

El uso ético de la inteligencia artificial de código abierto en educación universitaria

Christian Jonathan Ángel Rueda¹
Jorge Landaverde Trejo²

¹ Universidad Autónoma de Querétaro

² Universidad Marista de Querétaro

Correspondencia: cangel@uaq.mx ieducativa@umq.maristas.edu.mx

RESUMEN

Esta propuesta, de carácter exploratorio, invita a participar en un conversatorio: la emergencia de un nuevo paradigma que sea capaz de enfrentar los retos y riesgos derivados de una avalancha de intentos por descubrir patrones subyacentes al caos provocado por los desarrollos en torno a la generación, implementación y uso de la Inteligencia Artificial de Código Abierto (IA-CA). La perspectiva adoptada en estas reflexiones es abierta a la transdisciplinariedad para incidir en el ámbito universitario. El fin no es emitir soluciones, sino vislumbrar estructuras ético-hermenéuticas que fundamenten las tomas de decisiones de los diferentes actores en la vida universitaria, lo cual debería prevalecer más allá de la influencia dominante del determinismo tecnocrático. La novedad de esta preocupación radica no en el tema, sino en el enfoque analítico desde el cual se dimensiona el contexto de Stable Diffusion y su estrategia de lanzamiento de la IA-CA. El contexto al que nos referimos se denomina “aprendizaje profundo” mediante las IA (no confundirlo con las problemáticas derivadas de la IA simbólica). Es a partir de ahí desde donde se busca descubrir bondades y riesgos de un desarrollo que, si bien nos fascina, también puede llevarnos —tanto a docentes como a estudiantes— al lado oscuro de la IA por descuidar nuestra formación en la navegación por espacios cibernéticos. Esta propuesta es para debatir en torno a agendas acordes con necesidades de la academia, la curaduría de fuentes acreditadas de información y, especialmente, la organización de nuestra sociedad desde argumentos racionales y razonables.

Palabras Clave: aprendizaje profundo
código abierto, ética- hermenéutica, inteligencia artificial

ABSTRACT

This exploratory proposal invites you to participate in a conversation: the emergence of a new paradigm capable of facing the challenges and risks derived from an avalanche of attempts to discover patterns underlying the chaos caused by developments around generation, implementation and use of Open Source Artificial Intelligence (OS-AI). The perspective adopted in these reflections is open to transdisciplinarity to influence the university environment. The aim is not to issue solutions but to envision ethical-hermeneutical structures that support the decision-making of the different actors in university life, which should prevail beyond the dominant influence of technocratic determinism. The novelty of this concern lies not in the subject, but in the analytical approach from which the context of Stable Diffusion and its OS-AI launch strategy are dimensioned. The context to which we refer is called "deep learning" through AI (not to be confused with the problems derived from symbolic AI), and it is from there that we seek to discover the benefits and risks of a development that, although it fascinates us, can also lead us -both teachers and students - to the dark side of AI for neglecting our training in navigating cyber spaces. This proposal is to debate about agendas in accordance with the needs of the academy, the curation of accredited sources of information and, above all, the organization of our society from rational and reasonable arguments.

Key words: deep learning, open source, Ethics-Hermeneutics, artificial intelligence

Artículo arbitrado

Recibido:	Aceptado:
15 de octubre de 2022	16 de noviembre de 2022

1 INTRODUCCIÓN

El sistema educativo se encuentra en el umbral del uso de la Inteligencia Artificial de código abierto (IA-CA). La aparición de las IA-CA en las aulas universitarias ha generado una serie de controversias entre los profesionales. Entre ellas se encuentra la obligatoriedad de tener licencias específicas para implementar dicha tecnología y sobre el papel que deben jugar los estudiantes y los docentes al respecto.

Inmediatamente, salta el cuestionamiento sobre la obligatoriedad de hacer uso de las licencias Open RAIL-M¹ para trabajar con IA-CA en las aulas, lo cual enlaza con otra pregunta: ¿qué se debe y no se debe hacer con la IA-CA en los espacios áulicos? Las respuestas a dichas preguntas son multifacéticas. Desde una perspectiva empresarial o gubernamental, la respuesta suele enfocarse en estrategias de sustentabilidad para financiar las nuevas tecnologías y, en el caso de las instituciones educativas, ésta se concentra en encontrar mejores formas de aprovecharlas desde el principio ético de beneficencia para los diversos procesos pedagógicos. El presente trabajo trata de señalar que el desarrollo y uso de IA-CA dentro del sistema educativo, convendría hacerse con un enfoque ético-hermenéutico que conduzca a asumir la propia responsabilidad compartida. Esta investigación tiene como objetivo aportar en la construcción de un modelo ético para el uso de la IA-CA en las aulas universitarias.

Desde un enfoque ético-hermenéutico se considera que La Inteligencia Artificial de Código Abierto con restricciones de uso conductual (Open RAIL Artificial Intelligence), propicia la construcción de un nuevo paradigma de programación. Un ejemplo bastante particular del desarrollo y uso de IA es el que se da en las aulas universitarias, en donde la libertad en general, así como la libertad de los estudiantes en

particular, se encuentran en conflicto con los intereses corporativos y/o institucionales. Pero ¿cuál es el sentido de plantear esta polémica discusión?

