

Sección III

Desafíos éticos

Capítulo 7

Uso ético de la inteligencia artificial en la educación superior: Una revisión documental crítica

*Alberto Apolinar Vázquez Jácome
Universidad Autónoma de Baja California
Facultad de Humanidades y Ciencias Sociales*

<https://doi.org/10.61728/AE26001159>



Resumen

El presente capítulo analiza el uso ético de la inteligencia artificial (IA) en la educación superior a partir de una revisión documental crítica de literatura académica y documentos normativos publicados entre 2020 y 2025. El objetivo consiste en identificar los principales enfoques éticos presentes en el debate contemporáneo y examinar sus convergencias, tensiones y vacíos conceptuales. Metodológicamente, se adoptó una revisión narrativa estructurada con análisis temático, lo que permitió clasificar los estudios en cuatro orientaciones predominantes: descriptiva, normativa, crítica y formativa.

Los hallazgos evidencian que, si bien existe consenso en torno a la necesidad de subordinar la IA al respeto por la dignidad humana, la equidad y el bien común, persisten tensiones entre enfoques tecnocéntricos orientados a la eficiencia y perspectivas humanistas centradas en la formación integral. Asimismo, se identifican dilemas recurrentes vinculados con la autoría académica, los sesgos algorítmicos, la protección de datos y la posible deshumanización del proceso educativo.

El capítulo sostiene que el uso ético de la IA en la universidad requiere una articulación integrada entre principios normativos, reflexión crítica y práctica pedagógica. En consecuencia, se propone una perspectiva ético-pedagógica e institucional que reconoce la responsabilidad de las universidades en la formación de sujetos capaces de interactuar con tecnologías emergentes de manera consciente, responsable y socialmente comprometida.

Introducción

La incorporación de la IA en la educación superior se ha intensificado de manera acelerada durante la última década, particularmente a partir del desarrollo de sistemas de análisis de datos, aprendizaje automático y herramientas de inteligencia artificial generativa. Estas tecnologías

han comenzado a incidir de forma directa en los procesos de enseñanza, evaluación, investigación y gestión universitaria, reconfigurando prácticas académicas tradicionales y generando nuevos escenarios de actuación institucional. En este contexto, se ha señalado que “la inteligencia artificial está transformando la educación superior no solo en términos operativos, sino también en sus fundamentos pedagógicos y éticos” (Zawacki-Richter et al., 2020, p. 6).

Diversos autores coinciden en que la IA ofrece oportunidades relevantes para la personalización del aprendizaje, la mejora de la eficiencia administrativa y el apoyo a la toma de decisiones académicas. No obstante, la literatura también advierte que su adopción acrítica puede profundizar desigualdades existentes y comprometer principios fundamentales de la educación universitaria. Al respecto, se ha documentado que “las tecnologías algorítmicas tienden a reproducir y amplificar sesgos sociales cuando no se diseñan ni implementan bajo criterios éticos claros” (O’Neil, 2020, p. 21), lo cual plantea riesgos significativos en contextos educativos caracterizados por la diversidad y la equidad como valores centrales.

Desde una perspectiva ética, la discusión sobre la inteligencia artificial en la universidad no puede limitarse a aspectos técnicos o instrumentales. La ética de la IA implica cuestionar los fines educativos, el papel de los actores universitarios y la responsabilidad institucional frente al impacto social de estas tecnologías. En este sentido, se ha afirmado que “los principios éticos en la inteligencia artificial deben orientarse a la protección de la dignidad humana y al fortalecimiento de las capacidades de las personas” (Floridi et al., 2020, p. 4), particularmente en entornos formativos donde la educación cumple una función social y humanista.

Organismos internacionales han subrayado la urgencia de establecer marcos éticos para el uso de la inteligencia artificial en la educación. La UNESCO, por ejemplo, sostiene que “los sistemas de inteligencia artificial aplicados a la educación deben estar centrados en el ser humano y contribuir al bien común” (UNESCO, 2021, p. 7), enfatizando la necesidad de transparencia, equidad, inclusión y rendición de cuentas. Sin embargo, la literatura académica muestra que existe una tensión constante entre los enfoques normativos centrados en la regulación tecnológica y aquellos

que conciben la ética como un proceso formativo y cultural dentro de las instituciones educativas.

En el ámbito universitario, esta tensión se refleja en debates relacionados con la autoría académica, el uso de herramientas de IA por parte del estudiantado, la evaluación del aprendizaje y la protección de datos personales. Investigaciones recientes señalan que “la integridad académica enfrenta nuevos desafíos ante el uso de herramientas de inteligencia artificial que difuminan los límites entre apoyo legítimo y deshonestidad académica” (Eaton, 2023, p. 3). Estas problemáticas evidencian que la ética de la IA no puede abordarse únicamente desde la prohibición, sino desde la formación crítica y la construcción de criterios compartidos.

Desde una perspectiva pedagógica, distintos autores advierten que la incorporación de la IA debe analizarse a la luz del sentido educativo de la universidad: “la educación no puede reducirse a procesos de eficiencia o rendimiento, sino que debe orientarse a la formación de sujetos responsables en el mundo” (Biesta, 2021, p. 15). Esta visión resulta particularmente relevante para el análisis ético de la inteligencia artificial, en tanto permite cuestionar los efectos de la automatización sobre la relación pedagógica y la centralidad de la experiencia humana en el aprendizaje.

