativo (por ejemplo, el RGPD), los

estándares de acreditación y la competencia de otras instituciones (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; OECD, 2026).

Estos antecedentes influyen en los mediadores centrales de la utilidad percibida (UP), la facilidad de uso percibida (PEOU), la confianza y el compromiso (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; Glikson y Woolley, 2020). Estos, a su vez, facilitan una colaboración eficaz entre humanos e IA, resultando en un conjunto de resultados divididos en tres niveles (OECD, 2023; OECD, 2026):

- 1. Resultados del alumnado:** adopción, aprendizaje más profundo mediante el diálogo socrático, mayor autorregulación, mayor satisfacción y mejora de la retención (Ogunleye et al., 2024; Ouyang et al., 2022; OECD, 2026).
- 2. Resultados para el profesorado:** aumento de la eficacia docente, incremento de la productividad mediante la automatización de tareas "laboriosas" y mantenimiento de la satisfacción profesional mediante la preservación de la agencia humana (Toro-Dupouy & Erdmann, 2026; OECD, 2026; Roe & Perkins, 2024; OECD, 2026).
- 3. Resultados institucionales:** mejora de la calidad educativa a gran escala, de la eficiencia operativa y de la capacidad de innovación a largo plazo (Toro-Dupouy & Erdmann, 2026; Ouyang et al., 2022; OECD, 2026).



7.2

Moderadores clave y condiciones límite

La relación entre adopción y valor educativo se modera por el grado de autonomía del agente y la complejidad de la tarea (OECD, 2026; Glikson y Woolley, 2020; OECD, 2021). Mientras que los agentes asistentes en tareas de bajo riesgo (por ejemplo, la planificación de estudios) requieren menos supervisión, los agentes autónomos en áreas de mayor importancia (por ejemplo, la calificación) requieren una supervisión humana estricta y altos niveles de madurez en la gobernanza (OECD, 2021; OECD, 2026). Además, el ajuste Tarea–Tecnología modula el rendimiento: si se utiliza un agente para tareas que requieren un alto grado de juicio interpersonal, los resultados pueden ser indeseables (Glikson y Woolley, 2020; Goodhue & Thompson, 1995). Finalmente, la preparación institucional de la IA actúa como un filtro que determina si las condiciones facilitadoras son suficientes para un uso sostenido (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025).

7.3

Agenda de Investigación Futura

Para validar y refinar este marco, este informe propone las siguientes propuestas de investigación:

- **Proposición 1:** la utilidad percibida es un predictor más fuerte de la intención de adopción por parte de los agentes de IA que la facilidad de uso, especialmente para los profesores cuya adopción sigue un patrón escalonado basado en valores (Erdmann & Toro-Dupouy, 2025; Toro-Dupouy & Erdmann, 2026).
- **Proposición 2:** la confianza humana en los agentes de IA se basa principalmente en la calidad de la información (precisión y ausencia de alucinaciones) más que en la fluidez conversacional de la interfaz (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025; Glikson y Woolley, 2020).
- **Proposición 3:** la efectividad educativa es una función conjunta de la utilización y el ajuste tarea–tecnología; el uso frecuente de

agentes de IA para tareas poco adecuadas conducirá a una "pereza metacognitiva" y a un "espejismo de falso dominio" (Toro-Dupouy & Erdmann, 2026; OECD, 2026; Goodhue & Thompson, 1995).

- **Propuesta 4:** niveles más altos de alfabetización en IA de los estudiantes se correlacionan positivamente con una dependencia adecuada de los agentes de IA, lo que reduce los riesgos de sobredependencia y de violaciones de la integridad académica (Park et al., 2025; OECD, 2026).
- **Proposición 5:** el éxito en la adopción institucional depende de la preparación organizativa, en la que una infraestructura técnica sólida y el apoyo al liderazgo son necesarios, pero no suficientes, para la aceptación del profesorado (Erdmann & Toro-Dupouy, 2025; Toro-Dupouy & Erdmann, 2026).
- **Proposición 6:** a medida que aumenta la autonomía del agente, la necesidad de una gobernanza "humana en el bucle" se vuelve no lineal, lo que requiere un TPACK más sofisticado por parte del profesorado para mantener el control pedagógico (OECD, 2026).

