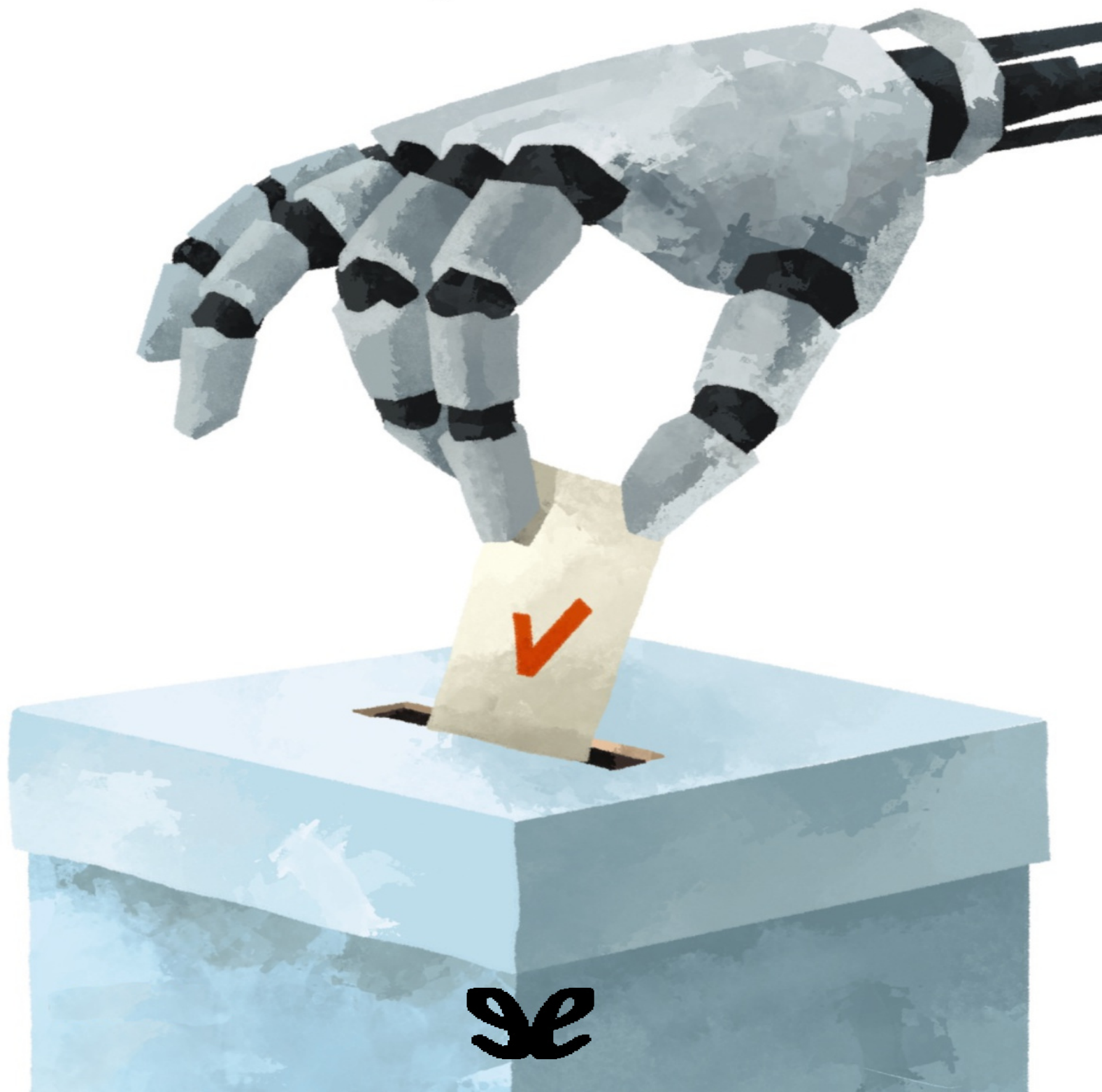


POR QUÉ LA IA DEBILITA LA DEMOCRACIA Y QUÉ HACER AL RESPECTO

Mark Coeckelbergh



En un contexto en el que la inteligencia artificial (IA) está pasando de los universos de la ciencia ficción a formar parte de nuestras vidas cotidianas, resulta imprescindible entender cuáles son las amenazas que esta plantea a corto y medio plazo, así como de qué manera afecta positiva y negativamente a nuestras instituciones. En este libro, Mark Coeckelbergh, reputado especialista en el campo del desarrollo y la regulación de la inteligencia artificial, se detiene a examinar las razones y los modos por los cuales una implantación desregulada de la IA como la que estamos viviendo afecta a nuestras democracias, debilitando principios tan importantes para su pervivencia como lo son los de la libertad individual y la igualdad entre los ciudadanos. Ahora bien, ¿cómo debemos actuar ante estas amenazas?, ¿por medio de qué estrategias podemos reforzar la democracia para resistir? Para responder a estas preguntas, el autor reivindica no solo la búsqueda y la implantación de tecnologías más democráticas desde su base, sino también la creación de nuevas instituciones políticas, así como un replanteamiento de la educación que recibimos en relación con la tecnología y su utilización para asegurarnos de que la inteligencia artificial termine resultando, en lugar de un peligro, un bien común para los seres humanos.



Mark Coeckelbergh

Por qué la IA debilita la democracia y qué hacer al respecto

ePub r1.0

Titivillus 17.03.2025

Título original: *Why AI Undermines Democracy and What To Do About It*

Mark Coeckelbergh, 2024

Traducción: Lucas Álvarez Canga

Editor digital: Titivillus

ePub base r3.0 (ePub 3)



Por la naturaleza de la red, puede que algunos de los vínculos a páginas web en el libro ya no sean accesibles.



Prefacio

Los tiempos que vivimos son de mucho trabajo para cualquier persona interesada en la ética y en la política de la inteligencia artificial (IA). En este momento, mientras me ocupo de los retoques finales del manuscrito, los grandes modelos lingüísticos para la generación de textos viven un momento de pleno auge (desde noviembre de 2022); los CEO de diversas empresas tecnológicas, y los «expertos» que los guían^[1], lanzan afirmaciones descabelladas sobre el impacto mundial y supuestamente catastrófico de la IA; en junio de 2022, el Parlamento Europeo aprueba el Acta de IA, que se presenta como su primera regulación mundial completa^[2]. ¿Qué está ocurriendo? ¿Es la IA una amenaza para el mundo? ¿Y cuál será su impacto en la legislación? Los especialistas en ética de IA están muy solicitados. Los debates públicos en torno a esta surgen por todas partes. La IA se convierte en el objeto de la lucha (geo)política. CEO de empresas tecnológicas como Sam Altman consiguen dominar el debate público. Practican lobby abiertamente y son recibidos en la Casa Blanca, así como por los jefes de Estado europeos^[3]. A los ciudadanos y a los gobiernos se nos dice que tenemos que desarrollar la IA con rapidez si no queremos que otros se nos adelanten. Se nos dice cómo y cuánto debe regularse, que puede dar lugar a efectos catastróficos y que debemos ralentizar su desarrollo^[4]. Los acontecimientos se suceden con rapidez, igual que el coro de los que dan bombo o son catastrofistas con respecto al crecimiento de la IA, impulsado a menudo por los medios de comunicación.

En medio de este revuelo, creado a veces con fines políticos concretos, no suelen oírse las voces de quienes analizan, detenida y metódicamente, el impacto de la IA en la sociedad. No obstante, se trata de una tarea de gran relevancia si queremos aprovechar las oportunidades que nos ofrece esta tecnología y utilizarla de manera responsable. Centrándose en la influencia de la IA en la democracia, este libro discute las políticas relativas a esta tecnología situándolas en un contexto intelectual más amplio y ofrece una visión de futuro para seguir avanzando. Hablaremos de la IA, por supuesto, pero también, y sobre todo, de la democracia. De lo vulnerables y resilientes que son y pueden ser nuestras formas de democracia actuales a la luz de poderosas fuerzas tecnológicas y antidemocráticas. Y, especialmente, de qué tipo de democracia queremos. Finalmente, reflexionaremos acerca de qué tipo de sociedad perseguimos y qué tipo de mundo deseamos dejar a las generaciones futuras. ¿Nos decantamos por la situación actual o intentamos asegurarnos de que las tecnologías avanzadas nos ayuden a trabajar por el bien común? Esto (y no los intereses corporativos a corto plazo ni las fantasías de ciencia ficción sobre la IA que destruyen la civilización) es lo que debería entrar en juego cuando discutimos acerca del futuro tecnológico.

KIOTO, 21 de junio de 2023

Índice de contenido

Prefacio

Capítulo 1. Introducción

Acerca de este libro

Capítulo 2. Una historia no tan democrática

Política y cibernética

Historias de la tecnología y la democracia

Capítulo 3. ¿Qué IA? ¿Qué democracia?

Inteligencia artificial (IA)

Democracia

Capítulo 4. Cómo debilita la IA los principios básicos de la democracia

Principios fundacionales de la democracia:

Libertad, igualdad, fraternidad, imperio de la ley y tolerancia

Cómo debilita la IA los fundamentos de la democracia

Capítulo 5. Cómo erosiona la IA el conocimiento y la confianza

El problema del totalitarismo revisado:

Las lecciones de Hannah Arendt

Algunos desafíos en relación con la IA,

El conocimiento y la democracia

Capítulo 6. Fortalecer la democracia y democratizar la IA

Fortalecer y crear nuevas instituciones democráticas

Regulación y supervisión

IA democrática

Capítulo 7. IA para la democracia y un nuevo Renacimiento

IA para la democracia

Un nuevo Renacimiento
Humanismo digital
Conclusión

Capítulo 8. El bien común y la comunicación

El bien común y los comunes
Comunicación y comunidad
Experiencia común y construcción de un mundo común
Resumen ejecutivo para responsables políticos

Sobre el autor

Capítulo 1

Introducción

Un fantasma se cierne sobre el panorama político actual, y es un fantasma horrible y peligroso: el autoritarismo. Las democracias se encuentran en peligro en todas partes, incluso en Occidente. A pesar de que a finales de los noventa éramos testigos de una ola de democratización, actualmente existen muchas tendencias antidemocráticas que a veces implican una caída en el autoritarismo. Un informe de 2021 muestra que las dictaduras están creciendo a nivel mundial. La polarización empeora, tiene lugar un aumento dramático de amenazas a la libertad de expresión, y el 70 % de la población mundial vive en una autocracia, en comparación con el 49 % que lo hacía en 2011 (Democracy Report, 2022). Las democracias occidentales no son inmunes a esta tendencia. Algunos hablan de un nuevo orden mundial, con poderosos actores intentando destruir el orden internacional constituido tras la Segunda Guerra Mundial, y con los Estados Unidos como presa de la polarización y la «decadencia» (Erlanger, 2022). También existen tendencias autocráticas en Europa, donde a menudo se presentan como bandazos hacia la (extrema) derecha; por ejemplo, en Hungría, Polonia y Serbia. En septiembre de 2022, un partido nacionalista sueco de derechas obtuvo más del 20 % de los votos. El

Reino Unido se volvió política y financieramente inestable después de que los populistas y, más tarde, los conservadores ultraliberales alcanzaran el poder. En el mismo año, los populistas de extrema derecha ganaron las elecciones en Italia: su líder Giorgia Meloni se convirtió en primera ministra. Como es bien sabido por la historia, los políticos antidemocráticos pueden alcanzar el poder a través de elecciones democráticas y, a continuación, debilitar o incluso abolir la democracia. Actualmente, y en determinados contextos, esto supone un peligro inminente.

Las tecnologías digitales como la inteligencia artificial (IA) ofrecen a la sociedad muchos beneficios y oportunidades. Pero también parecen desempeñar cierto papel en esas erosiones de las democracias, y en el ascenso y mantenimiento de los regímenes autoritarios y totalitarios. Se culpa a las redes sociales de ayudar a desestabilizar las democracias al destruir la verdad e incrementar la polarización. A la IA no le va mejor. Hoy en día, las historias acerca de esta versan a menudo sobre manipulación, polarización, discriminación, vigilancia, poder y represión.

Es posible que el autoritarismo no se encuentre en un horizonte inmediato, pero, incluso así, los riesgos para la democracia parecen muy reales. Los gobiernos y las organizaciones internacionales están preocupados. La administración Biden ha advertido recientemente en Estados Unidos sobre los peligros que representa la IA para la democracia, quejándose de que hay límites a lo que la Casa Blanca puede hacer para regular la tecnología^[1]. Una página web de la Comisión Europea llevaba como título «democracia en peligro» a la luz de un informe en torno a los riesgos que plantean las actuales tecnologías digitales, como la información falsa, la manipulación, la vigilancia y el aumento de poder de las entidades comerciales de las que dependemos en esta área y que establecen la agenda de nuestro futuro digital^[2]. Y antes, la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos advertía del impacto de la IA en tales derechos, el imperio de la ley y la

democracia^[3]. En otras palabras, la IA ha pasado a considerarse un problema, y cada vez se reconoce con mayor nitidez que plantea inconvenientes para la democracia.

Considérese el caso de Cambridge Analytica (Cadwalladr y Graham-Harrison, 2018), que incluyó la manipulación de los votantes basándose en el análisis de *big data*. Se recolectaron millones de datos de Facebook sin el consentimiento de la gente y se usaron en publicidad política dirigida a apoyar campañas políticas en Reino Unido y en los Estados Unidos. Se ha dicho que el uso de la IA en combinación con las redes sociales conduce a la polarización política y propaga divisiones en la sociedad que, a continuación, pueden ser explotadas por grupos en lucha por el poder (Smith, 2019), grupos que no son necesariamente democráticos. El auge del movimiento de extrema derecha QAnon en los Estados Unidos, que condujo a una irrupción violenta en el Capitolio, es un buen ejemplo. Parece que nos arriesgamos a encerrarnos en nuestras propias burbujas y cámaras de eco, asediados por algoritmos que intentan influenciarnos y separarnos.

También el programa ChatGPT, un extenso modelo lingüístico que recientemente ha adquirido gran popularidad, y que al mismo tiempo ha resultado polémico desde un punto de vista ético, ha sido relacionado con los intentos de debilitar la democracia. Algunos están preocupados por la posibilidad de que la IA se des controle y acabe asumiendo la toma de decisiones políticas^[4]. Esto puede parecer más bien rebuscado y, al menos, un problema para un futuro lejano. Pero existe también la preocupación de que en un futuro cercano la IA se *utilice* para influir en la toma de decisiones políticas. Para empezar, podría ser un poderoso instrumento de presión. Por ejemplo, podría redactar automáticamente un gran número de artículos de opinión y cartas al editor, enviar numerosos comentarios sobre publicaciones (*posts*) en redes sociales, y prestar ayuda para señalar a políticos y a otros actores relevantes: todo a gran velocidad y a nivel mundial. Esto podría incidir

significativamente en la toma de decisiones políticas (Sanders y Schneier, 2023). Asimismo, podría usarse para difundir propaganda, influyendo en las elecciones.

Sin embargo, la IA no se utiliza solamente para obtener poder, sino que también desempeña un papel cada vez más relevante en las instituciones de gobierno existentes. También aquí se ha mostrado una imagen negativa de la IA. Considérese el sistema automatizado de vigilancia del bienestar usado por los holandeses, que detuvo un tribunal porque consideró que violaba los derechos humanos y que vulneraba la privacidad de las personas: ¿se equiparaba el uso de este sistema con «espiar a los pobres» (Hentley y Booth, 2020)? En Austria hubo una controversia en torno a los perfiles algorítmicos de los solicitantes de empleo por parte del servicio de empleo público AMS, que fue acusado de discriminar injustamente algunas categorías de solicitantes de empleo (Allhutter *et al.*, 2020). La toma de decisiones judiciales basadas en IA también ha sido criticada por estar sesgada. En los Estados Unidos se ha dicho que el algoritmo COMPAS, usado por agentes que se encargan de asuntos relacionados con la libertad condicional y vigilada para juzgar el riesgo de reincidencia, discrimina negativamente a los acusados negros: un informe afirmaba que «era mucho más probable que tales acusados fueran considerados erróneamente, por lo que ofrecían un mayor riesgo de reincidencia que los acusados blancos» (Larsen *et al.*, 2016).

Mientras tanto, la IA también se ha vuelto popular en los gobiernos autocráticos. Los medios de comunicación occidentales informaron de que China había estado usándola para vigilar y reprimir. De acuerdo con el *New York Times*, sus ciudadanos se encuentran bajo una vigilancia constante: se rastrean los teléfonos, se monitorizan las compras, se censuran los chats. Se usa la policía predictiva para prever el crimen, pero también para reprimir a las minorías étnicas y a los trabajadores migrantes (Mozur, Xiao y Liu, 2022). Human Rights Watch afirma que el Gobierno chino recaba

mucha información personal y usa algoritmos para marcar a las personas que se consideran una amenaza potencial. Dice esa agencia que esta actividad ha conducido a restricciones de la libertad de expresión y la libertad de movimiento. Se envían algunas personas a campos de educación política (algoritmo de represión de China, 2019).

Pero el uso de la tecnología de vigilancia con IA no está restringido a China y, ni siquiera, a los regímenes autoritarios. Desarrollada y suministrada por China, pero también por países como los Estados Unidos, Francia, Alemania, Israel y Japón, está proliferando en todo el mundo e, incluso, en las democracias. De acuerdo con el índice de Carnegie AI Global Surveillance (AIGS) (Vigilancia Global IA), más de la mitad de las democracias desarrolladas del mundo usan sistemas de vigilancia con IA (Feldstein, 2019a). Incluso si tales tecnologías se usan en sistemas políticos que se llaman a sí mismos democráticos, siempre existe el riesgo de que se haga con fines represivos. En 2020, el presidente de extrema derecha de Brasil, Bolsonaro, fue acusado de «tecnoautoritarismo» por generar grandes conjuntos de datos e infraestructuras de vigilancia, en particular, el Registro Básico del Ciudadano [Citizen's Basic Register] que reúne datos de ciudadanos que van desde registros médicos hasta información biométrica (Kemeny, 2020). E incluso en Europa y los Estados Unidos, la pandemia del covid-19 se ha utilizado para movilizar la IA y otras tecnologías digitales al servicio de una aplicación más estricta de la ley y el control de la población. Por ejemplo, en 2020 Minnesota usó el rastreo de contactos para seguir la pista de manifestantes que participaron en las protestas a raíz del asesinato policial de George Floyd (Meek, 2020), y la IA se ha utilizado ampliamente para apoyar las medidas de control de gran alcance a propósito de la movilidad de la población.

Parece que nos encontramos ante la emergencia de nuevas formas de autoritarismo y totalitarismo apoyadas en tecnologías

digitales, por lo que ya es hora de reflexionar sobre la IA y la democracia, y de repetir la pregunta de Hannah Arendt en *Los orígenes del totalitarismo* (2017 [1951]) tras la Segunda Guerra Mundial, pero esta vez en el contexto de las tecnologías digitales. A pesar de las diferencias en los contextos históricos específicos, ¿está la democracia en peligro? ¿Y existen hoy en día las condiciones para que el totalitarismo vuelva a surgir? ¿De qué forma contribuyen las tecnologías digitales a estas condiciones? ¿Debilita la IA la democracia y conduce a nuevas formas de autoritarismo y totalitarismo (llamémoslas «autoritarismo digital» y «totalitarismo digital»)[5] y, si es así, cómo funciona y qué podemos hacer al respecto? Y de una forma más general, y más allá de la cuestión relacionada con el totalitarismo, ¿cuál es el impacto de la IA en la democracia? ¿Es la IA buena para la democracia?, y, si no, ¿qué se puede hacer al respecto? ¿Cómo podemos asegurarnos de que la IA apoya a la democracia?

ACERCA DE ESTE LIBRO

Este libro defiende que la IA, según se está desarrollando y usando actualmente, *debilita los principios fundamentales y la base del conocimiento sobre los cuales se construyen nuestras democracias, por lo que no contribuye al bien común*. Tras situar la cuestión en un contexto histórico y analizarla, guiado por las teorías político-filosóficas de la democracia, ofrece un itinerario de algunos de los principales riesgos que la IA supone para la democracia. Muestra que no es políticamente neutra, sino que conforma nuestros sistemas políticos de maneras que amenazan a la democracia y apoya tendencias antidemocráticas al debilitar los principios democráticos, al erosionar el conocimiento y la confianza necesarios en ella, propiciando el bien de unos pocos a expensas del de la mayoría.

Pero el libro también propone una salida. Aboga por reconocer plenamente el carácter político de la IA y señala la necesidad de la deliberación y el liderazgo públicos para caminar en una dirección más democrática, tanto en el caso de nuestras instituciones políticas como de la tecnología. La IA necesita ser interpretada de una manera menos dañina para la democracia y, preferiblemente, debería trabajar en favor de la democracia. Esto requiere (re)formar la tecnología en la fase de desarrollo e integrarla de modo que apoye, en lugar de debilitarlo, el sistema político actual. Sin embargo, el libro defiende que tal cambio en el liderazgo político y en el desarrollo democrático de esta tecnología solo puede tener éxito si existe una base político-epistémica adecuada. Tal base puede construirse alimentando una cultura y una educación inspiradas por el Renacimiento, el humanismo, el republicanismo y la Ilustración y, de forma más general, creando un conocimiento y una experiencia compartidos. Se defiende que, en última instancia, la democracia, así como la IA para la democracia, depende de la producción y la promoción del bien común, de la comunicación y de la construcción de un mundo (más) común. La IA y las tecnologías digitales pueden, y deberían, apoyar este proyecto en lugar de dificultarlo.

Permítaseme ofrecer una hoja de ruta rápida del libro:

Sin duda, la IA plantea serios riesgos para la democracia y es necesario que se comprendan mejor estos riesgos. Después de presentar una perspectiva histórica de la relación entre la tecnología y la democracia (capítulo 2), que muestra que las nuevas tecnologías han conducido a menudo a más centralización, pero que también subraya que no existe un determinismo por lo que respecta a la influencia de la tecnología en la política, el capítulo 3 analiza los dos conceptos principales en cuestión: ¿qué queremos decir con IA e (importante para esta investigación) qué queremos decir con *democracia*? Es necesario vincular las discusiones en torno a la IA con la filosofía y la teoría políticas, ya que el concepto de democracia

no es en modo alguno claro ni indiscutible. En lo que respecta a mi propia posición, me distancio de las definiciones de democracia en términos de votación y apunto hacia concepciones más ricas: argumento que los ideales deliberativos, participativos y republicanos de la democracia deberían guiarnos a la hora de discutir el problema de la IA y la democracia.

Los capítulos siguientes se concentran en cómo la IA, según se está utilizando y desarrollando actualmente, pone en peligro a la democracia: muestro cómo aquella tiene impacto sobre los principios liberaldemócratas y republicanos fundamentales, como la libertad, la igualdad, la fraternidad, el imperio de la ley y la tolerancia, con formas que socavan estos principios, poniendo así en peligro la democracia liberal y conduciendo potencialmente al autoritarismo y al totalitarismo (capítulo 4). Siguiendo los pasos de Hannah Arendt y en relación con el trabajo contemporáneo sobre la ética y la política de las tecnologías digitales, también muestro (en el capítulo 5) cómo su uso corre el riesgo de debilitar el conocimiento básico y la confianza en la democracia a través de la creación de asimetrías de poder, la manipulación, la erosión de la distinción entre lo real y lo falso, y la creación de burbujas epistémicas. Al final, la IA puede incluso destruir la confianza *entre* los ciudadanos y amenazar nuestra propia imagen en tanto que sujetos políticos autónomos, imagen que hemos mantenido como un bien tanpreciado desde la Ilustración.

Pero no hay necesidad de ser catastrofistas. Este libro ni es pesimista ni es antitecnología, y rechaza las opiniones tecnodeterministas que consideran que los desarrollos sociotecnológicos son autónomos y están más allá del alcance humano; aún existe tiempo y espacio para intervenir y mejorar las cosas. Después de un análisis de los problemas, considero cómo arreglar esto y explorar enfoques constructivos. ¿Cómo podemos hacer que la democracia sea más resiliente a la luz de la IA? Y, de forma más optimista, ¿qué puede hacer la IA *por* la democracia? En

los últimos capítulos (llamémoslo mi manifiesto «tecnodemocrático» o «IA democrática») argumento que necesitamos cambiar no solo nuestras instituciones políticas y regular la IA para que la democracia sea más fuerte, sino también *hacer que la IA sea más democrática*. La tecnología no debe darse por supuesta, podemos cambiarla: el desarrollo de la IA debería democratizarse. En el capítulo 6, argumento que para esto se necesitan cambios en el desarrollo de la tecnología y su vinculación, ya que se encuentra incrustada en ellas, con las instituciones políticas democráticas. Sin embargo, en el capítulo 7 subrayo que el proyecto de democratizar la IA no solo implica hacerla más democrática y a prueba de totalitarismo, sino también, de forma menos defensiva y más constructiva, crear una IA *para* la democracia. ¿Cómo lograr que apoye la democracia en lugar de erosionarla? Discutiré brevemente algunos trabajos que intentan hacer esto.

Con todo, al final del libro concluyo que rediseñar nuestras tecnologías y reformar nuestras instituciones políticas democráticas no es suficiente. Defiendo que el proyecto de una IA democrática solo puede tener éxito si se integra en un nuevo entorno cultural y educativo: una renovación de nuestra cultura política necesita un Renacimiento y una Ilustración nuevos, ayudados esta vez por las tecnologías digitales. Además, las políticas democráticas (y, ciertamente, la política tecnológica) también requieren un tipo de transformación normativa más profunda. En el capítulo final, el capítulo 8, defiendo que, si realmente nos preocupa la democracia, evitar de verdad las tendencias antidemocráticas y el auge del autoritarismo y el totalitarismo, necesitamos que la IA y otras tecnologías digitales nos ayuden a realizar y encontrar el bien común, y a *comunicar* y construir realmente un mundo común.

Capítulo 2

Una historia no tan democrática

Plantear preguntas en relación con la IA y la democracia puede parecer extraño, ya que normalmente no relacionamos los dos términos. Esto es así en parte porque tendemos a ver la IA y otras tecnologías como meras herramientas, como instrumentos. Las vemos como medios que no alcanzan los fines: nuestras metas y nuestros valores humanos. Muchas personas suponen que la tecnología misma no tiene mucho que ver con la política y la democracia. Se preguntan: «¿Seguramente todo depende de para qué se use? ¿De qué forma está conectada la IA con la democracia»? Pero, como sucede con otras tecnologías, la IA es algo más que una herramienta. En la ética de la tecnología, una forma común de expresarlo es reconocer la verdad que hay en el dicho «Las armas matan personas». Por supuesto, las personas matan a otras personas (con pistolas), pero la herramienta importa en sentidos que no dependen simplemente de la intención de las personas, pues permite y alienta la acción de matar. Sin pistolas, habría menos asesinatos. Lo mismo ocurre con la política de la tecnología. La IA no es simplemente un instrumento. Moldea nuestras acciones y nuestros fines. Influye en la sociedad. Beneficia a la sociedad, pero también crea riesgos que, a menudo, son

involuntarios e imprevistos. Ciertamente, su impacto depende en parte de lo que las personas hacen con ella, pero su influencia política es más profunda, y sus efectos políticos más «internos» de lo que normalmente se supone que es y hace la IA. Su impacto político no es simplemente el que tiene sobre la política propiamente dicha y lo que hacen los políticos, sino que también está relacionado con lo que hace y permite la *tecnología*.

Inspirado por muchas décadas de trabajo en la filosofía de la tecnología, he formulado esta visión de la siguiente manera: «la IA es política de principio a fin» (Coeckelbergh, 2022a: 5). La cuestión no está en que la IA sea cosa de un político concreto, sino más bien en que, en función de cómo la utilicemos los seres humanos, podría tener consecuencias políticas que trasciendan las previstas y que se extenderían más allá de la esfera de la ciencia y de la tecnologías, pues estas inteligencias tienen el potencial de conformar nuestras sociedades y las formas en que las gobernamos.

Considérese de nuevo ChatGPT: no es simplemente una herramienta de generación de textos, sino que es probable que transforme el modo en que escribimos, la forma en que trabajan los profesionales (considérese, por ejemplo, el periodismo), y, como ya he sugerido con mis ejemplos de la Introducción, la forma en que hacemos política. Estos efectos no siempre son deseados. Piénsese también en internet y cómo ha transformado ya nuestras sociedades: no solo la tecnología, también el hecho de que sus múltiples efectos no se habían previsto. En este sentido, no se trata simplemente de que los políticos utilicen IA, sino de que la propia IA es política en sí misma.

Este enfoque, que se irá aceptando a la vez que desgranando a lo largo de este libro, nos permite hacernos la siguiente pregunta: ¿en qué dirección política nos empuja la IA? ¿Hace que el mundo sea menos democrático y más autoritario y, si es así, quiere decir que estamos indefensos y entregados a sus fuerzas históricas? ¿Es inevitable el tecnoautoritarismo?

Para comprender mejor el problema concerniente a la IA y la democracia, para mostrar cómo la IA y los problemas con la democracia tienen más que ver una con la otra de lo que podríamos esperar, y para discutir con mayor profundidad la cuestión acerca de si la tecnología determina la política, propongo poner estas cuestiones en una perspectiva histórica (incluyendo parte de la historia de las ideas) y reflexionar de una forma más general sobre la relación entre la tecnología y la democracia.

POLÍTICA Y CIBERNÉTICA

Comencemos con la filosofía antigua. Cuando en *la República* se habla del arte y la ciencia de gobernar el Estado (2019: 488a-489e), Platón compara ese gobierno con dirigir y poner a navegar un barco. Usa el término *kybernetes* (en griego: κυβερνητής): el timonero o timonel del barco, el piloto, el que es bueno, ingenioso y habilidoso a la hora de dirigir y navegar. Dirigir un barco es un arte: requiere conocimiento experto y no se puede dejar a los marineros. También requiere el conocimiento de la navegación, por ejemplo, conocimiento de las estrellas. De forma similar, Platón argumenta que el hombre de Estado debería aprender el arte y la ciencia de dirigir el barco del Estado: es también una cuestión de cibernética. Debería saber cómo hacerlo: requiere pericia y no se puede dejar al arbitrio de la gente, a un puñado de ignorantes, pendencieros y, a menudo, borrachos inútiles.

A primera vista, la cibernética parece una metáfora atractiva y productiva para la política. Nuestro término contemporáneo «gobernanza» deriva de cibernética: se refiere a dirigir la sociedad. La cibernética depende de la retroalimentación y está orientada a objetivos. El piloto tiene que monitorizar continuamente los movimientos del barco y ajustarlos, si fuera necesario, para lograr su objetivo. Además, dirigir un barco antiguo requiere coordinar a

gente, por ejemplo, la coordinación de los remeros. Hoy en día, la navegación o coordinación de unidades políticas complejas y de gran tamaño, como los Estados nación y las sociedades impregnadas de tecnologías, parece ciertamente necesitar conocimiento experto. ¿Se puede dejar en manos de cualquiera? ¿No es eso peligroso? Si la política es cibernética, parece que al barco del Estado le va mejor cuando los expertos lo dirigen.

Sin embargo, actualmente la mayoría de la gente no relaciona de manera inmediata la cibernética con la política. Esto es así porque a mediados del siglo xx el profesor de Matemáticas americano Norbert Wiener usó el término para «la comunicación y control del animal y la máquina» (Wiener, 1948), introduciendo, así, el significado contemporáneo de la cibernética e influyendo significativamente en la historia de la informática. Actualmente muchas máquinas son cibernéticas en el sentido de que son capaces de dirigirse ellas mismas automáticamente, sin un control humano directo. Ya disponemos de pilotos automáticos en los aviones y, cada vez más, en los coches. ¿Deberíamos también poner un piloto automático en el barco del Estado? ¿La sociedad debería estar controlada como si fuera un coche autónomo gigante? ¿Deberíamos emplear la IA para guiar nuestras sociedades y coordinar a la gente para lidiar con desafíos locales y globales cada vez más complejos?

Esto no suena muy democrático, y tampoco lo era la idea platónica de un timonel guiando al Estado. Como ya sugiere mi resumen de esta perspectiva, Platón desconfiaba de la democracia. Reservaba el derecho de dirigir el Estado a una elite, preferiblemente un rey-filósofo o, más específicamente, una clase de reyes-filósofos. Estos *guardianes* serían virtuosos, captarían la verdad y estarían dedicados al bien de la ciudad-Estado. Hoy en día, esa idea puede tomar la forma de guardianes que usen la IA para gobernar los Estados nación. Esto no es democrático, sino autoritario. Equivale a la tecnocracia, o al menos a la tecnocracia en su forma autoritaria: el gobierno de los expertos. Basándonos en este análisis, resulta que

tanto la IA como la cibernética parecen estar en tensión con la democracia. Pero ¿cuál es, exactamente, el problema? ¿Es la IA *necesariamente* mala para la democracia?

Esta pregunta es oportuna. La idea de que la IA *asuma* el control político bien puede continuar siendo un escenario de ciencia ficción. Pero actualmente se está usando cada vez más para gobernar. Ya he mencionado que se utilizó durante la pandemia del covid-19 y, a menudo, de formas antidemocráticas. En China, Corea del Sur, Singapur y Taiwán, la IA se usó, por ejemplo, para respaldar la vigilancia de masas y el control de la población a través de teléfonos móviles que permitían a la policía rastrear los movimientos de las personas infectadas. Es más, las compañías tecnológicas llegaron a pedir al Gobierno estadounidense que les diera permiso para acceder a los datos de los ciudadanos (Consejo de Europa, 2020). Además, las decisiones con respecto al futuro de la tecnología y de la sociedad evaden a menudo los procesos y procedimientos democráticos. El papel que desempeña actualmente la IA en nuestra sociedad no es algo que se haya decidido por los ciudadanos o por gobiernos elegidos democráticamente. Por el contrario, la IA y las tecnologías relacionadas se han usado para manipular elecciones democráticas, y el futuro de la tecnología moderna se decide normalmente en las salas de juntas y en los laboratorios de Silicon Valley. Esto no es solo relevante para la política de los Estados Unidos. La influencia de las grandes tecnológicas tiene alcance mundial. Como señala Lucie Greene en *Silicon States* (2018), Silicon Valley es actualmente una «central eléctrica global». Lo mismo es cierto en el caso de otros centros de innovación y de las grandes empresas tecnológicas de China, donde el Gobierno desempeña un papel más importante a la hora de decidir el futuro. Pero el resultado es el mismo: la política tecnológica no es democrática. A través de sus tecnologías, una pequeña elite relacionada con unas pocas corporaciones globales decide cómo vivirá la gente su vida en Norteamérica, Europa, Sudamérica, Asia y África. La IA parece ser

parte de este imperio tecnológico. ¿Es la IA un enemigo de la democracia, como sugiere la historia del concepto de cibernética? ¿Hasta qué punto es resiliente la democracia ante la IA?

Para comprender mejor la cuestión, examinemos brevemente el problema más general de la relación entre la tecnología y la democracia. La tensión entre la IA y la democracia debe entenderse en el contexto de una historia más larga de relaciones problemáticas entre tecnología y democracia, una historia que se remonta hasta el amanecer de la civilización humana. Y no es una historia feliz.

HISTORIAS DE LA TECNOLOGÍA Y LA DEMOCRACIA

El hecho de mirar hacia la historia de la tecnología y la política no tiene como objeto mostrar que las tecnologías pueden ser, y siempre han sido, instrumentos de poder. Esta es una cuestión trivial. Las tecnologías siempre se han usado para gobernar a las personas, incluso si a menudo no eran conscientes de esta historia. Desde tiempos antiguos, se han utilizado en procedimientos políticos: como señala Risse, la democracia ateniense dependía del uso de tecnologías para la distribución, el control de horarios y la votación (2023: 51-52). Es más interesante el que las tecnologías también tengan y hayan tenido un impacto más amplio en la sociedad, incluyendo influencias no previstas. Las nuevas tecnologías facilitaron no solo nuevas formas de gobierno, sino nuevas formas de vida. En los tiempos antiguos, en Mesopotamia y Egipto se desarrollaron la metalurgia, las tecnologías de construcción de, por ejemplo, cuñas y palancas, la rueda, los sistemas de regadío, la escritura, el papel, los instrumentos de cirugía y, por supuesto, todo tipo de herramientas militares. Cada una de estas tecnologías servía para un propósito particular, pero también ayudaron a construir civilizaciones de forma satisfactoria, y fueron decisivas para ayudar a los gobernantes y a las elites a establecer y mantener su poder.

Considérese también el Imperio romano: su poder centralizado era posible gracias a 80 000 km de carreteras bien diseñadas, que conectaban la capital con las fronteras del Imperio. El dicho medieval «Todos los caminos llevan a Roma» puede interpretarse como una expresión de esta conexión entre tecnología y poder.

Sin embargo, no fue hasta principios de la era moderna cuando la relación entre la tecnología y el poder se hizo explícita a la luz de una nueva y moderna comprensión de la tecnología y del mundo. En *La Nueva Atlántida* (1999 [1626]), el político y filósofo inglés Francis Bacon (1999 [1626]), el político y filósofo inglés Francis Bacon describe una sociedad utópica en la que los humanos son capaces de controlar su destino y asumen el control de la naturaleza mediante la ciencia y la tecnología. La ciencia moderna se vio alimentada por esta aspiración. Sin embargo, el control de la naturaleza siempre ha venido acompañado del control de las personas. Las civilizaciones antiguas usaron la ciencia y la tecnología para recaudar impuestos y vigilar a la gente. La escritura misma, tan importante en las culturas humanísticas desde entonces, no era un regalo del dios Toth (el dios del conocimiento), como creían los antiguos egipcios, ni había sido robada al dios de la sabiduría, Enki, como creían los mesopotámicos, sino que, probablemente, fue inventada por los contables. Los contables y administradores de las ciudades antiguas como Uruk necesitaban la ayuda de la escritura y de los números^[1]. Estas herramientas permitían que la gente dirigiera estas economías antiguas, pero también permitía que algunas personas ejercieran el poder sobre otras. Desde entonces, todos los gobernantes han usado estos instrumentos de poder. Y a medida que crecieron las unidades políticas, la recolección de datos sobre los ciudadanos se volvió más importante y se utilizó para controlarlos. Las naciones Estado están especialmente hambrientas de datos. Mucho antes del gobierno contemporáneo (científico) mediante datos, las naciones Estado del siglo XIX ya recogían ya recogían datos de sus ciudadanos para disciplinarlos y gobernarlos.