En este contexto, el objetivo de este capítulo es analizar, mediante una revisión documental crítica, los principales enfoques éticos presentes en la literatura reciente sobre el uso de la inteligencia artificial en la educación superior. A partir del contraste entre distintas posturas académicas e institucionales, se busca identificar convergencias, tensiones y vacíos en el debate actual, con el propósito de construir una postura fundamentada que articule la ética, la pedagogía y la responsabilidad institucional como ejes orientadores del uso de la IA en la universidad contemporánea.

Metodología

El presente capítulo se desarrolló a partir de una revisión documental de carácter analítico y crítico, orientada a examinar los principales enfoques éticos sobre el uso de la inteligencia artificial en la educación superior. Este tipo de revisión permite no solo identificar tendencias conceptuales en la literatura especializada, sino también contrastar posturas teóricas y

normativas, así como evidenciar tensiones y vacíos en el debate académico. En este sentido, se reconoce que “la revisión documental crítica no se limita a sintetizar información, sino que implica interpretar, comparar y problematizar los enfoques existentes” (Snyder, 2019, p. 334).

El enfoque metodológico adoptado responde a la necesidad de analizar un fenómeno complejo y multidimensional, en el que convergen dimensiones éticas, pedagógicas, institucionales y sociopolíticas. Dado que el objetivo del capítulo no es medir impactos ni evaluar intervenciones específicas, sino comprender y analizar críticamente las perspectivas existentes, se optó por una revisión narrativa estructurada con análisis temático, la cual resulta pertinente cuando se abordan campos de estudio en consolidación o en rápida transformación. De acuerdo con Grant y Booth, “las revisiones narrativas estructuradas facilitan el análisis crítico de conceptos emergentes y debates en campos en rápida evolución” (Grant y Booth, 2009, p. 95).

Enfoque de la revisión documental

La revisión documental se orientó a identificar no solo consensos, sino también divergencias y contradicciones entre autores respecto al uso ético de la inteligencia artificial en la educación superior. Desde esta perspectiva, el análisis se centró en examinar cómo distintos estudios conceptualizan la ética de la IA, qué problemas priorizan y qué responsabilidades atribuyen a las instituciones universitarias.

Para ello, se integraron aportes provenientes de la literatura académica, documentos normativos y estudios críticos, lo cual permitió articular miradas interdisciplinarias desde la ética, la pedagogía y los estudios sobre tecnología. Esta integración responde a la premisa de que “el análisis ético de la inteligencia artificial requiere enfoques interdisciplinarios que superen la mirada exclusivamente técnica” (Mittelstadt, 2021, p. 501).

El enfoque adoptado permitió analizar la inteligencia artificial no como un conjunto homogéneo de herramientas, sino como un fenómeno sociotécnico cuya implementación en la universidad está mediada por decisiones institucionales, marcos normativos y concepciones pedagógicas. De este modo, la revisión documental se concibió como un proceso

interpretativo, orientado a comprender el sentido y las implicaciones éticas del uso de la IA en el ámbito universitario, más allá de una descripción funcional de sus aplicaciones.

Selección del conjunto de documentos revisados y criterios de análisis

La selección del corpus documental se realizó a partir de criterios explícitos de temporalidad, pertinencia temática y relevancia académica, con el propósito de asegurar coherencia metodológica y profundidad analítica. Se consideraron publicaciones realizadas entre 2020 y 2025, periodo en el cual se observa un crecimiento significativo de la producción académica sobre inteligencia artificial y ética en la educación superior. Esta delimitación temporal permitió concentrar el análisis en enfoques contemporáneos y debates actuales, en consonancia con la recomendación de que “la definición clara de los criterios de selección es fundamental para asegurar la coherencia y profundidad de una revisión documental” (Petticrew y Roberts, 2006, p. 39).

La selección estuvo conformada por artículos científicos arbitrados, libros académicos, capítulos de libro e informes de organismos internacionales, seleccionados por su contribución directa al análisis del uso ético de la inteligencia artificial en contextos universitarios. De manera específica, se priorizaron documentos centrados en la educación superior, la ética de la inteligencia artificial, la integridad académica, la equidad educativa y la gobernanza institucional, dado que estos ejes resultan centrales para el propósito del capítulo.

Las fuentes analizadas provienen de bases de datos académicas internacionales y repositorios institucionales, así como de documentos emitidos por organismos multilaterales especializados en educación y tecnología. La inclusión de distintos tipos de fuentes respondió a la necesidad de integrar perspectivas académicas, normativas y críticas, reconociendo que el debate ético sobre la inteligencia artificial no se construye exclusivamente desde la investigación empírica, sino también desde marcos regulatorios y reflexiones teóricas. En este sentido, se asume que la diversidad de fuentes fortalece el análisis, al permitir un diálogo entre distintos niveles de producción del conocimiento.

Estrategia de análisis y categorización temática

El análisis del corpus documental se llevó a cabo mediante una estrategia de análisis temático, orientada a identificar patrones recurrentes, tensiones conceptuales y diferencias en la manera en que la literatura aborda el uso ético de la inteligencia artificial en la educación superior. Este tipo de análisis permitió organizar la información de manera sistemática y facilitar el contraste entre autores y enfoques. De acuerdo con Braun y Clarke (2021), “la categorización temática permite identificar regularidades en los datos y construir interpretaciones analíticas que van más allá de la mera descripción” (p. 332).

