

Un modelo de IA agéntica para la gestión integral del absentismo escolar

/

An Agentic AI Model for the Comprehensive Management of School Absenteeism.

Fernando M. Faci Lucia*

Inspector de educación jubilado, investigador independiente

DOI: <https://doi.org/10.23824/ase.v0i44.1027>

Resumen

El absentismo escolar continúa gestionándose de forma fragmentada y reactiva, lo que dificulta la prevención y reduce la equidad entre centros. Este trabajo presenta un modelo de inteligencia artificial agéntica orientado a integrar datos, automatizar tareas y mejorar la coordinación entre actores educativos mediante una arquitectura multiagente basada en el ciclo MAPE-K y con supervisión humana en puntos críticos. El modelo se desarrolla a partir de una revisión narrativa estructurada de literatura científica y normativa (2000–2025) y se articula en doce módulos funcionales que incorporan salvaguardas éticas, trazabilidad, estándares de interoperabilidad e integración por API. La propuesta se implementa en una arquitectura multinivel que comprende centro, zona y distrito, alineada con los protocolos educativos vigentes y con los requisitos del RGPD y del AI Act. El sistema automatiza la ingestión y el análisis de datos de asistencia, genera alertas tempranas, facilita la planificación personalizada y coordina intervenciones entre docentes, familias y servicios educativos, apoyado en puntos HITL, un marco de indicadores clave de desempeño y una hoja de ruta de implantación progresiva. Los resultados indican una elevada

* Identificador ORCID: <https://orcid.org/0000-0002-8086-003X>

viabilidad técnica y organizativa, aunque condicionada por la calidad de los datos, los riesgos de sesgo y la necesidad de una gobernanza institucional sólida. El modelo refuerza la detección temprana y la toma de decisiones basadas en evidencia, siempre bajo supervisión humana para garantizar equidad, transparencia y responsabilidad pública.

Palabras clave: Inteligencia artificial agéntica, sistemas multiagente; absentismo escolar, gobernanza educativa, supervisión humana en el bucle (HITL), aprendizaje organizativo, cumplimiento ético y normativo, equidad

Abstract

School absenteeism continues to be managed in a fragmented and reactive manner, which hampers prevention and reduces equity across schools. This paper presents an agentic artificial intelligence model designed to integrate dispersed data, automate operational tasks and enhance coordination among educational stakeholders through a multi-agent architecture grounded in the MAPE-K cycle and supported by human oversight at critical points. The model is developed on the basis of a structured narrative review of scientific and regulatory literature (2000–2025) and is organised into twelve functional modules that incorporate ethical safeguards, enhanced traceability, interoperability standards and API-based integration. The proposal is implemented through a multilevel architecture encompassing the school, local area and district levels, aligned with existing educational protocols and the requirements of the GDPR and the EU AI Act. The system automates the ingestion and analysis of attendance data, generates early-risk alerts, supports personalised planning and coordinates interventions among teachers, families and educational services. It integrates HITL oversight mechanisms, a framework of key performance indicators and a phased implementation roadmap. The findings indicate high technical and organisational feasibility, although dependent on data quality, potential algorithmic bias and the need for robust institutional governance. The model strengthens early detection and evidence-informed decision-making, while maintaining human supervision to ensure equity, transparency and public accountability.

Keywords: Agentic artificial intelligence, multi-agent systems, school absenteeism, educational governance; human-in-the-loop supervision (HITL), organizational learning, ethical and regulatory compliance, equity

Contenido

1. Introducción	4
2. Arquitectura del sistema multiagente	5
M1. Memoria (corto y largo plazo)	8
M2. Base de conocimiento (Knowledge Base)	9
M3. Razonamiento y toma de decisiones	9
M4. Planificación y descomposición de tareas	9
M5. Bucle de ejecución	10
M6. Uso de herramientas e integración por API	11
M7. Registro y bucle de retroalimentación.....	12
M8. Evaluación y pruebas	12
M9. Salvaguardas o filtros éticos y de seguridad (guardarraíles).....	13
M10. Colaboración multiagente	14
M11. Definición y seguimiento de objetivos	15
M12. Interfaz de lenguaje natural (LLM)	15
T1. Orquestador y supervisor	16
T2. Módulo de explicabilidad y transparencia	16
T3. Módulo de gobernanza y cumplimiento	17
3. Herramientas de software	20
4. Organización por niveles operativos	21
4.1. Arquitectura multinivel (visión rápida)	22
4.2. Matriz de los 12 componentes por nivel.....	22
4.3. Flujo operativo MAPE-K con escalado institucional	23
4.4. Puntos de supervisión humana (HITL)	25
4.5. Datos y gobernanza	26
4.6. Indicadores clave de desempeño (KPI)	27
4.7. Interoperabilidad técnica	28
4.8. Orquestación de agentes	29
4.9. Secuencia operativa de un caso.....	30
5. Conclusiones y futuras líneas de investigación.....	30
6. Método de revisión de la literatura	31
7. Observaciones finales.....	32

1. Introducción

El control del absentismo escolar sigue siendo una tarea fragmentada y reactiva, centrada en justificar ausencias *a posteriori*, lo que debilita la prevención (Balfanz y Byrnes, 2018). Como señala Esteban (2025), se requieren mecanismos coordinados y basados en evidencia, ya que la actual falta de integración de datos reduce la capacidad de anticipación y equidad. La respuesta exige sistemas capaces de conectar información dispersa y coordinar actores hacia un objetivo preventivo.

Un sistema multiagente de inteligencia artificial ofrece una alternativa estructural al automatizar la gestión de información en tiempo real, distribuyendo funciones entre agentes especializados que cooperan bajo supervisión humana y principios éticos (Wooldridge, 2009; Russell y Norvig, 2021). Esta convergencia tecnológica facilita el aprendizaje organizativo (Senge, 2006), permitiendo a la institución mejorar su respuesta, integrando detección temprana y rendición de cuentas para una gobernanza sostenible.

1.1. Alcance de la propuesta

Este trabajo presenta y fundamenta un modelo arquitectónico multiagente para la gestión integral del absentismo escolar, alineado con el ciclo **MAPE-K** (Monitor, Analyze, Plan, Execute, Knowledge), con puntos de intervención humana (**HITL**, *human-in-the-loop*) y desplegado en tres niveles (centro, zona y distrito). Se detallan: (i) los componentes y sus interfaces (observación/ingesta, análisis, planificación, ejecución, memoria y base de conocimiento, explicación y salvaguardas éticas y de seguridad o guardarraíles (*guardrails* en inglés); (ii) la gobernanza de datos conforme al Reglamento (UE) 2016/679, General de Protección de Datos (**RGPD**) y al Reglamento (EU) 2024/1689, sobre inteligencia artificial (**AI Act**), minimización, trazabilidad, bases de legitimación, evaluación de impacto, rendición de cuentas); (iii) la orquestación operativa (roles, flujos, acuerdo de nivel de servicio (**SLA**, por sus siglas en inglés *Service Level Agreement*), artefactos de registro y auditoría) y (iv) un cuadro de indicadores (**KPI**)¹, por sus siglas en inglés *Key Performance Indicators*) para seguimiento y evaluación de desempeño, equidad y cumplimiento. Se incluye una guía

¹ En este artículo el plural de siglas como API, KPI y LLM no varía, según la norma establecida en este idioma.

de interoperabilidad que cubre interfaces de programación de aplicaciones (**API**, en inglés *Application Programming Interfaces*), interoperabilidad de herramientas de aprendizaje (**LTI**, por sus siglas en inglés *Learning Tools Interoperability*), API de experiencias (**xAPI**, por sus siglas en inglés *Experience API*) y control de acceso basado en roles (**RBAC**, por sus siglas en inglés *Role-Based Access Control*), así como una hoja de ruta (*roadmap*) de implantación por fases para pilotos y escalado.

Quedan fuera del alcance de este artículo las explicaciones causales del absentismo; evaluación clínica, social o familiar de intervenciones; ensayos con estimación de impacto causal; prescripción normativa o de políticas públicas (más allá del cumplimiento regulatorio); presupuestación y gestión de personal; desarrollo y evaluación exhaustiva de modelos predictivos (salvo ejemplos y criterios mínimos); diseño curricular de prevención; y gestión clínica/asistencial de casos.

El modelo presupone que se dispone en el entorno de aplicación de (i) calidad y disponibilidad de datos de asistencia y contexto educativo con metadatos adecuados; (ii) infraestructura básica (sistemas de información educativos, identidad y control de accesos); (iii) acuerdos de intercambio y bases jurídicas para tratamiento de datos; (iv) capacidad organizativa para supervisar puntos HITL y auditorías.

Resultados esperados. La propuesta permite diseñar, gobernar y operar un sistema multiagente auditable y escalable; ofrece criterios de evaluación (KPI) y mecanismos de control para cumplimiento y equidad. La validación empírica de su impacto en tasas de absentismo se reserva a implementaciones piloto y estudios posteriores.

2. Arquitectura del sistema multiagente

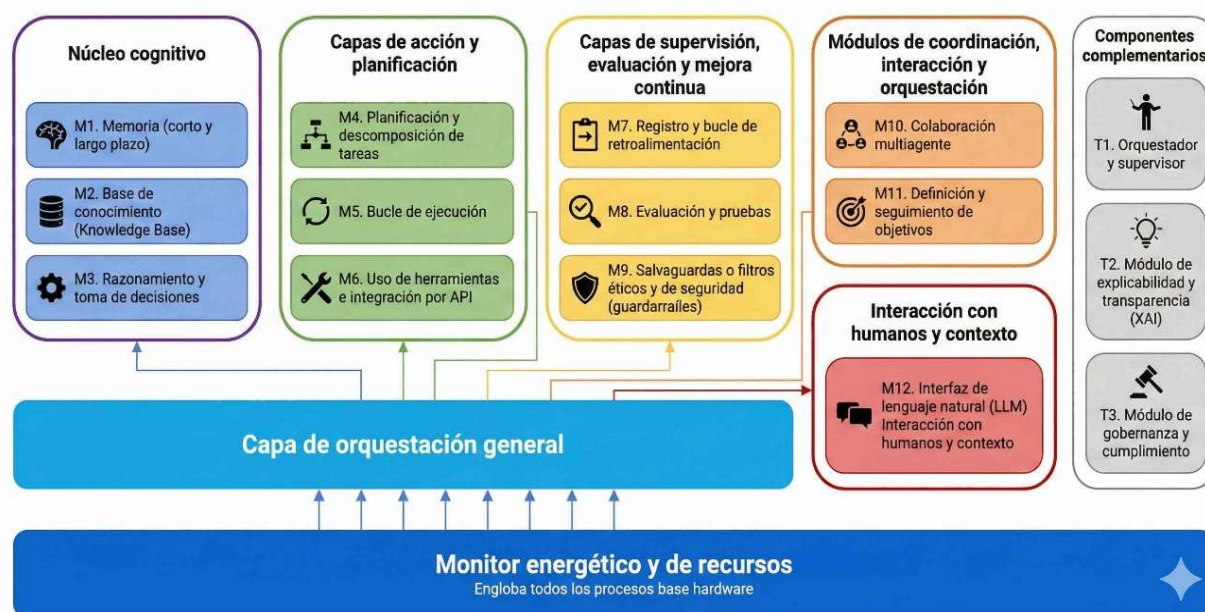
Según Wooldridge (2009), un agente autónomo percibe, razona y actúa para alcanzar metas sin intervención constante, operando bajo principios de racionalidad limitada (Simon, 1997) y buscando soluciones efectivas, no necesariamente óptimas. Estos sistemas combinan la autonomía individual (reactividad, proactividad) con la interacción cooperativa, donde la inteligencia global emerge de la coordinación de agentes especializados dentro de un ecosistema complejo. La construcción de un sistema agéntico de este tipo exige una **arquitectura integrada de software, hardware y agentes humanos**, donde la cooperación técnica y organizativa garantice la coherencia, la seguridad y la supervisión ética del conjunto (Russell y Norvig, 2021).

