


LA GOBERNANZA DE LA  
INTELIGENCIA ARTIFICIAL:  
DE SOLUCIONAR LOS PROBLEMAS  
A DIAGNOSTICARLOS

**URVASHI ANEJA**

Directora fundadora de Tandem Research  
([www.tandemresearch.org](http://www.tandemresearch.org)) e investigadora  
asociada de Catham House



La evidencia creciente de los múltiples riesgos y perjuicios que comportan los sistemas de toma de decisiones mediante algoritmos han urgido la cuestión de la gobernanza de la Inteligencia Artificial (IA). Los gobiernos, la industria y los organismos de la sociedad civil del planeta están actualmente desarrollando marcos de referencia para lograr un uso de la IA que sea más seguro, justo y responsable. El presente artículo ofrece un repaso de las visiones dominantes respecto a la gobernanza de la IA y mantiene que, si bien dichos enfoques son relevantes y necesarios, aún no abordan adecuadamente los retos sobre la ética, el poder y la comunidad. Como complemento a los enfoques centrados en encontrar “soluciones a los problemas” –basados en minimizar los daños causados por la IA– también necesitamos visiones centradas en el “diagnóstico de problemas”, que examinen el origen y las causas de los actuales imaginarios y de la economía basada en la IA. Pivotando entre estos dos enfoques, será más fácil identificar y llevar a cabo intervenciones sistémicas y emancipatorias, con vistas a fusionar los avances de la innovación en IA y el bienestar de la sociedad.

### **Solución de problemas mediante la autorregulación, los estándares técnicos y la defensa de los derechos humanos**

En respuesta a la desconfianza social que genera el impacto de la IA, y al aún incipiente marco regulatorio en este campo, un número creciente de compañías tecnológicas han optado por autorregularse, dotándose de un conjunto de principios éticos que guíen el desarrollo y el despliegue de la IA. En ese sentido, tan solo en 2018 y 2019, la organización *Algorithm Watch* ha contado más de cien declaraciones éticas empresariales sobre IA. Si bien difieren entre ellas en cuanto al articulado, casi todas convergen en los principios de transparencia, imparcialidad, equidad, responsabilidad y seguridad. Naturalmente, el problema principal de estos marcos éticos es que no son vinculantes ni obligatorios; resulta tan sencillo comprometerse a cumplir estos principios como violarlos. La mayoría de declaraciones se formulan generalmente en forma de promesas ambiguas, sin una concreción explícita de cómo deberán cumplirse o implementarse. Podría afirmarse que, en realidad, la mayoría de los marcos éticos parecen más un parapeto de la industria para sugerir que con la autorregulación es suficiente para minimizar los riesgos de la IA, y evitar así otras formas de regulación que les resten margen de autonomía y de control.

Otra pregunta que debemos plantearnos es qué tipo de ética está siendo privilegiada en dichos marcos; por qué, por ejemplo, determinadas interpretaciones de la equidad y la transparencia han sido incorporadas a las discusiones sobre la IA ética; y, del mismo modo, cuáles son los marcos éticos alternativos que, a consecuencia de ello, han quedado excluidos. Como señalan Noopur Raval y Amba Kak<sup>1</sup> con respecto a las diferencias entre los contextos geográficos “no podemos asumir que términos como imparcialidad, transparencia y responsabilidad tengan el mismo significado, o siquiera la misma centralidad, en las discusiones sobre IA, ética y gobernanza que tienen lugar [en el Norte] y en el Sur Global”.

1. Véase Raval, Kak y Calcaño, 2021.

Otro enfoque, surgido en el seno de la comunidad de la ciencia de la computación, defiende el diseño y la programación de sistemas algorítmicos más justos, responsables y transparentes. Sin embargo, este objetivo requiere de debates previos imprescindibles y no siempre sencillos. Por ejemplo, una investigación reciente acerca del algoritmo Compass, que es el que emplean los tribunales estadounidenses, sugiere que se alcanzarían resultados similares empleando menos de los 20 parámetros de información que se emplean actualmente. Al simplificar el algoritmo, éste sería más explicable y debido a ello, más fácil de calibrar y supervisar.

