

Consideraciones éticas y jurídicas de la IA en Educación Superior: Desafíos y Perspectivas ¹

Consideracions ètiques i jurídiques de la IA en Educació Superior: Desafiaments i Perspectives

Ethical and Legal Considerations of Artificial Intelligence in Higher Education: Challenges and Prospects

1

Isabel Cornejo-Plaza

Directora del grupo de investigación en Neuroderechos, Neuroética, Inteligencia Artificial, Metaverso y

Comportamiento Económico (Neurometa) en Instituto de Investigaciones en Derecho (IID), Universidad Autónoma de Chile.

E-mail: isabel.cornejo@uautonoma.cl

¹ El presente artículo deriva de la actividad de investigación llevada a cabo en el ámbito de los siguientes proyectos: INTACT– 101007623 – MSCA – Horizon 20202 - Research Executive Agency – European Union; Jean Monnet Centre of Excellence “Baldus” Building the Age of a Lawful and sustainable Data-Use – 101047644 – Erasmus+ - EACEA – European Union

¹ Directora del grupo de investigación en Neuroderechos, Neuroética, Inteligencia Artificial, Metaverso y Comportamiento Económico (Neurometa) en Instituto de Investigaciones en Derecho (IID), Universidad Autónoma de Chile. Correo: isabel.cornejo@uautonoma.cl ; ORCID 0000-0001-9116-0947

¹ Catedrático Jean Monnet de la Università degli Studi di Perugia (Italia); profesor titular y presidente del Comité Académico, Mestría Derecho Judicial, Escuela Judicial del Poder Judicial del Estado de Oaxaca (México) y de INDEPAC - Instituto de Estudios Superiores en Derecho Penal (México); Investigador asociado del Consiglio Nazionale delle Ricerche (Italia), CNR-IFAC. Correo: roberto.cippitani@unipg.it ; ORCID: 0000-0002-0615-2713

Roberto Cippitani

Catedrático Jean Monnet de la Università degli Studi di Perugia (Italia); profesor titular y presidente del Comité Académico, Mestría Derecho Judicial, Escuela Judicial del Poder Judicial del Estado de Oaxaca (Méjico) y de INDEPAC - Instituto de Estudios Superiores en Derecho Penal (Méjico); Investigador asociado del Consiglio Nazionale delle Ricerche (Italia), CNR-IFAC.

E-mail: roberto.cippitani@unipg.it

2

Resumen: El uso de Inteligencia Artificial, en adelante IA se ha extendido en la comunidad académica y estudiantil. ChatGPT es una herramienta de modelo de lenguaje generativo que presenta oportunidades y desafíos más allá de la regulación de derecho objetivo. Los principios éticos y un derecho prospectivo deben combinarse con una adecuada educación de la comunidad educativa a fin de utilizar la IA como herramienta en la generación de conocimiento y no de simple información descontextualizada que puede llevar al fracaso en los objetivos de la educación superior, así como los objetivos del desarrollo social más grande de la sociedad moderna, que es una educación inclusiva y solidaria.

Palabras claves: Educación, Inteligencia Artificial, Integridad científica, Libertad de Investigación, chat GPT

Abstract: The use of Artificial Intelligence, hereafter AI, has become widespread in the academic and student community. ChatGPT is a generative language modelling tool that presents opportunities and challenges beyond regulation. Ethical principles and a prospective law must be combined with a proper education of the educational community in order to use AI as a tool in knowledge generation and not mere decontextualised information that can lead to failure in the goals of higher education as well as the goals of the larger social development of modern society, which is an inclusive and solidarity-based education.

Keywords: Education, Artificial Intelligence, Scientific Integrity, Freedom of Research, GPT Chat

Resum: L'ús d'intel·ligència artificial, d'ara endavant IA s'ha estès a la comunitat acadèmica i estudiantil. ChatGPT és una eina de model de llenguatge generatiu que presenta oportunitats i desafiaments més enllà de la regulació de dret objectiu. Els principis ètics i un dret prospectiu s'han de combinar amb una educació adequada de la comunitat educativa per tal d'utilitzar la IA com a eina en la generació de coneixement i no de simple informació descontextualitzada que pot portar al fracàs en els objectius de l'educació superior, així com els objectius del desenvolupament social més gran de la societat moderna, que és una educació inclusiva i solidària.

Paraules clau: Educació, Intel·ligència Artificial, Integritat científica, Llibertat de Recerca, xat GPT

Sumario. 1. Introducción. 2. Oportunidades y desafíos de la IA aplicada a la educación superior. 3. Regulación ética de la IA en el sistema de la UE. 4. Principios generales de la IA aplicados a la educación. 5. Principios específicos de la IA en la UE aplicados a la educación. 6. Objetivos de una IA que cumple con los principios éticos. 7. Conclusiones. 8. Referencias

1. Introducción

El uso de IA se ha extendido en la comunidad académica y estudiantil. Piénsese en ChatGPT que es una herramienta de modelo de lenguaje generativo lanzada por OpenAI el 30 de noviembre de 2022, que permite al público interactuar sobre una amplia gama de temas (CIPPITANI & CASTROGIOVANNI, 2023).

El problema surge en cómo gestionar el uso adecuado de la IA como herramienta para lograr tanto los objetivos de la educación superior, así como los objetivos del desarrollo social más grande de la sociedad moderna, que es una educación inclusiva y solidaria (OTERMANS, SHARMA & ADITYA, 2022).

Según el filósofo y matemático inglés Alfred N. Whitehead, la principal función de la Universidad es “mantener el vínculo entre el conocimiento y el gusto de vivir, mediante la unión del joven y el viejo” (WHITEHEAD, 1967, p.6). El ideal de la Universidad no es tanto el conocimiento como el poder. Su función es convertir el conocimiento de la persona en un poder transformador de la humanidad. En este sentido, expresa Whitehead que la educación debe llevarnos a incrementar el valor de sentirnos vivos, pero de una manera satisfactoria (WHITEHEAD, 1985, p.48). Nótese que el filósofo matemático pensó la educación como el verdadero camino hacia el *enhancement*, o bien mejora u optimización, el cual en la actualidad se pretende emular de manera sintética a través de diversos dispositivos neurotecnológico, incluyendo los fármacos, aunque nuestras investigaciones nos han llevado a la conclusión que la educación es todavía el más eficaz optimizador humano (CORNEJO-PLAZA & SARACINI, 2023).

