

Inteligencia artificial y ética del lenguaje: del atomismo lógico a las redes neuronales artificiales

Sara Eugenia Madera Gómez

Universidad de Milán

sarymadera@hotmail.com

<https://orcid.org/0000-0003-0164-4695>

Introducción

La inteligencia artificial (IA) ha revolucionado numerosos campos, desde la medicina hasta la economía, planteando simultáneamente nuevos y complejos desafíos éticos. En este artículo, defiendo una tesis central: los problemas éticos fundamentales asociados con la IA generativa no son primariamente un problema de algoritmos, sino que derivan directamente de los conjuntos de datos (datasets) utilizados para su entrenamiento. Más aún, sostengo que esto es, en su esencia, un problema del lenguaje.

Mi argumento se basa en una analogía clave: así como el atomismo lógico de Bertrand Russell propone que el lenguaje debe reflejar la estructura de los hechos que describe, los datasets utilizados en el entrenamiento de IA funcionan como un ‘lenguaje’ que intenta representar la realidad. Sin embargo, al igual que el lenguaje puede ser impreciso o sesgado, los datasets pueden contener errores, sesgos y limitaciones que se transmiten directamente a los modelos de IA, generando consecuencias éticas significativas.

Esta perspectiva nos permite reinterpretar los desafíos éticos de la IA no como problemas técnicos aislados, sino como manifestaciones de las limitaciones inherentes a nuestros intentos de representar la realidad a través del lenguaje; en este caso, el lenguaje de los datos. Al reinterpretar los desafíos éticos de la IA a través de esta lente filosófica, este artículo no solo ofrece una nueva perspectiva sobre problemas ya tratados, sino que también proporciona un marco conceptual para abordar estos desafíos de manera más efectiva y contextual.

El presente artículo se estructura de la siguiente manera: primero, estableceré el marco teórico, explorando los principios del atomismo lógico y los fundamentos de la creación de datasets para IA. Luego, desarrollaré la analogía entre el atomismo lógico y la estructura de los datasets, mostrando cómo los datos individuales funcionan como ‘hechos atómicos’ y las etiquetas como ‘proposiciones atómicas’. A continuación, analizaré las implicaciones éticas de esta perspectiva, examinando cómo los sesgos y limitaciones en los datasets se traducen en problemas éticos concretos en las aplicaciones de IA.

Finalmente, propondré algunas direcciones para futura investigación y desarrollo, enfatizando la necesidad de un enfoque interdisciplinario que combine la filosofía del lenguaje, la ética y la ciencia de datos.

Definiciones y conceptos clave

Antes de profundizar en las implicaciones filosóficas y éticas, es crucial establecer una comprensión clara de los términos centrales que se utilizarán en el presente artículo:

- Inteligencia artificial (IA)

Campo de la informática enfocado en crear sistemas capaces de realizar tareas que típicamente requieren inteligencia humana.

- *Machine learning*

Una subcategoría de la IA que implica el desarrollo de algoritmos que permiten a las computadoras aprender de los datos y hacer predicciones o tomar decisiones basadas en esos datos.

- *Dataset*

Grandes conjuntos de datos que se utilizan para entrenar y probar algoritmos y modelos. Estos datasets emplean ‘etiquetas’ (labels) para categorizar la información.

- Red neuronal artificial

Modelos computacionales inspirados en el cerebro humano que se entrenan utilizando datasets. La red procesa estos datos para aprender patrones y hacer predicciones o clasificaciones (un proceso conocido como Deep Learning). La analogía neuronal, aunque poderosa, puede ser engañosa si se toma demasiado literalmente.

Atomismo lógico

El atomismo lógico, propuesto por Bertrand Russell y desarrollado por Ludwig Wittgenstein, representa un intento ambicioso de desentrañar la relación fundamental entre lenguaje y realidad. En esencia, esta teoría propone que existe una correspondencia estructural entre nuestras expresiones lingüísticas y los hechos del mundo. Russell (1914, p.8) argumentaba que “el lenguaje debe tener una estructura que refleje la estructura de los hechos que describe”. Esta afirmación, aparentemente simple, encierra profundas implicaciones filosóficas que nos servirán como punto de partida para nuestra analogía con los datasets en IA. El atomismo lógico se fundamenta en tres principios interrelacionados que merecen un análisis detallado.

