

La Inteligencia Artificial en contextos educativos: usos éticos de la información y alfabetización digital

Alfonso López-PulidoUniversidad Internacional de La Rioja (España)  **José Manuel Sánchez Ramírez**Universidad Internacional de La Rioja (España)  <https://dx.doi.org/10.5209/dere.102342>

Recibido: 12/12/2024 • Revisado: 13/01/2025 • Aceptado: 03/02/2025

ES Resumen. La proliferación de tecnologías digitales, incluyendo la Inteligencia Artificial (IA), ha modificado la manera en la que se accede, se procesa y se comparte la información. No obstante, esta mutación también ha intensificado problemas como la desinformación, las dificultades en la verificación de la información y los posibles usos no éticos de la IA, especialmente en contextos educativos y de investigación académica, así como el peligro de la disminución de las capacidades de expresión lingüística de los alumnos y de su independencia de pensamiento. El presente estudio explora cómo estas dinámicas afectan a la educación, los desafíos que deben superar el estudiantado y el profesorado, y las estrategias para mitigar estos riesgos, destacando la alfabetización digital.

Palabras clave. Inteligencia Artificial, ética, desinformación, alfabetización digital.

ENG Artificial Intelligence in educational contexts: ethical uses of information and digital literacy

ENG Abstract. The proliferation of digital technologies, including Artificial Intelligence (AI), has transformed the way we access, process and share information. However, this mutation has also intensified problems such as misinformation, difficulties in verifying information and possible unethical uses of AI, especially in educational and academic research contexts, as well as the danger of diminishing students' linguistic expression capabilities and their independent thinking. In this paper we explore how these dynamics affect education, the challenges that students and teachers must overcome, and strategies to mitigate these risks, highlighting digital literacy.

Keywords. Artificial Intelligence, ethics, disinformation, digital literacy.

Sumario. 1. Introducción. 2. La desinformación como reto global en la educación. 2.1 Desinformación con IA. 2.2 Usos No Éticos de la IA en Educación. 2.3 Falta de Reflexión Crítica. 3. Procedimientos para mitigar la desinformación en la educación con inteligencia artificial. 3.1 Desarrollo de Habilidades de Pensamiento Crítico. 3.2 Alfabetización Mediática y verificación de la información. 3.3 Marco de competencia digital docente. 3.4 Implementación Ética y Responsable de herramientas de la IA. 3.5 Uso de Sistemas de Tutoría Inteligente y Personalización. 3.6 Regulación del contenido digital y uso académico de la IA. 4. Conclusiones. 5. Referencias.

Cómo citar: López-Pulido, A. y Sánchez Ramírez, J. M. (2025). La Inteligencia Artificial en contextos educativos: usos éticos de la información y alfabetización digital. *Derecom* 38(1), 3-11. <https://dx.doi.org/10.5209/dere.102342>

1. Introducción

La IA ha experimentado una “aceleración exponencial” en poco tiempo, pasando de ser algo sólo reservado a estrechos círculos de expertos (Sanagustín, 2024, p. 3) a una quasi omnipresencia en todas las actividades cotidianas (Hashmi y Bal, 2024). La desinformación y las falsificaciones generadas por la IA están impactando sobre un elevado número

de facetas del mundo actual (Gregory, 2023), lo que acarrea una serie de consecuencias y la necesidad de adoptar, desde la sociedad, una respuesta crítica y firme (Shoaib et al., 2023). Aparte de los problemas éticos y legales (Öhman, 2020), los desafíos que se le plantean al periodismo y a los medios de comunicación en general (Wahl-Jorgensen y Carlson, 2021), la influencia sobre la seguridad individual y la priva-

cidad (Chauhan y Keprate, 2022; Ciftci et al., 2023), el desgaste de la confianza pública (Pawelec, 2022) o los efectos sobre la democracia y la opinión pública (Schiff et al., 2023), el impacto sobre el mundo educativo hace que se planteen multitud de interrogantes sobre los perjuicios y sobre los beneficios derivados del empleo de la IA. Así, los chatbots de IA, como es ChatGPT, desde que aparecieran en 2022, han hecho surgir una honda preocupación en el ámbito educativo (Dempere et al., 2023). Una muestra es la inquietud ante asuntos tales como la posibilidad de emplearlos para elaborar los trabajos de los estudiantes (Roose, 2022; Shankland, 2022), lo cual pone en riesgo el desarrollo de competencias esenciales (Shrivastava, 2022).

La importancia de este fenómeno, así como sus implicaciones, nos han llevado a plantearnos como objetivos, en el presente trabajo, la exposición de los problemas más relevantes (la desinformación como reto global en la educación, los usos no éticos de la IA en la misma y la falta de reflexión crítica), así como las soluciones (desarrollo de habilidades de pensamiento crítico, alfabetización mediática, verificación de la información, la implementación ética y responsable de herramientas de la IA, el uso de sistemas de tutoría inteligente y personalización, así como la regulación del contenido digital y uso académico de la IA).

2. La desinformación como reto global en la educación

La desinformación constituye un desafío fundamental en la sociedad actual, incrementando exponencialmente por el acceso masivo a plataformas digitales y a cualquier tipo de información de fuentes no fiables o contrastadas. En el contexto educativo, este fenómeno afecta a la aptitud de los alumnos para discernir qué información es información fiable y a su habilidad para distinguir hechos de opiniones o de información no válida académicamente, lo que compromete su desarrollo crítico y ético.

