

Adela Cortina

¿Ética o ideología de la inteligencia artificial?

El eclipse de la razón comunicativa
en una sociedad tecnologizada

1.ª edición, octubre de 2024

La lectura abre horizontes, iguala oportunidades y construye una sociedad mejor.
La propiedad intelectual es clave en la creación de contenidos culturales porque sostiene
el ecosistema de quienes escriben y de nuestras librerías.

Al comprar este libro estarás contribuyendo a mantener dicho ecosistema vivo y
en crecimiento.

En **Grupo Planeta** agradecemos que nos ayudes a apoyar así la autonomía creativa
de autoras y autores para que puedan seguir desempeñando su labor.
Dirígete a CEDRO (Centro Español de Derechos Reprográficos) si necesitas fotocopiar
o escanear algún fragmento de esta obra. Puedes contactar con CEDRO a través de la
web www.conlicencia.com o por teléfono en el 91 702 19 70 / 93 272 04 47.

© Adela Cortina Orts, 2024

© de todas las ediciones en castellano,

Editorial Planeta, S. A., 2024

Paidós es un sello editorial de Editorial Planeta, S. A.

Avda. Diagonal, 662-664

08034 Barcelona, España

www.paidos.com

www.planetadelibros.com

ISBN: 978-84-493-4279-0

Depósito legal: B. 14.112-2024

Maquetación: Realización Planeta

Impresión y encuadernación en CPI Black Print

Impreso en España – *Printed in Spain*



SUMARIO

Prólogo	11
1. LO QUE NOS HACE HUMANOS	15
1. Las máquinas no están hechas para entender una mentira	15
2. De la isla de Moreau a las leyes de Asimov y más allá . .	18
3. Un robot no debe dañar a un ser humano	23
4. El secreto de la máquina. Cuestiones cruciales	24
5. Cuidar la palabra, cultivar la democracia.	26
2. TRES TIPOS DE INTELIGENCIA ARTIFICIAL, TRES TIPOS DE ÉTICA	29
1. Políticos y empresarios virtuales.	29
2. ¿Qué es la inteligencia artificial?	33
3. La inteligencia artificial se dice de muchas maneras .	34
4. ¿Hasta dónde hemos llegado?	42
3. TRANSHUMANISMO Y POSTHUMANISMO: ¿PRONÓSTICO O IDEOLOGÍA?	43
1. Julian y Aldous Huxley: utopía y distopía.	43
2. ¿Qué es el transhumanismo? Religión sin revelación . .	44
3. El primer deber ético de las tecnociencias es no engañar	52
4. Silicon Valley, China, la Unión Europea. Geoestrategia del poder	55
5. ¿Desde qué ética?	58

4. INTELIGENCIA ARTIFICIAL CONFIABLE:	
¿UN CAMBIO DE ÉPOCA?	61
1. El «efecto Bruselas»	61
2. La ventaja competitiva de la Unión Europea	64
3. El coste de oportunidad en una carrera sin compasión	66
4. Principios éticos para la revolución 4.0.	69
5. ÉTICA ROBÓTICA	81
1. La robótica y la Cuarta Revolución Industrial	81
2. Los desafíos éticos de la robótica	86
3. La necesidad de una ética para las máquinas	88
4. ¿Roboética, ética de las máquinas o ética de agentes artificiales autónomos?	90
6. ¿QUÉ ÉTICA PARA LAS MÁQUINAS INTELIGENTES?	97
1. El experimento de los coches autónomos	97
2. Razones para integrar una ética en las máquinas	100
3. ¿Pueden los sistemas artificiales ser agentes morales?	102
4. Tres enfoques de teorías éticas	106
5. Cinco marcos éticos: deber, utilidad, virtud, capacidades, diálogo	108
7. LAS FRONTERAS DE LA PERSONA EN UN MUNDO TECNOCIENTÍFICO	123
1. ¿Singularidad de la humanidad?	123
2. Personas electrónicas: un debate en el mundo europeo	125
3. El círculo de personas se amplía más allá de los seres humanos	129
4. ¿Superando el biocentrismo?	134