Whitehead, se preocupó de los fines de la educación en su Inglaterra natal a principios del siglo XX, intentando “armonizar conceptualmente nuestra comprensión de la vida, la realidad y la cultura con los conocimientos científicos de Darwin y la física moderna” (MEYER, 2017, p.12). Señaló que una de las fallas de la educación de su época era la transmisión del conocimiento, la

imaginación, la creatividad y el pensamiento crítico el cual no debe confundirse con información y datos (WHITEHEAD, 1959).

T. S. Elliot premio nobel de literatura de 1948 señaló: “Where is the knowledge we have lost in information”, dónde está el conocimiento que se ha perdido en la información, tal como reza la poesía: “El primer coro de la roca” bellamente traducido por Jorge Luis Borges (ORTIZ, 2014). El contexto da significado a la información transformándola en conocimiento. Esto último envuelve una retórica, una audiencia, una práctica y una finalidad social (LOLAS, 1994).

A la luz de lo anterior, se presenta el riesgo de que la IA, necesita, se alimenta y genera el *big data*, es decir gran cantidad de datos e información descontextualizada desde el punto de vista humano. El desafío desde el punto de vista educativo es que por ahora esta forma generativa de IA (que parece generar innovación) pero no es tal porque esa información ya está dispersa en internet, y el algoritmo de la IA lo que hace es una predicción de una información en forma de datos desagregada o inconexa con otra (CORNEJO-PLAZA & GUIÑAZÚ, 2023).

El uso irresponsable de estas herramientas tecnológicas sin alfabetizar a la comunidad universitaria corre el riesgo de abrumar a la comunidad educativa con información no conducente a alcanzar un conocimiento, es decir, a la integración significativa de la información. Este no es un problema nuevo, ya lo advirtió Alfred N. Whitehead y el poeta T.S. Elliot, en los años 50 del siglo pasado. Lo que hoy llamamos IA se reducía a la regulación y control de procesos y mecanismos denominados cibernetica, una palabra que viene del griego *kybernetos* que significa timonel, que da cuenta del problema de la automatización de los procesos cognoscitivos. Norbert Wiener creador de la cibernetica, escribió el libro “God and Golem” donde señala que la cibernetica es la forma de IA de su época (WIENER, 1961).

En esa época Wiener tenía miedo de que las computadoras tan inteligentes, a quienes se les daba una instrucción sencilla, ganaran la guerra, pero a riesgo de que se arrasara con los enemigos y con los que daban las instrucciones.

Hoy en día existe aquel y otros riesgos a una escala mucho mayor frente a los algoritmos, lo cual trae aparejados problemas de protección de la privacidad y otros derechos fundamentales, así como determinación de la responsabilidad en el diseño, mal uso de estas máquinas, incluyendo la falta de educación a la población. El control de la IA ya salió de las manos de los científicos, radicándose en los grandes gigantes tecnológicos, donde existe una pugna por aplicar las reglas del mundo analógico a los entornos digitales (CORNEJO-PLAZA, 2023).

La educación superior abarca docencia, investigación científica posee tres dimensiones: invención conceptual, innovación de prácticas y transformación social (LOLAS, 2006). No siempre los que inventan los conceptos son los que los aplican, por ejemplo, Alexander Fleming descubrió la penicilina, pero no logró apreciar cómo su invención se aplicó en la II Guerra Mundial para salvar ciento de vidas humanas.

La invención conceptual es lo más importante en la ciencia. Hay entonces un dilema, porque es posible que la IA generativa cree conceptos, más no tendrá criterios para aplicarlos, es por esta razón que el humano debe educarse, supervisar y tomar el control de la IA.

Paralelamente es importante una ética y un derecho prospectivo que reflexiones contrafácticamente para idear escenarios posibles frente a los desarrollos de estas tecnologías (CORNEJO-PLAZA, 2021).

La disciplina de la UE en este ámbito está representada no solo por un Reglamento sobre la Inteligencia Artificial (en adelante “Ley de IA”) en fase de aprobación por el legislador supranacional, coexiste con otras fuentes normativas y documentos de reflexión que apoyan la interpretación².

En lo que sigue se analizarán las oportunidades y desafíos de la IA en la educación, así como las regulaciones éticas de la IA en este ámbito, dilucidando tanto los principios generales como específicos de la IA aplicados a la educación, para finalmente cerrar a manera de colofón con los objetivos de una IA conforme con los principios éticos que deben estar presentes en una regulación de IA que funcione como instrumento para una educación solidaria y una finalidad social.

2. Oportunidades y desafíos de la IA aplicada a la educación superior

Las aplicaciones de IA, como ChatGPT, pueden ser muy útiles para apoyar la actividad educativa, sea para los estudiantes como para profesores (DIJKSTRA et al. 2022;

² Resolución del Parlamento Europeo del 16 de febrero de 2017 con recomendaciones para la Comisión sobre “normas de Derecho civil sobre robótica”; la Comunicación de la Comisión Europea del 25 de abril de 2018 sobre “Inteligencia Artificial para Europa” (COM(2018) 237 final); la Comunicación de la Comisión Bruselas “Generar confianza en la inteligencia artificial centrada en el ser humano” del 8 de abril de 2019, COM(2019) 168 final y, más recientemente, en el Libro Blanco sobre Inteligencia Artificial del 19 de febrero de 2020 (COM(2020) 65 final). En el marco del Consejo de Europa, el Comité de Expertos en Intermediarios de Internet publicó el estudio Algoritmos y Derechos Humanos en marzo de 2018 y poco después, el 3 de diciembre de 2018, la Comisión Europea para la Eficiencia de los Sistemas de Justicia (CEPEJ) aprobó la “Carta ética sobre la utilización de la inteligencia artificial en los sistemas de justicia y su entorno”.

