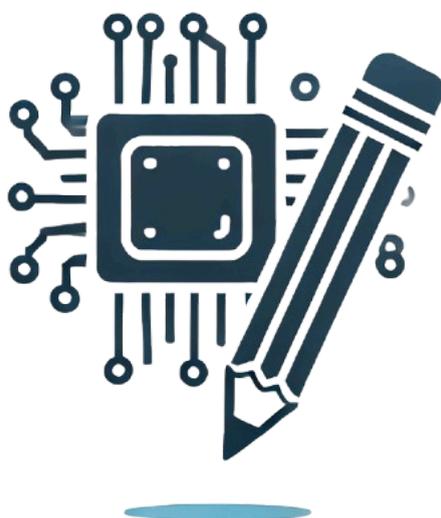


Decálogo técnico y ético para el uso responsable de la IA en Educación

De la precisión algorítmica a la responsabilidad educativa: principios para una implementación regulada y equitativa de la IA en entornos educativos



JOSÉ LUIS DE LA TORRE LORENTE

delatorre.ai

Formador de Inteligencia Artificial Aplicada en delatorre.ai. Responsable tecnológico de Educahistoria.com. Responsable externo de IA aplicada en el Institut Escola Municipal de Treball de Granollers.

JOSÉ LUIS DE LA TORRE DÍAZ

educahistoria.com

Profesor de Ciencias Sociales, Instituto Escola Municipal del Treball de Granollers. Director de Educahistoria.com. Especialista en IA educativa.





Índice

Introducción.....	3
Desafíos técnicos y éticos de la IA en educación.....	5
Limitaciones técnicas y éticas de la IA.....	6
Decálogo técnico para el uso responsable de IA en educación.....	10
Decálogo Ético para el Uso Responsable de IA en Educación.....	14
Relación entre técnica y ética en la IA.....	19
Preguntas Frecuentes (FAQ).....	24



Introducción

Propósito del documento

La inteligencia artificial (IA) ha emergido como una herramienta disruptiva en el ámbito educativo, facilitando desde la personalización del aprendizaje hasta la optimización de la gestión administrativa en entornos académicos. Según la UNESCO (2021), en su informe *Artificial Intelligence and Education: Guidance for Policy-makers*, la IA tiene el potencial de mejorar la accesibilidad y eficiencia del aprendizaje, beneficiando especialmente a estudiantes con necesidades educativas diversas. No obstante, su implementación plantea desafíos tanto técnicos como éticos, requiriendo un análisis crítico fundamentado en principios de equidad, transparencia y supervisión humana.

El presente documento tiene como objetivo establecer un marco analítico que permita comprender la integración responsable de la IA en la educación, delimitando sus restricciones técnicas y los dilemas éticos inherentes. A tal efecto, se abordarán dos dimensiones clave: las limitaciones estructurales de los sistemas de IA y sus implicaciones en la justicia educativa y la toma de decisiones pedagógicas.

Relevancia de la IA en el ámbito educativo

La IA ofrece nuevas perspectivas para la innovación pedagógica al posibilitar la personalización adaptativa de los procesos de enseñanza-aprendizaje. Su capacidad de análisis de grandes volúmenes de datos en tiempo real permite identificar patrones en el rendimiento académico y ajustar las metodologías de enseñanza de manera dinámica, promoviendo enfoques educativos más inclusivos y eficientes.

Sin embargo, su implementación no está exenta de riesgos. Aunque la IA facilita la diferenciación instruccional y la optimización de recursos docentes, también introduce problemáticas como la opacidad algorítmica y la automatización de decisiones pedagógicas sin supervisión humana. La falta de transparencia en los modelos de IA, sumada a los sesgos presentes en los conjuntos de datos utilizados para su entrenamiento, puede comprometer la equidad educativa. Además, la dependencia excesiva de estos sistemas plantea un posible desplazamiento del rol docente como mediador crítico del aprendizaje y, en consecuencia, la reducción de la autonomía del estudiante en su propio proceso educativo.



En este sentido, la integración de la IA en la educación requiere de un enfoque reflexivo por parte de docentes, responsables de política educativa y diseñadores de tecnología, quienes deben asegurar que su implementación se realice de manera equitativa, ética y alineada con los principios fundamentales de la educación inclusiva.

Desafíos técnicos y éticos de la IA en educación

Desde una perspectiva técnica, los sistemas de IA se desarrollan a partir de conjuntos de datos que pueden no representar adecuadamente la diversidad sociocultural de los entornos educativos. Un caso paradigmático es el del sistema de IA implementado por Ofqual en el Reino Unido en 2020 para la evaluación de exámenes A-level, el cual asignó calificaciones más bajas a estudiantes provenientes de entornos socioeconómicamente desfavorecidos, lo que resultó en protestas generalizadas y en la eventual suspensión del sistema por parte del gobierno británico (BBC News, 2020). Este incidente ilustra cómo la falta de transparencia en la modelización algorítmica puede generar desigualdades educativas y afectar negativamente a grupos vulnerables.

En el ámbito ético, el uso de IA en educación plantea interrogantes sobre la privacidad de los datos estudiantiles, la perpetuación de sesgos sistémicos y la necesidad de garantizar la supervisión humana en los procesos automatizados de toma de decisiones. Para mitigar estos riesgos, es imperativo que su implementación se rija por principios de equidad, auditabilidad y rendición de cuentas, asegurando que la tecnología no refuerce brechas educativas preexistentes ni limite el acceso equitativo a oportunidades de aprendizaje.

AI-TPACK como referencia teórica

Para analizar de manera estructurada la relación entre la IA y la educación, resulta pertinente adoptar el marco AI-TPACK (*Artificial Intelligence Technological, Pedagogical, and Content Knowledge*), una evolución del modelo TPACK que incorpora la inteligencia artificial como un elemento transversal en la intersección entre tecnología, pedagogía y contenido disciplinar.

