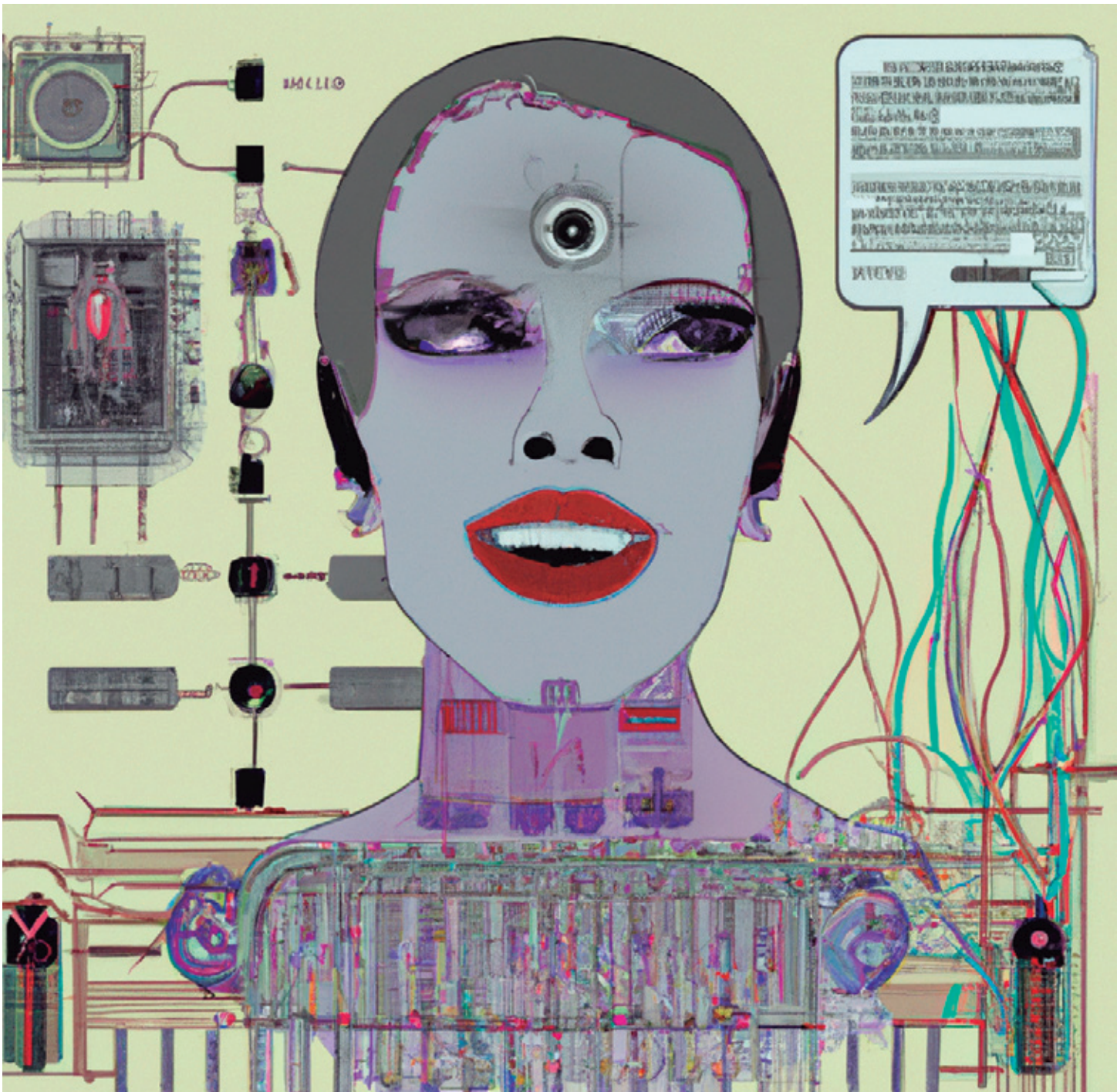


El filósofo de la tecnología Mark Coeckelbergh, profesor de la Universidad de Viena y miembro del grupo de expertos de alto nivel sobre IA de la Comisión Europea, conversó con el profesor de ciencia política de la UAM Juan Ignacio Criado sobre las ideas que recoge su libro *Ética de la inteligencia artificial* (Cátedra, 2021). Una suerte de manual de instrucciones para una convivencia saludable entre máquinas y humanos en todos los ámbitos cotidianos, desde el personal al político, que invita a construir una idea nueva de lo colectivo inspirada en el discurso sobre la buena vida de la filosofía antigua.