GABAJIWALA et al. 2022; KASNECI et. al. 2023). Esta herramienta debe estar sujeta a alfabetización de la comunidad estudiantil, la cual se debe acompañar de una enseñanza en valores de la ética en investigación, libertad de investigación e integridad científica (CORNEJO-PLAZA, 2023).

Es por eso que, en las fuentes de la Unión Europea, se pone en evidencia la importancia de la IA para la educación y para la sociedad en su conjunto³.

De hecho para el Parlamento europeo hay que “(...) aprovechar, promover y aumentar el valor añadido de la inteligencia artificial y (...) garantizar que las tecnologías de inteligencia artificial en la Unión sean seguras y contribuyan al bienestar y al interés general de sus ciudadanos, ya que pueden ayudar enormemente a la consecución del objetivo común de mejorar las vidas de los ciudadanos e impulsar la prosperidad dentro de la Unión, contribuyendo al desarrollo de mejores estrategias e innovación en diversos ámbitos y sectores”.

Las Instituciones de la UE opinan que “el uso de la IA en los sistemas educativos aporta un amplio abanico de posibilidades, oportunidades y herramientas para hacerla más innovadora, inclusiva, eficiente y cada vez más eficaz, introduciendo nuevos métodos de aprendizaje de alta calidad que son rápidos, personalizados y centrados en el estudiante”⁴.

La IA permite de “lograr una educación lo más individualizada posible, ofertando al alumno itinerarios académicos personalizados según sus fortalezas y debilidades y ofreciéndole materiales didácticos ajustados a sus características, mientras se mantiene la calidad educativa y el principio integrador de nuestros sistemas educativos”⁵.

Como destaca el documento Plan coordinado sobre la IA de la Comisión Europea del 2018 (véase el Anexo, párr. G) este escenario resulta estratégico crear competencias en el sector de la IA⁶. En efecto “para aprovechar plenamente las posibilidades que ofrece la inteligencia artificial

³ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), párr. “G”.

⁴ Parlamento europeo, Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual, 2021, párr. 31.

⁵ Ivi, párr. 33.

⁶ Comisión Europea, Comunicación, Plan coordinado sobre la Inteligencia Artificial, COM(2018) 795 final, del 7 de diciembre de 2018.

y hacer que los usuarios sean conscientes de las ventajas y los retos que entraña esta tecnología, es necesario incluir la IA o la alfabetización digital en la educación y la formación”⁷.

Sin embargo, como subraya el Libro Blanco sobre la Inteligencia Artificial, la IA, así como sucede con toda nueva tecnología, su uso presenta tanto oportunidades como riesgos materiales (para la seguridad y la salud de las personas, con consecuencias como la muerte, y menoscabos al patrimonio) como inmateriales (pérdida de privacidad, limitaciones del derecho de libertad de expresión, dignidad humana, discriminación en el acceso al empleo, etc.) y pueden estar vinculados a una gran variedad de riesgos.” (vid. párr. 5).

3. Regulación ética de la IA en el sistema de la UE

Para evitar los riesgos de la IA y aumentar sus beneficios, según las fuentes jurídicas europeas, la inteligencia artificial debe ser lícita, es decir cumplir las leyes y reglamentos aplicables⁸, y también posee el mandato de ser “ética”, en el sentido de que no debe afectar intereses y valores protegidos por el ordenamiento jurídico⁹ y por tanto ser respetuosa de un “marco ético” como lo llama el Parlamento Europeo¹⁰, frente a las nuevas cuestiones de naturales jurídica y ética que esta tecnología plantea¹¹.

La dimensión ética “no es un lujo ni algo accesorio: ha de ser parte integrante del desarrollo de la IA”¹².

Como en otros casos de reglamentación de la ciencia y de la tecnología, el Derecho de la UE debe “reforzar la protección de los derechos fundamentales a tenor de la evolución de la

⁷ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), párr. G.

⁸ Véase la Comunicación de la Comisión “Generar confianza en la inteligencia artificial centrada en el ser humano”, *passim*.

⁹ Comisión europea, Comunicación sobre la inteligencia artificial para Europa, COM (2018) 237 final, p. 1.

¹⁰ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, cit., vid. especialmente los párr.s 96, 110 y 115.

¹¹ Comisión europea, Generar confianza en la inteligencia artificial centrada en el ser humano, cit., párr. 1.

¹² Ivi, párr. 3. Sobre las cuestiones éticas de la IA, vid. por ejemplo FLORIDI, 2023.

sociedad, del progreso social y de los avances científicos y tecnológicos” (véase el preámbulo de la Carta de los Derechos fundamentales de la Unión Europea).

El carácter ético de la IA tiene una conexión directa con la educación, en cuanto los usuarios, los trabajadores y los profesionales de la IA tienen que ser capacitados de manera continua en las cuestiones éticas y se necesita de expertos que se dedican a este tema¹³. Para el Parlamento europeo la cuestión es particularmente importante, visto que propone la adopción de un sistema de educación sobre la IA a todos los niveles (escuela, universidad, capacitación profesional) y certificada por la UE¹⁴.

Todo lo anterior “implica que las aplicaciones de IA no solo deben ajustarse a la ley, sino también respetar unos principios éticos”¹⁵.

Esos principios son característicos de un sistema ético-jurídico específico¹⁶. Así, el sistema europeo, y todos aquellos que comparten una cultura jurídica análoga, puede ser más o menos diferente de otros sistemas como el de EE. UU.¹⁷ o (de manera más evidente) el de los países asiáticos (cfr. la comparación entre Europa, China e India en LADIKAS et al., 2015).