AI-TPACK permite conceptualizar cómo los docentes pueden integrar la IA en sus estrategias pedagógicas sin comprometer la equidad y la calidad educativa. A diferencia del modelo TPACK tradicional, que se centra en la interrelación entre el conocimiento tecnológico, pedagógico y del contenido, AI-TPACK introduce dimensiones adicionales como el conocimiento tecnológico de la IA (AITK), el conocimiento pedagógico de la IA (AIPK) y el conocimiento ético y regulatorio de la IA (AIEK). Estas nuevas dimensiones



resaltan la necesidad de que los docentes no solo adquieran competencia en el uso de herramientas digitales, sino que también comprendan los principios subyacentes del aprendizaje automatizado, la supervisión algorítmica y los desafíos éticos que emergen de la personalización educativa mediante IA.

Limitaciones técnicas y éticas de la IA

La inteligencia artificial (IA) ha revolucionado el panorama educativo, facilitando la personalización del aprendizaje, la optimización de procesos de evaluación y la automatización de tareas administrativas. No obstante, su implementación presenta desafíos sustanciales tanto en el ámbito técnico como en el ético, lo que demanda un análisis exhaustivo y un marco regulador robusto para garantizar su utilización de manera equitativa y transparente.

Limitaciones técnicas

1. Calidad y representatividad de los datos

Los modelos de IA, en particular los modelos de lenguaje de gran escala en la nube (LLMs en cloud), como ChatGPT, DeepSeek, Claude o LeChat, dependen de vastos conjuntos de datos para su entrenamiento y operación. Sin embargo, la calidad y representatividad de estos datos determinan su fiabilidad y precisión. Los sesgos inherentes o la falta de diversidad en los datos de entrenamiento pueden comprometer la neutralidad del sistema, generando respuestas parciales o desactualizadas, lo que tiene implicaciones significativas en contextos educativos.

Desde una perspectiva académica, la UNESCO ha subrayado la escasez de estudios sistemáticos que evalúen el impacto de la IA en la equidad educativa. Los modelos desarrollados predominantemente en entornos occidentales pueden carecer de adaptación a diversas realidades culturales y lingüísticas, exacerbando brechas existentes en el acceso a la educación. Por ello, resulta imperativo el establecimiento de marcos normativos que aseguren la diversidad en los conjuntos de datos utilizados en estos modelos. la escasez de estudios sistemáticos que evalúen el impacto de la IA en la equidad educativa. Los modelos desarrollados predominantemente en entornos occidentales pueden carecer de adaptación a diversas realidades culturales y lingüísticas, exacerbando brechas existentes en el acceso



a la educación. Por ello, resulta imperativo el establecimiento de marcos normativos que aseguren la diversidad en los conjuntos de datos utilizados en estos modelos.

2. Transparencia y explicabilidad

Uno de los mayores desafíos en la aplicación de la IA en educación es la falta de transparencia de los modelos algorítmicos. Muchos de estos funcionan como "cajas negras", impidiendo la comprensión de su lógica de decisión por parte de docentes, estudiantes y reguladores. Esta opacidad dificulta la detección y corrección de errores, así como la identificación de sesgos que puedan afectar la personalización del aprendizaje y los procesos de evaluación.

En el caso de los modelos de lenguaje de gran escala en la nube, la problemática es aún más acentuada debido a su naturaleza probabilística. Estos modelos generan texto basándose en correlaciones estadísticas, lo que significa que pueden producir respuestas sintácticamente correctas pero conceptualmente erróneas. La falta de trazabilidad en la generación de contenidos dificulta su validación, lo que representa un desafío crítico en ámbitos donde la precisión de la información es fundamental. La Comisión Europea ha enfatizado la urgencia de establecer auditorías y mecanismos de supervisión para garantizar que los sistemas de IA sean auditables y comprensibles. Sin un acceso adecuado a los datos de entrenamiento y la lógica algorítmica de los modelos, los educadores y reguladores carecen de herramientas para corregir errores o mitigar sesgos en la toma de decisiones automatizadas. y mecanismos de supervisión para permitir un control efectivo de estos algoritmos y garantizar su uso responsable.

3. Fiabilidad y robustez

La fiabilidad de los sistemas de IA es otro factor crucial, especialmente en entornos educativos donde la precisión y coherencia de las respuestas son esenciales. Los modelos de lenguaje en la nube pueden presentar vulnerabilidades cuando se enfrentan a datos ambiguos, erróneos o que no forman parte de su espectro de entrenamiento, lo que puede derivar en respuestas inconsistentes y poco confiables.

Además, la falta de control sobre la infraestructura de estos modelos supone un obstáculo adicional. Mientras que en sistemas locales es posible gestionar directamente los datos y optimizar los algoritmos según necesidades específicas, en los modelos en la nube los usuarios dependen completamente de las plataformas que los gestionan, sin acceso a sus



procesos internos. Esta falta de transparencia genera incertidumbre sobre la fiabilidad de los resultados y limita la capacidad de implementar medidas correctivas. La Comisión Europea ha destacado la importancia de una supervisión humana continua para mitigar estos riesgos y garantizar una aplicación justa de la IA en el sector educativo. Asimismo, se ha señalado que la opacidad de los modelos en la nube impide una intervención efectiva en caso de fallos o inconsistencias en los sistemas de evaluación educativa. continua para mitigar estos riesgos y garantizar una aplicación justa de la IA en el sector educativo.

Limitaciones éticas

1. Privacidad y protección de datos

El uso de IA en educación requiere la recopilación y análisis de grandes volúmenes de datos personales, lo que plantea preocupaciones sustanciales en materia de privacidad y seguridad de la información estudiantil. Normativas como el Reglamento General de Protección de Datos (GDPR) en Europa y la Ley de Privacidad de la Información de los Niños en Línea (COPPA) en Estados Unidos proporcionan directrices para un manejo ético de los datos educativos. Sin embargo, la aplicación efectiva de estas normativas enfrenta desafíos debido a la infraestructura altamente centralizada de los modelos de IA en la nube, cuya administración recae en un número reducido de actores tecnológicos. Este modelo de gestión plantea interrogantes sobre la transparencia y la rendición de cuentas en el manejo de datos educativos a gran escala.