En su obra sobre el conocimiento y el poder, el filósofo e historiador de las ideas Michel Foucault ha descrito cómo la Policía francesa y la Administración napoleónica usaron estos métodos para extender la vigilancia sobre individuos y grupos específicos a todos los ciudadanos (Foucault, 1980).

Tras esta primera explosión en el siglo XIX de recogida de de recogida de (*big*) *data*, los desarrollos actuales pueden considerarse la segunda revolución de los datos (Robertson y Travaglia, 2015), esta vez propulsada por las tecnologías digitales. En los años 1990, con el auge de internet, emergió el *e-gobierno*: los servicios públicos se prestan de forma digital (*online*) y se usan las tecnologías digitales entre los distintos departamentos. Más tarde, mediante los *smartphones*, se recogieron y compartieron cantidades masivas de datos de los ciudadanos. A menudo, terceras partes, principalmente del sector privado, se beneficiaron de ello. Los gobiernos también comenzaron a utilizar las redes sociales para informar y comunicarse con los ciudadanos en un intento de hacer que el sector público fuera más transparente. Al mismo tiempo, esta digitalización y gobierno basado en los datos hizo que estos últimos fueran más vulnerables ante los ciberataques. Pero también se incrementó la vulnerabilidad de los ciudadanos. Si el Gobierno tiene tus datos, te puede monitorizar y manipular. Y los gobiernos nacionales no son los únicos actores que ejercen poder sobre ti. Los Estados nación han estado perdiendo poder, han surgido poderes tecnocráticos nuevos, y se encuentran en el sector privado. ¿Cuán poderosos son aún los gobiernos y las naciones cuando las grandes empresas tecnológicas poseen muchos más datos y saben más acerca de sus ciudadanos de lo que pudieron soñar sus servicios secretos en el siglo XX? La tecnología, de nuevo, ha cambiado el orden social y se ha usado para obtener poder político.

Sin embargo, cuando se trata de contextualizar las preocupaciones en torno a la IA y la democracia, la cuestión más interesante no es simplemente que la tecnología y el poder hayan

sido siempre amigos, sino que las tecnologías han conducido sistemáticamente a formas *específicas* de orden y organización social y política: formas que implican un control centralizado y no democrático.

El relato habitual acerca de las relaciones entre la tecnología y la política, del que aún se hacen eco los libros de historia populares, como *Sapiens* (2015) de Yuval Harari, y que se retrotrae al menos hasta Marx, dice lo siguiente. Los primeros humanos vivían como cazadores-recolectores, y este tipo de vida y sus tecnologías relacionadas iban de la mano con una forma de organización social más o menos «democrática». Como la gente vivía en pequeños grupos, existía el trueque, el poder estaba descentralizado y era compartido por las personas; o, al menos, por los ancianos. No se creaban excedentes y los animales vivían en libertad, con lo que no tenía sentido intentar coger el grano o la vaca del vecino, pues, para empezar, prácticamente no había existencias. Quizás no hubiera una equidad o una ausencia de rivalidad perfectas (algunos cazadores o terrenos de caza eran mejores que otros, algunas tribus vivían mejor que otras), pero no había nada parecido a la autocracia: sin dictaduras, sin ni siquiera un Estado. La idea misma de un Estado solo se desarrolló una vez que hubo «propiedades» [*estates*], esto es, con la revolución agrícola. Una vez que se crean excedentes y se almacenan, alguien puede intentar robarlos o gravarlos con impuestos. El mayor peligro, entonces, dejó de estar en los animales salvajes y pasó a ser el del resto de seres humanos. Algunos afirmaron ofrecer seguridad, pero otros exigieron obediencia y bienes a cambio. Algunos ejercieron el poder sobre los demás dando lugar a menudo a la esclavitud, y a otras relaciones sociales y económicas desiguales y explotadoras.

Mientras tanto, la revolución agrícola hizo que la población aumentara, por lo que se desarrollaron comunidades políticas más grandes. Se centralizó el poder político. Se establecieron jerarquías. Las pirámides del antiguo Egipto no son solo edificios funcionales:

también constituyen metáforas sobre cómo se distribuía el poder en la sociedad. En las monarquías teocráticas del antiguo Egipto el poder estaba claramente centralizado y jerarquizado. La clase alta de miembros de la familia real, terratenientes, funcionarios del Gobierno y altos sacerdotes gobernaban sobre una clase media de mercaderes y artesanos quienes, a su vez, se encontraban por encima de una gran clase baja de trabajadores sin cualificar. Ya en tiempos antiguos esto creó tensiones entre las clases. Se usó la violencia para lidiar con ello. Y como algunos Estados se convirtieron en imperios, el hambre de más tierra y poder hizo emerger más conflictos. Existía una opresión dentro del Estado y guerras sangrientas entre Estados e imperios.

Esto no ha cambiado demasiado a lo largo la historia de la humanidad. Las ciudades-Estado de la antigua Grecia y los sistemas feudales de la Europa medieval fueron todavía ejemplos de sociedades centralizadas y jerárquicas. Se usaba el dinero y se desarrolló el comercio, lo cual, en principio, es bueno para la paz y la prosperidad: el dinero es un medio muy importante de tecnología y comunicación social, a menudo pasado por alto. Pero no igualó las relaciones sociales ni pudo evitar las muchas guerras que tuvieron lugar por disputas por la tierra. Las sociedades agrícolas necesitan territorio. Y los imperios comerciales necesitan tierras a fin de recabar los recursos necesarios para el comercio y para su expansión. Esto condujo a la colonización. Surgieron enormes imperios coloniales: los de Portugal y España en el siglo xv, más tarde también los de Francia, Holanda, Reino Unido, y así sucesivamente. El capitalismo evolucionó, por ejemplo, en el contexto de la acumulación de capital empresarial y la colonización de Indonesia por la Compañía Neerlandesa de las Indias Orientales (VOC). En tiempos modernos, las tecnologías de la Revolución Industrial permitieron una nueva forma de organización social: el capitalismo industrial moderno. Esto ayudó a abrir las sociedades tradicionales y facilitó una orientación hacia el futuro (Beckert,

2016); se destruyeron las estructuras feudales antiguas o, al menos, se debilitaron de forma significativa, pero también continuó habiendo, si bien no radicalizadas, relaciones sociales y políticas desiguales. Se utilizaron máquinas para explotar a los trabajadores y, más tarde, para reemplazarlos. De nuevo, hubo luchas y guerras, pero esta vez eran por los medios de producción *industrial* (por usar el lenguaje marxista) y los recursos naturales necesarios para ella. La evolución de las tecnologías, de nuevo, condujo a relaciones sociales y políticas desiguales. En particular, condujo a la acumulación y centralización del control por parte de los Estados y corporaciones, y lo que, en su *Manifiesto del Partido Comunista* de 1848, Marx y Engels denominaron «lucha de clases»: en el sistema capitalista, el proletariado oprimido y explotado lucha contra los propietarios de los medios de producción. Gradualmente, se instalaron las democracias, pero su poder era y está limitado dentro del contexto de estas fuerzas sociales y económicas abrumadoras, propulsadas por la historia de la tecnología.

¿Es inevitable esta historia de centralización, represión y explotación mediante la tecnología? ¿Y es cierta? La realidad es más compleja. En principio, es posible vivir en comunidades agrícolas que sean más igualitarias. En teoría, es posible organizar un Estado nación industrial de forma que otorgue más control a las unidades descentralizadas y a la mayoría (incluso si esta última no está exenta de problemas; señalaré más tarde que puede volverse tiránica). Esta descentralización tuvo lugar hasta cierto punto y en algunos lugares: actualmente, muchas personas en Occidente y en otras zonas viven en Estados nación representativos, aunque aún centralizados en gran medida, con *parte* del poder delegado a niveles inferiores. Y, quizás, la idea de sociedades agrícolas centralizadas no sea simplemente teoría. El antropólogo David Graeber y el arqueólogo David Wengrow han defendido en *The Dawn of Everything* (2021) [*El amanecer de todo*] que la agricultura y las ciudades no siempre

condujeron a jerarquías y a la dominación. Ellos afirman que las ciudades pueden existir sin un poder central.

Sus interpretaciones de la evidencia arqueológica arrojan cierta luz crítica acerca de historias como las ofrecidas por Harari y abren posibilidades para imaginar diferentes formas de organización social. En otras palabras, la relación entre la tecnología y las formas de organización social y política no es determinista. Hay excepciones. La tecnología no conduce *necesariamente* al poder centralizado, a la opresión y a la violencia. Esto ofrece una esperanza para la emancipación y el cambio social.

Sin embargo, desafortunadamente la versión dominante de la historia acerca de la tecnología, la civilización y el poder centralizado se suele ratificar a menudo. Considérese la historia más reciente de las tecnologías digitales actuales y su relación con la sociedad. Al principio, el futuro parecía prometedor. Las nuevas tecnologías prometían un futuro más democrático. En los noventa, cuando internet se empezó a utilizar ampliamente, la gente soñaba con sociedades descentralizadas en red, modeladas según la naturaleza en red descentralizada de las nuevas tecnologías, en particular, de internet. Esto suena bien en teoría. La idea sigue siendo algo popular en California, en donde las actitudes positivas anteriores hacia la tecnología coincidieron con la contracultura de los años 1960 y 1970, y contribuyeron a la historia de éxito del actual Silicon Valley. También continúa viva en los, por ejemplo, *Pirate Parties* (partidos piratas), que promueven formas más directas y participativas de *e-democracia*^[2].

Sin embargo, la esperanza de que las tecnologías digitales conduzcan a una mayor democracia ha continuado siendo, en gran parte, una forma de utopismo (a veces llamada ciberutopismo o web-utopismo). Lo que presenciamos en la realidad político-tecnológica diaria no es más democracia, sino una concentración de poder en las grandes empresas tecnológicas que deciden cómo vivimos nuestras vidas, además de en regímenes

autocráticos que usan las tecnologías digitales para mantener y aumentar su poder. Parece que más allá de las herramientas de los cazadores y recolectores, las tecnologías agrícolas, industriales y digitales más avanzadas respaldan tendencias no democráticas. Quienes tenían que remar en el barco del Estado de Platón eran esclavos. Actualmente, la mayor parte de nosotros rema en los barcos de las grandes tecnológicas y de los gobiernos autocráticos dirigidos por las tecnologías digitales y los humanos que las usan y las controlan. Queremos dirigir. Queremos dirigir nuestras propias vidas y ser los amos de las nuevas tecnologías. Pero lo que ocurre es lo contrario. No dirigimos, sino que nos dirigen. En el mejor de los casos, estamos en el asiento de los pasajeros mientras las decisiones las toman los Googles, OpenAIs, Facebooks y Twitters de este mundo. En el peor de los casos, somos los remeros: somos los motores y el combustible de la nueva economía digital, que crea los datos necesarios y se los entrega a quienes los usan para enriquecerse a nuestra costa.

Estos problemas relacionados con la tecnología digital se suman a los desequilibrios de poder existentes y a los problemas dentro de las democracias liberales. No se deben *simplemente* a la tecnología. No se puede culpar a la IA de todo, como acertadamente sugieren König y Wenzelburger (2020). Pero cuando nuestras democracias ya se encuentran debilitadas, los efectos negativos de las tecnologías digitales como la IA golpearán con más fuerza. Las democracias débiles son menos resilientes ante lo que la IA les puede hacer. Defenderé que esta amplifica los problemas y, de esta forma, apoya tendencias antidemocráticas que potencialmente podrían conducir al autoritarismo.

A pesar de que esto nunca es solo una cuestión tecnológica, podemos aprender de estas historias (y de historias de futuros utópicos) que existen vínculos claros entre la tecnología y las amenazas a la democracia. Esto también es cierto en el caso de la IA que, en forma de cibernética, tiene vínculos con ideas no

democráticas, y que actualmente contribuye a las mencionadas historias de centralización y dominación. Este hecho no es inevitable, pero es importante reconocer los problemas del pasado, los riesgos futuros y, ciertamente, lo que ya está ocurriendo. Estas historias muestran que la tecnología tiene consecuencias sociales y políticas y que es, al menos, muy *difícil* evitarlas. Desde una perspectiva marxista, se podría argumentar, por ejemplo, que, a medida que la IA acelera la sustitución del capital por trabajo, los que poseen capital se benefician de la nueva revolución tecnológica, mientras que otros se quedan atrás (Kaplan, 2016). De nuevo, es posible en teoría crear una forma de organización social que sea diferente. La influencia de la tecnología en la historia y la sociedad no es determinista. Las normas específicas de organización social no son ni necesarias ni inevitables. Dadas las fuerzas y desarrollos descritos, el cambio social es difícil pero no imposible. Si realmente queremos, parece que podemos crear una sociedad diferente. Por ejemplo, podemos rechazar el modelo platónico e intentar crear sociedades que sean más parecidas a las comunidades antiguas, más igualitarias (¿protoanarquistas?) que Graeber y Wengrow nos piden que imaginemos. Son posibles distintos tipos de comunidad política, distintas ciudades, distintos tipos de *polis*.

Sin embargo, los llamamientos en favor de más democracia y cambio político suenan a menudo como si el cambio social fuera solo cuestión de voluntad y acciones de los individuos. Consideremos el «*yes we can*» de Obama o el activismo climático contemporáneo, que apela a la (buena) voluntad de políticos individuales. Este énfasis en los individuos y en los humanos es engañoso. En su lugar, argumentaré en los últimos capítulos de este libro que, dado el importante papel de las tecnologías en la historia, la clave del cambio social no es solo el voluntarismo y el activismo humano e individual (a pesar de que es necesario algún tipo de liderazgo), sino también el cambio de las instituciones políticas y, especialmente, el cambio de las *tecnologías*. Si realmente queremos democracia, será

mejor que creemos tecnologías más democráticas. En particular, defenderé que necesitamos una «IA democrática» e, incluso, una «IA *para* la democracia».

Hablar de «IA democrática» suena extraño, ya que mezcla dos reinos diferentes que se supone que no se solapan: la política y la tecnología. Pero, como ya he mencionado antes, tal cambio conceptual (y ciertamente social) tiene mucho sentido una vez que comprendemos que las *tecnologías mismas son políticas*. A lo largo de este libro, la hipótesis es que las tecnologías no son meramente instrumentos neutrales para alcanzar objetivos humanos, sino también intentos de cambiar estos objetivos y los significados humanos ligados a ellos. Las historias mencionadas aquí apoyan ese punto: la tecnología ha ayudado a cambiar la perspectiva de civilizaciones enteras.

Reflexionar seriamente en torno a la IA implica aceptar que esta no es políticamente neutral y que tampoco es neutral con respecto a la democracia. Desde este punto de vista tiene sentido discutir y tomarse en serio la afirmación de que la IA actual socava la democracia, en lugar de decir simplemente que algunas *personas la usan* para socavar la democracia (que también es cierto, pero trivial). Al mismo tiempo, debe reconocerse que los humanos están siempre implicados y tienen cierta libertad para dirigir lo que está ocurriendo. Por lo tanto, la IA no conlleva *necesariamente* estos efectos y puede, incluso, contribuir a incrementar la resiliencia de la democracia. Quizás un nuevo tipo de IA pueda ayudar al cambio social y a apoyar la democracia. Sobre esta cuestión tendremos que decir más cosas, pero lo que ya no tiene sentido es continuar manteniendo el consabido «la IA puede usarse para lo bueno y para lo malo, simplemente depende de cómo se utilice». Tras décadas de filosofía de la tecnología y estudios sociales de la ciencia y de la tecnología, y tras considerar la aún más larga historia de la tecnología y de la política, pero también tras haber experimentado cómo internet y otras tecnologías digitales han transformado en

tiempos recientes rápida y profundamente las sociedades en las que vivimos, ya es hora de decir adiós a una concepción instrumental tan estrecha de la tecnología y de su relación con la política.

Capítulo 3

¿Qué IA? ¿Qué democracia?

Entonces, si la inteligencia artificial es política, ¿en qué dirección nos conduce la IA actual? Dados los desarrollos contemporáneos y la historia de la tecnología y de la democracia, la pregunta en relación con la IA y la democracia se plantea a menudo de una forma que subraya las amenazas: la cuestión es cómo de mala es la tecnología para la democracia. Este es también mi punto de partida en este libro. Es el punto de partida de mi análisis de los riesgos que la IA supone para la democracia y, ciertamente, de cómo ya la está socavando. Si bien no hay duda de que la IA también puede conllevar efectos positivos para la democracia. Por ejemplo, puede ayudar a los votantes a tomar decisiones al informarlos mejor; a detectar el fraude y la corrupción; y se puede utilizar contra el sesgo. Más adelante, añadiré algunas consideraciones acerca de cómo se puede utilizar la IA para respaldar a la democracia. Pero hoy día está ampliamente reconocido que existen problemas serios. Fenómenos como la manipulación de las elecciones, la desinformación, los filtros burbuja, los gobiernos que utilizan la IA para socavar el imperio de la ley, los sesgos contra algunas categorías de ciudadanos, y las asimetrías de poder entre las grandes tecnológicas y los ciudadanos ordinarios, todo ello ha dado

alas a un discurso acerca de la IA que la problematiza en relación con la democracia (Sudmann, 2019). Por ejemplo, señalando las nuevas posibilidades de manipulación («programar a la gente») y de hacer *nudging* político mediante el uso de *big data*, Helbing y sus colegas (2017) se preguntaron en su artículo de *Scientific American*, «¿sobrevivirá la democracia al *big data* y a la inteligencia artificial?». Los expertos están preocupados. Los ciudadanos están preocupados. Los políticos están preocupados. ¿Está en peligro la democracia y, si es así, por qué exactamente y qué se puede hacer al respecto?

La respuesta a la pregunta que hago al comienzo (¿socava la IA la democracia?) será «sí, lo hace, y, si no logramos abordar satisfactoriamente este problema, nuestras democracias no sobrevivirán en su forma actual, pero sí podemos hacer algo al respecto». Pero antes de responder, necesitamos en primer lugar analizar la pregunta: ¿Qué queremos decir por democracia? ¿Y por IA?

INTELIGENCIA ARTIFICIAL (IA)

Para comenzar, la IA puede significar muchas cosas. Técnicamente, puede referirse a un algoritmo basado en el razonamiento y la experiencia humanos que hace recomendaciones con un árbol de decisión, aplicando una fórmula del tipo «sí... entonces...» al input del usuario. Este conocimiento se introduce en el sistema por humanos: expertos humanos en un dominio particular. Considérense los sistemas expertos médicos: se introducen los síntomas en el sistema y, basándose en el conocimiento que expertos humanos introducen previamente en aquel, el ordenador puede dar una recomendación de diagnosis razonada. Esta es la denominada «IA a la antigua usanza» (*Good Old-Fashioned AI*, GOFAI). Está basada en reglas, y las reglas las introducen expertos humanos. La ventaja de este enfoque es que

permite crear sistemas muy transparentes: cualquiera con acceso a las reglas comprende lo que está haciendo el sistema de inteligencia artificial. Fue especialmente importante en los años 1980. Sin embargo, este tipo de IA recibió duras críticas por parte del filósofo Hubert Dreyfus (1972), ya que no lograba captar la inteligencia y la experiencia humanas. Defendía, entre otras cosas, que la mente humana no es un ordenador y que no se puede formalizar todo el conocimiento. Mucho antes, el pionero de la computación Alan Turing (1950) ya había señalado las limitaciones de los sistemas basados en reglas.

Sin embargo, la crítica de Dreyfus fue relevante desde el comienzo de la inteligencia artificial (principalmente en el congreso de Darmouth de 1956, donde investigadores como Minsky, McCarthy, Shannon y Simon discutieron el futuro de la computación)[1], cuando la IA estaba ligada al proyecto de simular la inteligencia humana. Hoy en día aún hay investigadores que trabajan en la denominada «IA fuerte» o «inteligencia artificial general» (IAG), que persigue replicar las funciones humanas. Algunos creen que la IA podría tener sensibilidad o consciencia en el futuro. A menudo, el debate sobre la IA sigue estando relacionado con (¿o secuestrado por?) este tipo de proyectos. Por ejemplo, tras el éxito de modelos de lenguaje como ChatGPT, la discusión ha surgido de nuevo. La IA se está volviendo cada vez mejor a la hora de imitar la inteligencia humana, incluso si no tiene LAG y carece de inteligencia por sí misma (lo cual levanta dudas acerca de la idoneidad de la misma denominación «inteligencia artificial» y la de plantear los desafíos de Darmouth).

Sin embargo, lo que ya existe, es que la IA se desempeña muy bien cuando se trata de tareas específicas (la llamada IA «débil» o «estrecha»). Se están generalizando las aplicaciones de este tipo. La IA se usa en el ajedrez y en otros juegos como el go y el *bridge*, donde ha alcanzado logros espectaculares[2], pero también, por ejemplo, en chatbots, motores de búsqueda, generadores de texto, vehículos autónomos, compras personalizadas en línea,

recomendaciones de obras de música y vídeo, filtros de *spam*, detección de enfermedades y bots en juegos de ordenador. Son estos buenos tiempos para la IA. Esta es parte de nuestra vida cotidiana, incluso sin que nos demos cuenta.

Una IA de esta especie ya no está basada (principalmente) en reglas. Actualmente, la expresión IA se refiere normalmente a aprendizaje automático (*machine learning*): basado en la estadística, el sistema tiene capacidad de discernir patrones entre los (normalmente muchos) datos. Esto se puede utilizar a continuación para clasificar y recomendar. Los anteriormente mencionados motores de búsqueda, los chatbots, los generadores de textos, y así sucesivamente, son sistemas de aprendizaje automático, algunos hasta un grado significativo. Consideremos también el sector médico: a partir del análisis de miles de imágenes de pulmones, la IA puede «aprender» a decir si existe una alta probabilidad de que una persona tenga cáncer de pulmón. A esto se lo denomina reconocimiento de imagen. El sistema aprende qué (patrones) se correlacionan en los datos visuales con el cáncer de pulmón. En este sentido, es posible utilizar la misma tecnología para la seguridad y la vigilancia, para recomendar vídeos o para reconocer imágenes de gatos. En la esfera política, la IA se puede alimentar con millones de datos de texto extraídos de sitios de las redes sociales y crear perfiles políticos de las personas. Puede decir con una alta probabilidad que, por ejemplo, eres políticamente de izquierdas. Esta información puede utilizarse en las campañas políticas e, incluso, para manipular tu voto. Imaginemos que el algoritmo descubre que eres lo que se conoce como un «votante indeciso» en un sistema político dominado por dos partidos: unas veces votas al partido A, otras al partido B. Esto es interesante para los políticos, pues estas son las personas que marcan la diferencia en unas elecciones y, por lo tanto, merece la pena gastar dinero en intentar influir en su comportamiento como votantes, por ejemplo, dirigiéndose a ellos con una propaganda política específica en la misma red social.

Pero la «IA» no es una sola cosa (si es que siquiera es una cosa): no es simplemente el algoritmo, sino también los datos. No es simplemente software, sino que también está relacionada con el hardware y con los recursos necesarios para producirlo. Tiene que ver con la tecnología, pero también con la infraestructura, la economía y la ecología. Debido a esta naturaleza relacional, es importante para la ética y la política de la IA considerar este panorama más amplio. Por ejemplo, el uso de la IA y los datos tiene consecuencias medioambientales y ecológicas importantes. Como ha defendido Kate Crawford en *Atlas of AI* (2021), la IA conlleva muchos costes ocultos que causan impacto en el planeta a través del uso de energía y de los recursos naturales necesarios. La IA también contribuye a las emisiones de carbono. Como todas las tecnologías digitales, parece ser limpia y virtual, opuesta a las tecnologías de la Revolución Industrial. Pero esto es engañoso. La IA puede impulsar la economía, pero ¿a qué precio medioambiental? ¿O puede ayudarnos la IA a mitigar el cambio climático?

Además, la IA es más que un sistema técnico y no es simplemente una cuestión de innovación y economía. Tiene que ver también con los humanos y con su cultura. La IA es asimismo las narrativas que contamos sobre ella, la gente y las organizaciones que la usan, y los significados que le atribuimos. La IA no solo tiene que ver con datos y ordenadores, sino también con nuestras esperanzas y sueños colectivos. Es un sinónimo de las mayores aspiraciones y miedos de la humanidad. Tiene que ver con nosotros.

Por ejemplo, la idea ya mencionada de construir una inteligencia similar a la humana es un sueño para algunos, una pesadilla para otros. De acuerdo con algunos, podría funcionar en ambos sentidos, dependiendo de cómo la desarrollemos y gestionemos. Nick Bostrom ha defendido en su libro *Superintelligence* (2014) que la IA a nivel humano llegará algún día y que, cuando esto ocurra, pronto aparecerán las máquinas superinteligentes que sobrepasarán a los cerebros humanos en capacidad, serán difíciles de controlar y,

probablemente, reemplacen a los humanos. En definitiva, la IA guarda relación con la historia y el futuro de la humanidad. Los transhumanistas como Bostrom simpatizan con la idea de que el fin de la humanidad tal como la conocemos se halla cerca, mientras que los críticos nos advierten ante lo que perciben como un distópico «mundo feliz». Otros hablan de la IA de una forma más concreta, personal y social, aunque no menos relevante desde el punto de vista existencial: tienen miedo de que pierdan sus trabajos cuando la IA asuma el mando. Otros la ven principalmente como una oportunidad de negocio. O esperan que la IA mejore... la democracia. Cuando las personas discuten la ética y la política de la IA, esto ocurre porque ella toca temas y desafíos que son importantes para nosotros actualmente. La IA no afecta únicamente a las tecnologías, sino también a nosotros como personas, como sociedades y como humanidad. Tiene que ver con nuestros valores y metas, con nuestro futuro y con qué futuro queremos.

DEMOCRACIA.

Democracia también puede significar muchas cosas. Significa literalmente el gobierno (*kratos*) del pueblo (*demos*), pero existen muchas ideas acerca de lo que significa. En su forma más «débil», la democracia es, simplemente, un procedimiento de agregación de votos, en particular, de la votación de la mayoría: en lugar de tener una persona que decide, se pide al pueblo que vote, y la mayoría de los votos gana. La democracia es, pues, una forma de gobierno en el que el procedimiento del voto de la mayoría se usa para tomar decisiones políticas importantes. Sin embargo, hoy en día esto no tiene lugar de forma directa. A la mayoría de nosotros no se nos pregunta respecto de asuntos políticos concretos. En su lugar, en la mayoría de las democracias liberales existentes, los ciudadanos son representados por políticos, quienes votan en el Parlamento en

nuestro nombre: esta es la democracia representativa. A los ciudadanos se les pide votar por los políticos una vez cada ciertos meses o años. Pero, en principio, también existen formas de democracia directa, en la que los ciudadanos votan ellos mismos con respecto a decisiones particulares, sin intermediarios. Este procedimiento se usa a veces a una escala menor en ciudades-Estado como en la antigua Atenas (a pesar de que muchas personas estaban excluidas de la votación), o en pequeñas naciones y regiones, como en Suiza. Jean-Jacques Rousseau, uno de los principales teóricos de la democracia, desarrolló en el siglo XVIII sus ideas en el contexto de la (entonces) ciudad-Estado de Ginebra. Influenciado por el republicanismo según se desarrolló en los tiempos antiguos y en el Renacimiento, defendió que el poder debería estar en manos del pueblo, preferentemente en la forma de una democracia directa. Aquí, la democracia no se concibe como una votación cuyo resultado es una suma de opiniones individuales, sino como *res publica*, relativa a los asuntos públicos y al bien común.

Esto nos conduce a concepciones de la democracia «más fuertes», en las que los gobiernos piden más a los ciudadanos que el mero voto. Esto es especialmente cierto en el caso de aquellas teorías de la democracia que surgieron en la tradición del republicanismo y del pensamiento ilustrado influenciado por esa tradición. El republicanismo clásico, desarrollado primero en la antigua Grecia y en Roma, presentaba una idea de la política y del gobierno que primaba el autogobierno y la participación de los ciudadanos (de forma opuesta al gobierno autoritario de los reyes y de los emperadores, por ejemplo). Pero también introdujo otras ideas que «engrasan» y enriquecen esta concepción de la democracia, como la virtud cívica, la educación y el bien común. En el republicanismo filosófico desde Aristóteles hasta Rousseau, a los ciudadanos no solo se les otorga el derecho de voto, sino que también se espera que sean buenos ciudadanos y que se preocupen y contribuyan al bien común.

Las teorías de la democracia deliberativa y participativa de hoy día subrayan menos la ética y la virtud, pero mantienen la idea de que los ciudadanos desempeñan un papel político activo (participación y deliberación en lugar de simplemente votar), con alguna orientación formativa hacia el bien común. Curiosamente, estas teorías permiten, pero también *necesitan*, que la gente promueva y reflexione acerca del bien común y participe en el proceso de toma de decisiones, no simplemente votando, sino también deliberando y discutiendo cómo debería dirigirse el Estado (Cohen, 1986; Estlund, 2009; Habermas, 1990; Landemore, 2020). Así, de acuerdo con estos puntos de vista, la democracia es más que un algoritmo de agregación de votos, y no digamos ya una simple expresión de opinión en las redes sociales. Más bien, requiere de la participación de los ciudadanos. Presupone un concepto de estos como sujetos autónomos e ilustrados que razonan y discuten lo que es mejor para la sociedad, que están dispuestos a hacer que sus puntos de vista se sometan a discusión con todas sus vulnerabilidades, y que se abren a revisarlos si resulta necesario. Inspirados por la tradición republicana, también podríamos tomar prestado el lenguaje de las virtudes y deberes cívicos, y hablar sobre el deber de participar en la sociedad y en la política. Siguiendo a Aristóteles, esto también guarda relación con la virtud, en particular, con la virtud de la justicia. De acuerdo con los pensadores de esta tradición, el compromiso activo de los ciudadanos no solo evita el abuso del poder, sino que también nos ayuda a desarrollarnos como personas virtuosas y como seres políticos. Esta no es solo una cuestión de naturaleza (somos animales políticos, para usar el término de Aristóteles), sino que además requiere cuidado y cariño: los ciudadanos necesitan desarrollarse, necesitan ser educados. Es necesaria la educación cívica.

Un ángulo interesante e importante desde el que considerar estas teorías de la democracia es usar el concepto de comunicación, que también se puede conectar con el bien común. De acuerdo con

Fuchs, la democracia está «organizada como ciertos procesos de comunicación en donde los humanos se informan, debaten y toman decisiones vinculantes de forma colectiva» (2023: 12). Esto necesita una esfera pública: una esfera social en donde la gente pueda discutir libremente problemas políticos y, así, formar un público y una opinión pública. Además, en la tradición filosófica republicana concebida de forma amplia, por ejemplo, en Rousseau, Dewey y Arendt, pero también ya antes en Cicerón y en San Agustín, existe a su vez un significado de comunicación más rico fomentado por la necesidad de que los ciudadanos discutan problemas comunes, coloquen sus intereses privados a un lado y se preocupen del bien común. En Rousseau (1997 [1762]), esto toma la forma de sometimiento a la voluntad general. Es esta una idea controvertida, ya que puede conducir a alguna forma de autoritarismo: una mayoría o un dictador que afirme encarnar la voluntad general podría imponer su voluntad sobre la de los demás y reprimir sus voces. Pero, dejando este concepto específico a un lado, la tradición republicana defiende una idea de democracia que se centra en la comunicación en el sentido de tener y crear algo en común. Dewey y Habermas subrayaron no solo la participación, sino también la comunicación. En *Democracy and Education* (1944), Dewey considera la comunicación un proceso en el cual los participantes aprenden a ver las perspectivas de los otros y a cooperar en problemas compartidos, lo que permite tanto el crecimiento personal como social. Se podría decir que conduce a una mayor comunidad, entendida como reunión. Para Habermas, la «acción comunicativa» (1984) trata de perseguir fines basados en una comprensión compartida. Arendt (1998 [1958]) defendía que la política necesita un espacio público compartido y la creación y mantenimiento de un mundo común. Así pues, estos pensadores republicanos defienden un enfoque más relacional de la democracia que conecta a la persona con el todo político; un enfoque en el que la democracia no es solo entendida de forma instrumental, sino que tiene valor

intrínseco: aprendemos a desarrollarnos y nos desarrollamos nosotros mismos como una comunidad.

Al final del libro profundizaré en estos conceptos que constituyen juntos una versión de la democracia rica y fuerte: defenderé que esta concepción de la democracia puede y debe guiar el desarrollo de la IA democrática y de la IA para la democracia. Por ahora, permítaseme primero hacer una advertencia sobre —y tomar distancia respecto de— una versión platónica de esta perspectiva.

Platón subrayaba que la política necesita conocimiento. La versión rica de la democracia que acabo de bosquejar está de acuerdo con esto en el sentido de que reconoce que esta forma de gobernar supone conocimiento por parte de los ciudadanos, que necesitan disponer del mismo para poder participar en el gobierno. Esto implica no solo conocimiento de los partidos políticos y de sus programas electorales, como sucede en todos los sistemas de representación que tenemos actualmente, sino un conocimiento y una educación políticos bastante más amplios. Todo el conocimiento y la capacidad necesaria para ser un *kybernetes*. Pero, en contraste con la cibernética política autoritaria de Platón, la versión democrática mantiene que *todos nosotros* deberíamos dirigir. Todos los ciudadanos deberían participar de las decisiones políticas. Y, al final, la democracia y la política no afectan en este sentido solo a la gobernanza, sino también a la comunicación y al trabajo en pro del bien común. Eso requiere conocimiento y comprensión, algo que todos nosotros deberíamos ejercitar y desarrollar antes, en y durante el proceso democrático.

Pero ¿es posible y deseable que todos nosotros dirijamos y adquiramos este conocimiento? ¿Cuán inclusiva es y debe ser la democracia? En la época de Platón, cuando aparecieron estas teorías republicanas de la democracia, solo gobernaban los ciudadanos, ya que se podían liberar del trabajo manual. Pero eso excluía al resto de la población: la clase de los ciudadanos incluía solo a una pequeña parte de la población ateniense. El análisis político de

Arendt en *La condición humana* (2005 [1958]), que celebra la acción política en cuanto que, opuesta al trabajo manual, está inspirada por este modelo antiguo y jerárquico de actividades basado a su vez en un modelo jerárquico de la sociedad: los ciudadanos varones de mayor edad hablan de política, mientras que el trabajo y las tareas del hogar los realizan los otros que están excluidos de esa forma de «democracia»: mujeres, esclavos y artesanos. En nuestros días se ha abolido la esclavitud (al menos, oficialmente), y la mayoría de nosotros tendemos a vivir en sociedades menos jerarquizadas con niveles de dominación más bajos. A su favor, Arendt rechazó el modelo jerárquico. Pero la cuestión continúa siendo relevante. ¿La democracia directa es posible cuando la mayoría de nosotros nos vemos sometidos a trabajos estresantes, actividades familiares extenuantes y un sistema socioeconómico exigente que deja poco tiempo y energía para las actividades políticas? Además, el conocimiento experto actual parece ser necesario para comprender lo que está pasando en sociedades del conocimiento complejas, con la ciencia y la tecnología desempeñando un papel importante. Esto crea un desafío para la democracia, al menos, en su forma directa. ¿Podemos y debemos todos nosotros tener este conocimiento?

La respuesta republicana tradicional es: sí. Pero el mundo ha cambiado. A pesar de que la educación propuesta por Dewey y otros es parte de la respuesta (volveré sobre esta cuestión más adelante), es improbable que pueda resolver el problema por completo. Gobernar sociedades modernas no es como gobernar una ciudad-Estado antigua: requiere un cierto nivel de pericia en un amplio rango de dominios. Y este es un problema para los ideales de democracia deliberativos, participativos y republicanos. Como ha defendido Zarkadakis (2020), los modelos deliberativos de autogobierno se enfrentan al problema de las «asimetrías de conocimiento»; existen brechas en el conocimiento entre expertos y no expertos. El plantea el problema como sigue: «los problemas complejos necesitan de conocimiento experto para poder resolverse,

pero en una democracia también necesitan la aprobación y el consentimiento del electorado» (2020: xxiv).

La teoría política contemporánea después del denominado giro epistémico (Landemore, 2017) y, ciertamente, todos los sistemas políticos contemporáneos que dependen cada vez más de la pericia se enfrentan, por lo tanto, al desafío de reconocer que la cibernética política moderna requiere de conocimiento y, al mismo tiempo, de asegurarse de que el sistema es y sigue siendo democrático. Una forma de expresar esto es encuadrar la cuestión en términos de verdad. Si los sistemas se orientan demasiado hacia el lado del conocimiento y de la pericia, se enfrentan a la objeción de que la verdad puede ser tiránica: puede conducir a la tecnocracia o, al menos, a la tecnocracia en su forma autoritaria. A primera vista, suena bien decir que el gobierno debería basarse en la verdad. Pero ¿podemos estar seguros de que una perspectiva particular es la verdad? ¿Quién lo decide? Quizás no exista una única verdad. Y existe el peligro de que un grupo de gente (el que afirma que mantienen la verdad) tiranice y oprima a otros en nombre de la verdad. En su ensayo sobre la verdad y la política, Arendt (1968) defiende que, considerada desde una perspectiva política, la verdad es despótica y dominadora, ya que no tiene en cuenta las opiniones de otras personas. Pertenece a un modo de pensar que impide el debate político y no necesita que imaginemos el punto de vista del otro. Si, por otra parte, los sistemas políticos se desplazan demasiado hacia el lado de la democracia (al menos, si la entendemos simplemente como el gobierno de la mayoría, lo que Arendt llama «contar narices»), la mayoría también se puede volver tiránica, y puede reinar la ignorancia. La afirmación de que no es necesario ni conocimiento ni pericia para gobernar también conduce al desastre. Platón nos recuerda lo problemático de que una sociedad esté dirigida por gente ignorante. Actualmente, vemos de nuevo el auge de formas de populismo antiintelectualistas (y anticiencia): piénsese en el trumpismo en los Estados Unidos. ¿Cómo

navegamos entre estos polos, especialmente a la luz de la IA? ¿Y estamos siquiera preparados para confrontar el problema?