8. LA LIBERTAD EN LA ERA DIGITAL	139
1. La tragedia de <i>Edipo rey</i>	139
2. El afán imperialista de las ciencias triunfantes por negar la libertad	140
3. Determinismo tecnológico	141
9. ECLIPSE DE LA RAZÓN COMUNICATIVA: UN RETO RADICAL PARA LA DEMOCRACIA.	145
1. Hacia dónde queremos ir.	145
2. La razón comunicativa: más allá de la razón instrumental	146
3. A los dictadores les gusta parecer demócratas	148
4. Cuidar la palabra frente a sus enemigos.	151
5. Tiempos de «posveracidad»	154
6. ¿Teoría de la justicia social crítica?	159
7. La palabra se vacía de significado	164
10. UN ESPACIO PÚBLICO LIBRE DE DOGMATISMOS.	167
1. El miedo al aislamiento: una mordaza para la razón comunicativa	167
2. Dos concepciones de la opinión pública	172
3. La «agresión moralista»	175
4. Lo moral en serio	179
11. LENGUAJE CLARO	183
1. Una mesa es una mesa.	183
2. Lenguaje inteligible.	188
3. Claridad es la plenitud de la vida	191
4. Oscuridad como arma ideológica	194
5. De la cortesía del filósofo al derecho de las personas.	197

12. EDUCAR EN LA ERA DE LA INTELIGENCIA ARTIFICIAL	203
1. Inteligencia artificial en educación, ¿para qué? . . .	203
2. Luces y sombras de la inteligencia artificial	207
3. La autonomía personal y la capacidad crítica en peligro.	212
4. ¿Qué hacer? Educar en una autonomía cordial . . .	218
Bibliografía.	223
Notas.	239

Capítulo 1

LO QUE NOS HACE HUMANOS

1. LAS MÁQUINAS NO ESTÁN HECHAS PARA ENTENDER UNA MENTIRA

Pero recordad, por favor, la ley conforme a la cual vivimos: no estamos hechos para entender una mentira...

Al comienzo de su novela *Machines Like Me*, título traducido al español como *Máquinas como yo y gente como vosotros*, el célebre novelista británico Ian McEwan recoge este breve texto de Rudyard Kipling, tomado de su relato «El secreto de las máquinas».

El consejo, expresado con estas palabras, resulta muy acertado en tiempos de supuesta posverdad, cuando se dan por buenas ese tipo de mentiras emotivas que alimentan diariamente a la gente como nosotros y pretenden dar forma a la opinión pública. Objetivo que suelen alcanzar con éxito, porque mentir no suele tener malas consecuencias para el mentiroso, y resultan raras, por escasas, noticias como la del pago de 785,5 millones de dólares a que se vio obligada Fox por sus mentiras. El abogado de Dominion Voting Systems, tras cerrarse el caso, afirmó: «La verdad importa». Ojalá tuviera razón.

Precisamente el hecho de que las máquinas supuestamente inteligentes sean incapaces de entender una mentira es la clave del libro de McEwan, en el que muestra magistralmente cómo resul-

ta imposible sobrevivir en nuestro mundo sin la capacidad de discernir entre veracidad y engaño, entre la comunicación veraz, que busca el entendimiento, como es propio de la palabra, tanto en la vida privada como en la pública, y la manipulación del lenguaje para obtener ventajas espurias. Y justamente es lo que nos está pasando en un tiempo en que se oscurece la labor de la razón comunicativa y pasa a primer plano la razón estratégica, hasta ocupar todo el espacio público. No puede decirse que ésta sea una buena noticia para fortalecer la democracia, que es una tarea tan urgente, porque está en peligro en todo el mundo.