Sin embargo, debemos tener presente que la equidad es un atributo de los sistemas sociales, no de los sistemas técnicos; solo se puede entender dentro de un contexto social. Incluso en el seno de la comunidad de científicos de la computación coexisten diversas definiciones de justicia/equidad. En determinados casos puede incluso ser necesaria la discriminación positiva para ajustar o corregir desigualdades sociales ya existentes. Así pues, el objetivo de construir mejores algoritmos debe estar precedido por un debate sobre si es realmente necesario o no

desplegar sistemas algorítmicos para desarrollar dichas tareas. Consideremos, por ejemplo, el caso de las aplicaciones para el reconocimiento facial; incluso si los algoritmos subyacentes son equitativos, transparentes y responsables, esto no desvanece las preocupaciones sobre el impacto que tenga esta tecnología sobre los derechos y libertades civiles.

Un tercer y más prometedor enfoque es la implementación de marcos regulatorios internacionales basados en el respeto a los derechos humanos. En este caso, se trata de marcos universales y vinculantes, que pasan a formar parte del corpus del derecho internacional; en este caso, las responsabilidades de gobiernos y empresas sí quedan bien delimitadas, y existe ya un conjunto de instancias jurídicas domésticas, regionales e internacionales pensadas para subsanar los incumplimientos. Esta área atrae cada vez más el interés de los gobiernos y de académicos, sumando cada vez más voces en pos de la adopción de este tipo de marco regulador de la inteligencia artificial basado en los derechos humanos. A pesar de ello, persisten aún tres escollos importantes por sortear. El primero se refiere a que dicho marco regulador sitúa al

individuo en el centro de su acción y prioriza los perjuicios individuales. No obstante, en el mundo de la IA los perjuicios son más a menudo sociales y estructurales que individuales. Tomemos, por ejemplo, los modelos de inteligencia artificial para la calificación del crédito que se nutren de datos de los usuarios obtenidos de las redes sociales y de sus hábitos de consumo; el problema no es solo de privacidad, o que determinados individuos sean discriminados por los amigos que tienen, o por cómo gastan su dinero, sino que con el tiempo y a partir de múltiples interacciones, están generando nuevas categorías del tipo de individuo que es confiable. Los algoritmos, según afirma Tarleton Gillespie<sup>2</sup>, “no son solo códigos con consecuencias” sino que están “íntimamente ligados a la producción de significado”. ¿Es suficiente un enfoque meramente regulatorio, cuando lo que está en juego es cómo conocemos y entendemos el mundo? La evidencia nos muestra que este tipo de cuestiones va más allá del citado marco regulador basado en los derechos humanos.

2. Véase Gillespie, 2013.



Personas “generadas” por un algoritmo.  
<https://thispersondoesnotexist.com>

Igualmente, una sola acción perjudicial puede pasar inadvertida, pero muchas instancias perjudiciales combinadas o persistentes en el tiempo pueden tener consecuencias sistémicas, que lleguen incluso a alterar las estructuras sociales. La reestructuración de industrias enteras mediante la automatización inteligente, por ejemplo, tendrá efectos adversos en el empleo. Las nuevas cotas de eficiencia y de automatización que conllevará la IA reducirán forzosamente la necesidad de mano de obra, y es muy probable que esto exacerbe la desigualdad global. Esta es en particular una preocupación para los países en vías de desarrollo, que ya tienen unos índices muy elevados de desempleo y subempleo, a los que deberán sumarse la relocalización de las industrias manufactureras. Bien es cierto que en el Sur Global están surgiendo nuevos empleos en el mercado laboral telemático, dedicados a proporcionar a las industrias globales de IA servicios de etiquetaje y categorización de datos necesarios para el buen funcionamiento de las aplicaciones, pero son empleos con retribuciones bajas y volátiles y con unas condiciones de trabajo precarias. Este tipo de perjuicios estructurales tienden, por lo general, a quedar al margen de los enfoques centrados en los derechos.