Aquí fracasa la analogía del barco o del avión. No es posible que todos dirijan el barco^[3]. Pero, ya para empezar, quizás gobernar no debería ser como dirigir un barco. Tal vez no debería ser una actividad cibernética en absoluto. Yendo más allá de Platón y de su autoritarismo, necesitamos metáforas mejores y, para ser honesto, aún no hemos encontrado una que resulte buena. Esto no ocurre solo debido a una falta de imaginación conceptual, sino también a que aún no vivimos en una democracia y carecemos de experiencia plena en un sistema de este tipo, si es que todos estamos en total acuerdo acerca de cómo sería una democracia plena y real. Parafraseando la respuesta de Mahatma Gandhi cuando se le preguntó qué pensaba de la civilización occidental, podemos responder de forma provisional la cuestión principal de este libro: «¿Democracia? Creo que sería una buena idea». Al preguntarnos por las relaciones entre IA y democracia, conviene no perder de vista que la democracia, especialmente en su versión más rica, continúa siendo un ideal que no se ha realizado por completo. Carecemos de la experiencia colectiva. Carecemos de ella porque en realidad no lo hemos intentado o, al menos, no lo suficiente. Esto se debe en parte al desacuerdo con respecto a lo que significa la democracia. Pero parte del problema consiste en que no sabemos lo que significa ni lo que debería significar en la actualidad. En realidad, no sabemos lo que queremos. Queremos democracia y queremos pericia, pero no sabemos qué significa la democracia y carecemos de una idea clara de cómo combinarla con la pericia.

Pero si queremos saber qué es la democracia, ¿qué tal si estudiamos esos sistemas existentes a los que llamamos «democráticos»? Esto es más problemático de lo que parece. Se pueden tomar en consideración las democracias occidentales, pero, desde una perspectiva global, no está nada claro que todas las instituciones políticas del mundo deban parecerse exactamente a lo

que tienen en mente los pensadores occidentales cuando hablan de democracia, a saber: los sistemas parcialmente democráticos que existen en Europa occidental y Norteamérica. Puede haber formas de democracia alternativas que sean igualmente buenas. ¿Y existen democracias «puras»? La situación actual es que en todas partes las fuerzas de la democracia y el autoritarismo están sujetas a continuas interacciones, negociaciones, fusiones y cambios. Existen fuerzas demo-créticas y antidemocráticas. Existen fusiones y acuerdos. Existen cambios. Actualmente, la dirección del cambio se dirige a menudo hacia el autoritarismo, incluso en Occidente. Nunes Da Costa (2022) usa el término «despotismos democráticos» para referirse a las tendencias hacia el despotismo en los contextos democráticos de hoy día.

Cuando se habla de autoritarismo, muchas personas en Occidente piensan en el pasado (la Alemania nazi, por ejemplo) o en las formas de autoritarismo en Rusia y China. Pero los problemas están más cerca de sus hogares, en donde han surgido tendencias antidemocráticas. El populismo de derechas y de extrema derecha ha progresado mucho, incluso en Europa. Nos vienen a la cabeza Trump en los Estados Unidos y Meloni en Italia. A pesar de que tales movimientos puede que no defiendan un autoritarismo en toda regla y usualmente afirman defender un *ethos* mayoritario (diciendo que harán lo que quiere la mayoría de la gente apelando, así, a una concepción débil de la democracia), tienden a hablar de boquilla de la democracia y bien pueden estar preparados para debilitarla una vez asuman el poder.

Sin embargo, la situación se vuelve aún más incómoda cuando nos damos cuenta de que el problema de los «despotismos democráticos» no solo tiene que ver con el autoritarismo: la democracia misma parece tener debilidades inherentes. En *La democracia en América* (2010 [1835]), Alexis de Tocqueville ya había avisado de la tiranía de la mayoría: sin controles, un grupo se puede volver dominante y oprimir a la minoría (1835: 306-307). Esta es

una razón por la que la democracia liberal necesita una base de derechos individuales para proteger a las personas frente a una mayoría opresora. Siempre existe también el peligro de que los intereses particulares de los individuos se antepongan al bien común y desactiven el debate en el espacio público: algo sobre lo que Arendt y otros filósofos de la tradición democrática republicana y deliberativa ya habían advertido con razón. Esto parece estar ocurriendo actualmente en el contexto de la IA, cuando los intereses personales dan forma a su futuro y cuando el debate público está plagado de tendencias no democráticas, incluyendo aquí la creación intencionada de polarización y los intentos de silenciar a los otros^[4] (me detendré mejor en estos efectos de la IA en los siguientes capítulos).

Así pues, la realidad (y, poco a poco, también la teoría) es más compleja que el binomio democracia versus autocracia (es importante tener en mente un contexto global con todas sus variedades de cultura política) y las democracias reales, existentes, tienen debilidades y desafíos inherentes. Sin embargo, y quizás también *debido* a los fenómenos y a la complejidad de los problemas mencionados, tiene sentido hacerse la pregunta presente con respecto a la democracia y la IA, con el término «democracia» entendido como un ideal que se realiza y ejecuta en diversos grados, y que puede tomar diferentes formas en Occidente y en otras partes. La cuestión de la democracia y de la IA es, pues, relevante para todas estas diferentes realizaciones parciales de la democracia, que se enfrentan (todas) a tendencias antidemocráticas (y, a veces, autoritarias) y a amenazas de diferentes grados y formas. Por ejemplo, tiene sentido preguntarse si, y hasta qué punto, los Estados Unidos de hoy en día son una democracia, y cómo su orden político incluidas sus tendencias antidemocráticas, está influenciado por la IA y otras tecnologías digitales.

Finalmente, está la cuestión de los límites del *demos* y de la *res publica*. Para empezar, ¿a quién se le permite acceder al barco? ¿Y, si

la metáfora no fuera democrática, quiénes son exactamente los que constituyen «el gobierno del pueblo»? ¿Se trata de ciudades, regiones, Estados nación, quizás de un colectivo global? ¿Pueden formar también comunidades políticas los no humanos? Por ejemplo, ¿pueden ser ciudadanos los animales, y deberían estar representados de alguna manera? ¿Cómo podemos reconocer conceptualmente el valor y significado político del medioambiente y del planeta? ¿Existe un buen argumento filosófico para restringir los límites de los territorios nacionales o de la humanidad? ¿Es sostenible la concepción antropocéntrica de la democracia^[5]? En este libro, limitaré la discusión a las democracias constituidas por humanos, pero, como ya he sugerido en otra obra sobre la política de la IA, merece la pena y es deseable explorar más a fondo esta área.

Para concluir, cuando se reflexiona acerca de la IA y la democracia, los debates sobre las relaciones entre esta última y la política que han mantenido ocupados a los filósofos políticos desde tiempos antiguos vuelven a resurgir, haciendo también aparecer a veces nuevas cuestiones, como las relativas al lugar de los conocimientos científicos en la sociedad o a los límites del *demos* a la luz de la crítica del antropocentrismo. Como sucede con la ética de la IA (o, al menos, con los enfoques filosóficos sobre este tema), que continúa alimentando discusiones en torno al significado y la naturaleza de la ética misma, hablar sobre política de la IA desde un punto de vista filosófico nos conduce, de forma inevitable, a antiguos debates filosóficos acerca de conceptos básicos usados en el razonamiento sobre la política, en particular, a los debates sobre la democracia y los conceptos filosófico-políticos relacionados. Ellos continuarán siendo el centro gravitatorio de este libro.

En los siguientes capítulos mostraré cómo y por qué la IA puede debilitar la democracia (con el término «democracia» entendido como un ideal y como algo solo realizado en la práctica de forma parcial). En primer lugar, defenderé que el problema no es

simplemente la manipulación de los votantes, sino que existen también otras formas, quizás incluso más importantes y profundas, en las que la IA corre el riesgo de erosionar los fundamentos mismos de nuestras democracias: los principios democráticos, como la libertad, la igualdad entre los ciudadanos, la fraternidad, la tolerancia y el imperio de la ley, pero también las bases del conocimiento y de la confianza que son necesarias para alcanzar una forma de democracia más rica, más estable y más resiliente, que pueda plantar cara a los desafíos del siglo XXI. A continuación, defenderé que podemos solucionar esto o, más bien, que podemos mitigar los riesgos fortaleciendo y transformando nuestras instituciones políticas y desarrollando una regulación sólida de la IA, pero también cambiando nuestras tecnologías, en este caso, la propia IA. Basándome en el reconocimiento de la no neutralidad profundamente política de la inteligencia artificial, propondré el desarrollo de una «IA democrática»: una forma de democratización tecnológica que intervenga de forma sistemática a nivel técnico. Lo que propongo es usar y *desarrollar* un nuevo tipo de IA que no solo proteja, sino que también fomente la democracia. Se trata de una IA *para* la democracia. Pero los cambios a nivel técnico no son suficientes: el desarrollo de la inteligencia artificial necesita estar integrado en una nueva cultura que sea completamente interdisciplinar y humanística. Ya que la IA no es simplemente algo técnico, sino también una narrativa y un sueño, propongo que también nosotros desarrollemos adicionalmente nuestra imaginación política y transformemos nuestra educación. Aquí, las artes y las humanidades, junto con la técnica y las ciencias sociales, tienen que desempeñar un papel clave. Hago un llamamiento a favor de un nuevo humanismo, un nuevo Renacimiento y una nueva Ilustración. El punto principal es que el futuro de la democracia y de la IA no debería dejarse en manos ni de expertos tecnológicos ni de políticos profesionales. Si nos tomamos la democracia seriamente, tanto lo que actualmente es como aquello en lo que se convierta la IA

debería estar determinado por todos nosotros. Pero esto requiere conocimiento, imaginación, educación y (como defenderé en el último capítulo) *comunicación* y una orientación hacia el *bien común*.

Capítulo 4

Cómo debilita la IA los principios básicos de la democracia

En marzo de 2018, un informante reveló a la prensa cómo una empresa de consultoría política recogió más de cincuenta millones de perfiles de votantes para predecir e influir en la votación en las elecciones de los Estados Unidos. Esa compañía, Cambridge Analytica, trabajó para el equipo electoral de Donald Trump y estaba dirigida entonces por Steve Bannon, el principal consejero de Trump. El escándalo no consistió solo en el hecho de que se recogieran tantos datos sin el consentimiento de la gente, sino también que estos datos fueran usados después para influir en el comportamiento del voto. En las propias palabras del informante: «Explotamos Facebook para recoger millones de perfiles de personas. Y construimos modelos para aprovecharnos de lo que sabíamos de ellos y apuntar a sus demonios internos» (Cadwalladr y Graham-Harrison, 2018). Se denominó a esto *microtargeting* con *big data* e IA, y fue posible apuntar a un gran número de votantes individuales basándose en sus perfiles políticos individuales. Se iniciaron algunas investigaciones, pero ya era demasiado tarde: Trump fue elegido en 2016. Más tarde, el Reino Unido votó

abandonar la Unión Europea (*brexít*). A pesar de que continúa siendo incierto hasta qué punto fue importante el papel de Cambridge Analytica, es plausible que hubiera una influencia significativa en estas elecciones.

Actualmente, el de Cambridge Analytica se considera un caso de estudio ejemplar sobre el uso político indebido de datos analíticos y de IA. Muestra cómo la IA puede influir en la política de forma directa, incluso a plena luz del día, en un sistema que se denomina a sí mismo democrático. La propaganda política siempre la han utilizado los políticos, tanto en democracias como en sistemas autoritarios, a través de, por ejemplo, publicidad en televisión y en los periódicos, pero existen diferencias significativas. Con las redes sociales y la IA, la escala y la velocidad son distintas a lo que lo eran antes: por ejemplo, Facebook tuvo 2 963 000 000 de usuarios activos en enero de 2023 (Kemp, 2023), y la comunicación es ahora personalizada. Al usar estas técnicas de tecnología punta, es posible orientar a los votantes en unas elecciones hacia una dirección particular. Como dijo un antiguo empleado de Cambridge Analytica, esto implica que nuestra democracia se halla bajo amenaza (Le Masurier, 2019). Normalmente, la manipulación de los votantes se considera antidemocrática. Sin embargo, ha ocurrido en lo que se consideran democracias como los Estados Unidos y el Reino Unido. Y puede ocurrir de nuevo: las plataformas de redes sociales como Facebook continúan recogiendo datos, manteniendo vivo el potencial para el tipo de análisis de datos y su manipulación por parte de empresas tales que Cambridge Analytica. Fuera de los focos de este caso particular, es probable que se continúe utilizando el *microtargeting* político en todo el espectro político.

Sin embargo, la democracia no consiste solo en votar. También existen formas más profundas en las que la tecnología de IA socava la democracia. «Más profundas», ya que tienen que ver con los fundamentos de la democracia. Los procedimientos de voto solo son uno de los elementos constitutivos de esta, pero existen más

elementos y fundamentos. Tenemos que indagar las condiciones para la democracia.

La mayoría de los pensadores políticos modernos sostienen que la democracia solo funciona si se basa en una serie de principios políticos fundamentales. Considérense los principios de libertad, igualdad y fraternidad, que eran tenidos por los principios básicos de la Revolución francesa de 1789. Los pensadores de la Ilustración y de las revoluciones americana y francesa pensaban que no podía haber democracia si el Estado oprime al pueblo, si el pueblo es tratado de forma radicalmente diferente de, por ejemplo, la nobleza, y si la sociedad está dividida y no conoce la justicia^[1]. Los principios de estas revoluciones continúan siendo claves para liberales y republicanos, y son las piedras angulares de la democracia moderna.

Revisaré primero los siguientes principios: la libertad, la igualdad, la fraternidad, el imperio de la ley y la tolerancia. A continuación, mostraré el impacto que la IA puede tener en estos principios.

PRINCIPIOS FUNDACIONALES DE IA DEMOCRACIA:

LIBERTAD, IGUALDAD, FRATERNIDAD, IMPERIO DE LA LEY Y TOLERANCIA

Un primer e importante principio es la *libertad*. Una de las figuras más prominentes de la Ilustración y de los teóricos de la democracia, Jean-Jacques Rousseau, defendió en *El contrato social* (1762) la entonces radical opinión de que todo el mundo nace libre. En contraste con Hobbes, Rousseau pensaba que el «estado de naturaleza» no era una situación horrible que necesitaba ser reemplazada por una autoridad, sino un estado bueno, que está corrompido por la sociedad. La democracia tiene que proteger y hacer tangible la libertad original del pueblo, en lugar de arrebatarla bajo la suposición de que, sin autoridad, el pueblo solo luchará y competirá entre sí. En lugar de someterse a un tirano,

defendía Rousseau, el pueblo debería gobernarse a sí mismo. Si acaso alguna sumisión es legítima, esa es la sumisión a la voluntad general del pueblo.

Esta idea ha sido muy influyente, desde la Revolución americana hasta la actualidad. Mientras reine la tiranía, no puede haber democracia. Las libertades personales de los ciudadanos se han tenido en cuenta en todas las cartas de derechos modernas (incluyendo los derechos humanos) y normalmente se mencionan en primer lugar. Por ejemplo, la Carta de Derechos de los Estados Unidos (las primeras diez enmiendas a la Constitución) comienza con las libertades personales. La libertad es una clave fundamental de cualquier democracia liberal. Además, la teoría contemporánea neorrepblicana (véase más abajo) subraya que la libertad no solo implica no interferencia y libertad de elección, sino también la no dominación: el no estar sujeto a un poder arbitrario, descontrolado (Pettitt, 1997; Skinner, 2008). Incluso si no se interfiere en el voto de las personas, puede que haya relaciones de dependencia en las que estas se hallen en *posición* de interferir en la votación o de obstaculizar su voto. Se puede defender que la democracia no encaja bien con tales relaciones: necesita asegurar el carácter no dominador de las relaciones y estructuras sociales.

Sin embargo, curiosamente, Rousseau también pensaba que la *desigualdad* era mala para la democracia. La igualdad de trato puede significar que todos tengan un voto y que cada voto cuente lo mismo. Pero en su *Discurso sobre el origen y los fundamentos de la desigualdad* (1755), Rousseau también defendía que la desigualdad económica radical es mala para la democracia, ya que conduce a que los ricos creen leyes que les interesen, vuelve a los pobres en contra de sí mismos, con lo que se encuentran divididos y no ejercen el poder contra los ricos, y convierte a los ricos en adictos al placer de dominar a los pobres sin preocuparse de la justicia. Más igualdad económica evita esta situación y, así, apoya a la democracia.

Además, el principio normalmente olvidado de la *fraternidad* necesita, entre otras cosas, que la esfera política no se encuentre dividida y que no haya solamente solidaridad, sino también justicia. Como ha defendido Puyol, la fraternidad es un principio político que puede ser movilizado en «la lucha contra las diferentes formas de exclusión social y política, subyugación, marginalization, explotación, discriminación, colonialismo y humillación» (2019: 2); se puede utilizar para condenar «el abandono social de los miembros más débiles de la sociedad» (*ibíd.*: 6)[2]. En otras palabras, una democracia que no se preocupe de una forma u otra por sus miembros más débiles no merece el nombre de democracia.

Para apoyar aún más la inclusión democrática, este principio puede combinarse con la noción neorrepblicana de la libertad como no dominación. Desarrollado de forma más sistemática por Philip Pettit, el neorrepblicanismo se inspira en el repblicanismo antiguo, pero lo reduce a una teoría de la no dominación que, a su vez, puede enmarcarse como una interpretación de lo que significa el autogobierno democrático. Parafraseando a Beckman y Rosenberg (2017), la idea es que todos los que puedan estar sujetos de forma potencial a la dominación por un poder público deberían disponer de la capacidad para controlar ese poder. Se puede argumentar que esto se debería aplicar especialmente a aquellos que se encuentran en la posición más débil, esto es, aquellos que se hallen en mayor riesgo de ser dominados. El cuidado y el bienestar están bien, pero la justicia y la no dominación son mejores o, al menos, *también* se necesitan. Un propietario de esclavos puede preocuparse por sus esclavos e, incluso, tratarlos bien, pero, aun así, no existe la libertad: continúa habiendo esclavitud y dominación. En una democracia no hay lugar para lo que podemos llamar «esclavitud ilustrada»: una contradicción en sus propios términos.

Otros principios de la democracia son el imperio de la ley, la tolerancia, la transparencia y los derechos humanos que, a su vez, están relacionados con los principios de las revoluciones francesa y

americana. Por ejemplo, el *imperio de la ley* incluye la idea clave de que la ley se aplica a todo el mundo de la misma manera (ya que todos los ciudadanos son iguales ante la ley) de forma justa, y que se respetan los derechos fundamentales. El Estado también necesita ser responsable y que los procedimientos sean transparentes. Lo opuesto a esto son las decisiones arbitrarias de un gobierno autoritario o el tratamiento desigual de los ciudadanos donde la elite tiene más privilegios. El imperio de la ley se halla en peligro cuando las personas son arrestadas de forma arbitraria o cuando un régimen particular practica formas de gobierno sesgadas contra personas que ya se encuentran en desventaja. El imperio de la ley también es una forma de honrar el principio neoliberal de que deberían evitarse el poder arbitrario y la dominación: el imperio de la ley, en lugar del imperio de los hombres, pretende asegurar la libertad en contra de la dominación (Sellers, 2015).

Defendido por el rival de Rousseau, Voltaire, y más tarde por John Stuart Mill, el principio de *tolerancia* se encuentra anclado a una concepción epistemológico-política particular: se basa en la idea de que podemos estar equivocados en lo que toca a nuestras creencias e ideas, y que, por lo tanto, es mejor que seamos tolerantes. En su *Diccionario filosófico* (2007 [1764]), Voltaire usó este argumento cuando criticaba la intolerancia de la Iglesia católica y en defensa de la libertad de prensa. En *Sobre la libertad* (2008 [1859]), Mill defendió que la tolerancia de las ideas de los demás es esencial para el desarrollo del conocimiento, ya que, de nuevo, nuestras creencias son falibles. Pueden no ser verdaderas. Por lo tanto, es mejor que permitamos la libre expresión de la opinión y sometamos nuestras ideas a la crítica. De acuerdo con Mill, la verdad necesita ser verdad viva en el sentido de que se discuta. De este modo, cuestionar y discutir es esencial tanto para la democracia como para el desarrollo de la ciencia. Mill pensaba que la verdad prevalecerá siempre que mantengamos una discusión libre y abierta sobre las ideas. En los Estados Unidos, esto se denomina a veces «el

mercado de las ideas». Dejemos competir a las ideas con libertad, y emergerá la verdad.

Esta idea de la libre discusión y de un libre mercado de ideas no está exenta de problemas. En particular, existe el desafío que describió Karl Popper (2010 [1945]) como la «paradoja de la tolerancia»: si una sociedad es completamente tolerante, otorga libertad a los intolerantes para destruir la sociedad tolerante. Así, la tolerancia ilimitada conduce a su desaparición. Por lo tanto, tiene que haber límites a la tolerancia (y, por consiguiente, al libre mercado de ideas). Sin embargo, desde un punto de vista liberaldemócrata y republicano tales límites requieren una justificación especial y fuerte (por ejemplo, se tendría que demostrar que un tipo particular de discurso es antidemocrático, que conduce a la dominación y, potencialmente, al autoritarismo), y los límites que imponer no deberían ser muy estrechos, dada la importancia democrática de la diversidad de opiniones, la apertura al debate y la libre deliberación. Las exigencias actuales de más censura *online* (desde la izquierda y desde la derecha) deberían evaluarse teniendo esto en cuenta (véase más adelante).

Anteriormente, y desde otro ángulo (la filosofía moral), Kant ya había defendido el uso público de la razón. La publicidad era para él un principio moral. Para usar tu propia razón se necesita libertad de expresión pública. En su famoso ensayo *¿Qué es la Ilustración?* defendía que «el uso público de la propia razón debe ser siempre libre, y solo esto puede lograr la ilustración entre los humanos» (Kant 2012 [1784]: 19). La Ilustración consiste en reflexionar por uno mismo, pero este reflexionar por uno mismo no debería entenderse de forma privada: es una actividad social y pública: necesita realizarse en comunidad con otros y con quienes nos *comunicamos*. Podemos usar nuestra imaginación, pero, en última instancia, también necesitamos el diálogo público: necesitamos hablar con las personas reales para poner a prueba nuestras ideas. Este concepto también se relaciona con la noción de *sentido común*

en *La crítica del juicio*, retomada más tarde por Arendt^[3]. El coraje del que habla Kant cuando recomienda usar la razón por uno mismo no solo se refiere a no aceptar ser guiado por la autoridad, sino también a exponer las propias creencias a su posible vulneración al comunicarlas y discutir las críticamente. De otra forma, solo existe la tiranía de la opinión privada. Además, junto con estos argumentos con respecto a la falibilidad y el uso público de la razón, imponerles la verdad a otros es también incorrecto en términos de libertad. Si no existe tolerancia, algunas personas pueden usar la violencia contra otras que tengan creencias diferentes. Voltaire señaló a autoridades religiosas intolerantes. Pero existen muchos otros ejemplos de violencia en nombre de las creencias. A raíz de la Revolución francesa, por ejemplo, los propios revolucionarios se volvieron intolerantes: tras la creación de la Primera República, tuvo lugar el periodo denominado del «Terror», con numerosas ejecuciones públicas de aquellos que estaban en desacuerdo con «la voluntad general». Por lo tanto, para evitar esto y para proteger las libertades de las personas, debemos ejercitar la tolerancia en la democracia. La alternativa es que algunos impongan sus opiniones sobre las de otros y (teniendo en mente el argumento epistemológico) que las creencias falsas florezcan porque nadie trate de rebatirlas.

¿Cuál es el impacto de la IA en el ejercicio y el florecimiento de estos principios democráticos?

CÓMO DEBILITA LA IA LOS FUNDAMENTOS DE LA DEMOCRACIA

La IA, según está siendo usada y desarrollada actualmente, corre el riesgo de debilitar estos principios democráticos de diferentes maneras. Comencemos por la *libertad*. Como hemos visto en el caso de Cambridge Analytica, la IA se puede usar para manipular las creencias políticas de las personas. La manipulación es una forma de

influencia que está oculta intencionadamente (en este caso, para influir en las creencias políticas). Pero ¿por qué, exactamente, es esto problemático? Un argumento que propongo es que interfiere con la libertad y la autonomía de los ciudadanos a la hora de conformar sus creencias^[4]. No se trata solo del habitual llamamiento a proteger la autonomía individual: en lugar de considerar la autonomía individual por sí misma, se añade aquí la dimensión epistémica del problema (la manipulación de las *creencias*; véase también el capítulo 5), y esto enlaza con una concepción particular de la *democracia*. En una democracia del tipo desarrollado hasta ahora, los ciudadanos deberían estar libres de las interferencias procedentes de la manipulación cuando se trata de tomar una decisión con respecto a quién van a votar. Este no es el caso cuando se pervierte su autonomía para tomar decisiones, cuando su formación de creencias se ve influenciada, por así decirlo, «bajo el radar», esto es, sin su conocimiento y haciendo uso de sus debilidades psicológicas. Esto ya se había hecho usando la propaganda política tradicional, pero con la IA (por ejemplo, en el caso de Cambridge Analytica) esto es mucho más rápido, alcanza a más personas y es especialmente mucho más *dirigido*, ya que está basado en perfiles individuales explotando a su máxima potencia las vulnerabilidades individuales: los mencionados demonios interiores. En otras palabras, tanto la función de la propaganda política, como el resultado de socavar la autonomía y la democracia, son mucho más poderosos y efectivos.

Además, la IA permite la automatización de la vigilancia, la política predictiva y las decisiones judiciales, que se traducen en riesgos para la libertad de los ciudadanos: basándose en estos usos de la IA, los ciudadanos pueden ser arrestados y detenidos injustamente. Un caso bien conocido es el de un hombre negro de los Estados Unidos que fue arrestado equivocadamente y sin explicación por un algoritmo de reconocimiento facial que lo identificó erróneamente (Porter, 2020). Este es un problema para la

libertad, pero también para el imperio de la ley, ya que cuando un algoritmo así toma decisiones sin juicio humano y una adecuada aplicación de la ley, arrestar a las personas se vuelve arbitrario, injustificado y basado en un proceso que no es transparente, lo que viola sus derechos fundamentales. También es un problema para la justicia, ya que usar tales algoritmos puede fomentar la discriminación contra individuos y grupos concretos de la sociedad (por ejemplo, en el contexto de los Estados Unidos, la discriminación estructural contra la gente de color), que, a veces, también tiene consecuencias para sus derechos a la libertad. Mientras que, en principio, se puede mejorar la justicia de los algoritmos, la ausencia de juicio humano y la no transparencia para con los sospechosos y los acusados continúa siendo problemática. Además, los regímenes autoritarios y quienes los apoyan pueden usar la tecnología para influir en las elecciones de países democráticos. A través de estas manipulaciones y amenazas a la libertad y al imperio de la ley, la IA también fomenta tendencias antidemocráticas y autoritarias en sistemas políticos que se supone que son democráticos.

Además, como ha mostrado Splichal (2022a) en su obra sobre la minería de opinión política (*political opinion mining*), el ideal ilustrado del uso público de la razón y la libre expresión de la opinión en la esfera pública (como lo expresaron Kant y Mill) han sido sustituidos por la individualización de la opinión pública y la extracción tecnológica de datos. La opinión pública se ha convertido en una colección de opiniones particulares, que, transformadas en datos, pueden usarse para manipular a los ciudadanos. En lugar de razonar, discutir y poner a prueba las ideas públicamente en comunidad con otros y, ciertamente, en lugar de *constituir* y mantener esa comunidad y esfera pública a través de tales discusiones y deliberaciones, la ciencia de datos y las redes sociales meramente agregan y calculan opiniones. Al tiempo crean una situación en la que afloran las opiniones privadas, pero no se ponen a prueba por su «valía pública», como habría querido Kant (Splichal,

2022b). Junto con las redes sociales, la minería de opinión mediante ciencia de datos difumina la distinción público/privado y es, por lo tanto, una amenaza para la formación de una opinión y un debate públicos tal y como los habían imaginado los pensadores ilustrados. Es la celebración y, a menudo, la tiranía, de la opinión privada.

En segundo lugar, es posible utilizar la IA de tal modo que se mantengan o generen *desigualdades* económicas. La automatización mediante la IA a menudo (pero no necesariamente) tiene como resultado la pérdida de empleos. Y si tienes un trabajo, el algoritmo puede decidir que lo haces lo suficientemente bien o que te comportas de formas que el algoritmo detecta como fraudulentas. Por ejemplo, Über gestiona sus conductores con IA, y se lo ha acusado de despedir a conductores injustamente basándose en el análisis de sus datos (HR-inform n.d.). Dado que este uso de la IA se dirige a grupos de trabajadores concretos (trabajadores de la economía colaborativa, quienes ya se encontraban en situación precaria), y dado que el proceso de decisión está automatizado, sin juicio humano, este problema se puede considerar una cuestión concerniente a la justicia y a la igualdad, junto con otro relativo a la dominación y la explotación. A través del sesgo y de la manipulación que explota la psicología humana, la IA también discrimina y causa impacto en las personas sin que sus resultados sean explicables (O'Neil, 2016). Por ejemplo, basándose en el análisis de *big data*, hay personas a las que no se les conceden créditos o empleos debido a su trasfondo étnico o a su código postal, a veces sin explicación y, a menudo, manteniendo la segregación, desigualdad y pobreza. Es este un problema para la igualdad, así como una amenaza para el imperio de la ley (no hay explicación ni transparencia) y al principio de fraternidad (los más débiles están aún más marginados).

Aquí también puede usarse el concepto de *justicia*, en particular, el de justicia como equidad. Existe en la actualidad una bibliografía emergente sobre la denominada desigualdad algorítmica (*algorithmic*

unfairness), y la equidad es un principio avalado en muchos documentos sobre la política de la IA^[5]. La idea es que, dado que las decisiones de la IA pueden conducir a resultados injustos, desiguales, sesgados y discriminatorios, como las desventajas basadas en la etnicidad o el género, es importante asegurarse de que se usa la IA de una forma justa y equitativa. Sin embargo, empíricamente hablando, existen diferentes percepciones con respecto a lo que es justo (Starke *et al.*, 2022), y en filosofía política existen diferentes opiniones sobre la justicia y la equidad que también se aplican a los debates acerca de la ética y la política de la IA (Coeckelbergh, 2022a). Por ejemplo, existe una tensión entre las concepciones universalistas de la justicia y las concepciones basadas en la identidad de grupo: ¿un tratamiento particular de (aquellos que usan) un algoritmo es injusto porque no trata a todos de la misma forma, o es injusto porque no logra, por ejemplo, evitar un sesgo contra un grupo históricamente en desventaja? Ya he mencionado el algoritmo COMPAS y la queja de que discrimina a los encausados de color, que podría justificarse sobre la base de una concepción universal de la justicia (tratar a todos de la misma manera, *independientemente* de su identidad personal o de grupo), o argumentando que el algoritmo se suma a la discriminación histórica, estructural y actual contra un grupo ya en desventaja en una sociedad concreta.

Téngase en cuenta también que la igualdad y la justicia como equidad son dos principios diferentes. La igualdad puede ser injusta en algunos casos y una distribución justa puede ser una distribución desigual. Por ejemplo, en su discusión de la inequidad del aprendizaje automático, Mittelstadt, Wachter y Russell (2023) se posicionan en contra de limitar la equidad y la justicia distributiva a una forma de igualitarismo estricto en la que solo cuenta la igualdad del resultado. Por ejemplo, rebajar a todo el mundo al nivel del peor parado, puede contar perfectamente como tratar a todo el mundo por igual, pero esto no es necesariamente justo y equitativo. Para la

democracia, ambos principios son relevantes, y el impacto en ella de la IA puede ser analizado igualmente en términos de su impacto en la justicia y la equidad (además, la justicia y la equidad también se pueden conectar con el principio de fraternidad, véase más adelante).

Otro concepto que se usa a menudo en relación con la igualdad es el de *poder*. Como ya han apuntado muchos comentaristas (Hawley, 2021; Zuboff, 2019), la IA aumenta las asimetrías de poder: el poder de las grandes tecnológicas amenaza a la democracia, al igual que lo hace lo que podríamos llamar el «capitalismo de IA». La IA no se ha desarrollado, ni tampoco usado, en un orden social neutro, sino que está enredada con el capitalismo. No es solo una tecnología, sino también una relación social y una forma de organización social. Dyer-Witheford y sus colegas hablan en su libro *Inhuman Power* acerca de «Capital-IA»: la IA es un instrumento del capital que conduce a la explotación del trabajo humano, al control de poblaciones y a una «concentración de la riqueza y del poder social en manos de los propietarios de las corporaciones de alta tecnología» (2019: 3). Esto significa no solo que la IA contribuye a la desigualdad general en la sociedad que, como hemos visto, es un problema para la democracia, sino también que la IA ejerce un impacto en la distribución del poder: el futuro de la IA lo deciden los directivos de empresa y sus accionistas. Las grandes empresas tecnológicas también toman otras decisiones políticas relevantes para la tecnología. Por ejemplo, Facebook, una compañía no elegida democráticamente, decide los límites de la libre expresión en su generalmente usada plataforma. Y si las grandes empresas tecnológicas tienen una influencia tan grande en nuestras vidas cotidianas, ¿es correcto que no sean responsables ante el bien público (véase también Brown, 2020)? Si las grandes tecnológicas son más poderosas que muchos países, ¿en qué sentido gobierna aún el *demos*, el pueblo? La situación actual no es democrática.

La crítica con respecto a la desigualdad y al poder también puede plantearse como la observación de una tensión fundamental entre la esfera pública y el capitalismo. Apoyándose en Habermas (1991) y subrayando la dimensión marxista, Fuchs (2023) ha defendido que la esfera pública tiene que ver con la idea de que todos se benefician de las instituciones, no solamente unos pocos privilegiados, que la esfera pública es un sistema comunicativo abierto a todos (2023: 273) y que se supone que es autónomo frente al capital (*ibíd.*: 277). Sin embargo, cuando nuestras prácticas comunicativas están mediadas por plataformas digitales en un entorno capitalista, tiende a producirse una «colonización capitalista de la esfera pública digital» (*ibíd.*: 281). Entre otros problemas tales que un acceso desigual, la debilitación de la solidaridad y la expansión del autoritarismo y de las noticias falsas, existe una fragmentación de la esfera pública. La IA, tal como se usa en y por las plataformas digitales, puede contribuir a dicha fragmentación cuando, por ejemplo, ayuda a crear filtros burbuja en los que las opiniones son homogéneas (*ibíd.*: 282). Esto nos conduce a la siguiente cuestión.

En tercer lugar, la IA puede conducir a la injusticia y a apoyar la división en la sociedad, socavando, así, la *fraternidad*. Por ejemplo, hay cierta preocupación de que la IA conduzca al sesgo y a la polarización. Ya he mencionado ambos problemas. Cuando la IA discrimina de forma injusta y ayuda a crear una sociedad que sea, a su vez, altamente injusta, la fraternidad y, por tanto, la democracia se ven debilitadas. Otra crítica importante se ha formulado en términos de los filtros burbuja (Pariser, 2011) y, posteriormente, las cámaras de eco y las burbujas epistémicas (Nguyen, 2020). Lo preocupante de las burbujas de filtros es que los algoritmos dotados de IA de plataformas como Facebook y Google presenten solo información y puntos de vista a los usuarios en función de sus preferencias y, por tanto, ofrezcan menos contacto con otros puntos de vista. Las burbujas epistémicas y las cámaras de eco tienen efectos similares:

Una burbuja epistémica es una estructura epistémica social en la que otras voces relevantes han quedado fuera, quizá accidentalmente. Una cámara de eco es una estructura epistémica social en la que otras voces relevantes han sido activamente excluidas y desacreditadas. Los miembros de una burbuja epistémica no están expuestos a información y argumentos relevantes. Por otra parte, los miembros de las cámaras de eco han sido preparados para desconfiar sistemáticamente de todas las fuentes externas. (Nguyen, 2020: resumen)

Se ha dicho que la IA y las redes sociales conducen a burbujas epistémicas y cámaras de eco: las personas no están suficientemente expuestas a otras opiniones y no confían en otras fuentes o desacreditan activamente otras voces. A esta alegación se le han presentado objeciones. Por ejemplo, cuando se trata de filtrar las noticias, la investigación empírica ha mostrado que la mayoría de la gente consume dietas mediáticas diversas, lo que ha conducido a que algunos defiendan que las burbujas epistémicas y las cámaras de eco son solo un problema para una minoría (Arguedas *et al.*, 2022; Fletcher, 2020). Además, el papel de los medios de comunicación tradicionales no debería pasarse por alto: por ejemplo, la televisión también puede conducir a la polarización, quizás en mayor medida (Hosseinmardi, 2022). Con todo, consumir diversas dietas de medios no implica necesariamente que las personas abandonen realmente sus burbujas epistémicas y políticas (pueden continuar estando expuestas principalmente a opiniones similares a las suyas en diversos medios). Asimismo, la preocupación sobre la polarización y la división continúa siendo relevante cuando, y en la medida en que, la gente continúe simplemente expresando sus opiniones y no logre interactuar realmente con otras personas y sus opiniones, ya sea en los entornos de las redes sociales o en cualquier otra parte. La fraternidad no está en peligro solo cuando la gente deja de estar *expuesta* a otras opiniones, sino también cuando, en palabras de Gunn y Lynch (2021), deja de tratar con respeto a otros agentes epistémicos y deja de querer aprender de

ellos. En otras palabras, la preocupación que se nos plantea aquí consiste en que los ciudadanos carecen de las virtudes epistémicas necesarias^[6] para la democracia y realmente no se comunican ni intentan comprender los puntos de vista de los demás: todo muy importante en las tradiciones republicanas e ilustradas.

Este problema también se puede formular como una amenaza para la *tolerancia*. Combinada con las redes sociales, la IA también puede apoyar la intolerancia cuando desalienta la escucha a los demás y la discusión de las ideas. Las guerras de opinión pueden conducir a la violencia, a ataques a la ciencia y a la búsqueda de la verdad. Y cuando la IA, como se usa en las redes sociales digitales, facilita la política de identidad (de la derecha y de la izquierda), que corre el riesgo de dividir y polarizar, puede conducir a más intolerancia y a una mayor fragmentación y destrucción de la comunicación y de la esfera pública. Esto es perjudicial para la democracia, de acuerdo con las tradiciones liberal y republicana.

