

Inteligencia Artificial en Educación Superior: Oportunidades y Riesgos

Artificial Intelligence in Higher Education: Opportunities and Risks

Oscar Cordon García 

Dpto. de Ciencias de la Computación e Inteligencia Artificial, e Instituto Interuniversitario en
Ciencia de Datos e Inteligencia Computacional (DaSCI)
Universidad de Granada (España)
ocordon@ugr.es

Recibido: 03/11/2023

Aceptado: 10/11/2023

Publicado: 1/12/2023

RESUMEN

En el momento actual, la Inteligencia Artificial (IA) está completamente integrada en nuestra sociedad y, gracias a los recientes desarrollos de la IA generativa, ha irrumpido en la Educación a un ritmo vertiginoso. El objetivo de este trabajo es reflexionar acerca de los beneficios que la IA puede aportar en esta área dentro de las Universidades Españolas, así como sobre los riesgos que puede conllevar. Para ello, se realizará un posicionamiento inicial del marco de trabajo actual; se enumerarán algunas de las posibles aplicaciones de la IA en Educación, desde un punto de vista realista y metodológico; se recopilarán algunas herramientas tecnológicas y casos de uso; y se plantearán una serie de claves para una implantación efectiva en el Sistema Universitario Español, incidiendo en los retos y en los riesgos a nivel tecnológico, pedagógico, organizativo, legislativo y ético.

PALABRAS CLAVE

Inteligencia Artificial; Enseñanza Superior; Nuevas Tecnologías; Innovación Pedagógica; Tecnología de la Educación.

ABSTRACT

At present, Artificial Intelligence (AI) is fully integrated in our society and, thanks to recent developments in generative AI, it has burst into Education at a dizzying pace. The aim of this paper is to reflect on the benefits that AI can bring in this area within Spanish Universities, as well as on the risks that it may entail. To this end, an initial positioning of the current framework will be made; some of the possible applications of AI in Education will be reviewed, from a realistic and methodological point of view; some technological tools and use cases will be compiled; and a series of keys for an effective implementation in the Spanish University System will be proposed, focusing on the challenges and risks at the technological, pedagogical, organizational, legislative, and ethical levels.

KEYWORDS

Artificial Intelligence; Higher Education; Information and Communication Technologies; Pedagogical Innovation; Educational Technologies.

CITA RECOMENDADA:

Cordón, O. (2023). Inteligencia Artificial en Educación Superior: Oportunidades y Riesgos. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 15, 16-27. <https://doi.org/10.6018/rriite.591581>

Principales aportaciones del artículo y futuras líneas de investigación:

- Se describe el marco de trabajo actual sobre el uso y la percepción de la IA en Educación Superior.
- Se muestran varias de las posibles aplicaciones de la IA en Educación, desde un punto de vista realista y metodológico.
- Se recopilan diversas herramientas tecnológicas y se revisan una variedad de casos de uso reales en distintas universidades a nivel internacional.
- Se plantean una serie de claves para una implantación efectiva en el Sistema Universitario Español, incidiendo en los retos y en los riesgos a nivel tecnológico, pedagógico, organizativo, legislativo y ético.

1. INTELIGENCIA ARTIFICIAL Y EDUCACIÓN. POSICIONAMIENTO INICIAL

“Artificial intelligence panic is spreading across the education sector. The newest wave might be among the biggest and fastest transformations to hit academia” (“El pánico a la Inteligencia Artificial (IA) se está extendiendo por el sector educativo. La última oleada podría ser una de las mayores y más rápidas transformaciones en el mundo académico”) (Mollick y Mollick, 2023). En los últimos meses estamos asistiendo a una gran cantidad de afirmaciones de esta naturaleza. El pasado refleja otros cambios similares de gran influencia, tales como la web 1.0 y 2.0, la Wikipedia, los buscadores inteligentes, etc. pero la gran diferencia es que la comunidad universitaria no los había apreciado tan rápido. En el caso actual influyen tanto la “democratización” de la IA, que permite un acceso rápido y sencillo para todos desde cualquier tipo de dispositivo, como el denominado efecto *Tik Tok*, que ha divulgado las aplicaciones de IA a una velocidad de vértigo. Las disrupciones anteriores eran conocidas de forma masiva por el estudiantado pero el profesorado las conocía en menor medida. Ahora el miedo proviene del hecho de que el profesorado también la conoce.

Durante los últimos años se habían ido realizando varias predicciones sobre la situación actual. En su trabajo publicado en el *Online Learning Journal* en 2019 (Picciano, 2019), el Profesor Anthony de la *City University of New York*, un pionero del área de *Digital Learning* y fundador del *Online Learning Consortium*, predecía que “durante la próxima década, las tecnologías digitales avanzarán en la capacidad de interactuar y ayudar en las actividades humanas. En la década de 2020 estas tecnologías serán más visibles, pero a partir de la de 2030 comenzarán a madurar, integrarse y tendrán su mayor impacto”. Por otro lado, el blog canadiense *Bakz T. Future* publicaba un tuit el 14 de diciembre de 2021, antes incluso de la publicación de *ChatGPT*, comentando que *“My prediction for 2022 will be that GPT-3 will go viral in schools and college campuses”* (“Mi predicción para 2022 es que *GPT-3* se hará viral en colegios y campus universitarios”). Más recientemente, la Profesora Mairéad Pratschke, *Chair in Digital Education* de la Universidad de Manchester, comentaba en una de sus charlas que: “Después de la COVID, muchos profesores universitarios quieren dar un paso atrás (*“back to normal”*) pero la disrupción que ha provocado la IA en el campo de la educación no va a permitirlo”.