4. Principios generales de la IA aplicados a la educación

Se pueden aplicar a la IA en el ámbito de la educación algunos principios generales del sistema jurídico europeo.

a) Autodeterminación: Las personas involucradas en cualquier tipo de actividad (comercial, científica, sanitaria, educacional

¹³ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, cit., párr. “G”.

¹⁴ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, cit., párr. 118 y 119.

¹⁵ Comisión europea, Generar confianza en la inteligencia artificial centrada en el ser humano, cit., párr. 1.

¹⁶ Sobre la función de los principios generales en la interpretación de un sistema jurídico, vid. CIPPITANI, 2015, p. 165 sigs.

¹⁷ Véase los principios de la ética médica elaborados por la National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research (1978). <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>, y elaborados por BEAUCHAMP & CHILDRESS, 1979.

etc.) tienen el derecho de expresar su consentimiento libre e informado en la aplicación del principio de autodeterminación. Incluso, cabe destacar que en las constituciones nacionales (véase por ejemplo el artículo 32 de la Constitución Italiana sobre el derecho a la salud), muchas fuentes europeas establecen el derecho al consentimiento como condición para desempeñar actividades concernientes a las personas: el tema está reglado por el artículo 3º párrafo 2 de la Carta de la UE; en la disciplina europea de la protección de los datos personales (véase el artículo 8 de la Carta de los Derechos Fundamentales de la Unión Europea y el Reglamento (UE) 2016/679); en el Convenio de Oviedo, que establece la reglamentación del consentimiento informado en las actividades en ámbito biomédico (véase especialmente el artículo 5)¹⁸.

Las fuentes europeas concernientes la IA no se refieren a la necesidad del consentimiento cuando por ejemplo establecen la obligación del respeto de los derechos fundamentales de las personas incluso la protección de los datos personales.

La preocupación que el uso de la IA sea basado en el consentimiento está expresado por los documentos europeos, así como por el Informe del Parlamento europeo sobre la IA en los sectores educativo, cultural y audiovisual¹⁹ donde se pide que “debe facilitarse a los usuarios una descripción que permita una comprensión general y adecuada de las funciones de que se trate, en particular sobre los datos utilizados, la finalidad del algoritmo y la personalización y sus resultados, siguiendo los principios de explicabilidad y equidad” donde se afirma que hay que desarrollar mecanismos de supervisión de los derechos de los consumidores, así como el consentimiento informado y a la libertad de elección.

b) Dignidad: En el derecho europeo la autodeterminación de la persona está sujeto a otro principio fundamental de carácter ético-jurídico: la dignidad.

La dignidad es un valor central en el marco ético-jurídico de la UE (vid. artículo 1 Carta de los Derechos Fundamentales de la UE; artículo 6 del Tratado UE), en particular en el ámbito de la ciencia y de la tecnología (véase el artículo 3 de la Carta sobre la investigación biomédica) y

¹⁸ El derecho al consentimiento importa también al derecho de rechazar el consentimiento en cualquier momento (véase la jurisprudencia TEDH, Evans c. Reino Unido del 10 de abril de 2007).

¹⁹ Parlamento europeo, Comisión de Cultura y Educación, INFORME sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual, 19 de abril 2021 - (2020/2017(INI)).

especialmente en materia de IA (véase por ejemplo el Libro Blanco sobre la inteligencia artificial; véase también el preámbulo de la propuesta de Reglamento).

Las aplicaciones de inteligencia artificial deben estar enfocadas en la persona y “empoderar a los ciudadanos a respetar sus derechos fundamentales. Su objetivo debe ser mejorar las capacidades de las personas, no sustituirlas, y permitir también el acceso de las personas con discapacidad”²⁰.

La IA, por lo tanto, debe ser pensada de manera de no afectar la dignidad humana²¹.

c) Solidaridad: Los sistemas constitucionales se basan en la protección de los derechos fundamentales en el contexto del deber de solidaridad (CIPPITANI, 2010), (véase, por ejemplo, el artículo 2 de la Constitución italiana de 1948; el artículo 10, párr. 2, de la Constitución española; la Constitución francesa de 1946, mencionada en el Preámbulo de la Constitución de la V República).

La solidaridad en su significado moderno (PECES-BARBA MARTINEZ G. et al., 2001) puede definirse como el deber de proteger los intereses de otras personas, en particular, de las personas vulnerables.

Dicho principio es aplicable en todas las relaciones sociales y jurídicas, tanto verticales (entre autoridades y ciudadanos) como horizontales (entre individuos, por ejemplo, relaciones familiares y contractuales), incluso con respecto a la propiedad.

La solidaridad también ha sido el principio subyacente al proceso de integración europea desde la Declaración de Schuman de 9 de mayo de 1950 (BOSCO, 1992) y hoy es uno de los principios fundadores de la Unión Europea (véase el artículo 2 del Tratado de la Unión Europea y la Carta de los Derechos Fundamentales) (CIPPITANI, 2022).

Desde el punto de vista de la ética de la ciencia y de la tecnología, el principio de solidaridad significa que se debe tener en cuenta el respeto a la dignidad de las personas, a sus derechos fundamentales, su bienestar y seguridad.

En los documentos de la UE se afirma que las aplicaciones de AI deben “velar por que los intereses de todos los ciudadanos, incluidos los que están marginados o en situación de

²⁰ Comisión Europea, Comunicación, Generar confianza en la inteligencia artificial centrada en el ser humano, cit., párr. 2.

²¹ Comisión Europea, Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza, cit., párr. 5.

vulnerabilidad, como las personas con discapacidad, sean tenidos en cuenta y estén representados debidamente”²².