Dada la antes mencionada «paradoja de la tolerancia», esto implica que deben establecerse algunos límites a la libertad de expresión *online*, a la vez que se mantiene la esfera pública tan abierta como sea posible y, ciertamente, *para mantenerla abierta*. La IA puede desempeñar un papel a la hora de implementar estos límites: hasta cierto punto, puede estar programada en algoritmos de las redes sociales. Pero esto también crea sus propios desafíos. ¿Quién decide los límites? ¿Son transparentes? ¿Y cuánta censura (porque se trata de censura) está justificada en una democracia?

La libertad de expresión *online* se ha debatido ampliamente. Por ejemplo, Warburton (2009) ha señalado sus peligros, tales como una falta de control de calidad y la comunicación antisocial: mientras que esto era posible antes, por ejemplo, en los periódicos, ahora puede alcanzar a millones de personas a la vez y hacer mucho más daño. Al mismo tiempo, defiende Warburton, la censura es más difícil de implementar en internet. Hoy en día, en la era de la IA, este tema es de nuevo muy relevante, y continuaré esta discusión en el

capítulo 5. Por ahora, señalaré que con los algoritmos de las redes sociales y con la IA (y la forma en que dominan lo que se suponía que era la esfera pública), la censura es realmente más fácil de implementar, y eso es peligroso para la democracia. Es más, continúa existiendo el desafío de la «paradoja de la tolerancia». Al discutir la libertad de expresión con respecto a la democracia (y al responder a demandas generalizadas a favor de una mayor censura *online*), necesitamos estar guiados por la consideración normativa de que necesitamos alcanzar un equilibrio entre, por una parte, asegurar la suficiente variedad de opiniones y, por otra, evitar la intolerancia que destruye la esfera pública y la democracia. Como sugerí anteriormente, la democracia liberal debe interpretarse, de acuerdo con las líneas republicanas, en el sentido de que exige que la mayoría del peso recaiga en la protección de la apertura y diversidad de la esfera pública. Si no, la IA se arriesga a ser el instrumento de censura que conducirá al fin de la sociedad abierta y, ciertamente, de la democracia.

Pero la IA no solo se arriesga a socavar las democracias liberales de estas diversas maneras: también apoya regímenes autoritarios, fortalece tendencias antidemocráticas y, a veces, autoritarias en otros regímenes. Como ya he sugerido anteriormente, el mundo político es más complejo de lo que se puede captar teniendo en cuenta una simple oposición binaria entre democracia y autoritarismo. Muchos sistemas políticos del mundo son sistemas «mixtos». Pero esto no hace que el argumento sea irrelevante: al contrario. Al fomentar la manipulación, la vigilancia, la división y la intolerancia, la IA fortalece el lado autoritario de los sistemas mixtos y les permite desalentar de forma más efectiva la libre discusión de las ideas y reprimir a sus ciudadanos. Además, los sistemas que ya eran autoritarios es probable que aumenten su autoritarismo. Los países como Rusia y China usan la IA y otras tecnologías digitales para la vigilancia de sus propios ciudadanos y para la censura, aumentando, así, su poder autoritario. Desde una perspectiva

liberaldemócrata, este es un problema para la libertad, pero también para otros principios, como la tolerancia y la fraternidad. La libertad se pierde cuando el Gobierno se inmiscuye en tu libertad y (por ejemplo) te arresta. Pero tu libertad también se ve en peligro cuando tienes que preocuparte todo el tiempo por si estás siendo monitorizado, cuando desconfías del Gobierno y, en última instancia, de otras personas, y cuando, en consecuencia, adaptas habitualmente tu comportamiento de formas que equivalen a conformarse con las reglas y las expectativas de otros. La IA es, en estas circunstancias, parte del denominado «autoritarismo digital» (Shabaz, 2018) en el que estas tecnologías facilitan que los regímenes autoritarios controlen a sus ciudadanos; y ayuda a crear un sistema más *totalitario*, una forma de gobierno autoritario en el que ya es prácticamente imposible escapar de la vigilancia y el control del Gobierno, y en el que (como resultado) los ciudadanos se vigilan de cerca unos a otros y participan en la violencia sancionada por el Gobierno contra otros (que, frecuentemente, les señala como no ciudadanos).

Por ejemplo, Feldstein (2019b) ha defendido que en países como China la IA ayuda a reprimir contragrupos, que sin la automatización suele ser normalmente costosa y laboriosa, ya que se trata de un trabajo intensivo, y permite controlar el descontento y la difusión de noticias falseadas. De una forma similar, Kendall-Taylor y sus colegas (2020) han defendido que los gobiernos autocráticos que adoptan las tecnologías digitales como la IA para la vigilancia se han vuelto más resilientes a las protestas. Además, el autoritarismo patrocinado por la IA se expande: los países que lo lideran (de nuevo, se usa a China como ejemplo) a menudo venden la tecnología a otros países que tienen aspiraciones similares.

Zeng (2020) ha defendido que no es la propia IA la que es un problema, sino ella en combinación con, por ejemplo, componentes ideológicos. Pero, de todas formas, no existe algo que sea una «IA en sí misma». La IA está siempre conectada con otros sistemas

humanos. Aquí, lo que yo defiendo es que la IA (tal como la usan los humanos e integrada en otros sistemas) no solo erosiona las democracias, sino que fortalece las formas ya existentes de autoritarismo y totalitarismo, violando de esta manera los principios de libertad y de fraternidad, y haciendo imposible la aplicación de los ideales republicanos y de la Ilustración, como la tolerancia y la participación en la esfera pública.

A pesar de que no hay un determinismo tecnológico y que los humanos y sus ideologías siempre desempeñan algún papel, parece incluso que se da una relación inherente entre el autoritarismo y la cibernética platónica, que es, en esencia, antidemocrática. Como describe el *Economist* a China, usando la metáfora de Platón: «el Partido [Comunista chino] no presta atención a los individuos disidentes o a las minorías infelices en mayor medida que un timonel [...] discute con los pasajeros la ruta más segura a través de una tormenta» (Rennie, 2022). La cibernética no es solo ciencia y tecnología: también es filosofía política y, de esta forma, una filosofía autoritaria y potencialmente totalitaria. Por ejemplo, como han defendido Reijers, Orgad y De Filippi (2022), el sistema de crédito social chino ilustra un movimiento hacia lo que ellos denominan «ciudadanía cibernética»: un nuevo modelo de gobernanza y un sistema sociotecnológico para rastrear, evaluar y modular el comportamiento de los ciudadanos. Dentro del marco de una cibernética de este tipo, la IA corre el riesgo de convertirse en el instrumento de un timonel platónico autoritario y potencialmente totalitario. El piloto gobierna y lo ve todo, por lo que gobernar la nave se vuelve un mero asunto de algoritmos: recibir instrucciones sobre lo que se debe hacer.

En el siguiente capítulo, me detendré en los orígenes del totalitarismo digital.

Capítulo 5

Cómo erosiona la IA el conocimiento y la confianza

La democracia no consiste solo en procedimientos, reglas y principios. Cuando examinamos sus condiciones resulta evidente que también tiene que ver con el conocimiento, las relaciones sociales y la confianza. Si queremos saber más acerca de la relación entre la IA y la democracia, y comprender mejor los riesgos (incluyendo los orígenes del totalitarismo digital), tenemos que indagar en las bases sociales y epistémicas de la democracia. Y hoy en día dichas bases están en crisis, debido, parcialmente, a las tecnologías digitales. El conocimiento se ve ampliado y mediado por los medios y tecnologías digitales, que han otorgado a la humanidad nuevas oportunidades gigantescas, pero que también la han conducido hacia problemas epistémicos serios. Por ejemplo, internet nos ofrece un exceso de información, existe desinformación en las redes sociales, los chatbots cometen errores enormes y, en un momento en el que todo el mundo puede buscar en Google, publicar una entrada y crear textos, se ve cuestionada la autoridad de la pericia científica. La IA se considera a menudo parte del problema. Algunos la ven como parte del gobierno de la elite y, por lo demás, ya no confían en los

expertos o en la democracia. Existe un auge del populismo, de los movimientos de extrema derecha y de las teorías conspirativas. La IA, así lo parece, corre el riesgo de empeorar esta crisis más que de resolverla. Pero ¿cuáles son, exactamente, los problemas potenciales (reales) y cuál es la contribución de la IA?

EL PROBLEMA DEL TOTALITARISMO REVISADO: LAS LECCIONES DE HANNAH ARENDT

Una forma de entender algunos problemas epistémicos relevantes para la democracia, y cómo pueden empeorar cuando se involucra en ellos a la IA, es revisar lo opuesto de la democracia: el autoritarismo y, en particular, el totalitarismo. A pesar de que en la actualidad las democracias occidentales no se encuentran en modo alguno cerca del totalitarismo (o del autoritarismo, en este caso), estudiar las condiciones del totalitarismo nos ayuda a subrayar con mayor claridad los riesgos relacionados con el conocimiento que plantea la IA para la democracia. El totalitarismo es una forma de gobierno que no solo entraña la centralización y la monopolización del poder político por una dictadura (autoritarismo), sino que ejerce un alto grado de control sobre los ciudadanos, destruyendo la libertad humana e interfiriendo en todos los aspectos de la vida. Esto conlleva también una dimensión que afecta al conocimiento: el totalitarismo impacta en el conocimiento de la sociedad.

La obra de Hannah Arendt es útil para comprender y evaluar el totalitarismo, y para estudiar sus implicaciones para la democracia y el conocimiento. Como muchos otros pensadores de su generación, Arendt defendió que el totalitarismo es peligroso porque arrebató las libertades. Por ejemplo, arrestar a la gente sin justificación es totalitario. Equivale a la destrucción total del imperio de la ley. Y, en contraste con otras formas de opresión, el totalitarismo opera usando el terror. Arendt describe cómo el totalitarismo en la forma

de nazismo y de estalinismo oprimieron a las masas y ejercieron un poder arbitrario sobre el pueblo. El totalitarismo niega y destruye la libertad humana.

Arendt describe formas de totalitarismo que surgieron en circunstancias históricas específicas que difieren de las actuales. Pero su análisis puede ser una lente a través de la cual podamos comprender mejor los riesgos antidemocráticos y potencialmente totalitarios que surgen con las nuevas tecnologías. El ejercicio del poder arbitrario es un ejemplo de tal riesgo. Con respecto a la IA, ya he mencionado la historia kafkiana del arresto injustificado y, al principio aparentemente inexplicable, de una persona sobre la base de la recomendación de un algoritmo de reconocimiento facial.

Pero el totalitarismo es algo más que una amenaza para la libertad. La IA también se puede usar para la gestión burocrática de las personas, dirigida e implementada por personas (y, quizás, en el futuro por máquinas) que simplemente siguen órdenes. Arendt denominó a esto «la banalidad del mal» (2006 [1963]). En su informe sobre el juicio al líder nazi Adolf Eichmann, defendió que este no era un monstruo sádico, sino un burócrata aburrido que afirmaba simplemente haber seguido órdenes. Así pues, lo que hizo fue cometer actos malvados sin intenciones malvadas. El problema era más bien la irreflexión, desconsideración y falta de empatía.

Ahora bien, este problema también se puede relacionar con la IA y la ciencia de datos. Usada de una forma burocrática, que aliena a las personas reales y desalienta el compromiso con sus consecuencias, la ciencia de datos puede tener efectos negativos sobre los individuos y la sociedad (incluso apoyando potencialmente el asesinato en masa y el genocidio sistemático), sin implicar necesariamente que los científicos de datos sean malvados o existan intenciones malvadas por parte de estos científicos. En *Privacy is Power*, Carissa Veliz hace una comparación entre la creación de registros de judíos por los nazis: «la recogida de datos puede matar» (2020: 114). Este caso histórico no solo ilustra la destrucción

total de la libertad y la fraternidad bajo condiciones totalitarias, sino que también muestra cómo puede contribuir la IA a este fenómeno de la banalidad del mal cuando se usa en un contexto social y político que promueve la irreflexión y el uso burocrático desvinculado de la tecnología. Actualmente, las circunstancias históricas son diferentes. Pero merece la pena tener en mente que la IA, en combinación con una burocracia y con un funcionamiento fluido, puede conllevar consecuencias devastadoras para las vidas de las personas, por ejemplo, cuando la gente pierde sus subsidios de desempleo y cae en la pobreza debido a una combinación de automatización facilitada por la IA y de irreflexión burocrática.

Sin embargo, curiosamente, Arendt también ofrece razones adicionales relacionadas con el conocimiento por las que el totalitarismo en particular es tan problemático moral y políticamente, si no malvado: destruye la verdad y se alimenta de las relaciones sociales deterioradas y de la desconfianza entre las personas. El problema no es solo el propio totalitarismo, sino también el mundo que lo hace posible: sus orígenes, incluyendo el mundo epistémico del que se alimenta. En primer lugar, las reglas totalitarias buscan explícitamente la destrucción de la verdad. Actualmente, la IA fomenta este objetivo cuando contribuye a la desinformación y permite manipulaciones en las elecciones (Helmus, 2022). Aquí, los *deepfakes* (vídeos en los que se cambian digitalmente las imágenes de personas para que parezca que son otras) resultan especialmente peligrosos, ya que la gente no puede detectarlos (Dobber *et al.*, 2021). Pero el problema no está solo en los casos específicos de manipulación y mentira. Existe un peligro socioepistémico mayor: en última instancia, estos tipos de desinformación y *deepfakes* intencionados, pero también la desinformación no intencionada, como un texto generado por IA que parece fidedigno aunque contenga falsedades, pueden conducir a un entorno epistémico en el que ya no sea posible distinguir la verdad de la falsedad. Esto es desastroso para la democracia: peor aún, es exactamente lo que

necesita el totalitarismo. En *Los orígenes del totalitarismo* (2017 [1951]), Arendt escribe que los gobiernos totalitarios crean un «mundo ficticio mintiendo de forma consistente» (2017: 499). Desdennan la verdad y la realidad, blindándose ellos mismos y sus seguidores «del impacto de la facticidad» (*ibíd.*: 549). En la medida en que la IA contribuye a crear esta situación, por ejemplo, al aumentar la cantidad de desinformación y producir un tipo de burbuja epistémica en la que se rechaza la información fáctica del exterior, debilita la democracia y contribuye, en última instancia, a crear las condiciones para el auge del totalitarismo.

En segundo lugar, desde Arendt y la tradición republicana en la que se inscribe podemos aprender que la democracia solo es posible si continuamos compartiendo un mundo y si no nos encontramos totalmente aislados unos de otros. La vigilancia y el terror totalitarios ejercidos por un poder estatal funcionan bien cuando las personas están ante todo socialmente desconectadas, aisladas, y ya no confían las unas en las otras (*ibíd.*: 628). Antes, Durkheim (1997 [1897]) había subrayado la importancia de la conexión social y mostrado que su ausencia en la sociedad moderna puede conducir al suicidio. Pero la desconexión social también puede conllevar consecuencias políticas: como ha defendido Robert D. Putnam en su libro sobre el colapso de la comunidad en los Estados Unidos, *Solo en la bolera* (2002), un declive en el tejido social socava el activo compromiso cívico necesario en una democracia fuerte. Con Arendt, también podemos añadir a esto que el aislamiento puede no solo debilitar la democracia, sino también crear las condiciones para el totalitarismo. Actualmente puede facilitar el uso totalitario de la IA para la vigilancia y el ejercicio del terror por parte del Estado. También en este sentido, el aislamiento mata. Además, de acuerdo con Arendt, el juicio político requiere sentido *común* (1968), estar interesado en el bien común y ser capaz de usar la imaginación propia para tomar contacto con las posturas de los demás. Si estamos simplemente defendiendo nuestra propia opinión, pero no

estamos interesados en comprender los puntos de vista de los demás, no podemos alcanzar un entendimiento mutuo. Cuando esto ya no es posible, se destruye la base para la democracia: no solo deja de haber fraternidad y tolerancia, sino que tampoco son posibles un entendimiento ni un mundo comunes. Si, y en tanto que la IA, en combinación con las redes sociales digitales, ayuda a crear un entorno político que pueda conducir al aislamiento, la desconfianza y la falta de interés por los puntos de vista de los demás, entonces se destruirán las condiciones para la emergencia del sentido común y, así, se facilitará el surgimiento de tendencias antidemocráticas y potencialmente totalitarias. La verdadera comunicación, en el sentido republicano e ilustrado de la palabra, ya no es posible (hablaré más en profundidad sobre la comunicación y la democracia en el capítulo 8).

Sin embargo, incluso si la IA no conduce al totalitarismo, existen más problemas que están asociados a ella, el conocimiento y la democracia.

ALGUNOS DESAFÍOS EN RELACIÓN CON LA IA, EL CONOCIMIENTO Y LA DEMOCRACIA

Mirando más allá de la teoría totalitaria y enlazando de nuevo con algunos de los puntos planteados en el capítulo anterior acerca de los peligros que se ciernen sobre la democracia, veo, al menos, los siguientes desafíos.

Primero, debido al tipo de conocimiento que producen y a los modos en que lo producen, la IA y la ciencia de datos pueden contribuir a aumentar las asimetrías de poder en la sociedad que, como hemos visto, es un asunto problemático para la democracia, pues se corre el riesgo de destruir la igualdad necesaria (o, al menos, un nivel mínimo de la misma), la justicia y la fraternidad. La IA crea una asimetría en términos de conocimiento entre, por una

parte, los gobiernos y las grandes corporaciones tecnológicas y sus expertos y, por otra, los ciudadanos (Couldry y Mejias, 2020; Zarkadakis, 2020). Esa asimetría en el conocimiento se traduce en una asimetría de poder: los expertos, las personas y las empresas en las que trabajan obtienen más poder. Esto es un problema por sí mismo, pero respecto a la democracia, amenaza sus principios fundamentales y es especialmente problemático para sus versiones deliberativa y participativa, que requieren la participación de todos los ciudadanos y no solo de los expertos. Es, incluso, un problema para la democracia definida en términos de votación: la gente ya no entiende lo que está ocurriendo en la sociedad, no digamos ya que puedan deliberar acerca del bien común sobre esta base. Y si uno no sabe lo que está ocurriendo, serán otros los que tomen las decisiones que le afecten sobre la sociedad y su futuro. El conocimiento es poder: cuando otros «lo saben prácticamente todo sobre nosotros» por medio de la IA, entonces tienen el poder (Veliz, 2020: 86). Así, los ciudadanos pierden la capacidad de actuar, tanto la agencia epistémica como la agencia política. Al no tener acceso al conocimiento, no toman decisiones con respecto a sus datos y al uso de la IA. Otros deciden: los gobiernos y las grandes tecnológicas.

La minería y el análisis de datos por parte de la IA es bueno para las empresas privadas cuyo modelo de negocio es vender publicidad basándose en el análisis de datos. Pero no venden solo publicidad. También recogen mucha información sobre nosotros, y esta información puede usarse potencialmente para manipularnos y controlarnos; por ejemplo, para influir de forma encubierta en lo que (queremos) comprar. Asimismo es tentador para los gobiernos emplear la IA y los *big data* para tomar decisiones y ofrecer servicios digitales. Ejemplos de lo que Zarkadakis (2020) denomina «automatizar el gobierno» son los casos de bienestar social que mencioné en el primer capítulo, pero también lo son cómo lidiaron los gobiernos de todo el mundo con la pandemia del covid-19. Usando LA, *big data* y conocimiento experto (médico), los gobiernos

crearon una situación de conocimiento asimétrica para la gobernanza de la pandemia, que se tradujo en una brecha de poder cada vez mayor entre los gobiernos y los ciudadanos, desplazando a las sociedades en la dirección de las tecnocracias, incluso quizás hacia el autoritarismo digital. Curiosamente, a menudo continuó habiendo un gobierno tecnocrático después de que pasara lo peor de la pandemia. Como ha señalado un estudio encargado por el Parlamento Europeo, la pandemia fue usada por los gobiernos para desplegar una gama de nuevas medidas de vigilancia y, con frecuencia, para reemplazar el juicio humano por la toma de decisiones algorítmica (Glowacka *et al.*, 2021).

Esto ha conducido, a su vez, a (más) desconfianza de los ciudadanos respecto a sus gobiernos, que tienen sus datos y toman decisiones políticas sobre ellos basándose en tales datos. Los movimientos de derechas y antigubernamentales se han beneficiado de esta desconfianza.

Sin embargo, el problema del conocimiento/poder no se debe solo al uso y a la apropiación de los datos, sino que reside también en cómo se crean los datos. A veces, los conjuntos de datos no son representativos y excluyen a (grupos de) personas. Cuando estos conjuntos de datos se usan para tomar decisiones con respecto a personas y cuando estas decisiones dan forma a la sociedad, el asunto se convierte en un problema político. Lo político en la IA, por tanto, se da desde el momento en el que se crea un conjunto de datos, pues esto ya implica exclusión. Como lo ha expresado Bartoletti, somos introducidos en conjuntos de datos por una «fuerza invisible» (2020: 38), lo cual implica que otras personas pueden tomar decisiones acerca de quién se vuelve parte del conjunto de datos y quién queda fuera. Esta es una decisión política. En una democracia, las decisiones con respecto a quién está «dentro» y quién «fuera» deberían hacerse de forma democrática, esto es, por parte de los ciudadanos o sus representantes. Pero son más bien otros, en particular los gobiernos y las (grandes) corporaciones

tecnológicas, las que toman estas decisiones en nuestro nombre basándose en el conocimiento del que disponen. Tales decisiones son tecnocráticas, no democráticas.

Además, podemos entender de nuevo estas asimetrías de poder como parte de una forma particular de capitalismo (Dyer-Witheford, Mikkola Kjosen y Steinhoff, 2019; Fuchs, 2023; Zubof, 2019). La actual economía de datos es de carácter capitalista, otorga poder a las grandes corporaciones tecnológicas. El conocimiento desempeña un papel central a la hora de mantener el capitalismo de IA. Pero, incluso en una sociedad no capitalista en la que los medios de producción sean colectivos, puede darse la tecnocracia si el poder está centralizado y se basa en un conocimiento que no esté, necesariamente, compartido con los ciudadanos. Podría decirse que este es el caso de los sistemas autoritarios del siglo xx denominados «comunistas» o «socialistas». En países poscomunistas como China y Rusia vemos una continuación de esta tradición política o, al menos, sus asimetrías de conocimiento/poder. En los regímenes autoritarios del siglo xxi, la IA ayuda a crear y a mantener las asimetrías de poder entre el gobierno autoritario centralizado y tecnocrático y sus ciudadanos.

Segundo, en una democracia los ciudadanos necesitan tener conocimiento político y formar de modo libre sus creencias políticas, pero es cuestionable si esto es posible cuando el uso de la IA contribuye a la manipulación, la desinformación y la polarización. Incluso si no es visible de forma inmediata una forma de gobierno totalitario, la democracia corre el riesgo de verse debilitada de diferentes maneras. Algunos de estos riesgos se pueden formular en términos de una falta de agencia epistémica que concierne al control que tienen los ciudadanos con respecto a sus creencias, su formación y su revisión potencial (Coeckelbergh, 2022b). Por ejemplo, si tus creencias políticas están manipuladas, ¿qué significa tu libertad para votar o tu derecho de deliberar y participar en discusiones públicas? Además, estos peligros también se pueden

conceptualizar en términos de una base epistémica dañada para los ideales republicanos e ilustrados más ricos de la democracia, por ejemplo, la democracia deliberativa y conceptos relacionados, como la diversidad de opiniones, la tolerancia y el mercado de ideas. Elaboraré un poco más este punto y me centraré en el fenómeno de la manipulación y de las burbujas epistémicas.

Como he defendido en capítulos anteriores, la *manipulación* es tanto un problema para la libertad y la autonomía como para el ideal de democracia relacionado que necesita de estos conceptos. Al influir en las personas sin su conocimiento y pasar bajo el radar de sus capacidades racionales autónomas, la IA no las respeta y viola la autonomía y libertad necesarias para la democracia, ya que supone que deliberamos con los demás y tomamos decisiones. Pero la manipulación no es solo un problema en relación con la libertad y la autonomía, es también un problema que tiene que ver con el *conocimiento* político necesario para la democracia. Si mi perfil está almacenado a propósito por una IA y, a continuación, se me sirve desinformación para influir en mi intención de voto, se está destruyendo a nivel individual la base del conocimiento necesaria para ejercer la ciudadanía democrática, ya que carezco de la suficiente agencia epistémica: soy incapaz de controlar y de formar de manera suficiente mis creencias políticas^[1].

Pero existen más problemas con respecto a la erosión de la base epistémica necesaria para la democracia, especialmente para la democracia deliberativa y participativa de la tradición republicana. Considérense la *desinformación* y las *noticias falseadas*. Si ya no sabemos qué es cierto y qué no, qué es real y qué no, ¿cómo podemos discutir con otros ciudadanos con respecto al bien común bajo tales condiciones o, siquiera, votar? Los programas como ChatGPT permiten producir textos repletos de información falsa. Los *deepfakes* pueden ofrecernos un discurso político que nunca tuvo lugar. Incluso pueden utilizarse para convencernos de que algo real es falso (Meikle, 2023). Esto es muy peligroso, especialmente en un

contexto donde la confianza ya se estaba erosionando, como en el Reino Unido (ejemplo de Meikle), en oposición, pongamos, a los países nórdicos, en donde la confianza pública aún tendía a ser relativamente alta (y que se refleja en su estrategia para la IA; véase Robinson, 2020). Cuando ya no podemos distinguir la verdad de la falsedad, si se destruye la confianza entre los ciudadanos, entonces la democracia no funciona.

Así pues, la cuestión no es que exista un derecho individual a la verdad^[2], sino más bien que la *democracia* necesita un tipo particular de entorno epistémico. Los ideales más ricos de la democracia, como los de la democracia deliberativa y sus principios, tales como la idea de la Ilustración de que las opiniones deben ponerse sobre la mesa para que entre todos descubramos la verdad, no funcionarán, y mucho menos prosperarán, si el uso de la IA conduce a la desinformación y a contenidos falseados. Este uso de la IA debilita y, en última instancia, destruye la posibilidad de la democracia: lo que, como defendió Arendt, puede conducir al totalitarismo.

Además, como ya se ha sugerido anteriormente, las *burbujas epistémicas* son un problema para la fraternidad y dividen a la sociedad. Pero también crean el tipo de base de conocimiento errónea para la democracia, ya que la diversidad de opiniones se reduce y se excluye conocimiento relevante. Para el voto y deliberación democráticos, es vital que uno se exponga a un abanico de distintas opiniones y que tenga suficiente conocimiento político. Una burbuja epistémica o una cámara de eco disminuye la diversidad de opiniones, fragmenta la esfera pública y evita la formación adecuada de mis creencias y opiniones políticas (o, al menos, hace difícil su revisión). Imaginemos, por ejemplo, que algunas personas tienen opiniones racistas, pero que quieren discutir y potencialmente revisar esas opiniones a la luz de otras distintas y de la evidencia científica. Para eso se necesita que estén lo suficientemente expuestos a otras opiniones y a la ciencia. Pero esto

es imposible si continúan en una burbuja racista. Tanto si continúan en ella a propósito como si no (tanto si el efecto es intencionado como si no), el resultado es la persistencia de una base epistémica estrecha para el ejercicio de su agencia democrática. En otras palabras, tienen agencia epistémica insuficiente para ejercitar su ciudadanía democrática.

Además, volviendo de nuevo más allá del nivel individual, cuando las burbujas epistémicas y las cámaras de eco alientan a la gente a aferrarse a sus creencias sin someterlas a una discusión abierta, hacen algo que va en contra de los ideales de la democracia deliberativa y de otros relacionados con los anteriores, como las ideas de Mili y las republicanas acerca de la virtud epistémica. En lugar de un mercado de ideas, tenemos desacuerdos que ya no se pueden salvar. No hay voluntad de abrir a debate las propias opiniones y no se hace ningún esfuerzo por comprender las de los demás, si es que uno se expone siquiera a ellas. En lugar de una comunidad humanística de discusión abierta, tenemos una sociedad dividida e intolerante. En lugar de Ilustración, nos encontramos con hostilidad frente al debate y la ciencia. En lugar de comunicación, obtenemos violencia.

El desacuerdo no es necesariamente malo. No solo es inevitable, sino que, incluso, puede ser productivo, como muestran las versiones pluralistas y agonistas de la democracia. Mouffe (2013) ha criticado los modelos habermasianos de la democracia deliberativa por su enfoque exclusivo sobre el consenso, y ha defendido que el desacuerdo y la lucha también pueden conducir a un cambio político muy necesario. Sin embargo, los desacuerdos y las tensiones descritos aquí no parecen ser productivos, al menos no para la democracia. Si es que conducen siquiera hacia el cambio social, lo hacen hacia el tipo erróneo de cambio social. Cuando, amparadas por entornos de IA y de tecnologías relacionadas, las personas expresan simplemente sus opiniones sin estar preparadas para discutir las públicamente, la violencia y la dominación se encuentran

a la vuelta de la esquina, ya que no se da ningún intento por entender el punto de vista de los demás y existe la tentación de imponer la opinión propia sobre la suya o de no tomarlos en cuenta. En ese caso, el desacuerdo conduce a la intolerancia y (teniendo en mente también las ideas de Arendt y de otros republicanos) a la destrucción en última instancia del sentido común y de la esfera pública, quizás incluso al totalitarismo.

Aunque, empíricamente hablando, el alcance de estos fenómenos epistémicos puede variar (por ejemplo, las burbujas epistémicas no siempre son un problema para todo el mundo; es posible que sean más urgentes en algunos contextos mediáticos que en otros), esta discusión aporta argumentos sobre por qué son problemáticas con respecto a la democracia.

Tercero, en la medida en que la IA ofrece oportunidades para la manipulación y conduce a los problemas epistémicos mencionados, la IA lleva a la desconfianza: la desconfianza de los ciudadanos hacia el Gobierno, pero también la desconfianza hacia los demás, incluyendo la desconfianza hacia los expertos y el conocimiento científico.

Comenzaré con la desconfianza hacia el Gobierno. La democracia en su forma representativa se encuentra con problemas cuando los ciudadanos ya no tienen la sensación de que el Gobierno pueda rendir cuentas y responda a sus necesidades y problemas. Cuando el Gobierno usa la IA y la ciencia de datos de formas que no son transparentes (porque se esconde de la visión pública o porque el sistema de IA utilizado no es en esencia transparente y no ofrece explicaciones suficientes)^[3], entonces hay un problema de responsabilidad. Un gobierno democrático necesita que las decisiones puedan ser explicables y que sean escudriñadas en el Parlamento y más allá de él. Además, también necesita que se puedan adscribir responsabilidades. Pero no siempre está claro quién es responsable cuando se usan tecnologías como la IA para automatizar decisiones, o cuando los usuarios intentan evadir la

responsabilidad al decir que el problema es culpa de la máquina. Esto también se aplica al uso de la IA por los gobiernos y por las instituciones estatales. Imaginemos, por ejemplo, que una institución de bienestar social interrumpe, de repente, a un ciudadano sus prestaciones sociales diciéndole que la IA ha decidido interrumpirlas. ¿Quién es responsable de esta decisión y cómo podemos evitar el hecho de que nadie rinda cuentas? De este modo, la automatización de decisiones puede erosionar la base de la rendición de cuentas y de la responsabilidad, y llevar a la desconfianza en el Gobierno y, al final, incluso a la desconfianza en todo el sistema político en su conjunto: la democracia.

Nótese, sin embargo, que la desconfianza de los ciudadanos hacia el Gobierno se basa a menudo en el sentimiento de que el Estado no satisface sus intereses *individuales*. Sin embargo, lo que cuenta *políticamente*, de acuerdo con la tradición democrática republicana, no son simplemente los intereses individuales, sino el bien común. Necesitamos distinguir la desconfianza que proviene de intereses particulares de la desconfianza que está justificada por completo democráticamente: una falta de confianza en que el Gobierno se preocupe del bien común. La forma actualmente predominante de desconfianza es más bien un síntoma del fracaso de la democracia y de la celebración equivocada de la supremacía de los intereses individuales. Además, no es suficiente señalar a los políticos profesionales cuando se trata de resolver el problema. En la tradición republicana, recuperar la confianza no es solo el deber de los políticos, sino de todos los ciudadanos, pues el gobierno se supone que es una cuestión de participación y de decisiones comunes. *Somos* los políticos. Recuperar la confianza, cambiar nuestras instituciones políticas hacia una dirección más democrática, y crear y emplear la IA de forma democrática (por ejemplo, la IA que es transparente y explicable) es la responsabilidad de todos nosotros.

Pero el problema de la desconfianza también resuena más allá de las relaciones ciudadano-Estado. Si Foucault (1980) está en lo correcto en cuanto a que las relaciones de poder también ocurren más allá del nexo ciudadano-Estado, la desconfianza política es asimismo un problema para las relaciones entre ciudadanos. Y estas relaciones también se verán afectadas por la IA y la ciencia de datos. Los ciudadanos pueden utilizar la IA de manera que se dañen las relaciones de confianza, por ejemplo, categorizando y manipulando a terceros a través de aplicaciones de redes sociales. Pero, a menudo, la erosión de la confianza no es intencionada, por ejemplo, cuando una burbuja epistémica creada sin intención alguna por los ciudadanos aumenta la desconfianza hacia otros que están fuera de la burbuja. También podría ocurrir que las relaciones de confianza entre los ciudadanos ya estuvieran dañadas por otras razones y que el uso de la IA las erosione aún más. Toda esta desconfianza en los demás y en la sociedad alimenta tendencias antidemocráticas, conduciendo potencialmente al autoritarismo y el totalitarismo, como defendía Arendt. Pero, incluso sin conducir de forma inmediata al autoritarismo, la desconfianza entre los ciudadanos, apoyada por la IA, es improbable que fomente las relaciones comunicativas, comprensivas, respetuosas y tolerantes entre ciudadanos que se necesitan para los ideales más ricos de democracia defendidos por Arendt, la tradición republicana y los pensadores ilustrados.

Finalmente, relacionado con la cuestión respecto a la agencia epistémica está el problema de que la IA y la ciencia de datos pueden erosionar la autoimagen y el autoconocimiento de los sujetos políticos. ¿Podemos continuar comprendiéndonos como sujetos autónomos a la luz de la IA y otras tecnologías digitales? Este es un problema para los ideales relevantes de la democracia. La suposición ilustrada era que somos sujetos autónomos que razonan unos con otros con respecto a la política en el espacio público. Pero si la IA me conoce mejor de lo que yo mismo me conozco (o, al

menos, esta es la afirmación de la industria tecnológica) porque rastrea mi comportamiento y categoriza e influye en mis creencias sin que yo tenga conocimiento de ello, entonces este tipo de práctica política parece estar desvirtuada. ¿Por qué molestarse con procedimientos democráticos y discusiones comunicativas si las opiniones y creencias políticas de los ciudadanos están influidas «bajo el radar» de su razón y control ilustrados, algo tan importante en las tradiciones filosóficas habermasiana, rawlsiana y kantiana? Y si las emociones y las «voces» son relevantes políticamente o, incluso, como han defendido Mouffe y otros pensadores importantes en política, ¿qué les ocurre entonces a la política y a la democracia si estas emociones y voces están fuertemente influenciadas por manipulaciones políticas que usan la IA y la ciencia de datos?

Siendo honesto, el problema no es completamente nuevo. La democracia moderna siempre se ha enfrentado al problema de la influencia social: cuando me decido a ser un ciudadano de una democracia, ya estoy siendo fuertemente influenciado por otros (intencionadamente o no): ¿qué dice eso de la calidad de mi voto, mi opinión y mi voz? Y, ciertamente, ¿qué dice sobre la posibilidad de la democracia? Estando hoy en día la influencia social mediada por tecnologías digitales, incluyendo la LA, no se vuelve más fácil lidiar con este problema. La cuestión no es que las cosas fueran perfectas antes de la IA, sino que esta las empeora, las amplifica. El Gran Hermano, que ya sabía de mí más de lo que me convenía, usa ahora la IA. O, mejor dicho, lo hace al menos en los países que se autodenominan democráticos: el Gran Hermano (Zuboff, 2015) me conoce mejor a través de la IA. Clasificado y manipulado por las grandes corporaciones tecnológicas y (en principio) por cualquiera que pueda usar aplicaciones que integren inteligencia artificial, el sujeto político que la Ilustración buscaba establecer parece ciertamente lejano.

Sin duda, la imagen de la Ilustración también era problemática y, desde luego, no era indiscutible. Pero, a pesar de que los

pensadores posmodernos y feministas han estado acertados al criticar la imagen de un sujeto completamente autónomo que niega su propia relacionalidad, y a pesar de que es probable que el autoconocimiento cambie en formas complejas bajo la influencia de la IA y de otras tecnologías digitales (algunas de las cuales pueden ser beneficiosas), normativamente hablando parece necesario para la democracia un grado mínimo de agencia epistémica y autonomía. Por lo tanto, es obligatorio crear entornos epistémicos que produzcan sujetos políticos con un grado de agencia y autonomía epistémicas que sea suficiente y adecuado para los ideales democráticos mencionados.

Como señalé antes, reconocer estos peligros no implica afirmar que la IA ocasione siempre y necesariamente estos efectos relacionados con el conocimiento. Pero mis argumentos ofrecen una forma de enmarcar las preocupaciones y explicar qué hay exactamente de problemático en estos fenómenos cuando se trata de *democracia*. Incluso si la IA tiene dichos efectos, no hay motivo para sentirse desesperado e indefenso con respecto a su impacto sobre la democracia: podemos hacer algo con estos problemas. Podemos hacer que la democracia sea más resiliente y podemos hacer una IA mejor, más democrática.

Capítulo 6

Fortalecer la democracia y democratizar la IA

Teniendo en cuenta los problemas indicados en los capítulos previos, ¿cómo podemos solucionarlos? ¿Qué podemos hacer para conseguir que la democracia sea más resiliente frente a la IA y a la revolución digital? Y, si no tomamos la IA como dada, sino que nos centramos definitivamente en su desarrollo, ¿podemos hacer que la IA sea más democrática? ¿Cuáles son los pasos para conseguirlo?

FORTALECER Y CREAR NUEVAS INSTITUCIONES DEMOCRÁTICAS

Un primer paso para hacer que la democracia sea más resiliente a la luz de la IA es algo que tiene poco que ver con la IA y más con las debilidades actuales de nuestras instituciones democráticas. Necesitamos fortalecer la democracia. La democracia es una gran idea, pero aún necesita que se plasme por completo. En particular, si nos distanciamos de la concepción mínima y débil de la democracia como procedimiento de voto y abrazamos un ideal más rico (inspirado en los enfoques republicano, ilustrado y deliberativo) que incluya la participación, la deliberación pública y la comunicación

social, tendremos la oportunidad de mitigar el impacto potencialmente negativo de la IA y de otras tecnologías. A pesar de que la IA es política de principio a fin (Coeckelbergh, 2022a), es improbable que, por sí misma, su uso destruya la democracia, a menos que esta se encuentre ya en serios problemas. Si fortalecemos y desarrollamos más nuestras instituciones democráticas, disminuimos los riesgos discutidos hasta ahora.