La inexistencia de una legislación rigurosa y de sistemas de vigilancia adecuados, ya que debería haber algún tipo de regulación (Bushwick y Mukerjee, 2022), permite, a través de Internet, la expansión de contenidos erróneos y/o sesgados generados por la IA, que adulteran esta fuente de conocimiento, lo que implica la aceptación, por parte de los alumnos, de una información no veraz y, a la par, comprometen los futuros modelos GPT, ya que estos se adiestrarán según los contenidos de la red que los mismos modelos GPT han originado y que, por ello, poseen asimismo sus sesgos y errores. No es superfluo insistir en la generación de conocimiento no fiable o con errores, dado que la IA, en realidad, no entiende el propio texto que ha producido. La información producida por los modelos GPT se encuentra mediatisada por los valores de los poseedores o/y generadores de los datos que emplean, por lo que nos hallamos ante el peligro de que se esté produciendo una reducción del abanico de expresión de ideas y opiniones, lo que conlleva que las posturas minoritarias sean excluidas, así como los grupos poblacionales que tengan una escasa presencia digital en la red (UNESCO, 2024). De ahí que sea imprescindible un desarrollo ético de la IA, reduciendo consciente-

mente el sesgo en los algoritmos y en los datos de entrenamiento, de tal forma que se asegure la equidad y se elimine cualquier tipo de discriminación (Henry y Oliver, 2022; Giovanola y Tiribelli, 2023).

Por otro lado, la proliferación de información falsa en redes sociales desjerarquiza contenidos y dificulta la identificación de fuentes confiables (Hangloo y Arora, 2022). Esto ha llevado a la necesidad de integrar estrategias educativas que promuevan la alfabetización mediática y digital, permitiendo a los estudiantes evaluar críticamente los datos que consumen (UNESCO, 2024; MEFPD, 2024). Por ello, esta alfabetización debe insistir en los riesgos de creer excesivamente en todo lo que nos proporcionan las tecnologías basadas en la IA, alertando especialmente sobre todo lo relativo al riesgo de plagio que poseen los chatbots, ya que seleccionan sus propuestas de los datos de entrenamiento (Ghosal, 2021). Pueden citarse, asimismo, como ejemplo, las alucinaciones, que son respuestas de la IA que no tienen ninguna explicación desde el punto de vista de los datos empleados en el entrenamiento (Cao et al., 2017; Falke et al., 2019). De ahí la importancia de insistir en que cualquier sistema de IA, por muy bien diseñado que esté, puede cometer errores, ya que depende de la calidad de los datos utilizados para entrenarlo (Kaplan y Haenlein, 2019).

2.1. Desinformación con la IA:

- La proliferación de *deepfakes* y la desinformación generada por la IA representan amenazas significativas para la integridad de los ecosistemas de información (Dempere et al., 2023). La posibilidad de alterar imágenes para crear otras nuevas, muy similares a las auténticas, pero falsas, con el objeto de llevar a cabo suplantaciones, favorece la difusión de conductas contrarias a la ética y que caen dentro del ámbito delictivo (UNESCO, 2024). Se han propuesto marcos integrados que combinan algoritmos de detección avanzados, colaboración entre plataformas y políticas para mitigar estos riesgos (Shoaib et al., 2023), siendo fundamental la investigación de cómo pueden crearse y ser detectadas (Mirskey y Lee, 2022).
- Las plataformas digitales modernas y las tecnologías emergentes, incluida la IA, introducen nuevos riesgos en la gestión de crisis. Se necesita un enfoque sistemático y coordinado para los esfuerzos sostenibles contra la desinformación (Hangloo y Arora, 2022; Karginshak y Jin, 2023).

2.2. Usos No Éticos de la IA en Educación

Problemas Éticos:

- Los problemas éticos más destacados, del uso de aplicaciones móviles impulsadas por la IA en la educación, incluyen la mala utilización de datos privados y la vigilancia de la vida de las personas (Klimova et al., 2023). Se recomienda lo que podríamos denominar como la “algoritmovigilancia”, con la finalidad de monitorear y prevenir los efectos adversos de los algoritmos, que podría basarse en la predicción de acciones individuales, partiendo de los datos

que se obtienen de las interacciones entre los usuarios y el chatbot, extremos que nos llevarían a exponer cuestiones sobre la igualdad y la libertad personal (Akgun y Greenhow, 2022). De ahí que los patrones que pueden identificar los chatbots, le aportan, al propietario del algoritmo, lo que podemos denominar como ventaja informativa (Murtaelia et al., 2021). Un posible ejemplo es que algunos centros de enseñanza superior pudiesen utilizar los chatbots para mediatizar los comportamientos de los alumnos, en el sentido de que optasen, de forma artificial, por cursar unos estudios en lugar de otros (Dempere et al., 2023).

- La rápida integración de la IA en el sistema educativo ha planteado desafíos éticos, como el impacto en las habilidades de pensamiento crítico, la autonomía y la toma de decisiones éticas de los estudiantes. Se propone a las instituciones educativas establecer políticas y directrices significativas para maximizar el uso beneficioso de la IA en las instituciones educativas (Saylam et al., 2023).

Falta de Reflexión Crítica:

- A pesar de los avances que se están produciendo, existe una falta de juicio crítico sobre los retos y peligros de la IA en la educación superior. La ética en los estudios de la IA es un área de investigación ignorada, a pesar de la penetración de dicha IA en todos los niveles de la educación (Bozkurt et al., 2021). La tenue conexión con los postulados pedagógicos teóricos y la necesidad de explorar puntos de vista que aúnén ética y educación en el empleo de la IA, son áreas que requieren mayor atención, de ahí que deberían centrarse los esfuerzos en estos segmentos (Zawacki-Richter et al., 2019).
- Sin embargo, la IA, especialmente el aprendizaje automático (Machine Learning), se ha utilizado para combatir la desinformación mediante diversas técnicas.

