

Involucrar a las comunidades para lograr una IA para todos



Tiempo de lectura: 7 min.

[Shobita Parthasarathy](#)

[Jared Katzman](#)

El futuro de la inteligencia artificial parece prometedor. Al fin y al cabo, la IA ya tiene muchas utilidades. Las herramientas de IA pueden proporcionar traducción de alta calidad en tiempo real para idiomas muy recurrentes como el inglés, un gran avance tecnológico. Los sistemas de IA también están preparados para aumentar la precisión de las pruebas de detección del cáncer y mejorar otros ámbitos de la salud. Sin embargo, gran parte del discurso en torno a la IA es más bien sombrío. No se trata solo de que la gente se preocupe por los posibles efectos de la IA generativa en su vida; las alteraciones en los empleos impulsadas por la innovación no son ni inusuales ni insuperables. Más preocupantes son las pruebas cada vez más numerosas de que los productos de los modelos de IA exacerbaban la desigualdad y la injusticia social

Es sabido que la tecnología de reconocimiento facial está demostrando ser una herramienta de opresión. Los informes de que la IA provoca detenciones erróneas de personas de raza negra se están convirtiendo en rutina. Los gobiernos de las ciudades estadounidenses están utilizando cámaras de reconocimiento facial para vigilar y controlar de forma agresiva a los residentes de viviendas de interés social, muchos de los cuales son negros. En contra de las esperanzas de que la IA redujera los prejuicios en la justicia penal, su uso hasta ahora ha magnificado las desigualdades estructurales del sistema. Mientras tanto, grandes empresas de IA, como OpenAI, explotan mano de obra en el extranjero para entrenar algoritmos. Y las herramientas de IA destinadas a beneficiar a las personas con discapacidad están teniendo el efecto contrario. Esta situación está perjudicando a las personas que ya se encuentran en situación de desventaja, al mismo tiempo que aumenta la desconfianza en la ciencia y el gobierno.

Las respuestas que se han puesto sobre la mesa a estos problemas de equidad y justicia suelen consistir en pequeños ajustes, a menudo de carácter técnico. La idea parece reducirse a que políticos, académicos y técnicos pueden resolver los problemas de la IA identificando sesgos estadísticos en conjuntos de datos, diseñando sistemas que sean más transparentes y explicables en su toma de decisiones, y ejerciendo supervisión. Por ejemplo, los expertos se preguntan cómo podrían las agencias gubernamentales evaluar la seguridad y eficacia de los algoritmos. Paralelamente, la industria tecnológica ha intentado educar a los desarrolladores sobre el impacto de los sesgos sociales en los algoritmos de IA y ha sugerido “soluciones de equidad” mínimas igualmente

centradas en los sesgos.

Debemos preguntarnos si realmente creemos que las personas marginadas deben contentarse con dejar sus destinos a merced de los tejesmanejos de gobiernos y empresas, cuando tales medidas han tenido escaso impacto en el pasado. ¿Dónde están las aportaciones, en esta ecuación, de las propias personas marginadas? Si nos preocupa la equidad en la era de la inteligencia artificial, ¿no deberían desempeñar un papel importante en la elaboración de la agenda de gobernanza quienes más tienen en juego?

Las mejores soluciones descansan en un ecosistema de innovación más inclusivo, en el que todos los actores –no solo los reguladores y los grupos de presión– asuman la responsabilidad de crear una IA equitativa y justa. Es importante que la IA no solo no discrimine, sino que se ponga proactivamente al servicio de toda la sociedad. En otras palabras, deberíamos pensar en cómo garantizar que la IA no esté aquí sólo para obtener ganancias, sino también para servir a todos, más allá de los beneficios económicos que puedan generar.

Con este fin, los financiadores públicos y filantrópicos de la investigación, las universidades y la industria tecnológica deberían buscar aliarse con las comunidades vulnerables para aprender lo que necesitan de la IA y construirlo. Los reguladores también deberían estar atentos al terreno, no solo a lo que dicen los altos cargos. Puede que los miembros típicos de una comunidad marginada –o, de hecho, de cualquier comunidad no experta– no conozcan los detalles técnicos de la IA, pero entienden mejor que nadie los desequilibrios de poder que dan lugar a las preocupaciones en torno al sesgo y la discriminación de esta tecnología. Así pues, es a las comunidades marginadas por la IA, y a los académicos y organizaciones que se dedican a comprender y mejorar las desventajas sociales, a quienes los diseñadores y reguladores más necesitan escuchar.

El avance hacia la equidad de la IA comienza en la fase de establecimiento de la agenda, cuando financiadores, ingenieros y líderes empresariales toman decisiones sobre las prioridades de investigación y desarrollo. Esto suele considerarse una tarea técnica o de gestión, que deben llevar a cabo expertos que conozcan el estado de la cuestión científica y las necesidades insatisfechas del mercado.