La disrupción que estamos experimentando se ha vinculado masivamente a la súbita aparición y desarrollo de la IA generativa pero no todo en IA y Educación es *ChatGPT*. Sucede algo parecido a lo acontecido con la irrupción general de la IA, una disciplina con casi 70 años de vida que, tras pasar por diferentes etapas, ha experimentado recientemente una eclosión de aplicaciones y de sistemas inteligentes con los que convivimos, dando lugar a una IA silenciosa que nos acompaña en nuestras actividades diarias, un éxito tecnológico sin precedentes en su breve historia. Gran parte de este reconocimiento se debe a la fase actual de desarrollo, la de la IA basada en datos y aprendizaje automático (la IA del *big data*), y en concreto a la propuesta de modelos neuronales avanzados, las redes neuronales profundas que realizan el denominado *Deep learning*. Este tipo de redes han demostrado un nivel de competencia igual o superior al de los seres humanos en un número cada vez mayor de actividades complejas tales como el reconocimiento de objetos en imágenes, el diagnóstico a partir de imágenes médicas, la restauración de imágenes, la genómica y la proteómica, el reconocimiento del habla, los sistemas de debate, etc. Igualmente destacable para el tema que nos ocupa es la propuesta de un modelo concreto llamado *Transformer* en 2017, el cual, combinado con los sistemas de aprendizaje por refuerzo humano, supuso un avance fundamental en la comprensión y generación de lenguaje natural, desplazando el foco desde el campo de la imagen al del lenguaje con los grandes avances en traducción automática y generación, clasificación y resumen de textos. Estas técnicas han derivado en la propuesta de los grandes modelos del lenguaje como *GPT-3.5* y *GPT-4* de *OpenAI* (empleados en el famoso agente conversacional *ChatGPT* lanzado hace menos de un año, el 30 de noviembre de 2022) y el resto de alternativas como *Bard* de *Google* o *LLaMA* de *Meta*. Con ello, se ha abierto la puerta a la tremenda disrupción de la IA generativa, competitiva en creatividad con el ser humano por primera vez en la historia y que ha llevado al Premio Nobel de Economía Paul Krugman a preguntarse “¿cuánto de lo que hacemos los seres humanos es realmente creativo o aporta una comprensión en profundidad?” (Krugman, 2022). El momento actual abre una nueva perspectiva multimodal de la IA Generativa combinando texto, imágenes y vídeo que promete resultados aún más espectaculares.

Aunque ni mucho menos todas las tecnologías de IA en Educación se basan en modelos del lenguaje, no me cabe ninguna duda de que *ChatGPT* ha supuesto a la vez un *boom* y una hecatombe en los entornos educativos. Por utilizar una metáfora “simpática”, me recuerda a cuando llevamos a mi hija al médico en época de pandemia y lo primero que hizo la médica fue hacerle un test de COVID. “O es COVID o es otra cosa”, nos dijo. Mi percepción actual del área es la misma: “O es *ChatGPT* o es otra cosa”.

Es como si de pronto, tanto profesores como estudiantes, tuviéramos la posibilidad de tener siempre a nuestro lado, un becario listo y con un carácter perfecto, a nuestra disposición las 24 horas del día y 7 días a la semana. Este becario puede hacer los deberes y exámenes mejor que la mayoría del estudiantado. Este becario puede hacer magníficos apuntes, discursos y conferencias. Sin embargo, el becario no era perfecto, presentaba errores y requería considerar los problemas legislativos asociados (González-Espejo, 2023).

En la misma línea, esta revolución tiene una vinculación directa con las últimas tendencias pedagógicas y concretamente con el término Educación 4.0, resultante del estudio de la evolución de las tendencias educativas con una perspectiva similar a la de la revolución industrial (Industria 4.0) y a la *world wide web* (*Web 4.0*), como puede verse por ejemplo en el trabajo “*May the Fourth Be with You: Creating Education 4.0*” (Salmon, 2019) de la Profesora Gilly Salmon, actualmente CEO y Consultora Educativa en la empresa *Education Alchemists Ltd* en el Reino Unido. La Educación 4.0 comprende metodologías

educativas de última generación como el aprendizaje activo, social, colaborativo y centrado en el estudiante; la evaluación auténtica, basada en proyectos y competencias; los itinerarios de aprendizaje personalizados, adaptables y flexibles; y los espacios de aprendizaje que fomentan la creatividad y la colaboración, todo ello soportado por tecnologías inteligentes.

Desafortunadamente, los expertos sitúan a la Educación actual en un punto intermedio entre la Educación 2.0 y la 3.0 en ese continuo y las reacciones recientes contra los avances de herramientas como la IA generativa, es decir, la prohibición de los trabajos realizados con el apoyo de *GPT*, se asemejan sospechosamente a las que se produjeron con la primera revolución industrial hace más de dos siglos, romper la máquina de producción, como en los disturbios de los talleres textiles de Nottingham en 1812 (Klein, 2019). Es bien sabido que los detectores de documentos generados con *GPT* no funcionan. Además, dan falsos positivos, lo que genera frustración y provoca problemas legales que pueden afectar a nuestras instituciones educativas.