De ellas podemos mencionar las siguientes:

- Clasificación automática. Es la forma más simple de llevar a cabo un análisis de la desinformación. Así, mediante un grupo de datos etiquetados, puede entrenarse un modelo de clasificación para discernir los contenidos verdaderos de los falsos (Nakamura et al., 2020). No obstante, presenta el hándicap de que no es exportable de unos dominios a otros (Shu et al., 2017)
- Extracción de características. Se focaliza en hallar características de desinformación que sean susceptibles de poder emplearse, con la finalidad de poder detectar aquellos contenidos que lo sean (Zhou et al., 2020).
- Fact-checking híbrido. Consiste en la búsqueda de contenidos engañosos mediante periodistas especializados. A pesar de sus buenos resultados, llegados a un estadio determinado acaban por convertirse en un obstáculo (Montoro-Montarroso et al., 2023).

- Creación de contenido sintético realista (Xu et al., 2023), si bien esto último, precisamente, a la par puede generar también más desinformación (Massod et al., 2022). Sin embargo, estas soluciones aún no han trascendido ampliamente los laboratorios de investigación, dada la limitación que impone el trabajar con grupos de datos de capacidad limitada. Por ello, la investigación debe dirigirse a la creación de sistemas fundamentados en la IA que sean confiables para poder detectar la desinformación y no depender de medios automatizados en su totalidad (Montoro-Montarroso et al., 2023).

3. Procedimientos para mitigar la desinformación en la educación con inteligencia artificial

Eliminar los efectos nocivos de la desinformación en la educación, mediante la IA, puede lograrse a través de varias estrategias clave, aparte de lo mencionado más arriba sobre el Machine Learning. Podemos destacar el desarrollo de habilidades de pensamiento crítico, la alfabetización mediática, la verificación de la información y la implementación ética y responsable de las herramientas de la IA.

3.1. Desarrollo de Habilidades de Pensamiento Crítico

El desarrollo de habilidades de pensamiento crítico, tales como el análisis, la evaluación y la generación de argumentos o información, basándose en evidencias sólidas y lógicas, es esencial para que los estudiantes y los investigadores puedan distinguir entre los mensajes veraces y la desinformación, a la vez que se mantiene la integridad académica. Es necesario el desarrollo de métodos de enseñanza específicos que fomenten el pensamiento crítico (August y Trostle, 2018), para apoyar al alumnado a lograr un mayor entendimiento de los contextos en los que se utiliza la IA y a emplearla de manera efectiva y ética. El empleo de la IA en contextos educativos e investigadores es altamente beneficioso, siempre que se preserven la integridad académica y no se menoscaben las habilidades de pensamiento crítico (Rusandi et al., 2023). De ahí que la IA pueda dar lugar a una mutación de la enseñanza a través de elementos tales como el aprendizaje autónomo o la individualización de los contenidos, lo que redundaría en garantizar la transparencia, la igualdad y la seguridad (Ortiz Muñoz, 2024).

3.2. Alfabetización Mediática y verificación de la información

La alfabetización mediática es crucial para proporcionar a los estudiantes las herramientas necesarias para identificar investigaciones creíbles y sólidas, de ahí que enseñarles a discernir cuándo se ha utilizado la IA y a identificar noticias falsas puede mejorar significativamente su capacidad para realizar investigaciones efectivas en la era digital (Black y Fullerton, 2020; Shoaib et al., 2023).

La alfabetización mediática, digital y en datos, es clave para contrarrestar la desinformación. Estas competencias permiten a los estudiantes no solo consumir información críticamente, sino también

participar activamente en la creación de contenido ético y veraz. Algunos estudios subrayan la importancia de enseñar habilidades para evaluar la validez de las fuentes, fomentar el pensamiento crítico y entender la lógica que hay detrás de la generación de noticias falsas (Kajimoto y Fleming, 2019).

3.3. Marco de competencia digital docente:

En España (Resolución de 4 de mayo de 2022), se ha establecido un marco de competencia digital docente por el que se establecen las “Áreas, competencias, etapas, niveles e indicadores de logro y exemplificación a través de afirmaciones sobre el desempeño correspondiente al nivel de cada competencia” (Anexo I, p. 67980).

Podemos señalar algunas de estas áreas:

Área 1. Compromiso profesional.

1.1 Comunicación organizativa: Utilizar las tecnologías digitales establecidas por las Administraciones Educativas o por los titulares del centro, para aplicar las estrategias de comunicación organizativa entre los agentes de la comunidad educativa y de estos con terceros.

1.2 Participación, colaboración y coordinación profesional: Utilizar las tecnologías digitales para participar en los órganos colegiados de gobierno y de coordinación docente del centro.

1.3 Práctica reflexiva: Reflexionar sobre la práctica pedagógica digital que se desarrolla en el aula y en el centro con la finalidad de aplicar las mejoras identificadas a través de este proceso.

1.4 Desarrollo profesional digital continuo: Desarrollar de forma continuada las competencias profesionales docentes a través de medios digitales. Mantener actualizada la competencia digital docente.

1.5 Protección de datos personales, privacidad, seguridad y bienestar digital: Utilizar de manera responsable, segura y saludable las tecnologías digitales para evitar riesgos laborales, personales y en el entorno y para garantizar el bienestar físico, psicológico y social del alumnado al utilizar las tecnologías digitales.

Área 2. Contenidos digitales

2.1 Búsqueda y selección de contenidos digitales: Localizar, evaluar y seleccionar contenidos digitales de calidad para apoyar y mejorar la enseñanza y el aprendizaje. Considerar, de forma específica, el objetivo de aprendizaje, el contexto, el enfoque pedagógico, el tipo de licencia y aspectos técnicos que garanticen la accesibilidad universal, la usabilidad y la interoperabilidad.