El desafío para la implementación –actual o futura– de IA-CA en la educación, genera serias distorsiones socio-éticas, pues involucra a toda una serie de "actores" en el ámbito de las nuevas tecnologías: desde profesionales de investigación y tecnología hasta agentes académicos del sistema educativo en sus diferentes contextos. Entonces ¿cómo adecuar los esquemas de desarrollo de IA-CA en las aulas universitarias a un sistema ético de responsabilidad compartida? En esta investigación se busca dar cuenta de cómo el desarrollo y uso de IA-CA dentro del sistema educativo, puede llevarse a cabo con un enfoque ético y con responsabilidad compartida al considerar las distintas relaciones entre la tecnología, las políticas educativas y la sociedad.

Resulta pertinente analizar el contexto de lo sucedido el 22 de agosto de 2022, cuando IA-CA de Stable Diffusion, creada por la empresa stability.ai, se liberó al uso público con la característica de adecuarse a los intereses del usuario con la flexibilidad de editar el código y desarrollar herramientas en otras áreas para las que no fueron diseñadas ni pensadas desde el inicio, sin restricciones éticas o morales. De esta manera, se abrió la posibilidad para que algunas acciones sean generadas contra el principio ético de no-maleficencia, por lo cual es importante tener una normatividad de cómo debe llevarse a cabo el uso de IA-CA en las aulas.

Es necesario considerar las dos esferas en las cuales conviene articular el modelo ético de responsabilidad compartida: primero, la esfera de la ciencia y, segundo, la de responsabilidad ética.

La factibilidad de una actitud ética que articule una praxis consciente de la responsabilidad compartida, implica abrir espacios de reflexión entre los actores de estos dos ámbitos, para así buscar posibles vías de

¹ Son licencias con restricciones de uso conductual que se pueden

aplicar a cualquier modelo de IA (licenses.ai, 2022).

negociación entre el campo académico, empresarial y gubernamental en torno a las formas de planeación y control del desarrollo de las propuestas tecnocientíficas como la IA-CA, entre otras.

Esta investigación pugna por una intervención, dentro de la dimensión ética del ser humano, para pensar el presente estado de la educación y contribuir al estudio sobre el desarrollo y uso de IA con licenciamientos de código abierto en las aulas universitarias. Se entiende que hay una serie de problemáticas socio-éticas y políticas que es necesario analizar para enfrentarlas.

2 METODOLOGÍA

El presente trabajo de investigación forma parte de un diseño múltiple de investigación, el cual se deriva de la línea de generación y aplicación del conocimiento denominada “Ética deliberativa y solidaria”, que se enfoca en el desarrollo de un modelo ético-hermenéutico para la promoción de una praxis de responsabilidad compartida que, en este caso, se refiere al uso de la IA en la enseñanza superior (*The multi-baseline research design*). Mediante el análisis de los principales escenarios educativos en curso y de las nuevas tecnologías emergentes, se abordan las distintas relaciones entre la tecnología, las políticas educativas y la sociedad.

El itinerario metodológico de este trabajo se fundamenta en el arco hermenéutico propuesto por Paul Ricoeur (2010), el cual se construye a través de tres fases: 1) se parte de la explicitación de una comprensión inicial del campo problemático que nos ocupa hasta llegar a la tipificación del problema en construcción; 2) se procede a dimensionar el impacto del problema en el ámbito educativo, mediante al análisis de las relaciones existentes entre las partes que componen el todo en busca de patrones encriptados y las consecuentes responsabilidades inherentes a la acción de indagación e intervención; y 3) se procura una síntesis o comprensión integral: el arco

hermenéutico de la comprensión inicial y el análisis explicativo se encuentran en un horizonte que permite vislumbrar un futuro posible y probable; y en donde el control en las tomas de decisiones lleve un sello que otorgue su debida importancia a la deliberación reglamentada con códigos éticos consensuados y compartidos, con el fin de brindar escenarios de igualdad solidaria entre los actores involucrados en el ejercicio ético-hermenéutico que implica tanto la investigación como la intervención.

La herramienta sociocultural “connected-papers” ha sido utilizada para identificar las tendencias en torno al problema que nos ocupa. El análisis de los textos se realizó por medio de la técnica de aproximación al corpus lingüístico-textual etiquetado, para descubrir los nodos semánticos que nos ayudarán a develar los patrones y tendencias de los sistemas de razón contemporáneos.

2.1 Comprensión inicial del uso de la IA-CA con apoyo en diversos autores

Para la comprensión inicial del campo problemático que nos ocupa, es necesario iniciar con una conceptualización del problema en un ejercicio paralelo para su delimitación y contextualización. Nos referimos, específicamente, al uso de la inteligencia artificial de código abierto en el ámbito de la educación.

Sobre el concepto de IA, en sentido general, se relaciona con aquellos programas que ‘imitan’ la conducta humana, basados en una amplia gama de datos; que brinda eficiencia y rapidez al análisis de un cúmulo de datos para la solución de problemas mediante un aprendizaje estratégico requiriendo o prescindiendo de la supervisión humana. Sin embargo, Haelen y Kaplan (2021) la definen como: “[...] un sistema que interpreta y aprende correctamente de datos externos, para lograr resultados específicos a través de una adaptación flexible” (Haelen y Kaplan, 2021 en Chan, 2022, p. 3).