En lugar de plantearnos esta resistencia a la innovación educativa, podríamos estudiar sus posibilidades para mejorar los procesos de enseñanza-aprendizaje. Esto me trae a la memoria la vieja y querida frase, previamente vinculada a muchas tecnologías educativas en las últimas décadas, “nuevas soluciones para viejos problemas de la educación”. La IA puede constituir una nueva oportunidad para resolver viejos retos y nuevos desafíos en Educación, empleando modelos de aprendizaje personalizado que proporcionen: 1) un papel más activo del estudiantado al conocer su evolución y poder optimizar su aprendizaje; 2) la identificación del estudiantado que requiere más apoyo (sistemas de IA para monitorización); 3) la evaluación e identificación de altas competencias (modelos predictivos de IA); 4) el tratamiento de estudiantes con diversidad funcional (analítica del aprendizaje, sistemas adaptativos basados en IA); 5) nuevos modelos de tutorización (sistemas inteligentes de tutorización); 6) sistemas de recomendación y retroalimentación; y 7) la predicción del fracaso temprano y la detección de estudiantado anómalo (sistemas de aprendizaje automático y evaluación de competencias).

Además, la IA puede conllevar otras ventajas como la reducción de tiempos en tareas más rutinarias (correcciones menores, preparación de tests, etc.) para poder dedicar más tiempo al estudiante, crear nuevo conocimiento y afrontar tareas pedagógicas más complejas. Finalmente, tal como expertos como Peter Norvig, Profesor de la Universidad de Stanford y *Google Director of Research* durante 20 años, han enunciado, las tecnologías inteligentes pueden aumentar la motivación del estudiante para el aprendizaje, concepto denominado “*BYOM: Bring your own motivation*” por analogía con el bien conocido de “*BYOD: Bring your own device*”.

Por supuesto, todas estas oportunidades no están exentas de riesgos, como las competencias digitales del docente o la legislación y el tratamiento ético del dato, como discutiremos más adelante.

2. APLICACIONES DE INTELIGENCIA ARTIFICIAL EN EDUCACIÓN SUPERIOR

En cualquier caso, las aplicaciones de IA en Educación llevan desarrollándose varias décadas y no se circunscriben a las recientes tecnologías de IA generativa. Como ejemplo, el área del *Learning Analytics*, vinculada a las tecnologías de aprendizaje automático, tiene más de 30 años de antigüedad. De hecho, el informe *2023 EDUCAUSE Horizon Report, Teaching and Learning Edition* (Pelletier et al., 2023), publicado en mayo del presente año, establece una diferencia de aplicaciones entre IA predictiva e IA generativa y presenta una revisión muy completa. Sin embargo, mi visión personal no coincide con esta

diferenciación, creo modestamente que la variable de decisión no debe ser la tecnología de IA considerada sino el problema a resolver desde el punto de vista educativo, dado que la mejora del proceso de enseñanza-aprendizaje debe ser el objetivo final de la aplicación.

En la literatura especializada, podemos encontrar otras segmentaciones o taxonomías que nos ayudan a encuadrar las distintas áreas de aplicación. Me resulta especialmente útil la presentada por la Profesora Merja Alanko-Turunen de la Universidad de Ciencias Aplicadas Haaga-Helia en Finlandia, que tras realizar una revisión detallada de la literatura reciente en el marco del consorcio europeo *Ulyseus European University (Analysis of scoping reviews on the AI higher education discourse, 2017-22)* (Alanko-Turunen, 2023), las clasifica en cuatro áreas principales (Zawacki-Richter et al., 2019):

- *Perfilado y predicciones* (por ejemplo, admisiones, planificación de los cursos, detección de abandono, etc.).
- *Tutores inteligentes* (por ejemplo, materiales de los cursos, *feedback*, evaluación del aprendizaje del estudiante, etc.).
- *Apoyo a la evaluación* (por ejemplo, corrección automática, evaluación del aprendizaje del estudiante, etc.).
- *Sistemas adaptativos y personalización* (por ejemplo, diseño de los materiales de los cursos, apoyo al profesorado y al diseño de aprendizaje, etc.).

Igualmente interesante me parece la segmentación realizada por el Dr. Raul V. Rodriguez y el Prof. Hemachandran Kannan, Vice Presidente y Director del *AI Research Centre* de la *Woxsen University* en India, respectivamente, directamente relacionada con los usos de *GPT* (Rodríguez y Kannan, 2023) pero centrada en una perspectiva más pedagógica:

- *Aprendizaje personalizado*: *GPT* tiene la capacidad de generar respuestas similares a las humanas a partir de *prompts* (peticiones) textuales, que podrían usarse para crear experiencias de aprendizaje personalizadas para los estudiantes.
- *Aprendizaje activo*: *GPT* puede usarse para generar preguntas para actividades basadas en debates, lo que puede ayudar a estimular el pensamiento crítico e involucrar a los estudiantes en un diálogo significativo. Puede crear simulaciones y escenarios en los que participen los estudiantes, proporcionando una experiencia de aprendizaje más inmersiva.
- *Tutorización*: *GPT* permite generar contenido educativo, como cuestionarios, evaluaciones y módulos interactivos, que pueden integrarse en plataformas de aprendizaje (LMSs) y otras plataformas educativas. También puede usarse para crear *chatbots* y asistentes virtuales que ayuden al estudiantado brindando apoyo y orientación personalizados.