Como parte del proceso analítico, se realizó una revisión de los objetivos de las investigaciones seleccionadas, con el fin de comprender las orientaciones predominantes en el campo de estudio. A partir de este ejercicio, se identificaron cuatro grandes tipos de objetivos en la literatura: descriptivos, normativos, críticos y formativos. Los estudios con objetivos descriptivos se centran en caracterizar los usos actuales de la inteligencia artificial en la educación superior y en mapear tendencias emergentes, al señalar que “el análisis del estado del arte resulta indispensable para comprender el alcance y las aplicaciones de la IA en contextos universitarios” (Zawacki-Richter et al., 2020, p. 8).

Por su parte, las investigaciones con objetivos normativos buscan proponer principios éticos, marcos regulatorios o lineamientos institucionales para orientar el uso responsable de la inteligencia artificial. En este tipo de trabajos se subraya que “la formulación de principios éticos constituye un paso necesario, aunque no suficiente, para garantizar un uso responsable de la inteligencia artificial” (Floridi et al., 2020, p. 6). En contraste, los estudios con objetivos críticos problematizan los riesgos y efectos adversos del uso de la IA, tales como la reproducción de sesgos, la afectación a la autonomía académica y la profundización de desigualdades educativas, al advertir que “los sistemas algorítmicos pueden reforzar desigualdades estructurales cuando no se cuestionan sus supuestos y condiciones de aplicación” (Benjamin, 2020, p. 12).

Los estudios con objetivos formativos conciben la ética de la inteligencia artificial como parte integral de la misión educativa de la uni-

versidad, destacando la necesidad de desarrollar capacidades críticas en estudiantes y docentes. Desde esta perspectiva, se sostiene que “la alfabetización ética y digital constituye una responsabilidad central de las instituciones de educación superior en contextos de automatización creciente” (Ramírez-Montoya, 2022, p. 4).

A partir de este análisis, se establecieron categorías conceptuales comunes que orientaron la organización del capítulo, entre las que se encuentran dignidad humana, equidad y justicia, transparencia algorítmica, responsabilidad institucional, integridad académica y formación ética. Estas categorías sirvieron como ejes analíticos para el desarrollo de los apartados posteriores, permitiendo articular los enfoques revisados y fundamentar el análisis crítico presentado.

Panorama del uso de la IA en la educación superior

El uso de la inteligencia artificial en la educación superior se ha expandido de manera heterogénea, tanto en alcance como en profundidad, dependiendo de los contextos institucionales, las políticas educativas y las concepciones pedagógicas subyacentes. La literatura reciente coincide en que la IA ya no constituye una innovación marginal, sino un componente estructural emergente en múltiples funciones universitarias. En este sentido, se ha documentado que “las aplicaciones de inteligencia artificial en la educación superior abarcan desde sistemas de tutoría inteligente hasta herramientas de apoyo a la gestión académica y administrativa” (Zawacki-Richter et al., 2020, p. 5).

No obstante, este crecimiento no ha sido uniforme ni exento de controversias. Mientras algunos autores destacan el potencial transformador de la IA para mejorar la calidad educativa, otros advierten sobre los riesgos de una adopción acelerada que no considere sus implicaciones éticas y pedagógicas. De acuerdo con Selwyn, “la inteligencia artificial en la educación suele presentarse como una solución técnica a problemas complejos que, en realidad, son de naturaleza social y pedagógica” (Selwyn, 2022, p. 18), lo cual plantea interrogantes sobre los fines reales de su implementación en la universidad.

La IA en la docencia universitaria: Usos y alcances

En el ámbito de la docencia, la literatura identifica como usos recurrentes de la inteligencia artificial la personalización del aprendizaje, los sistemas de recomendación de contenidos y la automatización de procesos de evaluación. Estas aplicaciones se sustentan en la promesa de adaptar los procesos educativos a las necesidades individuales del estudiantado. Al respecto, se ha señalado que “los sistemas basados en inteligencia artificial pueden ofrecer trayectorias de aprendizaje personalizadas a gran escala” (Holmes et al., 2022, p. 41).

Sin embargo, esta visión optimista es cuestionada por investigaciones que subrayan las limitaciones pedagógicas de dichos sistemas. Algunos estudios advierten que la personalización algorítmica puede reducir la complejidad del aprendizaje al priorizar patrones de rendimiento medibles, dejando de lado dimensiones formativas difíciles de cuantificar. En este sentido, se sostiene que “la educación universitaria no puede ser plenamente capturada por modelos algorítmicos centrados en la eficiencia y el rendimiento” (Biesta, 2021, p. 27).

Este contraste evidencia una tensión central en la literatura: mientras ciertos enfoques conciben la IA como un medio para optimizar la enseñanza, otros alertan sobre el riesgo de subordinar el sentido educativo a lógicas tecnocráticas.

Evaluación del aprendizaje y autoría académica en contextos de IA

Uno de los campos donde el impacto de la IA ha generado mayor debate es el de la evaluación del aprendizaje y la autoría académica. La aparición de herramientas de inteligencia artificial generativa ha desafiado las concepciones tradicionales de originalidad, autoría y trabajo académico. En este contexto, se ha documentado que “las tecnologías de IA generativa difuminan los límites entre la producción humana y la producción automatizada del conocimiento” (Kasneci et al., 2023, p. 3).