Un instrumento importante para preservar la democracia es la ley. Como Susskind (2022) ha argumentado en defensa de lo que llama «republicanismo digital», la ley debe preservar las instituciones democráticas, y frenar «el poder irresponsable de aquellos que diseñan y controlan la tecnología digital» y su individualismo de mercado, que cree equivocadamente que todo puede quedar en manos de los individuos y de sus empresas con sus propios intereses (Susskind, 2022: 11). La Constitución necesita desempeñar un papel, junto con principios universales, como los derechos humanos y los procedimientos y lugares para solucionar disputas (tradicionalmente, los juzgados). Por ejemplo, Nemitz (2018), en sus argumentaciones sobre la regulación de la IA en el contexto de la UE, conecta la democracia con el imperio de la ley y los derechos humanos. Y en las democracias liberales tenemos parlamentos que se supone que representan al pueblo (en lugar de a las grandes corporaciones tecnológicas) y que pueden promulgar leyes que limiten el poder del sector tecnológico. Pronto profundizaré en este asunto de la regulación de la IA.

Pero, además de conversar lo que tenemos ya, necesitamos construir algo nuevo. Necesitamos que cambien las instituciones políticas. Necesitamos transformar nuestras instituciones políticas en una dirección más participativa y comunicativa. Dado el actual desempeño insatisfactorio de nuestras instituciones políticas, no solo necesitamos proteger la democracia, también necesitamos crearla. En este aspecto, pueden ayudarnos ciertas ideas recientes de la teoría política. Por ejemplo, Landemore (2020) ha propuesto lo que

él llama un «minipúblico abierto»: un jurado de ciudadanos elegidos aleatoriamente que deliberan y crean las leyes. Esta no es la idea habitual de la democracia directa: preserva la representación, ya que continúa siendo una selección. Al mismo tiempo, los ciudadanos ordinarios están involucrados en el proceso de toma de decisiones, en lugar de las elites políticas o los millonarios de las tecnológicas. La IA también puede ayudar aquí (véase págs. 97-101). Sin embargo, no está claro cómo funcionaría un modelo de este tipo, pues en la actualidad se necesita bastante especialización, en particular con respecto a tecnologías digitales como la LA. También puede criticarse el elemento de aleatoriedad que hay aquí: ¿es aceptable lo que Guerrero (2014) llama «lotocracia»? ¿Y cómo de resistente contra el populismo y el extremismo es este modelo? ¿Qué otras intervenciones podrían ser (¿más?) productivas para una democracia basada en el conocimiento informado? Necesitamos reflexionar mejor acerca de qué equilibrios y negociaciones entre la democracia y la tecnocracia son necesarios en el siglo XXI.

Con respecto a la Constitución, necesitamos transformar nuestro orden constitucional de forma que otorgue poder suficiente y efectivo a los cuerpos elegidos democráticamente, en oposición a los poderes judiciales y a las grandes tecnológicas, a la vez que nos aseguramos de que haya suficientes aportaciones de los expertos. Y, si queremos construir una concepción de la democracia más republicana, se necesita también integrar las libertades individuales modernas orientándolas al bien común. A pesar de que esto plantea ciertas cuestiones con respecto a cómo definir el bien común^[1] (véase capítulo 5) y cómo equilibrar esta orientación normativa hacia el bien común con los derechos y libertades individuales, tal orientación puede proporcionar una base normativa y legal más robusta para nuestras democracias, haciéndolas más resilientes frente a cómo puedan afectarles la IA y otras tecnologías digitales. Además, necesitaremos también procedimientos mejores para resolver los conflictos entre ciudadanos, incluyendo aquí los

conflictos que tengan que ver con la IA y otras tecnologías digitales[2].

Otra institución democrática importante que debería mencionarse en este contexto es el periodismo de calidad. El periodismo, a menudo llamado cuarto poder, tiene la importante tarea de monitorizar y evaluar críticamente al Gobierno, a la vez que informar y educar al público. A la luz de tecnologías digitales como la IA y fenómenos relacionados, como las noticias falseadas, esto es ahora más necesario que nunca. Los periodistas ya tenían que comprobar los hechos y facilitar la difusión de información verdadera antes de que la IA se convirtiera en el centro de atención, pero en el actual contexto de desinformación promovido por el uso de la IA, este papel se vuelve más importante (Bimber y De Zúñiga, 2020). Más allá de quienes desempeñan este papel crítico en los medios de comunicación clásicos, como los periódicos y la televisión, necesitamos más «editores» y «moderadores» de discusiones en la esfera pública, al igual que hay mediadores en las redes sociales digitales: necesitamos personas que puedan ayudar a los ciudadanos con sus desafíos político-epistémicos. En este caso, la IA no es una amenaza, sino una ayuda, por ejemplo, presentando automáticamente artículos que perturben la fuente homogénea de noticias de un lector (Lin y Lewis, 2022) y desempeñando el papel de guardianes (Duberry, 2022) para proteger la democracia contra formas de populismo antidemocrático y el auge de diversos modos de autoritarismo. Como sabemos por el caso histórico de Hitler, pero también más recientemente por el de Trump, los líderes elegidos democráticamente pueden, una vez instalados en el poder, subvertir los procesos democráticos (Levitsky y Ziblatt, 2018) y (en el caso de Hitler) establecer incluso autocracias en toda regla. Los algoritmos de IA pueden utilizarse para mitigar el papel facilitador que desempeñan los medios digitales a fin de otorgar a las personas una voz y contribuir a su éxito. De forma más general, la IA y otras tecnologías digitales pueden ayudar a mejorar la calidad de la esfera

pública, teniendo en mente, por supuesto, que la libertad de expresión es también un valor importante en una democracia.

Esto no quiere decir que los algoritmos de IA deban ser los únicos guardianes y editores: los humanos son muy necesarios. Pero *qué* humanos es algo importante: la comprobación de los hechos y las decisiones relativos a los contenidos no deberían quedar en manos de actores no democráticos, como los CEO y los empleados de Facebook y Twitter. Existe un papel importante para el periodismo independiente, editores no corporativos, mediadores y medios de comunicación de propiedad pública. Los gobiernos deberían ofrecer incentivos y apoyar la educación y formación de tales editores y mediadores.

REGULACIÓN Y SUPERVISIÓN

Esto nos conduce a una segunda manera, relacionada con lo anterior, de abordar los desafíos concernientes a la IA y a la democracia, que también intenta reforzar el control de las instituciones políticas democráticas sobre el desarrollo y el uso de la tecnología: la regulación y la supervisión. Si actualmente existe un desequilibrio de poder entre las grandes corporaciones tecnológicas y las entidades políticas elegidas de forma democrática, no puede decirse que esta sea una situación inevitable o «natural». La creencia en el determinismo tecnológico (la opinión de que la tecnología da forma a la sociedad de acuerdo con su lógica interna y que, por consiguiente, determina nuestro futuro con solo un papel mínimo para los humanos) va a menudo de la mano con el conservadurismo político y el *laissez-faire*. Según señala Brevini, la fe en la tecnología se usa frecuentemente como «una poderosa apología del *status quo* y a favor de la estructura actual del capitalismo, sin dejar ningún espacio real para la crítica» (2022: 21). Pero si, por el contrario, dirigimos nuestra mirada más allá del

determinismo y el tecnosolucionismo, que cree erróneamente que existe una solución tecnológica para todos nuestros problemas sociales y políticos, podemos encontrar tal espacio para la crítica y, por supuesto, para la acción. Si hay en la actualidad un desequilibrio de poder, nosotros, como ciudadanos, podemos criticar, oponer resistencia y recuperar el poder. Mientras que, sin duda, el sector privado tiene un papel que desempeñar mediante la autorregulación y, debido a sus conocimientos y su posición actual en el panorama tecnológico y de los medios de comunicación, debe continuar siendo un interlocutor y actor cuando se trata de políticas relacionadas con la ética y la política de la IA de los próximos años. El futuro de nuestras sociedades y de las democracias es demasiado importante como para dejarlo por completo en las manos de organizaciones privadas y no democráticas. Su posición en términos de conocimiento y de poder no es ni natural ni inevitable. Para usar el mismo ejemplo que antes: dados su influencia y su éxito, plataformas como Facebook y Twitter, y los millonarios tecnológicos, como Mark Zuckerberg y Elon Musk, deciden actualmente qué información y qué opiniones acceden a la esfera pública. En una democracia, esto es inaceptable.

La mayoría de estos problemas son bien conocidos. Ya disponemos de análisis de los desafíos éticos y políticos y de las oportunidades de la IA. Se ha hecho mucho trabajo, tanto dentro como fuera de la academia. En la actualidad, se reconoce cada vez más por todas las partes implicadas que necesitamos políticas robustas y regulación efectiva que ayude a los ciudadanos, a las compañías, a los gobiernos y a otros actores sociales a lidiar con los riesgos potenciales y a aprovechar las posibilidades de la tecnología de una forma ética y políticamente sostenible. Esto no es importante solo desde un punto de vista legal para asegurar la responsabilidad legal (véase, por ejemplo, Katyal, 2022), sino también desde una perspectiva ética y política. En el momento en que escribo, hay varias propuestas nacionales para la regulación, como el *Blueprint*

for an AI Bill of Right^[3] (proyecto para una carta de derechos de la IA) de la administración Biden de la Casa Blanca. Y, siguiendo la propuesta de la Comisión Europea a favor de un marco regulatorio de la Unión Europea para la IA^[4], que se basa en parte en las recomendaciones de la comisión del grupo de expertos de alto nivel en IA, se ha aprobado en el Parlamento Europeo el Acta de la IA^[5]. Se dice que es «la primera ley de inteligencia artificial completa del mundo^[6]». Regula la IA de acuerdo con categorías de riesgo y también propone la supervisión.

Idealmente, la regulación y la ética van de la mano. No deberíamos tener que escoger entre ellas. Las regulaciones actuales de la IA y las propuestas de regulación de la IA están basadas tanto en principios éticos como legales. Por ejemplo, la UE dice que su regulación está basada en los valores europeos y en los derechos fundamentales, y las recomendaciones de la UNESCO tienen una lista de diferentes valores y principios, tales como el respeto de los derechos y de la dignidad humanos, las libertades fundamentales, el florecimiento del medioambiente y del ecosistema, la diversidad y la inclusividad, el no causar daño, la justicia y la no discriminación, la transparencia y la rendición de cuentas^[7]. A pesar de que las discusiones en torno a la política de la IA han presenciado diferentes opiniones acerca de qué principios éticos se deberían usar (por ejemplo, si los principios bioéticos deberían ser la base, como han propuesto Floridi y Cowls [2019]) y, a pesar de que los principios de alto nivel son insuficientes por sí solos para garantizar una IA ética (véase, por ejemplo, Mittelstadt, 2019), todos están de acuerdo en que la ética debería ser la base de la regulación de la IA. Sin embargo, hay que tener en cuenta que la ética se usa a veces para justificar una política del *laissez-faire* y la autorregulación. En ese caso, la ética se convierte en «lavado ético». Pero evitar la legislación de la IA es antidemocrático (Nemitz, 2018). Nos deja en una situación en la que los principios y valores clave de la

democracia continúan amenazados y en donde las reglas vienen hechas por quienes carecen de la legitimidad democrática para ello.

Además, es importante que el *proceso* regulatorio esté bien integrado en las instituciones democráticas mencionadas para evitar que sea secuestrado por intereses y poderes particulares. Si las grandes corporaciones tecnológicas no evitan la regulación, al menos intentan influenciarla profundamente. Ahora bien, constituir grupos de presión es también una forma de incorporar conocimiento experto, y los actores del sector privado son algunas de las partes interesadas relevantes. Pero en pro de la democracia y del bien común, esa influencia e implicación deben verse limitadas. Demasiada poca regulación y demasiado espacio para la influencia corporativa en las políticas tecnológicas conducen a la tiranía de los más fuertes. Este es hoy día un peligro inminente, cuando con frecuencia la IA se encuentra poco regulada en muchas partes del mundo y las grandes empresas tecnológicas intentan influir sustancialmente en la regulación de la IA, como sucede, por ejemplo, en Europa y en los Estados Unidos.

Por tanto, el problema no es cuál y cuánta regulación necesitamos. Como propuse en la sección anterior, también necesitamos reflexionar críticamente acerca de nuestras instituciones que, aparentemente, a duras penas son lo suficientemente resilientes frente a los desarrollos actuales. La democracia necesita limitar el poder. A la luz del poder de las grandes corporaciones tecnológicas, necesitamos revisar la forma en que el poder se distribuye actualmente en la sociedad. Necesitamos comprobaciones, equilibrios y nuevas instituciones. Por ejemplo, se podría considerar el establecimiento de una institución que conecte mejor los contextos de desarrollo de la tecnología con organismos elegidos democráticamente, y lo haga de formas en que redistribuya el poder hacia esos organismos elegidos (alejados de las grandes tecnológicas), mientras, al mismo tiempo, asegure un papel significativo para los expertos y una comunicación fluida con ellos.

Se podría también argumentar que algunas cosas no deberían dejarse en manos del mercado: que algunos datos, algunos bienes y servicios, algunas infraestructuras, y el papel de los guardianes o moderadores son demasiado importantes para dejarlos en manos de la iniciativa privada. Por ejemplo, una opción es convertir algunas de las grandes empresas tecnológicas en un servicio público, como ha propuesto Risse (2023: 71).

Esto nos conduce a la cuestión de hasta qué punto deberíamos permitir que la IA y otras tecnologías digitales se entremezclen con el capitalismo. ¿Debería democratizarse la tecnología (y sus ganancias y beneficios), en el sentido de que fueran de propiedad compartida o pública? Una razón por la que la regulación no está teniendo lugar actualmente, o se mantiene al mínimo, es la persistencia de la idea de que el desarrollo de la tecnología debería ser principalmente, si no exclusivamente, una cuestión de la iniciativa privada, de la propiedad privada y del capital privado. Pero no existe ninguna razón por la que (¿la versión actual del?) capitalismo debiera darse por sentada o vista como algo natural. Según muestran pensadores marxistas como Dyer-Witheford, Kjosén y Steinhoff (2019) y Fuchs, existen alternativas. Por ejemplo, Fuchs (2023) defiende una versión socialista de la democracia que define como la que «hace avanzar la economía común, el bien político y cultural para todos los humanos» (2023: 13), y que debería estar apoyada en las tecnologías digitales. Su versión socialista de lo que denomina «democracia digital» (*ibíd.*: 12) implica, entre otras cosas, que se ha de ofrecer más apoyo a los medios de comunicación públicos independientes (*ibíd.*: 297): esto ayudaría a los periodistas a desempeñar su papel de apoyo a la democracia mencionado antes. En este contexto, merece la pena señalar que, actualmente, no solo las redes sociales digitales, sino también los imperios privados de medios de comunicación tradicionales, como Fox News TV y la prensa sensacionalista británica, debilitan la democracia cuando

crean más polarización y destruyen la confianza en los medios y en la democracia.

Otra forma de reconsiderar el equilibrio entre la propiedad privada y pública de (los medios para) el desarrollo de la tecnología digital es otorgar más apoyo a la creación y mantenimiento del denominado dominio público digital. La idea es que las tecnologías y los recursos digitales, como también los datos, deberían ser de propiedad común y compartida. Piénsese, por ejemplo, en la Wikipedia o el *software* de código abierto que se utilizan libremente y con los que cualquiera puede colaborar.

Entre la bibliografía de la teoría crítica que promueve una mayor igualdad en la economía de datos existen al menos dos ideas diferentes sobre el tipo de propiedad deberían constituir los datos y cómo distribuirlos en una democracia. Una de ellas permanece dentro de la esfera liberal-individualista y promueve el empoderamiento de los ciudadanos individuales. Fischli (2022) ha abogado por una denominada «democracia propietaria de datos», una economía política en la que la propiedad de los datos (entendidos como capital) se distribuye entre los ciudadanos, los cuales tendrían así la posibilidad de participar en las decisiones con respecto a su utilización. Una idea diferente es el socialismo digital (Muldoorn, 2022), que implicaría una propiedad en común de datos, la participación de trabajadores y comunidades, y un control democrático de la inversión, con lo que se lograría una redistribución más radical del poder que va más allá del enfoque individualista.

Considerando las cosas desde una perspectiva global, se plantea la cuestión de si bastaría con un único sistema político-económico. Dada la diversidad de sistemas que hay en el mundo y de sus contextos y culturas políticas, necesitamos conectar la cuestión concerniente a la IA y la democracia con la reflexión con respecto a qué es lo que esta significa en relación con los sistemas y contextos específicos. Puede haber muchas versiones diferentes para hacer la IA más «común». De forma similar, deberíamos preguntarnos qué

significan la democracia y el uso de la IA a la luz de los desafíos ecológicos y climáticos, y si el pensamiento capitalista y de libre mercado es capaz de lidiar como es debido con estos desafíos. La teoría crítica (concebida ampliamente), combinada con la teoría medioambiental, puede ayudar a que se conciban sistemas alternativos. Pero, incluso desde una perspectiva de mercado, cabe defender la ruptura de los monopolios y oligopolios globales para preservar la competencia y mantener un campo de juego equilibrado. Las leyes antimonopolio pueden ser un instrumento para romper el oligopolio actual de un pequeño número de grandes empresas tecnológicas. La regulación medioambiental también sería aplicable dentro de un marco liberal.

Finalmente, dada la naturaleza global de estas tecnologías y los problemas tecnológicos y ecológicos, los esfuerzos regulatorios solo tienen sentido si existe algún tipo de armonización y coordinación a nivel *supranational* y global. La IA y sus consecuencias éticas y políticas no están limitadas a un país o a una región. Los datos y el *software* cruzan fronteras con facilidad, y lo mismo hacen los efectos medioambientales. Si el problema es global, también necesitamos soluciones a nivel global, además de a otros niveles^[8].

En el área de política de la IA, varias organizaciones internacionales han tomado iniciativas recientemente. Por ejemplo, en 2021 los Estados miembros de la UNESCO adoptaron un acuerdo global con respecto a la ética de la IA^[9], y el Consejo de Europa está trabajando en una política de la IA basada en los derechos humanos, el imperio de la ley y la democracia^[10]. Ya he mencionado la formulación de políticas por parte de la UE en esta área. Pero la UE tiene un aspecto *supranational*. Continúa siendo cuestionable hasta qué punto las organizaciones *internacionales* son eficaces en este campo, frente a lo que podría lograrse con un marco de gobierno realmente *supranational* para promover y dirigir el uso y desarrollo democrático de la IA. Como sucede con muchas otras cuestiones globales urgentes, tales que el cambio climático, la

pandemia y la guerra, cuando se trata del gobierno global de la IA aún nos encontramos en una especie de estado de naturaleza hobbesiana: no existe una autoridad de gobierno global, y el nacionalismo amenaza la necesaria y urgente obligación de coordinarse mejor. Dicho esto, la gobernanza global tiene sus propios desafíos, tales como negociar con diferentes culturas y tradiciones políticas y, por supuesto, asegurar y mantener el carácter democrático del gobierno a tal nivel. Por ejemplo, ¿cómo reconocer la diversidad en los enfoques asegurando, a la vez, la efectividad?

La solución a la ausencia de una autoridad política global no debería ser un leviatán hobbesiano, esto es, un Gobierno autoritario global, sino una democracia. Sin embargo, no está claro lo que esto implica a nivel global. Quizás necesitemos una república universal, como propuso Cicerón, que sea tal que reconozca diferentes niveles de poder. Pero ¿cómo aseguramos de que están protegidas las minorías y la diversidad cultural? ¿Y cómo llevarlo a la práctica? ¿Pueden integrarse las no democracias existentes en una estructura de gobernanza global democrática? ¿Deberían intervenir y ser forzadas a ser más democráticas, o existen otras vías? ¿Cómo asegurar la no dominación y evitar el imperio y el colonialismo? ¿Sería preferible una estructura federal? Se debe trabajar mucho más en este tema. Por desgracia, muchos filósofos políticos dan por sentados los límites de los Estados nación y de sus ciudadanos cuando se ponen a hablar sobre tecnología y democracia; y, ciertamente, cuando se trata de definir los límites de lo «común» en el «bien común» (un concepto que examinaré en mayor profundidad en los siguientes capítulos).

IA DEMOCRÁTICA

Hasta ahora, he propuesto un cambio regulativo e institucional para asegurar una política democrática de la IA. Sin embargo, un

tercer enfoque cruza la división entre la tecnología y la política. En lugar de considerar a la tecnología como algo dado y esperar hasta que ocurra algún contratiempo importante para, a continuación, medir el daño e intentar regular y resolver los problemas posteriormente, este enfoque proactivo y preventivo aborda el propio desarrollo tecnológico, y lo hace más ético y más políticamente responsable. El objetivo es intentar crear una IA que sea más democrática y esté más a prueba de totalitarismo.

A este propósito se pueden utilizar los métodos existentes para la ética de la tecnología y la innovación responsable, pero poniendo énfasis en la política: el objetivo reformulado de estos métodos es, por lo tanto, asegurarse de que no solo los valores y preocupaciones éticas, sino también las *políticas*, incluyendo a la democracia, se introducen en el proceso del desarrollo de la tecnología. Considérese, por ejemplo, el «diseño sensible a los valores» (Friedman y Hendry, 2019; Van Den Hoven, 2013), que propone métodos para tener en cuenta los valores humanos en el diseño de nuevas tecnologías; y el concepto de «investigación e innovación responsable» (Owen, Macnaghten y Stilgoe, 2012; Von Schomberg, 2013), que también busca tener en cuenta la ética al involucrar, entre otras cosas, a las partes interesadas en los procesos de investigación e innovación. De forma similar, y enteramente compatible con el espíritu de estos métodos y concepciones, propongo que, para el desarrollo *políticamente* responsable de la IA, se tengan en cuenta los valores y preocupaciones *políticas*, y que las partes interesadas que son políticamente relevantes estén implicadas en su desarrollo. Esta propuesta también está en Enea con el enfoque adoptado en un informe reciente del Grupo Europeo de Ética de la Ciencia y las Nuevas Tecnologías sobre la democracia en la era digital, que recomienda a la Comisión Europea la participación pública y el diseño sensible a los valores para proteger los valores fundamentales (2023: 49).

Este enfoque proactivo implica que los expertos en IA que están implicados en la elaboración de políticas y en la toma de decisiones con respecto a aquella tengan responsabilidad con perspectivas de futuro en la IA, así como en sus consecuencias éticas y políticas involuntarias. Como mínimo, esto implica que esos expertos deben estar informados por actores políticos y sociales (Hedlund y Persson, 2022). Pero aún mejor sería la *implicación* política y democrática en el nivel y el estadio de desarrollo de la tecnología.

La justificación que propongo es similar a los argumentos existentes en la ética de la tecnología. En la ética de la tecnología ya se ha defendido que la tecnología nunca es neutral y que posee una dimensión *moral* (Verbeek, 2011 y antes Latour, 2002), y que, por lo tanto, sería mejor que abriéramos y diésemos forma a los procesos de su desarrollo moral-tecnológico. Además, también se podría reconocer de forma similar, y en línea con el trabajo seminal de Winner (1986) y Feenberg en filosofía de la tecnología (por ejemplo, Feenberg y Hannay, 1995) y la obra de Latour (2004) sobre ciencia política, que la tecnología también tiene una dimensión inherentemente *política* y que, por tanto, sería mejor que diéramos forma de manera abierta y democrática a su desarrollo político-tecnológico. Si la tecnología está vinculada a acuerdos políticos y sociales y a ideologías y culturas políticas, si su desarrollo no solo se construye técnica sino socialmente, y si tiene consecuencias políticas, entonces el cambio real solo es posible si modificamos también las tecnologías, no solo las instituciones. Si el poder también proviene de la tecnología y no solo de lo que hacen personas en los parlamentos y otras instituciones democráticas, entonces una democratización completa del proyecto de la IA requiere que modifiquemos también la tecnología. Si la tecnología es «política proseguida por otros medios», como Latour (1983: 168) dice sobre la ciencia (a su vez una parodia del famoso aforismo de Clausewitz), entonces mejor sería que diéramos forma a esa política y a esa tecnología. Si realmente nos importan nuestras democracias

y queremos que sean más resilientes frente a los desafíos impuestos por la IA y otras tecnologías digitales, no solo deberíamos cambiar y fortalecer los procesos e instituciones y su control democrático de la tecnología, sino también cambiar la IA: deberíamos hacer que esta fuera más democrática. Creemos una IA a prueba de totalitarismo y una *IA democrática*.

En particular, propongo un *desarrollo democrático de la IA* (DDI) y una *democracia por diseño*^[11] destinada a apoyar lo que podríamos denominar una *tecnodemocracia*: no una tecnocracia, sino una democracia que continua y sistemáticamente reflexione acerca de sí misma a la vez que se transformen ella y sus tecnologías basándose en la conciencia de que las tecnologías como la IA son siempre políticas y que, por lo tanto, busque desarrollar tecnologías democráticas. Necesitamos *software*, procesos de innovación y regulaciones que conecten el desarrollo de la tecnología de formas más directas y efectivas con las discusiones con respecto a la política y la democracia y, preferiblemente, con procesos de toma de decisiones democráticos, procedimientos democráticos e instituciones democráticas.

A pesar de que implementar esta idea exige arrebatar algo de libertad a las empresas y a otros actores del sector privado en la sociedad que desarrolla la IA (al igual que hace la regulación *ex post*), al mismo tiempo los incluye desde el comienzo en los procesos como partes interesadas y actores clave. Además, también *apoya* a estos actores, ya que los *ayuda* a desarrollar tecnología políticamente responsable y democrática. Al apoyarles, ofrece más certeza de que están haciendo lo correcto y aumenta o mantiene la sostenibilidad ética, política y, en última instancia, también económica de sus prácticas y negocios.

Téngase en cuenta, sin embargo, que mi argumento para hacer que el desarrollo de la IA sea más «político» no implica la afirmación de que el desarrollo de la tecnología deba ser *politizado* en el mal sentido de la palabra: que implique la pérdida *total* de

independencia por parte de las empresas y que simplemente sustituya una forma de poder centralizado por otra. La cuestión no es, pongamos por caso, que Biden sustituya a Musk o Altman. Pero implica, al menos, que los desarrolladores y sus empresas deban tener en cuenta procesos democráticos y más inclusivos con respecto a la regulación de la tecnología cuando desarrollen IA e, idealmente, que la propia *conformación* de la tecnología implique también a instituciones y procesos democráticos, aunque sea de forma indirecta. Por ejemplo, la creación de estándares técnicos y de procedimientos de certificación que busquen hacer la IA más responsable ética y políticamente no debería dejar solos a los expertos, sino que estos necesitan estar relacionados con procesos e instituciones democráticos. Esto no quiere decir que los políticos, que no saben nada de tecnología, vayan a crear estándares y hagan las certificaciones. Los expertos son necesarios. Pero en el desarrollo y el debate en torno a los estándares tiene que haber *implicados* de forma sistemática y obligatoria ciertos organismos elegidos democráticamente y otras instituciones democráticas. El desarrollo de la IA no necesita tanto estar «politizado» como el ser políticamente *relational*, esto es, estar conectado democrática y comunicativamente con una sociedad más amplia.

Pero ¿cómo se articula entonces? ¿Y es suficiente? Teniendo en mente los puntos anteriores dedicados a las instituciones y la regulación, debemos ser más ambiciosos y más precisos. Si realmente queremos una IA más democrática en el sentido aquí esbozado, tenemos que pensar cómo conectar precisamente los procesos de investigación e innovación con las instituciones políticas. Dado que actualmente el desarrollo de la IA tiene lugar fuera del control democrático, necesitamos integrarlo adecuada y rigurosamente *en* la democracia. En lugar de deliberaciones *ad hoc* sobre nuevas tecnologías como la IA y otras más específicas (como es ahora el caso), y en lugar de depender de la buena voluntad de los líderes tecnológicos para incluir las opiniones de las partes

interesadas, necesitamos crear soluciones institucionales y más permanentes que puedan mediar entre la política y el desarrollo de la tecnología, y entre los expertos y los ciudadanos. Por ejemplo, una nueva institución permanente podría permitir a los ciudadanos como usuarios de IA y plataformas de redes sociales el hablar directamente o a través de mediadores con los expertos que desarrolla(ro)n estas plataformas y la IA relacionada. Los resultados de tales mediaciones y discusiones deberían retornar de forma regular al proceso de desarrollo.

A pesar de que esta idea necesita un desarrollo mayor, está claro que la situación actual es inaceptable desde un punto de vista democrático: necesitamos puentes efectivos e institucionalizados entre el mundo político (democrático) y el desarrollo tecnológico. Con Latour (2004), podemos decir que necesitamos embajadores que medien y viajen entre esos reinos. Pero lo que hacen no debería ser una cuestión de ayuda ocasional (por ejemplo, los comités actuales *ad hoc* y consejos asesores temporales), o continuar sin consecuencias vinculantes de ninguna clase: su papel debería estar institucionalizado e integrado en el marco político-legal. Si fuera necesario, deberíamos cambiar nuestras constituciones para permitirlo.

Además, no debemos dar por supuesto el actual estado epistémico y político de los ciudadanos. Es cierto que existe mucha ignorancia y, a veces, incluso antiintelectualismo como sucede, por ejemplo, en los Estados Unidos^[12]. La queja de Platón con respecto a las masas está lejos de ser irrelevante. Pero, de nuevo, esta situación no debería tomarse como algo dado o natural, y el problema no es completamente desesperanzador o inevitable. Una clásica receta republicana e ilustrada para el empoderamiento de la ciudadanía es la educación. La democracia, especialmente en sus formulaciones más ricas, bien puede *acabar* con el gobierno del pueblo, pero comienza con la educación.

En los capítulos anteriores he subrayado que la democracia necesita una base de conocimiento. La educación debería ayudarnos a construir esta base. Dewey, por ejemplo, pensaba que la educación es vital para la democracia. Es necesaria para votar, pero también para la democracia como una forma de vida y *una forma de convivencia*. Dewey diría que necesitamos educación para la cooperación y la resolución de problemas en sociedad. Teniendo en mente a Arendt y a Habermas, podríamos defender que es necesario para desarrollar sentido y comunicación comunes basados en el entendimiento y la deliberación también comunes. El conocimiento experto es una cuestión diferente, pero se podría defender que los ciudadanos necesitan, al menos, conocimientos y capacidades básicas para cumplir su papel en la democracia. Como mínimo, esto implica que todos los ciudadanos tengan educación democrática y que estén formados en el pensamiento crítico, en la deliberación con los demás, y que aprendan las capacidades necesarias para escucharse unos a otros y desarrollar un entendimiento común. Pero en el siglo *xxi* también se debería incluir la adquisición de un conocimiento básico de *tecnologías* como la IA y su papel en la sociedad, y el desarrollo de capacidades críticas y democrático-comunicativas en el contexto de las redes sociales digitales. Necesitamos formación para entendernos mejor unos a otros (pensando en Arendt: para imaginar el punto de vista de los demás) y aprender a comunicarnos mejor, no solo en general, sino también específicamente en el contexto de estos medios de comunicación y tecnologías que se han vuelto tan omnipresentes en nuestras vidas y en nuestras sociedades.

Además, por lo que respecta a la educación superior, aquellos que están aprendiendo a desarrollar la IA y otras tecnologías digitales no solo deberían estar educados y formados en su disciplina, sino también en la dimensión ética y *política* de su trabajo. A la luz de los problemas que tienen que ver con la IA y la democracia, necesitamos educar a los informáticos y a los ingenieros

acerca de cómo se relaciona la IA con valores y conceptos políticos tales como la democracia. O, mejor y más radicalmente expresado, abordar esta dimensión ética y política debería ser *parte de su disciplina*. Las iniciativas actuales en materia de ética integrada en la educación de la IA y en universidades de los Estados Unidos (por ejemplo, en Stanford y Harvard) y de Europa van en la dirección correcta. En cualquier caso, es importante que tal educación esté *integrada* en los procesos educativos, más que simplemente añadida a estos. Desarrollar buen *software* y buena IA debería incluir una noción más amplia de calidad, una que no solamente se refiera al funcionamiento (fines técnicos inmediatos), sino que abarque también la calidad de la vida y de la sociedad humanas, con fines democráticos y sociales más amplios, y orientada al bien común. Así pues, la IA democrática es IA que no solo hace lo que sea que se supone que hace (por ejemplo, hacer recomendaciones en las redes sociales), sino también que sea decisiva a la hora de fortalecer la democracia y apoyar la creación de entendimiento y de bien comunes.

Tal integración democrática del desarrollo de la IA en la sociedad promete fortalecer la democracia y hacerla más resiliente frente a estos cambios tecnológicos rápidos y omnipresentes: no mediante la regulación de la IA existente o fortaleciendo las instituciones democráticas (que, como he defendido, también es necesario), sino cambiando la tecnología y las formas en las que se desarrolla. La idea es que los riesgos políticos relacionados con la IA se tengan en cuenta en el desarrollo de la tecnología y que ya se discutan en la educación y en diferentes instituciones políticas y espacios públicos (nuevos y viejos). Todo ello supondría una importante mejora comparado con la situación actual, que no es lo suficientemente proactiva. Esto no es solo bueno para la democracia, sino que también permite que la tecnología sea más sostenible ética y políticamente, haciéndola más robusta en este sentido (de nuevo, una gran mejora comparada con la situación actual).

Sin embargo, como sugiere mi comentario acerca de los fines amplios y el bien común, también es posible concebir la IA democrática como *fomento* de la democracia, y no simplemente protección. Permítaseme añadir algo más con respecto a este enfoque constructivo en los siguientes capítulos.

Capítulo 7

IA para la democracia y un nuevo Renacimiento

Mi examen de la IA y la democracia se ha enmarcado hasta ahora de manera preferente en términos de peligro, riesgo, vulnerabilidad y resiliencia. Pero esto tergiversa de alguna manera mi proyecto de desarrollo de una IA democrática [DDI]. La DDI pretende una IA democrática en tanto que esta debe ser resistente al totalitarismo, pero también, de manera más positiva y constructiva, una IA *para* la democracia.

IA PARA LA DEMOCRACIA

La idea de una IA para la democracia consiste en disponer de una inteligencia artificial que fomente y fortalezca la democracia en lugar de debilitarla: desarrollar la IA como una verdadera *tecnología de comunicación* en el sentido republicano e ilustrado de la palabra. La IA, pues, debería *ayudarnos* a plasmar un ideal de democracia más rico y relacional: debería contribuir a construir un mundo común, a desarrollar un sentido común y a trabajar por el bien común.

La IA para la democracia podría significar fortalecer la base de conocimiento del ideal de democracia republicano y deliberativo. Por ejemplo, se podría imaginar un programa como ChatGPT que ayude a la educación cívica, pero que también lo haga en los procesos deliberativos ofreciendo (resúmenes) el conocimiento necesario para los ciudadanos. Por ejemplo, podría ayudar a difundir entre los participantes el conocimiento científico relevante y, de esta forma, contribuir al fortalecimiento epistémico de este pilar de la democracia. También podría informarles de las diferentes opiniones políticas, quizás dentro de unos límites fijados de antemano que limiten el rango de opiniones (véase de nuevo la «paradoja de la tolerancia» de Popper), y con algún procedimiento diseñado para detectar la desinformación y que contribuya de esta manera al mantenimiento de la diversidad y de la pluralidad necesarias para un debate público sano y rico. La IA también podría *ayudar* a detectar la desinformación y el sesgo en vez de contribuir a ellos.

La idea de emplear la IA para la democracia no es completamente nueva. Algunos investigadores ya han trabajado en modalidades técnicas para diseñar una inteligencia artificial democrática. Por ejemplo, Koster y sus colegas (2022) han optimizado un sistema de IA para preferencias humanas y han diseñado un mecanismo social que obtiene la mayoría de los votos en su experimento. Pero se podría ir más allá de la democracia como mayoría de votos. Tomando las sugerencias anteriores, se podría, por ejemplo, construir sistemas de recomendación que usen los ciudadanos para encontrar información de maneras que aseguren su diversidad y calidad, y eviten el sesgo, *mejorando* de este modo, en lugar de debilitarlas, tanto la agencia como la base epistémica para la deliberación y la participación democráticas. Ya existe investigación que busca aumentar la explicabilidad e interpretabilidad de la IA y que puede ayudar a construir la confianza necesaria para la democracia. Por ejemplo, Bezzaoui, Fegert y Weinhardt (2022) han propuesto un sistema de aprendizaje

automático que es entrenado para reconocer la desinformación y que (como IA explicable) informa de ella a los usuarios.

Sin embargo, muchos de los esfuerzos de investigación en IA e, incluso, en su ética, no están orientados específicamente a fortalecer la democracia. Además, como Henrik Stetra, Harald Borgebund y yo mismo hemos defendido (2022), la investigación técnica actual, que afirma estar haciendo esto, tiende a usar un significado de democracia muy aguado, y carece de una conexión con la filosofía política y su historia. Ya existe algún trabajo que busca conectar la ética de la IA con su desarrollo, tanto en la academia como en la industria. Sin embargo, se necesitan esfuerzos similares para la política de la IA, usando los recursos de la filosofía política (Coeckelbergh, 2022a) y conectándolos con la investigación y el desarrollo de la IA. Además, definir qué es una IA democrática no debería dejarse en manos de expertos, pues existe de nuevo el peligro de la tecnocracia. En su lugar, necesitamos *tecnodemocracia*. El desarrollo de una IA democrática ha de integrarse en procesos de participación equitativos, inclusivos y justos. Como ya he sugerido, aquí podemos aprender de la innovación responsable: sus metodologías son útiles para combatir y anticipar el impacto de la IA, para integrar valores éticos y políticos en su desarrollo y, en línea con la concepción de democracia adoptada en este libro, para hacer del desarrollo de la IA un proceso más inclusivo y participativo. Por ejemplo, en respuesta a las amenazas a la democracia, Bruno Frey ha propuesto que los ciudadanos deberían participar de forma activa en los procesos de toma de decisiones y Jeroen van den Hoven ha defendido que la innovación debe reflejar nuestros valores (Helbing *et al.*, 2017). Además, al desarrollar estas estrategias y métodos, se necesita poner más énfasis en qué conocimiento, experiencia, comprensión y comunicación (en el sentido más rico de crear comunidad y crear mundo) son necesarios en esos procesos de innovación responsable. Por ejemplo, necesitamos averiguar cuál debería ser el papel de los expertos y qué es necesario en términos

de capacidades de conocimiento y comunicación para que los ciudadanos puedan participar de forma significativa y efectiva y contribuir a los procesos tecnodemocráticos de este tipo.