Todas estas áreas de aplicación se sustentan en una batería de herramientas de apoyo al profesorado y al estudiantado basadas en IA. Podemos mencionar los generadores automáticos de distintos tipos de contenidos como *Jasper* (<https://www.jasper.ai/tools>); los *chatbots* avanzados con conexión web como *Perplexity* (<https://www.perplexity.ai/>); las herramientas de estudio personalizadas como *Studysmarter* (<https://www.studysmarter.es/>); las herramientas de generación de contenidos para enseñanza online como *Nolej AI* (<https://nolej.io/>); los generadores de planificación de cursos y de módulos docentes basados en *GPT* como *Teachermatic* (<https://teachermatic.com/>); los diseñadores avanzados de actividades de evaluación como *Quizziz* (<https://quizizz.com/>); las herramientas para desplegar *feedback* de evaluación personalizado basado en datos como *OnTaskLearning* (<https://www.ontasklearning.org/>) y los tutores inteligentes como *Khanmigo* (<https://openai.com/customer-stories/khan-academy>) o *Codetutor.ai* (<http://codetutor.ai/>).

Personalmente, me ha impresionado el caso de *Khanmigo*, el tutor basado en *GPT-4* desarrollado por la *Khan Academy*, que asiste tanto al docente como al discente, realiza preguntas individualizadas, motiva al aprendizaje, apoya en el desarrollo de materiales docentes, monitoriza el aprendizaje y facilita el estudio de la literatura o la historia jugando el rol de personajes literarios o históricos (<https://onx.la/b2d59>).

El informe *2023 EDUCAUSE Horizon Report, Teaching and Learning Edition* (Pelletier et al., 2023) recoge una serie de casos de uso operativos que considero muy relevantes al ser directamente trasladables a cualquier universidad española. Entre ellos, podemos destacar *LearningClues*, de la Universidad de Michigan, que extrae lo explicado o presentado visualmente durante las sesiones de clase y crea automáticamente enlaces a recursos relacionados en otras plataformas educativas en el ecosistema de aprendizaje del estudiantado, o el *Empathy Chatbot* de la *Western University* que desempeña el papel del paciente para las prácticas del examen clínico objetivo estructurado (ECO) de Medicina.

3. IMPLANTACIÓN EFECTIVA EN EL SISTEMA UNIVERSITARIO ESPAÑOL

Mi bagaje pasado como responsable de enseñanza online, tecnologías educativas y universidad digital en la Universidad de Granada y en CRUE Universidades Españolas me lleva a tener un sesgo en las aproximaciones a cualquier nueva tecnología educativa. Siempre me recuerda las diferencias entre un proyecto piloto y un servicio a nivel de universidad, así como la necesidad de una planificación rigurosa.

En el hipotético caso de implantación de un servicio de apoyo al uso de IA en Docencia en una universidad española, tengo claro que, como en otras ocasiones, solo funcionaría si se gestiona desde el más alto nivel: Rector, Equipo de Gobierno y/o Vicerrector. Además, requeriría realizar una planificación estratégica a varios niveles: individuo, unidad/departamento, institución y colaboración interuniversitaria. En este aspecto, me parece muy destacado el informe recientemente publicado (en septiembre de 2023) por EDUCAUSE, el *2023 EDUCAUSE Horizon Action Plan: Generative AI, Teaching and Learning Edition* (Robert y Muscanell, 2023), que proporciona una serie de pautas muy relevantes para el diseño de dicha planificación. Igualmente, la UNESCO ha publicado dos documentos muy relevantes en 2021, “Inteligencia Artificial y Educación. Guía Para las Personas a Cargo de Formular Políticas” (Miao et al., 2021), y 2023, “Oportunidades Y Desafíos de la Era de la Inteligencia Artificial Para la Educación Superior: Una Introducción Para los Actores de La Educación Superior” (Liu et al., 2023). En cualquier caso, esta implantación debería conllevar una obsesión con la calidad de servicio (experiencia) del estudiante, aspecto que, desgraciadamente, algunas veces minimizamos en nuestras universidades.

En mi opinión, el uso de aplicaciones relacionadas con *GPT* o con cualquier otro modelo de lenguaje puede resultar más factible como servicio, pero habría que estudiar el coste económico y la dependencia intensiva de una solución externa. Aunque contaríamos con la ventaja de la existencia de múltiples proveedores, a la vista del número creciente de actores en el mercado, deberíamos considerar el riesgo importante de que la solución adoptada podría dejar de funcionar o incluso desaparecer algún día. No considero factible un desarrollo individual por parte de una universidad española, dada la ingente cantidad de trabajo que requeriría. Podría ser una iniciativa interesante de desarrollo colaborativo universitario, a nivel andaluz o nacional vía CRUE, que no estaría exenta de una gran dificultad, dada la complejidad intrínseca de este tipo de desarrollos colaborativos y la específica del campo de aplicación: altos costes de desarrollo y ejecución, “alucinaciones” del modelo, necesidad de

gestión de los filtros, contramedidas y aspectos legales, etc. No olvidemos que grandes universidades internacionales como *Stanford* han abandonado proyectos de este tipo, en concreto el proyecto de diseño del modelo *Alpaca* basado en *LLaMA* de *Meta*, por las razones mencionadas.