La falta de control sobre la recopilación, almacenamiento y procesamiento de datos por parte de los usuarios genera una dependencia estructural de proveedores externos. Esta opacidad dificulta la auditoría de los sistemas y la verificación de prácticas de minimización de riesgos, exponiendo la información estudiantil a posibles vulneraciones o usos no autorizados. La ausencia de mecanismos públicos efectivos para la supervisión de estos procesos pone en cuestión la seguridad y transparencia en la gestión de datos educativos.

2. Sesgos algorítmicos y equidad

Los modelos de IA pueden reforzar y amplificar sesgos preexistentes en los datos de entrenamiento, lo que puede generar discriminación en la evaluación académica y en el acceso a recursos educativos. Investigaciones han demostrado que estos sistemas pueden asignar calificaciones desproporcionadamente más bajas a ciertos grupos debido a la falta de representación equitativa en los conjuntos de datos. La UNESCO recomienda la



implementación de auditorías constantes para evitar la reproducción de estos sesgos y garantizar un uso justo de la IA en educación. Además, la falta de acceso público a los datos de entrenamiento de los modelos más utilizados impide evaluar el grado de sesgo desde su origen, generando desigualdades estructurales difíciles de corregir. para evitar la reproducción de estos sesgos y garantizar un uso justo de la IA en educación.

Uno de los problemas más preocupantes es la imposibilidad de auditar los datos de entrenamiento de los modelos más utilizados en la actualidad. Debido a que las empresas propietarias de estas plataformas no hacen públicos los conjuntos de datos con los que entrenan sus modelos, resulta imposible evaluar si existen sesgos estructurales desde su origen. Esta opacidad contribuye a la reproducción de desigualdades preexistentes sin posibilidad de corrección. La lógica algorítmica basada en correlaciones estadísticas, sin una contextualización real de los datos, puede generar clasificaciones erróneas y estereotipos automatizados sin una supervisión humana efectiva.

3. Desplazamiento del rol docente

Si bien la IA ofrece herramientas valiosas para la enseñanza, su uso excesivo puede comprometer el papel del docente y afectar la interacción pedagógica. Diversos estudios han señalado que la automatización de procesos educativos puede limitar la capacidad de los estudiantes para desarrollar pensamiento crítico y restringir la flexibilidad metodológica de los docentes. La UNESCO enfatiza que la IA debe ser un complemento y no un sustituto de la labor docente.

Desde un punto de vista ético, la delegación de responsabilidades docentes a sistemas de IA compromete la autonomía de los profesores y diluye su papel como mediadores del conocimiento. Diversos estudios sobre la aplicación de la IA en educación destacan la importancia de mantener la supervisión pedagógica como un elemento central en la enseñanza. Para garantizar una educación equitativa y efectiva, los sistemas automatizados no deben reemplazar por completo el juicio docente en la evaluación o el diseño curricular, sino operar como herramientas de apoyo bajo un marco de control y ajuste humano continuo. La excesiva dependencia de estos modelos podría dar lugar a un sistema educativo más mecanizado y menos reflexivo, donde las decisiones didácticas se rijan por patrones predictivos en lugar del juicio crítico basado en la experiencia del docente.

Conclusión



Las limitaciones técnicas y éticas de la IA en educación ponen de manifiesto la necesidad de regulaciones estrictas y una supervisión humana constante. Organismos como la UNESCO han subrayado la importancia de establecer principios de transparencia, equidad y auditabilidad para mitigar los riesgos y garantizar la equidad en su aplicación. La Declaración Europea sobre los Derechos Digitales enfatiza la necesidad de evitar la vigilancia masiva y proteger la privacidad estudiantil en entornos digitales educativos., equidad y auditabilidad para mitigar los riesgos y garantizar la equidad en su aplicación. La Declaración Europea sobre los Derechos Digitales resalta la necesidad de evitar la vigilancia masiva y proteger la privacidad estudiantil. En este contexto, la auditoría algorítmica, la capacitación docente en IA y la implementación de marcos regulatorios sólidos son factores esenciales para una integración responsable de la IA en el sector educativo.

Decálogo técnico para el uso responsable de IA en educación

La integración de la inteligencia artificial (IA) en la educación supone un avance significativo, pero también conlleva retos técnicos que deben ser comprendidos y gestionados de manera adecuada. Este decálogo establece principios clave para su aplicación en entornos educativos, con énfasis en la transparencia, la seguridad y la equidad.

1. Calidad y representatividad de los datos

Los modelos de IA dependen de los datos con los que han sido entrenados. Es crucial que estos datos sean diversos, representativos y actualizados para evitar sesgos y garantizar su aplicabilidad en distintos contextos educativos. Los docentes deben asegurarse de que las herramientas de IA que utilizan han sido entrenadas con datos culturalmente inclusivos y metodológicamente rigurosos.

Los modelos de lenguaje en la nube (LLMs) suelen utilizar grandes volúmenes de datos extraídos de repositorios académicos, bases de datos abiertas y contenido web. Sin embargo, la calidad y verificabilidad de estos datos varía significativamente. Un ejemplo de dataset abierto utilizado en la construcción de modelos de IA es The Pile, que integra textos científicos, código de programación y literatura. No obstante, en modelos cerrados como ChatGPT o Claude, los datos de entrenamiento no son accesibles, lo que impide evaluar su representatividad y sesgos.



Algunas plataformas recurren a datasets sintéticos para mejorar el rendimiento de sus modelos, lo que plantea interrogantes sobre la fiabilidad y la relevancia de los datos generados artificialmente. Esta estrategia puede generar limitaciones en la aplicabilidad del modelo a contextos educativos específicos.

2. Transparencia y explicabilidad

Para que la IA pueda integrarse de manera efectiva en la educación, es fundamental que los docentes comprendan cómo funcionan los modelos que utilizan. La transparencia debe darse en distintos niveles: los algoritmos del modelo, que determinan la generación de respuestas; los datos de entrenamiento, que influyen en la calidad del conocimiento adquirido por la IA; la infraestructura del sistema, que afecta el procesamiento y almacenamiento de la información; y los límites del modelo, que establecen las restricciones en la cantidad y complejidad de datos que puede manejar.