2.2 Creación y modificación de contenidos digitales: Modificar y adaptar los contenidos educativos digitales respetando las condiciones de uso establecidas por cada licencia. Crear nuevos contenidos educativos digitales dentro de entornos seguros. Considerar, de modo específico, el objetivo de aprendizaje, el contexto, el enfoque pedagógico y los destinatarios al diseñar y crear o modificar los contenidos digitales.

2.3 Protección, gestión y compartición de contenidos digitales: Catalogar los contenidos educativos digitales y ponerlos a disposición de la comunidad educativa y del colectivo profesional

utilizando entornos seguros. Proteger eficazmente los contenidos digitales. Respetar y aplicar correctamente la normativa sobre propiedad intelectual y derechos de autor en la gestión y compartición de contenidos digitales.

Área 3. Enseñanza y aprendizaje

3.1 Enseñanza: Integrar en las programaciones didácticas el uso de las tecnologías digitales para mejorar la eficacia de las prácticas docentes. Gestionar y coordinar adecuadamente las intervenciones didácticas digitales. Desarrollar y experimentar con nuevos formatos y métodos pedagógicos para la enseñanza y para el aprendizaje.

3.2 Orientación y apoyo en el aprendizaje: Utilizar las tecnologías y servicios digitales, cumpliendo con las medidas de seguridad y protección de datos. Emplear las tecnologías digitales para ofrecer orientación y asistencia pertinente y específica. Desarrollar y experimentar nuevas vías y formatos para ofrecer orientación y apoyo, respetando los derechos digitales de todo el alumnado y evitando cualquier tipo de discriminación o sesgo.

3.3 Aprendizaje entre iguales: Seleccionar y utilizar tecnologías digitales seguras para mejorar el aprendizaje del alumnado a través de la colaboración.

3.4 Aprendizaje autorregulado: Utilizar las tecnologías digitales para favorecer en el alumnado la metacognición.

Área 4. Evaluación y retroalimentación

4.1 Estrategias de evaluación: Utilizar las tecnologías digitales para el diseño de los medios e instrumentos de evaluación e implementarlos cumpliendo con las medidas de seguridad y protección de datos personales.

4.2 Analíticas y evidencias de aprendizaje: Generar, almacenar, validar, seleccionar, analizar e interpretar las evidencias digitales sobre la actividad, el rendimiento y el progreso del alumnado con el fin de mejorar el proceso de enseñanza y el aprendizaje, respetando la normativa vigente en cuanto a protección de datos.

4.3 Retroalimentación y toma de decisiones: Utilizar las tecnologías digitales para ofrecer retroalimentación al alumnado. Adaptar las estrategias de enseñanza y proporcionar refuerzo específico a partir de los datos obtenidos. Informar al alumnado y a las familias y facilitar la comprensión de las evidencias de aprendizaje aportadas por las tecnologías digitales para que sean utilizadas en la toma de decisiones.

Área 5. Empoderamiento del alumnado

5.1 Accesibilidad e inclusión: Emplear las tecnologías digitales para facilitar el aprendizaje de todo el alumnado eliminando las barreras contextuales para su presencia, participación y progreso. Garantizar la accesibilidad física, sensorial y cognitiva a los recursos digitales. Adoptar medidas que promuevan la equidad y permitan reducir o compensar la brecha digital y el impacto de las desigualdades socioculturales y económicas en el aprendizaje.

5.2 Atención a las diferencias personales en el aprendizaje: Utilizar las tecnologías digitales para atender las diferencias del alumnado.

5.3 Compromiso activo del alumnado con su propio aprendizaje: Integrar las tecnologías digitales en estrategias pedagógicas que promuevan el compromiso activo del alumnado incentivando el desarrollo de operaciones cognitivas complejas y de competencias transversales, como el pensamiento crítico o la creatividad.

Área 6. Desarrollo de la competencia digital del alumnado

6.1 Alfabetización mediática y en el tratamiento de la información y de los datos.

6.2 Comunicación, colaboración y ciudadanía digital: Diseñar, implementar e integrar propuestas pedagógicas para el desarrollo y evaluación de la competencia digital del alumnado en la comunicación.

6.3 Creación de contenidos digitales: Diseñar, implementar e integrar, en los procesos de enseñanza y aprendizaje, propuestas pedagógicas para el desarrollo y evaluación de la competencia digital del alumnado en la creación y reelaboración de contenidos digitales.

6.4 Uso responsable y bienestar digital: Diseñar, implementar e integrar, en los procesos de enseñanza y aprendizaje, propuestas pedagógicas para el desarrollo y evaluación de la competencia digital del alumnado en el uso seguro, responsable, crítico, saludable y sostenible de las tecnologías digitales.

6.5 Resolución de problemas: Diseñar, implementar e integrar, en los procesos de enseñanza y aprendizaje, propuestas pedagógicas para el desarrollo y evaluación de la competencia digital del alumnado en la utilización de las tecnologías digitales para resolver problemas cotidianos y desenvolverse, como prosumidor, de forma creativa y crítica en un mundo digitalizado (Anexo I, pp. 67980-68024).

En el contexto educativo, es muy importante que tanto docentes como alumnos adquieran estas competencias como objeto mismo de aprendizaje (Ortiz Muñoz, 2024; Díaz y Nussbaum, 2024) ya que, como la propia resolución apunta:

Forman parte de la alfabetización básica de toda la ciudadanía en las etapas educativas obligatorias y de educación de adultos y constituyen un elemento esencial de la capacitación académica y profesional en las enseñanzas postobligatorias. Por otra, los docentes y el alumnado han de emplearlas como medios o herramientas para desarrollar cualquier otro tipo de aprendizaje (Anexo, p. 67979)

3.4. Implementación Ética y Responsable de herramientas de la IA

La implementación ética y responsable de herramientas de la IA en la educación es fundamental para mitigar los riesgos asociados con la desinformación. El profesorado debe ser consciente de los posibles sesgos en los algoritmos y los datos, así como de los problemas de privacidad (Chauhan y

Keprate, 2022) y la disminución de la interacción humana (Zawacki-Richter et al., 2019; Baidoo-Anu y Ansah, 2023; Dempere et al., 2023).