El servicio debería estar sustentado por un equipo especializado de PTGAS. Al igual que en el pasado con las unidades de innovación y/o de enseñanza online, es fundamental que dicho equipo sea fuertemente interdisciplinar. La Profesora Margarida Romero-Laurent Heiser, de la *Université cote d'Azur* en Francia, distingue entre cuatro categorías de expertos: de aprendizaje (pedagogos, tecnólogos educativos), IA, diseño de interfaces de usuario y científicos de la información (Romero, 2023). El Profesor Abelardo Pardo, de la *University of Adelaide* en Australia, distingue tres perfiles (diseñadores de aprendizaje, estadísticos y expertos en Tecnologías de la Información), proponiendo además un modelo estructural de *Learning Analytics* directamente implantable en una universidad (Pardo et al., 2022).

Otro aspecto importante para el diseño de este servicio es el de la formación para toda la comunidad universitaria, la llamada "*AI Literacy*", que en nuestro contexto debe dirigirse al PDI, PTGAS, estudiantado, equipos de gobierno y gestores y, por extensión, al público en general. Estos planes de formación son clave para la aceptación de las soluciones basadas en IA por parte de todos los implicados: la formación genera la confianza y la confianza lleva a la aceptación de las tecnologías disruptivas. Para conseguir una implantación exitosa, esta formación debe sustentarse en dos dimensiones distintas: 1) Enseñanza con IA; 2) Aprendizaje de la IA mediante un uso pedagógico creativo.

Esta formación entronca con otro aspecto fundamental, el cambio requerido en el modelo pedagógico. Como suele ser habitual, la tecnología por sí sola no proporciona la solución. Para incorporar la IA en nuestra docencia universitaria se requiere un diseño pedagógico adaptado al marco de la Educación 4.0. Por un lado, se debe considerar el cambio de competencias asociado (resolución de problemas del mundo real, habilidades de colaboración y trabajo en equipo, prácticas profesionales, capacidades futuras en el entorno de trabajo, habilidades de liderazgo, competencias tecnológicas, etc.) y emplear las metodologías docentes adecuadas para conseguir las (aprendizaje basado en problemas, aprendizaje social, actividades basadas en el desempeño, actividades basadas en habilidades, instrucción y la retroalimentación, alfabetización digital / formación en IA, etc.). Por otro, hay que hacer interaccionar la tecnología y la pedagogía de forma fluida buscando el objetivo final de la mejora del proceso de enseñanza-aprendizaje empleando modelos pedagógicos específicos avanzados, lo que varios expertos como Merja Alanko-Turunen denominan *AI Entangled Pedagogy* (algo así como "Pedagogía Enredada basada en la IA") (Alanko-Turunen, 2023), y dejando atrás la dicotomía Pedagogía-Tecnología (Fawns, 22).

Por supuesto, es necesario un cambio profundo en el modelo de evaluación, pasando a "evaluar el proceso y no (solo) el resultado". Estoy convencido de que prohibir el uso de *GPT* y herramientas similares, y perseguir su uso no es la solución. Es como "poner puertas al campo". Al contrario, deben incorporarse en el trabajo diario. Por ejemplo, pueden ayudar a los estudiantes a mejorar su pensamiento crítico y a articular sus ideas de mejora de forma clara y concisa, a través del empleo de ejercicios de evaluación en los que el modelo de lenguaje elabora una redacción basada en una petición (un *prompt*) y el o la estudiante la crítica y trabaja para mejorarla añadiendo nueva información, aclarando puntos, comprobando hechos y analizando el fondo de la cuestión (Mollick y Mollick, 2022).

Todas estas consideraciones deben estar definidas a priori en una “política de uso de la IA en el curso”, que incluya cláusulas sobre qué se puede y qué no se puede hacer, uso responsable de la IA (sesgos, seguridad, etc.), denuncia de malas prácticas y sus consecuencias, formación, y reconocimiento de su empleo por parte de la propia institución. Dicha política podría ser definida globalmente a nivel de la propia universidad.

Como siempre, la clave de todo es el diseño del aprendizaje, aprovechando las herramientas disponibles, combinando la Ciencia del Aprendizaje (cómo aprenden las personas) con las Tecnologías del Aprendizaje y el Conocimiento (herramientas para apoyar ese aprendizaje). Mi visión personal es que, aunque este proceso siempre requiere un gran esfuerzo por parte del docente, en este caso dicho esfuerzo será menor que en otros escenarios como en el de la enseñanza online. Esto se debe a la potencialidad de las herramientas de IA que pueden reducir la carga de trabajo apoyando la planificación didáctica, la creación de contenidos educativos, el diseño de mecanismos de evaluación, etc., llevándonos a un modelo colaborativo “*human-AI design and content co-creation*” (diseño y creación de contenidos cooperativo profesor-herramienta de IA).

4. REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL

La aplicación de la IA en la sociedad está generando una controversia legal y ética a nivel global. Concretamente, son varios los derechos que podrían verse vulnerados por las soluciones que aporta la IA en Educación, tales como la igualdad y no discriminación, la dignidad, la privacidad y la protección de datos, la propiedad intelectual, o el honor, entre otros. Estos desafíos han impulsado la definición del concepto de IA responsable, una IA que potencia valores deseados por la sociedad, tales como la equidad, la confiabilidad, la privacidad, la seguridad, la inclusión, la transparencia y la responsabilidad (Díaz-Rodríguez, et al., 2023).

