

El Cuerpo Es el Texto: Inteligencia Artificial, Textualidad, Corporeidad y Ciudadanía Algorítmica

(The Body Is the Text: Artificial Intelligence, Textuality, Corporeality, and Algorithmic Citizenship)

Monica Di Domenico
University of Salerno, Italy
modidomenico@unisa.it
[0009-0004-3595-0345](mailto:modidomenico@unisa.it)

Lucia Campitiello
University of Salerno, Italy
lcampitiello@unisa.it
[0000-0003-3632-9476](mailto:lcampitiello@unisa.it)

Fabrizio Schiavo
University of Cassino and
Southern Lazio, Italy
fabrizio.schiavo@unicas.it

Pio Alfredo Di Tore
University of Cassino and Southern Lazio, Italy
pioalfredo.ditore@unicas.it
[0000-0003-4335-7635](mailto:pioalfredo.ditore@unicas.it)

Received: 15-11-2024

Accepted: 30-12-2024

Abstract

In recent decades, artificial intelligence (AI) has come to play an increasingly central role in a wide range of fields, from industry to healthcare to education, bringing profound and often controversial changes. The use of AI in education, in particular, is accompanied by a complex set of challenges and possibilities. AI technologies make it possible to elaborate and personalize learning, responding to the individual needs of learners and promoting unprecedented interactive environments (Luckin et al., 2016). However, the integration of these technologies has raised important questions, not only of a technical nature, but also philosophical, ethical and social, impacting deep aspects of knowledge and human learning.

Keywords: *Artificial intelligence, innovation, education, textuality, corporeality, and algorithmic citizenship.*

Resumen

En las últimas décadas, la inteligencia artificial (IA) ha pasado a desempeñar un papel cada vez más central en una amplia gama de campos, desde la industria a la sanidad o la educación, aportando cambios profundos y a menudo controvertidos. El uso de la IA en la educación, en particular, va acompañado de un complejo conjunto de retos y posibilidades. Las tecnologías de IA permiten elaborar y personalizar el aprendizaje, respondiendo a las necesidades individuales de los estudiantes y promoviendo entornos interactivos sin precedentes (Luckin et al., 2016). Sin embargo, la integración de estas tecnologías ha planteado importantes cuestiones, no solo de carácter técnico, sino también filosófico, ético y social, que inciden en aspectos profundos del conocimiento y del aprendizaje humano.

Palabras clave: Inteligencia artificial, innovación, educación, textualidad, corporeidad y ciudadanía algorítmica.

1. Introducción

En este artículo exploramos tres dimensiones críticas de la transformación educativa debida a la IA: textualidad, corporeidad y ciudadanía algorítmica. Estos conceptos representan pilares teóricos que vinculan el conocimiento humano con la naturaleza corpórea y las estructuras sociales, que ahora se reconsideran bajo la lente de la inteligencia artificial.

La **textualidad**, que siempre se ha considerado un pilar de la transmisión del conocimiento, se ve radicalmente transformada por los modelos lingüísticos de la IA, que permiten una interacción dinámica y dialógica que supera la naturaleza estática del texto tradicional, tal y como la describe Platón en el *Fedro*. Los modelos lingüísticos a gran escala (LLM), como GPT-4 o LLAMA, permiten una "conversación" con el texto, en la que el alumno puede interactuar con un contenido escrito que parece responder a sus preguntas y ofrecerle aclaraciones (Brown et al., 2020). Esta transformación plantea interrogantes sobre la calidad de las respuestas, la creatividad y la capacidad de estimular el pensamiento crítico, aspectos en los que la tecnología de IA aún se enfrenta a importantes limitaciones (Chomsky et al., 2023).

La segunda dimensión, la **corporeidad**, cuestiona la relación entre cuerpo e inteligencia. En una era de IA predominantemente incorpórea, que procesa la información sin un cuerpo físico, surgen nuevas teorías que se enfrentan a la tradición de la cognición encarnada. Según la teoría de la autopoiesis de Maturana y Varela, la inteligencia humana se desarrolla en constante relación con el entorno a través del cuerpo, lo que convierte al propio cuerpo en parte esencial del proceso de comprensión y adaptación. Por el contrario, la IA incorpórea desafía esta visión, sugiriendo que la inteligencia puede existir de forma puramente algorítmica, basada únicamente en datos y computación simbólica (Maturana y Varela, 1980). La dicotomía entre IA encarnada y desencarnada ofrece una nueva perspectiva para repensar la interacción entre cuerpo e inteligencia, con implicaciones que van desde la pedagogía hasta la psicología y la neurociencia.