Desde una perspectiva crítica, diversos autores advierten que el uso de estas herramientas plantea riesgos significativos para la integridad

académica. Se ha señalado que “las políticas institucionales centradas exclusivamente en la detección del uso de IA resultan insuficientes para abordar los desafíos éticos de fondo” (Eaton, 2023, p. 5). En contraste, otros estudios proponen reconfigurar los modelos de evaluación para integrar el uso responsable de la IA como parte del proceso formativo, lo cual implica repensar los criterios de evaluación y el rol del estudiante como sujeto activo del aprendizaje.

Este debate refleja una divergencia clara en la literatura entre enfoques punitivos y enfoques formativos, lo cual tendrá implicaciones directas en la manera en que la universidad asume su responsabilidad ética frente al uso de la inteligencia artificial.

Gestión universitaria y toma de decisiones basadas en IA

Más allá del aula, la inteligencia artificial ha comenzado a incorporarse en la gestión universitaria, particularmente en procesos de admisión, análisis de datos institucionales, predicción de deserción y asignación de recursos. Estas aplicaciones se presentan como herramientas para mejorar la eficiencia y la toma de decisiones basada en datos. En este sentido, se ha indicado que “los sistemas algorítmicos permiten a las instituciones anticipar comportamientos y optimizar procesos administrativos” (Williamson y Eynon, 2020, p. 12).

No obstante, la literatura crítica advierte que estas prácticas pueden reforzar desigualdades estructurales si no se analizan sus supuestos y sesgos. Algunos estudios señalan que “los sistemas predictivos utilizados en contextos educativos tienden a reproducir patrones históricos de exclusión cuando se basan en datos sesgados” (Benjamin, 2020, p. 16). Esta situación resulta especialmente problemática en universidades comprometidas con la equidad y la inclusión social.

El contraste entre eficiencia institucional y justicia educativa constituye uno de los dilemas éticos más relevantes identificados en la revisión documental, evidenciando la necesidad de una gobernanza ética de la inteligencia artificial en el ámbito universitario.

Tendencias emergentes y tensiones en el uso universitario de la IA

A partir del análisis de la literatura, se identifican dos tendencias predominantes en el panorama del uso de la inteligencia artificial en la educación superior. Por un lado, una tendencia tecnocéntrica, que enfatiza la innovación, la eficiencia y la automatización de procesos; por otro, una tendencia crítica y humanista, que cuestiona los efectos de la IA sobre la formación integral, la autonomía académica y la justicia social.

Esta tensión ha sido sintetizada al señalar que “la adopción de la inteligencia artificial en la educación superior se encuentra marcada por una disputa entre visiones orientadas al mercado y enfoques centrados en el valor público de la educación” (Williamson, 2021, p. 9). En este escenario, la ética emerge como un eje articulador que permite evaluar no solo el cómo, sino el para qué del uso de la inteligencia artificial en la universidad.

Enfoques éticos identificados en la literatura sobre IA y educación superior

El análisis del corpus documental permitió identificar distintos enfoques éticos desde los cuales se ha problematizado el uso de la inteligencia artificial en la educación superior. Estos enfoques no son excluyentes, pero sí revelan prioridades, supuestos y tensiones que configuran el debate actual. La literatura coincide en que la ética de la IA en la universidad no constituye un campo homogéneo, sino un espacio de disputa conceptual entre visiones normativas, críticas y formativas.

Enfoques normativos y marcos éticos para el uso de la IA

Un primer enfoque ampliamente documentado es el normativo, centrado en la formulación de principios éticos generales para orientar el desarrollo y uso de la inteligencia artificial. Desde esta perspectiva, se propone

que la IA debe regirse por valores como la transparencia, la equidad, la rendición de cuentas y el respeto a los derechos humanos. Se ha señalado que “los marcos éticos basados en principios buscan establecer estándares mínimos que orienten el diseño y la implementación de sistemas de inteligencia artificial” (Floridi et al., 2020, p. 690).

En el ámbito educativo, este enfoque ha sido promovido principalmente por organismos internacionales, los cuales subrayan que la IA debe mantenerse bajo control humano y contribuir al bien común. En este sentido, se establece que “los sistemas de inteligencia artificial aplicados a la educación deben respetar la dignidad humana y promover la inclusión y la justicia social” (UNESCO, 2021, p. 8).

No obstante, diversos autores advierten que los principios, por sí solos, resultan insuficientes para garantizar un uso ético de la IA. Se ha argumentado que “la existencia de principios éticos no asegura su aplicación efectiva en contextos institucionales complejos” (Mittelstadt, 2021, p. 503), lo cual plantea la necesidad de mecanismos de implementación y evaluación más allá del plano declarativo.

Perspectivas críticas y sociopolíticas sobre la IA en la universidad

Un segundo enfoque identificado en la literatura es el crítico y sociopolítico, el cual problematiza los efectos estructurales de la inteligencia artificial en la educación superior. Desde esta perspectiva, la IA no es concebida como una herramienta neutral, sino como una tecnología inscrita en relaciones de poder, intereses económicos y dinámicas de exclusión. Se ha documentado que “los sistemas algorítmicos tienden a reproducir desigualdades sociales preexistentes cuando se basan en datos históricos sesgados” (Benjamin, 2020, p. 14).