En mi opinión, en este aspecto somos afortunados por vivir en un país europeo. Aunque hay movimientos para legislar la IA en muchos países, la Unión Europea lidera claramente estas iniciativas desde 2018, cuando el presidente francés Emmanuel Macron promovió la creación de un “Ecosistema global europeo en IA”. El 7 de diciembre de 2018, la Comisión presentó el “Plan Coordinado sobre la Inteligencia Artificial” (Comisión Europea, 2018) y el “Plan Coordinado sobre el Desarrollo y Uso de la Inteligencia Artificial “Made in Europe”” (Consejo de la Unión Europea, 2019). Dicho plan incide en una prioridad transversal: la IA responsable, confiable y ética, hecha en Europa, debe construirse sobre los derechos fundamentales en la Unión Europea.

Se necesita por tanto un marco legal para usar la IA de manera justa, transparente y responsable, y que aborde la naturaleza multicultural europea. Ese marco se está plasmando en el desarrollo de la Ley de Inteligencia Artificial Europea, comúnmente conocida como *AI Act* (Comisión Europea, 2021). Los riesgos que la *AI Act* busca abordar incluyen: i) sesgo y discriminación; ii) privacidad y seguridad; iii) responsabilidad y rendición de cuentas; y iv) transparencia y confianza pública. La *AI Act* establece un marco de trabajo que asigna las aplicaciones de IA a tres categorías de riesgo. En primer lugar, se prohíben las aplicaciones y los sistemas que crean un riesgo inaceptable. Estos riesgos inaceptables pueden incluir la pérdida de vidas humanas, la discriminación sistemática o la violación de derechos fundamentales, como por ejemplo una calificación social administrada por un gobierno del tipo de la aplicada en China. En segundo lugar, se categorizan las aplicaciones de alto riesgo, las cuales son algo menos críticas que las de riesgos inaceptables, pero aun así muestran un alto potencial de causar daño. Los sistemas de IA en esta categoría pueden incluir aplicaciones médicas, sistemas de transporte

autónomo, sistemas de clasificación para el acceso a servicios financieros o herramientas de escaneo del currículum vitae que clasifiquen a los solicitantes de empleo. Por último, se detallan las aplicaciones de bajo riesgo, que generalmente se utilizan en aplicaciones cotidianas que tienen un impacto limitado en la seguridad o los derechos fundamentales de las personas. Ejemplos de sistemas de IA de bajo riesgo incluyen motores de búsqueda en línea, aplicaciones de recomendación de contenido y *chatbots* de servicio al cliente. Aunque el nivel de riesgo es bajo, la *AI Act* aún impone requisitos para garantizar la transparencia y la información adecuada sobre estos sistemas.

La clasificación de riesgos en estas tres categorías es fundamental para determinar las obligaciones y requisitos específicos que se aplicarán a los proveedores de sistemas de IA en la Unión Europea. Cuanto mayor sea el riesgo, más rigurosos serán los controles y las medidas de mitigación requeridas por la regulación. Esta aproximación de escalonamiento permite adaptar la regulación a la gravedad de los riesgos asociados a los sistemas de IA y asegura que los sistemas críticos sean sometidos a una supervisión y control más estrictos.

Centrándonos en nuestro caso, la aplicación de la IA en el ámbito educativo presenta una serie de riesgos y desafíos, tales como:

- *Sesgo y discriminación*: los sesgos de los datos de entrenamiento o de los algoritmos podrían resultar en discriminación contra ciertos grupos de estudiantes, como aquellos de diferentes orígenes étnicos o socioeconómicos. La *AI Act* promueve la transparencia y la no discriminación, lo que significa que los sistemas de IA utilizados en la educación deben ser capaces de explicar sus decisiones y no perpetuar prejuicios.
- *Privacidad y seguridad de datos*: La recopilación y el procesamiento de datos personales de estudiantes plantea preocupaciones de privacidad. La *AI Act* establece reglas estrictas sobre la protección de datos, garantizando que las organizaciones educativas cumplan con los estándares de seguridad y privacidad, y que el estudiantado tenga el control sobre sus datos.
- *Acceso equitativo*: Los sistemas de IA en educación pueden generar desigualdades si no se utilizan de manera equitativa. Algunos miembros del estudiantado pueden no tener acceso a tecnología de alta calidad, lo que podría dejarlos en desventaja. La regulación busca garantizar que todo el estudiantado tenga un acceso equitativo a las oportunidades de aprendizaje, independientemente de su situación socioeconómica.
- *Calidad de la educación*: La automatización y la personalización a través de la IA pueden tener un impacto en la calidad de la educación. La *AI Act* fomenta la transparencia en la forma en que se utilizan los sistemas de IA en la enseñanza y asegura que estos sistemas se utilicen para mejorar la calidad de la formación y no para reemplazar a los docentes.
- *Evaluación*: Los sistemas de IA utilizados para evaluar al estudiantado pueden no ser precisos en algunos casos, lo que podría afectar sus oportunidades educativas futuras. La regulación exige que los sistemas utilizados para la evaluación estén sujetos a altos estándares de precisión y transparencia.