La investigación futura debería probar este marco integrador utilizando una variedad de metodologías, incluyendo encuestas a estudiantes y profesores, entrevistas semiestructuradas para explorar las experiencias de los usuarios y Modelado de Ecuaciones Estructurales (SEM) para analizar las relaciones entre antecedentes y resultados (Erdmann & Toro-Dupouy, 2025; Toro-Dupouy & Erdmann, 2026). Además, los estudios longitudinales y los ensayos controlados aleatorizados (ECA) son esenciales para distinguir los "efectos de novedad" a corto plazo de los avances de aprendizaje a largo plazo e identificar las condiciones específicas bajo las cuales el equipo humano-IA logra una sinergia genuina (Erdmann & Toro-Dupouy, 2025; OECD, 2026). En última instancia, el valor de los agentes de IA no es una propiedad inherente del software, sino un resultado emergente de un ecosistema educativo bien gobernado (Toro-Dupouy & Erdmann, 2026; OECD, 2026).

08

Capítulo 8.

Implicaciones para la gestión de las escuelas de negocios



La integración exitosa de agentes de IA en las escuelas de negocios requiere un cambio estratégico, pasando de tratar la tecnología como una herramienta periférica a considerarla como una capacidad institucional central. Los líderes deben reconocer que el objetivo estratégico no es la automatización por sí misma, sino la creación de un valor medible educativo y organizacional (Erdmann & Toro-Dupouy, 2025; OECD, 2026). Para lograrlo, la dirección debería adoptar una estrategia de despliegue basada en valores que coordine los cambios en tecnología, pedagogía, personas, procesos y gobernanza (Toro-Dupouy & Erdmann, 2026; Erdmann & Toro-Dupouy, 2025).

La implementación debe comenzar con casos de uso de bajo riesgo y alto volumen que generen ganancias inmediatas de eficiencia sin comprometer el rigor académico. Estos incluyen la automatización de flujos de trabajo administrativos, como la matrícula de cursos, la evaluación de expedientes académicos y las consultas de servicio al alumnado, así como la asistencia rutinaria del profesorado en la redacción del temario y la elaboración de ejercicios prácticos (Erdmann & Toro-Dupouy, 2025; OECD, 2026). Por el contrario, las escuelas deben mantener la responsabilidad humana en las decisiones de alto riesgo (OECD, 2023; UNESCO, 2021). Las decisiones que involucran calificación sumativa, progresión académica, acciones disciplinarias y asesoramiento sensible deben seguir siendo lideradas por humanos, con los agentes de IA actuando únicamente como apoyo o para el diagnóstico (OECD, 2026; Miao & Cukurova, 2024).

Para gestionar estas distinciones, las escuelas de negocios deben definir niveles claros de autonomía del agente y reglas de escalada (Parasuraman et al., 2000; OECD, 2023). Siguiendo el principio de "humano en el bucle", los marcos institucionales deberían especificar cuándo un agente puede actuar de forma independiente y cuándo debe ceder a un profesional humano (Miao & Holmes, 2023; UNESCO, 2021). La responsabilidad última por los resultados algorítmicos debe atribuirse siempre a las personas físicas o jurídicas, y no al propio sistema (UNESCO, 2021; Miao et al., 2024).