Por si faltara poco, esto ocurre cuando el avance de la IA impregna nuestra existencia y aumenta en todos los niveles los canales de lo que debería ser una comunicación llamada a resolver los grandes problemas de un universo que es ya cosmopolita. Por eso es urgente analizar la textura de este nuevo mundo y un buen punto de partida puede ser la mención del libro de McEwan, que da pie a un sinfín de reflexiones, tanto éticas como ontológicas, acerca de si los llamados «sistemas inteligentes», las máquinas supuestamente inteligentes, tienen una inteligencia como la humana —o pueden llegar a tenerla— y si son o pueden llegar a ser personas, seres autoconscientes y autónomos, a los que se debe proteger con derechos y a la vez exigir responsabilidades. La gran pregunta es si esas máquinas son instrumentos de los que los seres humanos podemos valernos para alcanzar diversas metas, o si irán sustituyendo paulatinamente a los seres humanos, de modo que se vaya poniendo fin a lo que se ha llamado «Antropoceno», la era del *Homo sapiens*.

La novela tiene un narrador, que forma parte de la «gente como nosotros», capaz de entender una mentira, es decir, de entender que se puede afirmar lo contrario de lo que se piensa con intención de engañar. El nombre del narrador y protagonista es Charlie, y cuenta cómo habiendo heredado algún dinero a la

muerte de su madre, decide emplearlo en comprar uno de los primeros humanos manufacturados, verdaderamente viables, que han sido puestos a la venta por 86.000 libras. En un primer lote se habían fabricado veinticinco individuos (es lo más neutral que podemos decir): trece Evas y doce Adanes. Charlie deseaba una Eva, pero las Evas se agotaron al cabo de una semana y tuvo que contentarse con un Adán. A pesar de la decepción, le consolaba algo saber que también había adquirido uno igual Alan Turing, el creador de la célebre máquina de Turing y del test de Turing, que permite descubrir la inteligencia de una máquina comprobando si sus respuestas son indistinguibles de las de un ser humano. Como es bien conocido, durante la Segunda Guerra Mundial Turing dio con las claves para descifrar códigos nazis, particularmente de la máquina Enigma. La intervención de Turing al final de la novela con un largo y emocionado discurso es decisiva.

Obviamente, a diferencia de los hijos biológicos, que nacen con sus características individuales (lo que John Rawls llamaría la «lotería natural» que a cada uno cabe en suerte, pero también la «lotería social», porque el feto va absorbiendo las enseñanzas de su entorno sociocultural y los entornos son muy diferentes según el país y el código postal), la personalidad de Adán tiene que ser programada, y por eso se vende al humano manufacturado con un manual de instrucciones de cuatrocientas setenta páginas. Llevar a cabo la tarea de programarlo es ineludible para ponerlo en funcionamiento —¿para darle la vida?—, pero a la vez es un tanto frustrante, porque, a pesar de la programación inicial, el aprendizaje de máquinas irá configurando su personalidad y condicionando sus decisiones a lo largo de su existencia. No digamos ya desde la aparición de la IA generativa, capaz de producir contenidos originales, inesperados en todas las actividades sociales.

Estas innovaciones parecen refrendar la sugerencia de algunos

autores de educar a las máquinas en aquellos valores morales que consideramos importantes, empezando el proceso educativo antes de que nazcan, insertándoles en el «cerebro» de silicio unos valores que ya irán rumiando desde el comienzo y que desarrollarán sin tener que esperar al largo proceso de la evolución.¹ A mi juicio, es hoy por hoy pura especulación y plantea una buena cantidad de problemas, como veremos más adelante.

Porque desde el comienzo de la obra de McEwan se plantea la gran pregunta: ¿es posible que la máquina tenga motivaciones, sentimientos subjetivos, autoconciencia, incluida la capacidad para la deslealtad y la traición?² A lo largo de la historia no se tiene noticia de ninguna máquina que haya sido desleal y traidora. Según el narrador, el nuevo Adán parece estar hecho para el bien y la verdad³ y, sin embargo, ¿puede hablarse de «inclinación al bien y a la verdad» en seres que no están dotados de las características que acabamos de mencionar?