El Parlamento europeo destaca la necesidad de utilizar la AI protegiendo a niños, niñas y adolescentes cuando se utiliza la AI en la educación y más en general opina que “los profesores deben controlar y supervisar cualquier implantación y uso de tecnologías de IA en escuelas y universidades, cuando interactúen con alumnos y estudiantes”²³, evidentemente con una función protectora.

d) Prevención y precaución: El principio de precaución es la base de un proceso de toma de decisiones de evaluación del riesgo. Este proceso tiene que identificar el grado de incertidumbre que conlleva la evaluación de la información científica y, a continuación, es necesario establecer el nivel de riesgo “aceptable” para la sociedad que, de todas maneras, no puede afectar a la dignidad y a la integridad de las personas. El riesgo debe estar presente (no es meramente hipotético) y no se puede descartar, aunque no sea seguro.

En el caso específico de la IA, las fuentes de la UE tienen en cuenta que el uso de la inteligencia artificial puede provocar riesgos desde varios puntos de vista “que pueden estar vinculados a ciber amenazas, a la seguridad personal (por ejemplo, con relación a nuevos usos de la IA, como en el caso de los aparatos domésticos) (Véase el Libro Blanco sobre la IA, párr. 5).

Según la antemencionada comunicación “Generar confianza en la inteligencia artificial centrada en el ser humano”, en el ámbito de este sistema de confianza, los actores (desarrolladores, proveedores y usuarios) deberían seguir algunas directivas entre las cuales la solidez técnica y la seguridad, que incluye la capacidad de resistencia a los ataques, un plan de repliegue y la seguridad general, precisión, fiabilidad y reproducibilidad.

La Ley sobre la IA tiene en consideración el riesgo potencial de la IA y regla sobre todo el caso de las obligaciones de los proveedores y de los sistemas de “alto riesgo” identificados en el Anexo III de la Propuesta de Reglamento, como los que se refieren a la identificación biométrica

²² Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), párr. 38.

²³ “Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual” del Parlamento europeo del 2021, párr. 43.

y categorización de personas; gestión y funcionamiento de infraestructuras esenciales; empleo, gestión de los trabajadores y acceso al autoempleo; acceso y disfrute de servicios públicos y privados esenciales y sus beneficios.

Especialmente se consideran de alto riesgo los sistemas para la educación y formación profesional.

Como se afirma en el “considerando” no. 35 de la Ley sobre la IA:

“Deben considerarse de alto riesgo los sistemas de IA que se utilizan en la educación o la formación profesional, y en especial aquellos que determinan el acceso o distribuyen a las personas entre distintas instituciones educativas y de formación profesional o aquellos que evalúan a las personas a partir de pruebas realizadas en el marco de su educación o como condición necesaria para acceder a ella, ya que pueden decidir la trayectoria formativa y profesional de una persona y, en consecuencia, afectar su capacidad para asegurar su subsistencia”. De hecho “Cuando no se diseñan y utilizan correctamente, estos sistemas pueden violar el derecho a la educación y la formación, y el derecho a no sufrir discriminación, además de perpetuar patrones históricos de discriminación”.

e) Proporcionalidad: Otro principio general del Derecho de la UE y que aplica también a la IA es la “proporcionalidad” por el cual la tecnociencia debe llevarse a cabo a través de medios no excesivos respecto a los fines que se deben alcanzar. Los medios y los fines deben ser legítimos y deben evitar riesgos que pueden afectar los intereses protegidos por el ordenamiento jurídico.

Los documentos europeos, por ejemplo, afirman que el uso de sistemas de inteligencia artificial para luchar contra la delincuencia es admisible solo si son necesarios y proporcionales²⁴.

En particular, en el ámbito educativo la aplicación de la AI debería ser proporcional, en el sentido que no debería usarse de manera excesiva o inútil, cancelando otras metodologías de educación o hasta (como se afirma de manera quizás exagerada) remplazando los profesores.

²⁴ Véase el Libro Blanco sobre la Inteligencia artificial del 2020, cit., párr. 6

El Parlamento europeo afirma que los sistemas de IA “no deben sustituir a las relaciones educativas que impliquen a profesores, y (...) no se debe prescindir de las formas tradicionales de educación”²⁵.

De hecho “no pueden emplearse las tecnologías de IA en detrimento o a expensas de la educación presencial, ya que los profesores no deben ser sustituidos por la IA o por tecnologías conexas”²⁶.

5. Principios específicos de la IA en la UE aplicados a la educación

Asimismo, el marco normativo que está diseñando la Unión Europea identifica principios específicos que pueden derivar de las aplicaciones de IA, como en la educación.

i) No exclusividad y no automaticidad. Para las fuentes de la UE, una decisión no debe basarse exclusivamente y de manera automática sobre el resultado de un dato técnico o científico, sino pasar por el filtro del decisor (humano).

Este principio está expresado por el artículo 22 del Reglamento (UE) que en su primer párrafo establece: “Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”.

La necesidad de una vigilancia humana está prevista en la propuesta de Reglamento sobre la IA (Véase el “considerando” 48 y el artículo 14). En la Comunicación de 2019 “Generar cofinancia...” (párr. 2.2.i) se destaca como “La supervisión humana ayuda a garantizar que un sistema de IA no socave la autonomía humana ni cause otros efectos adversos”. La supervisión debe realizarse a través de mecanismos de gobernanza, tales como el enfoque de la participación humana (*human-in-the-loop*), la supervisión humana (*human-on-the-loop*), o el control humano (*human-in-command*).

En el caso de la IA aplicada a la educación, el principio del cual se está hablando puede ser útil en el caso de la evaluación del resultado de la educación por parte de los docentes en examen

²⁵ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), párr. 118.

²⁶ “Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual” del Parlamento europeo del 2021, párr. 43.

de materias, de grado, de doctorado, etc. Incluso en este caso no podría confiar en los algoritmos para establecer la evaluación sin la participación de los profesores y de los estudiantes.