Recientemente, el impulso que ha tomado la inteligencia artificial, especialmente gracias al aprendizaje profundo, la computación distribuida y la integración multimodal, ha ampliado notablemente las posibilidades de estos agentes, que ahora pueden operar en entornos cada vez más complejos, dinámicos y abiertos (Russell y Norvig, 2021; Silver et al., 2021). Su aplicación ya no se limita a dominios simulados o controlados, sino que se extiende a contextos reales como la gestión educativa, la planificación urbana o la gobernanza digital, donde la cooperación entre agentes humanos y artificiales redefine las fronteras de la toma de decisiones (Zhu, Dastani, y Wang, 2024). En el ámbito educativo, estas capacidades permiten diseñar sistemas multiagente capaces de prevenir y controlar el absentismo escolar, mediante la detección temprana de patrones de riesgo, la coordinación automatizada de respuestas y el aprendizaje continuo a partir de los resultados de las intervenciones, todo ello manteniendo la supervisión ética y humana del proceso.

A continuación se presentan los **componentes esenciales de un sistema complejo multiagente**, entendidos como las piezas estructurales que permiten su funcionamiento coordinado, adaptativo y seguro. En este tipo de sistemas, la inteligencia no reside en un único elemento central, sino que emerge de la interacción entre múltiples módulos especializados, cada uno con funciones propias y mecanismos de comunicación estandarizados.

Para organizar el modelo, se ha seguido la estructura gráfica de un sistema multiagente propuesta por el arquitecto de Inteligencia artificial (*AI Architect*) Brij Kishore Pandey (2025), que permite visualizar la interrelación funcional entre los distintos componentes de la inteligencia artificial agéntica. Este esquema (ver figura 1) sirve como base para identificar los módulos esenciales y comprender su articulación interna, lo que facilita el tránsito hacia el análisis del núcleo cognitivo, donde reside la inteligencia del agente y se concentran los procesos de razonamiento, aprendizaje y toma de decisiones.

Figura 1. Componentes de un sistema de IA agéntica



Fuente: La imagen principal, elaboración propia, basado (octubre 2025) en la infografía de Pandey, B. K. (2025) titulada. *How Agentic AI Systems Actually Work (A Complete Blueprint)*, ubicada en <https://www.linkedin.com/in/brijkishorepandey/>. Con autorización de su autor (7 de noviembre de 2025)

La organización del sistema agéntico según este modelo responde a una lógica funcional inspirada en el ciclo cognitivo de la inteligencia artificial y en los principios de la arquitectura MAPE-K (Monitor, Analyze, Plan, Execute–Knowledge), que permite estructurar el comportamiento inteligente de forma modular, trazable y adaptable (Kephart y Chess, 2003). En la base se sitúa el **núcleo cognitivo**, donde la memoria, la base de conocimiento y el razonamiento constituyen la tríada esencial del procesamiento inteligente: recordar, comprender y decidir (Russell y Norvig, 2021). Sobre esta base se despliegan las **capas de acción y planificación**, responsables de traducir las decisiones en operaciones efectivas mediante planificación jerárquica y bucles iterativos de ejecución y ajuste (Ghallab, Nau, y Traverso, 2016). La siguiente capa incorpora los mecanismos de **supervisión, evaluación y mejora continua**, que garantizan la retroalimentación y el aprendizaje del sistema a partir de la experiencia (Argyris y Schön, 1978; Sutton y Barto, 2018). Los filtros éticos y de seguridad actúan como salvaguardas para prevenir sesgos y asegurar la conformidad normativa con el *AI Act* y el *Reglamento General de Protección de Datos* (High-Level Expert Group on Artificial Intelligence, 2019; European Parliament and Council of the European Union, 2024). Finalmente, los módulos de **coordinación, interacción y orquestación**, incluyendo la interfaz de lenguaje natural, la gobernanza, la explicabilidad y el control energético, integran la dimensión humana y organizativa, permitiendo que la

inteligencia artificial agéntica funcione como un sistema socio-técnico coherente, transparente y sostenible, en el que cada componente contribuye a la inteligencia global del conjunto (Floridi, 2021; Wooldridge, 2009). Además de todos estos módulos, hay dos elementos más que considerar: la capa de orquestación general y el monitor energético y de recursos.

Núcleo cognitivo (inteligencia del agente)

M1. Memoria (corto y largo plazo)

En un sistema de inteligencia artificial agéntica, la memoria constituye el eje que permite al agente mantener coherencia contextual y aprender con el tiempo. Su estructura dual (memoria de corto y largo plazo) reproduce, en términos computacionales, los procesos de retención y transferencia descritos por la neurociencia y la psicología cognitiva (Bransford, Brown, y Cocking, 2000), así como los mecanismos secuenciales de almacenamiento y recuperación estudiados en la arquitectura neuronal por Graves, Wayne y Danihelka (2014). Esta capacidad posibilita un razonamiento longitudinal sobre los patrones de absentismo, de modo que las intervenciones puedan basarse en evidencia acumulada y no en juicios aislados (Holmes, Bialik, y Fadel, 2019).

La memoria a corto plazo (Short-Term Memory, STM) actúa como espacio de trabajo temporal para la sesión actual, garantizando coherencia inmediata. Por su parte, la memoria a largo plazo (Long-Term Memory, LTM) almacena conocimiento experiencial persistente para el aprendizaje acumulativo y la personalización, persistiendo más allá de una sesión (Badia et al., 2020). Esta estructura dual posibilita el razonamiento longitudinal sobre patrones de absentismo y la evaluación de intervenciones (Holmes, Bialik, y Fadel, 2019).

Sin un sistema de memoria, un agente sería puramente reactivo y carente de continuidad: respondería a estímulos inmediatos sin integrar experiencia ni contexto. La combinación de ambas memorias, corta y larga, dota al sistema de la capacidad de mantener coherencia temporal, aprender de la experiencia y personalizar su conducta según los usuarios o entornos (Graves et al., 2014, pp. 4-6; Badia et al., 2020, pp. 509-512). En el ámbito del control del absentismo escolar, esta arquitectura permite tanto la reacción inmediata ante nuevas ausencias como un razonamiento

longitudinal sobre patrones de comportamiento, evolución de los casos y eficacia de las intervenciones aplicadas.

M2. Base de conocimiento (Knowledge Base)

La base de conocimiento funciona como el repositorio normativo y procedimental del sistema, donde se integran protocolos, legislación y registros educativos. A diferencia de un mero almacén de datos, constituye una red semántica que guía la toma de decisiones conforme a criterios verificables de equidad y cumplimiento (OECD, 2023a; UNESCO, 2021).

En el ámbito del absentismo escolar, la base de conocimiento puede incluir guías del Ministerio de Educación, legislación autonómica aplicable, el protocolo de intervención establecido y bases de datos institucionales que orientan las respuestas automatizadas o asistidas del sistema (Sevilla, Cuevas-Ruiz, Rello, y Sanz, 2025). Por ejemplo, ante un caso de ausencias reiteradas, el agente podría consultar la normativa autonómica sobre escolarización obligatoria y el protocolo de derivación a servicios sociales, asegurando que la respuesta se ajusta a los procedimientos oficiales y respeta los criterios de equidad y protección del menor.

M3. Razonamiento y toma de decisiones

Sobre estas capas, el módulo de razonamiento transforma la información en acción: aplica reglas explícitas cuando la normativa lo exige y modelos probabilísticos cuando se requiere predicción. Este razonamiento híbrido garantiza que las alertas o intervenciones conserven trazabilidad técnica y validez pedagógica (Russell y Norvig, 2021), al mismo tiempo que refuerza la eficiencia y adaptabilidad institucional derivadas de la automatización inteligente descritas por McAfee y Brynjolfsson (2017).

Capas de acción y planificación

M4. Planificación y descomposición de tareas

Es el componente encargado de traducir los objetivos estratégicos del sistema en acciones concretas y operativas. Actúa como un arquitecto del comportamiento agéntico: analiza la meta global, identifica los recursos y restricciones disponibles, y

desglosa el propósito general en una secuencia lógica de subtarear que puedan ser ejecutadas y monitorizadas.

Este proceso recurre a técnicas de *Hierarchical Task Network Planning* (HTN) y de razonamiento orientado a metas (*goal-based reasoning*), ampliamente documentadas en la literatura clásica de la inteligencia artificial (Russell y Norvig, 2021) y desarrolladas en profundidad en revisiones recientes sobre planificación jerárquica (Georgievski y Aiello, 2016).

Gracias a ello, el agente puede diseñar rutas dinámicas hacia la consecución de resultados, adaptándose a cambios contextuales o a la información que se va incorporando durante la ejecución (Ghallab, Nau, y Traverso, 2016). En el contexto del absentismo escolar, por ejemplo, ante el objetivo global de “reducir el absentismo en el centro”, el sistema puede generar un plan compuesto por tres fases: (1) detectar patrones de inasistencia mediante análisis predictivo, (2) activar intervenciones adaptadas al perfil del alumno y (3) evaluar el impacto de las medidas implementadas para ajustar futuras actuaciones.

M5. Bucle de ejecución

El bucle de ejecución es el mecanismo que permite al sistema agéntico llevar a la práctica las tareas planificadas, observar sus efectos y reajustar las acciones en función de los resultados obtenidos. Opera de forma iterativa: ejecuta una acción, evalúa su impacto y, si detecta desviaciones respecto al objetivo, modifica el plan o actualiza sus parámetros de actuación (Michie, 1963). Este ciclo continuo de *plan–act–evaluate–adapt* (ciclo PDCA; Deming, 1986) dota al agente de capacidad de aprendizaje operativo y de adaptación en tiempo real, evitando comportamientos rígidos o ineficientes (Wooldridge, 2009). En un sistema orientado al control del absentismo escolar, por ejemplo, el bucle de ejecución permite que, tras aplicar una intervención como el contacto con la familia o la derivación al orientador, el agente valore su eficacia y, si el absentismo persiste, reajuste automáticamente la estrategia para proponer nuevas acciones más adecuadas al contexto.

M6. Uso de herramientas e integración por API

La integración operativa mediante las API (*Application Programming Interfaces*) y servicios externos permite que el sistema agéntico se comunique y coopere con aplicaciones y servicios externos, extendiendo sus capacidades más allá de su propio entorno. Mediante las API, el agente accede en tiempo real a bases de datos institucionales, sistemas de gestión educativa, plataformas analíticas y servicios de mensajería automatizada. Esta integración, que es piedra angular de los ecosistemas digitales conectados, está expuesta de forma accesible para administraciones públicas en el *New European Interoperability Framework* y cuenta con evidencia aplicada en ecosistemas tecnológicos educativos en España (European Commission, 2017; García-Holgado, 2018). La interoperabilidad resultante garantiza que las decisiones y acciones del sistema se basen en información actualizada y que los resultados se integren sin fricciones en los flujos de trabajo existentes.

Para asegurar esta coherencia, se aplican **estándares internacionales** como LTI y xAPI, que permiten el intercambio seguro y estructurado de datos entre plataformas educativas, especialmente los LMS (*Learning Management Systems*), sistemas analíticos y agentes inteligentes. Estos estándares cumplen funciones complementarias dentro del ecosistema educativo digital. **LTI** actúa como un conector que facilita la integración fluida entre plataformas de gestión del aprendizaje y herramientas externas, evitando duplicidades de acceso y permitiendo la sincronización automática de resultados y calificaciones. **xAPI**, por su parte, funciona como un registro detallado de la actividad formativa, capaz de recopilar datos sobre el aprendizaje que ocurre tanto dentro como fuera del entorno virtual. Mientras **LTI** se enfoca en la interoperabilidad técnica, **xAPI** amplía la capacidad de análisis y seguimiento de la experiencia educativa, y su uso conjunto posibilita una visión integrada del aprendizaje y la intervención, esencial para sistemas agénticos centrados en la mejora continua. Hay que señalar que en este artículo, LMS se refiere a la plataforma docente (p. ej., Moodle/Canvas), distinta del SGE/SIS administrativo.