2. DE LA ISLA DE MOREAU A LAS LEYES DE ASIMOV Y MÁS ALLÁ

Para ir respondiendo a estas cuestiones retrocederemos en principio algunas décadas en ese ámbito de la literatura que puede llamarse en sentido amplio «de ciencia ficción» y que tantas verdades desvela de la vida cotidiana, y haremos pie en el siglo XIX. En ese tiempo florecen un buen número de novelas inquietantes, estrechamente vinculadas al progreso en los descubrimientos científicos, que se han convertido con el tiempo en referencias indiscutibles de la literatura, el cine y el amplio mundo de la cultura. Una de ellas es *La isla del doctor Moreau*, de H. G. Wells, una obra a la que he recurrido en distintas ocasiones porque descubre de forma palmaria las dificultades que plantea intentar

transformar de forma artificial animales no humanos en humanos, tomando como una de las claves esenciales la de inculcarles una mente moral.⁴ Moralización y hominización —no sólo humanización— irían entonces de la mano. A mi juicio, algo semejante se intenta hoy de nuevo, pero dando un paso más allá, al tratar de hominizar las máquinas inteligentes programando en ellas una suerte de mente moral. Pasaríamos entonces del sustrato biológico, propio de animales y seres humanos, a un sustrato de silicio.

Todavía en el ámbito biológico, en 1896 publica Wells *La isla del doctor Moreau*, de la que por cierto ha habido excelentes versiones cinematográficas, y en ella deja patente cómo entiende el autor el proceso de hominización llevado a cabo, no por evolución natural, sino de modo artificial, y cómo fracasa rotundamente. Moreau se propone transformar animales no humanos en humanos y para ello cuenta con dos claves: la vivisección, practicada a los animales en un recinto que recibe el nombre de Casa del Dolor y que permite transformarlos fisiológicamente en «humanimales», y la mentalización, la conversión de las mentes animales en mentes humanas. La obra de Wells recibió un buen número de críticas precisamente por los pasajes dedicados a la vivisección, pero lo apasionante es preguntarse por qué le pareció imprescindible no sólo modificar la fisiología de los animales, sino también introducir en sus mentes una ley humana, la «ley de la humanidad». Un «recitador de la ley» recuerda diariamente a los humanimales unas leyes que son propias de la conducta humana y se sustancian en las siguientes: no caminarás a cuatro patas, no sorberás la bebida, no comerás carne ni pescado, no cazarás a otros hombres. Y en cada ocasión, al acabar su cantinela, lanza al recitador la pregunta persuasiva: «¿Acaso no somos Hombres?».

Pero no, no eran seres humanos, y de ahí que la novela acabe en el rotundo fracaso del experimento. Moreau no logra transformar a los animales en humanos y, por si faltara poco, de tanto

violentar su naturaleza, sólo consigue que incendien los barracones en los que se llevan a cabo los experimentos, asesinen al doctor y regresen a la selva, que es a fin de cuentas su lugar natural. El relato es sin duda estremecedor y deja en el aire la gran cuestión: ¿por qué fracasó el intento de hominización?

Tal vez la respuesta venga de algo tan sencillo, pero a la vez tan poco reflexionado, como es el hecho de que comprender realmente una ley de la humanidad, una orientación moral, exige contar con un *cuero humano* y, estrechamente ligado a él, con ese sentido que permite sintonizar con la ley, que ha recibido el nombre de «sentido moral» o «conciencia» y que lleva entrañados en su seno un conjunto de elementos que más adelante intentaremos desgranar. Por el momento recordaremos cómo Kant en uno de sus libros tardíos, *La metafísica de las costumbres*, afirmaba que en los seres racionales hay unas «prenociones estéticas de la receptividad del ánimo para los conceptos del deber en general», que son de la mayor importancia, porque se refieren a disposiciones morales que, si no se tienen, tampoco hay un deber de adquirirlas.⁵ Forman parte de la naturaleza humana y, aunque son estéticas, se las puede calificar a la vez de morales, porque preparan a la persona para escuchar la voz del deber, para comprenderla y para poder obedecerla o desoírla. Esas predisposiciones son, según él, el sentimiento moral, la conciencia moral, el amor al prójimo y el respeto por sí mismo, es decir, la autoestima. Constituyen una especie de antenas de la naturaleza que permiten percibir el deber.