- Hechos atómicos y proposiciones atómicas

En el corazón del atomismo lógico yace la idea de que la realidad está compuesta por hechos atómicos, unidades irreducibles que consti-

tuyen los bloques fundamentales del mundo. Correspondiendo a estos hechos atómicos, Russell postula la existencia de proposiciones atómicas en el lenguaje. Como él mismo explica: “Los hechos atómicos son las unidades últimas del mundo externo y las proposiciones atómicas son las unidades últimas de nuestro conocimiento del mundo externo” (Russell, 1918, p. 199).

Esta correspondencia entre hechos atómicos y proposiciones atómicas no es meramente una conveniencia lingüística, sino que representa un intento de alinear nuestra comprensión conceptual con la estructura misma de la realidad. Sin embargo, esta propuesta plantea interrogantes fascinantes: ¿hasta qué punto nuestro lenguaje puede capturar la complejidad del mundo?, ¿existen hechos atómicos que eludan toda representación lingüística?

Descomposición y análisis

El atomismo lógico sostiene que cualquier proposición compleja puede ser descompuesta en proposiciones atómicas. Este proceso de análisis lógico promete una comprensión clara y precisa de las afirmaciones complejas. Wittgenstein (1921/1961, 3.25), en su *Tractatus Logico-Philosophicus*, afirma con característica concisión: “ Toda proposición tiene un análisis completo”.

Este principio de descomposición, aunque atractivo en su simplicidad, plantea desafíos significativos. ¿Es realmente posible descomponer todo enunciado en unidades atómicas sin perder matices o significados emergentes? La idea de que toda complejidad puede reducirse a componentes simples tiene un atractivo intuitivo, pero ¿no corremos el riesgo de caer en un reduccionismo excesivo?

Lenguaje y representación

Finalmente, el atomismo lógico propone que el lenguaje tiene una estructura lógica que refleja la estructura del mundo. Las proposiciones

atómicas son las piezas fundamentales de este lenguaje y su veracidad depende de su correspondencia con los hechos atómicos. Wittgenstein lo expresa de manera poética: “La proposición es una figura de la realidad” (1921/1961, 4.01). Esta metáfora pictórica es poderosa, pero también problemática. ¿Hasta qué punto puede el lenguaje ‘pintar’ una imagen fiel de la realidad?, ¿no está nuestro lenguaje inevitablemente teñido por nuestras percepciones, cultura y limitaciones cognitivas? Estas preguntas adquieren una urgencia renovada en la era de la IA.

La importancia del atomismo lógico para nuestro análisis va más allá de su valor histórico en la filosofía del lenguaje. Nos proporciona un marco para examinar críticamente cómo construimos representaciones de la realidad, ya sea a través del lenguaje natural o de los lenguajes formales de la computación. Al aplicar estos principios a los *datasets* utilizados en IA, podemos obtener perspectivas novedosas sobre cómo estos *datasets* representan (o fallan en representar) la realidad, y cómo estas representaciones pueden llevar a problemas éticos en los sistemas de IA.

Inteligencia artificial y conjuntos de datos

La inteligencia artificial (IA), en su manifestación contemporánea, representa un intento audaz de emular y potencialmente superar capacidades cognitivas humanas a través de sistemas computacionales. Sin embargo, a diferencia de la inteligencia humana, que emerge de procesos biológicos y experiencias acumuladas, la IA se construye sobre una base de datos estructurados y algoritmos de aprendizaje.

Como señala acertadamente Andrew Ng (Marr, 2018), uno de los pioneros en el campo del *Deep Learning*, “[l]os datos son el nuevo petróleo para la IA. Cuantos más datos de alta calidad tengamos, mejores modelos de IA podemos construir”. Esta metáfora del “petróleo de datos” es reveladora en múltiples niveles. Por un lado, enfatiza el valor y la centralidad de los datos en el desarrollo de la IA. Por otro, nos invita a reflexionar sobre las implicaciones éticas y ambientales de esta nueva ‘economía de datos’.

La creación de datasets para IA no es un proceso neutral o meramente técnico. Implica decisiones cruciales sobre qué aspectos de la realidad se capturan, cómo se categorizan, etiquetan y qué se deja fuera. Estas decisiones tienen profundas implicaciones filosóficas y éticas que a menudo se pasan por alto en las discusiones técnicas sobre IA. Esta generación de datasets de alta calidad para IA se rige por varios principios que, aunque presentados como directrices técnicas, tienen profundas raíces filosóficas:

- Relevancia y representatividad

¿Qué constituye una muestra ‘representativa’ de la realidad? Esta pregunta nos lleva al corazón de debates epistemológicos sobre la naturaleza del conocimiento y la posibilidad de una representación objetiva.