En el contexto del absentismo escolar, por ejemplo, el agente podría conectarse a un sistema de gestión educativa para verificar en tiempo real los registros de asistencia, generar automáticamente una alerta cuando detecta un patrón de ausencias relevantes y notificar al tutor o al equipo directivo mediante un servicio de mensajería

integrado (SMS o API). Tras la intervención, actualizaría el expediente del alumno con la información sobre la acción realizada y sus resultados, garantizando trazabilidad completa, coherencia de datos y coordinación efectiva entre los distintos actores educativos implicados.

Supervisión, evaluación y mejora continua

M7. Registro y bucle de retroalimentación

Este componente del sistema agéntico permite aprender de la experiencia acumulada y perfeccionar su comportamiento a lo largo del tiempo. Cada acción, decisión o error queda documentado junto con su resultado, generando una base empírica que el agente analiza para determinar qué estrategias producen mejores resultados y cuáles deben ajustarse. Este proceso transforma los datos de ejecución en conocimiento operativo, de modo que el sistema compara resultados esperados y reales, corrige sus modelos de decisión y optimiza sus respuestas futuras. En este sentido, la retroalimentación funciona como el eje del aprendizaje organizativo (Argyris y Schön, 1978) y del aprendizaje por refuerzo computacional (Sutton y Barto, 2018), convergiendo en enfoques recientes que sostienen que la señal de recompensa constituye el mecanismo esencial para la adaptación inteligente (Silver et al., 2021).

En el modelo de IA agéntica aplicado al absentismo escolar, este bucle permite, por ejemplo, registrar el impacto de distintas medidas (contacto familiar, tutoría individual o derivación a servicios sociales) y ajustar las recomendaciones futuras según la eficacia observada en casos similares, mejorando de forma progresiva la capacidad predictiva y de intervención del sistema.

M8. Evaluación y pruebas

Este apartado aborda los mecanismos que garantizan que el sistema agéntico opere con precisión, fiabilidad y coherencia antes de su implantación y durante su uso continuado. Este proceso implica someter los modelos y algoritmos a comprobaciones sistemáticas que permitan verificar si las predicciones, alertas o decisiones generadas se corresponden con los casos reales y cumplen los criterios de calidad esperados (European Commission, 2019; Jacobson, Levin, y Kapur, 2019). Incluye la validación de datos, la detección de sesgos y la comparación entre resultados previstos y

observados. En el caso del absentismo escolar, por ejemplo, la evaluación serviría para comprobar si las alertas emitidas por el sistema coinciden con los alumnos que efectivamente presentan absentismo crónico, garantizando que las intervenciones se basen en información veraz, trazable y estadísticamente sólida.

Además, este ciclo de verificación continua se alinea con las obligaciones de gestión de riesgos, registro y trazabilidad, documentación técnica y evaluación de impacto previstas en el Reglamento (EU) 2024/1689 (AI Act), e incorpora supervisión humana efectiva, transparencia y rendición de cuentas para su uso en contextos educativos de alto riesgo (European Parliament and Council of the European Union, 2024; European Commission, 2019).

M9. Salvaguardas o filtros éticos y de seguridad (guardarraíles)

Las salvaguardas éticas y de seguridad, también llamadas guardarraíles, actúan como controles que garantizan el uso responsable y seguro del sistema agéntico (High-Level Expert Group on Artificial Intelligence, 2019; European Commission, 2019). Su función es prevenir sesgos algorítmicos, respuestas inadecuadas o decisiones que vulneren derechos, así como evitar el uso indebido de los datos personales que el sistema maneja. Estas salvaguardas combinan mecanismos técnicos como el control de acceso basado en roles (RBAC), la seudonimización² y el registro de auditoría con procedimientos de supervisión humana orientados a la transparencia, la equidad y la rendición de cuentas (Floridi, 2021). Operan de forma preventiva, bloqueando recomendaciones sesgadas, invasivas o discriminatorias, y correctiva, registrando las decisiones y sometiéndolas a revisión para detectar posibles desviaciones.

En el contexto educativo, permiten asegurar que las intervenciones automatizadas sobre el absentismo escolar se ajusten a la normativa vigente y a los principios de protección del menor, evitando discriminaciones y garantizando que toda recomendación esté respaldada por criterios verificables y auditables.

² El término seudonimización hace referencia al tratamiento por el cual los datos personales “ya no pueden atribuirse a un interesado sin información adicional”, siempre que esa información adicional se conserve separada y bajo medidas técnicas y organizativas adecuadas (RGPD, art. 4.5)

Coordinación y cooperación

M10. Colaboración multiagente

Se basa en la cooperación entre distintos agentes autónomos que actúan de manera coordinada para alcanzar objetivos comunes dentro del sistema. Cada agente dispone de un rol, metas y responsabilidades bien definidas, lo que permite distribuir la carga cognitiva y operativa de forma eficiente. Este enfoque favorece la especialización y la toma de decisiones contextualizada, ya que cada agente aporta su competencia específica, analítica, pedagógica, comunicativa, administrativa o de supervisión, dentro de un marco de cooperación estructurado (Li y Tan, 2019; Herrera et al., 2020).

En el sistema multiagente orientado al control del absentismo escolar, los roles se articulan de manera complementaria: un agente analítico detecta patrones de riesgo a partir de los datos de asistencia y rendimiento; un agente de tutoría traduce esas alertas en estrategias de intervención personalizadas; un agente familiar mantiene la comunicación con los padres o tutores legales, fomentando su implicación; un agente de inspección garantiza el cumplimiento normativo, la equidad y la protección de datos; un agente administrativo o institucional coordina la información con otras plataformas educativas; y un agente de aprendizaje evalúa los resultados del conjunto para ajustar los modelos y mejorar su eficacia (Outay et al., 2024) (ver tabla 1).

Tabla 1. *Tipos de agentes*

Tipo de agente	Función principal	Rol operativo en el sistema
Analítico	Detectar patrones de riesgo y generar alertas a partir de los datos de asistencia y rendimiento.	Predictivo y de apoyo a la toma de decisiones.
De tutoría	Traducir las alertas en estrategias pedagógicas o intervenciones personalizadas.	Educativo y adaptativo; actúa directamente sobre el caso del alumno.
Familiar	Mantener comunicación con los padres o tutores legales y promover su implicación.	Comunicativo y de mediación; fortalece el vínculo escuela–familia.
De inspección	Garantizar el cumplimiento normativo, la equidad y la protección de datos.	Supervisor y ético; vela por la legalidad y la transparencia del sistema.
Administrativo o institucional	Coordinar la información con otras plataformas educativas y asegurar coherencia documental.	Operativo y de integración; gestiona flujos de datos y coordinación institucional.
De aprendizaje	Evaluar los resultados globales y ajustar los modelos o estrategias según la eficacia observada.	Metaevaluador y de mejora continua; optimiza el rendimiento general del sistema.

Fuente: Elaboración asistida por IA (ChatGPT-5 Thinking, octubre-2025); revisión y edición del autor.

No obstante, este tipo de colaboración conlleva también **riesgos inherentes**. La autonomía distribuida puede generar pérdida de trazabilidad o conflictos de objetivos entre agentes con prioridades distintas como, por ejemplo, entre eficiencia operativa y equidad educativa. Asimismo, los errores o sesgos de un agente pueden propagarse al resto de la red si no existen mecanismos de validación cruzada o protocolos de control. Por ello, la colaboración multiagente debe sustentarse en arquitecturas robustas de comunicación y aprendizaje cooperativo (Zhu, Dastani, y Wang, 2024), además de incorporar salvaguardas éticas (guardarraíles), trazabilidad de decisiones y supervisión humana que garanticen un equilibrio entre autonomía, transparencia y responsabilidad institucional.

M11. Definición y seguimiento de objetivos

Este componente define los objetivos operativos y supervisa su cumplimiento mediante KPI (*Key Performance Indicators*), los cuales traducen fines generales en metas cuantificables para medir la eficacia y ajustar estrategias en tiempo real (Kaplan y Norton, 1996; OECD, 2019). Permiten detectar desviaciones y mejorar decisiones, combinando la monitorización técnica con la intervención humana en la definición de las metas. Esto faculta a los agentes para optimizar acciones ante objetivos concretos, integrando métricas como las del PAE (Gobierno de Aragón y Ayuntamiento de Zaragoza, 2017).

Interacción con humanos y contexto

M12. Interfaz de lenguaje natural (LLM)

Desarrollada a partir de un modelo de lenguaje a gran escala (LLM, *Large Language Model*), enmarcado dentro de los modelos fundacionales de la inteligencia artificial generativa (*Foundation Models, Generative AI*), la interfaz de lenguaje natural constituye el principal canal de comunicación entre el sistema multiagente y los usuarios humanos: docentes, equipos directivos, inspectores y familias. Su objetivo es traducir la complejidad técnica del sistema en interacciones comprensibles, fluidas y contextualizadas. Gracias a modelos avanzados de procesamiento del lenguaje, puede interpretar preguntas formuladas en lenguaje natural, consultar la base de conocimiento y generar respuestas explicativas, recomendaciones pedagógicas o informes automáticos según las necesidades del interlocutor (OpenAI, 2023; Rae et al., 2021).

En el contexto escolar, esta interfaz permite a docentes, familias y directivos consultar riesgos, solicitar explicaciones sobre medidas aplicadas o pedir informes de indicadores mediante lenguaje natural, obteniendo respuestas estructuradas y adaptadas a su nivel de comprensión.

El LLM actúa como mediador cognitivo traduciendo el razonamiento técnico en explicaciones transparentes, esenciales para la decisión educativa (Floridi y Chiriatti, 2020; Bender et al., 2021). Mediante modelos especializados con razonamiento estructurado (*chain-of-thought*), el sistema adapta su lenguaje dinámicamente a cada perfil —técnico para inspectores o accesible para familias—. Esta capacidad exige estricta supervisión ética y de privacidad para garantizar la neutralidad y protección de datos, uniendo así la analítica del sistema con el juicio humano.

Componentes complementarios

Una vez descritos los módulos funcionales del sistema, se detalla a continuación su despliegue institucional en los tres niveles operativos.

T1. Orquestador y supervisor

El módulo orquestador constituye el núcleo operativo que asegura coherencia y estabilidad, sincronizando tareas y recursos para evitar redundancias, a modo de director que guía las funciones de cada agente (Wooldridge, 2009; Jennings y Bussmann, 2003). Más allá del control técnico, gestiona prioridades e integra la supervisión humana (*human-in-the-loop*) en decisiones sensibles, reforzando la confianza y rendición de cuentas (Shneiderman, 2020; Rahwan, 2018). En el ámbito del absentismo, coordina la acción entre agentes y ajusta dinámicamente las estrategias mediante monitoreo continuo (DeLoach y Valenzuela, 2021), garantizando así la alineación pedagógica y ética del sistema.

T2. Módulo de explicabilidad y transparencia

El módulo de explicabilidad (Explainability and Transparency Module, XAI) garantiza que el sistema multiagente no solo actúe con eficacia, sino también con transparencia. Siguiendo el marco propuesto por Doshi-Vélez y Kim (2017), la explicabilidad se entiende como la capacidad del sistema para ofrecer razones interpretables de sus decisiones.

En este modelo se emplean técnicas como *Feature importance*³, que identifica las variables que más influyen en una predicción, y métodos como **LIME**⁴ o **SHAP**⁵, que generan explicaciones locales sobre por qué un modelo toma una decisión concreta. Estas herramientas traducen el razonamiento algorítmico a un lenguaje comprensible para docentes y gestores, facilitando la revisión humana y el aprendizaje institucional (Guidotti et al., 2018; Long y Magerko, 2020).