La inteligencia artificial en la educación superior puede generar sesgos, plagio (Ghosal, 2021), micro-gestión y problemas de privacidad (Chauhan y Keprate, 2022), además de afectar a la ética, la creatividad y el pensamiento crítico (Meckler y Verma, 2022; Ivanov, 2023).

Es importante que los miembros del cuerpo docente utilicen proactivamente y de manera ética las herramientas de la IA (Bouteraa et al., 2024) para mitigar riesgos como el fraude académico. Chat-GPT y otras IA generativas ofrecen beneficios como apoyo a la investigación, la evaluación mecanizada en los estudios superiores y una optimización de las interacciones ser humano/computadoras (Brito et al., 2019; Rusandi et al., 2023; Cámara Molina, 2024), pero también presenta riesgos como fraude académico (Meckler y Verma, 2022), seguridad de la evaluación en línea y otro tipo de consecuencias socioeconómicas, tales como brechas en la alfabetización digital, desplazamiento laboral y la ansiedad provocada por la propia IA (Dempere et al., 2023).

No obstante, están apareciendo herramientas, como los métodos de verificación basados en la IA, que se revelan como aliados importantes para identificar información falsa en entornos educativos.

3.5. Uso de Sistemas de Tutoría Inteligente y Personalización

Los sistemas de tutoría inteligente y la personalización pueden ayudar a los estudiantes a recibir apoyo académico adaptado a sus necesidades individuales, lo que puede mejorar su comprensión y reducir la dependencia de fuentes de información no verificadas. La investigación sobre la IA en la educación superior se centra en la elaboración de perfiles, la evaluación y la personalización, pero es necesaria una reflexión más crítica sobre los desafíos, los riesgos y los enfoques críticos (Zawacki-Richter et al., 2019).

Sin embargo, un número importante de investigadores sostienen que los chatbots no poseen cualidades humanas imprescindibles como la empatía, la asertividad o la capacidad de captar las emociones (Murtarellia et al., 2021). Por ello, no es conveniente que la IA reemplace a los profesores, ya que es incapaz de proporcionar a los estudiantes la humanidad de los docentes, ni tampoco experiencias de aprendizaje valiosas como la reflexión existencial o el rigor ético (Felix, 2020). No obstante, desde otra perspectiva se sostiene que, cada vez más, la línea que separa lo humano de la máquina se está haciendo muy difusa (Guzmán y Lewis, 2022), lo que conlleva reestructurar constantemente qué papeles deben atribuirse (Hevner y Storey, 2023), de ahí que se haga apremiante que los supuestos éticos pasen a un primer plano (Alasadi y Balz, 2023).

Teniendo en cuenta ambas posturas, lo que se colige es que estamos ante el umbral de una nueva forma de entender la enseñanza, tanto en los modos de aprendizaje, como en qué se le va a pedir impartir al profesorado (Cámara Molina, 2024).

3.6. Regulación del contenido digital y uso académico de la IA

En el ámbito académico debe tenerse en cuenta que nos encontramos ante cambios importantes y, a la par, delicados, puesto que surgen tanto peligros como ventajas y es preciso manejar todo ello desde un enfoque ético (Cámara Molina, 2024).

Es muy necesario desarrollar políticas que promuevan la transparencia en la creación de contenido digital y el etiquetado de materiales generados por la IA. Los riesgos se hallan, principalmente, en la extendida y errónea idea de que, tanto los datos como la codificación, en lugar de ser valores relativos y cuestionables, son fundamentos absolutos, ya que la tendencia es reducir, tanto la educación como el propio aprendizaje, a un mero razonamiento estadístico, prescindiendo de los valores intrínsecos al ser humano. La problemática educativa y la trascendencia del hecho educativo poseen una complejidad tal, que aún no es posible aprehenderlas y solucionarlas desde el punto de vista tecnológico (Selwyn, 2016).

El Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado (INTEF), organismo dependiente del Ministerio de Educación, Formación Profesional y Deportes de España (MEFPD), ha elaborado una *Guía sobre el uso de la Inteligencia Artificial en el ámbito educativo* (2024). En ella, además de orientar sobre las posibilidades y aplicaciones de la IA en educación, dedica un capítulo completo a las propuestas prácticas para abordar la relación que debe establecerse entre la ética y la IA con el alumnado, destacando cómo tienen que trabajarse con los estudiantes los sesgos algorítmicos, las *deepfakes*, la privacidad y la seguridad de los datos, la equidad y la inclusión (MEFPD, 2024).

También, en este sentido, es fundamental que el uso de la inteligencia artificial, en el aprendizaje académico y las tareas que deben realizar tanto docentes como alumnos se realicen de forma ética y responsable, verificando que todas las aplicaciones de la IA respeten la privacidad (Chauhan y Keprete, 2022), promuevan la equidad y transparencia, y eviten la desinformación (Pedró et al., 2019; Henry y Oliver, 2022; Cámara Molina, 2024). Estudiantes y docentes deben utilizar estas tecnologías de apoyo como herramientas complementarias, sin delegar en ellas su responsabilidad académica y garantizando siempre la autoría original y la integridad del trabajo presentado.