Este enfoque cuestiona la adopción de la IA bajo discursos de eficiencia e innovación, señalando que dichas narrativas pueden invisibilizar impactos negativos sobre grupos vulnerables. En el contexto universitario, se advierte que “la automatización de decisiones académicas puede afectar de manera desproporcionada a estudiantes provenientes de contextos socialmente desfavorecidos” (Williamson y Eynon, 2020, p. 229).

Desde esta postura, la ética de la IA implica una reflexión sobre justicia social, gobernanza democrática y responsabilidad pública de las universidades, más allá del cumplimiento normativo.

La dimensión pedagógica y humanista en la ética de la IA

Un tercer enfoque relevante es el pedagógico-humanista, el cual sitúa la ética de la inteligencia artificial en relación directa con el sentido formativo de la educación superior. Desde esta perspectiva, la IA debe evaluarse en función de su impacto en la formación integral del estudiante y en la relación pedagógica. Se ha sostenido que “la educación universitaria no puede reducirse a procesos de optimización técnica sin comprometer su dimensión ética y formativa” (Biesta, 2021, p. 29).

Este enfoque enfatiza que la universidad tiene la responsabilidad de formar sujetos críticos capaces de comprender y cuestionar el uso de tecnologías emergentes. En este sentido, se plantea que “la alfabetización ética y digital constituye una dimensión central de la formación universitaria en contextos de inteligencia artificial” (Ramírez-Montoya, 2022, p. 5).

A diferencia de los enfoques normativos, esta postura subraya la importancia de integrar la ética de la IA en los procesos educativos, no solo como regulación externa, sino como práctica formativa cotidiana.

Tensiones y convergencias en los enfoques éticos sobre IA

El contraste entre los enfoques identificados evidencia una tensión persistente entre la regulación técnica, la crítica sociopolítica y la formación ética. Mientras los enfoques normativos buscan establecer límites claros al uso de la IA, los enfoques críticos cuestionan las estructuras que la sostienen, y los enfoques pedagógico-humanistas enfatizan la dimensión formativa de la ética.

No obstante, la literatura también muestra puntos de convergencia, particularmente en el reconocimiento de que la inteligencia artificial debe mantenerse subordinada a los fines educativos y al respeto por la

dignidad humana. Esta convergencia permite sostener que la ética de la IA en la educación superior requiere una articulación integrada entre principios, crítica social y práctica pedagógica.

Dilemas éticos recurrentes en el uso de la IA en la universidad

El análisis de la literatura especializada permite identificar una serie de dilemas éticos recurrentes asociados al uso de la inteligencia artificial en la educación superior. Estos dilemas no se presentan de manera aislada, sino que emergen de la interacción entre prácticas pedagógicas, decisiones institucionales y desarrollos tecnológicos. La literatura coincide en que dichos dilemas ponen en tensión principios fundamentales de la educación universitaria, como la integridad académica, la equidad, la autonomía y la centralidad de la persona.

Autoría académica e integridad intelectual ante el uso de la IA

Uno de los dilemas más debatidos en la literatura reciente se relaciona con la autoría académica y la integridad intelectual, particularmente a partir del uso de herramientas de inteligencia artificial generativa. Diversos estudios señalan que estas tecnologías cuestionan las nociones tradicionales de originalidad y producción del conocimiento. En este sentido, se advierte que “la inteligencia artificial generativa desafía las concepciones convencionales de autoría al producir textos que no pueden atribuirse de manera clara a un solo sujeto” (Kasneci et al., 2023, p. 4).

Desde una perspectiva ética, se ha documentado que el uso no regulado de estas herramientas puede debilitar los principios de honestidad académica. Al respecto, se sostiene que “la integridad académica en contextos de inteligencia artificial requiere redefinir qué constituye una contribución legítima del estudiante” (Eaton, 2023, p. 6). Este dilema revela la necesidad de transitar de enfoques punitivos hacia modelos formativos que orienten el uso ético de la IA en los procesos de aprendizaje.

Sesgos algorítmicos y equidad educativa en contextos universitarios

Otro dilema central identificado en la literatura se refiere a los sesgos algorítmicos y su impacto en la equidad educativa. Diversos autores han demostrado que los sistemas de inteligencia artificial pueden reproducir desigualdades sociales cuando se entrenan con datos históricos que reflejan patrones de exclusión. En este sentido, se ha señalado que “los algoritmos no son neutrales, sino que incorporan valores y sesgos presentes en los datos y en las decisiones de diseño” (O’Neil, 2020, p. 28).

En el contexto universitario, este dilema adquiere especial relevancia cuando la IA se utiliza para procesos de admisión, evaluación o predicción del desempeño académico. Investigaciones recientes advierten que “el uso de sistemas predictivos en educación puede reforzar desigualdades estructurales si no se someten a evaluaciones éticas y de impacto” (Williamson, 2021, p. 11). Esta problemática interpela directamente la misión social de la universidad y su compromiso con la justicia educativa.