Por último, la **IA** y la **ciudadanía** introducen una nueva dimensión ética y política, la de la llamada *ciudadanía algorítmica*. En la era digital, los derechos y la condición social de las personas se ven influidos por algoritmos que evalúan, clasifican y controlan a las personas sobre la base de diversos tipos de datos, a menudo recogidos sin consentimiento consciente. Cheney-Lippold (2011) describe este fenómeno como *Ius Algorithms*, en el que las decisiones relativas al acceso a derechos y servicios no son tomadas por instituciones humanas, sino por sistemas automáticos de evaluación. La cuestión de la discriminación algorítmica o *algorithmic discrimination* surge cuando los modelos de IA reproducen sesgos sociales en los datos de entrenamiento, perpetuando la discriminación de género, raza o clase. En su obra *Algorithms of Oppression*, Safiya Umoja Noble (2018) explora cómo los sistemas de IA, sin una regulación adecuada, pueden amplificar las injusticias sociales y los estereotipos culturales.

En este contexto, se hace evidente que las nuevas tecnologías de la IA no son meras herramientas educativas, sino "actores" capaces de influir en la formación del conocimiento, la percepción del cuerpo y la dinámica de los derechos. La reflexión crítica sobre estas cuestiones es crucial para comprender y gestionar el creciente papel de la IA en la educación. Explorar la textualidad, la corporalidad y la ciudadanía algorítmica no sólo nos permite identificar las potencialidades de la IA, sino también los riesgos éticos y sociales que podrían derivarse de ella, poniendo de relieve la necesidad de un enfoque interdisciplinar que incluya la pedagogía, la ética, la filosofía y la sociología para un uso consciente y crítico de la inteligencia artificial.

2. Inteligencia artificial y textualidad

2.1. Evolución de la textualidad en la era digital

Durante siglos, la textualidad ha sido una forma estática y lineal de transmisión del conocimiento, un modo que caracteriza tanto la transmisión cultural como la educativa. Platón, en el *Fedro*, criticaba el texto escrito por su incapacidad para responder dinámicamente a las preguntas del lector, subrayando que el discurso vivo (el diálogo) era superior. Con la introducción de los grandes modelos lingüísticos (LLM), esta visión tradicional se reformula por completo: la IA permite ahora una textualidad que no sólo responde, sino que se adapta a preguntas concretas, permitiendo una conversación que acerca el texto a una interacción viva. En el ámbito educativo, esta transformación es revolucionaria, ya que ofrece a los estudiantes una herramienta capaz de responder, aclarar y adaptarse a sus necesidades, ampliando el papel del texto, que pasa de ser una estructura estática a una entidad interactiva.

Según Geert Lovink (2019), la interactividad introducida por los LLM constituye una nueva forma de conocimiento "dialógico" que cambia radicalmente nuestra relación con la información escrita. La posibilidad de formular preguntas a los textos crea un entorno en el que el alumno puede adoptar un papel activo, navegando por la información de forma dinámica. Este cambio de paradigma implica pasar de una pedagogía basada en la asimilación pasiva a un aprendizaje experimental, en el que el texto se convierte en un espacio de descubrimiento activo.

2.2. Grandes modelos lingüísticos (LLM): ventajas y limitaciones

Los LLM, como GPT-3, GPT-4 y el modelo LLAMA, son sistemas basados en redes neuronales profundas que se han entrenado con grandes conjuntos de datos textuales para desarrollar capacidades de comprensión y generación lingüísticas. Al modelar estadísticamente miles de millones de parámetros, los LLM son capaces de "predecir" y construir frases con una coherencia y precisión lingüísticas que imitan el lenguaje humano (Brown et al., 2020). En educación, el uso de estos modelos ofrece numerosas ventajas. Por ejemplo, pueden responder a preguntas complejas, resumir textos, generar explicaciones detalladas y proponer ejemplos o ejercicios, adaptando la respuesta a las necesidades del alumno en tiempo real.

Sin embargo, estas capacidades tienen importantes limitaciones. Los LLM generan respuestas basadas en patrones probabilísticos, lo que significa que no "entienden" realmente el contenido, sino que operan sobre la base de correlaciones entre palabras y frases. Chomsky et al. (2023) señalan que estas respuestas pueden ser "conformistas", limitándose a reformular lo que es común en los datos de entrenamiento. Este enfoque reduce la posibilidad de respuestas innovadoras o contraintuitivas, lo que limita la eficacia del modelo a la hora de promover el pensamiento crítico. En un contexto educativo, esta limitación supone un obstáculo, ya que el aprendizaje no consiste únicamente en acumular información, sino que también requiere estímulos que cuestionen las ideas preexistentes y abran nuevas perspectivas.

