

JUAN RODRÍGUEZ MEDINA

La de la inteligencia artificial no es una violencia visible, explica la filósofa chilena Gabriela Arriagada, no es estridente o inmediata: “No hay gritos, no hay golpes. Es una violencia mucho más silenciosa, y precisamente por eso, más difícil de detectar. Es la que ocurre cuando los datos, que parecen objetivos e imparciales, comienzan a definir quién cuenta y quién no dentro del mundo digital”, contesta la autora al preguntarle por el título de su nuevo libro: “La violencia discursiva de los datos” (Ediciones UC).

Especializada en éticas aplicadas y profesora de Ética de IA y Datos en la PUC, Arriagada ya había reflexionado sobre esta violencia en su libro anterior, “Los sesgos del algoritmo” (La Pollera, 2024), en un sentido amplio, es decir, sesgos de género, raza y clase. Esta vez, como indica el subtítulo de su nuevo ensayo, el asunto es la “Discapacidad, IA y ética”.

Ya sabemos que en nuestro presente en todo lugar y todo el tiempo estamos nutriendo con datos a los levitanes digitales y sus algoritmos. “La violencia discursiva”, dice Arriagada, “aparece cuando esos datos traducen la diversidad humana a categorías estrechas, muchas veces diseñadas desde una noción dominante de ‘normalidad’. Por ejemplo, si un sistema de inteligencia artificial ha sido entrenado mayoritariamente con información de personas sin discapacidad, tenderá a interpretar otras formas de habitar el cuerpo o la mente como anomalías, errores o déficits. No porque ‘quiera’ discriminar, sino porque fue educado en un marco que no reconoce esa diversidad como parte legítima del mundo”.

“Lo preocupante”, agrega la filósofa, “es que, cuando esa lógica se automatiza, adquiere una apariencia de neutralidad técnica. La discriminación deja de parecer un acto humano y pasa a verse como el resultado ‘objetivo’ de un cálculo matemático. Y eso la hace más difícil de cuestionar”.

—Los datos no son la realidad.

“Los datos no solo describen la realidad: la construyen. Determinan qué problemas se consideran relevantes, qué cuerpos se diseñan como estándar y cuáles quedan fuera del diseño. Si no examinamos críticamente qué valores están inscritos en esos datos, la tecnología puede terminar perfeccionando sistemas de exclusión bajo el discurso de la eficiencia y la innovación. Hablar de violencia discursiva es, en el fondo, una invitación a mirar detrás del algoritmo. A preguntarnos: ¿quién fue considerado al construir este sistema?, ¿quién quedó fuera?, ¿qué idea de ser humano está siendo reforzada?”.

Justicia social

Arriagada aboga por que la IA sea una herramienta que nos ayude a construir una sociedad más justa, humana, inclusiva. El asunto, en términos históricos, es si alguna vez una tecnología disruptiva se ha ajustado a nuestros buenos deseos o si más bien somos como aprendices de brujos; y, por otro lado, al menos hasta ahora, los grandes desarrollos de la IA provienen de corporaciones privadas, cuyo norte es la rentabilidad.

“Hay una tensión estructural entre innovación guiada por el mercado y justicia social”, reconoce Arriagada. “Pero reducir el problema únicamente a la propiedad sería simplificarlo. La historia muestra que las tecnologías disruptivas nunca han sido moralmente neutras ni han seguido automáticamente ‘buenos deseos’. La imprenta, la electricidad, internet: todas ampliaron libertades y, al mismo



ENTREVISTA | Gabriela Arriagada, filósofa:

“La tecnología nunca ha sido NI SERÁ NEUTRA”

La especialista en ética aplicada e inteligencia artificial publica “La violencia discursiva de los datos”, un ensayo en que reflexiona sobre las parcialidades y sesgos que se ocultan bajo la aparente objetividad de los algoritmos. Y sobre la necesidad, entonces, de preguntarnos cómo desarrollamos las tecnologías, quiénes lo hacen y para qué.

La imprenta, la electricidad, internet: todas ampliaron libertades y, al mismo tiempo, generaron nuevas formas de concentración de poder”.

tiempo, generaron nuevas formas de concentración de poder. No somos aprendices de brujo por el mero hecho de innovar; lo somos cuando renunciamos a gobernar lo que creamos”.