Privacidad, protección de datos y prácticas de vigilancia académica

La recopilación masiva de datos constituye otro dilema ético relevante en el uso universitario de la inteligencia artificial. La literatura señala que muchas aplicaciones de IA dependen del análisis de grandes volúmenes de información personal del estudiantado, lo cual plantea riesgos asociados a la privacidad y la vigilancia. Al respecto, se ha documentado que “la expansión de sistemas de análisis de datos educativos incrementa la posibilidad de prácticas de vigilancia que afectan la autonomía y la confianza en el entorno universitario” (Selwyn, 2022, p. 63).

Desde una perspectiva ética, se advierte que la protección de datos no debe entenderse únicamente como un requisito legal, sino como una condición para preservar la dignidad y los derechos de las personas. En este sentido, se sostiene que “la gobernanza ética de los datos educativos es un componente esencial del uso responsable de la inteligencia artificial” (UNESCO, 2021, p. 15).

Dependencia tecnológica y riesgos de deshumanización del proceso educativo

Un dilema adicional identificado en la literatura es el riesgo de dependencia tecnológica y la posible deshumanización de los procesos educativos. Algunos estudios advierten que una confianza excesiva en sistemas automatizados puede debilitar la relación pedagógica y reducir el papel del docente a funciones de supervisión técnica. Se ha señalado que “la automatización de decisiones educativas puede desplazar el juicio pedagógico y erosionar la dimensión relacional de la enseñanza” (Biesta, 2021, p. 31).

Se pone de manifiesto la necesidad de mantener a la inteligencia artificial como un medio al servicio de la educación, y no como un sustituto de la interacción humana. Desde esta perspectiva, la ética del uso de la IA implica reflexionar sobre los límites de la automatización y el valor insustituible de la experiencia educativa compartida.

Contrastes y tensiones entre posturas académicas sobre la ética de la IA en la educación superior

La revisión documental evidencia que el debate sobre el uso ético de la inteligencia artificial en la educación superior se caracteriza por la coexistencia de posturas divergentes, las cuales responden a distintas concepciones sobre el papel de la universidad, la tecnología y la formación humana. Estas tensiones no deben entenderse como incompatibilidades absolutas, sino como expresiones de un campo en construcción donde confluyen perspectivas normativas, críticas y pedagógicas.

Entre la regulación técnica y la formación ética: Enfoques en tensión

Una de las tensiones más frecuentes identificadas en la literatura se presenta entre los enfoques que privilegian la regulación técnica del uso de la IA y aquellos que subrayan la formación ética como eje central.

Desde la primera postura, se sostiene que el establecimiento de marcos normativos y principios claros constituye la base para un uso responsable de la inteligencia artificial. En este sentido, se afirma que “la definición de principios éticos comunes permite orientar el desarrollo y la implementación de sistemas de inteligencia artificial en contextos educativos” (Floridi et al., 2020, p. 692).

En contraste, otros autores advierten que la sola formulación de principios resulta insuficiente si no se acompaña de procesos formativos y culturales dentro de las instituciones. Se ha señalado que “los principios éticos carecen de eficacia cuando no se traducen en prácticas institucionales y en capacidades críticas de los actores educativos” (Mittelstadt, 2021, p. 504). Esta tensión pone en evidencia la necesidad de articular la regulación con la educación ética, evitando enfoques reduccionistas.

Innovación tecnológica, eficiencia institucional y justicia educativa

Otra tensión relevante se manifiesta entre las narrativas que promueven la inteligencia artificial como motor de innovación y eficiencia y aquellas que alertan sobre sus efectos en la justicia educativa. Desde una perspectiva tecnocéntrica, se argumenta que la IA permite optimizar procesos educativos y administrativos, mejorando la calidad institucional. Al respecto, se ha documentado que “las tecnologías basadas en datos ofrecen nuevas oportunidades para la toma de decisiones informadas en la educación superior” (Zawacki-Richter et al., 2020, p. 9).

Sin embargo, las posturas críticas cuestionan esta narrativa al señalar que la eficiencia no garantiza equidad. Se advierte que “los sistemas algorítmicos utilizados en educación pueden reforzar desigualdades sociales si se implementan sin una evaluación ética rigurosa” (Benjamin, 2020, p. 18). Este contraste revela un dilema central: la universidad debe decidir si la innovación tecnológica se orienta únicamente a la optimización de resultados o si se subordina a principios de justicia social e inclusión.

Control institucional y autonomía académica en el uso de la IA

La literatura también evidencia una tensión entre el control institucional del uso de la IA y la autonomía académica de docentes y estudiantes. Algunas investigaciones subrayan la necesidad de establecer políticas claras para regular el uso de herramientas de inteligencia artificial, particularmente en contextos de evaluación y producción académica. En este sentido, se sostiene que “las instituciones deben asumir un rol activo en la definición de límites y responsabilidades en el uso de la IA” (Eaton, 2023, p. 7).

No obstante, otros autores advierten que un control excesivo puede afectar la autonomía académica y la libertad pedagógica. Se ha señalado que “la gobernanza algorítmica en educación corre el riesgo de sustituir el juicio profesional por decisiones automatizadas” (Selwyn, 2022, p. 71). Esta tensión obliga a repensar modelos de gobernanza que equilibren la responsabilidad institucional con el respeto a la autonomía universitaria.