La Universidad Internacional de La Rioja, partiendo de su Responsabilidad Social Corporativa y adaptándose a los requerimientos de una sociedad en continua evolución, ha sido una precursora en el uso crítico y responsable de la IA, publicando la *Declaración UNIR para un uso ético de la Inteligencia Artificial en Educación Superior* (27/02/2024). Su objetivo es el de “sentar unas bases de consenso que guíen a todos los actores que forman parte del grupo Proeduca (profesores, estudiantes, personal de gestión, junta de gobierno, investigadores, etc.) en el uso, aplicación, e incluso desarrollo de soluciones basadas en IA”, ya que es ineludible compaginar los usos positivos de la IA con la práctica docente e investigadora (Burgos, 2024).

Dicha declaración se articula sobre los siguientes principios:

- Principio de Contribución Social
- Principio de Equidad
- Principio de Capacitación
- Principio de Supervisión
- Principio de Ética
- Principio de Confidencialidad
- Principio de Transparencia
- Principio de Sostenibilidad
- Principio de Conocimiento
- Principio de Trazabilidad

En suma, “UNIR se compromete a utilizar la inteligencia artificial de manera ética y responsable, a potenciar el conocimiento, promover la innovación y beneficiar a la sociedad en su conjunto” (Burgos, 2024).

Por último, merece señalarse que, en un nivel más amplio, la UNESCO ha publicado siete pasos que las instituciones de gobierno pueden dar para la regulación de la IA generativa, aprovechando sus virtudes y ejerciendo sobre ella un control público:

“Paso 1: Adscribirse a regulaciones generales de protección de datos internacionales o regionales, o desarrollar normativas nacionales.

Paso 2: Adoptar/revisar y financiar estrategias sobre IA para todo el sector gubernamental.

Paso 3: Consolidar e implementar normativas específicas sobre la ética de la IA.

Paso 4: Adaptar o hacer cumplir las leyes de derechos de autor existentes para regular los contenidos generados por IA.

Paso 5: Elaborar marcos regulatorios sobre la IA generativa.

Paso 6: Construir capacidades para el uso adecuado de la IAG en educación e investigación.

Paso 7: Reflexionar sobre las implicaciones a largo plazo de la IAG en para la educación y la investigación” (2024, pp. 19–20).

4. Conclusiones

La desinformación y los usos no éticos de la IA en la educación son problemas complejos que requieren enfoques multidisciplinares y colaborativos. La implementación de políticas éticas, la vigilancia de algoritmos y la educación sobre el empleo de la IA con responsabilidad son esenciales para mitigar los riesgos asociados. Además, es crucial desarrollar sistemas de IA confiables y colaborativos para la detección temprana de desinformación y fomentar una reflexión crítica sobre los desafíos éticos en la educación.

Para eliminar la desinformación en la educación con la IA es crucial desarrollar habilidades de pensamiento crítico, fomentar la alfabetización mediática y asegurar una implantación desde el punto de vista ético y con responsabilidad de las herramientas de la IA. Estas estrategias pueden ayudar a los estudiantes y educadores a manejarla de forma consciente y segura en el contexto digital actual, reduciendo la propagación de desinformación y mejorando la calidad de la educación.

Es necesaria una profunda reflexión, desde diversas posturas y puntos de vista, que aborde la on-

tología de las experiencias y creaciones que genera la IA, así como un refuerzo de nuestros parámetros de conocimiento y discernimiento de lo que es auténtico o falso, todo ello con vistas a devolver a las personas una autonomía que se ve cercenada por representaciones de la IA que se dan sin su consentimiento.

Por último, deben señalarse los innumerables beneficios y potencialidades, en los contextos educativos y de la investigación, que posee una IA sabiamente gestionada, que no se superponga al pensamiento crítico y a la integridad académica, y que posibilite cohabitar éticamente con las formas de conocimiento que se venían llevando a cabo.