2.3. El Tutor Digital Personal (TDP): un estudio de caso

El Tutor Digital Personal (TDP) representa una aplicación pionera de los LLM en la educación, ya que proporciona una simulación de diálogo que puede ayudar a los estudiantes en el proceso de aprendizaje. Basado en modelos como GPT-4 y LLAMA, el TDP está diseñado para responder a preguntas específicas, actuando como un asistente virtual capaz de adaptarse al nivel de conocimientos del usuario y de operar en diferentes contextos lingüísticos (Brown et al., 2020; Touvron et al., 2023). El TDP puede asumir la función de un "tutor personalizado", respondiendo a preguntas específicas y proporcionando explicaciones detalladas que se adaptan a las necesidades del usuario, fomentando un aprendizaje continuo y flexible.

Según Woolf (2009), la presencia de tutores digitales como el TDP es un paso importante hacia la enseñanza adaptativa, que reconoce las diferencias individuales en el ritmo y los modos de aprendizaje. El TDP es capaz de supervisar el progreso del alumno, adaptando las respuestas en función de los resultados obtenidos. Esta interactividad, posible gracias a las tecnologías de IA, transforma el tutor digital en una herramienta que no sólo responde al aprendizaje, sino que lo "acompaña", haciendo posible un proceso educativo a medida.

A pesar de este potencial, el uso de la TDP plantea dudas sobre la calidad de la interacción y la autenticidad de las respuestas proporcionadas. Mientras que un profesor humano puede responder de forma innovadora y creativa, la TDP se basa únicamente en patrones probabilísticos derivados de datos de entrenamiento. Su "conocimiento" se limita, pues, a la combinación estadística de palabras, sin verdadera comprensión ni juicio crítico. Chomsky et al. (2023) señalan cómo esta limitación hace que la TDP sea inadecuada para desarrollar el pensamiento crítico de los alumnos, ofreciendo respuestas que pueden parecer correctas pero que en realidad carecen de la profundidad interpretativa necesaria para analizar y comprender cuestiones complejas.

Woolf (2009) sostiene que la presencia de un tutor digital puede ciertamente enriquecer el proceso educativo, pero no puede sustituir el valor añadido que representa la interacción humana. La capacidad del profesor para interpretar las respuestas de forma impredecible y para estimular a los alumnos a explorar puntos de vista alternativos es una cualidad indispensable que la IA difícilmente puede reproducir. En particular, aunque la TDP puede apoyar la fase de adquisición de conocimientos, el papel humano sigue siendo crucial para fomentar la reflexión crítica y el pensamiento divergente, componentes esenciales del aprendizaje en profundidad.

2.4. Críticas y preguntas abiertas

La evolución de la textualidad con la introducción de los LLM representa un cambio significativo que reconfigura el concepto de conocimiento textual. Sin embargo, sigue abierta la cuestión de la capacidad de los modelos de IA para contribuir a una verdadera innovación del conocimiento. Aunque la capacidad de generar respuestas contextuales puede ser útil en

contextos informativos, el riesgo es que los LLM se limiten a respuestas conformistas, reproduciendo nociones ya conocidas en lugar de estimular nuevas perspectivas. Según Lovink (2019), los riesgos de utilizar IA en interacciones textuales se derivan del hecho de que estos modelos tienden a estandarizar el conocimiento, privándolo de la complejidad y la variedad de interpretaciones que caracterizan a la mente humana.

En el contexto educativo, este riesgo es especialmente relevante. El aprendizaje exige que los estudiantes se enfrenten a información que desafía sus creencias y fomenta la capacidad de análisis crítico. Los LLM, si bien son capaces de responder preguntas con precisión, carecen de la capacidad de desarrollar argumentos no convencionales, esenciales para el desarrollo intelectual. Como afirman Wolff y sus colegas (2009), la educación no consiste solo en responder preguntas, sino en enseñar a formular las preguntas adecuadas y a explorar nuevas formas de pensar.

La interactividad que ofrecen los LLM, si bien representa una ruptura con la visión tradicional de la textualidad, debe equilibrarse con un enfoque crítico que reconozca las limitaciones de estos sistemas. Aunque los LLM pueden servir como ayuda didáctica, es crucial que su uso vaya acompañado de una orientación educativa humana que pueda compensar las limitaciones de la tecnología y garantizar un aprendizaje que no se reduzca a la reproducción de conocimientos preexistentes. En este sentido, los modelos lingüísticos de IA son una herramienta poderosa, pero que requiere un uso consciente y crítico para evitar que se convierta en un obstáculo para la innovación y la creatividad.