Pero, ¿estos expertos comprenden realmente qué necesidades deben atender? Cuando los expertos dirigen la innovación, están decidiendo qué problemas son importantes y cómo deben entenderse y resolverse. A menudo, los problemas considerados importantes son aquellos que, al ser abordados, producen beneficios. Pero a veces cuando los desarrolladores intentan resolver problemas sociales, normalmente lo hacen con una aportación mínima de las poblaciones más afectadas. Esto conduce a diagnósticos erróneos. Tal es la historia del programa One laptop per child (Una laptop por niño), por ejemplo. Desarrollada por el MIT Media Lab y financiada por donantes internacionales, la iniciativa buscaba mejorar la educación de los niños de familias con bajos ingresos de todo el mundo, garantizándoles el acceso a computadoras conectadas a internet. Pero el proyecto fracasó porque las computadoras no eran fáciles de usar para los niños, se rompían con frecuencia y eran difíciles de reparar, y dependían de electricidad que, en el mejor de los casos, estaba disponible de forma intermitente. Incluso cuando las laptops funcionaban, los contenidos que incluían contribuían poco a la consecución de los objetivos educativos locales.

Centrar a las comunidades marginadas en el establecimiento de la agenda de la IA ayudaría a evitar estos resultados. Un ejemplo alentador es el de la Universidad Carnegie Mellon, donde los informáticos trabajaron con residentes de Pittsburgh, ciudad natal de la institución en Pensilvania, para crear una tecnología que controlara y visualizara la calidad del aire local. La colaboración comenzó cuando los investigadores asistieron a reuniones comunitarias en las que escucharon a residentes que sufrían los efectos de la contaminación atmosférica de una fábrica cercana. Los residentes habían tenido dificultades para llamar la atención de las autoridades locales y nacionales porque no podían aportar el tipo de datos que motivaran el interés por su caso. Los investigadores se pusieron a trabajar en prototipos de sistemas que pudieran producir los datos necesarios y perfeccionaron su tecnología en respuesta a las aportaciones de la comunidad. Al final, su sistema reunió información heterogénea,

como informes de olores, vídeos de las chimeneas de las fábricas y datos sobre la calidad del aire y el viento, que los residentes enviaron a las autoridades. Tras examinar los datos, los administradores de la Agencia de Protección del Medio Ambiente de Estados Unidos accedieron a revisar el cumplimiento de la normativa por parte de la fábrica, y al cabo de un año la empresa matriz de la fábrica anunció el cierre de las instalaciones.

El diseño de abajo hacia arriba, en Pittsburgh y en otros lugares, requiere apertura y humildad por parte de los investigadores, reconocimiento de la experiencia de la comunidad y el deseo de capacitar a las personas marginadas. Este enfoque no solo ayuda a satisfacer necesidades reales, sino que también fomenta la confianza en la ciencia y la tecnología entre poblaciones sometidas a malos tratos y abandono.

Tal y como está constituida actualmente, la industria de la IA está mal equipada para respetar los conocimientos de las comunidades marginadas. Los líderes de la IA, tanto en el mundo empresarial como en el académico, constituyen un grupo demográficamente reducido, poco influido por el interés público. Los programas de responsabilidad social se tratan como complementarios y no como fundamentales para la misión, además son los primeros en desaparecer cuando las empresas y las universidades recortan los presupuestos.

¿Cómo garantizar que la actual generación de tecnólogos sea la última en actuar de este modo? ¿Cómo inculcar en los ingenieros –y en sus líderes– un interés genuino por los beneficios humanitarios de la IA y una mayor sensibilidad ante sus daños potenciales?

Una respuesta, aunque no la única, es la educación. Las universidades pueden renovar la forma en que enseñan ingeniería, integrando las humanidades y las ciencias sociales en el plan de estudios básico. En la actualidad, las universidades imponen una clara separación entre la ingeniería y el bien social; pueden exigir que los estudiantes de STEM cursen una única asignatura sobre ética profesional, mientras que el resto del plan de estudios enseña a los alumnos que la tecnología es neutral desde el punto de vista político y moral. Más útil sería diseñar cursos introductorios de ciencia e ingeniería que ayuden a los estudiantes a comprender los supuestos sociales y políticos que subyacen a decisiones aparentemente técnicas, así como las consecuencias de estos supuestos.

El conocimiento de abajo hacia arriba y la humildad para seguir aprendiendo de los necesitados son las herramientas para garantizar una IA responsable, pero también para aprovechar el inmenso potencial de esta tecnología emergente. La IA puede agravar los problemas sociales, pero también puede utilizarse para resolverlos. Además de su obligación de prevenir los daños, los responsables políticos, los financiadores de la investigación, los líderes tecnológicos y universitarios y los profesionales de STEM tienen la oportunidad de fomentar la equidad a través de la innovación. Ahí es donde reside la verdadera promesa de la IA.

Este texto es un extracto de un artículo publicado originalmente en Issues in Science and Technology, revista trimestral de la National Academies of Sciences, Engineering, and Medicine y Arizona State University

Letras Libres

No.279 / diciembre 2024

<https://letraslibres.com/ciencia-tecnologia/parthasarathy-katzman-comunidades-inteligencia-artificial/>

[ver PDF](#)

Copied to clipboard