El vocablo "sistema" hace referencia, más

que a los sistemas tecnológicos -llámense algoritmos o conjuntos de datos de la GPT-3- a los sistemas socio-tecnológicos implicados en la generación y uso de las IA. El sistema no se limita a los sistemas tecnológicos (como el hardware, los algoritmos y los conjuntos de datos de GPT-3).

El vocablo “sistema” hace referencia a los sistemas socio-tecnológicos implicados en la generación, implementación y uso de la IA. Esto significa considerar a los actores humanos como esenciales en el desarrollo y regulación de la IA, especialmente en los modelos de lenguaje de IA-CA.

En relación con los usos de la IA, se considera que pueden orientarse hacia la automatización, la predicción y el control de procesos, los cuales pretenden incidir en innovación, competitividad y producción óptima. En el informe de investigación 451 de Hewlett Packard Enterprise (2019) se describen dos categorías de IA, a saber:

- La *IA Basada en Funcionalidad*, la cual echa mano de información acumulada para presentar mejores opciones, pero no tiene capacidad para aprender de acciones pasadas. Un ejemplo de este tipo de IA es el localizador GPS.
- La *IA Basada en Capacidades*, la cual presenta tres modalidades: 1) la *Inteligencia Artificial Estrecha*, que combina una memoria reactiva para realización de tareas simples; 2) la *Inteligencia Artificial General*, la cual tiene capacidad de aprender, enseñar, entender y simular al humano; y 3) la *Super Inteligencia Artificial*, cuyo desempeño, además del procesamiento de datos y su almacenamiento, es capaz de tomar decisiones similares al del ser humano promedio.

Una de las características más destacadas del desarrollo de máquinas con IA, es la relacionada con el Aprendizaje Automático con base en un

software algorítmico, el cual permite predecir y reaccionar ante situaciones imprevistas sustentados en resultados anteriores. Esto se realiza a través del reconocimiento de patrones y modelos analíticos para realizar ajustes sin previa programación de la máquina. Dicho aprendizaje automático, dependiendo del etiquetado de los datos, puede ser supervisado, semi-supervisado o no supervisado por el humano.

A partir del modelo de aprendizaje automático se ha llegado hasta los llamados *modelos de aprendizaje profundo*, los cuales procesan grandes cantidades de datos, generalmente no supervisados, mediante una amplia gama de redes neuronales artificiales. Estas IA superan las capacidades humanas (en cuanto al procesamiento del lenguaje natural) y realizan sus tareas mediante el análisis de conversaciones y el reconocimiento de imágenes a partir de videos usados para vigilar cuestiones operativas y, de esta manera, detectar identidades y lugares geográficos.

Por otra parte, es importante identificar el grado de responsabilidad ética de los diversos actores involucrados en la generación y desarrollo de máquinas con IA, esto en cuanto al uso de las mismas para beneficio y no maleficio de todas las partes involucradas. Aquí es donde se cuestiona la postura de quienes ignoran la necesidad de que, la sociedad en general y los expertos en particular, definan su postura ética y reconozcan su grado de responsabilidad en el desarrollo de tecnologías y su repercusión en las vidas humanas y demás seres que cohabitan en el planeta Tierra. Se trata de una responsabilidad que necesita de un estudio a fondo sobre: 1) el número, tipo y características de las máquinas que se van a implementar; 2) la aplicación que se le dará a las mismas y el objetivo beneficioso al que sirven; y 3) el grado de automatización, autonomía y su uso competitivo.

Como lo menciona Chan (2022), la mala aplicación de IA-CA puede debilitar la posibilidad de la autonomía personal o grupal al alentar a las personas a actuar con fines que no son los suyos o por razones que no han construido. En el caso de

la IA-CA, su uso puede conducir a una serie de productos y servicios que promueven el desplazamiento de trabajadores por la automatización (las máquinas programadas sin apoyo humano: robots), la disposición para la reestructuración de otros sectores industriales (comercio) y los cambios en la lógica mental del consumidor. Incluso, se abre la posibilidad de relevar a los docentes por algoritmos.

De lo anterior, surge el riesgo de un uso indebido de la IA. A esto se suma el tipo de diseño humano y aplicación de los programas basados en IA que, junto con la carencia de formación ética por parte de los actores involucrados con dicha tecnología, puede socavar la autonomía al alentar a las personas a actuar con fines que no son los suyos o por razones que no son de su elección.

De acuerdo con Chan (2022), el determinismo tecnológico está estrechamente relacionado con una percepción poco informada de los alcances del desarrollo tecnocientífico y, en el caso específico de este trabajo, del desarrollo de la inteligencia artificial. Dicha autora, hace la distinción entre una visión utópica ingenua, la cual sólo aprecia las bondades de la tecnología y, en el otro extremo, una visión distópica que solo ve amenazas a las libertades humanas derivadas de una tecnología sin control.