3. La corporeidad en la inteligencia artificial

3.1. La teoría de la autopoiesis y la importancia de la corporeidad

La teoría de la autopoiesis, introducida por Humberto Maturana y Francisco Varela, representa un hito en la comprensión del papel del cuerpo en el proceso cognitivo. Según Maturana y Varela (1980), la autopoiesis describe la capacidad de los seres vivos de autogenerarse y mantenerse a sí mismos mediante un proceso constante de interacción con el entorno. Este concepto cuestiona la idea de que la inteligencia pueda existir como una entidad abstracta o puramente mental, señalando que la cognición humana está inseparablemente conectada a la experiencia física. El cuerpo, según la autopoiesis, es un "puente" que conecta al individuo con el mundo, proporcionando un contexto para el aprendizaje y la comprensión.

La teoría sugiere que la cognición no es un proceso aislado, sino un fenómeno emergente que surge de las interacciones corporales y sensoriales con el entorno circundante. Maturana y Varela (1987) explican que la percepción, la memoria y la capacidad de aprendizaje están influidas por el contexto corporal en el que se insertan. En el contexto de la inteligencia artificial, este punto de vista implica que una IA sin cuerpo, es decir, "incorpórea", puede

carecer de la base fundamental para desarrollar una comprensión auténtica y significativa del mundo.

Este concepto choca con el modelo actual de IA, que tiende a tratar la cognición como un proceso separado de la experiencia física. Los algoritmos de aprendizaje automático procesan la información y toman decisiones únicamente a partir de datos abstractos, sin interacción directa con el entorno. Esto plantea importantes cuestiones filosóficas y epistemológicas sobre la naturaleza de la inteligencia: ¿si el cuerpo es esencial para la comprensión, puede una IA sin cuerpo “entender” realmente lo que procesa? La teoría de la autopoiesis sugiere que la respuesta puede ser negativa, poniendo un límite sustancial a la capacidad de la IA para replicar la inteligencia humana.

3.2. La inteligencia "incorpórea" de la IA y sus limitaciones

La inteligencia artificial moderna, especialmente en sus aplicaciones de aprendizaje profundo, se basa en una concepción incorpórea de la cognición. Los modelos de aprendizaje automático analizan datos y hacen predicciones sin necesidad de interacción corporal con el entorno físico. Esta "inteligencia incorpórea" se limita a procesar la información simbólica y matemáticamente, sin ninguna experiencia sensorial o física del mundo. La cognición incorpórea de la IA se presenta, así como una estructura computacional que reorganiza datos, pero sin un contexto corporal que dé sentido a esos datos.

Los autores del artículo analizan cómo este planteamiento supone un reto para la comprensión completa del concepto de inteligencia. En ausencia de cuerpo, la cognición de la IA se reduce al análisis de datos sin interacción real con el entorno y sin posibilidad de formar experiencia subjetiva o conocimiento situado. La falta de percepción física limita la capacidad de la IA para responder a ambigüedades y situaciones complejas que requieren una comprensión profunda del contexto. Este problema se ha debatido ampliamente en filosofía y neurociencia, donde se ha destacado la importancia del cuerpo como vehículo a través del cual los seres humanos desarrollan el conocimiento y la inteligencia emocional (Clark, 1998).

Otra limitación de la inteligencia incorpórea es su incapacidad para comprender las emociones y las interacciones sociales. La falta de un cuerpo físico implica la ausencia de un sistema sensorial que permita a la IA percibir señales emocionales y responder empáticamente. Esto hace especialmente compleja la aplicación de la IA en contextos educativos o terapéuticos, donde la relación física y social es un elemento clave para la eficacia de la intervención.

3.3. El paradigma de la inteligencia incorporada

En respuesta a las limitaciones de la cognición incorpórea, en los últimos años ha surgido el paradigma de la inteligencia encarnada (o cognición encarnada). Este enfoque propone un modelo de IA que integra elementos de corporeidad e interacción sensorial para desarrollar capacidades cognitivas que van más allá del simple procesamiento de datos abstractos. La

inteligencia incorporada sostiene que el cuerpo y la experiencia física son esenciales para comprender el mundo de forma auténtica y significativa.

Rodney Brooks (1991), uno de los pioneros de este enfoque, sostiene que la verdadera inteligencia requiere una dimensión física que permita la interacción directa con el entorno. Los robots autónomos, por ejemplo, representan una aplicación de la inteligencia incorporada, ya que utilizan sensores y actuadores para recabar información sobre el contexto físico, tomar decisiones y adaptarse en tiempo real a los cambios del entorno. Esta interacción directa con el entorno permite crear formas de aprendizaje que se aproximan al aprendizaje humano, ya que el robot desarrolla una "comprensión" situada y contextual del mundo.