El segundo escollo para desarrollar un marco regulador centrado en que los derechos humanos se basa en que la existencia de unos derechos implica también la de unos deberes de garantía que, en este caso, recaen sobre todo en el Estado y otras organizaciones vicarias. Y los acontecimientos recientes evidencian que los estados, incluyendo las democracias liberales, están dotándose de tecnologías de vigilancia basadas en la IA a un ritmo más rápido que nunca antes, con el propósito de supervisar y vigilar a los ciudadanos. ¿Hasta qué punto, pues, pueden estos mismos estados proporcionar una protección significativa contra las vulneraciones de los derechos humanos? Uno de los puntos débiles del enfoque legalista o regulador a escala global es asumir que en caso de infracción existen también las instancias de reparación. Lo cierto es que en muchos lugares, particularmente en los países en vías de desarrollo, las instituciones de defensa de los derechos humanos

tienden a ser más débiles y mucho menos capaces de proteger a los más vulnerables. Tal y como afirma Chinmayi Arun<sup>3</sup>, “cuando las empresas despliegan estas tecnologías en los países del Sur, existen muy pocos recursos e instituciones que puedan proteger los derechos de las personas marginalizadas. Las democracias jóvenes carecen aún de la estabilidad institucional necesaria; (...) y del tiempo que requiere institucionalizar las prácticas democráticas”.

## **En lugar de preocuparse solo por ganar tanto dinero como sea posible, las empresas de tecnología pueden y deben preocuparse por diseñar el mundo en línea de manera que contribuya al bienestar de las personas**

El tercer y último escollo se basa en que, como dice Anna Laureen Hoffman<sup>4</sup>, los marcos basados en los derechos se concentran en un número determinado de actores *perverseos*, desatendiendo a las estructuras más amplias que les dan cabida. Un ejemplo de ello es el documental de Netflix “El gran *hackeo*”, producido en el 2019 y focalizado en la empresa Cambridge Analytica, que vilipendia a Facebook y a su fundador, Mark Zuckerberg, sin recalcar que la lógica subyacente a sus actos —*clusterizar* a los usuarios para influir sobre preferencias con objetivos tanto comerciales como políticos— es una práctica que se ha extendido de manera generalizada en todo el sector y que está siendo aceptada y aplaudida por muchos. Además, como argumenta Hoffman, los debates sobre los derechos humanos

no contemplan cómo se crean las estructuras de privilegios y cómo esto puede producir resultados injustos. Tampoco contribuyen a facilitar la relación entre las estructuras cambiantes de poder y de agencia. Considérese la cuestión de la supervisión en el lugar de trabajo: aunque se solicite el consentimiento del trabajador para la recopilación de datos para evaluar su rendimiento, e incluso respetando escrupulosamente los estándares de protección de datos, tiende a prestarse poca atención al cambio fundamental que esto produce en la dinámica de poder entre patrones y empleados, como consecuencia de los sistemas de monitorización, incluso cuando estos cuentan con el consentimiento expreso de los afectados.

### **Desentrañando el imaginario y la economía de la IA**

Resulta indudable que las diversas aproximaciones citadas a los retos de la IA (la autorregulación, los estándares técnicos y muy especialmente la defensa de los derechos humanos) jugarán un papel crucial a la hora de fijar estándares, establecer normas y responsabilidades, y encauzar la investigación sobre la IA hacia el beneficio social. Pero, de manera complementaria, es necesario adoptar también un enfoque más estructural, que atienda a los relatos, los valores, los intereses y las historias subyacentes a las actuales tendencias de la IA y que ponga de nuevo en el centro los debates sobre el poder y la comunidad.

Hasta el momento, las estrategias nacionales de IA de los gobiernos de todo el mundo se caracterizan por un relato en el que resuena el principio de que “el ganador se lo lleva todo”, convencidos como están de que es necesario asegurarse un lugar preeminente en la carrera de la IA, con vistas a asegurar el crecimiento económico y la prosperidad futuros. Son muchos los documentos políticos que exudan la ansiedad de “ser dejado atrás”.

Por el contrario, en los países en vías de desarrollo, la IA se asocia mayormente a la “modernidad” y al “progreso”, y se contempla como una tecnología que puede ayudar a estos países a “ponerse al día” y a superar los desafíos al desarrollo de largo recorrido. Si atendemos pues a este futuro imaginario, podemos extrapolar que la IA podría supuestamente catapultar a los países en vías de desarrollo hacia nuevas cotas de crecimiento,