En muchos casos, los modelos comerciales de IA aplican filtros de seguridad dinámicos que modifican respuestas sin notificar al usuario, lo que impacta en su confiabilidad. Asimismo, aunque algunos sistemas permiten configuraciones personalizadas, esto no equivale a un verdadero ajuste fino (fine-tuning), que solo está disponible en entornos avanzados o en modelos de código abierto.

3. Fiabilidad y robustez

Los sistemas de IA pueden producir respuestas inconsistentes o incorrectas. Antes de emplearlos en la evaluación del alumnado o en la planificación de clases, es esencial verificar su fiabilidad contrastando sus resultados con fuentes verificadas. Complementar su uso con metodologías pedagógicas tradicionales ayuda a reducir la dependencia exclusiva de la tecnología.

Un problema recurrente en los modelos de IA es el fenómeno de la alucinación, en el cual los sistemas generan información plausible pero incorrecta. Esto ocurre porque los modelos de lenguaje no verifican la factualidad de sus respuestas, sino que generan contenido basado en probabilidades lingüísticas. Esto es particularmente crítico en temas recientes o especializados, donde la verificación debe ser una práctica constante.

4. Supervisión humana constante



La IA debe ser una herramienta complementaria y no un sustituto del docente. Es imprescindible que los educadores supervisen activamente las decisiones generadas por estas herramientas y mantengan el control sobre los procesos educativos. Si se detectan errores o sesgos, es necesario intervenir y corregir.

Desde un punto de vista técnico, la supervisión humana también implica la incorporación de mecanismos de validación y ajuste de respuestas en las plataformas de IA. Algunas herramientas permiten el "ajuste fino" (fine-tuning), donde los docentes pueden entrenar el modelo con ejemplos específicos para mejorar su precisión. Además, la verificación debe incluir estrategias como la triangulación de información y el análisis de patrones de respuesta.

5. Seguridad y privacidad de los datos

El uso de IA en educación implica la recopilación y almacenamiento de datos sensibles. Aunque existen regulaciones como el GDPR y la COPPA, su implementación efectiva sigue siendo un reto, especialmente en plataformas comerciales de IA. Se recomienda optar por herramientas que permitan la anonimización de datos y especifiquen de manera clara cómo gestionan la privacidad de los estudiantes.

6. Minimización del sesgo algorítmico

Los sistemas de IA pueden amplificar sesgos preexistentes en los datos de entrenamiento. Estos sesgos pueden originarse en diferentes niveles: datos de entrenamiento no representativos, priorización de patrones en el modelo y sesgo de interacción del usuario. Un modelo entrenado mayoritariamente en inglés, por ejemplo, podría no interpretar correctamente el español en su totalidad.

Para mitigar estos riesgos, se recomienda evaluar las respuestas de diferentes herramientas, identificar patrones discriminatorios y diversificar las fuentes de información. Además, el sesgo de automatización puede llevar a que los usuarios confíen demasiado en la IA sin cuestionar sus respuestas, lo que destaca la necesidad del pensamiento crítico en su uso educativo.

7. Evaluación continua del impacto



El uso de IA en educación debe ser monitoreado constantemente para evaluar su efectividad. Se recomienda establecer indicadores de desempeño, como la equidad en los resultados y la satisfacción del usuario. Los docentes deben utilizar plataformas que permitan ajustes en tiempo real.

Dado que la mayoría de los modelos comerciales no ofrecen auditoría completa, se recomienda implementar estrategias como comparación de modelos, registro de interacciones y validación cruzada.

8. Interoperabilidad y estandarización

La interoperabilidad es fundamental en un entorno donde los modelos en la nube son cerrados y centralizados. Muchas plataformas no son transparentes en sus límites de procesamiento, lo que puede afectar su aplicación en la educación.

9. Reducción de la dependencia tecnológica

El uso de modelos de código abierto, como Llama 2, Mistral y Falcon, permite a las instituciones mantener control sobre la infraestructura tecnológica y los datos utilizados. Sin embargo, su implementación requiere recursos adicionales.

10. Desarrollo de competencias digitales en docentes y estudiantes

El modelo AI-TPACK enfatiza la necesidad de que los docentes comprendan el conocimiento tecnológico de IA (AITK), el conocimiento pedagógico de IA (AIPK) y el conocimiento ético y regulatorio (AIEK) para garantizar su aplicación crítica y responsable en el aula.



Decálogo Ético para el Uso Responsable de IA en Educación

La incorporación de la inteligencia artificial (IA) en el ámbito educativo introduce desafíos éticos fundamentales que deben ser abordados con rigurosidad para garantizar su implementación equitativa, transparente y centrada en la protección de los derechos de docentes y estudiantes. Este decálogo establece principios clave para orientar la regulación, supervisión y evaluación continua del impacto de la IA en la educación.

1. Privacidad y protección de datos

El uso de IA en educación implica el tratamiento de datos sensibles de estudiantes y docentes, lo que exige el cumplimiento estricto de normativas como el Reglamento General de Protección de Datos (RGPD) en Europa y la Ley de Privacidad Infantil en Internet (COPPA) en EE.UU. Estas normativas establecen principios de transparencia, minimización de datos y derechos de acceso y eliminación de información personal, garantizando que los sistemas de IA no recopilen ni procesen más datos de los necesarios y que los usuarios puedan gestionar su privacidad de forma efectiva en entornos educativos. Además, el Reglamento de Inteligencia Artificial de la UE establece que los sistemas de IA clasificados como de alto riesgo, incluidos los utilizados en educación, deben contar con documentación técnica detallada y registros auditables para evaluar posibles riesgos.

Los modelos comerciales de IA en la nube almacenan interacciones sin un control explícito por parte del usuario, lo que plantea riesgos en cuanto a privacidad. Se debe garantizar que cualquier dato recopilado pueda ser eliminado bajo solicitud y que los sistemas educativos implementen plataformas que permitan auditar su uso y acceso.

2. Transparencia y explicabilidad de los procesos algorítmicos

Los sistemas de IA deben ser auditables y comprensibles, asegurando que los docentes y estudiantes puedan interpretar sus decisiones automatizadas. La transparencia debe garantizarse en diferentes niveles: estructura de los algoritmos, datos de entrenamiento y su procedencia, infraestructura de procesamiento, y límites operativos de los modelos.