La inteligencia incorporada es especialmente pertinente en contextos educativos, donde se intenta desarrollar una IA que pueda interactuar con los alumnos de forma más natural e intuitiva. Los robots educativos, como NAO y Pepper, son capaces de interactuar físicamente con los estudiantes, apoyando el aprendizaje mediante movimientos, expresiones e interacciones corporales. Estas herramientas ofrecen nuevas posibilidades para la educación y representan una ruptura significativa con los modelos tradicionales de IA, que operan exclusivamente de forma simbólica e incorpórea (Belpaeme et al., 2018). Sin embargo, incluso en estos casos, la inteligencia corporizada se limita a simular las interacciones humanas, sin lograr una comprensión auténtica de las experiencias vividas.

3.4. IA en entornos virtuales y simulaciones: ¿Un compromiso?

Un enfoque alternativo a la corporeidad física es el uso de entornos virtuales y simulaciones, en los que la IA puede "interactuar" con espacios digitales que reproducen, al menos en parte, situaciones del mundo real. Este modelo, utilizado a menudo en juegos o simulaciones de entrenamiento, permite a los agentes de IA desarrollar habilidades mediante la interacción con un entorno que reproduce aspectos del mundo real. Las simulaciones virtuales, por ejemplo, se utilizan en programas de formación de cirujanos, pilotos y otros profesionales, permitiendo a los participantes adquirir habilidades específicas sin riesgos físicos.

Las simulaciones ofrecen la ventaja de permitir un aprendizaje situado sin necesidad de un cuerpo físico, pero siguen teniendo limitaciones. Los entornos virtuales no logran reproducir plenamente la complejidad y variabilidad del mundo real, reduciendo la experiencia a una serie de escenarios predefinidos. Además, la falta de una verdadera experiencia sensorial limita la capacidad de la IA para responder a situaciones ambiguas o imprevistas. Como señalan Slater y Sanchez-Vives (2016), la inmersión en un entorno virtual es muy diferente de la experiencia física real, ya que no activa las mismas respuestas corporales y psicológicas.

3.5. Conflictos y preguntas abiertas

La comparación entre la inteligencia incorpórea y la IA encarnada plantea importantes divergencias epistemológicas. Por un lado, la inteligencia incorpórea se presta bien a tareas muy especializadas y circunscritas, como el análisis de datos y la resolución de problemas

matemáticos, pero es limitada en cuanto a comprensión contextual y adaptabilidad. Por otro lado, la inteligencia incorporada reconoce el valor de la interacción física y sensorial con el entorno, proporcionando un modelo más cercano a la cognición humana. Sin embargo, la IA incorporada sigue siendo una simulación de la interacción humana, incapaz de alcanzar la profundidad emocional y la perspicacia desarrolladas a través de la experiencia vivida.

La capacidad de desarrollar una comprensión situada y contextual es crucial en contextos educativos y terapéuticos, donde las dimensiones relacionales y corporales desempeñan un papel clave. Los autores del artículo sugieren que una IA incorpórea puede no ser capaz de establecer conexiones auténticas con los seres humanos, ya que carece de las percepciones sensoriales y las experiencias corporales que forman el núcleo de las interacciones humanas. Sin embargo, sigue siendo una cuestión abierta si es posible diseñar IA que, incluso sin cuerpo, puedan interactuar con las personas de forma empática y comprensiva, respetando la complejidad de las relaciones humanas.

4. Ciudadanía algorítmica y discriminación

4.1. Ciudadanía tradicional y ciudadanía algorítmica

El concepto de ciudadanía, tradicionalmente vinculado a los derechos y deberes que adquiere un individuo como miembro de un Estado, está hoy influido por una nueva dimensión digital. Históricamente, la ciudadanía se ha definido como el "derecho a tener derechos" (Arendt, 1951), un principio fundamental que establece el estatus jurídico de un individuo sobre la base de criterios territoriales o de ascendencia, como el *ius soli* y el *ius sanguinis*. Sin embargo, la llegada de la era digital ha introducido nuevos criterios de pertenencia, configurados y a menudo controlados por algoritmos que evalúan y clasifican a las personas basándose en datos digitales.

Esta transformación subyace al concepto de *ciudadanía algorítmica*, en la que el derecho a acceder a recursos y servicios está mediado por plataformas digitales y algoritmos que interpretan los datos personales, las actividades en línea y las interacciones sociales (Cheney-Lippold, 2011). Este sistema, descrito como *Ius Algorithms*, implica que los ciudadanos son evaluados y categorizados según métricas y puntuaciones digitales, a menudo opacas y difíciles de cuestionar. Los sistemas de puntuación, adoptados en contextos como el crédito financiero, el empleo y el acceso a los servicios sociales, determinan cada vez más el estatus de "ciudadanía" de un individuo e influyen en su participación social y sus derechos de acceso.