3. Véase Arun, 2020.

4. Véase Hoffman, 2019.

soltando el lastre de los problemas políticos, económicos y sociales que actualmente les atañen. Esta particular visión de la IA ha sido alimentada por las empresas tecnológicas; como escribe Yarden Katz<sup>5</sup>, la IA es una herramienta de *marketing* que encarna el mismo ímpetu que hace una década se depositaba en el análisis de datos a gran escala y los centros de datos empresariales. También ha generado un fértil espacio para que empresas del sector privado comercialicen y vendan productos basados en la IA con el pretexto de mejorar la eficiencia o resolver un problema crítico de desarrollo. El grueso del desarrollo global de la IA lo dirige un reducido grupo de “macroempresas tecnológicas” que tienen acceso a la mayoría de datos digitales globales; el crecimiento de la actual industria de la IA se basa de hecho en una enorme concentración de poder que se retroalimenta. Los argumentos acerca de una “IA socialmente beneficiosa” son a menudo promovidos por las mismas empresas tecnológicas, en un intento de desviar la atención de los intereses comerciales que impulsan la innovación en IA, o para ocultar sus usos y efectos más perniciosos.

El aprendizaje automático o *machine learning*—la técnica dominante en la IA actual— es esencialmente una forma de estadística computacional que implica dos procesos fundamentales: la clasificación o reconocimiento de patrones, y la predicción para la generación de patrones. Estos patrones los genera una correlación estadística y no indican la existencia de una relación de causalidad. Es por ello que Matteo Pasquinelli<sup>6</sup> sostiene que el aprendizaje automático, como una modalidad de estadística, representa una forma de “compresión de la información” que al mismo tiempo genera “ocultaciones de información”, al desestimar datos que son relevantes. La primera es la que hace posible que las empresas obtengan beneficios, mientras que la segunda comporta a menudo “una pérdida de la diversidad cultural del mundo”.

Los datos nunca son neutrales y siempre son parcialmente representativos. Dado que los sistemas de aprendizaje automático funcionan a partir de la inferencia de datos *históricos* (de sucesos que ya han tenido lugar), los patrones o predicciones producidos por dichos sistemas tenderán a reproducir

y a ampliar desigualdades socioeconómicas ya existentes. El requerimiento de datos masivos para nutrir al aprendizaje automático está propiciando las dinámicas extractivas de información de la economía de datos, a expensas de la privacidad y los derechos individuales.

Si percibimos la IA en estos términos —como una forma de estadística computacional que, como sostiene Cathy O’Neil, “reproduce el futuro basándose en el pasado”— nos veremos forzados a moderar nuestras expectativas de “una IA para el Desarrollo” o “una IA para el Bien Social”.

Los modelos estadísticos son ciertamente útiles, ya que pueden ayudarnos a identificar tendencias, nos proporcionan apoyo informativo, y son una herramienta útil en la toma de decisiones, en la planificación de los recursos y en la automatización de determinados procesos. Sin embargo, debemos ser conscientes de sus límites y sus ocultaciones, así como de los intereses comerciales que empujan la innovación actual y de sus mecanismos de actuación. Esto es particularmente importante en el caso de los países en vías de desarrollo, en los que la IA aún se percibe como la panacea a los problemas del desarrollo más complejos y donde, en paralelo, el Estado está perdiendo capacidad de maniobra frente al entramado cada vez más complejo de nuevas tecnologías y en el espacio que ello genera para que tecnócratas y grandes empresas tecnológicas influyan sobre las políticas tecnológicas estatales.

El célebre investigador en IA, De Kai, ha definido el momento actual como “una siesta de la IA”, ya que los modelos estadísticos de datos masivos como el aprendizaje automático y el aprendizaje profundo (*deep learning*) aún no representan bien la noción de “inteligencia”. La referencia habitual a que el aprendizaje automático se asemeja a enseñarle a un niño mediante la repetición de un mismo conjunto de *inputs* —por ejemplo, mostrándole una y otra vez un gato para que lo identifique— es inadecuado, ya que la máquina carece de la capacidad de abstracción que le permite al niño identificar un gato tras un breve número de ejemplos. Es posible que aquí resida una oportunidad para adentrarnos en un futuro diferente de la IA, en base a desarrollar otras técnicas computacionales capaces de crear máquinas verdaderamente inteligentes y menos reduccionistas, que

no dependan de un volumen masivo de datos y de recursos tan solo al alcance de un reducido grupo de empresas tecnológicas.