- Calidad y precisión

La búsqueda de datos ‘precisos’ presupone una teoría de la verdad y la correspondencia entre nuestras representaciones y la realidad.

- Diversidad y balance

El imperativo de la diversidad en los datasets refleja preocupaciones éticas sobre la equidad y la inclusión, pero también plantea preguntas ontológicas sobre la naturaleza de las categorías y las diferencias.

- Suficiencia de datos

¿Cuántos datos son ‘suficientes’? Esta pregunta tiene resonancias con debates filosóficos sobre la inducción y la justificación del conocimiento.

- Variedad y escalabilidad

La búsqueda de variedad en los datos refleja una comprensión de la complejidad y multidimensionalidad de la realidad que desafía las simplificaciones excesivas.

- Anonimización y privacidad

Las preocupaciones sobre la privacidad de los datos nos llevan a cuestiones éticas fundamentales sobre la autonomía individual y los límites del conocimiento público.

- Consistencia y estabilidad

La búsqueda de consistencia en los datos refleja un compromiso con principios lógicos fundamentales, como la no contradicción.

- Evaluación y validación continua

Este principio reconoce la naturaleza provisional y falible de los datos utilizados.

Al examinar estos principios, se hace evidente que los *datasets* no son meras colecciones de información, sino intentos estructurados de capturar y representar aspectos de la realidad. Cada bit de información en un dataset se convierte, en efecto, en un ‘hecho atómico’ (en el sentido del atomismo lógico) y las etiquetas asociadas funcionan como ‘proposiciones atómicas.’ Gitelman (2013, p. 2) nos recuerda que “los datos brutos son un oxímoron”, enfatizando que incluso los datos aparentemente más básicos ya están imbuidos de decisiones interpretativas y contextuales. Esta observación tiene profundas implicaciones para nuestra comprensión de los *datasets* en IA ya que las etiquetas en los conjuntos de datos de IA, al igual que las proposiciones atómicas en el atomismo lógico, no son meras descripciones neutrales, sino que tienen el poder de dar forma a nuestra comprensión de la realidad.

Kitchin (2014, p. 2) señala que “los datos no existen independientemente de las ideas, instrumentos, prácticas, contextos y conocimientos utilizados para generarlos, procesarlos y analizarlos”. Esta advertencia nos recuerda los peligros de un reduccionismo excesivo, tanto en filosofía como en IA. En última instancia, la creación de datasets para IA nos obliga a confrontar preguntas filosóficas fundamentales sobre la naturaleza de la realidad, el conocimiento y la representación. Preguntas que no son me-

ras abstracciones académicas, sino que tienen consecuencias prácticas y éticas profundas en un mundo cada vez más mediado por sistemas de IA.

Conexión entre atomismo lógico y conjuntos de datos

La aparente disparidad entre el atomismo lógico, una teoría filosófica del siglo XX y los conjuntos de datos modernos utilizados en inteligencia artificial puede, a primera vista, parecer insalvable. Sin embargo, un análisis más profundo revela paralelismos sorprendentes y fructíferos entre ambos conceptos. Esta sección explora estas conexiones, argumentando que existe una analogía significativa entre los principios del atomismo lógico y la estructura de los conjuntos de datos utilizados en IA.

Como ya se señaló previamente, el atomismo lógico postula la existencia de hechos atómicos como las unidades básicas e indivisibles de la realidad. De manera análoga, los datos individuales en un conjunto de datos de IA pueden considerarse como ‘hechos atómicos’ en el sentido de Russell. Cada punto de datos, ya sea un píxel en una imagen, una palabra en un texto, o un valor numérico en una tabla, representa una unidad básica de información. Al igual que los hechos atómicos de Russell, estos datos individuales son las unidades fundamentales sobre las cuales se construye todo el edificio del aprendizaje automático.

Esta analogía no es meramente superficial. Tanto los hechos atómicos del atomismo lógico como los datos individuales en IA comparten una característica crucial: se presentan como elementos objetivos y no interpretados de la realidad. En el atomismo lógico, las proposiciones atómicas son expresiones lingüísticas que corresponden directamente a los hechos atómicos. De manera similar, las etiquetas (*labels*) en los conjuntos de datos de IA funcionan como ‘proposiciones atómicas’, proporcionando una interpretación lingüística o categórica de los datos brutos.

Por ejemplo, en un conjunto de datos de imágenes, la etiqueta ‘gato’ asociada a una imagen particular funciona de manera análoga a una proposición atómica que afirma ‘esto es un gato’. Esta etiqueta, al igual

que una proposición atómica, hace una afirmación sobre la realidad (en este caso, el contenido de la imagen) que puede ser verdadera o falsa.