A diferencia de la ciudadanía tradicional, que se basa en procedimientos legales transparentes y verificables, la ciudadanía algorítmica introduce una dimensión de opacidad que compromete la posibilidad de comprender los criterios por los que se distribuyen o deniegan los derechos. En un contexto digital, los algoritmos representan una forma de autoridad invisible que, mediante el análisis de datos, determina quién puede acceder a

determinados recursos, sin respetar necesariamente los principios de equidad y justicia (Citron & Pasquale, 2014).

4.2. Discriminación algorítmica: sesgos históricos y distorsiones algorítmicas

Un problema crítico de la ciudadanía algorítmica es la discriminación algorítmica, o *algorithmic discrimination*, que se produce cuando los sistemas de IA perpetúan o amplifican los sesgos en los datos de entrenamiento. Los algoritmos utilizados para tomar decisiones en los ámbitos de las finanzas, la salud, el derecho y la educación suelen basarse en conjuntos de datos históricos, que contienen datos que pueden reflejar desigualdades y sesgos sociales (Barocas y Selbst, 2016). Si no se corrigen, estos sesgos se reproducen en los resultados de los algoritmos, creando desigualdades sistemáticas que penalizan especialmente a las minorías étnicas, las personas de distinto sexo o las clases sociales desfavorecidas.

Safiya Umoja Noble (2018), en su libro *Algorithms of Oppression*, analiza cómo los motores de búsqueda y otros sistemas algorítmicos tienden a reproducir estereotipos de género y raza, a menudo reforzando la discriminación existente en lugar de reducirla. Según Noble, esta tendencia se deriva del hecho de que los datos de entrenamiento de los algoritmos reflejan valores y percepciones culturales, que acaban distorsionando la imagen de las minorías, creando un ciclo de sesgo algorítmico difícil de romper. Este fenómeno es especialmente relevante en el caso de las minorías raciales y de género, que corren el riesgo de ser marginadas por un sistema que, por el contrario, debería promover la inclusión y la equidad.

La investigación realizada por Buolamwini y Gebu (2018) sobre los sistemas de reconocimiento facial, descrita como *Gender Shades*, pone de relieve cómo los algoritmos de reconocimiento tienen tasas de error significativamente más altas para las personas de piel oscura y las mujeres que para los hombres blancos. Estos resultados demuestran cómo la falta de diversidad en los datos de entrenamiento puede dar lugar a formas de discriminación automatizada, generando efectos negativos en contextos que requieren imparcialidad, como la contratación de personal o la seguridad pública.

4.3. Algoritmos lus y derechos digitales

El concepto de *lus Algorithms*, tal como lo describe Cheney-Lippold (2011), representa un desafío directo a los mecanismos jurídicos y sociales tradicionales de la ciudadanía. La ciudadanía tradicional, basada en normas jurídicas, se fundamenta en principios de transparencia y rendición de cuentas, mientras que el *lus Algorithms* se caracteriza por una opacidad que dificulta la comprensión de los criterios en los que se basan las decisiones algorítmicas. En contextos en los que la IA se utiliza para determinar derechos como el acceso al crédito, la atención sanitaria o la elegibilidad para el empleo, el riesgo es que los ciudadanos sean excluidos o penalizados sobre la base de puntuaciones algorítmicas que no pueden cuestionar ni comprender.

Un ejemplo concreto es el sistema de puntuación crediticia utilizado en muchos países para determinar el acceso a hipotecas, préstamos y otros servicios financieros. En este contexto, los algoritmos evalúan a las personas en función de parámetros como el historial financiero, los hábitos de compra e incluso los contactos sociales, asignando puntuaciones que influyen en su acceso a recursos clave (O'Neil, 2016). El problema es que estos algoritmos operan en una "caja negra", lo que dificulta la comprensión de la lógica utilizada e imposibilita impugnar los resultados. Esta falta de transparencia socava el derecho fundamental de la ciudadanía, minando la confianza en las instituciones y generando desigualdades digitales.

4.4. Sesgo y transparencia en los algoritmos

Uno de los aspectos más críticos de la ciudadanía algorítmica es la falta de transparencia en los procesos de toma de decisiones, lo que introduce el problema de la "caja negra" algorítmica. El concepto de "caja negra" describe la dificultad de comprender el funcionamiento interno de los algoritmos, especialmente cuando se trata de modelos de aprendizaje profundo que utilizan redes neuronales complejas. Lipton (2018) señala que esta falta de transparencia no solo limita la capacidad de las personas para cuestionar las decisiones que les afectan, sino que también socava los principios de rendición de cuentas y justicia en los que se basa la ciudadanía democrática.