Si un ser careciera de estas disposiciones no contaría con la base natural indispensable para escuchar la voz de la ley moral y seguirla o rechazarla. Podemos decir, recurriendo a la distinción indispensable en el mundo filosófico entre base y fundamento, que la ley moral sería el fundamento de la obligación moral, pero las disposiciones naturales estéticas serían la base sin la que resulta

imposible escuchar la voz de la ley y seguirla.⁶ Se trataría, como dice Jesús Conill, de una «estética de la libertad»;⁷ es una predisposición del ánimo, estética, pero natural, que consiste en la razón práctica que muestra al hombre su deber en el caso concreto, absolviéndolo o condenándolo. De modo que forma parte de su naturaleza biológica moral la predisposición a juzgarse a sí mismo.

Justamente esta conciencia, como herencia biológica, podría ser el punto de contacto entre la naturaleza y la libertad, al igual que las otras tres prenociones estéticas. ¿Contaban los humanimales con esas bases naturales que hacen posible comprender las exigencias y las invitaciones de la moralidad y que constituyen un requisito imprescindible no sólo para la humanización, sino también para la hominización? Algo semejante venía a decir Charles Darwin cuando afirmaba en *El origen del hombre*:

Suscribo totalmente la opinión de aquellos autores que sostienen que de entre todas las diferencias entre el hombre y los animales inferiores, el sentido moral o conciencia es, con mucho, la más importante. Este sentido, tal como señala Mackintosh, «tiene una supremacía legítima sobre cualquier otro principio de acción humana».⁸

La conciencia o sentido moral no sería, pues, según Darwin, una dimensión intranscendente de la vida humana, sino que, a su juicio, marca la diferencia más relevante entre el hombre y el animal y tiene supremacía sobre cualquier otro principio humano; una afirmación en la que insiste reiteradamente a lo largo de su obra.

Es verdad que un buen número de animalistas insisten en afirmar que entre ciertos animales no humanos, como los chimpancés, y los seres humanos existe un conjunto de disposiciones comunes que podríamos llamar «protomorales» o, por decirlo con

Frans de Waal, los *building blocks* de la moralidad, entre los que señala la reciprocidad, el consuelo, la aversión a la inequidad, la empatía y el seguimiento de reglas de conducta reforzado por otros.⁹ Como también es verdad que en algunos casos es discutible que también los animales estén dotados de esos *building blocks*. Pero lo cierto es que, aunque así fuera, no les capacitarían todavía para desarrollar el sentido moral o conciencia, sino que hace falta algo más. Según Darwin, nuestro sentido moral o nuestra conciencia se origina en los instintos sociales, es conducido en gran parte por la aprobación de nuestros semejantes, regido por la razón, el interés propio y, en los últimos tiempos, por sentimientos religiosos, y confirmado por la instrucción y el hábito.¹⁰

Al leer estos textos no podemos dejar de preguntarnos si en el cambio de siglo del xx al xxi no hemos entrado en una nueva fase en el intento de hominizar seres no humanos, recurriendo en esta ocasión como materia prima a las máquinas. A ello se referiría ese rótulo tan frecuente hoy en día, «Ética para máquinas» o «Ética de las máquinas», un membrete que nació, al parecer, hacia 1991, cuando James Gips comparó diferentes teorías éticas con el fin de incorporarlas a los robots en el texto «Towards the Ethical Robot», y después, en 2000, Storrs Hall introdujo la expresión *machine ethics* en «Ethics for Machines». Desde entonces al menos proliferan los intentos de dotar a los sistemas inteligentes (robots, algoritmos, máquinas) de leyes morales de uno u otro cuño. En principio, la meta consistía en lograr que se comporten de modo que respeten los intereses de los seres humanos y ayuden a satisfacerlos, pero, evidentemente, la empresa ha desbordado sobradamente esta pretensión.