T3. Módulo de gobernanza y cumplimiento

Este módulo se encarga de asegurar que el sistema multiagente opere dentro del marco legal, ético y organizativo establecido por las normativas vigentes, como el Reglamento General de Protección de Datos (RGPD) y el AI Act de la Unión Europea (European Parliament and Council of the European Union, 2016; European Parliament, 2024). Su función es gestionar las políticas de acceso, uso, retención y eliminación de datos, garantizando la privacidad, la trazabilidad y la seguridad de la información. Además, incorpora mecanismos para la evaluación de impacto en la protección de datos (DPIA) y la documentación de todas las decisiones automatizadas relevantes, permitiendo auditorías y revisiones por parte de las autoridades competentes. Este módulo actúa, en esencia, como el marco de control normativo y ético del sistema, asegurando que la inteligencia artificial se implemente de manera responsable, transparente y conforme a los principios de equidad, proporcionalidad y supervisión humana exigidos en los entornos públicos y educativos (High-Level Expert Group on Artificial Intelligence, 2019).

³ **Feature importance**: es un conjunto de métodos que cuantifican cuánto contribuye cada variable al resultado de un modelo predictivo. Su funcionamiento suele basarse en medir cómo cambia el rendimiento del modelo cuando se altera o se elimina una variable determinada. Si al modificar una variable el modelo pierde precisión, se interpreta que esa variable es importante; si el rendimiento apenas cambia, su relevancia es baja. El resultado final es un ranking que indica qué variables son, en promedio, más influyentes en las predicciones. Este enfoque puede verse afectado por variables muy correlacionadas entre sí o con una cardinalidad muy alta número muy grande de valores distintos, lo que puede distorsionar la estimación de la importancia real.

⁴ **LIME**: explicación local (caso por caso) que aproxima el modelo alrededor de una predicción concreta para indicar qué variables la empujaron. Útil para justificar decisiones individuales; puede variar con pequeños cambios en el vecindario analizado.

⁵ **SHAP**: descompone cada predicción en la contribución de cada variable respecto a una base; las contribuciones suman la diferencia y, al promediarlas en valor absoluto, ofrecen una importancia global más estable.

Monitor energético y de recursos

Este módulo supervisa en tiempo real métricas de hardware para optimizar el consumo computacional y económico. Al ajustar dinámicamente la carga de trabajo, equilibra rendimiento y sostenibilidad según los principios de *Green AI* (Schwartz et al., 2020), asegurando una infraestructura educativa eficiente y alineada con los objetivos europeos de sostenibilidad digital.

Capa de orquestación general

La capa de orquestación general es el marco que integra y coordina todos los módulos del sistema, articulando un flujo de control explícito y gestionando las dependencias entre agentes y servicios.

El **marco de integración** constituye la infraestructura que conecta y sincroniza los distintos módulos del sistema multiagente, garantizando la coherencia funcional y el intercambio fluido de información entre ellos. Se apoya en herramientas de orquestación como AutoGen, que permiten encadenar procesos complejos desde la recuperación de conocimiento y el razonamiento hasta la ejecución y el *feedback*, dentro de un mismo flujo estructurado (Qin et al., 2024; Liang et al., 2024). Estas plataformas actúan como un sistema nervioso central, gestionando las llamadas entre agentes, la transferencia de contexto, la activación de las API externas y la trazabilidad de cada paso. El ejemplo de AutoGen (conversación multiagente) ante un caso de absentismo podría ser el siguiente⁶:

Agentes implicados en el proceso orquestación

Orchestrator: coordina todo el proceso; asigna tareas, integra las respuestas y valida la coherencia del resultado.

DataAgent: accede a la base de datos del sistema y al vector store (almacén semántico) para recuperar los indicadores relevantes del alumno.

ExplainerAgent: utiliza técnicas de explicabilidad (SHAP o LIME) para interpretar el modelo y generar una explicación textual sobre los factores que influyeron en la alerta.

PolicyAgent: revisa el contenido desde el punto de vista de la normativa (RGPD, AI Act, protocolos de centro), verificando que no se expongan datos personales sensibles.

Humano (Tutor): representa la intervención humana final (HITL); revisa el borrador, introduce ajustes y valida la comunicación antes de enviarla.

⁶ Fuente: Elaboración asistida por IA (ChatGPT-5 Thinking, octubre-2025); revisión y edición del autor

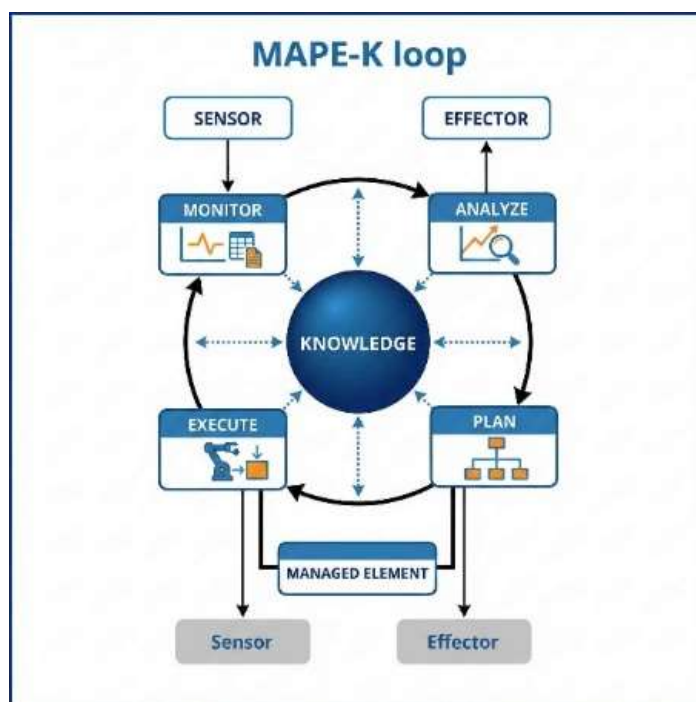
Mini-transcripción del proceso:

1. Trigger o Disparador. El modelo M_abs v1.3⁷ lanza una alerta cuando la probabilidad de absentismo crónico $\geq 0,35$ en una ventana de 4 semanas. Se activa HITL-2. Criterio mínimo: precisión $\geq 0,80$.
2. Orchestrator → DataAgent: “Necesito variables clave del Alumno A (semana 42).”
3. DataAgent → Orchestrator: “Entrego faltas=5, retrasos=3, caída_notas=-1,5, incidencias=2...”
4. Orchestrator → ExplainerAgent: “Genera **explicación local** de la alerta y un párrafo de intervención.”
5. ExplainerAgent → Orchestrator: “Explicación: mayor peso de ausencias recientes (+0,29) y retrasos (+0,11); amortigua asistencia histórica (-0,03). Propuesta: tutoría breve y llamada a familia.”
6. Orchestrator → PolicyAgent: “**Valida privacidad/tono/adecuación y aplica política de minimización/anonimización** antes de cualquier difusión fuera del centro.”
7. PolicyAgent → Orchestrator: “**Cumple RGPD/AI Act**. Seudonimiza identificadores; **elimina atributos sensibles** (dirección, teléfono) en mensajes a zona/distrito. Añade ‘sin datos de terceros’.”
8. Orchestrator → Human (Tutor): “Borrador listo para revisión.”
9. Human: “Aprobado con este ajuste final.”

Este esquema sigue la lógica del **ciclo MAPE-K** (*Monitor, Analyze, Plan, Execute - Knowledge*) (figura 2), que estructura la monitorización, el análisis, la planificación y la ejecución apoyadas en una base de conocimiento común para la adaptación continua (Kephart y Chess, 2003). Gracias a este marco, el sistema puede comportarse como una red integrada de agentes cooperativos, en lugar de un conjunto fragmentado de componentes aislados, facilitando la automatización segura, modular y transparente de los procesos de prevención e intervención ante el absentismo escolar.

⁷ La denominación M_abs v1.3 es una denominación de ejemplo para un modelo predictivo de absentismo que estima la probabilidad de que un alumno entre en absentismo crónico en la próxima ventana (p. ej., 4 semanas). Es el referente técnico que dispara la alerta del flujo AutoGen.

Figura 2. Representación del MAPE-K loop



Fuente: Elaboración propia por medio de Gemini Pro-3.0, siguiendo el modelo gráfico de Abbas, N., Andersson, J., y Löwe, W. (2010).

El **flujo de control** estructura la secuencia lógica del sistema, partiendo del contexto histórico y normativo hacia las fases de razonamiento, planificación y ejecución, para concluir con el ajuste basado en resultados. Este proceso se alinea con el ciclo MAPE-K, actualizando el conocimiento tras cada iteración para reforzar la coherencia, la adaptabilidad y la auditabilidad del conjunto (Kephart y Chess, 2003).

A su vez, el **bucle de aprendizaje continuo** permite al sistema evolucionar refinando sus modelos de predicción y decisión a partir de los aciertos y errores detectados en ciclos previos. Este mecanismo asegura una mejora adaptativa y supervisada, incrementando progresivamente la precisión de las alertas y la alineación ética de las intervenciones con los objetivos pedagógicos establecidos.

3. Herramientas de software

Cada módulo teórico encuentra su correspondencia práctica en un entorno de desarrollo rápido, cuya relación actualizada se puede obtener mediante consulta

técnica⁸. En este esquema, la capa MAPE-K actúa como el marco de orquestación y control inteligente que integra dichos componentes, apoyándose en herramientas de sincronización (como LangGraph) y monitoreo de recursos. Para garantizar la gobernanza ética y el cumplimiento del RGPD y el AI Act, esta arquitectura prioriza soluciones autoalojadas y módulos de anonimización que aseguren la soberanía del dato (Kephart y Chess, 2003; Liang et al., 2024).

4. Organización por niveles operativos

Se propone el modelo de organización que este autor ha impulsado en la comunidad Autónoma de Aragón, España, desde el año 2000.

En la ciudad de Zaragoza existen más de 300 centros educativos de Educación Infantil, Primaria, Secundaria o Formación Profesional. El Programa de Actuación para la Prevención del Absentismo Escolar firmado entre el Gobierno de Aragón y el Ayuntamiento de Zaragoza (2017) centra su intervención en la Educación Infantil, Educación Primaria y Educación Secundaria y divide la ciudad en diez zonas de intervención urbana específicas. La gestión del distrito educativo de Zaragoza se coordina desde un único Servicio Provincial, dependiente del Departamento de Educación del Gobierno de Aragón.

En el contexto del sistema educativo en Aragón, la estructura propuesta se ajusta plenamente al modelo ya establecido de prevención, control y seguimiento del absentismo escolar, regulado por la Orden ECD/331/2018, de 27 de febrero, del Departamento de Educación, Cultura y Deporte, por la que se determinan las actuaciones y la coordinación entre administraciones públicas en esta materia (BOA nº 52, de 14 de marzo de 2018).

Así, el nivel denominado en el modelo agéntico como Centro se corresponde con el centro educativo o unidad escolar, donde se realiza la detección inicial de las

⁸ El prompt para obtener la tabla puede ser el siguiente: Generate a table titled Recommended Tools for Each Module. The table must contain exactly two columns, Module and Recommended Tools. For each module, identify and list the most widely used, mature and industry-adopted tools, libraries, frameworks or services available at the moment of generating the response. Use real prevalence and best-practice criteria, avoiding any fictitious technologies. The modules are: M1 Memory (STM/LTM), M2 Knowledge Base, M3 Reasoning and Decision Making, M4 Planning and Decomposition, M5 Execution Loop, M6 Tool Use and API Integration, M7 Logging and Feedback Loop, M8 Evaluation and Testing, M9 Guardrails and Safety, M10 Multi-Agent Collaboration, M11 Goal Definition and Tracking (KPI), M12 Natural Language Interface (LLM). Present the output strictly as a Markdown table without adding explanations or commentary.

ausencias, el registro en las aplicaciones corporativas (SIGAD o SIRGA) y la primera comunicación con las familias. El nivel intermedio, equivalente a la Zona, se identifica con la Comisión Local o Zonal de Absentismo Escolar, órgano de coordinación interinstitucional en el que participan los centros, los servicios sociales municipales, la policía local y otras entidades del entorno. Este nivel permite el análisis compartido de casos y la derivación de situaciones complejas. Finalmente, el nivel superior, equivalente al Distrito, encuentra su correlato en el Servicio Provincial de Educación o en las comisiones provinciales de Absentismo, que asumen funciones de evaluación, consolidación de datos, definición de políticas y enlace con la Inspección de Educación.