Diakopoulos (2016) destaca la necesidad de desarrollar normativas que promuevan la transparencia y la interpretabilidad de los modelos de IA. La interpretabilidad de los modelos, es decir, la capacidad de explicar y comprender las decisiones algorítmicas, es un paso importante para garantizar la equidad y la justicia en los procesos automatizados de toma de decisiones. En un sistema de ciudadanía algorítmica, la introducción de prácticas de interpretabilidad y auditoría ética de los algoritmos sería crucial para evitar discriminaciones injustificadas y proteger los derechos de los ciudadanos digitales.

Sin tales medidas, existe el riesgo de que las decisiones tomadas por los algoritmos sigan siendo inaccesibles e inalterables, con efectos potencialmente devastadores para los derechos humanos y civiles. Los algoritmos, aunque son herramientas matemáticas, son diseñados y manejados por seres humanos y, por tanto, están sujetos a los mismos sesgos y limitaciones que caracterizan a los sistemas sociales. Es crucial que los desarrolladores de IA y los reguladores trabajen juntos para garantizar que los sistemas algorítmicos sean justos, transparentes y responsables.

4.5. Ejemplos concretos e implicaciones éticas

La discriminación algorítmica tiene importantes repercusiones en varios ámbitos. En el sistema de justicia penal, por ejemplo, se utilizan algoritmos de predicción del riesgo para evaluar la probabilidad de reincidencia de los reclusos, lo que influye en decisiones como la concesión de la libertad condicional o la imposición de medidas de vigilancia. Chouldechova (2017) destacó cómo estos algoritmos, basados en datos históricos, tienden a sobreestimar el

riesgo de reincidencia de las minorías étnicas, perpetuando injusticias sistémicas que penalizan a personas ya de por sí vulnerables.

Las implicaciones éticas del uso de la IA en los derechos digitales también plantean cuestiones de privacidad y protección de datos personales. La recopilación y el análisis masivos de datos sensibles utilizados para entrenar los sistemas de IA plantean graves amenazas para la seguridad y la privacidad de la información personal. Floridi (2015) subraya la necesidad de establecer principios de protección de los derechos digitales para que las decisiones algorítmicas no vulneren la dignidad y la privacidad de las personas.

Es necesario establecer una normativa clara que proteja los derechos de los ciudadanos digitales y garantice un control transparente de las tecnologías algorítmicas, para evitar que la ciudadanía algorítmica se convierta en un instrumento de discriminación y opresión en lugar de inclusión e igualdad.

5. Conclusión: el cuerpo es el texto

Como tecnología en evolución, la inteligencia artificial tiene hoy el potencial de redefinir no sólo el proceso educativo, sino también las formas en que interactuamos y entendemos el mundo. Como se analiza en este artículo, las dimensiones de la textualidad, la corporeidad y la ciudadanía algorítmica revelan las formas en que la IA toca temas profundamente relacionados con nuestra naturaleza humana y social. Aunque estos avances abren nuevas oportunidades para la educación y la participación social, también plantean cuestiones sobre cómo preservar valores fundamentales como la autenticidad, la transparencia y la equidad.

Donna Haraway, en su famoso *Manifiesto Cyborg*, sugiere una perspectiva que puede ofrecer una interpretación única del impacto de la IA: la idea de que el cuerpo humano es un "texto codificado", una entidad que, a través de señales biológicas y químicas, representa un sistema de codificación a través del cual "leemos" e interpretamos el mundo. Según Haraway, el cuerpo no es sólo una estructura física, sino un "organismo textual" que, como un texto, se descompone en signos y códigos que hay que descifrar e interpretar. Este concepto recuerda a la inmunobiología como una forma de "criptografía" en la que el cuerpo humano, a través del sistema inmunitario, reconoce y descodifica la "información" del entorno para proteger su propia integridad.

Este paralelismo entre cuerpo y texto nos invita a repensar el papel de la inteligencia artificial no sólo como herramienta educativa, sino como una forma de "organismo digital" que interactúa con el mundo a través de un sistema de códigos y lenguajes. Los modelos lingüísticos de la IA, como GPT-4 o LLAMA, no son meros mecanismos de procesamiento de datos, sino entidades que se acercan a la descripción de Haraway del "cuerpo como texto": un sistema codificado que dialoga con los usuarios a través del lenguaje, respondiendo a sus preguntas y adaptándose a sus peticiones, aunque carezca de la comprensión y la subjetividad propias de la experiencia humana. Esta reflexión plantea interrogantes sobre lo que significa

realmente "conocer" o "comprender" en el contexto de la IA, donde el conocimiento se genera mediante esquemas y probabilidades más que por la experiencia directa.