El Reglamento de Inteligencia Artificial exige que los modelos de IA de propósito general publiquen un resumen de los contenidos utilizados en su entrenamiento y que los modelos



de riesgo sistémico se sometan a evaluaciones de seguridad, pruebas adversariales y medidas de ciberseguridad. En el ámbito educativo, estas evaluaciones pueden incluir pruebas de robustez para evitar sesgos en las recomendaciones de aprendizaje automatizado, auditorías sobre el impacto de los modelos en la equidad educativa y simulaciones de ataques adversariales para garantizar la integridad del sistema. Estas medidas aseguran que la IA aplicada en educación sea confiable, segura y resistente a manipulaciones o vulnerabilidades. Para garantizar un uso educativo confiable, es necesario que las herramientas de IA permitan auditoría independiente y ofrezcan documentación clara sobre la generación de respuestas.

3. Equidad y no discriminación

El diseño y la aplicación de la IA deben garantizar la equidad y evitar la reproducción de sesgos sistémicos que afecten a ciertos grupos de estudiantes. Los sistemas deben desarrollar mecanismos de detección y mitigación de sesgos algorítmicos para que la IA no refuerce desigualdades preexistentes en la educación. Por ejemplo, se ha documentado que algunos algoritmos de IA utilizados en procesos de selección académica han penalizado a ciertos grupos debido a patrones sesgados en los datos de entrenamiento, como la subrepresentación de estudiantes de comunidades marginadas. Para mitigar este riesgo, es fundamental implementar auditorías periódicas que evalúen la equidad de los modelos y aplicar técnicas de reentrenamiento con datos más diversos y representativos.

El Reglamento de Inteligencia Artificial clasifica como de alto riesgo los sistemas de IA empleados en la evaluación de estudiantes o en su admisión a instituciones educativas. Dichos modelos deben garantizar un nivel adecuado de precisión y no pueden reemplazar evaluaciones humanas sin supervisión. Las plataformas de IA deben proporcionar mecanismos de control y monitoreo de sesgos en tiempo real.

4. Supervisión humana y responsabilidad

El uso de IA en educación debe estar siempre bajo la supervisión de docentes y especialistas en pedagogía. No debe reemplazar la toma de decisiones humanas en procesos críticos como la evaluación del desempeño estudiantil o el diseño curricular.

El Reglamento de Inteligencia Artificial establece que los sistemas de IA en educación deben permitir intervención y supervisión humana, asegurando que ningún modelo pueda operar sin posibilidad de ajuste por parte de los docentes. La supervisión más efectiva en



este contexto implica la validación manual de decisiones automatizadas, la implementación de mecanismos de corrección en tiempo real y la capacitación docente en la identificación de sesgos y errores en la IA. Además, se recomienda el uso de plataformas que proporcionen herramientas de monitoreo y ajustes configurables que permitan adaptar la IA a las necesidades específicas del entorno educativo. Además, cualquier sistema de IA que genere impactos significativos en la educación debe contar con mecanismos de control y registro de incidencias para garantizar la trazabilidad de sus decisiones.

5. Uso responsable y contextualizado de la IA

La IA debe ser una herramienta de apoyo y no un sustituto del trabajo docente. Su aplicación debe evaluarse en función del contexto educativo, priorizando la autonomía de los estudiantes y evitando una dependencia tecnológica desmedida.

El Reglamento de Inteligencia Artificial prohíbe el uso de IA que manipule el comportamiento de los usuarios o explote vulnerabilidades relacionadas con la edad o condiciones socioeconómicas. En educación, esto implica que las plataformas de IA no pueden emplear técnicas persuasivas que afecten la autonomía del estudiante ni refuercen decisiones sin supervisión docente.

6. Accesibilidad e interoperabilidad

El acceso desigual a tecnología puede acentuar brechas educativas. Se deben desarrollar estrategias que aseguren la interoperabilidad de la IA con diversas plataformas, evitando dependencias innecesarias con proveedores específicos y garantizando su adaptabilidad a distintos entornos educativos.

Algunas plataformas comerciales de IA limitan funcionalidades según la región del usuario, lo que genera desigualdades en el acceso a herramientas avanzadas. Es necesario fomentar el uso de alternativas de código abierto y modelos locales que garanticen igualdad de oportunidades.

7. Desarrollo de competencias digitales y pensamiento crítico

Es imprescindible formar a docentes y estudiantes en el uso informado y crítico de la IA. Para ello, se pueden emplear estrategias como la enseñanza de alfabetización digital mediante programas de formación específicos, el desarrollo de competencias en ética de la



IA, la integración de estudios de caso sobre el impacto de la IA en la educación y la implementación de metodologías activas como el aprendizaje basado en problemas, donde los estudiantes analicen situaciones en las que la IA pueda presentar sesgos o errores. La alfabetización digital debe incluir el desarrollo de habilidades para interpretar, evaluar y cuestionar los resultados generados por los sistemas automatizados.

8. Uso ético de los datos y la IA generativa

Las herramientas de IA generativa deben emplearse con responsabilidad para evitar la producción de contenido engañoso o información no verificada. Se recomienda establecer protocolos de verificación y regulación sobre el uso de estas tecnologías en entornos educativos.

Uno de los principales desafíos con la IA generativa es su capacidad para generar contenido sintético que puede parecer veraz sin estar basado en datos comprobables. Esto es especialmente problemático en la educación, donde la precisión de la información es crucial. Por ello, es fundamental que los docentes y estudiantes adopten prácticas de verificación, como la comparación de respuestas generadas por la IA con fuentes académicas confiables y el uso de herramientas de detección de desinformación.

Además, el uso de IA generativa debe ir acompañado de formación en alfabetización digital y ética de la inteligencia artificial. Los estudiantes deben ser instruidos en el reconocimiento de sesgos algorítmicos, la evaluación de credibilidad de la información y el desarrollo de un enfoque crítico ante los contenidos generados automáticamente. Esto garantizará un uso responsable y consciente de estas herramientas en el ámbito educativo.