3. UN ROBOT NO DEBE DAÑAR A UN SER HUMANO

La idea de que es necesario introducir en las «mentes» de los robots normas éticas que los lleven a actuar de forma que coincidan con los intereses de los seres humanos está presente desde el nacimiento de los robots, pero será en 1942, en el cuento de ciencia ficción «Círculo vicioso», cuando el científico visionario Isaac Asimov expondrá las célebres tres leyes de la robótica, a las que recurrirá a menudo en sus escritos. Estas leyes, como es bien conocido, son las siguientes:

1. Un robot no debe dañar a un ser humano o, por su inacción, dejar que un ser humano sufra daño.
2. Un robot debe obedecer las órdenes que le son dadas por un ser humano, excepto cuando estas órdenes están en oposición con la primera ley.
3. Un robot debe proteger su propia existencia, hasta donde esa protección no esté en conflicto con la primera o segunda ley.

Manual de robótica
56.^a edición, año 2058

Más adelante añadirá una ley cero, que precederá a todas las demás: «Un robot no puede dañar a la humanidad, o, por inacción, permitir que la humanidad sufra daños».

Evidentemente, las Leyes de la Robótica de Asimov hubieran planteado una gran cantidad de problemas a los robots para actuar éticamente en los casos concretos, como veremos más adelante, porque para aplicarlas es necesario entender los contextos e interpretarlos y porque a menudo entran en contradicción. Por poner un ejemplo, si por salvar a seres humanos el robot se ve obligado

a dañar a otro ser humano que los amenaza, tiene que cumplir y violar a la vez la primera ley. Y como las leyes son formulaciones matemáticas, impresas en los senderos positrónicos del cerebro, el robot se destruye cuando incumple alguna de ellas. El tema ha generado una gran cantidad de discusiones, pero, en cualquier caso, con las leyes se abría la posibilidad y necesidad de una ética de los robots, una ética robótica, una roboética, una ética de la IA o una ética de las máquinas, que han proliferado en nuestros días.

4. EL SECRETO DE LA MÁQUINA. CUESTIONES CRUCIALES

Al hilo de estos sintagmas se plantean al menos cinco cuestiones cruciales:

1. ¿Se trata realmente de una «ética de la IA» en el sentido del genitivo subjetivo por el que entendemos que son las máquinas las que tienen una ética desde la que actúan, unas normas o unos valores desde los que deberían actuar, como podrían ser las leyes de Asimov o cualquier otro modelo ético? ¿O más bien se trata de la ética con la que los seres humanos deben actuar para servirse de la IA?
2. O, dicho de otro modo: ¿una ética de la IA es la que deben practicar los sistemas inteligentes desde sus propios valores o es la que los seres humanos deberíamos adoptar para servirnos de los sistemas inteligentes?
3. El mundo de la IA es ya irreversible y es preciso vivir en él de forma inteligente desde el punto de vista humano. ¿Se trata de que los seres humanos utilicemos los sistemas inteligentes como *instrumentos* o de que los sistemas inteligentes *sustituyan* a los seres humanos?

4. Como ya hemos comentado, tomar las soluciones que ofrecen los algoritmos, aplicados a muy diversas esferas de la vida social, y asumirlas en nuestras decisiones sin mayor reflexión ni crítica, ¿no implicaría caer en un determinismo algorítmico, renunciando a la libertad?¹¹
5. ¿Qué tipo de ética sería en cada uno de los dos casos, teniendo en cuenta que hay diversas teorías éticas y muy distintas culturas en la vida cotidiana de un universo multicultural?
6. En el caso de que las inteligencias artificiales pudieran integrar en sus «mentes» normas éticas, ¿se las podría considerar como seres dotados de características, tradicionalmente tenidas por humanas, como sentido moral, conciencia y autoconciencia, autonomía, libertad y responsabilidad? ¿Tendríamos que reconocerlas como personas dotadas de dignidad y de derechos, incluso como miembros de la comunidad social y política?
7. No parece entenderlo así Kipling en el texto con el que hemos dado comienzo a este capítulo, tomado de «El secreto de las máquinas», porque con todo acierto el autor concluye revelando a las máquinas ese secreto, que no es sino el siguiente:

Mas recuerda la ley que rige nuestras vidas;
 No estamos construidas para entender mentiras,
 No podemos amar, llorar ni perdonar.
 Si nos manejas mal, encontrarás la muerte.
 Somos mayores que el Hombre o los Reyes.
 Sé humilde cuando te arrastres por debajo de nuestras bielas,
 Nuestro toque puede cambiar todas las cosas creadas.
 Somos todo en el mundo, salvo los dioses.
 Aunque nuestro humo oculte los cielos de sus ojos,
 Desaparecerá y las estrellas volverán a brillar;
 Porque pese a nuestra fuerza, peso y magnitud,
 No somos más que el producto de su intelecto.