La regulación en la Unión Europea busca por tanto equilibrar la promoción de la innovación tecnológica en la educación con la protección de los derechos del estudiantado y la calidad de la enseñanza. Al abordar los riesgos mencionados, la *AI Act* persigue garantizar que la IA en el ámbito educativo se utilice de manera ética y responsable, contribuyendo positivamente a la educación y el desarrollo de dicho estudiantado.

Otro aspecto relevante es el del *copyright* de la IA generativa, que está provocando mucho debate y que puede constituir una amenaza para el uso de estos sistemas en el futuro. En este campo hay varias perspectivas enfrentadas. Podemos destacar, por ejemplo, que mientras Europa quiere proteger al autor, otros países como Japón han dictaminado recientemente que el entrenamiento de modelos de IA no tiene derechos de autor (Growcoot, 2023), aunque el tema está todavía sujeto a una discusión ardua (Fukuoka et al., 2023).

Las universidades y el resto de las instituciones educativas deben contemplar todas las cuestiones legales mencionadas en sus sistemas de IA para apoyo a la docencia. Es fundamental destacar que las aplicaciones de la IA en la Educación están etiquetadas con riesgo alto en la *AI Act*, lo que provocará que dichos sistemas sean sometidos a una supervisión y un control más estrictos.

Aunque esos procedimientos de control no están definidos aún, existiendo dos tendencias principales, *check-lists* o auditorías, su propia indefinición se convierte en una amenaza para la implantación efectiva de la IA en nuestras universidades. En mi modesta opinión, en el Sistema Universitario Español tenemos un déficit de personal especializado en legislación relacionada con las Tecnologías de la Información y la Transformación Digital y, lógicamente, con los aspectos legales de la IA que son mucho más recientes y están aún por desarrollar en muchos de los casos. Para conseguir el objetivo también habría que reforzar nuestras plantillas o subcontratar estos servicios de asesoría legal.

5. CONCLUSIONES

En la actualidad, la IA se considera la tecnología de propósito general que impulsa la cuarta revolución industrial y su impacto en la sociedad se compara con el de la electricidad a principios del siglo pasado. La IA es una herramienta fundamental para alcanzar los retos marcados en los 17 Objetivos de Desarrollo Sostenible de Naciones Unidas (Montes, et al., 2021). Pensamos que en la Educación (ODS 4) puede provocar una disrupción significativa, que ya se viene observando en los últimos tiempos.

La IA ha madurado como tecnología, lo que plantea la necesidad de establecer marcos de trabajo para una IA responsable, justa, inclusiva, confiable, segura y transparente. Esta IA debe estar al servicio de las personas para el bien social y la sostenibilidad. A nuestra sociedad le toca conocer sus limitaciones, usos e implicaciones y, a partir de este conocimiento, hacer que su evolución y desarrollo sea en nuestro beneficio. Vivimos la transición de “una época de cambio a un cambio de época” que nos conducirá a un escenario donde el ser humano deberá convivir con sistemas inteligentes que le ayudarán a tomar decisiones y que realizarán muchas de las tareas que ahora desarrollan los humanos.

En el campo de la Educación, podríamos incluso evolucionar hacia a un nuevo híbrido: “humano + ordenador (IA)”, el que el ser humano ejerza el pensamiento crítico, la creatividad, la colaboración, el liderazgo, etc.; mientras que la IA, aun requiriendo una alfabetización y conllevando una serie de riesgos, pueda proporcionar un apoyo fundamental en un entorno de Enseñanza (diseño del aprendizaje con herramientas de IA) y Evaluación (basada en las mismas actividades del aprendizaje) con IA. Los enfoques actuales de *copilot* en programación de ordenadores y de co-creación de contenidos docentes parecen dirigirnos de forma adecuada hacia ese camino.