Lo confirma el Parlamento europeo hablando de AI y educación, para el cual “es necesario tener en cuenta sus aspectos tecnológicos, reglamentarios y sociales, con unas garantías adecuadas y un enfoque centrado en el ser humano que garantice en última instancia que los seres humanos puedan controlar y corregir en todo momento las decisiones del sistema”. En particular “los sistemas de IA no deben tomar ninguna decisión final que podría afectar a las oportunidades educativas, como la evaluación final de los alumnos, sin una completa supervisión humana” así que “las decisiones automatizadas sobre personas físicas basadas en la elaboración de perfiles que tengan repercusiones jurídicas o similares deben estar estrictamente limitadas y siempre exigir el derecho a la intervención humana y a una explicación de conformidad con el GDPR”²⁷.

ii) Transparencia. La legislación vigente requiere que las personas tengan el derecho de conocer cuáles son las razones que fundamentan las decisiones de las autoridades. Este derecho a la transparencia y a la información completa está previsto en muchas normas como los artículos 13, 14 y 15 del GDPR y del artículo 41 de la Carta Europea de Derechos Fundamentales de la UE. La Comunicación del 2019 sobre “Generar confianza en la inteligencia artificial centrada en el ser humano” (véase párr. 2.2.iv) prevé que “es importante registrar y documentar tanto las decisiones tomadas por los sistemas como la totalidad del proceso (incluida una descripción de la recogida y el etiquetado de datos, y una descripción del algoritmo utilizado) que dio lugar a las decisiones”. Hay que aplicar medidas para la explicabilidad del proceso de toma de decisiones algorítmicas.

Por otra parte, en la Resolución del Parlamento sobre la robótica se afirma que la “transparencia y la inteligibilidad de los procesos decisorios” se deben considerar como los principales principios éticos aplicables. De hecho “siempre ha de ser posible justificar cualquier decisión que se haya adoptado con ayuda de la inteligencia artificial y que pueda tener un impacto significativo sobre la vida de una o varias personas (...); siempre debe ser posible reducir los cálculos del sistema de inteligencia artificial a una forma comprensible para los humanos; (...) los robots avanzados deberían estar equipados con una ‘caja negra’ que registre los datos de

²⁷ Parlamento Europeo, “Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual”, cit., párr. 43.

todas las operaciones efectuadas por la máquina, incluidos, en su caso, los pasos lógicos que han conducido a la formulación de sus decisiones” (párr. 12).

iii) Responsabilidad. En coherencia con el principio de no exclusividad, las decisiones que se basan en datos técnico-científicos deben ser atribuibles a sujetos jurídicos y especialmente a personas. Esos sujetos tendrán la responsabilidad (civil, administrativa y penal) en caso de incumplimiento de las reglas y principios²⁸.

En la Comunicación “Generar confianza en la inteligencia artificial centrada en el ser humano” (véase párr. 2.2.vii) donde a este respecto se afirma: “Deben instaurarse mecanismos que garanticen la responsabilidad y la rendición de cuentas de los sistemas de IA y de sus resultados, tanto antes como después de su implementación. La posibilidad de auditar los sistemas de IA es fundamental, puesto que la evaluación de los sistemas de IA por parte de auditores internos y externos, y la disponibilidad de los informes de evaluación, contribuye en gran medida a la fiabilidad de la tecnología”.

La Resolución de 2017 del Parlamento concerniente la robótica (véase párr. 49-59) exige las personas deben ser capaces de minimizar los riesgos y gestionar el impacto negativo (párr. 55); además se afirma la necesidad de un esquema de seguros obligatorios (párr. 57) y de un fondo en caso de ausencia del seguro (párr. 58), como ya está previsto para los vehículos.

iv) La responsabilidad social. En aplicación del principio general de solidaridad, en ámbito científico y tecnológico las fuentes internacionales y europeas plantean la “responsabilidad social” y la “rendición de cuentas” de los investigadores y de los docentes.

La función social de la ciencia y de la enseñanza implica derechos (vid. el artículo 13 de la Carta de los Derechos Fundamentales), así como de deberes.

La responsabilidad social implica un uso eficiente de los recursos y, por tanto, el cumplimiento de principios como la transparencia y la rendición de cuentas (véase el párrafo 22 de la Recomendación de la UNESCO de 1997).

²⁸ Véase ampliamente el Libro Blanco sobre la Inteligencia artificial de 2020, especialmente p. 15; véase también el Libro Blanco “Inteligencia artificial para Europa”, espec. el párr. 3.3 “Garantizar un marco ético y jurídico adecuado”; véase la Resolución del Parlamento europeo del 2017, puntos L y M.

En este contexto, las fuentes europeas afirman que la IA y las tecnologías conexas tienen que ser utilizadas de manera socialmente responsable teniendo el papel de buscar soluciones, pero promoviendo los valores fundamentales y el Estado de Derecho²⁹.

6. Objetivos de una IA que cumple con los principios éticos.

La construcción de una IA aplicada al ámbito educativo, que cumple con los principios éticos antemencionados, tiene como objetivo principal proteger los intereses fundamentales del ordenamiento jurídico.

En particular la IA debe respetar los derechos humanos y otros intereses considerados fundamentales por el ordenamiento jurídico de la UE (véase el “considerando” 28 de la propuesta de Reglamento sobre la IA; véase también la Resolución del Parlamento sobre la robótica, párr. 10).

Como afirma la Comisión europea en la exposición de los motivos de la propuesta de Reglamento sobre la IA (vid. párr. 3.5) la legislación sobre la IA debe reforzar la aplicación y la promoción de los derechos fundamentales en la UE previstos por las constituciones nacionales, la Convención de Roma del 1950, y la Carta UE: el respeto de la vida privada y familiar y la protección de datos de carácter personal (artículos 7 y 8), la no discriminación (artículo 21) y la igualdad entre hombres y mujeres (artículo 23); la libertad de expresión (artículo 11) y de reunión (artículo 12), y garantizar el derecho a la tutela judicial efectiva y a un juez imparcial, la presunción de inocencia y los derechos de la defensa (artículos 47 y 48), así como el principio general de buena administración; los derechos de los trabajadores a unas condiciones de trabajo justas y equitativas (artículo 31), un elevado nivel de protección de los consumidores (artículo 28), los derechos del niño (artículo 24) y la integración de las personas discapacitadas (artículo 26). El derecho a un nivel elevado de protección del medio ambiente y la mejora de su calidad (artículo 37). Además, los sistemas de IA no deben afectar el derecho a igualdad y por tanto deben evitar discriminaciones.