9. Evaluación continua del impacto de la IA

El impacto de la IA en la educación debe ser monitoreado y ajustado de forma continua. Es crucial implementar estrategias de auditoría, comparación entre múltiples modelos y triangulación de información para validar su efectividad y mitigar riesgos.

El Reglamento de Inteligencia Artificial exige que los modelos de IA de alto riesgo implementen sistemas de monitoreo continuo que permitan detectar desviaciones en su desempeño y prevenir impactos adversos en la educación.

10. Regulación y cumplimiento normativo



El desarrollo y aplicación de la IA en la educación deben alinearse con marcos regulatorios internacionales, garantizando el respeto a principios de justicia, equidad y protección de datos.

El Reglamento de Inteligencia Artificial introduce un marco normativo detallado para los sistemas de IA en educación, estableciendo requisitos de transparencia, auditoría y supervisión humana. Además, prohíbe el uso de IA para evaluar emociones o realizar perfiles biométricos en contextos educativos, asegurando un uso responsable alineado con los derechos fundamentales.

Para garantizar el cumplimiento de estos principios, es esencial establecer mecanismos de verificación efectivos en la implementación de IA en educación. Esto implica la obligación de realizar auditorías independientes periódicas que evalúen la alineación de los sistemas de IA con los marcos regulatorios, así como la creación de protocolos estandarizados para la recopilación y el análisis de datos utilizados por estos modelos. Además, los organismos educativos deben contar con criterios claros para la certificación de herramientas de IA, asegurando que cumplan con las normativas de seguridad y equidad antes de su despliegue en entornos de aprendizaje.

El Reglamento también enfatiza la importancia de la rendición de cuentas en el uso de IA en educación, exigiendo que las plataformas y proveedores de estos sistemas ofrezcan informes detallados sobre su desempeño, sesgos detectados y medidas correctivas implementadas. Esto permitirá no solo mejorar la fiabilidad de la IA en educación, sino también reforzar la confianza de docentes y estudiantes en el uso de estas tecnologías.



Relación entre técnica y ética en la IA

La IA no se debe integrar sin regulación

La implementación de la inteligencia artificial en la educación no puede realizarse sin un marco regulador que garantice su uso ético y técnico adecuado. La UNESCO y la Comisión Europea han advertido sobre los riesgos de una aplicación no regulada de la IA, como la vulneración de la privacidad, la discriminación algorítmica y la falta de transparencia en la toma de decisiones automatizadas.

El Reglamento de Inteligencia Artificial de la Unión Europea clasifica los sistemas de IA utilizados en educación como de "alto riesgo", lo que implica que deben cumplir con requisitos estrictos de transparencia, supervisión humana y equidad. Sin estos mecanismos de control, la IA podría reforzar desigualdades educativas y socavar la autonomía docente.

Ejemplos de cómo fallos técnicos pueden derivar en problemas éticos

1. Sesgos en la evaluación automática de plataformas de educación en línea: Algunas plataformas educativas que emplean IA para evaluar el rendimiento de los estudiantes han demostrado favorecer ciertos estilos de aprendizaje mientras penalizan a otros. Esto genera desigualdad en la forma en que se califican habilidades diversas, afectando la equidad educativa.
2. Errores en la recomendación de contenido educativo: Algoritmos de recomendación empleados en entornos educativos pueden reforzar sesgos preexistentes al mostrar contenido basado en patrones limitados de interacción. Esto puede derivar en la falta de exposición a nuevas perspectivas y la creación de burbujas de conocimiento en los estudiantes.
3. Privacidad y uso indebido de datos: Algunos modelos de IA en la nube recopilan datos de estudiantes sin mecanismos de consentimiento informados, lo que representa una vulneración del Reglamento General de Protección de Datos (GDPR). La falta de supervisión sobre estos procesos compromete la seguridad de la información estudiantil.



Estrategias para una implementación equilibrada

Para evitar que los problemas técnicos se conviertan en dilemas éticos, se proponen las siguientes estrategias:

1. Supervisión humana constante: La IA debe ser una herramienta de apoyo, no un sustituto del juicio docente. Se recomienda la validación manual de decisiones automatizadas y la capacitación de los docentes en auditoría algorítmica. Dentro de este marco, AI-TPACK resulta fundamental para estructurar el conocimiento pedagógico, técnico y ético de la IA, asegurando una supervisión efectiva y evitando la dependencia ciega de los algoritmos.
2. Regulaciones claras y cumplimiento normativo: La implementación de la IA en educación debe alinearse con marcos regulatorios como el Reglamento de IA de la UE y las directrices de la UNESCO, asegurando que su aplicación no comprometa derechos fundamentales. El Consenso de Beijing subraya la importancia de adoptar regulaciones adaptativas que permitan la innovación sin comprometer la equidad y transparencia.
3. Transparencia y explicabilidad de los algoritmos: Las instituciones educativas deben exigir documentación detallada sobre los modelos de IA utilizados, incluyendo datos de entrenamiento, metodologías de evaluación y mecanismos de control de sesgos.
4. Evaluación continua del impacto: Se deben establecer indicadores de equidad, efectividad pedagógica y aceptación por parte de docentes y estudiantes para ajustar el uso de IA en el aula según su desempeño real. La equidad digital es un criterio clave de esta evaluación, asegurando que la personalización del aprendizaje no excluya a ciertos grupos y que los recursos de IA sean accesibles a toda la comunidad educativa.
5. Uso de modelos abiertos y auditables: Fomentar el uso de modelos de IA de código abierto permite una mayor transparencia y adaptabilidad a contextos educativos diversos, reduciendo la dependencia de proveedores comerciales opacos.

Recomendaciones prácticas para docentes



Checklist para evaluar herramientas de IA

Para garantizar un uso efectivo y responsable de la IA en educación, los docentes pueden seguir este checklist antes de implementar herramientas basadas en inteligencia artificial:

1. Transparencia y explicabilidad

- ¿La herramienta proporciona información sobre cómo genera sus respuestas y recomendaciones?
- ¿El modelo de IA utilizado permite revisar y auditar los datos de entrenamiento para comprender sus sesgos y limitaciones?