En un estudio analítico sobre las soluciones éticas, en respuesta a los riesgos y beneficios que representan los modelos de lenguaje de inteligencia artificial abierta, tales como el Transformador pre-entrenado para generar lenguaje similar al humano, versión 3 (Generative Pre-trained Transformer-3: GPT-3), o el modelo de lenguaje Transformador Instruido y pre-entrenado para generar lenguaje similar al humano (*Instruct Generative Pre-trained Transformer: InstructGPT*). La autora Chan (2022) critica el determinismo tecnológico representado por dos visiones contrapuestas: por un lado, se presenta una visión utópica que considera la tecnología como una herramienta de control humano que facilita la acción autónoma.

La lógica subyacente a la visión utópica es

de carácter efficientista y mercantilista. Le subyace una lógica interna de "eficiencia, economía mercantil, innovación, progreso" que justifica y perpetúa esta visión. Como Colman (2005) señala, desde un enfoque ético, es posible construir una visión más amplia de la evolución de la tecnología en todos los ámbitos y su impacto económico y social sin recurrir a metáforas reduccionistas ni maniqueas. La tecnología es sólo un instrumento y es el ser humano el único responsable de la innovación liberadora.

Finalmente, se presenta una visión distópica que considera a la tecnología como una amenaza imposible de enfrentar a no ser que se le obstaculice su avance.

2.2 Análisis de las relaciones existentes entre las partes que componen el todo del problema en busca de patrones subyacentes y responsabilidades inherentes a la acción de indagación e intervención

Con el fin de construir una visión más amplia de la evolución de la tecnología y llegar a dimensionar su impacto en ámbitos como el social, político, económico y, en el caso de este estudio, el ámbito educativo, es preciso identificar las fuerzas conflictivas subyacentes, así como asumir las responsabilidades derivadas de los planes de indagación e intervención pertinentes para actuar con razón y razonabilidad.

En el caso de un método operativo basado en la ética-hermenéutica, donde se considera que la intervención del ser humano, supone la creación de espacios de debate que opten por una agenda temática acorde con los intereses y necesidades del ámbito correspondiente (educativo); la curaduría de las fuentes de información y la organización del entorno de trabajo-estudio de acuerdo con argumentos racionales y razonables.

Se observa que los procesos automatizados presentan el riesgo de generar problemas éticos en las instituciones educativas, como la falta de transparencia y explicabilidad de los algoritmos de la AI (Floridi et al. (2019), en LaGrandeur, 2021,

pág. 699) lo cual se manifiesta en la toma de decisiones por parte de la IA o en la reducción del control humano. También puede surgir una evasión de responsabilidad por parte de los "operadores", como es el caso de un programador quien no tiene disposición para evaluar las consecuencias que genera su acción. El problema es que, en caso de una desviación del comportamiento previsible, la IA podría aplicar criterios diferentes a los esperados por el programador o, en su caso, por los operadores. El operador ejercería su control de acuerdo con los valores humanos, pero estos pueden ser distintos a los programados en la IA. El programador necesita identificar los datos relacionados con el valor y entender su alcance. Si estos valores son diferentes, entonces el programador podría desviarse de su propósito original. El operador puede preferir una evaluación basada en la "razonabilidad" del comportamiento humano, por ejemplo: si la respuesta es errónea en un preciso momento, se puede prever que va a tener una forma similar de reaccionar en el futuro, como lo mencionan Nallur, Lloyd y Pearson (2021).

En este escenario tecnológico, nos encontramos con una IA-CA capaz de desplegar aprendizaje profundo sin límites éticos controlados en su uso; en consecuencia, tendremos que intentar resolver retos imprevistos en las instituciones educativas. Los problemas de eficiencia y libertades individuales en el caso de una IA-CA también conducen a preguntas éticas importantes referidas a la autonomía. Esto significa una concentración del poder en los operadores, quienes crean valores que luego son compartidos por la IA-CA. En este escenario, no se puede entender por qué la IA-CA desarrollaría una visión de valores diferentes del operador como lo analiza Fernández Fernández (2021). Desde esta perspectiva, es fundamental una educación ética en los estudiantes para ser operadores con perspectivas éticas. Los aportes de esta visión son dos:

- Los alumnos y docentes pueden asumir una parte de la responsabilidad en la vigilancia

de la IA-CA y, a su vez, la sociedad puede generar una vigilancia de las instituciones educativas que la operan.

- La participación de los alumnos y docentes en la toma de decisiones puede permitir que ambos logren construir su autonomía con el uso de la IA-CA. Asimismo, esto se convertiría en una apuesta a su desarrollo como seres humanos, ya que cada alumno podría involucrarse en el manejo de estos temas.

Entonces surge la pregunta sobre cómo plantear una perspectiva humanista digital para abonar a la ética en el uso de la IA-CA. Es importante aclarar que debido al escaso avance hacia el despliegue de una IA-CA configurada de acuerdo con los parámetros éticos planteados por la comunidad internacional, resulta necesario retroalimentar o complementar la filosofía ética y moral existente, ya que actualmente no es suficiente para controlar la IA-CA.

2.3 Comprensión integral para vislumbrar un futuro posible y probable

En la tercera fase del arco hermenéutico se plantea acceder a una comprensión profunda del problema para el advenimiento de un adecuado uso ético de la IA-CA; para nuestro caso, el ámbito de la educación universitaria.