Sin embargo, esta analogía también revela las limitaciones y peligros potenciales de nuestros sistemas de etiquetado. Como advierte Bowker (2014, p. 65), “las formas en que categorizamos y clasificamos dan forma al mundo en que vivimos”. Las etiquetas en los conjuntos de datos de IA, al igual que las proposiciones atómicas en el atomismo lógico, no son meras descripciones neutrales, sino que tienen el poder de dar forma a nuestra comprensión de la realidad.

El atomismo lógico sostiene que cualquier proposición compleja puede ser descompuesta en proposiciones atómicas. De manera similar, el proceso de creación de conjuntos de datos para IA implica la descomposición de fenómenos complejos en unidades discretas y etiquetadas de información. Este proceso de descomposición y análisis es fundamental tanto para el atomismo lógico como para la creación de conjuntos de datos de IA. Sin embargo, como advierte Kitchin (2014, p. 2), “los datos no existen independientemente de las ideas, instrumentos, prácticas, contextos y conocimientos utilizados para generarlos, procesarlos y analizarlos”. Esta advertencia nos recuerda los peligros de un reduccionismo excesivo, tanto en filosofía como en IA. La analogía entre el atomismo lógico y los conjuntos de datos de IA tiene implicaciones prácticas y éticas significativas:

- Limitaciones de la representación

Al igual que el atomismo lógico ha sido criticado por su incapacidad para capturar ciertos aspectos de la realidad (Quine, 1951), los conjuntos de datos de IA pueden fallar en representar adecuadamente la complejidad del mundo real.

- Sesgos implícitos

Así como las proposiciones atómicas del atomismo lógico pueden contener sesgos lingüísticos y culturales, las etiquetas en los conjuntos de datos de IA pueden perpetuar y amplificar sesgos sociales existentes.

- Poder ontológico

Tanto el atomismo lógico como los conjuntos de datos de IA tienen el poder de dar forma a nuestra comprensión de la realidad, influyendo qué consideramos como 'hechos' y cómo los categorizamos.

- Necesidad de contextualización

Al igual que las proposiciones atómicas necesitan ser entendidas en un contexto más amplio, los datos individuales en IA requieren una contextualización cuidadosa para ser interpretados correctamente.

En conclusión, la analogía entre el atomismo lógico y los conjuntos de datos de IA nos proporciona una nueva lente a través de la cual podemos examinar críticamente nuestras prácticas de creación y uso de datos en IA. Nos recuerda que, al igual que el lenguaje en la filosofía, los datos en IA no son un reflejo neutral de la realidad, sino una construcción que incorpora decisiones filosóficas profundas sobre la naturaleza de la realidad y nuestro conocimiento de ella.

Implicaciones éticas

Al considerar los conjuntos de datos como representaciones de la realidad, similares a las proposiciones atómicas en el atomismo lógico, nos enfrentamos a cuestiones fundamentales sobre la objetividad, la representación y la justicia en los sistemas de IA. Un problema ético central que emerge de esta analogía es el del sesgo de representación. Al igual que el lenguaje en el atomismo lógico puede fallar en capturar toda la complejidad de la realidad, los conjuntos de datos en IA a menudo presentan una visión parcial o sesgada del mundo. Como señala O'Neil (2016), estos sesgos pueden tener consecuencias devastadoras cuando se incorporan en sistemas de IA que toman decisiones que afectan la vida de las personas. Por ejemplo, un conjunto de datos de imágenes médicas que no representa adecuadamente la diversidad de tonos de piel podría

llevar a diagnósticos erróneos para ciertos grupos étnicos, perpetuando y exacerbando las desigualdades existentes en la atención médica.

La falta de diversidad cultural en los conjuntos de datos presenta otro desafío ético significativo. Al igual que el atomismo lógico de Russell y Wittgenstein surgió de un contexto filosófico occidental específico, los conjuntos de datos de IA a menudo reflejan los sesgos culturales de sus creadores. Esta falta de diversidad puede llevar a sistemas de IA que malinterpreten o subestimen las expresiones lingüísticas y las necesidades de diferentes grupos culturales. Esto no solo plantea problemas de equidad, sino que también limita la eficacia y aplicabilidad global de los sistemas de IA.