Convergencias en torno a la centralidad de la persona en la educación universitaria

A pesar de las tensiones identificadas, la revisión documental muestra convergencias significativas entre las distintas posturas académicas. En particular, existe un consenso creciente en torno a la necesidad de mantener la centralidad de la persona en el uso de la inteligencia artificial en la universidad. Desde distintos enfoques se coincide en que “la inteligencia artificial debe estar subordinada a los fines educativos y al respeto por la dignidad humana” (UNESCO, 2021, p. 10).

Esta convergencia permite sostener que la ética de la IA en la educación superior no se reduce a una dicotomía entre tecnología y humanismo, sino que requiere una integración crítica que reconozca el potencial de la innovación sin renunciar a los valores fundamentales de la educación universitaria.

Síntesis crítica y posicionamiento ético en el contexto universitario

El análisis documental realizado a lo largo de este capítulo pone de manifiesto que el debate sobre el uso ético de la inteligencia artificial en la educación superior se ha desarrollado desde enfoques diversos, los cuales, si bien comparten preocupaciones comunes, difieren en sus prioridades y alcances. Desde los enfoques normativos, se sostiene que la ética de la inteligencia artificial debe apoyarse en principios orientadores, en tanto que “la formulación de marcos éticos compartidos resulta esencial para garantizar que el desarrollo y la aplicación de la inteligencia artificial se orienten al bien común” (Floridi et al., 2020, p. 692). En la misma línea, se enfatiza que “los sistemas de inteligencia artificial deben diseñarse y utilizarse de manera que respeten la dignidad humana, la equidad y los derechos fundamentales” (UNESCO, 2021, p. 8).

Por su parte, las perspectivas críticas advierten sobre los riesgos estructurales asociados a la automatización y al uso de sistemas algorítmicos en contextos educativos, particularmente en lo relativo a la reproducción de desigualdades y a la concentración del poder. Desde esta postura, se señala que “las tecnologías algorítmicas tienden a reforzar desigualdades preexistentes cuando se insertan en estructuras sociales desiguales” (Benjamin, 2020, p. 18), así como que “el uso de datos y sistemas automatizados en educación plantea riesgos significativos para la autonomía institucional y el control democrático” (Williamson y Eynon, 2020, p. 229).

De manera complementaria, los enfoques pedagógico-humanistas sitúan la ética de la inteligencia artificial en relación directa con el sentido formativo de la universidad, subrayando que la incorporación de tecnologías emergentes debe evaluarse a la luz de su impacto en la formación integral del estudiantado y en la relación pedagógica. Desde esta perspectiva, se sostiene que la educación universitaria no puede reducirse a criterios de eficiencia técnica sin comprometer su dimensión ética y social, en tanto que “la educación debe orientarse a la formación de sujetos responsables en el mundo” (Biesta, 2021, p. 29).

Si bien estos enfoques aportan elementos valiosos para comprender la complejidad del fenómeno, la revisión documental permite identificar

vacíos y tensiones no resueltas en el debate actual. En particular, se observa que los marcos normativos tienden a centrarse en la formulación de principios generales, pero ofrecen orientaciones limitadas sobre su implementación pedagógica e institucional. Al respecto, se ha advertido que “los principios éticos, por sí solos, no garantizan prácticas éticas efectivas en contextos organizacionales complejos” (Mittelstadt, 2021, p. 504). De igual forma, las perspectivas críticas, aunque fundamentales para visibilizar riesgos y desigualdades, no siempre proponen estrategias claras para integrar la inteligencia artificial de manera formativa en la vida universitaria (Benjamin, 2020, p. 20).

A partir de estos hallazgos, el análisis sugiere la necesidad de articular los enfoques éticos existentes en una perspectiva integradora que reconozca la complejidad del contexto universitario. Esta articulación implica comprender la ética de la inteligencia artificial no únicamente como un conjunto de normas externas, sino también como un proceso formativo e institucional. Desde una perspectiva educativa, se sostiene que “la alfabetización ética y digital constituye una responsabilidad central de las instituciones de educación superior, en tanto permite a los estudiantes desarrollar una comprensión crítica del uso de tecnologías emergentes” (Ramírez-Montoya, 2022, p. 5). De manera complementaria, se enfatiza que “la integridad académica en contextos de inteligencia artificial requiere una aproximación formativa que involucre tanto a estudiantes como a docentes” (Eaton, 2023, p. 6).

Desde esta perspectiva, la articulación ética propuesta se sustenta en tres ejes complementarios. En primer lugar, la responsabilidad institucional en la definición de políticas claras y transparentes sobre el uso de la inteligencia artificial, dado que “las instituciones educativas tienen un papel clave en la regulación y orientación ética del uso de tecnologías emergentes” (UNESCO, 2021, p. 10). En segundo lugar, la formación ética del estudiantado y del profesorado como parte de la misión educativa universitaria, en tanto que la inteligencia artificial “debe mantenerse subordinada a los fines educativos y al respeto por la dignidad humana” (Floridi et al., 2020, p. 694). Finalmente, el reconocimiento de la centralidad de la persona como principio orientador del desarrollo tecnológico articula estos ejes y refuerza la necesidad de un enfoque ético-pedagógico e institucional integrado.