Ahora bien, si analizamos y proponemos una reglamentación en el uso ético de la IA-CA en educación universitaria dentro de un contexto global, podemos observar distintas variantes de acuerdo con el contexto regional, como son los aspectos social, político y económico con una perspectiva globalizada en una construcción ética y moral de la IA-CA.

El código ético de la IA-CA en un contexto global establece principios para la formación de una reglamentación ética y social sobre el uso de la inteligencia artificial, tales como: el respeto a los derechos humanos, las libertades y la promoción de igualdad de oportunidades.

Un factor importante a tomar en cuenta para la construcción de una reglamentación ética

para la IA-CA en educación universitaria –y desde una perspectiva global– es identificar por qué no son fáciles de aplicar a una escala mundial. Por lo tanto, se propone sacar provecho de la globalidad de las universidades como un lugar para debatir sobre el derecho a la información y el acceso a la información curada, los métodos y las técnicas ad hoc a los nuevos entornos. Como lo mencionan Siqueira De Cerqueira y sus colaboradores (2021), también es necesario proponer la traducción de los principios éticos a la práctica y, de esta manera, constituir la formación ética de la IA que se base en el respeto, la responsabilidad, el compromiso cívico, el compromiso con las mayorías y minorías.

Hagendorff (2020) menciona la tensión entre los principios éticos y los intereses y objetivos más amplios de la investigación, la industria y los negocios; lo cual puede explicarse mediante la teoría sociológica. Todos estos sistemas tienen sus propios códigos, sus propios valores objetivos y sus propios tipos de capital (económico o simbólico) a través de los cuales se construyen y se toman decisiones. La intervención ética en estos sistemas solo puede realizarse a una escala muy limitada. En un contexto global, las universidades deben promover la aplicación de una ética de la IA-CA, para que esta sea una piedra angular en cada una de las diferentes profesiones y en la labor investigativa. De este modo, nuestros alumnos y profesores serán capaces de hacerse cargo de las nuevas tecnologías transformadoras, tales como: IA-CA, cloud computing, big data, etc. En consecuencia, se avanzará en el desarrollo de una ética de la inteligencia artificial cuando las universidades canalicen sus competencias, con apoyo en un código ético de la IA-CA, que tenga como objetivo principal no solo servir a ciertos intereses de lucro y/o poder sino contribuir al crecimiento del conocimiento y la expresión creativa de lo humano.

Como las tecnologías basadas en el uso de la IA siguen posicionándose en todos los ámbitos de la vida humana, en paralelo se han ido formulando propuestas de normatividad para

regular el uso de aplicaciones online predictivas, la incorporación de bots a aparatos electrónicos de uso doméstico, el reconocimiento facial y la videovigilancia.

Por su parte, la Comisión Europea consciente de los riesgos que supone la aplicación de la IA en las actividades ciudadanas, el 21 de abril de 2021 publicó su Propuesta de Reglamento IA, en el marco de la “Shaping Europe’s Digital Future”, para impulsar la transformación digital y la creación de nuevas oportunidades a empresas, garantizando máximo respeto por los derechos de los ciudadanos.

Mediante este reglamento se trata de establecer un marco legal sobre las implicaciones que conlleva el uso de la IA, para lo cual establece la siguiente clasificación sobre los posibles usos de la IA (ver Tabla 1):

Tabla 1. Tipos de aplicación de IA

Tipo de aplicación	Descripción
Prohibida	destinada a manipular, explotar información, clasificar evaluaciones sociales, control por vigilancia de forma indiscriminada
Sujeta a autorización	identificación biométrica remota en espacios públicos, sometida a autorización administrativa
Con requisitos específicos	con reglas concretas para usos como de un “chatbot” o para el empleo de sistemas “deep-fake”
De “alto riesgo”	el uso de IA para el funcionamiento de infraestructuras críticas requerirá previa verificación de un tercero independiente

Tabla 1. Tipos de aplicación de IA. Fuente: adaptado de Comisión Europea (CE), 2021.

Otros usos de IA como las aplicaciones

predictivas (por ejemplo, la admisión de estudiantes y la concesión de créditos) o aplicaciones para evaluación de riesgos, requieren de una declaración responsable.

Esta propuesta de reglamento IA “establece ciertas obligaciones de transparencia con los usuarios y consumidores, así como la obligación de garantizar una adecuada formación y conocimientos a las personas encargadas de gestionar y supervisar las aplicaciones que impliquen el uso de la IA (Andersen Global, pág s/n”. Asimismo, establece multas administrativas para empresas que incumplan sus obligaciones derivadas del uso de la IA. Finalmente, este marco regulatorio es una oportunidad para desarrollar nuevas aplicaciones de IA con seguridad jurídica, aunando innovación con altos estándares de seguridad y todo ello en favor de los derechos de usuarios y consumidores (Andersen Global, 2021).