Una prioridad principal en la gestión es invertir en la alfabetización en IA del profesorado y de los estudiantes (OECD, 2026; Toro-Dupouy &

Erdmann, 2026; Erdmann & Toro-Dupouy, 2025). Para el profesorado, esto implica un desarrollo profesional estructurado para construir el Conocimiento Tecnológico de Contenidos Pedagógicos (TPACK), lo que les permite actuar como orquestadores del aprendizaje potenciado por IA en lugar de simples supervisores (Toro-Dupouy & Erdmann, 2026; OECD, 2026; Miao & Cukurova, 2024). Para los estudiantes, la formación debe centrarse en la evaluación crítica de los resultados de la IA para evitar la "pereza metacognitiva" y el "espejismo de falso dominio" (OECD, 2026; Miao et al., 2024).

Una implementación eficaz requiere, además, que los agentes de IA se integren con los sistemas institucionales existentes, como los Sistemas de Gestión del Aprendizaje (LMS) y los almacenes de datos de información estudiantil, en lugar de desplegarse como herramientas aisladas (OECD, 2021; OECD, 2023; OECD, 2026). Esta integración permite análisis más ricos y evita la fragmentación de la experiencia estudiantil (OECD, 2021; OECD, 2026). Esta preparación técnica debe ir acompañada de marcos de gobernanza robustos que aborden la privacidad, la seguridad de los datos, el sesgo algorítmico y la integridad académica (Kasneji et al., 2023; Miao & Holmes, 2023; OECD, 2021; Zawacki-Richter et al., 2019). Los mecanismos de transparencia son esenciales para garantizar que el razonamiento detrás de las recomendaciones agénticas sea visible tanto para los profesores como para los estudiantes (Miao & Holmes, 2023; OECD, 2026).

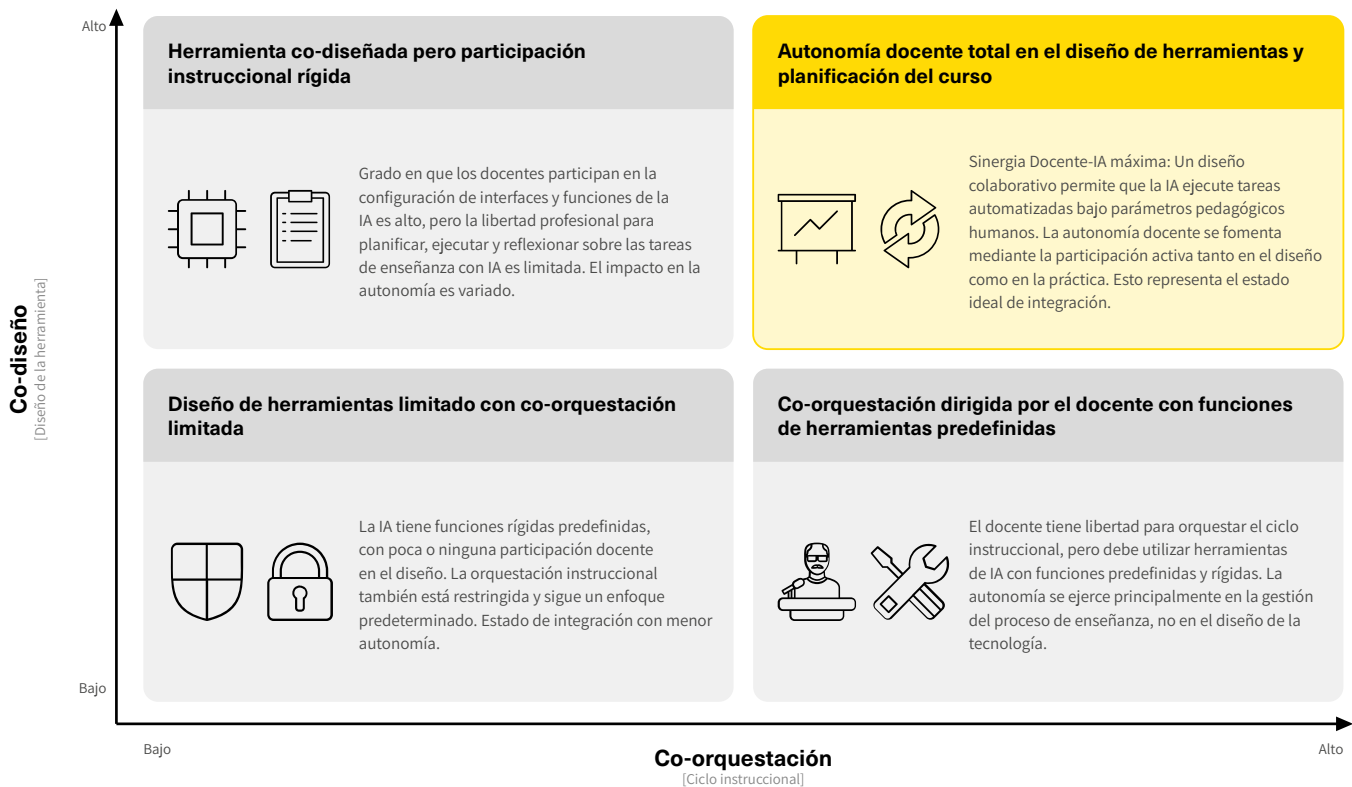
Por último, las escuelas de negocios deberían pilotar, evaluar y escalar de forma gradual (Toro-Dupouy & Erdmann, 2026; Miao & Holmes, 2023). El despliegue debe basarse en una evaluación de la preparación institucional y en un plan claro de gestión del cambio que tenga en cuenta la aceptación de las partes interesadas (Toro-Dupouy & Erdmann, 2026; OECD, 2021). El éxito debe evaluarse en función de KPIs educativos y organizativos específicos, que incluyen la velocidad de aprendizaje, la satisfacción del alumnado, la reducción de la carga de trabajo del profesorado, las tasas de retención y la eficiencia de costes (OECD, 2026; OECD, 2021). Al abordar a los agentes de IA como una capacidad coordinada, las escuelas de negocios pueden aprovechar su potencial transformador mientras preservan la interacción humana, el pensamiento crítico y la agencia del aprendiz que son fundamentales