De ese modo, también hemos de considerar el tipo de “inteligencia” que prioriza la innovación actual en IA. Los paradigmas actuales promueven una interpretación particular de la persona individual, como un agente libre y racional, que busca maximizar sus deseos. Diana Myers señala que estas concepciones atomizadas del individuo están dissociadas de las fuerzas sociales y políticas y que “se identifican más bien con la racionalidad instrumental del mercado”. Los valores que generalmente se asocian a la inteligencia son la capacidad de razonar y la de hacer inferencias. Pero el supuesto de la racionalidad queda desmentido casi diariamente por la experiencia; la forma en que tanto gobiernos como la opinión pública han respondido ante la pandemia es un buen ejemplo de ello. La maximización individual del deseo también es potencialmente perjudicial. La actual política de vacunación o la resistencia a ponerse la mascarilla son ejemplos en este sentido. En esta definición dominante de la inteligencia artificial también se mezclan diferentes tipos de inteligencia, particularmente los relacionados con la inteligencia emocional o con cualidades como la empatía o el cuidado de los demás. Una vez más, la pandemia ha demostrado lo importantes que son estos otros tipos de inteligencia; actualmente los grupos de ayuda mutua son los que mantienen viva a la gente en India, y estos grupos se basan en el deber percibido de cuidar, no en un cálculo racional de la maximización del interés propio.

Junto con el desarrollo de nuevas técnicas computacionales para la inteligencia mecánica hemos de considerar otros tipos de persona, de inteligencia y de ética para trazar futuros alternativos a la IA. Los sistemas de conocimiento autóctonos, por ejemplo, no sitúan a los humanos al margen o por encima de los sistemas naturales. Los paradigmas relacionales basados en la idea de la sostenibilidad social y ambiental están detrás del desarrollo de muchas tecnologías surgidas de las comunidades indígenas. Esta ética relacional es un principio básico en muchas tradiciones culturales de todo el mundo, un principio que enfatiza el deber de cuidar de los demás, el deber de cuidar de las personas y del medio ambiente. En contraposición al modelo occidental de racionalidad implícito en la concepción

5. Véase Katz, 2017.

6. Véase Pasquinelli, 2019.



re]cognition

GALLERY AT TATE BRITAIN INFO SUBSCRIBE

← BACK TO GALLERY [SHARE](#)

**AI STATEMENT**  
 A MAN WEARING A HAT.  
 HADS HOURS IN THE YOUNGER BEARING A HAT

**1/1/2018**  
 A SUPPORTER OF LEBANON'S RESISTANCE LEADER  
 SAAYED HASSAN NABHALLAH HAS HIS PICTURE  
 ON HIS HEAD DURING A PUBLIC APPEARANCE BY  
 NABHALLAH AT A RELIGIOUS PROCESSION  
 AUTHOR: KYLE TAYLOR  
 SOURCE: REUTERS

**CONTEXT 18.6%**

TAGS  
 MAN, PERSON, COOKING, INDOOR, WEARING

**1946**  
 A MAN IN A BLACK CAP  
 AUTHOR: JOHN BURT FOSTER  
 SOURCE: TATE

**CONTEXT 16.6%**

TAGS  
 PERSON, HAT, MAN, WEARING, LOOKING

**CONTEXT 16.6%**

TAGS  
 MAN WITH GLASSES ON FACE  
 MAN WITH A HAT ON  
 MAN HAS A BEARD  
 THE MAN HAS GLASSES  
 THE EYES OF A MAN  
 BLACK HAT ON THE MAN  
 BLACK HAT ON THE MAN

**FACES 9.09%**

**SUBJECT 1**  
 AGE: 39  
 GENDER: MALE

**SUBJECT 2**  
 AGE: 70  
 GENDER: MALE

**COMPOSITION 86.6%**

**FACES 9.09%**

**SUBJECT 1**  
 AGE: 48  
 GENDER: MALE

**COMPOSITION 86.6%**

**FACES 9.09%**

**SUBJECT 1**  
 AGE: 48  
 GENDER: MALE

predominante y la ética de la IA, Sabelo Mhlambi<sup>7</sup> aporta la noción de *ubuntu*, propia del África Subsahariana, que se focaliza en el carácter relacional de las personas, como una marco referencial desde el que encarar la lógica estructural dañina que generan los sistemas de IA. Al tiempo, otros investigadores están explorando cómo la ética más asociada a los países asiáticos, centrada en el comunitarismo y en valores como la empatía y la compasión, puede moldear la gobernanza de la IA.