Por otra parte, el reporte “Guías éticas para una AI confiable” (emitido por la Comisión Europea) tiene como propósito el proveer de un repositorio para evaluar el desarrollo responsable de sistemas AI que sean benéficos por la humanidad y el medio ambiente. En resumen, destacan siete elementos esenciales para lograr una AI confiable (ver Tabla 2):

Tabla 2. Siete elementos esenciales para lograr una AI confiable

Elementos	Descripción
1 Agencia humana y supervisión	Los sistemas de IA deben permitir sociedades equitativas mediante el apoyo a la agencia humana y derechos fundamentales y no disminuir, limitar o desorientar la autonomía humana.
2 Robustez y seguridad	La IA confiable requiere que los algoritmos sean seguros, confiables y lo suficientemente robustos como para lidiar con

	errores o inconsistencias durante todas las fases del ciclo de vida de los sistemas de IA.
3 Privacidad y gobernanza	Los ciudadanos deben tener control total sobre sus propios datos, mientras que los datos que les conciernen no se utilizarán para dañarlos o discriminarlos.
4 Transparencia	Debe garantizarse la trazabilidad de los sistemas de IA.
5 Diversidad	No discriminación y equidad: los sistemas de IA deben tener en cuenta toda la gama de capacidades humanas, habilidades y requisitos para garantizar la accesibilidad.
6 Bienestar social y ambiental	Los sistemas de IA deben utilizarse para mejorar el cambio social positivo y mejorar la sostenibilidad y la responsabilidad ecológica .
7 Rendición de cuentas	Deben establecerse mecanismos para garantizar la responsabilidad y la rendición de cuentas de los sistemas de IA y sus resultados.

Tabla 2. Siete elementos esenciales para lograr una AI confiable. Fuente: adaptado de Floridi, Luciano, 2019.

Una vez revisadas las propuestas de normatividad para asegurar un uso apegado a los principios éticos de beneficencia y no maleficencia, de justicia y de respeto a la autonomía mediante la explicabilidad suficiente, a continuación, se ofrece una propuesta de reglamentación del uso ético de la IA-CA en educación universitaria en un contexto global, la cual se presenta en la Tabla 3.

Tabla 3. Reglamentación del uso ético de la IA-CA en contextos universitarios a nivel global

Aspectos	Descripción
1 Fines	La IA debe ser utilizada en beneficio y salvaguardia del ser humano ante todo tipo de riesgos.
2 Principios	Los principios éticos inherentes a la IA-CA deben ser respetados y promovidos.
3 Etapas	La reglamentación deberá: partir de los principios; establecer una fase de investigación, detección de fallas y vulnerabilidades; y una fase de desarrollo hasta su aplicación contextualizada.
4 Necesidades	Debe tenerse en cuenta la responsabilidad del usuario en su manejo de la IA.
5 Aplicación	Se trata del trabajo cooperativo y concertado entre universidades las cuales deben suscribir el Código ético de la IA-CA.
6 Evaluación	Se refiere a la valoración del grado de aplicación de la IA en cada institución, su contribución y otros factores que contribuyan a una mejora en la calidad y eficiencia educativa.
7 Derechos del usuario	Tales como el derecho a la privacidad de datos personales y sus propiedades intelectuales.

8 Derechos de los estudiantes	Se refiere al interés por buscar, recibir y expresar su opinión en relación a las aplicaciones que se emitan ética y pedagógicamente.
9 Temática	En este caso, señalar la relación entre la IA-CA, los beneficios educativos y la privacidad de los alumnos.

Tabla 3. Reglamentación del uso ético de la IA-CA en contextos universitarios a nivel global. Fuente: elaboración propia.

Los desarrollos en IA-CA tienen un gran potencial para mejorar la vivencia de experiencias de aprendizaje e investigación, pues proporcionan acceso a materiales educativos, recursos o información clave para el desarrollo multidimensional de los estudiantes. Aquí el problema radica en que la incorporación inapropiada o errónea de la IA puede resultar perjudicial para la educación. Sólo si el uso asume la perspectiva de que es necesario proveer acceso a una infraestructura abierta y conectada, basado en el uso ético de los recursos en un entorno transparente, podrá ser positivo para la educación universitaria. El sentido de *transparencia* es en relación con los algoritmos libres de intencionalidad maléfica.

En la actualidad, el mundo universitario aún se encuentra lejos de tener la suficiente conciencia y preocupación acerca del uso ético de la IA-CA en el ámbito educativo. El uso inapropiado o errado podría generar problemas tales como: que una persona consiga una gran cantidad de atención que no es debida e, incluso, la exclusión injusta o la discriminación por razones no deseadas.

Es necesario plantear una reglamentación adecuada al contexto geográfico de cada una de las instituciones educativas, la cual tome en cuenta lo que señala Chan (2022) sobre los investigadores de IA-CA, quienes deberían tener cierta responsabilidad sobre cómo se presenta su

investigación al público, con una interpretación transparente y comprensible de la tecnología y la divulgación. Asimismo, implica una responsabilidad ética que permita analizar y reflexionar sobre el papel de la tecnología en nuestra sociedad y el impacto que tiene en nuestras vidas. Dada la complejidad de una inteligencia artificial que se usa para el objetivo educativo, es indispensable una reglamentación contextualizada en instituciones educativas a nivel regional.