5. Referencias

- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: addressing ethical challenges in K-12 settings. *AI Ethics* 2, 431-440. <https://doi.org/10.1007/s43681-021-00096-7>
- Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in Education and Research: opportunities, concerns, and solutions. *Journal of Chemical Education*, 100, 8, 2965-2971. <https://pubs.acs.org/doi/10.1021/acs.jchemed.3c00323>
- August, E., & Trostle, J. A. (2018). Using writing assignments to promote critical thinking, learning and professional identity: the epidemiology workplace writing repository. *Journal PublicHealth*, 40, 419-422.
- Baidoo-Anu, D., & Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7, 52-62. <https://doi.org/10.61969/jai.1337500>
- Black, J., & Fullerton, C. (2020). Digital Deceit: Fake News, Artificial Intelligence, and Censorship in Educational Research. *Open Journal of Social Sciences*, 8, 7, 71-88. <https://doi.org/10.4236/jss.2020.87007>
- Bouteraa, M., Bin-Nashwan, S.A., Al-Daihani, M., Dirie, K.A., Benlahcene, A., Sadallah, M., Zaki, H.O., Lada, S., Ansar, R., Ming Fook, L., Chekima, B. (2024). Understanding the diffusion of AI-generative (ChatGPT) in higher education: Does students' integrity matter? *Computers in Human Behavior Reports*, 14, 100-127. <https://doi.org/10.1016/j.chbr.2024.100402>
- Bozkurt, A., Karadeniz, A., Bañeres, D., Guerrero-Roldán, A., & Rodríguez, M. (2021). Artificial Intelligence and Reflections from Educational Landscape: A Review of AI Studies in Half a Century. *Sustainability*, 13, (2), 1-14. <https://doi.org/10.3390/SU13020800>
- Brito, C., Ciampi, M., Sluss, J., & Santos, H. (2019). Trends in engineering education: a disruptive view for not so far future. 2019 18th international conference on information technology based higher education and training. *ITHET*, 1-5.
- Burgos, D. (2024). *Declaración UNIR para un uso ético de la Inteligencia Artificial en Educación Superior*. Universidad Internacional de La Rioja. <https://www.unir.net/actualidad/responsabilidad-social-corporativa/declaracion-unir-para-un-uso-etico-de-la-inteligencia-artificial-en-educacion-superior/>
- Bushwick, S., & Mukerjee, M. (2022). ChatGPT explains why AIs like ChatGPT should be regulated. Disponible en: <https://www.scientificamerican.com/article/chatgpt-explains-why-ais-like-chatgpt-should-be-regulated/>
- Cámara Molina, J. C. (2024). El uso de la IA como herramienta para la investigación académica: políticas editoriales y condicionantes éticos en el camino hacia su normalización. *Derecom*, 37, 33-46. <https://dx.doi.org/10.5209/dere.98112>
- Cao, Z., Wei, F., Li, W., & Li, S. (2017). Faithful to the original: fact aware neural abstractive summarization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32 (1), 4784-4791. <https://doi.org/10.1609/aaai.v32i1.11912>
- Chauhan, S., & Keprate, A. (2022). Standards, Ethics, Legal Implications & Challenges of Artificial Intelligence. *2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 1048-1052. <https://ieeexplore.ieee.org/document/9989614>
- Ciftci, U. A., Yuksek, G., & Demir, I. (2023). My face my choice: Privacy enhancing deepfakes for social media anonymization. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 1369-1379. <https://ilkedemir.weebly.com/publications.html>
- Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, 8, 1-13 <https://doi.org/10.3389/feduc.2023.1206936>
- Díaz, B. & Nussbaum, M. (2024). Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. *Computers & Education*, 217, 1-19. <https://www.sciencedirect.com/science/article/abs/pii/S036013152400085X>
- Dirección General de Evaluación y Cooperación Territorial. Resolución de 4 de mayo de 2022, de la por la que se publica el Acuerdo de la Conferencia Sectorial de Educación, sobre la actualización del marco de referencia de la competencia digital docente. *Boletín Oficial del Estado*, núm. 116, de 16 de mayo de 2022, 67979-68026. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2022-8042
- Falke, T., Ribeiro, L., Utama, P., Dagan, I., & Gurevych, I. (2019). Ranking generated summaries by correctness: an interesting but challenging application for natural language inference. *Proceedings of the 57th annual meeting of the Association for Computational Linguistics*, 2214-2220. <https://aclanthology.org/P19-1213/>
- Felix, C. (2020). The role of the teacher and AI in education. En E. Sengupta, P. Blessinger and M. S. Makhanya (Eds.), *International perspectives on the role of Technology in Humanizing Higher Education Innovations in higher education teaching and learning* (pp. 33-48). Emerald Publishing Limited.
- Ghosal, S. (2021). ChatGPT on Characteristic Mode Analysis. *Journal of Latex Class Files*, 14, 8, 1-5.

- Giovanola, B., & Tiribelli, S. (2023). Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. *AI & Society*, 38, 2, 549–563.
- Gregory, S. (2023). Fortify the Truth: How to Defend Human Rights in an Age of Deepfakes and Generative AI. *Journal of Human Rights Practice*, 15, 3, 702–714. <https://academic.oup.com/jhrp/issue/15/3>
- Guzman, A. L. & Lewis, S. C. (2022). Artificial intelligence and communication: A Human-Machine Communication research agenda. *New Media & Society*, 22, (1), 70–86. <https://journals.sagepub.com/doi/abs/10.1177/1461444819858691?journalCode=nmsa>
- Hangloo, S., & Arora, B. (2022). Combating multimodal fake news on social media: methods, datasets, and future perspective. *Multimedia systems*, 28, 6, 2391–2422. <https://doi.org/10.1007/s00530-022-00966-y>
- Hashmi, N., & Bal, A. S. (2024). Generative AI in higher education and beyond. *Business Horizons*, 67, 5, 607–614. <https://doi.org/10.1016/j.bushor.2024.05.005>
- Henry, J.V., & Oliver, M. (2022) Who will watch the watchmen? The ethico-political arrangements of algorithmic proctoring for academic integrity. *Postdigit Sci Educ.*, 4, (2), 330–353. <https://link.springer.com/article/10.1007/s42438-021-00273-1>
- Hevner, A., & Storey, V. (2023). Research Challenges for the Design of Human-Artificial Intelligence Systems (HAIS). *ACM Transactions on Management Information Systems*, 14, (1), 1–18. <https://doi.org/10.1145/3549547>
- Ivanov, S. (2023). The dark side of artificial intelligence in higher education. *The Service Industries Journal*, 43, 1055–1082. <https://www.tandfonline.com/doi/abs/10.1080/02642069.2023.2258799>
- Kajimoto, M., & Fleming, J. (2019). News Literacy. *Oxford Research Encyclopedia of Communication*. <https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-848>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62, (1), 15–25. <https://www.sciencedirect.com/science/article/abs/pii/S0007681318301393>
- Karinshak, E., & Jin, Y. (2023). AI-driven disinformation: a framework for organizational preparation and response. *Journal of Communication Management*, 4, 539–562. <https://doi.org/10.1108/jcom-09-2022-0113>
- Klimova, B., Pikhart, M., & Kacel, J. (2023). Ethical issues of the use of AI-driven mobile apps for education. *Frontiers in Public Health*, 10, 1–8. <https://doi.org/10.3389/fpubh.2022.1118116>.
- Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., Malik, H. (2022). Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. *Applied intelligence*, 54, 3974–4026. <https://doi.org/10.1007/s10489-022-03766-z>
- Meckler, L., & Verma, P. (2022, diciembre 28) Teachers are on alert for inevitable cheating after the release of ChatGPT. *The Washington Post*. <https://www.washingtonpost.com/education/2022/12/28/chatbot-cheating-ai-chatbotgpt-teachers>
- Ministerio de Educación, Formación Profesional y Deportes de España. (2024). *Guía sobre el uso de la Inteligencia Artificial en el ámbito Educativo*. Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado. <https://intef.es/Noticias/guia-sobre-el-uso-de-la-inteligencia-artificial-en-el-ambito-educativo/>
- Mirsky, Y., & Lee, W. (2022). The creation and detection of deepfakes. *ACM computing surveys*, 54, 1, 1–41. <https://doi.org/10.1145/3425780>
- Montoro-Montarroso, A., Cantón-Correa, J., Rosso, P., Chulvi, B., Panizo-Lledot, Á., Huertas-Tato, J., Calvo-Figueras, B., Rementeria, M., & Gómez-Romero, J. (2023). Fighting disinformation with artificial intelligence: fundamentals, advances and challenges. *El Profesional de la información*. 32, 3, 1–16 <https://doi.org/10.3145/epi.2023.may.22>)
- Murtarella, G., Gregory, A., & Romentia, S. (2021). A conversation-based perspective for shaping ethical human-machine interactions: the particular challenge of chatbots. *Journal Business Research*, 129, 927–935. <https://ideas.repec.org/a/eee/jbrse/v129y2021icpp927-935.html>
- Nakamura, K., Levy, S., & Wang, W. Y. (2020). Fakeddit: a new multimodal benchmark dataset for fine-grained fake news detection. En N. Calzolari (Ed.), *Proceedings of the 12th International conference on language resources and evaluation* (pp. 6149–6157). European Language Resources Association <http://www.lrec-conf.org/proceedings/lrec2020/index.html>
- Öhman, C. (2020). Introducing the pervert's dilemma: a contribution to the critique of Deepfake Pornography. *Ethics and Information Technology*, 22, 2, 133–140. <https://link.springer.com/article/10.1007/s10676-019-09522-1>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2024). *Guía para el uso de IA generativa en educación e investigación*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000389227>
- Ortiz Muñoz, F. J. (2024). La Inteligencia Artificial como elemento disruptivo para consolidar el cambio del paradigma educativo, *Derecom*, 36, 65–85. <http://www.derecom.com/secciones/articulos-de-fondo/item/534-la-inteligencia-artificial-como-elemento-disruptivo-para-consolidar-el-cambio-del-paradigma-educativo>
- Pawelec, M. (2022). Deepfakes and democracy (theory): how synthetic audiovisual media for disinformation and hate speech threaten core democratic functions. *Digital Society*, 1, 2, 1–29. <https://link.springer.com/journal/44206/volumes-and-issues/1-2>

- Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*. UNESCO.
- Roose, K. (2022, diciembre 9). The brilliance and weirdness of ChatGPT. *The New York Times*. <https://sites.hackleyschool.org/ai/2022/12/15/hello-world/>
- Rusandi, M., A., Saripah, I., Khairun, D., & M. (2023). No worries with ChatGPT: building bridges between artificial intelligence and education with critical thinking soft skills. *Journal of public health*, 4, 3, 602-603. <https://doi.org/10.1093/pubmed/fdad049>
- Sanagustín, M. V. (2024). Los algoritmos de la des/esperanza. Una reflexión sociológica. *XXXVI Congreso Internacional AISOC: Organizaciones, cambios tecnológicos y sociedades inclusivas*.
- Saylam, S., Duman, N., Yildirim, Y., & Satsevich, K. (2023). Empowering education with AI: Addressing ethical concerns. *London Journal of Social Sciences*, 6, 39-48. <https://doi.org/10.31039/ljss.2023.6.103>
- Schiff, K. J., Schiff, D. S., & Bueno, N. (2023). *The liar's dividend: The impact of deepfakes and fake news on trust in political discourse*. <https://ideas.repec.org/p/osf/socarx/x43ph.html>
- Selwyn, N. (2016). *Is technology good for education?* Polity Press.
- Shankland, S. (2022). Why Everyone's obsessed with ChatGPT, a mind-blowing AI Chatbot. *CNET*. <https://www.cnet.com/tech/computing/why-everyones-obsessed-with-chatgpt-a-mind-blowing-ai-chatbot/>
- Shrivastava, R. (2022, diciembre 12). Teachers fear ChatGPT will make cheating easier than ever. *Forbes*. <https://www.forbes.com/sites/rashishrivastava/2022/12/12/teachers-fear-chatgpt-will-make-cheating-easier-than-ever/?sh=5b0ab2ac1eef>
- Shoaib, M., Wang, Z., Ahvanooy, M., & Zhao, J. (2023). Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models. *ArXiv*, abs/2311.17394, 1-8. <https://doi.org/10.48550/arXiv.2311.17394>
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: a data mining perspective. *ACM SIGKDD explorations newsletter*, 19, 1, 22-36. <https://doi.org/10.1145/3137597.3137600>
- Wahl-Jorgensen, K., & Carlson, M. (2021). Conjecturing fearful futures: journalistic discourses on deepfakes. *Journalism Practice*, 15, 6, 803-820. <https://experts.umn.edu/en/publications/conjecturing-fearful-futures-journalistic-discourses-on-deepfakes>
- Xu, F., Sheng, V. S., & Wang, M. (2023). A unified perspective for disinformation detection and truth discovery in social sensing: a survey. *ACM computing surveys*, 55, 1, 1-33. <https://doi.org/10.1145/3477138>
- Zawacki-Richter, O., Mariñ, V., 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, (1), 1-27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake news early detection: A Theory-driven Model. *Digital Treats Research and Practice*, 1, 2, 1-25. <https://doi.org/10.1145/3377478>