Otras fuentes para desarrollar una IA para la democracia son los experimentos de democracia directa asistida por aquella. También pueden ayudarnos a reflexionar acerca de las posibilidades de una democracia directa que no existía anteriormente, pero que ahora es posible con la IA y otras tecnologías digitales. Por ejemplo, en Taiwán los ciudadanos ya disponen de posibilidades para la participación y la discusión *online* (como entregar peticiones digitales y hacer comentarios sobre el presupuesto) y se ha utilizado la IA en una plataforma participativa *online* para gestionar la conversación (Tang, 2019). Además, en *Cyber Republic*, Zarkadakis (2020) ha defendido una combinación de inteligencia de las máquinas y asambleas ciudadanas para resolver el problema de las asimetrías de conocimiento y salvar la democracia liberal. La idea es seleccionar un grupo de ciudadanos cuya diversidad refleje la de un grupo social más amplio que, a continuación, se divide en grupos, cada uno de los cuales se centra en una parte del problema. Cada grupo interactúa con un conocimiento base: repositorios de datos, expertos humanos y medios digitales. La IA puede ayudar a automatizar (parcialmente) el proceso. Por ejemplo, el aprendizaje automático permite seleccionar frases populares y ayudar así a los ciudadanos a alcanzar un consenso (Zarkadakis, 2020: 67-84). La IA también puede facilitar el descubrimiento de conocimiento, ayudar a las personas a decidir si deberían confiar en una información particular, traducir textos si fuera necesario y ayudar en el control de horarios (*ibíd.*: 85-100).

Esta propuesta muestra mucha confianza en la capacidad de los ciudadanos para gobernar, suponiendo que se les ha ofrecido una base de conocimiento. Esto va en línea con el enfoque deliberativo y republicano. Se trata de una forma tranquila de intentar implementar un modelo de democracia directa en un mundo que parece menos

apropiado que nunca para hacerlo. Al involucrar algunos elementos de representación, es, quizás, menos vulnerable a algunos argumentos clásicos contra la democracia directa. Sin embargo, puede que continúen existiendo asimetrías de poder tanto dentro de las asambleas como de los grupos. Algunos tienen más conocimiento y capacidades previas que otros: esto tiende a influir en la opinión que unos y otros defiendan. Por ejemplo, las diferencias en conocimiento técnico o las capacidades retóricas es probable que se traduzcan en diferencias de poder.

Obsérvese también que este es un enfoque muy estructurado y dirigido. Se necesita mucho de esto. Desafortunadamente, no podemos depender por completo de lo que ya hacen los ciudadanos con la tecnología digital y la democracia. Por ejemplo, Coleman (2017) observa cómo interactúan los ciudadanos *online* unos con otros y propone que nos basemos en esas prácticas de comunicación horizontal entre pares para apoyar la democracia. Pero esto no está exento de problemas. Como reconoce Coleman, la información puede distorsionarse. Considérense de nuevo problemas tales como la desinformación y las burbujas epistémicas. El enfoque de Coleman muestra demasiada confianza en las promesas democráticas que se hicieron en los años 1990 y 2000 con respecto a internet y las redes sociales. Entretanto, conocemos más cosas y sabemos que esto no es así: estas tecnologías y las prácticas relacionadas con ellas no han conducido a un mundo más democrático. Y, a pesar de que simpatizo con las organizaciones de base y las iniciativas de la sociedad civil, tengo miedo de que las iniciativas de abajo arriba, ellas solas, no funcionen. Reconozco que, sin duda, tienen potencial. Se pueden crear bienes comunes de datos, por ejemplo (véase también el próximo capítulo). Y Savaget, Chiarini y Evans (2019) muestran que la sociedad civil puede emplear la IA para influir en la política. Los datos pueden usarse para controlar el gasto público y sacar a la luz la corrupción. La IA también permite fomentar la participación. Pero el caso de Brasil también muestra que existen

intereses poderosos que operan contra ella, y esto junto con las resistencias institucionales en las organizaciones relevantes. Por lo tanto, también es necesario el cambio institucional de arriba a abajo. Actualmente, la mayoría de nosotros vivimos en colectividades más grandes y bajo condiciones políticas más duras que requieren una acción urgente para salvaguardar la democracia. Junto con el apoyo a las iniciativas sociales existentes y facilitando los cambios de abajo arriba que están en marcha, también necesitamos visión y liderazgo político para iniciar nuevos cambios institucionales y realizarlos en la forma en la que se desarrolla y usa la tecnología.

UN NUEVO RENACIMIENTO

A la luz del desafío de desarrollar una conciencia política de la IA como algo que sirve a la democracia, es importante recordar que la inteligencia artificial no es solo una tecnología, sino también un sueño y un ideal. Necesitamos, por tanto, una perspectiva más amplia del futuro de nuestras sociedades y del planeta que nos permita guiar los cambios institucionales y políticos mencionados anteriormente. Para ello, no solo necesitamos mejores tecnologías, sino también mejores narrativas e imaginación. Y necesitamos personas que puedan hacer esto. En otras palabras, necesitamos educar a nuevas generaciones de personas que, a la luz de los riesgos y posibilidades de la IA, puedan crear nuevas historias e imágenes de un futuro en el que merezca la pena vivir.

No se trata solo de imaginar un futuro tecnológico mejor, *de* tecnología. También podemos *usar* las tecnologías digitales para crear estas historias nuevas. Está bien que recordemos aquí que la tradición política del republicanismo, que inspiró muchas de las ideas discutidas en este libro, ha estado conectada históricamente con el humanismo clásico según se desarrolló en la época antigua y en el Renacimiento. Los humanistas del Renacimiento construyeron

comunidades de aprendizaje y hablaban con colegas académicos y actores políticos cruzando las fronteras de sus respectivos países. De esta forma, las ideas se expandieron, nacieron nuevos movimientos políticos y mundos enteros cambiaron para siempre. La tecnología ha desempeñado un papel importante en estas *comunicaciones*. Los medios de comunicación y las tecnologías que impulsaron el Renacimiento fueron los libros y la imprenta. Ahora, en la época digital, necesitamos construir nuevas comunidades de conocimiento y comunicación basadas en las tecnologías digitales. Si queremos una revolución política realmente democrática necesitamos una buena base epistémica. La producción y comunicación del conocimiento son los impulsores principales de la resiliencia y del florecimiento democráticos. Necesitamos promocionar y construir una esfera pública de mejor calidad (ciertamente, crear un público mejor, como sugiere Splichal), y fortalecer la confianza en el conocimiento y en la sociedad. La IA y las tecnologías digitales pueden ayudarnos en esto; propulsar una nueva revolución cultural y educativa: un nuevo Renacimiento y una nueva Ilustración basados en las tecnologías actuales (pero adaptadas para encajar en el nuevo objetivo) y completamente inter- y transdisciplinares.

HUMANISMO DIGITAL

Hasta cierto punto, esta «revolución silenciosa», como la llamarán algunos, ya está teniendo lugar de diferentes maneras. En la academia, algunos esgrimen el estandarte del «humanismo digital» (Werthner *et al.*, 2022), y la idea de un nuevo Renacimiento está ganando terreno incluso en el mundo empresarial (Recke, 2017). El de humanismo digital^[1] es un término vago, pero lo utilizo aquí para decir que, en lugar de usar el humanismo meramente para *criticar* las tecnologías digitales, se adopta un enfoque constructivo: el objetivo es desarrollar tecnologías en línea con los valores

humanísticos y con el objetivo de lograr una buena vida y, ciertamente, la democracia. Otro componente consiste en que la relación entre los humanos y las tecnologías debería ser estudiado de formas interdisciplinarias. De este modo, el humanismo digital conecta las orientaciones normativas de la filosofía antigua, el humanismo renacentista y la Ilustración, pero utiliza estas ideas como inspiración para el desarrollo de la tecnología digital contemporánea y generar debate en torno a él. Desarrollando y discutiendo la IA en estos términos podríamos avanzar en dirección a una *inteligencia artificial humanista*.

Algunas ideas del humanismo digital son bien conocidas entre quienes ya usan este enfoque más constructivo de la tecnología y practican el diálogo entre las humanidades y las ciencias técnicas desde hace décadas, por ejemplo, en los campos de la filosofía de la tecnología, los estudios sociales de la ciencia y la tecnología, la ética computacional y la ética de la ingeniería. Pero el concepto de humanismo digital puede utilizarse de forma estratégica para promover y desarrollar este enfoque más allá, sobre todo si se parte de una base intelectual más amplia (un objetivo al que contribuye este libro).

Por supuesto, el propio término «humanismo» no es políticamente neutral y se emplea para todo tipo de propósitos. Como dije, a menudo se usa para criticar la tecnología, por ejemplo cuando en el contexto de los debates sobre la IA se moviliza y defiende lo humano contra la amenaza de la máquina. Por ejemplo, Nida-Rümelin (2022), quien, junto con Weidenfeld introdujo por primera vez el término «humanismo digital» en el mundo germanoparlante y apeló al humanismo, pero también a la Ilustración, abogaba por fortalecer y proteger la responsabilidad humana más que delegarla en la IA. Sin embargo, las afirmaciones de este tipo corren el riesgo de abandonar formas más complejas en las que la IA es capaz, no obstante, de influir en nuestro mundo moral y político, y tienden a desestimar formas más positivas en las

que la automatización de la IA contribuye a la sociedad, incluyendo a la democracia. A menudo el humanismo ha estado demasiado a la defensiva cuando se trata de evaluar la tecnología. El humanismo digital puede contribuir a estimular enfoques más constructivos y proactivos.

Además, por lo que respecta a la educación, el movimiento humanista digital también necesita más apoyo desde dentro de las humanidades. A pesar de que actualmente más y más informáticos e ingenieros comienzan a comprender que sus tecnologías tienen consecuencias éticas y sociales, con frecuencia los estudiantes y los académicos de la rama de humanidades continúan sin ser conscientes de la dimensión profundamente política de las tecnologías, y de las maneras en que conforman nuestro conocimiento del mundo y nuestra relación con otros. Aquí también se necesita un cambio en la educación. El humanismo clásico, que se centra en el discurso y rechaza sus *media* y tecnologías, ya no es suficiente. La biblioteca debe completarse con el laboratorio; los estudiantes y académicos de humanidades necesitan intercambios, espacios y proyectos más interdisciplinarios (y, en especial, más comunicaciones interdisciplinarias) sobre la política de la IA y otras tecnologías digitales.

Otra barrera para el crecimiento del humanismo digital como proyecto político-académico es el hecho de que, paradójicamente, se trata de humanos. Los poshumanistas críticos, como Haraway (1991) y Braidotti (2019), pero también los filósofos medioambientales y los pensadores posmodernos como Foucault, han criticado al humanismo por excluir a algunos humanos (mujeres, esclavos, poblaciones nativas) y por ser antropocéntrico. En respuesta a estas críticas cabe argumentar que cualquier forma de humanismo digital contemporáneo debe ser lo suficientemente inclusiva y considerar, al menos, si su enfoque ha de ampliarse hasta los no humanos. Demasiado a menudo, el humanismo digital es un mero sinónimo de los llamados enfoques «humanocéntricos» de la IA. y de otras

tecnologías digitales que son populares en las facultades técnicas y en los documentos políticos. Esto ha sido útil para promocionar la ética de la IA. Pero, como digo, tiene que haber una base intelectual más amplia, y una forma de lograrlo es cuestionando el antropocentrismo del humanismo digital[2].

Para que funcione una democracia del siglo XXI, estas nuevas formas de humanismo, Renacimiento e Ilustración tienen que implementarse en la educación de las personas en todo el espectro disciplinar, incluyendo tanto a los informáticos *como* a los filósofos. En la educación superior, el nivel necesario de interdisciplinariedad para el humanismo digital necesita un cambio radical en la estructura de nuestras universidades, más que unos meros cambios cosméticos. Tal vez lleve tiempo, pero, cuando todo esto sea posible, el humanismo digital, el nuevo Renacimiento y la nueva Ilustración que aquí se defienden serán de gran ayuda para seguir desarrollando la visión e imaginación políticas necesarias para crear una IA para la democracia y otras tecnologías digitales que la mejoren.

CONCLUSIÓN

Así pues, el desafío de esta revolución tecnocultural es nada menos que crear un nuevo orden social, político y cultural que esté entremezclado con las tecnologías digitales de formas que sean sostenibles ética, política y ecológicamente. Tal desafío no atañe solo a la IA, sino que implica una transformación social e intelectual más amplia y profunda. A la luz del ideal de democracia aquí empleado y tomando inspiración del poshumanismo, la cuestión se convierte en el desafío de crear un sentido común, crear comunidad y construir un mundo (más) común dada la situación tecnológica, política y medioambiental actual. Prestar especial atención a las tecnologías y a cómo cambian la sociedad y el entorno planetario es clave en todo

esto. Si el uso de las tecnologías digitales de nuestros días como la IA muestra tendencias que van en contra de crear un mundo común, y tienden a la fragmentación y exclusión epistémica y política, tenemos que hacer que nuestras democracias sean más resilientes y mejores, y asegurarnos de que la educación ayude a crear una base común y las capacidades necesarias para actuar en la esfera pública; y que nuestras tecnologías cambien de tal forma que fomenten la *comunicación* en el sentido de compartir un mundo, comprenderse unos a otros y construir comunidad: potencialmente también a nivel global y más allá de lo humano.

Con respecto a la construcción de una comunicación y comunidad humanas, el asunto no es que debamos convertirnos todos en lo mismo o en uno: ni a nivel local ni global. El origen de la palabra «comunicación» no está relacionado con «unidad», sino con «hacer común» y «compartir» (Peters, 2008). Cuando compartimos, continuamos siendo personas únicas. Y, como personas, podemos beneficiarnos de la comunidad y, por supuesto, retener nuestros propios intereses y (teniendo en mente la Ilustración) nuestros propios derechos individuales. Pero el giro de comunicación democrático al que nos referimos aquí tiene que ver con lo que le debemos a la comunidad, en lugar de lo que la comunidad o el Estado debería hacer por nosotros. En la antigua Roma, *munus* significaba un servicio público ofrecido por un ciudadano a su comunidad: tenía que ver con obligaciones públicas. Actualmente, el desafío consiste en orientar y transformar nuestras tecnologías e instituciones para que faciliten un servicio de intercambio público real: no la mera defensa de los derechos e intereses del individuo, ni en «compartir» las opiniones individuales, sentimientos de ofensa o expresiones de legitimación, sino en la creación y participación en un mundo común. No es cuestión de reclamar únicamente que el Estado haga cosas por nosotros, ni de limitarse a expresar sentimientos y opiniones personales, sino de afrontar la creación de una *res publica*, una esfera pública, a través de la deliberación, la

cooperación, la comunicación y la comprensión comunes. Como ciudadanos somos corresponsables de crear ese mundo compartido, que a su vez generará confianza y sentido común y ayudará a realizar la democracia.

Para la IA y las tecnologías que son parte de su ecosistema, como las redes sociales digitales, este ideal democrático y educativo implica que necesitamos asegurarnos de que no solo permitan la búsqueda de intereses particulares y la expresión de opiniones individuales, sino que también fomenten la comunicación y la construcción de comunidad: no se trata de eliminar la diferencia, sino de alentar el verdadero intercambio y la construcción de confianza y entendimiento. Además, necesitamos comunidades e instituciones que formen y cultiven las capacidades comunicativas y las virtudes epistémicas de los ciudadanos, tales como estar dispuesto a escuchar las perspectivas de los demás y practicar la tolerancia, estar preparado para que las propias opiniones sean vulnerables frente a la discusión pública, y estar dispuesto a imaginar y a trabajar en dirección del bien y del futuro comunes. Todo esto requiere cambios no solo individuales, sino también institucionales, por ejemplo, en la educación, cambios técnicos en la IA. y la forma en que la utilizamos, y cambios en las redes sociales y en sus infraestructuras. Tanto el sector público como el privado tienen un papel que desempeñar en todo esto. Si no logran emprender estas tareas, los primeros se arriesgan a volverse irrelevantes y los segundos, antidemocráticos. El futuro de nuestras sociedades es demasiado importante como para dejarlo en manos de tecnócratas o a desarrolladores de IA y a sus empresas: una democratización de la IA y de otras tecnologías digitales es el único camino hacia delante.

Como ya he defendido a lo largo de este libro, esta democratización de la IA debería integrarse dentro de un esfuerzo más amplio por reformar nuestras democracias, y basarse en una discusión más profunda y amplia del tema de lo que normalmente se

ofrece en los debates acerca de la ética de la IA. Como correctamente ha defendido Himmelreich (2022), las exigencias actuales para democratizar la IA dependen con frecuencia de una concepción de la democracia demasiado estrecha, que esencialmente la reducen a votación y participación. Aquí y en los capítulos previos he demostrado que la democracia (y cómo impacta la IA sobre ella) implica mucho más que eso, pues tiene que ver también con cuestiones como la libertad, la igualdad, la justicia y el poder, así como con la cuestión de qué tipo de conocimiento necesitamos en una democracia y cuál debería ser el papel de ese conocimiento. Tiene que ver con la reflexión acerca de nuestras instituciones políticas y sociales, con la forma en que organizamos nuestras economías y con la educación. Y, en última instancia, con la comunicación y con (re)hacer un mundo común. Necesitamos cambios en todas estas áreas.

Si no logramos acometer estos cambios, nos arriesgamos a que la IA siga erosionando y corrompiendo las democracias existentes. Eso sería un desastre. No solo equivaldría a abandonar un conjunto de ideales magnífico y valioso, incluyendo la democracia, el humanismo, el republicanismo y la Ilustración. También sería una pérdida terrible para la humanidad, que se arriesga a volverse indefensa y a perder las muchas oportunidades ofrecidas por la IA. De forma más importante y concreta, sería una enorme pérdida para nosotros, los ciudadanos, como seres humanos limitados y vulnerables. De hecho, esto quiere decir que los ciudadanos que ya están experimentando algunas amenazas y problemas creados por la erosión de la democracia (en contraste con aquellos políticos y oligarcas tecnológicos que se benefician de la situación actual) corren el riesgo de tener que pagar el precio más alto por la pérdida de la democracia: la vulnerabilidad socioeconómica y el aislamiento. Esta situación puede desembocar en respuestas antidemocráticas y, potencialmente, en los horrores del autoritarismo y el totalitarismo.

Incluso si no fuera así, conduciría a menos democracia, más miseria, menos florecimiento humano y menos bien común.

Capítulo 7

El bien común y la comunicación

Para concluir, permítaseme acercarme a un concepto político-filosófico que ha sido parte de mi caja de herramientas normativa durante el tiempo en que he estado escribiendo este libro. En varios puntos de los capítulos anteriores he defendido que la IA debería apoyar el bien común. También he rastreado este argumento hasta algunas fuentes de la historia de la filosofía, y he mencionado conceptos relacionados como el sentido común. Permítaseme finalizar este libro revisitando y clarificando esta orientación normativa hacia el bien común que considero que es una parte importante de la concepción fuerte y rica de la democracia que he empleado. También lo relacionaré con otros conceptos «dirigidos hacia lo común» (comunes [*commons*], comunicación, comunidad, experiencia común y mundo común), y exploraré en mayor profundidad sus implicaciones para (la política de) la IA y la democracia.

EL BIEN COMÚN Y LOS COMUNES

Apelar al bien común con respecto a la IA no es algo nuevo: a pesar de no disponer de la atención que se merece, el concepto ha

desempeñado durante los últimos años algún papel en el discurso de la ética de la IA. Por ejemplo, el informe de la Comisión de Inteligencia Artificial de la Cámara de los Lores británica (Cámara de los Lores, 2018) propuso como uno de los principios fundamentales para establecer un código de la IA que «debería desarrollarse para el bien común y en beneficio de la humanidad» (2018: 125). Y en un artículo de opinión del Carnegie Council, Kaspersen y Wallach se preguntan: «¿Cómo podemos asegurarnos de que, actualmente, las tecnologías están siendo desarrolladas teniendo en cuenta el bien común, en lugar de hacerlo en beneficio de unos pocos elegidos?» (2023). Como se desprende de mis comentarios a lo largo de este libro, no podría estar más de acuerdo con esta opinión.

Pero ¿qué quiere decir el bien común? Cuando se menciona el término en documentos de ética y política de la IA, normalmente no se define (Berendt, 2019). A menudo significa simplemente que todos deberíamos beneficiarnos de la IA, no solo unos pocos. Apoyo esta opinión, pero hay más significados y hay mucho más que decir sobre la cuestión. En filosofía política, el término es algo controvertido: como ya he explicado, está enraizado en una teoría particular (o conjunto de teorías), y no todo el mundo cree que deba ser una meta política central. Para clarificar lo que *yo* quiero decir con el término, y para desarrollar en mayor profundidad y concluir esta parte de mi argumento, permítaseme (re)visitar la discusión en torno al bien común en la filosofía política y señalar lo que está en juego para el debate en torno a la IA y la democracia.

En los capítulos anteriores ya he mencionado que el término «bien común» desempeña un papel en las teorías republicanas tradicionales de la democracia, por ejemplo, en Aristóteles, Rousseau, Tocqueville, Dewey y Arendt, pero también en teorías democráticas deliberativas, participativas, epistémicas y neorepublicanas más recientes. Permítaseme, pues, desentrañar más aún el concepto.

En la tradición aristotélica, el bien común no es la agregación de bienes individuales o, como en el utilitarismo, la mayor cantidad de placer o satisfacción preferente para la mayor cantidad de personas. Más bien, la idea es que la política (democrática) no debería solo, o no principalmente, abordar el fomento del interés personal y el bien individual, sino lo que es compartido y beneficioso para todos los miembros de una comunidad política dada. La idea debería estar dirigida al florecimiento de la *polis*. Ya he señalado que los ciudadanos de la *polis* se benefician del bien común, pero también se supone que contribuyen a él. Como señala Hussain (2018), tienen una «obligación relacional» de preocuparse por el bien común. El énfasis se pone en la dimensión pública y común de la política y de la democracia. Aristóteles defendió incluso en la *Política* (1998) que el bien de la ciudad-Estado debería tener prioridad. En la toma de decisiones políticas y en su vida pública, se supone que los ciudadanos trascienden sus intereses particulares y se preocupan por el bien común. Esto no solo lo promueve (y, de acuerdo con Aristóteles, la virtud y la vida buena), sino que también evita que haya políticos y gobiernos corruptos. En última instancia, protege contra la tiranía.

Esta idea del bien común ha sido muy influyente en la historia de la filosofía política. Por ejemplo, Tomás de Aquino, dentro de la tradición aristotélica, defendió que la ley debería tener como objetivo el bien común. Esto influyó en el pensamiento renacentista (por ejemplo, en Maquiavelo) y, más tarde, dicho concepto reaparece en las obras de Rousseau, Smith, Marx, Mill, Keynes, Rawls y muchos otros pensadores. También influyó en la política moderna. James Madison, uno de los Padres fundadores de los Estados Unidos que desempeñó un papel clave al redactar el borrador de su constitución, instó a sus conciudadanos a cooperar en pos del bien común y consideró la distribución desigual de la propiedad un factor clave que divide la sociedad. Los *Federalist*

Papers argumentan a favor de los gobernantes que tienen la sabiduría para discernir y perseguir el bien común^[1].

En la filosofía política contemporánea, la noción del bien común desempeña un papel en la ya mencionada teoría democrática neorrepblicana. Pettit ha defendido que el fin de la democracia es promover el interés público y servir al bien común. Vincula esta noción con su concepto central: la no dominación. En una democracia republicana, defiende Pettit, el Estado tiene justificación para inmiscuirse en las vidas de las personas actuando a favor del bien común, pero no de una forma dominadora (2004: 15) o de una forma que sea arbitraria. De esta manera, el principio de la libertad, uno de los pilares de la democracia, se interpreta en términos de no dominación y esto limita, para Pettit, cómo se puede apoyar el bien común. Ese límite ayuda a mantener a la democracia de forma liberal. Es un desafío importante de la política moderna: en una democracia liberal, trabajar en favor del bien común necesita que este se mantenga en equilibrio con el principio de la libertad (individual).

Pero ¿cuál es el contenido del bien común (o, digamos, del interés público)? Los teóricos políticos contemporáneos han estado en desacuerdo con respecto a si el término puede y debe definirse de una forma sustantiva sin mayor justificación (con la democracia tratando, en consecuencia, de realizar este denominado «ideal regulativo»), o si debería ser considerado como el resultado de los procedimientos democráticos participativos y deliberativos. Se supone que tales procedimientos rastrean el bien común: un bien que no se define de antemano, pero que necesita descubrirse mediante un concienzudo proceso deliberativo. En opinión de Dewey, es una cuestión de ensayo y error, de experimento: no existe una concepción *a priori* del bien común. También es una cuestión de deliberación real: la atención no se centra demasiado en los ideales^[2]. Esto resuelve el problema de que no solo estamos a menudo en desacuerdo con respecto a lo que significa la

democracia, sino de que tal vez no *sepamos siquiera qué queremos* cuando decimos que queremos democracia. La respuesta de Dewey es: descubrámoslo juntos.

Ciertamente, una objeción principal a las teorías sustantivas es que pasan por alto que podemos discrepar acerca de lo que es el bien común. Se podría responder que esta es principalmente una preocupación teórica y que, en la práctica, a menudo hay acuerdo. Por ejemplo, en el caso de la IA, muchas personas consideran que supone un problema cuando conduce a una sociedad altamente desigual e injusta en la que el poder se concentra en manos de las grandes corporaciones tecnológicas y beneficia a unos pocos en lugar de a la mayoría, por lo que en última instancia creen que no promueve el bien común. De forma similar, existe actualmente un consenso cada vez mayor que defiende que parte de lo que significa «bien común» es tener una economía sostenible, y un entorno natural limpio y no contaminado. El problema, pues, no reside tanto en conocer o definir el bien común, como más bien en que algunos de nosotros no compartimos este objetivo y no apoyamos el trabajo en pos del mismo porque no está en su interés inmediato el hacerlo. En otras palabras, muchas personas pueden estar de acuerdo en lo que es y significa el bien común y, aun así, seguir persiguiendo sus propios intereses. En el mejor de los casos, creen que son otros los que deberían resolver el problema: este es el denominado «problema del gorrón». Sin embargo, en el peor de los casos, no les importa el bien común (incluso si saben lo que se quiere decir con él), o incluso pueden intentar engañar a los otros de modo que piensen que lo que ellos hacen contribuye al bien común. Como defendieron Marx y Engels en *La ideología alemana* (2014 [1864]), la clase dominante intenta típicamente presentar sus intereses como el bien común. Desde otro ángulo, Pettit también reconoce que un sistema es susceptible de malinterpretar ciertos intereses privados como públicos (2004: 166). Y existe el peligro de que una persona o grupo de personas intente imponer a otros su visión del bien común,

por ejemplo, cuando en un sistema autoritario los gobernantes usan narrativas utópicas y propaganda colectivista para este fin.

Con todo, incluso en una democracia que funcione bien es inevitable que haya desacuerdos, no solo con respecto al bien común y a cómo deberían equilibrarse este y los intereses y los derechos individuales, sino también acerca de lo que significa en la práctica apoyar al bien común. Por ejemplo, la mayoría de nosotros estaremos de acuerdo en que la sostenibilidad es un fin, pero eso no significa que estemos de acuerdo en cómo lograr ese fin ni en cómo la consecución de ese fin debería guardar cierto equilibrio con los intereses privados. ¿Significa en este caso la consideración del bien común que, por ejemplo, deberíamos dejar de usar coches y de comer carne? ¿Debería ser obligatorio o forzado? Es posible que algunos no estemos dispuestos a abandonar nuestra libertad, un valor que, como ya he defendido, también es fundacional para la democracia, y que está relacionado con la noción de no dominación de Pettit. En cualquier caso, se necesita reconocer que existen diferentes intereses, no solo, por ejemplo, los intereses del colectivo en general, sino también los intereses individuales y del grupo. También se debe tener en cuenta de alguna forma el pluralismo cuando se trata de valores y principios políticos. Existen muchos valores e, incluso si estamos de acuerdo en respetar algunos de ellos (algo que no es obvio en un mundo multicultural), también podemos estar en desacuerdo con respecto a qué valor deberíamos priorizar en un caso particular y qué significa plasmar ese valor en un contexto particular. Esto hace que sea problemático el término «bien común» en su forma sustantiva.

A la luz de estos desafíos, los enfoques procedimentales son atractivos, ya que parecen ser más compatibles con una visión pluralista de la sociedad y de la democracia. Si no estamos de acuerdo acerca de lo que significa e implica en la práctica el bien común (pero, aun así, estamos de acuerdo *en que* deberíamos promoverlo), hablemos de ello y descubramos juntos lo que significa

en la práctica, aquí y para nosotros: discutamos y deliberemos qué es bueno para nuestra comunidad política. Desde los puntos de vista kantianos, habermasianos y rawlsianos, de lo que se trata es de la deliberación y el razonamiento públicos (a pesar de que tales consideraciones no son puramente procedimentales)[3]. La idea es deliberar y encontrar la(s) razón(es) común(es) que pueda(n) ser compartida(s) por los ciudadanos con diferentes trasfondos dentro de una sociedad pluralista. Sin embargo, en opinión de Dewey, el enfoque procedimental se vuelve más experimental: averigüemos lo que funciona para nosotros, experimentemos. Dewey rechaza los enfoques fundacionalistas (y la apelación a la razón abstracta) y propone aplicar métodos experimentales a la democracia. Desde esa perspectiva, la democracia tiene que ver con aprender y crecer en común, aprender y crecer al nivel del colectivo: no simplemente a nivel individual. En el pensamiento relacional republicano, ambos niveles están conectados. Además, para asegurarnos de que están incluidos diferentes intereses y valores, se podría hacer que los procesos deliberativos y participativos fueran más inclusivos. Esto se relaciona con la cuestión de *quién* está invitado a la mesa: una cuestión de poder. Haciendo un guiño a Mouffe (2013), se podría y se debería permitir la oportunidad de rebatir una definición particular del bien común. Si esto no es posible, el peligro del autoritarismo se encuentra de nuevo a la vuelta de la esquina.

Sin embargo, estos enfoques no se excluyen por completo. A pesar de que desde una perspectiva deliberativa experimental (digamos, una perspectiva deweyana), se tendría que rechazar un enfoque fundacionalista del bien común, sin embargo, aún se podría estar de acuerdo en trabajar por el bien común basándose en un consenso amplio de lo que significa antes incluso de someterlo a deliberación. De este modo, se podrían incluir elementos sustantivos como puntos de partida y limitaciones del debate, y usar el proceso deliberativo para discutir cómo alcanzar el objetivo general (el bien común) y para equilibrar diferentes valores e intereses. También se

ha defendido que los elementos sustantivos resultan necesarios, ya que un procedimiento, por sí solo, no *garantiza* la realización del bien común (Blum, 2016), a pesar de que esto pide de nuevo la cuestión de qué es lo que *puede* hacerlo. Un enfoque procedimental y experimental general evita este problema, y la versión deweyana puede vivir con el hecho de que no haya garantías. La democracia necesita fundamentos (en este libro he enumerado varios principios y conocimientos necesarios) pero no ese tipo de fundamentos: dentro de una perspectiva deweyana, no existe tal fundamento o garantía.

Para la IA, esto implica que necesitamos la elaboración de políticas relativas a ella que busquen explícitamente el bien común y la elaboración deliberativa de políticas tecnológicas para averiguar qué significa e implica el bien común en el contexto del uso y desarrollo de esta tecnología, a la luz de los desafíos sociales actuales. Esto supone el desarrollo de la deliberación pública en torno a la IA y el bien común, pero también experimentar con el desarrollo tecnológico: descubrir qué implica la IA para el bien común probando diferentes versiones. Quizás incluso la(s) legislación(es) que la tenga por sujeto debería(n) considerarse un experimento, especialmente a la luz de los rápidos cambios tecnológicos (considérese lo rápido que se desarrollaron los grandes modelos de lenguaje) y el impacto incierto de las regulaciones. Reflexionemos sobre las formas de hacer que el proceso regulatorio sea, a la vez, más rápido (por ejemplo, el Acta de IA en la UE ha necesitado varios años para completarse, y entretanto han cambiado muchas cosas) y más flexibles.

Sin embargo, la cosa no va solo de las políticas *sobre* la IA. Como ya he mencionado, esta también puede *ayudar* e incluso mejorar los procesos deliberativos y contribuir a monitorizar y estudiar los a veces desalentadores, pero también excitantes y potencialmente muy beneficiosos, experimentos sociotecnológicos en que se han convertido la IA y la tecnología digital. Hay una sensación de que ya

estamos viviendo el experimento. Puesto que muchos de nosotros usamos, por ejemplo, motores de búsqueda basados en IA y generadores de texto, estamos viviendo el experimento de la IA y todos nosotros participamos en él. Sin embargo, lo hacemos, desafortunadamente, sin haber dado nuestro consentimiento individual o colectivo, y se está haciendo en una forma no democrática, dirigida por las grandes tecnológicas. El desafío es dirigir este experimento sociotecnológico de una forma más democrática de lo que lo es ahora y de forma que vaya encaminado hacia el bien común.

Obsérvese que esta discusión acerca del bien común y la democracia también se encuadra, de nuevo, como un problema relacionado con el conocimiento y la democracia. Ya he defendido que el conocimiento es necesario para la democracia. Pero ¿qué tipo de conocimiento?, y ¿conocimiento para quién? Con respecto al bien común, se puede ir en una dirección más platónica, como lo hace la teoría epistocrática, y poner el énfasis en la necesidad de pericia y de conocimiento del bien común, lo cual necesita la implicación de expertos que, como guardianes platónicos, conocen de alguna manera el bien común y cómo alcanzarlo. O se puede subrayar la dimensión participativa de las soluciones procedimentales y defender que sea lo que sea el bien común (y que sea lo que sea lo que digan los expertos y los filósofos sobre el asunto), lo que importa si queremos *democracia* es que necesitamos un procedimiento democrático deliberativo y participativo legítimo a fin de definir lo que significa para nosotros y cómo puede y debe alcanzarse, en este tiempo y en esta sociedad y *polis*.

Afortunadamente, esas direcciones no son por necesidad mutuamente excluyentes. A lo largo del libro he sugerido que necesitamos ambas dimensiones. En la sociedad moderna precisamos de expertos: su pericia necesita confluir en los procesos deliberativos, especialmente cuando se trata de responder a la cuestión de *cómo* alcanzar el bien común. Tecnologías tales como la

IA pueden ayudar en esto. Pero deberíamos evitar una tecnocracia si eso implica un sistema en el que gobiernen *solo* los expertos, ayudados por la ciencia y la tecnología: un sistema en el que estos afirmen saber qué es el bien común (en oposición a los ciudadanos ignorantes) y, a continuación, imponer su definición a la sociedad. Al mismo tiempo, deberíamos evitar un sistema en el que los expertos ya no sean bienvenidos y no se confíe en ellos en absoluto: esto también destruiría la democracia en el sentido más rico apoyado en este libro.

Finalmente, cuando se define el bien común en contraste con los intereses personales, se podría, de una forma más positiva, defender que deberíamos buscar la protección y la creación de lo que haya literalmente en común: los *bienes* comunes (plural). Esto nos conduce al concepto de los *comunes*.

El término «comunes^[4]» se refiere a los recursos que se comparten y que están disponibles para todos los miembros de una sociedad o comunidad particular: recursos que son «gratis» en ese sentido. A menudo, los recursos naturales como el aire y el agua son considerados comunes, a pesar de que, con frecuencia, estén privatizados y monetizados. Los bienes culturales como el arte y la arquitectura también son a menudo definidos como parte de los comunes. Considérese, por ejemplo, el programa Patrimonio de la Humanidad de la UNESCO: algunos sitios históricos, edificios y pueblos son considerados parte del patrimonio de la humanidad. Se podría discutir también si los medios de producción, las tierras dedicadas a la agricultura, la información y las tecnologías deberían formar parte de los comunes. Ya he mencionado la idea de los comunes digitales. Internet también puede considerarse como parte de los comunes (Lessig, 2001): en los años 1990, internet se asociaba a menudo con la libertad (Barker, 2015). ¿Y no es la tierra como un todo un bien común (global), una tierra que es *gratis*^[5]?

Una forma de proteger los comunes es hacerlos propiedad del Estado: hacer que lo que los romanos llamaban la *res communis* fuera

parte de la *res publica*. Pero la propiedad pública no es necesaria para que un bien sea considerado parte de los comunes. Por ejemplo, el aire bien puede ser regulado en contextos particulares (por ejemplo, en respuesta a la contaminación atmosférica), pero él mismo no es propiedad del Gobierno. Y la Wikipedia, por ejemplo, aunque puede considerarse como parte de los comunes digitales, no es propiedad de ningún Estado.

Pero, para empezar, ¿por qué necesitan protección los comunes? Un concepto bien conocido en filosofía política y en economía es el llamado «tragedia de los comunes^[6]»: cuando sus usuarios individuales persiguen simplemente el propio interés, esto conduce a la extinción de los recursos comunes (finitos). La degradación del medio ambiente natural se considera a veces en estos términos. ¿Es este también un problema para los comunes digitales, o no padecen el problema de la extinción? Dulong de Rosnay y Stalder (2020) han defendido que, mientras que, por ejemplo, los datos y la información no se vean afectados por la sobreutilización o la exclusividad, puede haber aún diversas maneras de degradar y de contaminar (piénsese de nuevo en la desinformación, por ejemplo) y que, en lugar de compartir y de la autogobernanza, hay una tendencia hacia la falta de participación, la infrarrepresentación de ciertos grupos, la apropiación y el capitalismo de vigilancia.