Por otra parte, la Comunicación “Generar confianza en la inteligencia artificial centrada en el ser humano” (párr. 2.2.v) destaca el peligro que los sistemas de IA sean afectados por sesgos involuntarios.

²⁹ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), párr. 38.

La protección efectiva de los derechos humanos está asociado al principio de transparencia antemencionado, especialmente en presencia de situaciones de vulnerabilidad, como en caso de la gestión de “la migración, el asilo y el control fronterizo afectan a personas que con frecuencia se encuentran en una situación especialmente vulnerable y dependen del resultado de las actuaciones de las autoridades competentes” (véase. el “considerando” 39 de la propuesta de Reglamento sobre la IA).

Por lo que se refiere al uso de la IA en la educación el tema se encuentra abordado por el antes citado “Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual” del Parlamento europeo del 2021, especialmente en el caso de las discriminaciones de género.

Desde este punto de visto, se recuerda como “las experiencias pasadas, especialmente en los ámbitos técnicos, han demostrado que los avances y las innovaciones a menudo se basan principalmente en datos masculinos y que las necesidades de las mujeres no se reflejan plenamente” (párr. “G”) y que “ las desigualdades de género, los estereotipos y la discriminación también se pueden crear y reproducir mediante el lenguaje y las imágenes difundidas a través de los medios de comunicación y las aplicaciones basadas en la IA” (párr. “G”) especialmente en la educación y en los programas culturales que tienen una influencia considerable en la formación de las creencias y los valores de las personas.

El Informe destaca también que “una de cada diez mujeres de la Unión ya ha sufrido alguna forma de ciber-violencia desde los 15 años y que el ciberacoso sigue siendo motivo de preocupación en el desarrollo de la IA, también en el ámbito de la educación”.

Finalmente, entre los derechos que hay que respetar, en el ámbito educativo los sistemas de AI tienen que ser construidos teniendo en cuenta derechos específicos como el a la educación de los alumnos y la libertad académica de los docentes (CIPPITANI, 2015; CIPPITANI, 2014).

Como se ha mencionado anteriormente, la AI por un lado debe fomentar las capacidades y la creatividad de los estudiantes y por otro lado mejorar la calidad de la enseñanza. Al contrario, en este sistema de valores, no se puede admitir un uso de la AI que sea de baja calidad y que sustituye la esencial función de los profesores.

7. Conclusiones

La IA y en especial la IA generativa como el Chat GPT entre otras, ha causado una discusión en ciernes en el ámbito educativo. Consideramos que la IA y los sistemas de lenguaje generativo son herramienta que podrían constituir una herramienta eficaz para lograr el conocimiento

siempre y cuando no se transforme en un vector de mera información y acumulación desestructurada de datos. Podría a su vez transformarse en un mecanismo para el acortamiento de las brechas sociales en educación. Por lo anterior, se hace imperativo una alfabetización transversal de estas herramientas en los entornos educacionales, a fin de entender cómo se utiliza, reforzando los esfuerzos en ética en investigación, integridad científica, así como el respeto a los derechos fundamentales en juego. Es necesario además entender que la IA generativa debe subordinarse al sistema ético y de derecho objetivo de la comunidad internacional y de la UE, el cual es el más desarrollado en este ámbito, lo cual se puede apreciar en los distintos marcos y sistemas normativos analizados en esta investigación. Finalmente consideramos que la IA podría convertirse en un instrumento eficaz para una educación más solidaria, inclusiva y eficaz conforme a un fin social de transformación e integración del conocimiento de cara a los objetivos del desarrollo sostenible.

8. Referencias

- BEAUCHAMP, T. L.; CHILDRESS J. F. (1979). *Principles of Biomedical Ethics*, Oxford University Press
- BOSCO G. (1992). *Lezioni di diritto internazionale*, Milano: Giuffrè
- CIPPITANI, R.; CASTROGIOVANNI, L. (2023). *ChatGpt a la prueba del derecho de la Unión Europea: protección de datos personales y otras cuestiones ético-jurídicas*, en Cornejo Plaza; M. I.; Isler Soto, E., (Editors), Temas actuales sobre consumo, inteligencia artificial, plataformas digitales y neuroderechos. Santiago del Chile. Rubicón Editores.
- CIPPITANI, R. (2022). *La solidaridad institucional europea frente a las situaciones de crisis*. En (Eds) Carlos Molina del Pozo, Virginia Saldaña Ortega, Hacia la construcción de un verdadero proyecto federal para la Unión Europea, A Coruña: Colex, pp. 64-87
- CIPPITANI, R. (2016). *Interpretación y Derecho de la Integración*. Buenos Aires-Bogotá-Porto Alegre: Astrea.
- CIPPITANI, R. (2015) *Academic Freedom as a Fundamental Right*, en 1st International Conference on Higher Education Advances, HEAD'15, Valencia: Universitat Politècnica de València, pp. 552-558, doi: <http://dx.doi.org/10.4995/HEAD15.2015.1522>
- CIPPITANI, R. (2014), *La libertad de cátedra y de investigación en el ámbito de la autonomía universitaria*, en Buenrostro Ceballos, A.F., La libertad de cátedra y de investigación en el

ámbito de los derechos humanos. (pp. 129-188), Mexicali: Universidad Autónoma de Baja California.

CIPPITANI, R. (2010) *La solidarietà giuridica tra pubblico e privato*, Roma-Perugia: ISEG.

Comunicación de la Comisión Europea Inteligencia Artificial para Europa, del 25 de abril de 2018, (COM (2018) 237 final).