para una educación de alta calidad (Erdmann & Toro-Dupouy, 2025; OECD, 2026; Miao et al., 2024).

Los líderes de escuelas de negocios deben utilizar la matriz de la Figura 8 para determinar el equilibrio adecuado entre la agencia humana y la orquestación automatizada.

FIGURA 8

Matriz de autonomía docente y orquestación de IA
Fuente: elaboración propia a partir de OECD (2026), OECD Digital Education Outlook 2026.



9

Capítulo 10.

Conclusiones



La revisión bibliográfica realizada en este informe técnico revela que los agentes de inteligencia artificial representan un profundo cambio estructural en el panorama de la educación superior, que va más allá de la mera entrega de contenido digital y se orienta a la creación de un ecosistema educativo receptivo y centrado en el estudiante (Ouyang et al., 2022; OECD, 2026). Un hallazgo principal de esta revisión es que, aunque los agentes de IA demuestran un potencial transformador significativo, su valor educativo es contingente y no automático (Ouyang et al., 2022; OECD, 2026). El mero despliegue de tecnología avanzada no garantiza una mejora del aprendizaje; en cambio, el éxito es producto de una compleja alineación sociotécnica (OECD, 2026; Ouyang et al., 2022).

La evidencia sugiere que la efectividad educativa está determinada por la sinergia entre la calidad tecnológica, la alineación pedagógica y la adaptación tarea-tecnología (Ouyang et al., 2022; OECD, 2026). Estos factores tecnológicos están mediados por características individuales, específicamente la alfabetización en IA, la autoeficacia y la confianza, así como por dimensiones organizativas como la preparación institucional, la supervisión humana y marcos de gobernanza sólidos (Erdmann & Toro-Dupouy, 2025; Miao & Holmes, 2023; OECD, 2023). Para las escuelas de negocios, la adopción no es un proceso uniforme; mientras que los estudiantes suelen centrarse en la utilidad y la comodidad, la adopción por parte del profesorado sigue un camino más gradual y escalonado que requiere evidencia clara de integridad pedagógica y de valor profesional (Toro-Dupouy & Erdmann, 2026).