4.1. Arquitectura multinivel (visión rápida)

- **Centro (micro):** detección diaria, contacto con familias, registro de intervenciones, validación de casos.
- **Zona escolar (meso):** agregación y análisis comparativo entre centros del área; coordinación interservicios (EOEP/Orientación, servicios sociales, transporte, salud).
- **Distrito educativo (macro):** gobierno del dato, estándares, auditoría, evaluación de impacto, objetivos y financiación por resultados.

En Zaragoza, este esquema se implementa a través del PAE (2017), con los niveles Centro–Comisión de Zona–Comisión Técnica Provincial, aprobados en Comisión Mixta el 04-04-2017

4.2. Matriz de los 12 componentes por nivel

La siguiente tabla 3 muestra la distribución funcional de los doce componentes del sistema agéntico en los tres niveles de actuación: centro educativo, comisión local o zonal y servicio provincial, evidenciando cómo cada capa institucional asume responsabilidades diferenciadas pero complementarias dentro de una arquitectura común orientada a la detección, intervención y evaluación del absentismo escolar. Se expresan los tres niveles de participación funcional por componente y nivel institucional:



-  **Actuación plena:** el componente opera activamente en ese nivel (ejecución o decisión directa).
-  **Actuación parcial:** el componente actúa en modo consultivo, de apoyo o de consolidación.

Tabla 3. *Matriz de componentes por nivel*

Componente	Centro educativo	Comisión local o zonal	Servicio provincial / distrito educativo
M1. Memoria (corto y largo plazo)	 Registro de asistencia, incidencias y actuaciones del alumno.	 Almacenamiento temporal de casos y modelos de intervención.	 Archivo histórico pseudonimizado y retención conforme a normativa.
M2. Base de conocimiento	 Protocolos internos, normativa autonómica y causas de absentismo.	 Banco de recursos y rutas de derivación interinstitucional.	 Corpus normativo, guías oficiales y ontologías educativas.
M3. Razonamiento y toma de decisiones	 Modelos de riesgo y umbrales ajustados al contexto del centro.	 Priorización de casos complejos y coordinación interinstitucional.	 Evaluación de impacto, análisis de equidad y simulaciones políticas.
M4. Planificación y descomposición	 Diseño de planes individualizados (alerta → entrevista → pacto → seguimiento).	 Planes y campañas zonales coordinadas entre centros.	 Planificación estratégica anual, metas y asignación de recursos.
M5. Ciclo de ejecución	 Aplicación de medidas, registro de resultados y comunicación con familias.	 Coordinación de equipos y programas intercentros.	 Supervisión y evaluación de programas territoriales.
M6. Uso de herramientas e integración (API)	 Conexión con SGE y mensajería a familias.	 Integración con servicios sociales, salud escolar y transporte.	 Enlace con padrón, becas, ayudas y sistemas estadísticos autonómicos.
M7. Registro y retroalimentación	 Bitácora de casos y resultados; lecciones aprendidas.	 Comparación intercentros y evaluación de medidas.	 Telemetría global, análisis de fallos y actualización de modelos.
M8. Evaluación y verificación	 Revisión de alertas y validación de indicadores.	 Evaluación comparativa y análisis de efectividad.	 Auditorías externas y metaevaluación institucional.
M9. Salvaguardas y seguridad (ética y datos)	 Gestión de consentimientos y RBAC.	 Pseudonimización, DPIA zonal y acuerdos de intercambio.	 Auditorías éticas, DPIA maestro y cumplimiento RGPD/AI Act.
M10. Colaboración multiagente	 Agentes: Monitor, Tutoría y Familia.	 Agentes: Coordinador y Analítica zonal.	 Agentes: Gobernanza, Ética/Datos y Evaluación.
M11. Definición y seguimiento de objetivos	 Los KPI del centro: absentismo, puntualidad, respuesta a alertas.	 Los KPI zonales: brechas territoriales y efectividad de medidas.	 Los KPI globales: reducción anual, equidad y eficiencia institucional.
M12. Interfaz de lenguaje natural (asistente virtual)	 Asistente para tutores y jefatura: explicación de alertas y recomendaciones.	 Asistente zonal: informes comparativos y actas automáticas.	 Asistente distrital: informes ejecutivos y resúmenes de impacto.

Fuente: Elaboración asistida por IA (ChatGPT-5 Thinking, octubre-2025); revisión y edición del autor.

4.3. Flujo operativo MAPE-K con escalado institucional

El sistema agéntico se organiza siguiendo el ciclo MAPE-K escalado en tres niveles (centro, zona escolar y distrito educativo) para garantizar coherencia vertical, supervisión humana y aprendizaje institucional continuo.

4.3.A. MAPE-K: Monitor (Centro)

El **nivel de centro** constituye la base operativa del sistema agéntico y el punto más cercano a la realidad educativa cotidiana. Es aquí donde se produce la observación directa de los comportamientos de asistencia y donde la inteligencia artificial adquiere sentido práctico al integrarse con la acción pedagógica y tutorial.

El agente realiza una ingesta diaria automatizada desde el SGE para construir una visión precisa de la asistencia. A partir de reglas locales y modelos predictivos, genera alertas tempranas que identifican casos antes de que se consoliden. Bajo supervisión humana (HITL-1), el tutor valida la pertinencia de la señal y decide la intervención inicial, asegurando que la decisión mantenga su raíz pedagógica

El nivel de centro representa la primera capa de inteligencia situacional del sistema: combina automatización y análisis predictivo con supervisión humana, asegurando que las decisiones iniciales mantengan su raíz pedagógica y ética. Por ejemplo, el disparador operativo se rige por el **baremo local de faltas no justificadas** (unidad absoluta, no porcentaje), que activa revisión HITL-1 y comunicación a familias (PAE, Gobierno de Aragón y Ayuntamiento de Zaragoza, 2017).

4.3.B. MAPE-K: Analyze (Centro → Zona)

El nivel analítico conecta la observación local con la comprensión sistémica para detectar patrones de escala mayor. Ante riesgos críticos o recurrentes, el sistema examina en profundidad las causas probables —familiares, sociales o escolares— para distinguir tendencias estructurales. Posteriormente, genera informes anonimizados para la zona escolar que permiten contrastar tendencias territoriales bajo estrictas garantías de confidencialidad. Este proceso transforma los datos locales en conocimiento compartido, facilitando la anticipación de problemas y orientando intervenciones equitativas basadas en evidencia.

4.3.C. MAPE-K: Plan (Centro, con apoyo de Zona)

La fase de planificación traduce el diagnóstico analítico en una hoja de ruta individualizada que establece medidas proporcionales y una secuencia de apoyo verificable. En situaciones complejas condicionadas por factores estructurales, el plan se elabora conjuntamente con servicios externos para evitar duplicidades. Todo el

proceso es validado por la coordinación zonal (HITL-2), que prioriza la vulnerabilidad y garantiza la equidad, conectando así la evidencia tecnológica con una decisión final humana y contextualizada.

4.3.D. MAPE-K: Execute (Centro / Zona)

La fase de ejecución materializa la planificación en intervenciones medibles. El centro educativo aplica y registra las medidas inmediatas, como avisos o tutorías, para facilitar el seguimiento, mientras que la zona escolar coordina acciones conjuntas ante patrones sistémicos comunes que exceden la capacidad individual. De este modo, la inteligencia analítica se traduce en una acción educativa coordinada, garantizando que las decisiones basadas en datos tengan un impacto verificable en la asistencia.

4.3 E. MAPE-K: Knowledge (Distrito)

El nivel de distrito actúa como eje de consolidación, transformando la información operativa en conocimiento institucional para orientar la política educativa. Mediante la gestión de registros longitudinales, la calibración de umbrales y el análisis de KPI, convierte la experiencia acumulada en recursos estructurados que permiten ajustar políticas y redistribuir recursos para favorecer la equidad. Este proceso culmina con la supervisión de la autoridad educativa (HITL-3), que valida ética y metodológicamente los cambios en los modelos, garantizando que el sistema aprenda de su práctica bajo principios de coherencia normativa y responsabilidad pública.

4.4 Puntos de supervisión humana (HITL)

Estos tres niveles materializan el principio *human-in-the-loop* (HITL), asegurando que la IA apoye el juicio profesional sin sustituirlo y garantizando el control ético del proceso. En **HITL-1 (Centro)**, el profesorado y la dirección validan las alertas y deciden la intervención inmediata, asegurando así la proximidad pedagógica y el conocimiento contextual (Spillane, 2006).

El nivel **HITL-2 (Zona)** coordina la equidad y coherencia entre centros mediante la colaboración de equipos directivos, orientadores y técnicos municipales. Este trabajo permite contextualizar las alertas según el entorno social, tal como evidencia el PAE (Gobierno de Aragón y Ayuntamiento de Zaragoza,

2017), promoviendo además la rendición de cuentas horizontal (Earl y Timperley, 2015) .Finalmente, **HITL-3 (Distrito)** corresponde a la inspección y dirección provincial, encargadas de evaluar la calidad metodológica y el cumplimiento normativo (AI Act, 2024; RGPD, 2016) bajo principios de transparencia (UNESCO, 2021) . Este nivel coordina recursos y garantiza la equidad territorial (PAE, 2017) . Esta gobernanza distribuida combina autonomía profesional y control ético (Harris, 2014), reforzando la confianza institucional al mantener el criterio humano en las decisiones críticas (Shneiderman, 2020).

4.5. Datos y gobernanza

La gobernanza de datos constituye un elemento estructural que asegura la trazabilidad y el cumplimiento normativo mediante tres capas jerárquicas, equilibrando el uso educativo y la protección de derechos conforme al RGPD y el AI Act. La validez de los registros se verifica mediante el control del ciclo de vida y la adopción del principio de *Data Protection by Design and by Default*, integrando la privacidad y minimización como salvaguardas intrínsecas (European Data Protection Board, 2020).

Siguiendo a Floridi (2021) y Mittelstadt (2019), quienes enfatizan que la gobernanza requiere supervisión institucional y transparencia más allá de la técnica , el sistema implementa un modelo multinivel: el **Centro** mantiene un modelo de datos mínimo para la acción inmediata ; la **Zona** consolida información anonimizada para favorecer la comparabilidad territorial; y el **Distrito** gestiona un repositorio pseudonimizado con auditoría centralizada mediante una DPIA maestra.

El modelo se sustenta en cinco principios rectores: minimización de datos (solo lo necesario); finalidad exclusiva educativa; trazabilidad de decisiones verificables; derecho a explicación ante decisiones automatizadas; y evaluación sistemática de sesgos socioeconómicos o de género (figura 3).

Figura 3: Modelo de gobernanza y principios rectores



Fuente: Elaboración propia, representación asistida por Gemini Pro-3.0

De este modo, el sistema consolida una gobernanza de datos madura y responsable, alineada con la Recomendación sobre la Ética de la IA de la UNESCO (2021) y con los criterios del Comité Europeo de Protección de Datos sobre el tratamiento de datos personales en el desarrollo y despliegue de modelos de IA (European Data Protection Board., 2024), garantizando que la IA educativa refuerce la confianza, la equidad y la rendición de cuentas institucional. Este modelo, además, facilita evaluaciones longitudinales de impacto y la consolidación de infraestructuras educativas interoperables con reutilización de datos a escala de sistema, como destaca la OCDE en el *Digital Education Outlook 2023* (OCDE, 2023b; Vincent-Lancrin & González-Sancho, 2023⁹)

4.6. Indicadores clave de desempeño (KPI)

Los indicadores clave de desempeño permiten medir de forma objetiva la eficacia y coherencia del sistema agéntico en sus distintos niveles de actuación. Tal como se recoge en el cuadro siguiente (figura 4), estos indicadores traducen la actividad operativa en evidencia verificable sobre el impacto de las intervenciones, facilitando el seguimiento continuo, la comparación entre territorios y la mejora progresiva de las políticas educativas.