En el marco de una institución educativa, la implementación de esta reglamentación contextualizada requiere una explicación pertinente del contexto conceptual que gira en torno al impacto de IA-CA en la educación. La IA-CA puede ayudar a mejorar procesos de educación para reducir el costo de la enseñanza, así como a ampliar el alcance de los programas. Puede aprovechar también que se están creando instituciones multidisciplinarias y desafiantes (I+D+I) para generar evidencias comunicativas eclécticas. Esto dinamiza las capacidades tecnológicas necesarias para llevar a cabo dichos proyectos digitales.

Es posible también tomar en cuenta lo que comenta Ekambaram (2022) para plantear un código ético de IA-CA en educación universitaria en un contexto regional (que ciertamente vale para estas instituciones), que es aprovechar lo más eficazmente posible la ciencia de los datos y las tecnologías para impulsar el desarrollo de las capacidades cognitivas de los estudiantes. Es necesario construir un código ético de manera que funcione de manera integral y unitaria tanto para el software como las experiencias de aprendizaje.

Entre los puntos importantes de este código hay que señalar varios: el código ético debe respetar el principio del cumplimiento a la norma (*discarding the rule*) y en caso contrario, señalar si es válida o no. El código ético debe enfatizar que el desarrollo de la inteligencia artificial es algo que se trata de promover, más no someter a una crítica pedagógica. Es necesario

incluir una mejora del software para apoyar el aprendizaje y el desarrollo cognitivo; además, debe ser transparente sobre los datos y las técnicas consultadas para lograr este objetivo. En cuanto al sesgo inherente a la IA-CA, hay que aclarar que esta se denomina como sesgo porque la inteligencia artificial tiene una tendencia clara de poder predecir con más o menos acierto los resultados de un modelo (también denominado *tendency bias*), sin embargo, no siempre se tiene claro si es el software o el usuario el que hace esta toma de decisiones. Si la decisión es del usuario, el código ético debe incluir una explicación clara sobre los resultados en cuanto a elaboraciones de datos y herramientas de análisis. En caso de que se trate del software, hay que mencionar este factor y verificar si los resultados son tendenciales o no, pues esto puede afectar al usuario.

A continuación, se presenta en la Tabla 4 una propuesta de reglamentación del uso ético de la IA-CA en educación universitaria en un contexto regional.

Tabla 4. Reglamentación del uso ético de la IA-CA en contextos universitarios nivel regional

Aspecto	Descripción
1 General	El código ético para el uso de la IA-CA debe iniciar con una declaración del contexto y la definición de los términos del lenguaje con el que se opera en la IA. Esto permite que los usuarios sepan cuáles son las implicaciones subyacentes al tipo de software utilizado.
2 Seguridad	Cualquier cambio en la reglamentación debe contemplar los temas de seguridad, pues con la IA puede verse en riesgo la

	privacidad y la protección de datos.
3 Desarrollo de Software	La IA-CA debe ser explícita respecto a su uso y los derechos que otorga al usuario del programa.
4 Metodología	La IA-CA debe ser transparente y señalar sus herramientas de análisis, ya que esto puede reflejarse en el desarrollo del código.
5 Reglamentación formal	Debe ser clara en cuanto a su procesamiento, es decir, en cuanto a los datos o patrones que se usan para conseguir los resultados esperados.
6 Evaluación continua	Del software, es necesario revisar la utilidad y funcionalidad del programa para evitar fallas y aspectos vulnerables.
7 Aspecto educativo	Una evaluación constante de la IA-CA puede dar señales de su utilidad para optimizar los procesos de enseñanza-aprendizaje.
8 Aspecto ético y pedagógico	Ambos se ven implicados para asegurar que el software sea suficientemente versátil para ser utilizado de manera flexible sin limitar su potencial.
9 Responsabilidad	Del usuario al utilizar IA-CA en instituciones educativas.

Tabla 4. Reglamentación del uso ético de la IA-CA en contextos universitarios nivel regional. Fuente: elaboración propia.

4.- Reflexiones finales

En este artículo se han propuesto diferentes aspectos del uso de la IA-CA que deben considerarse desde una visión prospectiva, de tal modo que, si la IA-CA en educación universitaria se utilizara en su justa medida, su uso ético-hermenéutico contribuirá al desarrollo de la investigación, al desarrollo de nuevas capacidades a partir de los datos, la selección automática libre de favoritismos del perfil del estudiante, etc. De esta manera se puede afirmar que el uso ético de la IA-CA en educación universitaria contribuye a la formación profesional y al propósito de servir a la humanidad en cada persona.

El interés que subyace a este trabajo se encamina a la construcción de códigos de ética en el uso de la IA-CA en educación universitaria para impulsar la elaboración consensuada de una reglamentación ética y social sobre el uso de la IA contextualizada. La actividad de las universidades en investigación, desarrollo y aplicación para un uso ético de la IA-CA, requiere fundamentalmente de confianza entre los estudiantes y los profesores. De esta manera, las universidades se convertirán en centros de futuros profesionales que manejen la IA-CA con agencia ética-hermenéutica. Para el desarrollo de esta capacidad, es fundamental que las universidades cuenten con mecanismos de participación ciudadana en sus instituciones para acceder a un uso de la IA-CA que promueva el respeto, el reconocimiento y la reciprocidad en el cuidado mutuo dentro de entornos digitales.