Estas exploraciones son cruciales. Sin embargo, el problema obvio aquí es la debilidad de las instituciones de conocimiento del Sur Global, que ha quedado una vez más en evidencia por la pandemia. Los análisis anticipatorios en los países en vías de desarrollo –incluso de países de renta media como India– han sido muy pobres, cuando no inexistentes; la cobertura en los medios de comunicación ha sido bastante superficial y las comunidades académica y política no han sabido interpretar adecuadamente la situación. Una vez más, la información más fiable y oportuna ha sido la que han aportado expertos foráneos, de los países industrializados, estudiando el mundo en vías de desarrollo. Y de nuevo, tenemos ejemplos de la asimetría del conocimiento en los debates actuales acerca de la ética de la IA; resulta llamativo, por ejemplo, que los debates sobre la nueva *colonización* a caballo de la IA tengan lugar mayormente en universidades y centros de investigación de las economías industrializadas. Resulta imperativo, pues, desarrollar una agenda emancipadora de la IA, que contemple mayores inversiones en la generación de conocimiento propio en los países en vías de desarrollo.

Invertir en dicha capacidad puede también proporcionar un equilibrio crítico entre los impulsos del sector privado y la atracción que sienten los gobiernos hacia las tentaciones de la IA. El rechazo que se está produciendo en contra de los sistemas de IA en algunos sectores de la sociedad civil de EEUU y Europa es un buen ejemplo; académicos, investigadores y organizaciones de juristas han desempeñado un papel fundamental identificando, documentando y estudiando los aspectos nocivos de la IA, y sus estudios son ahora la base de la acción gubernamental. Esta es obviamente una solución a largo plazo; aunque estas inversiones se hiciesen hoy, su impacto solo sería visible mucho más adelante.

A corto plazo, de una forma más inmediata, dos aspectos son críticos. El primero es la necesidad de dotarnos de nuevas organizaciones de la sociedad civil para la custodia, gestión y supervisión de los datos; ya hay varios experimentos en marcha en todo el mundo, como las cooperativas de datos y los fideicomisos de datos. Ceder a las personas el control sobre sus datos es fundamental: quien estructura, analiza y se beneficia de los datos puede fácilmente convertirse en un árbitro de quien tiene poder, estatus y reconocimiento en la sociedad.

En este sentido, debemos poner límites al alcance de la actual economía de datos extractiva; dicho de otro modo, no se trata tan solo de asegurar el uso ético y responsable de los datos, sino de plantearse en primer lugar si realmente necesitamos una recopilación exhaustiva de datos personales. En realidad, existen pocas evidencias de que la “personalización” mejora la rentabilidad empresarial. Es más, existe un número creciente de empresas líderes a nivel mundial que se alejan del *marketing* conductual para aproximarse a otras formas de publicidad más contextual. Del mismo modo, contamos con multitud de ejemplos de que formas alternativas y menos intrusivas de recolección de datos pueden ser igualmente rentables.

El segundo aspecto crítico a contemplar es la necesidad de abordar el dominio sobre el mercado que ejercen las grandes empresas tecnológicas. Las políticas de defensa de la competencia tendrían que actualizarse para considerar como indicador del bienestar del consumidor, no solo el precio de las cosas, sino también su control sobre los datos que genera. Exigir la neutralidad de las plataformas y un mayor control de las fusiones y adquisiciones podría crear un mercado más competitivo. Unos estándares abiertos podrían contribuir a derribar los terrenos vedados de las grandes empresas tecnológicas y a crear un ecosistema digital más descentralizado. Ninguna de estas soluciones es fácil y algunas de ellas podrían incluso tener otras consecuencias indeseables, pero todas están al alcance de los instrumentos legales tradicionales. Estas intervenciones políticas son especialmente importantes para los países en vías de desarrollo; en India, por ejemplo, los gigantes de la tecnología global están copando sectores propios del mercado y del Estado, y ejercen una cantidad enorme de poder cívico, en un momento en que los mecanismos de control y de rendición de cuentas brillan por su ausencia.

## ¿Cuáles son nuestras prioridades sociales o nuestras visiones de lo que constituye una “buena vida” y qué papel queremos que jueguen en esta vida los sistemas automatizados?