Los datos empíricos resumidos en la Tabla 2 destacan que, aunque tanto estudiantes como profesores valoran la IA, sus limitaciones en la adopción difieren significativamente.

TABLA 2

Comparación intergrupar de resiliencia, percepción de valor e intención de adopción (Toro-Dupouy & Erdmann, 2026).

Fuente: (Toro-Dupouy & Erdmann, 2026). ITA = Intención de Adopción de la IA; PV = Percepción de Valor de la IA (Evaluación Global del Valor de la IA); RES = Resiliencia Institucional Percibida.

Estudio 1	Media	Desv. típica	Prueba t (valor p)
RES			4,74 (< 0,001)
Muestra completa	3,628	0,869	
Profesorado	3,388	0,894	
Estudiantes	3,825	0,798	
PV			1,27 (0,205)
Muestra completa	3,641	0,844	
Profesorado	3,577	0,854	
Estudiantes	3,693	0,834	
ITA			-1,83 (0,069)
Muestra completa	3,913	0,896	
Profesorado	4,010	0,909	
Estudiantes	3,833	0,879	

La integración de agentes de IA ofrece oportunidades significativas para mejorar la experiencia universitaria. Estas incluyen la personalización a gran escala, el apoyo académico 24/7, la retroalimentación formativa inmediata y contextual, y la automatización de tareas administrativas "laboriosas" para aumentar la productividad del profesorado (Erdmann & Toro-Dupouy, 2025; Roe & Perkins, 2024; OECD, 2021; OECD, 2023; OECD, 2026). A nivel institucional, los sistemas de agentes pueden mejorar los servicios estudiantiles y la eficiencia operativa mediante asesoramiento basado en datos y flujos de trabajo administrativos optimizados (OECD, 2026).

Por el contrario, estos beneficios deben equilibrarse con los riesgos sistémicos. La literatura destaca preocupaciones persistentes sobre el sesgo y la sobredependencia en la automatización, que pueden conducir a la "pereza metacognitiva" y a la erosión del pensamiento crítico (OECD, 2026).

Además, la falta de transparencia en la toma de decisiones algorítmicas, el potencial de alucinaciones, los riesgos para la privacidad, el sesgo algorítmico y la responsabilidad siguen, en gran medida, sin resolverse (Park et al., 2025; Miao & Holmes, 2023; OECD, 2023). Las decisiones de alta consecuencia que implican calificación, progresión y acciones disciplinarias requieren una gobernanza estricta de las personas en el bucle para evitar la pérdida de la agencia humana (Miao & Holmes, 2023; OECD, 2026; Miao & Cukurova, 2024).

Por lo tanto, las escuelas de negocios deberían priorizar el valor educativo y la agencia humana por encima de la automatización en sí misma (Kasneci et al., 2023; OECD, 2026). Los agentes de IA deberían implementarse como una capacidad para complementar, en lugar de reemplazar, a los educadores humanos, que son fundamentales en los aspectos sociales y relacionales del aprendizaje (OECD, 2026). La contribución del marco teórico integrativo propuesto radica en su capacidad para conectar estos factores tecnológicos, individuales, pedagógicos, organizativos y de gobernanza con los resultados reales de adopción y de aprendizaje (Erdmann & Toro-Dupouy, 2025).