En consecuencia, el análisis desarrollado en este capítulo permite sostener que el uso ético de la inteligencia artificial en la educación superior no puede abordarse desde enfoques fragmentados o exclusivamente técnicos. Por el contrario, requiere una visión integral que articule principios éticos, reflexión crítica y práctica pedagógica, reconociendo a la universidad como un espacio privilegiado para la formación de sujetos capaces de interactuar con la tecnología de manera consciente, responsable y socialmente comprometida.

Conclusiones

El análisis desarrollado a lo largo de este capítulo permitió examinar el uso ético de la inteligencia artificial en la educación superior desde una revisión documental crítica, evidenciando la coexistencia de enfoques normativos, críticos y pedagógico-humanistas en la literatura especializada. En conjunto, estos enfoques coinciden en reconocer que la incorporación de la inteligencia artificial en el ámbito universitario plantea desafíos éticos significativos que no pueden ser abordados únicamente desde una perspectiva técnica o instrumental. En este sentido, se reafirma que “la ética de la inteligencia artificial debe orientarse al respeto de la dignidad humana y al bien común” (Floridi et al., 2020, p. 689), principio que atraviesa de manera transversal los estudios revisados.

La revisión documental permitió identificar que los enfoques normativos han contribuido de manera relevante a la formulación de principios y marcos éticos orientadores, particularmente en lo relativo a la transparencia, la equidad y la rendición de cuentas. No obstante, también se evidenció que dichos marcos presentan limitaciones cuando no se acompañan de procesos formativos e institucionales que garanticen su implementación efectiva. Tal como se advierte en la literatura, “la existencia de principios éticos no asegura por sí misma prácticas éticas en contextos organizacionales complejos” (Mittelstadt, 2021, p. 504), lo cual resulta especialmente pertinente en el ámbito universitario.

Por su parte, las perspectivas críticas analizadas aportan una lectura indispensable sobre los riesgos estructurales asociados al uso de la inteligencia artificial, particularmente en relación con la reproducción

de desigualdades, los sesgos algorítmicos y la concentración del poder en sistemas automatizados. Estas posturas permiten problematizar la adopción acrítica de tecnologías emergentes, al señalar que “los sistemas algorítmicos tienden a reforzar desigualdades preexistentes cuando se insertan en estructuras sociales desiguales” (Benjamin, 2020, p. 18). Desde esta óptica, la ética de la inteligencia artificial en la educación superior exige una reflexión constante sobre justicia educativa e inclusión social.

Asimismo, el análisis de los enfoques pedagógico-humanistas subrayó la centralidad de la formación ética como parte de la misión universitaria, destacando que la incorporación de la inteligencia artificial debe evaluarse en función de su impacto en la formación integral del estudiantado y en la relación pedagógica. En este marco, se reconoce que “la educación universitaria no puede reducirse a criterios de eficiencia técnica sin comprometer su dimensión ética y social” (Biesta, 2021, p. 29), lo cual refuerza la necesidad de situar la ética de la IA en el corazón del proyecto educativo institucional.

A partir de la articulación crítica de estos enfoques, el capítulo permite concluir que el uso ético de la inteligencia artificial en la educación superior requiere una visión integrada que combine principios normativos, análisis crítico y práctica pedagógica. Esta integración supone reconocer la responsabilidad institucional en la definición de políticas claras, promover la formación ética y digital de estudiantes y docentes, y mantener la centralidad de la persona como eje orientador del desarrollo tecnológico. En consonancia con ello, se reafirma que “las instituciones de educación superior tienen la responsabilidad de orientar el uso de la inteligencia artificial hacia fines educativos y socialmente responsables” (UNESCO, 2021, p. 10).

En suma, el presente capítulo contribuye a la comprensión del debate contemporáneo sobre la ética de la inteligencia artificial en la universidad, al ofrecer una revisión documental crítica que evidencia tanto los aportes como las tensiones existentes en la literatura. Lejos de proponer respuestas cerradas, el análisis subraya la necesidad de continuar reflexionando sobre el papel de la inteligencia artificial en la educación superior desde una perspectiva ética, formativa y socialmente comprometida, reconociendo a la universidad como un espacio clave para la construcción de prácticas tecnológicas responsables.

Referencias

- Benjamin, R. (2020). *Race after technology: Abolitionist tools for the new Jim code*. Polity Press.
- Biesta, G. (2021). *World-centred education: A view for the present*. Routledge.
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.
- Eaton, S. E. (2023). *Academic integrity in the age of artificial intelligence*. University of Calgary Press.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... Vayena, E. (2020). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). ChatGPT for education: Opportunities and challenges. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Mittelstadt, B. (2021). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>
- O’Neil, C. (2020). *Weapons of math destruction: How big data increases inequality and threatens democracy (Updated ed.)*. Crown.
- Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing.
- Ramírez-Montoya, M. S. (2022). Ética, innovación educativa y transformación digital en la educación superior. *Revista de Educación a*

- Distancia*, 22(69), 1–6. <https://doi.org/10.6018/red.487021>
- Selwyn, N. (2022). *Should robots replace teachers? AI and the future of education*. Polity Press.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- UNESCO. (2021). *Recomendación sobre la ética de la inteligencia artificial*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.
- Williamson, B. (2021). Education, technology and the politics of data. *Educational Philosophy and Theory*, 53(12), 1–14. <https://doi.org/10.1080/00131857.2020.1855187>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(39), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>