### Nuestros controvertidos futuros

Los debates acerca de la gobernanza de la IA se dirimen a menudo en los salones del poder, entre decisores políticos, líderes empresariales, intelectuales públicos influyentes y científicos destacados. Sin embargo, estos son debates que atañen a nuestro futuro compartido y en liza: ¿cuáles son nuestras prioridades sociales o nuestras visiones de lo que constituye una “buena vida” y qué papel queremos que jueguen en esta vida los sistemas automatizados? Podemos fijarnos en los movimientos en pro de la justicia social de todo el mundo para entender las necesidades, los deseos y las prioridades de la gente; estos movimientos a menudo reúnen o hablan de las preocupaciones de los miembros más marginalizados de nuestra sociedad. Nuestras visiones colectivas del futuro tienen que centrarse en estas preocupaciones o, de lo contrario, continuaremos reproduciendo los patrones existentes de injusticia y desigualdad social.

7. Véase Mhlambi, 2020.



Me vienen a la mente en este sentido cuatro movimientos activos hoy en India. Los agricultores del país han estado protestando contra un nuevo proyecto de ley agraria que privatizaría la agricultura. Más allá de los méritos o deméritos del proyecto de ley, el motivo central del descontento de los agricultores de la India es el hecho de no haber sido consultados en las decisiones tomadas sobre ellos y sobre su futuro. El segundo movimiento se inició a finales de 2019, cuando ciudadanos de todas las religiones y de todas las clases protestaron contra una propuesta de enmienda a la ley de ciudadanía india. Una vez más, y sin entrar directamente en el contenido político de la propia ley, las protestas pusieron claramente de manifiesto que las personas tienen identidades múltiples y fluidas, y reivindicaron que la fluidez de la identidad es importante tanto para la seguridad física como para la movilidad social. Por otro lado, los trabajadores temporales de las plataformas de reparto y transporte protestaron en todo el país contra unas condiciones de trabajo injustas y explotadoras. Hay un reconocimiento evidente de la distribución flagrantemente desigual del valor entre las plataformas y los trabajadores temporales a expensas de su bienestar. Y, finalmente, desde la propia juventud india ha surgido un movimiento que reclama atención sobre el cambio climático y sobre la salud de las generaciones futuras.

En estos cuatro movimientos sociales, los ciudadanos están tratando de gestionar el futuro y discutiendo qué constituye una “buena vida”; sus reivindicaciones giran en torno a la agencia, la identidad, el trabajo y el medio ambiente. Hemos de conectar todos estos debates, ver qué nos están diciendo estos movimientos sobre las diferentes visiones de una futura buena vida y encontrar de qué maneras la IA puede contribuir a hacer realidad estos futuros sin amenazarlos.

## REFERENCIAS BIBLIOGRÁFICAS

- Arun, Ch. “AI and the Global South: Designing for Other Worlds”, en: Dubber, M.; Pasquale, F. y Das, S. (eds.). *The Oxford Handbook of Ethics of AI*, julio de 2020.
- Gillespie, T. “The Relevance of Algorithms”. Presentación en el marco del seminario *Governing Algorithms*, 17 de mayo de 2013. Accesible en línea: <http://opentranscripts.org/transcript/relevance-of-algorithms/>
- Hoffman, A. L. “Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse”. *Information, Communication & Society*, vol. 22, nº 7: Data Justice, 2019.
- Katz, Y. “Manufacturing an Artificial Intelligence Revolution”. Harvard Medical School; Berkman Klein Center for Internet & Society, noviembre de 2017. Accesible en línea: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3078224](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3078224)
- Mhlambi, S. “From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance”. *Ethics and Governance of AI*, julio de 2020. Accesible en línea: <https://cyber.harvard.edu/story/2020-07/rationality-relationality-ubuntu-ethical-and-human-rights-framework-artificial>
- Pasquinelli, M. “How a Machine Learns and Fails - A Grammar of Error for Artificial Intelligence”. *Spheres, Journal for Digital Cultures: Spectres of AI*, n.º 5, noviembre de 2019. Accesible en línea: <https://spheres-journal.org/contribution/how-a-machine-learns-and-fails-a-grammar-of-error-for-artificial-intelligence>
- Raval, N.; Kak, N. y Calcaño, A. “A New AI Lexicon: Responses and Challenges to the Critical AI discourse”. *AI Now Institute en Medium*, 2021. Accesible en línea: <https://medium.com/a-new-ai-lexicon/a-new-ai-lexicon-responses-and-challenges-to-the-critical-ai-discourse-f2275989fa62>