Por último, queda claro que la evidencia sobre el agente de IA en la educación superior sigue surgiendo (Roe & Perkins, 2024; OECD, 2026). La mayoría de los estudios actuales son pilotos exploratorios o de corto plazo; por lo tanto, se necesita más investigación longitudinal, comparativa y causal para distinguir los aumentos temporales del rendimiento de la adquisición de habilidades duraderas (Ouyang et al., 2022; OECD, 2026). Este marco proporciona una base rigurosa para futuras investigaciones primarias que involucren a estudiantes, al profesorado y a responsables institucionales. Es probable que los agentes de IA se vuelvan cada vez más relevantes en la modalidad de e-learning, pero su adopción sostenible dependerá de la capacidad de una institución para combinar con éxito la innovación tecnológica con la calidad pedagógica, la supervisión humana continua y una gobernanza responsable (OECD, 2026; Miao & Holmes, 2023; Miao & Cukurova, 2024).

Al-Fraihat, D., Joy, M., Masa'deh, R., and Sinclair, J. (2020). Evaluating e-learning systems success: An empirical study. *Computers in Human Behavior*, 102, 67–86.

Batista, J., Mesquita, A., and Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, 15(11), 676.

Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.

DeLone, W. H., and McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9–30.

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., and Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology. *Information Systems Frontiers*, 21, 719–734.

Erdmann, A., and Toro-Dupouy, L. (2025). The influence of the institutional environment on AI adoption in universities: identifying value drivers and necessary conditions. *European Journal of Innovation Management*, 28(9), 4365–4398.

Garrison, D. R., Anderson, T., and Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2–3), 87–105.

Glikson, E., and Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.

Goodhue, D. L., and Thompson, R. L. (1995). Task–technology fit and individual performance. *MIS Quarterly*, 19(2), 213–236.

Holmes, W., Bialik, M., and Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.

Kasneci, E., Sessler, K., Küchemann, S., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.

- Lo, C. K. (2023).** What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410.
- Long, D., and Magerko, B. (2020).** What is AI literacy? Competencies and design considerations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*.
- Martin, F., Sun, T., and Westine, C. D. (2020).** A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, 159, 104009.
- Miao, F., & Cukurova, M. (2024).** *AI competency framework for teachers*. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- Miao, F., & Holmes, W. (2023).** *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Miao, F., Shiohira, K., & Lao, N. (2024).** *AI competency framework for students*. UNESCO. <https://doi.org/10.54675/JKJB9835>
- Mishra, P., and Koehler, M. J. (2006).** Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., and Qiao, M. S. (2021).** Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- OECD. (2021).** *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD Publishing. <https://doi.org/10.1787/589b283f-en>
- OECD. (2023).** *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- OECD. (2026).** *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024).** A systematic review of generative AI for teaching and learning practice. *Education Sciences*, 14(6), 636.
- Ouyang, F., Zheng, L., and Jiao, P. (2022).** Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27, 7893–7925.
- Panadero, E. (2017).** A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422.

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A*, 30(3), 286–297.

Park, J. (2025). A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. *Computers and Education: Artificial Intelligence*, 100487.

Roe, J., & Perkins, M. (2024). Generative AI and agency in education: A critical scoping review and thematic analysis. *arXiv preprint arXiv:2411.00631*.

Salloum, S. A., Alhamad, A. Q. M., Al-Emran, M., Monem, A. A., and Shaalan, K. (2019). Exploring students' acceptance of e-learning through the development of a comprehensive technology acceptance model. *IEEE Access*, 7, 128445–128462.

Tlili, A., Shehata, B., Adarkwah, M. A., et al. (2023). What if the devil is my guardian angel? ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15.

Toro-Dupouy, L., & Erdmann, A. (2026). Institutional resilience and AI adoption in a turbulent economic context: perceptions across stakeholders. *Transformations In Business & Economics*, 25(2 (68)), 55-82.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. <https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence>

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Zawacki-Richter, O., Marín, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial-intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, 39.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70.

OBS Business
School



www.obsbusiness.school

RED DE EDUCACIÓN SUPERIOR

 Planeta Formación y Universidades