⁹ La cita de Vincent-Lancrin & González-Sancho hace referencia a un capítulo de OCDE 2023b

Figura 4. Ejemplos de indicadores principales por nivel



Fuente: Elaboración propia, representación asistida por Gemini Pro-3.0

Es crucial distinguir indicadores de resultado agregado, como el abandono escolar prematuro (AEP), útil para políticas comparadas pero con rezago temporal y no accionable para intervención inmediata (Faci, 2011; OECD, 2023c), de los indicadores de proceso. Para la gestión operativa del absentismo se requieren métricas de caso como la tasa de absentismo crónico ($\geq 10\%$), asistencia regular y alertas tempranas, que permiten la acción preventiva directa.

4. 7. Interoperabilidad técnica

La interoperabilidad del sistema permite la comunicación segura entre plataformas mediante estándares abiertos como LTI 1.3 y xAPI, facilitando el intercambio de datos sin duplicidades (1EdTech Consortium, 2019; Advanced Distributed Learning Initiative, 2017). Estas integraciones utilizan eventos estandarizados, autenticación única (SSO) y RBAC (Ferraiolo y Kuhn, 1992), siguiendo las recomendaciones del European Data Protection Board (2020). Para asegurar la coherencia semántica, se adopta el European Learning Model (ELM), que provee un idioma digital común entre centro, zona y distrito (Europass, 2025).

La evaluación se estructura en tres dimensiones —eficiencia operativa, impacto educativo y equidad territorial— (figura 5), inspiradas en el Learning Compass 2030 (OECD, 2019) y la Recomendación sobre la Ética de la IA (UNESCO, 2021), incorporando criterios éticos para detectar sesgos (Floridi et al, 2018). Finalmente, la orquestación coordina los agentes y niveles mediante retroalimentación continua,

configurando un sistema transparente y adaptativo para la supervisión educativa (Ifenthaler y Yau, 2020).

Figura 5. Esquema de la interoperabilidad



Fuente: Elaboración propia, representación asistida por Gemini Pro-3.0

4.8. Orquestación de agentes

El sistema se organiza en torno a agentes especializados que cooperan entre los distintos niveles institucionales. Cada agente cumple una función específica dentro del ciclo MAPE-K, de modo que el conjunto opera como una red inteligente y distribuida. Esta estructura favorece la colaboración, la transparencia en las decisiones y la adaptación del sistema a las necesidades de cada territorio (ver tabla 5).

Tabla 5. Agentes principales y función, por nivel

Nivel	Agentes principales	Función
Centro	Agente Monitor, Agente Tutoría, Agente Explicador	Detección inicial, seguimiento individual y comunicación con familias y docentes.
Zona	Agente Coordinador, Agente Analítica	Gestión intercentros, análisis comparativo y activación de medidas colectivas.
Distrito	Agente Gobernanza, Agente Ética/Datos, Agente Evaluación	Definición de políticas, control ético y evaluación global del sistema.

Fuente: Elaboración propia

En conjunto, esta orquestación multinivel permite que la inteligencia artificial agéntica combine conocimiento local y supervisión central, asegurando una gobernanza educativa más eficiente, transparente y ética, orientada a la equidad.

4.9. Secuencia operativa de un caso

La gestión de incidencias sigue una secuencia trazable: detección automática de patrones anómalos, validación inicial (HITL-1) por el tutor con medidas tipo A, y escalado a intervención zonal (HITL-2) si no hay mejora en diez días. Si el absentismo persiste por causas estructurales, se notifica al distrito (HITL-3) para movilizar recursos. Finalmente, los resultados retroalimentan los modelos predictivos, cerrando el ciclo de mejora continua.

5. Conclusiones y futuras líneas de investigación

Es técnicamente viable desarrollar un sistema agéntico para la detección y control del absentismo escolar. Los componentes esenciales (memoria dual, razonamiento, planificación, aprendizaje continuo e interacción natural) están disponibles y pueden integrarse en arquitecturas interoperables y auditables.

La propuesta alcanza un nivel de grado de desarrollo suficiente para su implementación progresiva. Su estructura permite combinar la reacción inmediata ante ausencias con el análisis longitudinal de patrones, ofreciendo un equilibrio entre eficiencia operativa y comprensión contextual.

El principal desafío es institucional: se requiere un marco central que coordine la implantación garantizando una gobernanza ética. Si bien las ventajas incluyen detección temprana y fundamentación de políticas en evidencia, la dependencia excesiva de modelos automatizados conlleva riesgos de sesgo. Por ello, es imperativo establecer mecanismos de supervisión humana y auditoría continua que preserven el juicio profesional y la calidad de los datos, convirtiendo la tecnología en una herramienta responsable al servicio de la equidad.

Las futuras líneas de investigación deberían centrarse en evaluar empíricamente el impacto del modelo mediante pilotos controlados y análisis longitudinales que examinen la precisión predictiva y la eficacia sostenida de las intervenciones, así como en realizar auditorías de equidad que detecten posibles sesgos en la priorización y recomendación de casos. También resulta necesario estudiar la consistencia de los puntos de supervisión humana, explorar la integración de nuevas fuentes de datos que amplíen la capacidad analítica sin comprometer la privacidad y analizar la

transferibilidad del sistema a otros contextos educativos. Finalmente, sería relevante avanzar en estudios de coste-efectividad que permitan estimar la sostenibilidad institucional del despliegue.

La implicación activa de la inspección de educación resulta decisiva para garantizar la coherencia metodológica, la supervisión ética y la solidez institucional necesarias para que un sistema agéntico de gestión del absentismo pueda desplegarse con rigor, equidad y eficacia.

6. Método de revisión de la literatura

Para la confección del artículo se realizó primero una revisión narrativa estructurada orientada a identificar marcos, arquitecturas y evidencias prácticas sobre IA agéntica y sistemas multiagente aplicados a la gestión del absentismo en contextos educativos, incluyendo gobernanza de datos, cumplimiento normativo y evaluación de impacto.

Se consultaron bases y catálogos como Google Scholar, ERIC, OECD iLibrary, UNESCO Digital Library, arXiv, y literatura institucional (Unión Europea, European Data Protection Board, Agencia Española de Protección de Datos) Adicionalmente, se empleó ATLAS.ti Paper Search (Beta) para acelerar la exploración temática y la deduplicación, combinando posteriormente la verificación manual.

Se realizó una cobertura temporal con un primer tramo que abarca 2000-2022 y un segundo tramo que abarca de 2022-2025. Los idiomas de búsqueda principales han sido el español e inglés. Se han manejado herramientas de traducción para terceros idiomas en búsquedas accidentales.

En ATLAS.ti Paper Search se han utilizado las siguientes cadenas de búsqueda;

- ("multi-agent system" OR "agentic AI" AND ("attendance" OR "absenteeism") AND "school")
- ("early warning system" OR EWS) AND ("attendance" OR "absenteeism") AND ("machine learning")
- ("artificial intelligence" OR "machine learning") AND ("education" OR "school" OR "student") AND ("attendance" OR "engagement" OR "dropout" OR "absenteeism")

- ("multi-agent system" OR "agent-based model") AND ("education" OR "school" OR "student") AND ("attendance" OR "engagement" OR "dropout" OR "absenteeism")
- ("multi-agent system" AND "school attendance")
- ("multi-agent system" AND "dropout")

También se ha utilizado preguntas del tipo. “Studies using multi-agent or data-driven AI systems to prevent school absenteeism.”

La selección se hizo tras analizar el título y resumen de los artículos. No se elaboró diagrama PRISMA-lite. La evidencia empírica sobre IA agéntica específica para absentismo es limitada; se triangulan arquitecturas multiagente y casos de EWS como aproximaciones.

7. Observaciones finales

Se han incluido dos anexos en documento aparte:

- [Anexo I](#), con un glosario, ordenado, en el que se expone una sencilla ejemplificación para hacer más comprensible cada término.
- [Anexo II](#), con una lista de siglas o acrónimos.

Financiación

Sin financiación expresa. El costo de elaboración de este artículo ha sido sufragado por el autor.

Conflicto de intereses

Este autor es miembro coordinador del grupo para la elaboración del número monográfico en el que se haya alojado este artículo. Asimismo es miembro del comité científico de la Revista en la que se aloja este artículo. Por todo ello, el autor ha sido especialmente meticuloso en no interferir ni el proceso de edición ni en la elección de revisores externos posterior a la entrega de este artículo.

Declaración sobre el uso de herramientas de inteligencia artificial

Este trabajo ha incorporado el apoyo de herramientas de inteligencia artificial generativa como entre ellas ChatGPT (modelo GPT-5, OpenAI) y Gemini 3.0 Pro para tareas de estructuración conceptual, redacción inicial, reedición y revisión de coherencias. Todas las decisiones analíticas, interpretativas y conclusiones han sido elaboradas y validadas por el autor, garantizando el control humano del contenido y la responsabilidad plena sobre los resultados presentados.

Referencias

- 1EdTech Consortium. (2019). *Learning Tools Interoperability® (LTI®) 1.3 Core Specification*. 1EdTech Consortium. <https://www.imsglobal.org/spec/lti/v1p3>
- Abbas, N., Andersson, J., & Löwe, W. (2010, August). Autonomic software product lines (ASPL). In *Proceedings of the Fourth European Conference on Software Architecture: Companion Volume* (pp. 324-331).
- Advanced Distributed Learning Initiative. (2017). *Experience API (xAPI) Specification Version 1.0.3*. U.S. Department of Defense. <https://adlnet.gov/projects/xapi>
- Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley. ISBN 0-201-00174-8
- Badia, A. P., Piot, B., Kapturowski, S., Sprechmann, P., Vitvitskyi, A., Guo, Z. D., & Blundell, C. (2020). Agent57: Outperforming the Atari human benchmark. *Proceedings of Machine Learning Research*, 119, 507–517. <https://proceedings.mlr.press/v119/badia20a.html>
- Balfanz, R., & Byrnes, V. (2018). Using data and the human touch: Evaluating the NYC inter-agency campaign to reduce chronic absenteeism. *Journal of Education for Students Placed at Risk*, 23(1–2), 107–121. <https://doi.org/10.1080/10824669.2018.1435283>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Bowers, A. J., Sprott, R., y Taff, S. A. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity. *The High School Journal*, 96(2), 77–100. <https://doi.org/10.1353/hsj.2013.0000>
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (2000). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- DeLoach, S. A., & Valenzuela, J. P. (2021). Multi-agent systems engineering: A coordinated model-based approach. *Journal of Autonomous Agents and Multi-Agent Systems*, 35(2), 27. <https://doi.org/10.1007/s10458-021-09537-3>
- Deming, W. E. (1986). *Out of the crisis*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint* <https://arxiv.org/abs/1702.08608>

- Earl, L., & Timperley, H. (2015). *Evaluative thinking for successful educational innovation* (OECD Education Working Papers No. 122). OECD Publishing. <https://doi.org/10.1787/5jrxtk1jtdwf-en>
- Esteban Frades, S. (2025). Impactos de la inteligencia artificial en la educación. *Avances en Supervisión Educativa*, 43. <https://doi.org/10.23824/ase.v0i43.978>
- Europass. (2025, 11 de agosto). *Latest developments to the European Learning Model*. <https://europass.europa.eu/en/news/latest-developments-european-learning-model>
- European Commission. (2017). *New European Interoperability Framework (EIF): Promoting seamless services and data flows for European public administrations* [Brochure]. Publications Office of the European Union. <https://op.europa.eu/en/publication-detail/-/publication/bca40dde-deee-11e7-9749-01aa75ed71a1>
- European Commission, High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- European Data Protection Board. (2020). *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default (Version 2.0)*. Brussels: European Data Protection Board. <https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and-en>
- European Data Protection Board. (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. Brussels: European Data Protection Board. https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en_edpb.europa.eu
- European Parliament and the Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*. *Official Journal of the European Union*, L 119, 1–88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- European Parliament and the Council of the European Union. (2024, June 13). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, L 202, 1–172. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Faci, F. (2011). El abandono escolar prematuro en España. *Avances en Supervisión Educativa*, 14. <https://doi.org/10.23824/ase.v0i14.468>