En concreto, “más allá de las declaraciones de buena voluntad”, Arriagada cree, primero, que hay que fortalecer la gobernanza pública: “Si la infraestructura digital se ha transformado en un componente estructural de nuestras sociedades, comparable a la energía o al transporte, entonces requiere marcos regulatorios acordes con su impacto”.

“En segundo término, es indispensable redistribuir poder en el diseño tecnológico. Mientras la arquitectura de la inteligencia artificial permanezca concentrada en un número reducido de actores privados, la definición de sus prioridades éticas será ne-

cesariamente parcial. Incorporar a las comunidades directamente afectadas, incluidas las personas con discapacidad, en las etapas de diseño, prueba y supervisión no constituye un gesto simbólico, sino una estrategia concreta para corregir sesgos estructurales. La inclusión no puede limitarse al acceso o al uso, debe abarcar también la producción y la toma de decisiones”.

Un tercer elemento, propone Arriagada, es potenciar la infraestructura pública y los modelos alternativos de desarrollo de estas tecnologías, de hecho, ejemplifica, existen consorcios universitarios y proyectos financiados por el Estado “que demuestran que el progreso tecnológico no depende exclusivamente del capital privado”.

Ahora, como el asunto es ético, como tiene que ver con preguntarnos qué es ser humano y cómo vivir bien, también se requiere un cambio cultural en la forma en que evaluamos el éxito tecnológico: “Hoy predominan indicadores como escalabilidad, rapidez de adopción y retorno financiero. Si no incorporamos como métricas centrales el impacto so-

“No somos aprendices de brujo por el mero hecho de innovar; lo somos cuando renunciamos a gobernar lo que creamos”.

cial, la equidad y la accesibilidad, la dimensión ética continuará siendo marginal”, cree Arriagada. “La experiencia histórica es elocuente. Las tecnologías disruptivas no se han alineado espontáneamente con ideales de justicia. Han sido moldeadas mediante regulación, liberación pública y conflicto político. La inteligencia artificial no será justa por simple inercia. Lo será únicamente si existe voluntad colectiva para someterla a principios democráticos”.

—El asunto no es solo técnico.

“El desafío es político y axiológico. Se trata de decidir si la inteligencia artificial será únicamente una herramienta de optimización económica o si la concebiremos como infraestructura social orientada por valores compartidos. La tecnología no determina por sí sola nuestro destino, pero sí amplifica las decisiones que adoptamos como comunidad”.

—Hace algunos días se presentó en Chile Latam-GPT, “un modelo de lenguaje abierto diseñado con identidad propia, a diferencia de otros modelos globales”, según destacó el Gobierno. ¿Es una buena noticia?

“Lo que destaco en particular, más que simplemente la existencia del modelo, es lo que significa regionalmente y lo que proyecta: un llamado a la pluralidad epistemológica y cultural. Cuando se anuncia que Latam-GPT es un modelo abierto y con identidad propia, se está señalando algo relevante. La mayoría de los grandes modelos de lenguaje han sido entrenados desde contextos anglosajones, con corpus, prioridades y marcos culturales específicos. Eso influye en cómo entienden el mundo, qué referencias privilegian y qué experiencias quedan en segundo plano. Ahora bien, lo más interesante no es solo el origen geográfico, sino el modelo de desarrollo. Cuando hablamos de proyectos abiertos, colaborativos y con vocación comunitaria, estamos frente a un intento de modificar la arquitectura de poder en la producción tecnológica. Eso es coherente con la crítica a la violencia discursiva de los datos: no basta con cambiar el contenido, es necesario cambiar quién define qué datos se usan, qué problemas se priorizan y qué comunidades participan”.

—¿La tecnología no es neutra, entonces?

“La afirmación de que ‘la tecnología es neutra’ parte de una concepción instrumental: la tecnología sería un medio vacío y el juicio moral recaería exclusivamente en quien la emplea. Sin embargo, mi trabajo en ética de la inteligencia artificial se sustenta precisamente en lo contrario: la tecnología nunca ha sido ni será neutra. La IA no es intrínsecamente buena o mala; carece de agencia moral propia. Pero tampoco es neutra. Está configurada desde posiciones de conocimiento situadas, atravesada por relaciones de poder y por supuestos epistemológicos que deben hacerse explícitos. La tarea de la ética de la IA, tal como la concibo, no consiste solo en regular usos, sino en desnaturalizar la pretendida neutralidad técnica y examinar críticamente las condiciones de posibilidad de la propia tecnología”.



Gabriela Arriagada Bruneau es académica de la PUC.

THE NEW YORK TIMES