Para la IA, crear y cuidar del bien común, entendido en términos de los comunes, podría implicar la creación y mantenimiento de datos comunes y de la parcial o completa desprivatización de las tecnologías e infraestructuras necesarias para la IA. La idea es, entonces, que la IA (por ejemplo, extensos modelos de lenguaje o motores de búsqueda) y los datos que utiliza sean accesibles para todos y que todos se pudieran beneficiar de ellos: una idea que continúa inspirando a algunas personas del mundo tecnológico de la IA^[7]. Además, deberían evitarse las diferentes formas de degradación y apropiación. Estas metas podrían alcanzarse a través de regulaciones y, si fuera necesario, haciendo de los comunes

tecnológicos una propiedad pública de la inteligencia artificial, de los datos y de las infraestructuras digitales necesarias. Ya existen iniciativas de comunes de datos^[8]. Mazzucato (2018) ha defendido que los datos deberían convertirse en bienes públicos, permitiendo a los ciudadanos conformar la economía digital, en lugar de que lo hagan los gigantes tecnológicos.

Para concluir, en este libro doy mi apoyo a una versión de la democracia dirigida al bien común (incluyendo también, quizás, la creación, el mantenimiento y la gestión de los comunes), y que tiene procedimientos deliberativos y experimentales para rastrear y crear el bien común. Esto último es importante, ya que puede no saberse por completo de antemano cuál es el bien común, dado que existen diferentes opiniones con respecto al bien común, y diferentes intereses y valores conectados con la discusión, habida cuenta de que en una democracia necesitamos evitar que un actor (por ejemplo, una empresa tecnológica y su propietario) imponga su idea de lo que es el bien común. Por lo tanto, es a la vez deseable y necesario deliberar y experimentar: necesitamos involucrar a los ciudadanos en el proyecto de la IA para el bien común, y necesitamos usar el ensayo y error, al menos hasta cierto punto. En consecuencia, la IA necesita desarrollarse y ser utilizada dentro de un marco democrático que debería ayudar a apoyarla, en lugar de debilitarla. Y, si en un sentido la IA ya es un experimento sociotecnológico a gran escala, necesitamos hacer que el experimento sea más democrático.

El bien común así definido supone no solo una guía para el uso bueno y democrático de la IA y de su desarrollo: él mismo es también parcialmente el *resultado* del desarrollo de la IA *para* la democracia. Además, la IA y otras tecnologías digitales pueden desempeñar un papel en este proceso hacia el (mayor) bien común. Pueden y deben ayudar a crear el bien común, el conocimiento común y la experiencia común, y pueden y deben ayudar a que los procesos deliberativos y experimentales conjuntamente descubran el

bien común. Potenciar el bien común también implica proteger y crear, por ejemplo, modelos de lenguaje y datos comunes.

Finalmente, cuando se aborda el establecimiento de los límites del colectivo que toma las decisiones, necesitamos asegurarnos de que los mencionados procesos democráticos de votación, deliberación y experimentación son suficientemente inclusivos. Los procesos democráticos también deberían tener lugar a distintos niveles: dentro de las comunidades locales y de los Estados nación, pero preferiblemente también a nivel global. Dado el alcance global de la IA y de las tecnologías relacionadas, parece necesaria alguna forma de gobernanza global. Necesitamos también procedimientos para lidiar con las potenciales tensiones entre el bien común de distintos colectivos, por ejemplo, entre el bien común de las comunidades locales y el bien común definido en el contexto de un Estado nación o de la humanidad (sin embargo, no continuaré desarrollando aquí este punto).

COMUNICACIÓN Y COMUNIDAD

Otros conceptos relevantes «dirigidos hacia lo común» ya mencionados anteriormente en este libro son la comunicación y la comunidad. Están vinculados con la concepción relacional de la democracia defendida por los pensadores republicanos. Pero ¿qué significan y qué implican para el proyecto de la IA a favor del bien común?

Técnicamente, por lo normal, la comunicación se entiende en términos de enviar y recibir información, es decir, de intercambiarla. Esta es la noción dominante en los debates sobre «tecnologías de la información y de la comunicación», y en la cibernética está relacionada con el control y la (auto)regulación. Sin embargo, como ya he sugerido, existe otro significado de «comunicación» que se puede resumir en «crear lo comunitario» o «crear comunidad». En

lugar de centrarse en la cantidad de comunicación que, obviamente, producen las tecnologías digitales y cuyo principio puede contribuir al conocimiento básico de la democracia, deberíamos considerar la cualidad de la comunicación, donde la cualidad se evalúa en términos de la contribución a la creación de comunidad y al bien común. Desde el punto de vista de una república, crear relaciones comunicativas entre ciudadanos es importante en sí mismo, perspectiva bajo la cual la comunicación pasa a concernir a las personas y no tanto a la información y el control. Y, como señalé antes, podemos añadir con Arendt que es importante en la comunicación política democrática intentar comprender los puntos de vista de los demás y generar sentido común. Esto ofrece una perspectiva más relational tanto de la comunicación como de la democracia.

Sin embargo, desde una perspectiva democrática liberal, mencionar el término «comunidad» nos conduce hasta un campo de minas político-filosófico que emergió en la filosofía política contemporánea en los años 1980, cuando los denominados comunitaristas desafiaron al liberalismo clásico al poner el énfasis a la comunidad. Apelar a la comunidad plantea muchas cuestiones. ¿Cuál debería ser su rol en una sociedad democrática liberal, si es que es sostenible un sistema así? ¿Cuál debería ser la relación o el equilibrio entre el individuo y el colectivo, la persona y la comunidad? ¿El énfasis en la comunidad conduce necesariamente a los sistemas tradicionales o autoritarios? ¿Cuál puede y debe ser el rol de la comunidad en la modernidad?

Al igual que Aristóteles, los comunitaristas como MacIntyre, Sandel y Taylor defienden que la comunidad es lo principal y el individuo lo secundario. Charles Taylor, por ejemplo, defendió que una sociedad liberal solo funciona si los ciudadanos tienen un sentido de comunidad basado en su concepción compartida del bien común. Esto implica, según él, que los ciudadanos necesitan alguna forma de vida en común que sea considerada «un bien

supremamente importante» (1985: 213). Pero ¿esta forma de pensar conduce a adoptar una forma de colectivismo, quizás el tipo de colectivismo que vemos en algunos regímenes autoritarios o totalitarios? ¿Cuáles son los derechos y libertades de los individuos dentro de un orden político comunitario? Los liberales están preocupados con respecto a este punto, y con razón, pero, con Taylor y otros, se puede igualmente preguntar, dado que somos seres sociales y relacionales, hasta qué punto es sostenible una democracia liberal *sin* otorgar ningún papel a la comunidad. En la *Política*, Aristóteles escribe de manera célebre sobre los «animales políticos»: somos el tipo de ser que hace (necesita hacer) cosas juntos y vive (necesita vivir) en comunidades. De acuerdo con él, esto es importante para la política. Desde esta perspectiva, la política no tiene que ver tanto con los individuos como con el bien de la *polis* y con lo que tenemos en común. Necesitamos la *polis* y la *polis* nos necesita: somos seres políticos en este sentido.

Visto bajo la luz de la historia, y teniendo en cuenta las cosmovisiones no occidentales, esta forma de pensar es más corriente que excepcional: en cualquier caso, es la forma occidental de individualismo lo que es excepcional. Muchas otras culturas y civilizaciones también tienen y tuvieron concepciones relacionales de los seres humanos y la política. Esta es una dirección atractiva, pero también supone desafíos para las democracias liberales. Por una parte, existe mucha sabiduría en esta forma de pensamiento relational con respecto a los humanos y el vivir conjuntamente, y me parece que una democracia que rechace la relacionalidad social y política básica de los humanos no es viable. Por otra parte, es importante integrar estas nociones de comunidad y relacionalidad en el pensamiento político moderno de formas que, al mismo tiempo, respeten la libertad y los derechos individuales. Esto también plantea la importante y desafiante pregunta de cómo son y deberían estar conectados los derechos individuales y el bien común. Por ejemplo, con Pettit podríamos defender que esforzarse por el bien común

debería estar constreñido por el principio de libertad y la no dominación. Sin embargo, eso no resuelve el problema de hasta qué punto son importantes las libertades y derechos individuales frente al bien común. Quizás la cuestión misma no se pueda responder por completo por reyes-filósofos, sino que debería someterse a la deliberación democrática en situaciones particulares, dado que las respuestas y equilibrios específicos pueden diferir entre contextos político-culturales distintos y cuando se confrontan con problemas colectivos diferentes.

Este no es el lugar para embarcarse en una larga discusión sobre el liberalismo frente al comunitarismo (ni tampoco sobre el pensamiento relacional). Sin embargo, permítaseme señalar que la tradición republicana que he estado dibujando en este libro tiene claramente una dimensión comunitaria. El bien común es bueno para *nosotros*, esto es, para una comunidad política dada. Podemos discutir la naturaleza y el límite de esta comunidad, preguntarnos cuán sustancial tiene que ser y hasta qué punto poner el énfasis en ella resulta compatible con la libertad del individuo, pero deberíamos reconocer, al menos, que, de acuerdo con esta dirección de pensamiento republicano, la comunidad no es necesariamente antidemocrática. Más bien, la idea republicana consiste en que la política y la democracia son *también* como mínimo cuestión de comunidad política. Y me parece que esa comunidad tiene que compartir *algunos* valores y metas para funcionar en tanto que comunidad política y, desde luego, como una democracia, aunque, quizás, no necesariamente todos los valores y metas (esto plantea de nuevo el problema del pluralismo y de la diversidad). En cualquier caso, en una democracia tal y en un contexto pluralista moderno, el bien común puede funcionar como una meta general y una orientación normativa que organice la comunidad, y facilite y limite la cooperación, la deliberación y la experimentación democrática.

Afirmar que la IA debería promocionar el bien común implica defender que contribuya en algún grado a la creación de comunidad,

o al menos que la tenga como ideal u objetivo; sin embargo, y en contraste con las nociones tradicionales y conservadoras de la idea de comunidad, aquí esta no implica oponerse a principios ilustrados modernos, como la libertad y los derechos individuales, sino más bien ser compatible o, incluso, estar integrada en ellos (independientemente de lo desafiante que este proyecto pueda llegar a ser). Además, la comunidad de la que se habla aquí no está totalmente dada (incluso si existen tradiciones) ni disponible de antemano ni ligada a una identidad fija. Está dada parcialmente, pero también necesita crearse y, mientras está en el proceso de creación (y quizás siempre lo esté, es un proceso continuo), opera al mismo tiempo como un ideal regulador junto con la noción del bien común. Pero la comunidad y el bien común nunca son simplemente ideales abstractos: están encarnados en las prácticas, en las formas de cooperación y en las formas de vida. Sin embargo, no son solo el punto de partida, sino también un resultado. Son también cuestión de ensayar cosas e improvisar frente a problemas y situaciones particulares. Tienen una dimensión abierta y experimental.

Así, al igual que la IA, la propia comunidad democrática republicana es un experimento. No se conoce de antemano cómo puede ser una comunidad así. Y continúa habiendo desafíos serios. Por ejemplo, es un desafío reunir el bien común y los derechos individuales, y es un desafío conectar las nociones tradicionales de la comunidad, mientras, al mismo tiempo, nos comprometemos a *crear democracia* en respuesta a la situación actual (que está conformada en gran medida por la tecnología). Existen elementos tradicionales (por ejemplo, la sabiduría tradicional y las formas tradicionales de hacer las cosas), pero en una sociedad moderna el experimento no se limita a estos elementos, ni está constreñido por ellos. En la democracia que tengo en mente, la comunidad y la comunicación deben permanecer abiertas y permitir la pluralidad.

Vista desde esta perspectiva, la IA necesita ser una «tecnología de la comunicación» no solo en el sentido de lidiar con los datos y

ofrecernos información con respecto a los patrones que hay en estos datos, sino también en el sentido de trabajar por el bien común y ayudar a crear (más) comunidad como parte de crear democracia. Lo que esto implique en la práctica no está claro de antemano ni en abstracto, pero se necesita que se *haga* y se experimente con ello en contextos particulares y en respuesta a problemas sociales específicos.

Un factor que es importante a la hora de crear una comunidad y una democracia de este género es el conocimiento y la experiencia comunes. La IA y las tecnologías digitales pueden ayudarnos a la creación de *comunes epistémicos y experienciales* y, al contribuir al conocimiento y a la experiencia comunes, a que nos unamos. Esto debería adoptar siempre la forma de una ayuda y no de una sustitución, ya que la sabiduría humana, es decir, el conocimiento humano arraigado en experiencias concretas, relacionales y comunitarias en tanto que seres políticos en un determinado contexto, también es necesaria. Sin embargo, si se integran bien en procesos epistémicos y comunicativos pertinentes, la IA y las tecnologías digitales, en principio, complementan o nos ayudan a construir esa base de conocimiento y experiencia.

Sin embargo, actualmente la IA y las tecnologías digitales (en tanto que están integradas en formas específicas de distribución de poder, como el capitalismo del *laissez-faire* y el cuasiautoritarismo de las grandes corporaciones tecnológicas) parecen más bien apoyar un sistema que busca el bien de algunos individuos (algunos inversores tecnológicos, propietarios de tecnología y sus compañías) en lugar del bien común. Un sistema que privatiza tanto los recursos como los beneficios de lo que podría considerarse común, compartido y *comunicado*. En lugar de crear más comunidad y enriquecer los comunes, las tecnologías parecen dividir a la gente y resultan en menos confianza, menos repartimiento y menos bien común. Esto no quiere decir que no exista un bien común o que no haya comunicación en absoluto, lo cierto es que el sistema actual los

permite únicamente hasta cierto punto y dentro de una comunidad más bien pequeña de propietarios millonarios de compañías tecnológicas (y algunos de sus empleados y profetas transhumanistas). Lo que es o deja de ser el bien común en el contexto de la IA y de las tecnologías digitales se ve definido en la actualidad por gente como Elon Musk, Sam Altman y Mark Zuckerberg antes que de maneras democráticas, y, a menudo, el acuerdo resultante no esconde sino la mera y despiadada persecución de intereses particulares. La categoría de quienes se benefician de la IA también está limitada socioeconómica, cultural y geográficamente. Necesitamos más conocimiento y experiencia común que atraviese estas divisiones, y necesitamos usar y desarrollar la IA de forma que contribuya al bien común y a la construcción de comunidad en muchos lugares y contextos. Esto exige la regulación de la IA y de los datos de una manera que los haga más comunes, en oposición a privados, y, más ampliamente, una política de la tecnología y, desde luego, una economía que reorganicen de forma significativa y con valentía cómo se desarrolla y utiliza la tecnología en la sociedad para que se dirija hacia el bien común y la comunidad. Un cambio tan sustancial necesita liderazgo político y elaboración de políticas democráticas en instituciones de nivel nacional e internacional, pero las iniciativas de abajo arriba también necesitan más apoyo. Ideas como la plataforma de cooperativismo (Scholz, 2017), por ejemplo, que implica que los trabajadores, en vez de los capitalistas, posean plataformas digitales, puede suponer una intervención interesante y útil en la forma actual del capitalismo.

Para concluir, considero la comunicación no solo como un procedimiento técnico o como parte de una cibernética (política), sino como la creación del bien común y de comunidad. El liberalismo necesita estar, como mínimo, suplementado con, si no transformado por, algún reconocimiento de la relacionalidad y de la comunidad implicadas en la política y en los seres humanos de los que depende,

mientras mantiene (otros) valores democráticos importantes como el pluralismo, la libertad y la tolerancia. Debería pedirnos que nos esforcemos por realizar y crear el bien común a la vez que reconozca que podemos (y deberíamos tener permitido) disentir en cosas; y toda vez que mantenemos la libertad y la no dominación, uno de los principios sobre los que se construye la democracia liberal. Por tanto, la IA necesita llegar a ser una tecnología de comunicación completamente democrática también en este sentido: que contribuya a la construcción de comunidad, y esté guiada por la sabiduría y la experiencia (incluyendo el conocimiento tradicional que sea parte de la comunidad y forma de vida existente), a la vez que respeta los principios democrático-liberales básicos como la libertad y la no dominación. La IA puede y debe ayudar a la comunicación en el sentido de crear comunidad y debería estar dirigida hacia el bien común, pero esa comunicación (entendida en este sentido más profundo), ha de estar sujeta a límites normativos y epistémicos: debería estar guiada por los principios básicos liberaldemócratas y el entorno epistémico debería estar regulado de forma que apoye lo común (en lugar de dividir), mantenga la diversidad y el pluralismo, aliente el experimento y la apertura, y evite tendencias antidemocráticas. Si la IA y las tecnologías digitales se usan y gobiernan de esta forma, conduciendo hacia nuevas experiencias comunes y creando bien común, no existirían herramientas de control (como en la cibernética autoritaria), sino tecnologías de *comunicación*.

Experiencia común y construcción de un mundo común

Finalmente, esta comprensión de la comunicación es como interpreto la afirmación de Hannah Arendt de que debemos crear y mantener un mundo común. Arendt defendía en *La condición*

humana (1998 [1958]) que la experiencia de compartir un mundo común con otros nos ayuda a desarrollar un sentido común. Así pues, «Mundo» se refiere aquí no solo a una colección de cosas (a pesar de que Arendt también reconocía que los artefactos humanos, por ejemplo, una mesa, nos pueden unir), sino también a la experiencia común: compartir un «mundo» en el sentido fenomenológico de la palabra, según se puede encontrar, por ejemplo, en las obras de Heidegger y Merleau-Ponty. Compartir la experiencia es significativo política y democráticamente. Al combinar el concepto de Arendt de compartir un mundo común con todo lo dicho hasta ahora con respecto a la comunicación, se podría afirmar que la comunicación en el sentido específico que propongo aquí es la creación de un mundo común a través de la creación de *bienes experienciales comunes*, lugares de encuentro experiencial donde se unen diferentes perspectivas y experiencias, y donde las propias perspectivas y la experiencia subjetiva no son solo trascendentes de manera imaginativa, sino también comunicativa.

La cuestión es que, al igual que es imposible tener una forma de democracia resiliente sin trabajar por la meta del bien común (y descubrir juntos lo que es y cómo se puede alcanzar) y sin tener *algún* grado de comunidad, es imposible tener un sistema democrático sostenible sin (crear) un mundo común: sin crear, compartir y usar conocimiento y experiencia comunes, y sin trascender la perspectiva y experiencia propias.

Con respecto a la IA y a otras tecnologías digitales, esto implica que deberíamos usar y desarrollar las tecnologías y los medios de comunicación de tal forma que no conduzcan a una situación en la que cada uno tenemos nuestro propio mundo (o donde cada grupo tiene su propio mundo), y en donde tenemos dificultades a la hora de trascender nuestra propia perspectiva. Debemos asegurarnos de que ayudan a crear un mundo común.

Se trata de algo especialmente importante a la luz de los riesgos político-epistémicos relacionados con la IA. Ya he mencionado el

problema que reside en que su uso pueda conducir a la desinformación y a la polarización. Esta desintegración epistémica crea una falta de confianza, pero también divisiones sociales y políticas: conduce a la soledad política sobre la que ya se ha advertido, a la ausencia de comunidad y a estar encerrado en la propia experiencia. La democracia necesita experiencia común y un mundo común. Y ya que, como en el caso de la comunidad, un mundo común de este tipo nunca es algo simplemente dado, y no está dado, desde luego, en el mundo moderno, la democracia en tanto que *hacer* democracia tiene que ver con *crear* ese mundo común. La IA debería ayudarnos a crear ese mundo común, más que a dividirlo o a obstaculizar su creación.

¿Disponemos de una alternativa viable? Se puede rechazar el ideal de la democracia aquí articulado y tomar en consideración utilizar la IA en combinación con una forma débil de democracia (votación por mayoría y, desde luego, nuestra actual «democracia»). Se podría también imaginar una IA que use la noción utilitaria del bien común para gobernarnos: este guardián significaría el fin de la democracia. Sin embargo, si nos preocupa mantener y crear un tipo de democracia mejor y más rico, que sea menos vulnerable a las formas antidemocráticas de populismo y no conduzca a las pesadillas totalitarias, podemos y debemos hacerlo mejor: podemos y debemos imaginar y desarrollar la IA como una tecnología de *comunicación*, entendida en el sentido de comunicación desarrollado aquí, más rico, *político* y democrático.

Sin embargo, me preocupa, tal y como están las cosas, que estemos haciendo más bien lo contrario. Si la tradición filosófica republicana está en lo cierto con respecto a las condiciones necesarias para la democracia, me temo que, con la IA según se está utilizando y desarrollando actualmente, estamos socavando nuestros sistemas democráticos. En última instancia, esto puede conducir a crear las condiciones para el totalitarismo: un sistema político al que solo o principalmente le interese el control y que

origine un mundo en el que las personas ya no tengan nada en común, no estén preparadas para considerar las perspectivas de los demás, no quieran ya trabajar por el bien común y no se *comuniquen*. La pérdida de esta comunicación y la necesidad de (re)construir un mundo común era la preocupación de Arendt tras la Segunda Guerra Mundial y tras los horrores del nazismo y el estalinismo. Hoy en día, en la situación política actual y bajo la luz de las posibilidades y peligros creados por la LA, también debería ser nuestra preocupación.

RESUMEN EJECUTIVO PARA RESPONSABLES POLÍTICOS

Dado que los usos actuales de la IA se encuentran a menudo vinculados con la manipulación y con el poder autoritario, ¿es buena para la democracia? La LA, según se está desarrollando y usando actualmente, entraña el riesgo de socavar los principios fundamentales y la base de conocimiento sobre la que se construyen nuestras democracias, y no contribuye al bien común. Pero si adoptamos una noción más rica, más republicana de la democracia y comprendemos mejor la dimensión política más amplia de la tecnología, se abren las siguientes vías hacia una IA más democrática:

1. Regular, pero también, y de forma más fundamental, reajustar el equilibrio de *poder* entre las corporaciones privadas y las instituciones democráticas; e integrar los valores y principios políticos en el *desarrollo de la tecnología y en los procesos de innovación*. Por ejemplo, que los parlamentos, en lugar de las empresas privadas, decidan qué límites éticos se deben establecer para la generación automática de textos y *chatbots*, introducir más tecnologías y datos de IA en los comunes, y hacer que sea un requisito legal (junto con incentivos como la certificación) el que las compañías evalúen el impacto ético y

político de sus nuevas tecnologías antes de que salgan al mercado.

2. Fortalecer las instituciones democráticas existentes para salvaguardar los principios y derechos democráticos fundamentales. Pero también crear *nuevas instituciones* que sean más deliberativas y participativas, facilitar y guiar la experimentación socio-tecnológica. Que estas nuevas instituciones sean apoyadas por la IA y por otras tecnologías digitales, y permitan la mediación entre los expertos y la toma de decisiones democráticas de forma más duradera, sistemática, proactiva y rápida. Por ejemplo, usar minipúblicos ayudados por IA para deliberar acerca de su uso y conformar consejos permanentes formados tanto por expertos como por legos que ofrezcan recomendaciones al Gobierno con respecto a la IA y a otras tecnologías digitales nuevas y emergentes, y que sean capaces de responder con rapidez a los nuevos desarrollos tecnológicos. Crear procedimientos legislativos acelerados para una regulación tecnológica más sensible y potencialmente experimental que pueda ajustarse de una forma más sencilla.
3. Invertir en educación y facilitar nuevas comunidades de aprendizaje interdisciplinar y democrático. No solo necesitamos cambios tecnológicos e institucionales, sino también cambios en la educación y en la cultura, quizás un nuevo humanismo digital y un nuevo Renacimiento y una nueva Ilustración. Por ejemplo, apoyar iniciativas de debate de abajo arriba, así como acerca del pirateo informático y su repercusión para la democracia, integrar la educación interdisciplinar en todos los niveles, tanto en disciplinas técnicas como en humanidades y ciencias sociales, además de invertir en la investigación académica y facilitar discusiones públicas orientadas a repensar el humanismo bajo la luz de la IA y las tecnologías digitales.

4. Aspirar al bien común y fortalecer las condiciones para la democracia centrándose no solo en el control, sino también, y especialmente, en facilitar la comunicación, entendida como creación de comunidad y construcción de un mundo común. Considérese, por ejemplo, la creación de políticas tecnológicas a largo plazo que busquen el bien común, el desarrollo tecnológico cooperativo y de medidas técnicas que vayan contra las burbujas epistémicas, la polarización, la desconfianza y el odio. Habría que propiciar que la educación no solo transfiera conocimiento científico y competencias lingüísticas, sino que también permita desarrollar capacidades críticas frente a tecnologías digitales como la IA, y que forme a la gente para tener en cuenta las opiniones de los demás y discutir el futuro de la tecnología (y de cualquier otra materia) de una forma respetuosa y constructiva. Así, no solo se apoyaría a la democracia, sino que, en última instancia, se contribuiría a que viviéramos mejor juntos.

Agradecimientos

Agradezco a los revisores anónimos sus comentarios, pues me ayudaron a corregir y actualizar el borrador inicial de este manuscrito, y especialmente a mis editores, Mary Savigar y Stephanie Homer, por su apoyo continuo durante el proceso. También quiero agradecer a Zachary Storms su ayuda a la hora de darle forma a este libro, así como a todas las personas que, a lo largo de su confección, han sugerido bibliografía, se han sumado a la conversación y me han ofrecido su inestimable amistad y compañerismo.

Bibliografía^[*]

ALESSA, H., «The Role of Artificial Intelligence in Online Dispute Resolution: A Brief and Critical Overview», *Information & Communications Technology Law* 31(3), 2022, págs. 319-342.

Algoritmos de represión de China, «Human Rights Watch», 2019; disponible en: <<https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass>>.

ALLHUTTER, D., Cech, E, Fischer, E, Grill, G., y Mager, A., «Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics are Made Effective», *Frontiers in Big Data* 3(5), 2020; disponible en: <<https://www.frontiersin.org/articles/10.3389/fdata.2020.00005/full>>.

ARENDT, H., *Between Past and Future: Eight Exercises in Political Thought*, Londres, Penguin Books, 1968 [disponible en castellano como *Entre el pasado y el futuro: ocho ejercicios sobre la reflexión política*, Barcelona, Península, 2018].

—*The Human Condition*, 2.^a ed., Chicago University of Chicago Press, 1998 [1958] [disponible en castellano como *La condición humana*, Barcelona, Paidós Ibérica, 2005].

—*Eichmann in Jerusalem: A Report on the Banality of Evil*, Penguin, 2006 [1963] [disponible en castellano como *Eichmann en Jerusalén*, Barcelona, Debolsillo, 2006].

—*The Origins of Totalitarianism*, Penguin, 2017 [1951] [disponible en castellano como *Los orígenes del totalitarismo*, Madrid,

Alianza, 2006].

ARGUEDAS, A. R., ROBERTSON, C. T., FLETCHER, R., y NIELSEN, R. K., «Echo chambers, Filter Bubbles, and Polarisation: A Literature Review», Reuters Institute for the Study of Journalism, 2022; disponible en: <<https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review>>.

ARISTÓTELES, *Politics*, ed. de C. D. C. Reeve, Indianápolis, Hackett Publishing Company, 1998 [disponible en castellano como *Política*, Madrid, Gredos, 2022].

BACON, E., «New Atlantis», en Susan Bruce (ed.), *Three Early Modern Utopias*, Oxford, Oxford University Press, 1999 [1626] [disponible en castellano como *La nueva Atlántida*, Buenos Aires, Losada, 2017].

BARKER, Joshua, «Guerilla Engineers», en Sheila Jasanoff y Sang Hyun-Kim (eds.), *Dreamscapes of Modernity*, Chicago, University of Chicago Press, 2015, págs. 199-218.

BARTOLETTI, I., *An Artificial Revolution*, Toronto, Indigo, 2020.

BECKERT, J., *Imagined Futures*, Cambridge, MA, Harvard University Press, 2016.

BECKMAN, L., y Rosenberg, J. H., «Freedom as Non-Domination and Democratic Inclusion», *Res Publica* 24, 2017, págs. 181-198.

BERENDT, B., «AI for the Common Good?! Pitfalls, Challenges, and Ethics Pen-testing», *Paladyn, Journal of Behavioral Robotics* 10(1), 2019, 44-65; disponible en: <<https://doi.org/10.1515/pjbr-2019-0004>>.

BERES, L. R., «Trump's America: Anti-Intellectual and Proud of It». *YaleGlobal Online*, 17 de octubre de 2017; disponible en: <<https://archive-yaleglobal.yale.edu/content/trumps-america-anti-intellectual-and-proud-it>>.

BEZZAOUI, I., Fegert, J., y Weinhardt, C., «Distinguishing Between Truth and Fake: Using Explainable AI to Understand and Combat Online Disinformation», en Olga Levina, Lasse

- Benutzen y Samantha Papavasiliou (eds.), *Proceedings of the 16th International Conference on Digital Society (ICDS, 2022)*, International Academy Research and Industry Association (IARIA), 2022.
- BIMBER, B., y DE ZÚÑIGA, H. GIL, «The Unedited Public Sphere», *New Media & Society* 22(4), 2020, 700-715.
- BLUM, C., «Determining the Common Good: A (Re-) Constructive Critique of the Proceduralist Paradigm», *Phenomenology and Mind* 3, 2016, págs. 140-148.
- BOSTROM, N., *Superintelligence*, Oxford, Oxford University Press, 2014 [disponible en castellano como *Superinteligencia*, Teell, 2016].
- BRAIDOTTI, R., *Posthuman Knowledge*, Polity Press, 2019 [disponible en castellano como *El conocimiento posthumano*, Cambridge, Gedisa, 2020].
- BREVINI, B., *Is AI Good for the Planet?*, Cambridge, Polity Press, 2022.
- BROWN, L., «Using Technology to Defeat Democracy», en James P. Steyer (ed.), *Which Side of History? How Technology is Reshaping Democracy and Our Lives*, Common Sense Media, 2020.
- CADWALLADR, C., y GRAHAM-HARRISON, E., «Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach», *Guardian*, 17 de marzo de 2018; disponible en: <<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>>.
- Cámara de los Lores, «AI in the UK: Ready, Willing and Able?», *Informe del Comité Especial sobre Inteligencia Artificial de la sesión 2017-2019*, 2018; disponible en: <<https://publicadons-parliament-uk/pa/ld201719/ldselect/ldai/100/100.pdf>>.
- COECKELBERGH, M., *Green Leviathan or the Poetics of Political Liberty: Navigating Freedom in the Age of Climate Change and Artificial*

- Intelligence*, Londres, Routledge, 2021.
- *The Political Philosophy of AI: An Introduction*, Cambridge, Polity Press, 2022a [disponible en castellano como *Filosofía política de la inteligencia artificial*, Madrid, Cátedra, 2023].
- «Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence», *AI & Ethics*, 2022b; disponible en: <<https://doi.org/10.1007/s43681-022-00239-4>>.
- COHEN, J., «An Epistemic Conception of Democracy», *Ethics* 97(1), 1986, págs. 26-38.
- «Truth and Public Reason», *Philosophy and Public Affairs* 37(1), 2009, 1-42.
- COLEMAN, S., *Can the Internet Strengthen Democracy?*, Cambridge, Polity Press, 2017.
- COULDRY, N., y MEJIAS, U. A., *The Cost of Connection*, Stanford, Stanford University Press, 2020.
- Consejo de Europa, «AI and Control of Covid-19 Coronavirus», 2020; disponible en: <<https://www.coe.int/en/web/artificial-intelligence/ai-and-control-of-covid-19-coronavirus>> [en castellano como «La IA y el control del coronavirus Covid-19»; disponible en: <<https://www.coe.int/en/web/artificial-intelligence/la-ia-y-el-control-del-coronavirus-covid-19>>.]
- CRAWFORD, K., *Atlas of AI*, Yale University Press, 2021 [disponible en castellano como *Atlas de IA*, Barcelona, Nuevos emprendimientos editoriales, 2023].
- Democracy Report 2022, V-Dem Institute, 2022; disponible en: <https://v-dem-net/media/publications/dr_2022.pdf>.
- DEWEY, J., *Democracy and Education*, Los Angeles, The Free Press, 1944 [disponible en castellano como *Democracia y educación*, Las Rozas (Madrid), Morata, 2004].
- DOBBE, T. *et al.*, «Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?», *International Journal of Press!Politics* 26(1), 2021, págs. 69-91.

- DONALDSON, S., y Kymlicka, W, *Zoopolis: A Political Theorey of Animal Rights*, Oxford, Oxford University Press, 2011 [disponible en castellano como *Zoopolis, una revolución animalista*, Madrid, Errata naturae, 2018].
- DREYFUS, H., *What Computers Cant Do*, Cambridge, MA, MIT Press, 1972.
- DUBERRY, J., *Artificial Intelligence and Democracy: Risks and Promises of AI-Mediated Citizen-Government Relations*, Chaltherham, UK, Northangton, MA, Edward Elgar Publishing, 2022.
- DULONG DE ROSNAY, M. y STALDER, E, «Digital Commons», *Internet Policy Review* 9(4), 2020; disponible en: <<https://doi.org/10.14763/2020.4.1530>>.
- DURKHEIM, E., *Suicide: A Study in Sociology* trad, de J. A. Spaulding y G. Simpson, Los Angeles, The Free Press, 1997 [1897] [disponible en castellano como *El suicidio*, Tres Cantos (Madrid), Akal, 2012].
- DYER-WITHEFORD, N., MIKKOLA KJOSEN, A., y STEINHOFF, J., *Inhuman Power*, Londres, Pluto Press, 2019.
- ERLANGER, S., «Democracy under Siege from Autocrats, Social Media and Its Own Failures», *New York Times*, 27 de mayo de 2022; disponible en: <<https://www.nytimes.com/2022/09/27/world/europe/democracy-politics-threats.html>>.
- ESTLUND, D., «Epistemic Proceduralism and Democratic Authority», en *Does Truth Matter?*, Nueva York, Springer, 2009, págs. 15-27.
- FEENBERG, A., y Hannay, A., *Technology and the Politics of Knowledge*, Bloomington, Indiana University Press, 1995.
- FELDSTEIN, S., «The Global Expansion of AI Surveillance», Carnegie Endowment for International Peace, 17 de septiembre de 2019a; disponible en:

- <<https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847>>.
- «The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression», *Journal of Democracy* 30(1), 2019b, págs. 40-52.
- FISCHLI, R., «Data-owning Democracy: Citizen Empowerment through Data Ownership», *European Journal of Political Theory*, 2022; disponible en: <<https://doi.org/10.1177/14748851221110316>>.
- FLETCHER, R., «The Truth behind Filter Bubbles: Bursting Some Myths», Reuters Institute for the Study of Journalism, 24 de enero de 2020; disponible en: <<https://reutersinstitute.politics.ox.ac.uk/news/truth-behind-filter-bubbles-bursting-some-myths>>.
- FLORIDI, L., y COWLS, J., «A Unified Framework of Five Principles for AI in Society», *Harvard Data Science Review*, 2019; disponible en: <<https://doi.org/10.1162/99608f92.8cv550dl>>.
- FOUCAULT, M., *Power! Knowledge: Selected Interviews and Other Writings, 1972-1979*, ed. de C. Gordon, Nueva York, Pantheon, 1980.
- FRIEDMAN, B., y HENDRY, D. G., *Value Sensitive Design: Shaping Technology with Moral Imagination*, Cambridge, MA, MIT Press, 2019.
- FUCHS, C., *Digital Democracy and the Digital Public Sphere*, Londres Routledge, 2023.
- GARNER, R., *A Theory of Justice for Animals: Animal Rights in a Nonideal World*, Oxford, Oxford University Press, 2013.
- GLOWACKA, D., YOUNGS, R., PINTEA, A., y WOEOSIK, E., «Digital Technologies as a Means of Repression and Social Control», Estudio encargado por el DROI subcomité del Parlamento Europeo, 2021; disponible en: <[https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653636/EXPO_STU\(2021\)653636_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653636/EXPO_STU(2021)653636_EN.pdf)>.

- GRAEBER, D., y Wengrow, D., *The Dawn of Everything: A New History of Humanity*, Londres, Penguin/Allen Lane, 2021 [disponible en castellano como *El amanecer de todo: una nueva historia de la humanidad*, Barcelona, Ariel, 2022].
- GREENE, L., *Silicon States*, Berkeley, Counterpoint Press, 2018.
- Grupo Europeo de Ética en la Ciencia y las Nuevas Tecnologías (GEE), *Opinion on Democracy in the Digital Age*, Comisión Europea, junio de 2023; disponible en: <https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/8bl1albc-0f21-l1ee-bl2e-01aa75ed71a1>.
- GUERRERO, A. A., «Against Elections: The Lottocratic Alternative», *Philosophy & Public Affairs* 42(2), 2014, págs. 135-178.
- GUNKEL, D., *Robot Rights*, Cambridge, MA, MIT Press, 2018.
- GUNN, H., y LYNCH, M. P., «The Internet and Epistemic Agency», en Jennifer Lackey (ed.), *Applied Epistemology*, Oxford, Oxford University Press, 2021, págs. 389-409.
- HABERMAS, J., *The Theory of Communicative Action*, vol. 1, trad. de T. McCarthy Bacon, 1984 [disponible en castellano como *Teoría de la acción comunicativa*, Madrid, Cátedra, 2011].
- Moral Consciousness and Communicative Action*, trad. de C. Lenhart y S. W. Nicholson, Cambridge, MA, MIT Press, 1990 [disponible en castellano como *Conciencia moral y acción comunicativa*, Madrid, Trotta, 2008].
- The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*, Cambridge, MA, MIT Press, 1991.
- HARARI, Y., *Sapiens: A Brief History of Humankind*, Nueva York, Harper, 2015 [disponible en castellano como *Sapiens: una breve historia de la humanidad*, Barcelona, Debolsillo, 2023].
- HARAWAY, D., «A Cyborg Manifesto: Science, Technology, and Socialist-feminism in the Late Twentieth Century», en *Simians*,

- Cyborgs and Women: The Reinvention of Nature*, Londres, Routledge, 1991, págs. 149-181.
- HARDIN, G., «The Tragedy of the Commons», *Science* 162, 1968, 1243-1248. Harford, T, «The Earliest Accountants Were the True Authors of Writing», *Accounting and Business*, 2017; disponible en: <<https://www.accaglobal.com/my/en/member/member/accounting-business/2017/10/insights/cuneiform-writing.html>>.
- HAWLEY, J., *The Tyranny of Big Tech*, Washington D.C., Regnery Publishing, 2021.
- HEDLUND, M., y PERSSON, E., «Expert Responsibility in AI Development», *AI & Society*, 2022; disponible en <<https://doi.org/10.1007/s00146-022-01498-9>>.
- HELBING, D. *et al.*, «Will Democracy Survive Big Data and Artificial Intelligence?», *Scientific American*, 25 de febrero de 2017; disponible en: <<https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/>>.
- HELMUS, T, *Artificial Intelligence, Deepfakes and Disinformation: A Primer*, 2022; disponible en: <www.jstor.org/stable/resrep42027>.
- HENLEY, J., y BOOTH, R., «Welfare Surveillance System Violates Human Rights, Dutch Court Rules», *Guardian*, 5 de febrero de 2020; disponible en: <<https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules>>.
- HIMMELREICH, J., «Against Democratizing AI», *AI & Society*, 2022; disponible en: <<https://doi.org/10.1007/s00146-021-010357-z>>.
- HOSSEINMARDI, H., «It's Not Just Social Media: Cable News Has Bigger Effect on Polarization», *Ars Technica*, 10 de agosto de 2022; disponible en: <<https://arstechnica.com/science/2022/08/its->

not-just-social-media-cable-news-has-bigger-effect-on-polarization/>.