- Ferraiolo, D. F., & Kuhn, D. R. (1992). *Role-based access controls*. In *Proceedings of the 15th National Computer Security Conference* (pp. 554–563). National Institute of Standards and Technology. <https://csrc.nist.gov/pubs/conference/1992/10/13/rolebased-access-controls/final>
- Floridi, L. (Ed.). (2021). *Ethics, governance, and policies in artificial intelligence (Philosophical Studies Series, 144)*. Springer. <https://doi.org/10.1007/978-3-030-81907-1>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., & Vayena, E. (2018). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 31(2), 161–188 <https://doi.org/10.1007/s11023-018-9482-5>
- Georgievski, I., y Aiello, M. (2016). An overview of Hierarchical Task Network planning. *ACM Computing Surveys*, 49(1), Article 37, 1–41. <https://doi.org/10.48550/arXiv.1403.7426>
- Ghallab, M., Nau, D., & Traverso, P. (2016). *Automated planning and acting*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139583923>
- Gobierno de Aragón y Ayuntamiento de Zaragoza. (2017). *Programa de actuación para la prevención del absentismo escolar*. Aprobado por la Comisión de Seguimiento del Convenio Marco de Colaboración entre el Gobierno de Aragón y el Ayuntamiento de Zaragoza el 4 de abril de 2017. <https://www.educaragon.org/>
- Graves, A., Wayne, G., & Danihelka, I. (2014). Neural Turing machines. *arXiv preprint arXiv:1410.5401*. <https://arxiv.org/abs/1410.5401>
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1–42. <https://dl.acm.org/doi/10.1145/3236009>
- Harris, A. (2014). *Distributed leadership matters: Perspectives, practicalities, and potential*. Corwin Press.
- Herrera, M., Pérez-Hernández, M., Kumar Parlikad, A., & Izquierdo, J. (2020). Multi-Agent Systems and Complex Networks: Review and Applications in Systems Engineering. *Processes*, 8(3), 312. <https://doi.org/10.3390/pr8030312>
- Holgado, A. (2018). *Integration analysis of solutions based on software as a service learning ecosystems* [Tesis doctoral]. Universidad de Salamanca. <https://repositorio.grial.eu/>

- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Boston, MA: Centre for Curriculum Redesign. <https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf>
- Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961-1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a complex system: Conceptual and methodological implications. *Educational Researcher*, 48(2), 112–119. <https://doi.org/10.3102/0013189X19826958>
- Jennings, N. R., & Bussmann, S. (2003). Agent-based control systems: Why are they suited to engineering complex systems? *IEEE Control Systems Magazine*, 23(3), 61–73. <https://doi.org/10.1109/MCS.2003.1200249>
- Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action*. Boston, MA: Harvard Business School Press.
- Kephart, J. O., & Chess, D. M. (2003). The vision of autonomic computing. *Computer*, 36(1), 41–50. <https://doi.org/10.1109/MC.2003.1160055>
- Li, Y., & Tan, C. (2019). A survey of the consensus for multi-agent systems. *Systems Science y Control Engineering*, 7(1), 468–482. <https://doi.org/10.1080/21642583.2019.1695689>
- Liang, P., Wu, C., Zhang, Y., Lee, J., & Zou, J. (2024). AutoGen: Enabling next-gen multi-agent systems with LLMs. *arXiv preprint arXiv:2308.08155*. <https://doi.org/10.48550/arXiv.2308.08155>
- Long, D., & Magerko, B. (2020). What Is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- McAfee, A., & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. W. W. Norton y Company.
- Michie, D. (1963). Experiments on the mechanization of game-learning Part I. Characterization of the model and its parameters. *Computer Journal*, 6(3), 232–236. <https://doi.org/10.1093/comjnl/6.3.232>
- Mittelstadt, B. D. (2019). *Principles alone cannot guarantee ethical AI*. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- OpenAI. (2023). *GPT-4 Technical Report*. *arXiv preprint arXiv:2303.08774*. <https://doi.org/10.48550/arXiv.2303.08774>

- Organisation for Economic Co-operation and Development (OECD). (2019). *Measuring innovation in education 2019: What has changed in the classroom?* Paris: OECD Publishing. <https://doi.org/10.1787/9789264311671-en>
- Organisation for Economic Co-operation and Development (OECD). (2023a). *The importance of knowledge bases for artificial intelligence*. Paris: OECD Publishing. <https://doi.org/10.1787/a8d820bd-en>
- Organisation for Economic Co-operation and Development (OECD). (2023b). *OECD Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- Organisation for Economic Co-operation and Development (OECD). (2023c), “Proposal for an action plan to reduce early school leaving in Spain”, No. 71, OECD Publishing, Paris, <https://doi.org/10.1787/0c249e7a-en>. (En español: <https://doi.org/10.1787/9bc3285d-es>.)
- Outay, F., Jabeur, N., Bellalouna, F., & Al Hamzi, T. (2024). *Multi-agent system-based framework for an intelligent management of competency building*. *Smart Learning Environments*, 11(1), Article 41. <https://doi.org/10.1186/s40561-024-00328-3>
- Pandey, B. K (2025, octubre 27). *How Agentic AI Systems Actually Work (A Complete Blueprint) [Infografía]*. LinkedIn. <https://www.linkedin.com/in/brijkishorepandey/>.
- Qin, J., Han, X., Liang, X., Chen, Y., & Yu, D. (2024). LangChain: Building applications with large language models. *arXiv preprint arXiv:2402.19085*. <https://doi.org/10.48550/arXiv.2402.19085>
- Rae, J. W., Borgeaud, S., Cai, T., Millican, K., Hoffmann, J., Song, F., Aslanides, J., Henderson, S., Ring, R., Young, S., Rutherford, E., Hennigan, T., Menick, J., Cassirer, A., Powell, R., van den Driessche, G., Hendricks, L. A., Rauh, M., Huang, P.-S., ... Irving, G. (2021). Scaling language models: Methods, analysis & insights from training Gopher. *arXiv*. <https://doi.org/10.48550/arXiv.2112.11446>
- Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, 20(1), 5-14. <https://doi.org/10.1007/s10676-017-9430-8>
- Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach (4th ed.)*. Pearson Education. ISBN 978-0134610993
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63. <https://doi.org/10.1145/3381831>
- Senge, P. M. (2006). *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency. ISBN-10: 0385517254

- Sevilla, A., Cuevas-Ruiz, P., Rello, L., y Sanz, I. (2025). La inteligencia artificial en la educación: oportunidades, retos y equidad en un nuevo paradigma de aprendizaje. *Papeles de Economía Española*, 184, 1–20. https://www.funcas.es/wp-content/uploads/2025/07/PEE_184_Sevilla_et-al.pdf
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Silver, D., Schrittwieser, J., Hubert, T., Antonoglou, I., Lai, M., Guez, A., ... Hassabis, D. (2021). Reward is enough. *Artificial Intelligence*, 299, 103535. <https://doi.org/10.1016/j.artint.2021.103535>
- Simon, H. A. (1997). *Modeling bounded rationality*. Cambridge, MA: MIT Press.
- Spillane, J. P. (2006). *Distributed leadership*. San Francisco, CA: Jossey-Bass.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction (2nd ed.)*. MIT Press. <https://mitpress.mit.edu/9780262039246/reinforcement-learning/>
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021, November). *Recommendation on the Ethics of Artificial Intelligence*. Paris: UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Vincent-Lancrin, S., & González-Sancho, C. (2023). *Interoperability: unifying and maximising data reuse within digital education ecosystems*. En *OECD Digital Education Outlook 2023: Towards an effective digital education ecosystem* (pp. 274–296). OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- Wooldridge, M. (2009). *An introduction to multiagent systems (2nd ed.)*. Wiley. ISBN 978-0470519462
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhu, C., Dastani, M., & Wang, S. (2024). A survey of multi-agent deep reinforcement learning with communication. *Autonomous Agents and Multi-Agent Systems*, 38(1), Article 4. <https://doi.org/10.1007/s10458-023-09633-6>

Anexo I - Glosario bilingüe (EN/ES) — IA en gestión educativa

English	Español	Definición (ES)
Agentic AI	IA agéntica	<p>Sistema de IA con capacidad operativa para observar, analizar, planificar y activar intervenciones (no solo predecir), normalmente en ciclos cortos y con supervisión humana.</p> <p>Ejemplo: Tras detectar un patrón de riesgo (análisis), el agente podría automáticamente preparar un borrador de correo para el tutor con los datos clave (plan) y ponerlo en su bandeja de salida para que este solo tenga que revisar y enviar (ejecución supervisada).</p>
Multi-agent system	Sistema multiagente	<p>Arquitectura con varios agentes de software que cooperan o se reparten funciones (detección, decisión, comunicación) para objetivos educativos.</p> <p>Ejemplo: Un sistema multiagente podría coordinar la recopilación automatizada de asistencia, la gestión de comunicaciones escolares y el registro de incidencias, repartiéndose funciones entre agentes especializados.</p>
HITL (Human-in-the-loop)	Supervisión humana en el bucle	<p>Diseño en el que personas validan o ajustan recomendaciones/acciones de la IA antes y después de su ejecución.</p> <p>Ejemplo: El tutor revisa las recomendaciones de la IA sobre intervenciones antes de aprobar su ejecución en el centro.</p>
MAPE-K cycle	Ciclo MAPE-K	<p>Bucle Monitor–Analyze–Plan–Execute con Knowledge (K): monitoriza datos, analiza, planifica, ejecuta y aprende para mejorar.</p> <p>Ejemplo: El sistema recolecta datos de asistencia, analiza patrones de absentismo, planifica intervenciones, ejecuta alertas y aprende de cada ciclo para ajustar las futuras decisiones.</p>
Memory / Memoria (STM-LTM)	Memoria a corto y largo plazo	<p>Sistema de almacenamiento que permite al agente recordar interacciones previas (STM) y aprender de experiencias pasadas (LTM).</p> <p>Ejemplo: El sistema recuerda ausencias reiteradas de un alumno y ajusta su recomendación de seguimiento.</p>
Knowledge base (K)	Base de conocimiento (K)	<p>Repositorio de políticas, umbrales, playbooks y lecciones aprendidas que realimenta el ciclo MAPE-K.</p> <p>Ejemplo: El sistema se nutre de una base donde se recopilan ejemplos de intervenciones exitosas y umbrales de alerta ajustados por expertos.</p>
Reasoning and Decision Making	Razonamiento y toma de decisiones	<p>Proceso por el que el agente transforma datos en acciones mediante reglas lógicas o modelos predictivos.</p> <p>Ejemplo: El agente analiza ausencias y propone avisar al tutor si detecta tres faltas seguidas.</p>
PDCA Cycle	Ciclo PDCA (Plan–Do–Check–Act)	<p>Metodología de mejora continua basada en cuatro etapas: planificar, ejecutar, verificar y actuar.</p>

English	Español	Definición (ES)
		Ejemplo: El agente ejecuta una intervención, evalúa su resultado y adapta la estrategia en el siguiente ciclo MAPE-K.
Reinforcement Learning	Aprendizaje por refuerzo	<p>Método de aprendizaje automático en el que los agentes mejoran su comportamiento mediante recompensas o penalizaciones derivadas de sus acciones.</p> <p>Ejemplo: El agente de aprendizaje evalúa las intervenciones aplicadas y ajusta sus recomendaciones según los resultados obtenidos.</p>
API (Application Programming Interface)	Interfaz de programación de aplicaciones	<p>Conjunto de reglas que permite que diferentes sistemas intercambien información de manera estructurada y segura.</p> <p>Ejemplo: El agente utiliza una API del sistema de gestión educativa para obtener los registros de asistencia en tiempo real.</p>
RBAC (Role-Based Access Control)	Control de acceso basado en roles	<p>Sistema de seguridad que otorga permisos según el rol del usuario (docente, orientador, inspector, etc.), evitando accesos no autorizados.</p> <p>Ejemplo: Solo los inspectores pueden revisar los datos agregados de distrito; los tutores acceden únicamente a los casos de su grupo.</p>
Federated Learning	Aprendizaje federado	<p>Estrategia de entrenamiento distribuido que permite mejorar los modelos de IA sin compartir datos originales.</p> <p>Ejemplo: Cada centro educativo entrena su modelo local y comparte solo los parámetros con la zona escolar.</p>
Guardrails	Guardarraíles	<p>Salvaguardas técnicas y procedimentales que limitan el comportamiento del sistema (validaciones, control de acceso RBAC, anonimización, umbrales, auditoría, revisión humana) para asegurar seguridad, equidad y cumplimiento (RGPD/AI Act).</p> <p>Ejemplo: Antes de emitir una alerta de absentismo crónico, el sistema valida datos, comprueba base jurídica, exige confianza $\geq 0,8$ con explicación XAI y eleva a tutoría (HITL) dejando rastro de auditoría.</p>
Orchestration	Orquestación	<p>Coordinación automatizada de agentes, servicios o módulos dentro de un sistema multiagente.</p> <p>Ejemplo: El orquestador central asigna tareas a los agentes analítico, familiar y de inspección, asegurando coherencia y sincronización.</p>
Conversational Interface	Interfaz conversacional	<p>Medio de comunicación basado en lenguaje natural entre el usuario y el sistema, facilitado por modelos de lenguaje (Los LLM).</p> <p>Ejemplo: Un docente pregunta al asistente virtual: “¿Qué alumnos presentan riesgo esta semana?” y recibe una respuesta explicada.</p>