El problema del sesgo de selección en la creación de conjuntos de datos refleja las limitaciones del atomismo lógico en su intento de representar la realidad a través de proposiciones atómicas. Así como la elección de qué proposiciones considerar como ‘atómicas’ puede ser arbitraria o culturalmente sesgada, la selección de datos para incluir en un conjunto de datos de IA puede introducir sesgos significativos. Mittelstadt (*et al.*, 2016) señalan que estos sesgos pueden ser particularmente perniciosos cuando se trata de decisiones automatizadas en áreas sensibles como la justicia penal o la evaluación crediticia.

El impacto de los conjuntos de datos sesgados en las decisiones automatizadas es quizás la implicación ética más urgente de esta analogía. Al igual que las proposiciones atómicas en el atomismo lógico se consideraban los bloques de construcción de afirmaciones más complejas sobre la realidad, los datos individuales en los conjuntos de datos de IA se convierten en la base sobre la cual se toman decisiones automatizadas. Barocas y Selbst (2016, p. 671) argumentan que esto puede llevar a una “discriminación por diseño”, donde los sesgos históricos y sociales se codifican y amplifican a través de sistemas de IA aparentemente objetivos.

El reforzamiento de estereotipos a través de conjuntos de datos sesgados es otra consecuencia ética preocupante. De la misma manera que el lenguaje puede perpetuar estereotipos y prejuicios, los con-

juntos de datos de IA pueden reforzar y amplificar las desigualdades existentes. Noble (2018) demuestra cómo los algoritmos de búsqueda, entrenados en conjuntos de datos sesgados, pueden perpetuar estereotipos raciales y de género, influyendo en la percepción pública y las oportunidades individuales.

Estas implicaciones éticas nos obligan a reconsiderar fundamentalmente nuestro enfoque de la creación y uso de conjuntos de datos en IA. Al igual que el atomismo lógico nos invita a examinar críticamente la relación entre lenguaje y realidad, debemos escrutinar cuidadosamente cómo nuestros conjuntos de datos representan y dan forma al mundo.

Conclusiones

El análisis de la relación entre el atomismo lógico y los conjuntos de datos de IA nos permite comprender mejor los desafíos éticos que enfrentamos en el desarrollo y aplicación de la inteligencia artificial. Este enfoque filosófico-lingüístico nos ofrece una nueva perspectiva para abordar y mitigar los problemas éticos en IA, enfatizando la importancia de la diversidad, representatividad y equidad en la creación de conjuntos de datos.

En última instancia, la analogía entre el atomismo lógico y los conjuntos de datos de IA nos recuerda que la creación de conocimiento, ya sea a través del lenguaje filosófico o de los datos computacionales, nunca es un proceso neutral. Nos enfrentamos al desafío ético de crear sistemas de IA que no solo sean eficientes y precisos, sino también justos, inclusivos y respetuosos de la diversidad humana. Este desafío requiere una colaboración continua entre filósofos, éticos, científicos de datos y desarrolladores de IA para garantizar que nuestros sistemas tecnológicos reflejen nuestros valores éticos más elevados.

Propongo que futuros desarrollos en IA deberían considerar explícitamente esta conexión entre lenguaje, representación y ética, para crear sistemas más justos y equitativos. Esto podría implicar nuevas metodologías

para la creación de conjuntos de datos, así como un escrutinio más riguroso de los principios filosóficos que subyacen a nuestras prácticas de IA.

En conclusión, al reinterpretar los desafíos éticos de la IA a través de la lente del atomismo lógico y la filosofía del lenguaje, podemos obtener nuevas perspectivas sobre cómo crear sistemas de IA más éticos y equitativos. Este enfoque nos recuerda que la IA, al igual que el lenguaje, es una herramienta poderosa para dar forma a nuestra comprensión del mundo, y como tal, debe ser desarrollada y utilizada con un profundo sentido de responsabilidad ética.

Referencias bibliográficas

- Barocas, S. y Selbst, A. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671-732.
- Bowker, G. C. (2014). The Theory/Data Thing. *International Journal of Communication*, 8, 1795-1799.
- Gitelman, L. (Ed.). (2013). "Raw Data" Is an Oxymoron. MIT Press.
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. SAGE Publications.
- Marr, B. (2018, May 21). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. *Forbes*. <https://bit.ly/41EgGTA>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S. y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishers.
- Quine, W. V. O. (1951). Two Dogmas of Empiricism. *The Philosophical Review*, 60(1), 20-43.
- Russell, B. (1914). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. Open Court Publishing Company.
- Russell, B. (1918). The Philosophy of Logical Atomism. *The Monist*, 28(4), 495-527.
- Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). Routledge & Kegan Paul.