HR-inform (n. d.), «Uber Sued over “Robo-firing” Algorithms»; disponible en: <<https://www.hr-inform.co.uk/news-article/uber-sued-over-robo-firing-algorithms>>.

HUGHES, J., «The Politics of Transhumanism», 2001; disponible en: <http://www.changesurfer.com/Acad/TranshumPolitics.htm#_ftn2>.

HUSSAIN, W, «The Common Good», en Edward N. Zalta y Uri Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*, primavera de 2018; disponible en: <<https://plato.stanford.edu/archives/spr2018/entries/common->

KANT, I., «An Answer to the Question: What is Enlightenment?», en Pauline Kleingeld (ed.), *Toward Perpetual Peace and Other Writings on Politics, Peace, and History*, New Haven, CT, Yale University Press, 2006 [1784], págs. 17-23 [disponible en castellano como *Contestación a la pregunta: ¿Qué es la Ilustración?*, Taurus, 2012].

KAPLAN, J., *Artificial Intelligence: What Everyone Needs to Know*, Oxford University Press, 2016 [disponible en castellano como *Inteligencia artificial: lo que todo el mundo debe saber*, Teell, 2017].

KASPERSEN, A., y WALLACH, W, «Now is the Moment for a Systemic Reset of AI and Technology Governance», *Carnegie Council*, 24 de enero de 2023; disponible en: <<https://www.carnegiecouncil.org/media/article/systemic-reset-ai-technology-governance>>.

KATYAL, S. K., «Democracy & Distrust in an Era of Artificial Intelligence», *Daedalus*, 151(2), 2022, 322-334; disponible en: <<https://www.jstor.org/stable/10.2307/48662045>>.

KEMENY, R., «Brazil is Sliding into Techno-authoritarianism», *MIT Technology Review*, 19 de agosto de 2020; disponible en: <<https://www.technologyre>

view.com/2020/08/19/1007094/brazil-bolsonaro-data-privacy-cadastro-base/>.

KEMP, S., «Facebook Statistics and Trends», *Datareportal*, 19 de febrero de 2023; disponible en: <<https://datareportal.com/essential-facebook-stats>>.

KENDALL-TAYLOR, A., FRANTZ, E., y WRIGHT, J., «The Digital Dictators: How Technology Strenghtens Autocracy», *Foreign Affairs*, marzo-abril de 2020; disponible en: <<https://www.foreignaffairs.com/articles/china/2020-02-06/digital-dictators>>.

KITCHER, P., *Science in a Democratic Society*, Búfalo, NY, Prometheus Books, 2011.

KÖNIG, P. D., y Wenzelburger, G., «Opportunity for Renewal or Disruptive Force? How Artificial Intelligence Alters Democratic Politics», *Government information Quarterly* 37(1), 2020, 1-11.

KOSTER, R., *et al.*, «Human-centred Mechanism Design with Democratic AI», *Nature Human Behaviour*, 2022; disponible en: <<https://doi.org/10.1038/s41562-022-01383-x>>.

LANDEMORE, H., «Beyond the Fact of Disagreement? The Epistemic Turn in Deliberative Democracy», *Social Epistemology* 31(3), 2017, págs. 277-295.

— *Open Democracy*, Princeton, Princeton University Press, 2020.

LARSEN, J., MATTU, S., KIRCHNER, L., y ANGWIN, J., «How We Analyzed the COMPAS Recidivism Algorithm», *ProPublica*, mayo de 2023 [2016]; disponible en: <<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>>.

LATOUR, B., «Give Me a Laboratory and I Will Raise the World», en Karin D. Knorr-Cetina y Michael Mulkay (eds.), *Science Observed: Perspectives on the Social Studies of Science*, Housand Daks, CA, Sage, 1983.

— «Morality and Technology: The End of the Means», *Theory, Culture & Society* 19(5/6), 2002.

- Politics of Nature: How to Bring the Sciences into Democracy*, Cambridge, MA, Harvard University Press, 2004 [disponible en castellano como *Políticas de la naturaleza*, Barcelona, RBA, 2013].
- LE MASURIER, J., «“Our Democracy is under Threat”, Former Cambridge Analytica Employee Tells FRANCE 24», *FRANCE 24*, 28 de noviembre de 2019; disponible en: <<https://www.france24.com/en/20191128-interview-brittany-kaiser-cambridge-analytica-targeted-ads-facebook-democracy-under-threat>>.
- LESSIG, L., *The Future of Ideas: The Fate of the Commons in a Connected World*, Nueva York, Random House, 2001.
- LEVITSKY, S., y ZIBLATT, D., *How Democracies Die*, Nueva York, Crown, 2018 [disponible en castellano como *Cómo mueren las democracias*, Barcelona, Ariel, 2018].
- LIN, B., y LEWIS, S., «The One Thing Journalistic AI Just Might Do for Democracy», *Digital Journalism*, 2022; disponible en: <<https://doi.org/10.1080/21670811.2022.2084131>>.
- MARX, K., y ENGELS, E., *Marx and Engels’s «German Ideology Manuscripts»: Presentation and Analysis of the «Feuerbach Chapter»*, Londres, Palgrave Macmillan, 2014 [1864] [disponible en castellano como *La ideología alemana*, Tres Cantos (Madrid), Akal, 2014].
- MAZZUCATO, M., «Let’s Make Private Data into a Public Good», *MIT Technology Review*, 2018; disponible en: <<https://www.technologyreview.com/s/611489/lets-make-private-data-into-a-public-good/>>.
- MEEK, A., «Minnesota is Now Using Contact Tracing to Track Protestors, as Demonstrations Escalate», *BGR*, 30 de mayo de 2020; disponible en: <<https://bgr.com/lifestyle/minnesota-protest-contact-tracing-used-to-track-demonstrators/>>.
- MEIKLE, G., *Deepfakes*, Cambridge, Polity Press, 2022.

- MILL, J. S., *On Liberty*, Indianápolis, Hackett Publishing, 1978 [1859] [disponible en Castellano como *Sobre la libertad*, Madrid, Tecnos, 2008].
- MITTELSTADT, B., WACHTER, S., y RUSSELL, C., «The Unfairness of Fair Machine Learning: Levelling Down and Strict Egalitarianism by Default», 2023; disponible en: <<https://arxiv.org/ftp/arxiv/papers/2302/2302.02404.pdf>>.
- MOUFFE, C., *Agonistics: Thinking the World Politically*, Londres, Verso, 2013 [disponible en castellano como *Agonística: pensar el mundo políticamente*, México, Fondo de Cultura Económica, 2022].
- MOZUR, P, XIAO, M., y LIU, J., «“An Invisible Cage”: How China is Policing the Future», *The New York Times*, 25 de junio de 2022; disponible en: <<https://www.nytimes.com/2022/06/25/technology/china-surveillance-police.html>>.
- MULDOORN, J., «Data-Owning Democracy or Digital Socialism?», *Critical Review of International Social and Political Philosophy*, 2022, disponible en: <<https://doi.org/10.1080/13698230.2022.2120737>>.
- NEMITZ, P., «Constitutional Democracy and Technology in the Age of Artificial Intelligence», *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 376, 2018, págs. 1-14.
- NGUYEN, C. T, «Echo Chambers and Epistemic Bubbles», *Episteme* 17(2), 2020, págs. 141-161.
- Nida-Rümelin, J., «Digital Humanism and the Limits of Artificial Intelligence», en Hannes Werthner *et al.* (eds.), *Perspectives on Digital Humanism*, Nueva York, Springer, 2022.
- NUNES DA COSTA, M., *Democratic Despotisms*, Newcastle upon Tyne, Cambridge Scholars, 2022.
- O’NEIL, C., *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Nueva York, Broadway

- Books, 2016 [disponible en castellano como *Armas de destrucción matemática*, Madrid, Capitán Swing, 2018].
- OSTROM, E., *Governing the Commons: The Evolution of Institutions for Collective Action: The Political Economy of Institutions and Decisions*, Cambridge, Cambridge University Press, 1990 [disponible en castellano como *El gobierno de los bienes comunes: la evolución de las instituciones de acción colectiva*, México, Fondo de Cultura Económica, 2011].
- OWEN, R., MACNAGHTEN, P., y STILGOE, J., «Responsible Research and Innovation: From Science in Society to Science for Society, with Society», *Science and Public Policy* 39(6), 2012, págs. 751-760.
- PARISER, E., *The Filter Bubble*, Londres, Penguin, 2011 [disponible en castellano como *El filtro burbuja*, Madrid, Taurus, 2017].
- PETERS, J. D., «Communication: History of the Idea», *International Encyclopedia of Communication*, 2018; disponible en: <<https://onlinelibrary.wiley.com/doi/10.1002/9781405186407.wbiecc075>>.
- PETTIT, P., *Republicanism: A Theory of Freedom and Government*, Oxford, Clarendon Press, 1997 [disponible en castellano como *Republicanism: una teoría sobre la libertad y el gobierno*, Barcelona, Paidós Ibérica, 1999].
- «The Common Good», en Keith Dowding, Robert E. Goodin y Carole Pateman (eds.), *Justice and Democracy: Essays for Brian Barry*, Cambridge, Cambridge University Press, 2004, págs. 150-169.
- PLATÓN, *Republic*, en John M. Cooper (ed.), *Plato: Complete Works*, Hackett, Indianápolis, 1997, págs. 971-1223 [disponible en castellano como *República*, Madrid, Tecnos, 2019].
- POPPER, K., *The Open Society and its Enemies* (Vol. 1: *The Spell of Plato*), Londres, Routledge, 1995 [1945] [disponible en castellano como *La sociedad abierta y sus enemigos*, Barcelona, Paidós Ibérica, 2010].

- PORTER, J., «A Black Man Was Wrongfully Arrested because of Facial Recognition», *The Verge*, 24 de junio de 2020; disponible en: <<https://www.theverge.com/2020/6/24/21301759/facial-recognition-detroit-police-wrongful-arrest-robert-williams-artificial-intelligence>>.
- PUTNAM, R. D., *Bowling Alone: The Collapse and Revival of American Community*, rev. y actualizado edn., Nueva York, Simon & Schuster, 2020 [disponible en castellano como *Solo en la bolera: colapso y resurgimiento de la comunidad norteamericana*, Barcelona, Galaxia Gutenberg, 2002].
- PUYOL, A., *Political Fraternity: Democracy beyond Freedom and Equality*, Londres, Routledge, 2019.
- RECKE, M., «What is Digital Humanism?», NEXT Conference, 21 de noviembre de 2017; disponible en: <<https://next.conf.eu/2017/II/what-is-digital-humanism/#gref>>.
- REIJERS, W, ORGAD, L., y DE FILIPPI, P, «The Rise of Cybernetic Citizenship», *Citizenship Studies*, 2022; disponible en: <<https://doi.org/10.1080/13621025.2022.2077567>>.
- RENNIE, D., «An Uncertain Future», *Economist*, 15 de octubre de 2022; disponible en: <<https://www.economist.com/special-report/2022/10/10/for-western-democracies-the-price-of-avoiding-a-clash-with-china-is-rising>>.
- RISSE, M., *Political Theory of the Digital Age*, Cambridge, Cambridge University Press, 2023.
- ROBERTSON, H., y TRAVAGLIA, J., «Big Data Problems We Face Today Can be Traced to the Social Ordering Practices of the Nineteenth Century», 2015; disponible en: <<https://blogs.lse.ac.uk/impactofsocialsciences/2015/10/13/ideological-inheritances-in-the-data-revolution/>>.
- ROBINSON, S. C., «Trust, Transparency, and Openness: How Inclusion of Cultural Values Shapes Nordic National Public Policy Strategies for Artificial Intelligence (AI)», *Technology in Society*

- 63, 2020; disponible en: <[https:// doi.org/10.1016/j.techsoc.2020.101421](https://doi.org/10.1016/j.techsoc.2020.101421)>.
- ROUSSEAU, J.-J., «Of the Social Contract», en Victor Gourevitch (ed.), *The Social Contract and Other Later Political writings*, Cambridge, Cambridge University Press, 1997 [1762], págs. 39-152 [disponible en castellano como *El contrato social o principios de derecho político*, Madrid, Tecnos, 2007].
- SÆTRA H. S., «A Hobbesian Argument for World Government», *Philosophies* 7(3), 66, 2022.
- BORGE BUND, H., y COECKELBERGH, M., «Avoid Diluting Democracy by Algorithms», *Nature Machine Intelligence*, 2022; disponible en: <[https:// doi.org/10.1038/s42256-022-00537-w](https://doi.org/10.1038/s42256-022-00537-w)>.
- SANDERS, N. E., y SCHNEIER, B., «How ChatGPT Hijacks Democracy». *New York Times*, 15 de enero de 2023; disponible en: <<https://www.nytimes.com/2023/01/15/opinion/ai-chatgpt-lobbying-democracy.html>>.
- SAVAGET, P., CHIARINI, T., y EVANS, S., «Empowering Political Participation through Artificial Intelligence», *Science and Public Policy* 46(3), 2019, 369-380; disponible en: <<https://doi.org/10.1093/scipol/scy064>>.
- SCHOLZ, T., «"Platform Cooperativism" vs. the Sharing Economy», en Nicolas Douay y Annie Wan (eds.), *Big data & Civic Engagement*, Planum Publisher, 2017, págs. 47-52.
- SELLERS, M. N. S., «Republicanism: Philosophical Aspects», en Neil J. Smelser y Paul B. Baltes (eds.), *International Encyclopedia of the Social & Behavioral Sciences*, Amsterdam, Elsevier, 2015, págs. 477-82.
- SHABAZ, A., *The rise of Digital Authoritarianism*, Freedom House, 2018; disponible en: <<https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism>>.
- SINGER, P., *Animal Liberation*, Nueva York, HarperCollins, 2009 [1975] [disponible en castellano como *Liberación animal*, Madrid,

- Trotta, 1999].
- SKINNER, Q., «Freedom as the Absence of Arbitrary Power», en Cecile Laborde y John Maynor (eds.), *Republicanism and Political Theory*, Hoboken, NY, Blackwell Publishing, 2008, págs. 83-101.
- SMITH, R. E., «My Social Media Feeds Look Different from Yours and It's Driving Political Polarization», *USA Today*, 2 de septiembre de 2019; disponible en: <https://eu.usatoday.com/story/opinion/voices/2019/09/02/social-media-election-bias-algorithms-diversity-column/2121233001/>>.
- SPLICHAL, S., *Datafication of Public Opinion and the Public Sphere: How Extraction Replaced Expression of Opinion*, Londres, Anthem Press, 2022a.
- «In Data We (Don't) Trust: The Public Adrift in Data-Driven Public Opinion Models», *Big Data & Society*, enero-junio de 2022b, págs. 1-13.
- STARKE, C., BALEIS, J., KELLER, B., y MARCINKOWSKI, E., «Fairness Perceptions of Algorithmic Decision-making: A Systematic Review of the Empirical Literature», *Big Data & Society* 9(2), 2022; disponible en: <<https://doi.org/10.1177/20539517221115189>>.
- SUDMANN, A. (ed.), *The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms*, Bielefeld, Transcript Verlag, 2019.
- SUSSKIND, J., *The Digital Republic: On Freedom and Democracy in the 21st Century*, Londres, Bloomsbury, 2022.
- TANG, A., «Inside Taiwan's New Digital Democracy», *Economist*, 12 de mayo de 2019; disponible en: <<https://www.economist.com/open-future/2019/03/12/inside-taiwans-new-digital-democracy>>.
- TAYLOR, C., «Alternative Futures: Legitimacy, Identity and Alienation in Late Twentieth Century Canada», en Alan Cairns y Cynthia Williams (eds.), *Constitutionalism, Citizenship, and Society in*

- Canada, Toronto, University of Toronto Press, 1985, págs. 183-229.
- TOCQUEVILLE, A. de, *Democracy in America*, Indianápolis, Liberty Fund, 2012 [1835] [disponible en castellano como *La democracia en América*, Madrid, Trotta, 2010].
- TURING, A., «Computing Machinery and Intelligence», *Mind* 49(236), 433-1950, 460; disponible en: <<https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>>.
- UNESCO, «Recommendation on the Ethics of Artificial Intelligence», 2021; disponible en: <<https://unesdoc.unesco.org/ark:/48223/pf0000381137>>.
- VAN DEN HOVEN, J., «Value Sensitive Design and Responsible Innovation», en Richard Owen, John Bessant y Maggy Heintz (eds.), *Responsible Innovation*, Nueva York, John Wiley & Sons, 2013, págs. 75-83.
- VELIZ, C., *Privacy is Power*, Londres, Bantam Press, 2020 [disponible en castellano como *Privacidad es poder*, Barcelona, Debate, 2021].
- VERBEEK, P P, *Moralizing Technology*, Chicago, Chicago University Press, 2011. Vermuele, A., *Common Good Constitunioalism*, Cambridge, Polity Press, 2022. Voltaire, *Philosophical Dictionary*, Nueva York, Basic Books, 1962 [1764] [disponible en castellano como *Diccionario filosófico*, Tres Cantos (Madrid), Akal, 2007].
- VON SCHOMBERG, R., «A Vision of Responsible Research and Innovation», en Richard Owen, John Bessant y Maggy Heintz (eds.), *Responsible Innovation*, Hoboken, NY, Wiley, 2013, págs. 51-74.
- WARBURTON, N.J «Free Speech in the Age of the Internet», en *Free Speech: A Very Short Introduction*, Oxford, Oxford University Press, 2009, págs. 81-95; disponible en: <<https://doi.org/10.1093/actrade/9780199232352.003.0005>>.

- WERTHNER, H., PREM, E., LEE, E. A., y GHEZZI, C., *Perspectives on Digital Humanism*, Nueva York, Springer, 2022; disponible en: <<https://link.springer.com/book/10.1007/978-3-030-86144-5>>.
- WIENER, N.J. *Cybernetics: Or, Control and Communication in the Animal and the Machine*, Cambridge, MA, MIT Press, 1948 [disponible en castellano como *Cibernética: o el control y comunicación en animales y máquinas*, Barcelona, Tusquets editores, 1985].
- WINNER, L., «Do Artifacts Have Politics?», *The Whale and the Reactor: A Search for Limits in an Age of High Technology*, Chicago, University of Chicago Press, 1986, págs. 19-39.
- ZARKADAKIS, G., *Cyber Republic: Reinventing Democracy in the Age of Intelligent Machines*, Cambridge, MA, MIT Press, 2020.
- ZENG, J., «Artificial Intelligence and Chinas Authoritarian Governance», *International Affairs* 96(6), 2020, 1441-1459.
- ZUBOFF, S., «Big Other: Surveillance Capitalism and the Prospects of an Information Civilization», *Journal of Information Technology* 30(1), 2015, págs. 75-89.
- The Age of Surveillance Capitalism*, Profile Books, 2019 [disponible en castellano como *La era del capitalismo de la vigilancia*, Barcelona, Paidós Ibérica, 2020].



MARK COECKELBERGH (nacido en 1975) es un filósofo belga de la tecnología. Ph.D. de la Universidad de Birmingham. Es profesor de Filosofía de los Medios y la Tecnología en el Departamento de Filosofía de la Universidad de Viena y ex presidente de la Sociedad de Filosofía y Tecnología .

Obras:

Liberation and Passion: Reconstructing the Passion Perspective on Human Being and Freedom, 2002.

The Metaphysics of Autonomy: The Reconciliation of Ancient and Modern Ideals of the Person, 2004.

Imagination and Principles: An Essay on the Role of Imagination in Moral Reasoning, 2007.

Growing Moral Relations: Critique of Moral Status Ascription, 2012.

Human Being @ Risk: Enhancement, Technology, and the Evaluation of Vulnerability Transformations, 2013.

Environmental Skill: Motivation, Knowledge, and the Possibility of a Non-Romantic Environmental Ethics, 2015.

Money Machines: Electronic Financial Technologies, Distancing, and Responsibility in Global Finance, 2015.

New Romantic Cyborgs: Romanticism, Information Technology, and the End of the Machine, 2017.

Moved by Machines: Performance Metaphors and Philosophy of Technology, 2019.

Introduction to Philosophy of Technology, 2019.

AI Ethics, 2020.

Green Leviathan or the Poetics of Political Liberty, 2021.

Self-improvement: technologies of the soul in the age of artificial intelligence, 2022.

The political philosophy of AI: an introduction, 2022

Robot ethics, 2022.

Digital Technologies, Temporality, and the Politics of Co-Existence, 2023.

Why AI Undermines Democracy and What To Do About It, 2024.

Notas

Prefacio

[1] Véase, por ejemplo, en: <<https://www.bbc.com/news/uk-65746524>>. <<

[2] Disponible en: <<https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>>. <<

[3] En mayo de 2023, el CEO de Open AI, Sam Altman, visitó a reguladores europeos amenazando con dejar de operar en Europa si la UE regulaba demasiado la IA (y no en los términos que él planteaba); disponible en: <<https://techcrunch.com/2023/05/25/sam-altman-european-tour/>>. <<

[4] Publicada en 2023, la Carta abierta del Future of Life Institute (Instituto de la Vida Futura), y también firmada por Elon Musk, afirma que la IA avanzada puede cambiar la historia de la vida en la Tierra y llegar a estar fuera de control, y argumenta que, por lo tanto, deberíamos ralentizar el entrenamiento de sistemas de IA fuerte durante al menos seis meses; disponible en: <<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>>. <<

Capítulo 1. Introducción

[1] Disponible en: <www.wsj.com/articles/white-house-warns-of-risks-as-ai-use-takes-off-d4cc217f>. <<

[2] La página web hace referencia a un informe reciente sobre la democracia en la era digital realizado por el Grupo Europeo de Ética de la Ciencia de las Nuevas Tecnologías (2023); disponible en: <<https://euraxess.ec.europa.eu/worldwide/africa/news/democracy-peril-commissions-ethics-group-stresses-need-and-ways-deepen>>. <<

[3] Disponible en: <<https://www.npr.org/2021/09/16/10379023314/the-u-n-warns-that-ai-can-pose-a-threat-to-human-rights>>. <<

[4] Véase, por ejemplo, Risse (2023) y, de nuevo, la carta que pidió ralentizar el desarrollo de IA potentes. <<

[5] Por autoritarismo y totalitarismo digital me refiero a un sistema de organización social y (en términos más generales) de control que depende de las tecnologías digitales y que somete a las personas a la autoridad (autoritarismo) y ejerce un control total sobre ellas (totalitarismo). Este lo puede llevar a cabo el Gobierno, pero también actores corporativos. Sin embargo, el propósito final de este libro no es establecer y ocuparse del concepto académico, sino el de investigar cómo la IA pone en riesgo el hecho de que se socave la democracia (a menudo, de formas que no conducen necesariamente al autoritarismo o al totalitarismo) y de lo que se puede hacer al respecto. <<

Capítulo 2

[1] Para un buen resumen de esta cuestión, véase Harford (2017). <<

[2] También se podrían considerar aquí algunos transhumanistas. A pesar de que a menudo la creencia de estos últimos en la llegada de una tecnología que sobrepase a la humanidad vaya en direcciones elitistas y tecnocráticas, por ejemplo, en la forma del denominado «largoplacismo», algunos han defendido una versión democrática. Aquí es relevante la obra de James Hughes; mientras muchos otros transhumanistas adoptan el libertarismo, él defiende lo que denomina «transhumanismo democrático», que es liberal, social y democrático (Hughes, 2001). <<

Capítulo 3

[1] Disponible en: <https://en.wikipedia.org/wiki/Dartmouth_workshop>. <<

[2] En 1951, Alan Turing ya había publicado un artículo sobre un programa de ajedrez. En 1997, el superordenador de IBM Deep Blue ganó al entonces campeón del mundo de ajedrez Garry Kasparov. En 2016 Alphago de Google DeepMind venció a Lee Sedol en el juego de go. En 2022, una IA venció al ocho veces campeón del mundo de *bridge*. <<

[3] Sin embargo, Aristóteles defendió en la *Política* (2022: 1276b19-30) que, al igual que cada marinero es responsable de la seguridad del viaje, también cada ciudadano es responsable de la seguridad de la comunidad. <<

[4] Nótese que silenciar no solo y no necesariamente es un problema relacionado con tendencias autoritarias de derechas. Por ejemplo, Nunes Da Costa defiende que en el *wokeismo* (movimiento de quienes se mantienen alerta frente al racismo) actual, la cultura de la cancelación y la política identitaria, que emergieron en un contexto liberal democrático dando voz a preocupaciones legítimas en torno a la justicia social, tienden a convertirse tanto en

antidemocráticos como en no liberales cuando intentan silenciar a personas con las que están en desacuerdo, amenazando, así, las libertades democráticas fundamentales. No seguiré discutiendo aquí esta afirmación, pero obsérvese que existen políticas identitarias tanto de derechas como de izquierdas, y que, a pesar de que, ciertamente, existen intentos problemáticos de lo que se viene a llamar «censura», no siempre está claro lo que cuenta como «silenciar»; es importante distinguir entre la crítica abierta a ciertas opiniones y el silenciar. Lo primero es bienvenido en una democracia, lo segundo no. <<

[5] A pesar de que hasta hace poco la filosofía política ha sido más bien una cuestión antropocéntrica, la teoría política contemporánea ha estado explorando lo que podría implicar extender los límites de lo político más allá de lo humano. Tras la teoría de la liberación animal de Singer (2009 [1975]), que rechazó el «especieísmo» y que puede interpretarse como una teoría política, ha habido académicos de diferentes direcciones políticas que han intentado incluir a los animales. Por ejemplo, Donaldson y Kymlicka (2011) han discutido la ciudadanía para los animales, y Garner (2013) ha propuesto una teoría de la justicia para los animales. Los poshumanistas críticos y (otras) personas que han defendido los enfoques relacionales también han cuestionado el antropocentrismo de la teoría política tradicional. Algunos autores hasta han considerado incluir a las máquinas: Gunkel (2018), por ejemplo, ha explorado la cuestión de los derechos de los robots. <<

Capítulo 4

[1] A pesar de que entonces no se aplicaron estos principios a las mujeres, a la gente de color y a los nativos, el fracaso al aplicarlos de manera universal no disminuye su valor normativo para la democracia, que requiere su vigilancia universal y rigurosa (lo cual, a

su vez, y, dependiendo de la teoría de la justicia de cada uno, puede incluir una mayor consideración por las injusticias infligidas a grupos particulares). <<

[2] Puyol interpreta el principio de la diferencia de Rawls como una aplicación del principio de fraternidad. <<

[3] Sin embargo, nótese que para Arendt esto implicaba sopesar el propio juicio con los juicios posibles y no necesariamente reales de los otros, en particular, de la humanidad. <<

[4] Véase también mi trabajo sobre IA, democracia y agencia epistémica (Coeckelbergh, 2022b). <<

[5] Por ejemplo, en documentos de la OCDE y de la Comisión Europea. <<

[6] La epistemología de la virtud es un enfoque reciente de la epistemología que puede conectarse de forma fructífera con esta discusión acerca de la democracia y la IA. <<

Capítulo 5

[1] Esta falta de agencia epistémica tiene impacto, a su vez, sobre mi autonomía: si, sin mi conocimiento, se está ejerciendo una influencia sobre mí, careceré de autonomía para tomar una determinada decisión, algo que era crucial para los filósofos de la Ilustración. Ciertamente, la agencia epistémica y la autonomía deben entenderse de una forma relacional: existen muchas influencias en mis creencias debido a mi relación con el mundo y con los otros y, por lo tanto, no existe una agencia epistémica completa y no es posible una autonomía epistémica absoluta. Pero, si algún grado de autonomía epistémica no solo es deseable para la democracia, sino también necesario, entonces, si mis creencias políticas están

manipuladas por la IA, existe claramente un problema de agencia epistémica y de autonomía y, por tanto, un obstáculo para la democracia. <<

[2] Un derecho así sería difícil de defender por diferentes razones (véase Risse, 2023), a pesar de que uno tenga que disponer de protecciones para la agencia epistémica individual. <<

[3] Este tiende a ser el caso de los sistemas de aprendizaje profundo. <<

Capítulo 6

[1] Por ejemplo, con este «constitucionalismo del bien común» Vermuele (2022) ha propuesto una versión conservadora basada en el *ins commune*, en la que se fija lo que necesita una comunidad política floreciente. En línea con la dirección normativa propuesta en este libro, se prefiere un orden constitucional más abierto y dinámico. <<

[2] Considérense, por ejemplo, las violaciones a la privacidad o las cuestiones relacionadas con el *copyright* que surgen cuando se usa la IA para crear textos y obras de arte. Aquí debemos considerar el establecimiento de tribunales especiales que traten con casos tecnológicos. Ya existen muchos tribunales especiales, ¿por qué no añadir uno especializado en disputas que tengan que ver con las tecnologías digitales? Además, dado que los tribunales y los jueces humanos se hallan a menudo sobrecargados de trabajo, que el apoyo legal tiende a ser caro y que los procedimientos legales existentes para resolver los conflictos no siempre son muy efectivos, eficientes y fáciles de comprender para la mayoría de los ciudadanos, la IA y las propias tecnologías digitales también ofrecen oportunidades para formas alternativas de resolución de conflictos, por ejemplo, para la mediación y la resolución de litigios. Ya existen

sistemas de resolución de disputas *online* y mediante algoritmos que se presentan como alternativas a los tribunales de justicia. Esto relaja algo la carga de los tribunales y, más importante, ayuda a aumentar el acceso a soluciones legales para los ciudadanos. Si se hace bien, también podría aumentar la transparencia. Ambas cosas, el acceso a la ley y la transparencia, son buenas para la democracia. Pero también existe aquí el riesgo de la tecnocracia, en particular, la promoción de un tipo de enfoque algorítmico para aplicar la ley que no sea lo suficientemente sensible a las características únicas de cada caso y no logre tomar en consideración la experiencia personal de las personas. Todo lo cual ya es un problema en el sistema jurídico actual donde la gente puede verse tentada a abrirse camino a través de los casos de una forma «algorítmica», pero que puede verse agravada si el proceso se delega a las máquinas. Por ejemplo, Alessa (2022) ha argumentado que los sistemas basados en la IA no siempre tienen en cuenta las respuestas emocionales de los ciudadanos. Tales limitaciones socavan la confianza de estos en el sistema y, por tanto, pueden suponer un peligro para la democracia y el sistema legal como un todo y propiciar el éxito de la IA en la sociedad. Si se usa en los tribunales, los jueces pueden depender demasiado de la automatización, sin aplicar su propio juicio. Existe también el problema de que la parte jurídica del sistema se vuelva demasiado poderosa en relación con los cuerpos elegidos democráticamente. Este siempre ha sido un peligro en las democracias, pero en ausencia de un grado suficiente de regulación tecnológica, los problemas que suscita la IA es probable que acaben en el escritorio de un juez antes que en el debate entre los miembros de un cuerpo democrático. <<

[3] Disponible en: <<https://www.whitehouse.gov/ostp/ai-bill-of-rights/>>. <<

[4] Disponible en: <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX: 52021PC0206>>. <<

[5] Disponible en: <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex% 3A52021PC0206>>. <<

[6] Disponible en: <<https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>>. <<

[7] Disponible en: <<https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196>>. <<

[8] Filosóficamente, un enfoque más global de este tipo puede sustentarse tanto en argumentos hobbesianos (Ssetra, 2022) como no hobbesianos (Coeckelbergh, 2021). Hobbes justificó el Gobierno de forma célebre diciendo que, de lo contrario, la gente se encontraría en un estado de naturaleza que es desagradable y brutal. Para evitar el conflicto, se necesita una autoridad estatal que elimina la libertad de las personas, pero proporciona seguridad y paz. De forma similar, es posible defender que nos encontramos actualmente en un estado de naturaleza con respecto a la regulación tecnológica a nivel global: cada Estado nación hace lo que quiere. Esto crea situaciones peligrosas y arriesgadas con respecto a los efectos éticos y políticos de la tecnología. En respuesta a esto se podría defender que necesitamos una autoridad global que pueda implementar una regulación efectiva. Sin embargo, el problema con el que se encuentra el enfoque hobbesiano es el del autoritarismo. Necesitamos un sistema democrático de gobierno global de la tecnología. En mi libro *Green Leviathan* (2021), discuto esta tensión entre la libertad y la autoridad/autoritarismo en el contexto del

cambio climático y la IA, y de formas que llevan la formulación del problema más allá de Hobbes. <<

[9] Véase UNESCO (2021). <<

[10] La agenda se actualiza con regularidad y se encuentra disponible en: <<https://www.coe.int/en/web/artificial-intelligence/work-in-progress#01EN>>. <<

[11] Recientemente, el término «democracia por diseño» también se ha propuesto por el Grupo Europeo sobre la Ética en las Ciencias y las Nuevas Tecnologías en su informe sobre la democracia en la era digital (2023). <<

[12] En los años 1960, Hofstadter se quejó del antiintelectualismo en los Estados Unidos. En la América (¿pos-?) Trump, la ignorancia y el antiintelectualismo impregnan la vida pública (Beres, 2017). <<

Capítulo 7

[1] Téngase en cuenta que el humanismo digital no debería confundirse con las humanidades digitales: esto último es el estudio de las humanidades con métodos digitales. Las humanidades digitales pueden, por supuesto, ser parte del proyecto del humanismo digital, por ejemplo, ayudando a conectar a las humanidades con las disciplinas técnicas. <<

[2] No hay solo humanos en este planeta, y los no humanos también tienen intereses. El enfoque humanista clásico necesita complementarse con una sensibilidad y consciencia poshumanistas hacia una perspectiva más holista, ecológica y planetaria. De forma similar, los ideales de la Ilustración no deberían proteger solo a los humanos. A pesar de que este proyecto de extender la política y la democracia hasta los no humanos plantea claramente desafíos

filosóficos (por ejemplo, los animales no humanos no pueden participar en la deliberación), al menos deberíamos explorar qué significan la libertad, igualdad, justicia y democracia una vez que seamos completamente conscientes de la naturaleza relacional de la humanidad y el valor de los no humanos en sus entornos. Esta perspectiva podría contribuir al desarrollo democrático de la IA y, de forma más general, al desarrollo democrático de las tecnologías digitales. <<

Capítulo 8

[1] Véase *Federalist Papers*, núm. 57; disponible en: <https://avalon.law.yale.edu/18th_century/fed57.asp>. <<

[2] Kitcher (2011), por contra, se refiere a Dewey, pero se centra en el respaldo ideal y no en el real. <<

[3] Por ejemplo, Cohen (2009) ha desarrollado un relato que intenta integrar alguna noción de verdad. <<

[4] La diferencia entre «bien común» [*common good*] y los «comunes» [*commons*] reside en que la primera expresión se refiere a un conjunto de cosas más general con una orientación moral y política hacia el bien de la comunidad, mientras que la segunda se refiere a algo concreto, el compartir bienes específicos como, por ejemplo, datos. [*N. del T.*] <<

[5] Me refiero aquí a *Mare Liberum* de Grocio, que defendió la no privatización de los mares. Para profundizar, véase Risse (2023). <<

[6] La frase «la tragedia de los bienes comunes» se encuentra en el famoso artículo de Hardin en *Science* (1968). Da el ejemplo de un pastor que está totalmente sobrecargado en su oficio porque se encuentra a disposición de todo el mundo y, por lo tanto, el resto de

pastores lo carga con más y más animales. Como señala Ostrom (1990), Aristóteles ya había observado en su *Política* que de lo común nadie se preocupa, ya que, en su lugar, no piensan más que en sus intereses privados. <<

[7] Nótese, por ejemplo, que OpenAI desarrolló primero su tecnología de IA como tecnología «abierta», pero que, luego, cambió su enfoque. <<

[8] Véase, por ejemplo, en: <<https://opendatacommons.org>>; y <<https://creativecommons.org/>>. <<

Bibliografía

[*] La fecha de última consulta de los hipervínculos que figuran en la siguiente bibliografía, así como en las notas al pie de esta edición, es del 8 de julio de 2024. <<

Índice de contenido

Prefacio

Capítulo 1. Introducción

Acerca de este libro

Capítulo 2. Una historia no tan democrática

Política y cibernética

Historias de la tecnología y la democracia

Capítulo 3. ¿Qué IA? ¿Qué democracia?

Inteligencia artificial (IA)

Democracia

Capítulo 4. Cómo debilita la IA los principios básicos de la democracia

Principios fundacionales de la democracia:

Libertad, igualdad, fraternidad, imperio de la ley y tolerancia

Cómo debilita la IA los fundamentos de la democracia

Capítulo 5. Cómo erosiona la IA el conocimiento y la confianza

El problema del totalitarismo revisado:

Las lecciones de Hannah Arendt

Algunos desafíos en relación con la IA,

El conocimiento y la democracia

Capítulo 6. Fortalecer la democracia y democratizar la IA

Fortalecer y crear nuevas instituciones democráticas

Regulación y supervisión

IA democrática

Capítulo 7. IA para la democracia y un nuevo Renacimiento

IA para la democracia
Un nuevo Renacimiento
Humanismo digital
Conclusión

Capítulo 8. El bien común y la comunicación

El bien común y los comunes
Comunicación y comunidad
Experiencia común y construcción de un mundo común
Resumen ejecutivo para responsables políticos

Agradecimientos

Bibliografía[*]

Sobre el autor