Comunicación de la Comisión Bruselas, Generar confianza en la inteligencia artificial centrada en el ser humano del 8 de abril de 2019, COM (2019) 168 final

Comisión europea, Libro Blanco sobre Inteligencia Artificial, del 19 de febrero de 2020 (COM (2020) 65 final).

Comisión Europea, Comunicación, Plan coordinado sobre la Inteligencia Artificial, del 7 de diciembre de 2018, COM(2018) 795 final

DIJKSTRA, R., GENÇ, Z., KAYAL, S., & KAMPS, J. (2022). *Reading comprehension quiz generation using generative pre-trained transformers*. https://e.humanities.uva.nl/publications/2022/dijk_read22.pdf.

CORNEJO-PLAZA, MI & SARACINI, C. (2023). On pharmacological neuroenhancement as part of the new neurorights' pioneering legislation in Chile: a perspective. *Front. Psychol.* 14:1177720, (pp. 1-9). Doi: 10.3389/fpsyg.2023.1177720

CORNEJO-PLAZA, M. I. & ISLER SOTO, E. (2023). (Editors), *Temas actuales sobre Consumo, Inteligencia Artificial, Plataformas Digitales y Neuroderechos*, Santiago: Rubicón Editores.

CORNEJO-PLAZA, M. I. & GUINAZÚ, Flavia (2023), Límites éticos y legales (neuroderechos), a la utilización de las interfaces cerebro computadora (bcis) para aumentar las capacidades cognitivas, en (Ed) Azuaje, M. *Introducción a la ética y el derecho de la inteligencia artificial*, (pp. 101-115). Editorial La Ley.

CORNEJO-PLAZA, M. I. (2023). Consideraciones Éticas de la Neurociencia del consumidor. En (ed.) Barona, S. *Derecho del consumo y protección del consumidor sustentable en la sociedad digital del siglo XXI* (pp. 117-132). Santiago de Chile: Ed. Universidad Autónoma de Chile. DOI: <http://doi.org/10.32457/UA.167>

CORNEJO-PLAZA, M. I. (2023). *Libertad de investigación: un derecho distinto de la libertad de expresión y subordinado a la integridad científica*. En (Coods.) Martínez Gómez, J.; Cornejo Plaza, M.I. y Fernández Pichardo, G., *Bioética. Violencia y Conflictos Internacionales* (pp. 225-

- 242) México, Tirant Lo Blanch.
- CORNEJO-PLAZA, M.I. (2021). *Derecho y Neurociencias: El caso del mejoramiento neural farmacológico*, Tesis Doctoral, Facultad de Derecho, Universidad de Chile.
- FLORIDI, L. (2023), *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*, Oxford: Oxford University Press.
- GABAJIWALA, E., MEHTA, P., SINGH, R., & KOSHY, R. (2022). *Quiz maker: Automatic quiz generation from text using NLP*, en *Futuristic trends in networks and computing technologies* (pp. 523–533). Singapore: Springer.
- KASNECI, E., SESSLER, K., KÜCHEMANN, S., BANNERT, M. DEMENTIEVA, D., FISCHER, F., GASSER, U., GROH,G., GÜNNEMANN, S., HÜLLERMEIER, E., KRUSCHE, S., KUTYNIOK, G., MICHAELI, T., NERDEL, C., PFEFFER, J., POQUET, O., SAILER, M., SCHMIDT, A., SEIDEL, T., STADLER, M., WELLER, J., KUHN, J., KASNECI, G., (2023) *ChatGPT for good? On opportunities and challenges of large language models for education*, in *Learning and Individual Differences*, Volume 103, April 2023, 102274 DOI: [10.1016/j.lindif.2023.102274](https://doi.org/10.1016/j.lindif.2023.102274).
- LADIKAS, M.; CHATURVEDI, S.; ZHAO, Y.; STEMERDING, D. (2015) *Science and Technology Governance and Ethics: A Global Perspective from Europe, India and China*, Springer: Charm
- LOLAS, F. (1994). *Medical praxis: an interface between ethics, politics, and technology*, en *Social Science and Medicine* (Oxford) 39: 1-5.
- LOLAS, F. *La innovación como tarea social. El ejemplo de las tecnologías de la representación-representatividad*, en Lolas, F. (Editor) *Ética e innovación tecnológica*. (pp.191-208), Santiago del Chile: Centro Interdisciplinario de Estudios en Bioética, Universidad de Chile.
- MEYER. (2017). *Human agency in the twenty-first century: the views of P. S. Davies, R. Niebuhr, and A. N. Whitehead*, en *International Journal for Philosophy of Religion*, 82(2), 119–134. <https://doi.org/10.1007/s11153-017-9629-y>
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research (1978). <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>
- ORTIZ, R. (2014). *Las deslecturas de Borges: Eliot, el traductor argentino y la tradición*,

- en *Variaciones Borges*, 37, 37–52. <http://www.jstor.org/stable/24882826>
- OTERMANS, P. C. J., SHARMA, S., & ADITYA, D. (2022). *Challenges Faced by Students with Illiterate Parents: Insights from Teachers - An Interview Study Using Thematic Analysis*, en *Educational Quest*, 13(2), 95-110. <https://doi.org/10.30954/2230-7311.2.2022.2>
- Parlamento Europeo, Resolución del 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))
- Parlamento europeo, Comisión de Cultura y Educación, Informe sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual, 19 de abril 2021 - (2020/2017(INI))
- PECES-BARBA MARTINEZ, G et al. (2001). *Historia de los Derechos fundamentales, t. II, Siglo XVII, vol. I, El contexto social y cultural de los derechos. Los rasgos generales de evolución*, Madrid: Dikinson.
- WIENER, N. (1961). *Cybernetics or Control and Communication in the animal and the machine*, Cambridge (Mass): MIT Press.
- WHITEHEAD, A. (1967). *The Aims of Education and Other Essays*. New York: The Free Press.
- WHITEHEAD, A. (1985), *La función de la razón*, Madrid: Tecnos.