6. REFERENCIAS BIBLIOGRÁFICAS

- Alanko-Turunen, M. (2023). *Towards Entangled, Sustainable Higher Education Artificial Intelligence Solutions. Webinar: AI in Education, revolutionizing learning and teaching, potential uses and implications of AI in the field of education*. Ulyseus. <https://n9.cl/j6pxi>
- Comisión Europea. (2018). *Plan Coordinado sobre la Inteligencia Artificial*. Comisión Europea. <https://n9.cl/wdzbe>
- Consejo de la Unión Europea. (2019). *Conclusiones relativas al Plan Coordinado sobre el Desarrollo y Uso de la Inteligencia Artificial «Made in Europe»*. Consejo de la Unión Europea. <https://n9.cl/wubm6r>
- Comisión Europea. (2021). *Regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*. Comisión Europea. <https://n9.cl/kgmr2>
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E. y Herrera, F. (2023). Connecting the Dots in Trustworthy Artificial Intelligence: From AI Principles, Ethics, and Key Requirements to Responsible AI Systems and Regulation. *Information Fusion*, 99. <https://doi.org/10.1016/j.inffus.2023.101896>
- Fawns, T. (2022). An Entangled Pedagogy: Looking Beyond the Pedagogy—Technology Dichotomy. *Postdigital Science and Education*, 4, 711–728. <https://doi.org/10.1007/s42438-022-00302-7>
- Fukuoka, S., Murata, T. y Mizuguchi, A. (2023). *Legal Issues in Generative AI under Japanese Law – Copyright*. *Robotics / Artificial Intelligence Newsletter*. Nishimur & Asahi. <https://onx.la/1b438>
- González-Espejo, M.J. (2023). *ChatGPT debe revolucionar, de una vez por todas, la enseñanza del Derecho*. *Hay Derecho*. <https://onx.la/b541b>
- Growcoot, M. (2023). *Japan Declares AI Training Data Fair Game and ‘Will Not Enforce Copyright’*. PetaPixel. <https://onx.la/ff8bd>
- Klein, C. (2019). The Original Luddites Raged Against the Machine of the Industrial Revolution. *History*. <https://n9.cl/9gu0b>
- Krugman, P. (2022). Does ChatGPT Mean Robots Are Coming For the Skilled Jobs? *New York Times*. <https://n9.cl/2xy5r>
- Liu, B.L., Morales, D., Roser-Chinchilla, J., Sabzalieva, E., Valentini, A., Vieira do Nascimento, D. y Yerovi, C. (2023). *Oportunidades y desafíos de la era de la inteligencia artificial para la educación superior: una introducción para los actores de la educación superior*. UNESDOC. <https://n9.cl/kqa70>
- Miao, F., Holmes, W., Huang, R. y Zhang, H. (2021). *Inteligencia Artificial y Educación. Guía Para las Personas a Cargo de Formular Políticas*. UNESCO. https://n9.cl/u-n-e-s-c-o_g_p_l_p
- Mollick, E.R. y Mollick, L. (2022). *New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments*. SSRN. <https://ssrn.com/abstract=4300783>
- Mollick, E.R. y Mollick, L. (2023). *Why All Our Classes Suddenly Became AI Classes Strategies for Teaching and Learning in a ChatGPT World*. Harvard Business Publishing. <https://n9.cl/6jppqz>
- Montes, R., Melero, F.J., Palomares, I., Alonso, S., Chiachío, J., Chiachío, M., Molina, D., Martínez-Cámara, E., Tabik, S. y Herrera, F. (2021) *Inteligencia Artificial y Tecnologías Digitales para los ODS. Real Academia de Ingeniería*. <https://n9.cl/24ypu>
- Pardo, A., Mirriahi, N., Gašević, D. y Dawson S. (2022). A Model for Learning Analytics to Support Personalization in Higher Education. En R. Sharpe, S. Bennett y T. Varga-Atkins, *Handbook of Digital Higher Education* (pp. 26—37). ElgarOnline. <https://doi.org/10.4337/9781800888494.00012>

- Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., Grajek, S., Birdwll, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R. y Zimmern, J. (2023). *EDUCAUSE Horizon Report, Teaching and Learning Edition*. EDUCAUSE. <https://n9.cl/ji3n2>
- Picciano, A. (2019). Artificial Intelligence and the Academy's Loss of Purpose. *Online Learning Journal*, 23(3). <https://doi.org/10.24059/olj.v23i3.2023>
- Robert, J. y Muscanell, N. (2023). *2023 Horizon Action Plan: Generative AI. Teaching and Learning Edition*. EDUCAUSE. <https://onx.la/56e80>
- Rodríguez, R. y Kannan, H. (2023). *Active Learning, AI Style: The Role of Agent GPT in the Classroom*. SwissCognitive. <https://n9.cl/vka1h>
- Romero, M. (2023). Creative Uses of AI in Education. Webinar: AI in Education, revolutionizing learning and teaching, potential uses and implications of AI in the field of education. Ulysseus. <https://n9.cl/j6pxi>
- Salmon, G. (2019). May the Fourth Be with you: Creating Education 4.0. *Journal of Learning for Development*, 6(2), 95—115. <https://doi.org/10.56059/jl4d.v6i2.352>
- Zawacki-Richter, O., Marín, V.I., Bond, M. y 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. <https://doi.org/10.1186/s41239-019-0171-0>

7. AGRADECIMIENTOS

Esta investigación ha contado con el apoyo de MCIN/AEI/10.13039/501100011033 y de los Fondos Europeos FEDER "Una manera de hacer Europa" en el marco del proyecto CONFIA (PID2021-122916NB-I00).

INFORMACIÓN SOBRE EL AUTOR

Óscar Cerdón García

Universidad de Granada

Catedrático en el Dpto. de Ciencias de la Computación e Inteligencia Artificial en la Universidad de Granada. Fundó su Centro de Enseñanzas Virtuales en 2001, ha sido Delegado de la Rectora (Vicerrector) para la Universidad Digital de 2015 a 2019 y Presidente del grupo de trabajo *Formación online y tecnologías educativas* (FOLTE) de la sectorial TIC de CRUE Universidades Españolas de 2016 a 2020. Fue miembro del Grupo de Expertos que desarrolló la Estrategia Española de I+D+I en IA para el Ministerio de Ciencia, Innovación y Universidades en 2018-19. Ha recibido varios premios nacionales e internacionales, destacando el Premio Nacional de Informática ARITMEL en 2014. Está incluido en el 1% de investigadores más citados en el mundo (Web of Science) y en el Top 2% de investigadores más citados en el mundo en Inteligencia Artificial (Universidad de Stanford-Elsevier).



Los textos publicados en esta revista están sujetos a una licencia de Reconocimiento 4.0 España de Creative Commons. Puede copiarlos, distribuirlos, comunicarlos públicamente y hacer obras derivadas siempre que reconozca los créditos de las obras (autoría, nombre de la revista, institución editora) de la manera especificada por los autores o por la revista. La licencia completa se puede consultar en: [Licencia Creative Commons Atribución-NoComercial-Compartir por igual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).