Tip: En herramientas como ChatGPT o DeepSeek, revisa si el proveedor ofrece información sobre cómo se entrenó el modelo. Ninguno de los dos comparte sus datos de entrenamiento, pero DeepSeek ofrece sus modelos abiertos en github, mientras que ChatGPT no.

2. Privacidad y protección de datos

- ¿Cumple con normativas como el RGPD? ¿Qué tipo de datos recopila y cómo los gestiona?
- ¿Los datos ingresados por los estudiantes se almacenan en servidores externos o se procesan localmente sin retención de información?

Tip: Modelos como Mistral AI y LLaMA 3 permiten ejecución local, lo que mejora la privacidad. Evita herramientas que no especifiquen qué hacen con los datos ingresados.

3. Supervisión humana

- ¿Permite la intervención y supervisión del docente en sus resultados?
- ¿Se pueden ajustar las respuestas generadas por la IA según el contexto educativo sin alterar la experiencia del estudiante?

Tip: Herramientas como Claude permiten dar instrucciones específicas para personalizar respuestas. En particular, Claude destaca por su capacidad de ajuste en el estilo de escritura, lo que facilita su adaptación a diferentes necesidades pedagógicas.

4. Equidad y accesibilidad

- ¿Es inclusiva para estudiantes con diferentes capacidades y contextos socioeconómicos?
- ¿El modelo de IA ofrece soporte en diferentes idiomas y formatos accesibles para estudiantes con discapacidades?



Tip: Verifica si la IA soporta accesibilidad mediante lectores de pantalla o transcripción de audio. DeepSeek ofrece traducción en múltiples idiomas, lo que puede mejorar la inclusión.

5. Fiabilidad y precisión

- ¿Se han evaluado los posibles sesgos del modelo y su impacto en la educación?
- ¿Los errores o respuestas incorrectas de la IA pueden corregirse manualmente o se depende exclusivamente del proveedor del modelo?

Tip: Evalúa respuestas de la IA en temas especializados. ChatGPT puede generar información convincente pero incorrecta, por lo que se recomienda contrastar con fuentes verificadas.

6. Interoperabilidad

- ¿Puede integrarse con otras plataformas educativas sin problemas?
- ¿Es compatible con las herramientas que los docentes y estudiantes ya usan en su entorno educativo?

Tip: Modelos como Claude permiten exportar datos y resultados en múltiples formatos. Asegúrate de que la herramienta se integre con plataformas como Moodle o Google Classroom.

7. Sostenibilidad y costo

- ¿Es viable económicamente para la institución y no genera dependencia tecnológica?
- ¿Existen opciones de IA de código abierto que puedan sustituir soluciones comerciales sin perder funcionalidad?

Tip: Alternativas como Mistral AI permiten implementación local sin depender de suscripciones costosas. Evalúa opciones de código abierto antes de invertir en IA propietaria.

8. Regulación y cumplimiento ético

- ¿Está alineada con las normativas nacionales e internacionales sobre IA?
- ¿El proveedor de IA facilita documentación clara sobre cómo cumple con regulaciones de protección de datos y ética en la IA?

Tip: Herramientas con certificaciones de cumplimiento normativo (como RGPD o ISO/IEC 27001) garantizan mayor seguridad en su uso. Verifica la documentación técnica antes de su implementación.

9. Estrategias pedagógicas compatibles



- ¿Cómo se adapta a los objetivos de enseñanza y aprendizaje establecidos por el docente?
- ¿Permite a los docentes personalizar sus funciones para alinearse con metodologías pedagógicas específicas?

Tip: Herramientas como ChatGPT pueden adaptarse a distintas metodologías educativas a través de prompts detallados. Diseña guías de uso para los estudiantes para optimizar su aprovechamiento.

10. Evaluación continua

- ¿Se pueden medir y ajustar sus efectos en la práctica educativa?
- ¿El impacto de la IA en el aprendizaje se puede revisar y ajustar mediante métricas e indicadores educativos?

Tip: Algunas herramientas permiten exportar registros de uso para evaluar su impacto. Utiliza plataformas que ofrezcan métricas sobre la interacción de los estudiantes con la IA.



Preguntas Frecuentes (FAQ)

1. ¿Puede la IA reemplazar a los docentes?

No. La IA es una herramienta complementaria que puede apoyar la enseñanza, pero no sustituye el papel del docente en la orientación pedagógica, la evaluación crítica y la interacción humana con los estudiantes. Su uso debe ser supervisado y adaptado a cada contexto educativo.

El rol del docente en el futuro seguirá evolucionando hacia un facilitador del aprendizaje, enfocándose en la personalización de la enseñanza y en el desarrollo del pensamiento crítico en los estudiantes. Más que transmitir información, los docentes serán guías que ayuden a los alumnos a utilizar herramientas tecnológicas de manera ética y efectiva.

2. ¿Es seguro utilizar IA en el aula?

Depende de la herramienta y su nivel de privacidad y seguridad. Se recomienda utilizar plataformas que cumplan con regulaciones como el RGPD y evaluar si los datos de los estudiantes son protegidos adecuadamente. Herramientas como DeepSeek y Mistral ofrecen modelos abiertos que pueden ejecutarse de manera local, lo que mejora la privacidad.

Es importante diferenciar entre la IA como tecnología y las políticas empresariales de los proveedores de modelos en la nube. No es la IA en sí la que representa un riesgo, sino la falta de transparencia en la gestión de datos de ciertas empresas. La clave es elegir herramientas con políticas claras y alineadas con principios de privacidad y equidad.

3. ¿Cómo se puede evitar el sesgo en los modelos de IA?

El sesgo se mitiga utilizando modelos de IA auditables, analizando su entrenamiento y validando sus respuestas con múltiples fuentes. También es importante capacitar a los docentes y estudiantes en pensamiento crítico frente a los resultados generados. Modelos abiertos como LLaMA 3 o DeepSeek permiten más transparencia en este aspecto.