Sin embargo, ¿tiene razón Kipling o más bien la revolución 4.0 ha superado de tal modo la potencia de la Revolución Industrial, inaugurada con la máquina de vapor, que las nuevas máquinas —las máquinas inteligentes— son o serán capaces de ejercer esas competencias que Kipling les niega?

Incluso podría llegar a ocurrir que las ejercieran de un modo muy superior en comparación con los seres humanos, como aseguran buen número de investigadores que consiguen atraer inversiones millonarias para sus proyectos. ¿Quién —o quiénes— tienen que decidir cómo gestionar este mundo, que ya no es tan nuevo?

5. CUIDAR LA PALABRA, CULTIVAR LA DEMOCRACIA

Porque queda en el tintero una cuestión clave, reflejada en el subtítulo del presente libro, tal vez la cuestión esencial. En nuestra sociedad tecnologizada el inusitado aumento de los canales de conexión a través de plataformas y redes sociales está haciendo posible un progreso en la comunicación entre los seres con capacidad de ejercerla más allá de los confines de la tierra, en el mundo infinito del ciberespacio. Con estas alforjas sería posible construir el presente y el futuro desde un «nosotros» democrático, que alcanzara a todos los seres con capacidad de decidir. Pero el mayor problema consiste en que, para lograrlo, no basta contar con las redes, conectarse no es comunicarse, es preciso cuidar con esmero la palabra, que para lograr la comunicación tiene que ser, entre otras cosas, veraz. Y, desgraciadamente, son malos tiempos para la veracidad, no tanto para la verdad, que continúa conviviendo con la presunta «posverdad». Lo que peligra sobre todo es la veracidad, y conviene recordar que lo contrario de la verdad es el error, lo contrario de la veracidad es la mentira.

Las máquinas no pueden entender una mentira, dirá Kipling, pero los seres humanos sí, por eso podemos romper los puentes de comunicación desde el engaño sistemático, desde la prostitución de la palabra. El mayor peligro no consiste en que robots y algoritmos distorsionen el entendimiento mutuo, porque, hoy por hoy, no son sino el producto de nuestro intelecto. Pero ¿y mañana?

Por desgracia, en este siglo XXI se viene reproduciendo lo que los autores de la primera generación de la Escuela de Fráncfort llamaron en el siglo pasado un «eclipse de la razón», producido por el triunfo de la razón instrumental y estratégica que colonizaba todo el ámbito del espacio público. Con el agravante de que en aquel tiempo fue posible superar el eclipse de la razón al descubrir la existencia de otro tipo de razón, la razón comunicativa, que nos permitiría entendernos y construir juntos un mundo nuevo. Pero en nuestra sociedad tecnologizada, atravesada por infinitas conexiones, cuando las decisiones deben tomarse con tal celeridad que no hay tiempo para calibrar si las noticias que llegan a través de ellas son *deepfakes*, creaciones de ChatGPT en alguna de sus versiones o son veraces y bien informadas, es justamente la razón comunicativa la que está quedando eclipsada, cuando es indispensable para construir una vida humana justa y feliz, y, por supuesto, sociedades democráticas.

Es urgente cuidar la palabra, es preciso cultivar la democracia en nuestro tiempo. ¿Aprovechando los grandes beneficios que puede aportar la IA o poniendo en sus manos la construcción de un mundo mejor, como sugieren un buen número de autores?

La pregunta es, evidentemente, retórica. El bien máspreciado es la libertad, la libertad que puede ser disfrutada por todos los seres humanos y que, en su forma política, se expresa de modo óptimo en la democracia, a pesar de todas las limitaciones.