El profesional del conocimiento en una sociedad abierta debe pensar, proponer y debatir tanto al interior de la academia como al exterior con la sociedad. No solo importan los papeles de estudiante y profesor sino, sobre todo, el actuar con responsabilidad social, mismo que se renueva a través de la relación en el proceso formativo y al expandirse las fronteras del conocimiento. Esta labor es de la máxima importancia en una sociedad abierta.

Es posible que haya un desarrollo benéfico

de la IA, a la par de aplicaciones en riesgo, por lo que el uso ético debe encontrar un marco jurídico y social donde se ejerciera un control y una valoración ético-hermenéutica de su uso. De esta manera, se presenta la posibilidad de desarrollar códigos éticos de la IA-CA.

La responsabilidad de los agentes académicos, sean estudiantes, profesores o administrativos, necesita tomar en cuenta que la IA debe ser un sistema sociocultural de servicio y no una herramienta tecnológica para un propósito de dominación y exclusión de las personas.

La IA debe ser una parte fundamental de la innovación, no sólo en cuanto a su contenido en el proceso formativo, sino, preferentemente en su influencia en el desarrollo de una ciudadanía digital con responsabilidad compartida respecto del cuidado de las personas y del medio ambiente. La innovación en la educación es un factor clave para el futuro. En este sentido, los códigos de ética-hermenéutica no son un límite a la investigación innovadora, sino un andamiaje respuestas nuevas a contextos inéditos.

5. Referencias

- Andersen Global (2021). La Propuesta de Reglamento sobre Inteligencia Artificial, nuevo marco legal para la innovación. Obtenido de <https://es.andersen.com/es/la-propuesta-de-reglamento-sobre-inteligencia-artificial-nuevo-marco-legal-para-la-innovacion.html>
- Chan, A. GPT-3 and InstructGPT: technological dystopianism, utopianism, and “Contextual” perspectives in AI ethics and industry. *AI Ethics* (2022). DOI: <https://doi.org/10.1007/s43681-022-00148-6>
- Colman, A. (2005). Un/becoming digital: The ontology of technological determinism and its implications for art education. *Journal of Social Theory in Art Education*, 25(1), pp. 278-305.
- Comisión Europea (CE) (2021). *ReasonWhy*. Obtenido de <https://www.reasonwhy.es/actualidad/comision-europea-propuesta-regulacion-inteligencia-artificial>
- Comisión Europea (CE), (2021). Propuesta de Reglamento sobre Inteligencia Artificial en el marco de la estrategia digital europea denominada “Shaping Europe’s Digital Future”.
- Ekambaram, V. (2022). Creating a Code of Ethics for Open-Source Intelligence Applications. In *Proceedings of the Wellington Faculty of Engineering Ethics and Sustainability Symposium*. DOI: <https://doi.org/10.26686/wfeess.vi.7648>
- Fernández-Fernández, J. L. (2021). Hacia el Humanismo Digital desde un denominador común para la Ciberética y la Ética de la Inteligencia Artificial. *Disputatio. Philosophical Research Bulletin*, 10(17), pp. 107-130.
- Floridi, L. (Junio de 2019). Establishing the rules for building trustworthy AI. En: *Nature Machine Intelligence*, 1, p. 261-262. Recuperado el junio de 2022, de <http://www.nature.com/natmachintell>
- Goldie, M. B. (2015) “An Early English Rutter: The Sea and Spatial Hermeneutics in the Fourteenth and Fifteenth Centuries,” *Speculum*, 90(3), pp. 701–727.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), pp. 99-120.
- LaGrandeur, K. (2021). How safe is our reliance on AI, and should we regulate it? En *AI and Ethics*, 1(2), pp. 93-99.
- licenses.ai (2022) En www.licenses.ai. Recuperado el 15 de octubre de 2022 de <https://www.licenses.ai/blog/2022/8/26/bigscience-open-rail-m-license>
- N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan (2021) “A survey on bias and fairness in machine learning,” *ACM Computing Surveys (CSUR)*, 54(6), pp. 1–35.
- Nallur, V., Lloyd, M., & Pearson, S. (2021). Automation: An Essential Component Of Ethical AI?. *arXiv preprint arXiv:2103.15739*.
- Özkardes-Cheung, Y. (2021). Hewlett Packard Enterprise (2019). Investigación 451. *Entrepreneur en español*. Recuperado de <https://www.entrepreneur.com/es/marketing/la-guia-completa-de-inteligencia-artificial-para-empresas/>
- Ricoeur, Paul, (2010). *Del texto a la acción. Ensayos de hermenéutica II*. Argentina: Fondo de Cultura Económica.
- Siqueira De Cerqueira, J. A., Dos Santos Althoff, L., Santos De Almeida, P., & Dias Canedo, E. (2021, January). Ethical perspectives in AI: A two-folded exploratory study from literature and active development projects. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (p. 5240).
- Stable Diffusion. (s.f.). En Wikipedia. Recuperado el 15 de octubre de 2022 de https://es.wikipedia.org/wiki/Stable_Diffusion