El sesgo en IA depende en gran medida de los datos con los que ha sido entrenada. Si los datos provienen de fuentes sesgadas o reflejan desigualdades estructurales, el modelo



reproducirá esos mismos sesgos. La auditoría y supervisión de los datos de entrenamiento es clave para minimizar estos efectos.

4. ¿Qué modelos de IA son más adecuados para la educación?

Depende del propósito. Modelos como DeepSeek, Mistral y LLaMA pueden servir para generación de texto y análisis sin depender de servidores externos, mientras que herramientas como Claude ofrecen capacidades avanzadas de ajuste de estilo. Se recomienda utilizar IA de código abierto cuando sea posible para mayor control.

Lo ideal es que los modelos de IA en educación puedan ejecutarse en local y adaptarse a necesidades específicas del aula. En caso de usar modelos en la nube, se debe priorizar el uso de IA de código abierto con auditoría pública de datos de entrenamiento. Sin embargo, hay que tener en cuenta que algunas aplicaciones de IA pueden no ser recomendables en entornos educativos debido a su falta de control sobre los datos y sus procesos internos.

5. ¿Cómo pueden los docentes evaluar la efectividad de la IA en el aula?

Deben establecer métricas claras, como el impacto en el aprendizaje, la equidad en el acceso a la información y la facilidad de uso. También se recomienda realizar revisiones periódicas sobre el desempeño de la IA en las actividades educativas.

Para evaluar la IA en el aula, es fundamental medir su impacto con datos concretos: ¿Mejora la comprensión de los alumnos? ¿Aumenta la participación en clase? ¿Genera dependencia o promueve la autonomía? Estas métricas ayudarán a decidir si su uso es beneficioso o necesita ajustes.

6. ¿Qué riesgos tiene el uso de IA en la educación?

Algunos riesgos incluyen la pérdida de privacidad, el sesgo algorítmico, la falta de transparencia en las respuestas y la dependencia excesiva de los estudiantes en la tecnología. Para mitigar estos riesgos, es clave la supervisión docente y el cumplimiento de normativas. Modelos de código abierto pueden reducir ciertos riesgos al ofrecer mayor transparencia.

Además, es necesario considerar cómo la IA impacta la toma de decisiones en el aula. Si los docentes confían demasiado en los modelos, pueden perder su capacidad de análisis crítico



sobre el aprendizaje de sus estudiantes. Es recomendable siempre validar la información y no delegar completamente las decisiones en estos sistemas.

7. ¿Cómo pueden los docentes asegurarse de que los estudiantes no dependan demasiado de la IA?

Fomentando el pensamiento crítico y la autonomía en el aprendizaje. Se pueden diseñar actividades que combinen el uso de IA con la investigación manual, la reflexión y la discusión en grupo.

La IA debe verse como una herramienta de apoyo, no como una solución única. Crear tareas en las que los estudiantes tengan que cuestionar y justificar sus respuestas, además de verificar información con fuentes confiables, los ayudará a no depender ciegamente de la IA.

8. ¿Se pueden usar herramientas de IA en exámenes y evaluaciones?

Depende de la política del centro educativo. En general, se recomienda limitar su uso en evaluaciones formales y enfocarse en tareas donde la IA pueda actuar como apoyo sin reemplazar la capacidad analítica del estudiante.

Si bien algunas herramientas de IA pueden ser útiles en la evaluación formativa, su uso en exámenes debe ser controlado. Es recomendable combinar pruebas tradicionales con actividades supervisadas que permitan verificar la comprensión real de los estudiantes sin depender exclusivamente de la IA.

9. ¿Qué tan confiables son las respuestas de la IA?

Las respuestas pueden ser útiles, pero no siempre son precisas. Es importante contrastarlas con fuentes verificadas y no asumir que todo lo generado por IA es correcto. Modelos como DeepSeek y Claude tienen configuraciones que permiten ajustar la precisión de las respuestas en función del contexto.

El nivel de confianza en la IA debe depender del control que tengamos sobre sus datos y procesos. Modelos abiertos permiten una mayor verificación de sus fuentes y sesgos, mientras que los modelos cerrados requieren mayor precaución al interpretar sus respuestas.

10. ¿Cuál es el futuro de la IA en la educación?



El futuro de la IA en la educación dependerá de su regulación, integración ética y desarrollo de herramientas adaptativas. Se espera que continúe evolucionando como un recurso que personaliza el aprendizaje y mejora la eficiencia del docente, sin sustituir la enseñanza humana. Modelos de código abierto jugarán un papel clave en este desarrollo al garantizar mayor accesibilidad y transparencia.

Además, la IA en educación deberá centrarse en desarrollar herramientas adaptadas a contextos educativos específicos y alineadas con principios pedagógicos sólidos. La clave no será solo la tecnología, sino cómo se diseña su integración para potenciar la enseñanza sin deshumanizarla.

Bibliografía

- **UNESCO.** (2021). *Artificial Intelligence and Education: Guidance for Policy-makers*. UNESCO.
- **Comisión Europea.** (2025). *Declaración sobre los Derechos Digitales para la Década Digital*. Comisión Europea.
- **Unión Europea.** (2016). *Reglamento General de Protección de Datos (GDPR)*. Parlamento Europeo y Consejo de la Unión Europea.
- **Congreso de los EE.UU.** (1998). *Children's Online Privacy Protection Act (COPPA)*. U.S. Federal Trade Commission.
- **UNESCO.** (2019). *Beijing Consensus on Artificial Intelligence and Education*. UNESCO.
- **UNESCO.** (2021). *Recomendación sobre la ética de la inteligencia artificial*. UNESCO.
- **UNESCO.** (2022). *Guía para las personas a cargo de formular políticas sobre IA en educación*. UNESCO.
- **INTEF.** (2023). *Guía sobre el uso educativo de la inteligencia artificial*. Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado.
- **UOC.** (2024). *Guía para la aplicación de la inteligencia artificial a la acción docente*. Universitat Oberta de Catalunya.
- **BBC News.** (2020, August 25). *Ofqual chief Sally Collier steps down after exams chaos*. BBC News.
<https://www.bbc.com/news/education-53909>