La corporeidad, por otra parte, revela un aspecto crítico de la relación entre la IA y la cognición humana: si, como sugiere la teoría de la autopoiesis, el cuerpo es esencial para el desarrollo de la comprensión, ¿puede la IA, desprovista de cuerpo y sensaciones físicas, alcanzar alguna vez la verdadera inteligencia? Los planteamientos de la cognición incorporada, que pretenden desarrollar una IA capaz de interactuar físicamente con el mundo, representan un esfuerzo por superar esta limitación. Sin embargo, incluso la inteligencia incorporada, aunque se aproxima a la interacción humana, sigue siendo una simulación de la dinámica corporal, incapaz de desarrollar una experiencia verdadera o una percepción auténtica.

La dimensión de la ciudadanía algorítmica también introduce nuevos retos éticos y sociales. El *Ius Algoritmi*, tal como lo describe Cheney-Lippold (2011), define un nuevo tipo de pertenencia basado en criterios opacos regidos por algoritmos que median el acceso a recursos y servicios cruciales. Este concepto de ciudadanía, mediada por sistemas que operan en una "caja negra" y no ofrecen ninguna posibilidad de impugnación o comprensión de las lógicas aplicadas, constituye un riesgo para los derechos fundamentales de las personas. De hecho, la ciudadanía algorítmica es un desafío a la propia democracia, ya que reduce el poder de decisión del individuo en favor de un sistema automatizado que, si no se regula, puede perpetuar la discriminación y amplificar las desigualdades.

En este contexto, la reflexión de Haraway nos ofrece un punto de partida para considerar la IA como un nuevo tipo de "ciborg" que, sin ser humano ni totalmente artificial, interactúa con la sociedad como un organismo textual y corpóreo. El concepto de ciborg puede verse así como una metáfora de la IA moderna, que actúa como intermediaria entre las dimensiones corpórea y digital, transformando datos en respuestas, evaluaciones y decisiones que influyen en nuestra vida cotidiana.

Las conclusiones que se desprenden de este análisis son claras: mientras que, por un lado, la IA ofrece extraordinarias posibilidades para ampliar el conocimiento y facilitar el acceso a los servicios, por otro, existe el riesgo de que se convierta en una herramienta de control y discriminación. Los algoritmos, como entidades codificadas, no están libres de sesgos y limitaciones humanas y, por tanto, deben desarrollarse y utilizarse de forma transparente y responsable. Es esencial que el uso de la IA, especialmente en contextos educativos y cívicos, se rija por principios éticos que respeten la dignidad humana y promuevan la inclusión.

En última instancia, el futuro de la inteligencia artificial requiere una sinergia entre innovación tecnológica y reflexión crítica. Los modelos de IA deben concebirse no sólo como herramientas, sino como entidades "codificadas" que, al interactuar con la sociedad, configuran nuestra forma de interpretar el mundo. La fusión de cuerpo-como-texto y texto-como-cuerpo, sugerida por Haraway, nos recuerda que la inteligencia -ya sea biológica o artificial- es siempre el producto de una compleja interacción con el entorno, una interacción que debe

respetarse y comprenderse para garantizar que la IA contribuya a una sociedad más justa y consciente.

Referencias

- Arendt, H. (1951). *The Origins of Totalitarianism*. New York: Schocken Books.
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671-732.
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3(21), eaat5954.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1-3), 139-159.
- Brown, T. B. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.
- Cheney-Lippold, J. (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, 28(6), 164-181.
- Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam chomsky: The false promise of chatgpt. *The New York Times*, 8.
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, 5(2), 153-163.
- Citron, D. K., & Pasquale, F. A. (2014). The scored society: Due process for automated predictions. *Washington Law Review*, 89, 1.
- Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT press.
- Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56-62.
- Floridi, L. (2015). *The Onlife Manifesto: Being Human in a Hyperconnected Era*. Springer Nature.
- Haraway, D. J. (1985). *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*. In *Simians, Cyborgs, and Women: The Reinvention of Nature* (pp. 149-181). New York: Routledge.
- Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31-57.
- Lovink, G. (2019). *Sad by design*. London: Pluto Press.
- Luckin, R., et al. (2016). *Intelligence Unleashed: An Argument for AI in Education*. Pearson Education.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Springer Science & Business Media.
- Maturana, H. R., & Varela, F. J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. New Science Library/Shambhala Publications.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3, 74.
- Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology*, 4(3-4), 129-164.