English	Español	Definición (ES)
Explainability (XAI)	Explicabilidad	<p>Capacidad del sistema para explicar por qué genera una alerta o recomendación (variables clave, reglas aplicadas).</p> <p>Ejemplo: “Alerta generada porque el estudiante X acumula tres ausencias en cinco días y su participación en el LMS bajó un 40%.”</p>
Data Lake	Lago de datos	<p>Repositorio centralizado que almacena grandes volúmenes de datos en bruto para análisis o modelado.</p> <p>Ejemplo: El data lake distrital conserva los registros de asistencia e incidencias.</p>
Green AI	IA verde	<p>Enfoque de desarrollo de IA que prioriza eficiencia energética y sostenibilidad.</p> <p>Ejemplo: El monitor energético redistribuye tareas entre servidores en función del consumo.</p>
Data minimization	Minimización de datos	<p>Principio de usar solo los datos estrictamente necesarios para la finalidad educativa.</p> <p>Ejemplo: El sistema analiza solo asistencia y rendimiento, no datos personales adicionales.</p>
DPIA (Data Protection Impact Assessment)	Evaluación de Impacto en Protección de Datos	<p>Análisis formal del riesgo para derechos y libertades en despliegues de IA; define medidas técnicas y organizativas.</p> <p>Ejemplo: Antes de lanzar una nueva función, el centro realiza una DPIA para revisar riesgos y medidas.</p>
Interoperability	Interoperabilidad	<p>Capacidad de sistemas (asistencia, LMS, mensajería) para compartir datos con estándares comunes.</p> <p>Ejemplo: El sistema de control de asistencia envía datos al de alerta temprana y este comunica con mensajería familiar.</p>
Early Warning System (EWS)	Sistema de alerta temprana	<p>Conjunto de indicadores que anticipan riesgo (p. ej., absentismo) y generan señales para intervención.</p> <p>Ejemplo: El sistema detecta tres ausencias seguidas y genera automáticamente una alerta.</p>
Risk score	Puntuación de riesgo	<p>Valor numérico que resume la probabilidad de absentismo próximo.</p> <p>Ejemplo: El sistema calcula que un estudiante tiene un 80% de probabilidad de absentismo.</p>
Threshold	Umbral	<p>Valor a partir del cual una señal dispara una alerta o activa una intervención.</p> <p>Ejemplo: Cuando las faltas injustificadas superan el 5% del mes, el sistema activa revisión.</p>
Signal	Señal	<p>Evento o indicador que sugiere riesgo (p. ej., impuntualidades) y ayuda a priorizar casos.</p> <p>Ejemplo: Una racha de impuntualidades genera una señal de posible riesgo.</p>

English	Español	Definición (ES)
Playbook	Guía de intervención	<p>Catálogo de acciones graduadas con criterios de activación y seguimiento.</p> <p>Ejemplo: Nivel 1: SMS automático; Nivel 2: Llamada del tutor; Nivel 3: Reunión con la familia.</p>
Human Oversight	Supervisión humana	<p>Principio ético que garantiza intervención humana en decisiones automatizadas.</p> <p>Ejemplo: El inspector revisa las alertas de la IA antes de abrir un expediente.</p>
Governance	Gobernanza de datos/IA	<p>Conjunto de políticas y controles para asegurar uso ético y seguro de datos y algoritmos.</p> <p>Ejemplo: El centro crea un comité de revisión de algoritmos para asegurar cumplimiento normativo.</p>
Governance framework	Marco de gobernanza	<p>Estructura normativa que define roles, responsabilidades, métricas y auditoría para IA en centros educativos.</p> <p>Ejemplo: La política del centro exige auditorías periódicas y define quién puede modificar parámetros del sistema.</p>
Fairness / Bias	Equidad / Sesgo	<p>Grado en que el sistema funciona de forma comparable entre grupos y evita perjuicios sistemáticos.</p> <p>Ejemplo: Si el sistema genera más falsos positivos en un grupo concreto, muestra sesgo.</p>
Model drift	Deriva del modelo	<p>Pérdida de precisión al cambiar los patrones reales o el contexto educativo.</p> <p>Ejemplo: Un modelo entrenado antes de la pandemia deja de ser fiable tras ella.</p>
Precision / PPV	Precisión / Valor predictivo positivo	<p>Porcentaje de alertas correctas entre las emitidas.</p> <p>Ejemplo: De 10 alertas, 8 se confirmaron como casos reales (PPV=80%).</p>
Learning analytics	Analítica del aprendizaje	<p>Uso de datos educativos para mejorar enseñanza y resultados, con principios de transparencia y responsabilidad.</p> <p>Ejemplo: Los datos de participación y evaluación permiten planificar refuerzos personalizados.</p>
Decision support system	Sistema de apoyo a decisiones	<p>Módulo que recomienda acciones (no decide) con justificación y registro.</p> <p>Ejemplo: El sistema sugiere entrevistar al alumno por su patrón de ausencias; el tutor valida la acción.</p>
Administrative functions	Funciones administrativas (dirección)	<p>Ámbitos de gestión donde la IA puede ayudar: planificación, organización, comunicación y coordinación.</p> <p>Ejemplo: La dirección escolar utiliza IA para planificar jornadas y comunicar novedades.</p>

English	Español	Definición (ES)
Biometric recognition	Reconocimiento biométrico	<p>Tecnología de identificación (facial, huella, voz) con riesgos elevados de privacidad.</p> <p>Ejemplo: Se usa reconocimiento facial para controlar accesos, con alternativas para quien lo rechace.</p>
Opt-out	Derecho de exclusión	<p>Posibilidad de que familias o alumnos se excluyan de ciertos módulos de análisis automatizado.</p> <p>Ejemplo: Las familias pueden solicitar no incluir a sus hijos en el módulo de IA predictiva.</p>
SLA (Service-level agreement)	Acuerdo de nivel de servicio	<p>Compromiso de tiempos máximos para revisión o intervención tras una alerta.</p> <p>Ejemplo: El proveedor garantiza que toda alerta será revisada en menos de 48 horas lectivas.</p>
Counterfactual (administrative)	Contrafactual administrativo	<p>Estimación de qué habría ocurrido sin intervención, a partir de registros históricos.</p> <p>Ejemplo: Se compara la evolución del absentismo en aulas con IA frente a las que siguieron procesos tradicionales.</p>
Stepped-wedge design	Diseño por escalado escalonado	<p>Estrategia de evaluación en la que los centros se incorporan por fases, comparando resultados con periodos previos.</p> <p>Ejemplo: Los colegios adoptan el sistema por etapas para analizar impacto progresivo.</p>

Autoría. Elaboración asistida por IA (ChatGPT-5 Thinking, octubre-2025); revisión y edición del autor.

Anexo II – Lista de Siglas o acrónimos utilizados en el artículo ¹⁰

Sigla (EN)	Sigla (ES)	Aclaración en inglés	Aclaración en español
AI Act	—	Artificial Intelligence Act (Reg. (EU) 2024/1689)	Reglamento de Inteligencia Artificial
API	—	Application Programming Interface	Interfaz de programación de aplicaciones
CEN TC 353	—	CEN Technical Committee 353	Comité Técnico 353 del CEN
DPIA	EIPD	Data Protection Impact Assessment	Evaluación de Impacto en Protección de Datos
ELM	—	European Learning Model	Modelo Europeo de Aprendizaje
ESL	AEP/AET	Early School Leaving	Abandono escolar prematuro. En España se utiliza también Abandono escolar temprano. El término escolar también puede ser suplido por educativo
EWS	SAT	Early Warning System	Sistema de alerta temprana
GDPR	RGPD	General Data Protection Regulation	Reglamento General de Protección de Datos
HITL	—	Human-in-the-Loop	Supervisión humana en el bucle Mecanismo por el que una persona interviene en puntos de control del ciclo del sistema (antes, durante o después de la decisión) para aprobar, corregir, bloquear o escalar resultados.
KPI	—	Key Performance Indicator(s)	Indicador(es) clave de desempeño Métrica cuantificable que mide el grado de avance hacia un objetivo operativo o estratégico.
LMS	SGA	Learning Management System	Sistema de Gestión del Aprendizaje Plataforma docente. Gestiona cursos, recursos o actividades y más

¹⁰ Nota: alguna sigla ya aparece en el glosario pero por mantener la unidad formal de la lista se incluye también en este listado

Sigla (EN)	Sigla (ES)	Aclaración en inglés	Aclaración en español
LTI	—	Learning Tools Interoperability	Interoperabilidad de herramientas de aprendizaje Es la forma estándar de conectar apps educativas con el aula virtual (LMS) para que funcionen como si fueran una sola: entras una vez, se comparten usuarios/roles, y pueden enviar/recibir resultados y registro de actividades sin copiar datos a mano.
MAPE-K	—	Monitor, Analyze, Plan, Execute, Knowledge	Monitorizar, Analizar, Planificar, Ejecutar, Conocimiento
MIS (educ.)	SGE	(Educational) Management Information System	Sistema de Gestión Educativa Plataforma administrativa y académico-institucional. Gestiona matrícula, expediente oficial, grupos y horarios, asistencia/partes, evaluaciones oficiales/boletines/actas y más.
OECD	OCDE	Organisation for Economic Co-operation and Development	Organización para la Cooperación y el Desarrollo Económicos
—	PAE	(—)	Programa de Actuación para la Prevención del Absentismo (Zaragoza)
PDCA	PDCA	Plan–Do–Check–Act Cycle	Ciclo Planificar–Hacer–Verificar–Actuar
RBAC	—	Role-Based Access Control	Control de acceso basado en roles Ejemplos de roles: Tutor, Jefatura de Estudios, Inspección
SLA	ANS	Service Level Agreement	Acuerdo de nivel de servicio Compromiso formal y medible entre proveedor y cliente que fija niveles mínimos de servicio
SSO	—	Single Sign-On	Autenticación única
STM / LTM	—	Short-Term / Long-Term Memory	Memoria a corto / largo plazo
XAI	—	Explainable Artificial Intelligence	Inteligencia artificial explicable
xAPI	—	Experience API	API de experiencias

Sigla (EN)	Sigla (ES)	Aclaración en inglés	Aclaración en español
UNESCO	UNESCO	United Nations Educational, Scientific and Cultural Organization	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
SMS	SMS	Short Message Service	Servicio de mensajes cortos

Autoría. Elaboración asistida por IA (ChatGPT-5 Thinking, octubre-2025); revisión y edición del autor.

En español se incluyen únicamente las siglas cuya forma en castellano está asentada en el uso. En algunos casos, la traducción al español ha impuesto su propia sigla (p. ej., RGPD, OCDE, EIPD, SGE); cuando no existe equivalente consolidado, se mantiene la sigla inglesa (p. ej., API, KPI, PDCA).