# ¿QUÉ TIPO DE CÍBORGS QUEREMOS SER?

## ÉTICA E INTELIGENCIA ARTIFICIAL

### CONVERSACIÓN JUAN IGNACIO CRIADO MARK COECKELBERGH



**JUAN IGNACIO CRIADO**

En 1969 el político y premio Nobel de Economía Herbert Simon escribió *The Sciences of Artificial*, un libro que remarcaba la importancia de lo artificial en las vidas humanas y nos recordaba que, vistos como sistemas de comportamiento, los seres humanos éramos relativamente simples. La aparente complejidad de nuestra conducta se explica fundamentalmente como el fruto de una continua reflexión acerca de la complejidad de nuestro entorno; es decir, para Simon somos simples máquinas algo más inteligentes de lo normal, pero no demasiado. Películas como *Blade Runner* (1982) nos sugerían un mundo donde robots y humanos convivían sin apenas diferenciarse y en el que nuestros sueños se confundían con una realidad inquietante; una especie de *La vida es sueño* en tiempos del punk. Su secuela reciente nos emplaza a 2049, año que da título a la película, cuando aparecen de forma descarnada algunos de los impactos más virulentos de la bioética y el cambio climático que desafían de manera definitiva la vida humana en la Tierra.

La tecnología ha acompañado a los seres humanos a lo largo de la historia, pero esta es la primera vez que desafía nuestra hegemonía sobre el control del planeta y que nos permite imaginar un mundo netamente distinto, libre de cadenas. Sin embargo, frente a esta posibilidad de ensueño, aparecen escenarios reales, donde los peligros derivados de la creación de «datacracias» provenientes de Oriente, de la «gobernanza algorítmica» o del «capitalismo de vigilancia», nuevo paradigma en el Gobierno y la Administración pública, nos conducen al rescate de la filosofía, las humanidades y, en general, las ciencias humanas y sociales como eje clave para orientar la construcción de un futuro que no está escrito en las estrellas.

Además de plantear dilemas filosóficos y de desarrollar argumentos contrapuestos ante una misma cuestión, algo muy significativo de su manera de reflexionar, en *Ética de la inteligencia artificial* Mark Coeckelbergh aborda también cuestiones de alcance político y práctico, como el diseño de estrategias de IA a pequeña escala, las distintas interpretaciones de las nuevas formas digitales del trabajo o cómo entender la vida buena en un contexto dominado por la integración creciente entre máquinas y humanos. Antes de pasar a hablar del contenido del libro, ¿qué te impulsó a escribirlo? ¿Qué laguna querías cubrir?

**MARK COECKELBERGH**

Desde 2017 sentía un interés creciente por la IA y la ética, pero apenas existían estudios desde un enfoque filosófico de la tecnología. Pensé que era necesario abordar esta cuestión desde el punto de vista de la filosofía de la tecnología y situarla en un contexto más amplio.

**JUAN IGNACIO CRIADO**

Al inicio del libro hablas de «ir más allá del hip y de las pesadillas sobre la IA». ¿Cómo defines tu propio pensamiento sobre esta cuestión y en qué medida influye en tu libro el enfoque filosófico?

**MARK COECKELBERGH**

Mi intención era posicionarme frente al transhumanismo y las personas que piensan que la IA va a definir el futuro en términos

**Tenemos que pensar cómo transformar las redes sociales para recrear una esfera pública. Es muy difícil, pero estamos en un momento que nos brinda esa oportunidad.**

MARK COECKELBERGH

**Es la primera vez que la tecnología desafía nuestra hegemonía sobre el control del planeta y que nos permite imaginar un mundo netamente distinto.**

JUAN IGNACIO CRIADO

de superinteligencia. No tengo nada en contra de la ciencia ficción ni de la literatura. Las dos películas que has nombrado, *Blade Runner* y *Blade Runner 2049*, nos ayudan a pensar en el futuro, pero yo intento abordar el debate actual sobre la IA que forma parte de nuestra vida cotidiana. Otro eje del libro, aunque se presenta como una introducción, es la filosofía antigua y su discurso sobre la buena vida y la buena sociedad, que para mí sigue siendo una fuente de inspiración tremendamente útil.

**JUAN IGNACIO CRIADO**

En el segundo capítulo del libro introduces el debate acerca de términos como «singularidad tecnológica», «superinteligencia» o «transhumanismo». En particular, tratas la rivalidad entre la creación artificial y la de los seres humanos, un tema con muchas implicaciones y posiciones contradictorias. Parafraseando la expresión «complejo de Frankenstein» acuñada por Asimov, hablas del «nuevo monstruo de Frankenstein». ¿Cómo podemos ir más allá de estas narraciones competitivas entre la IA y la inteligencia humana y qué papel juegan en ello las culturas no occidentales?

**MARK COECKELBERGH**

El pensamiento occidental hace una distinción muy estricta entre los humanos y el resto del mundo. En ética hablamos de derechos humanos y los distinguimos de los derechos de los animales y de otras entidades. La reflexión sobre la tecnología, o sobre otras cuestiones como la ecología, implica explorar perspectivas no occidentales, conocer otras conceptualizaciones de nuestra manera de relacionarnos con el mundo que no pasan exclusivamente por la absoluta primacía de lo humano. Si queremos avanzar desde el punto de vista ético, necesitamos una ontología relacional.

**JUAN IGNACIO CRIADO**

Una de las reflexiones más interesantes de tu libro es la que muestra cómo la tensa relación entre humanos e IA puede ayudarnos a ampliar nuestro conocimiento de la mente, la conciencia, la creatividad y el conocimiento humanos. Tú analizas tres oposiciones: Ilustración-Romanticismo, Ilustración-no Ilustración y humanismo-poshumanismo. ¿Qué opinas del potencial emergente entre humanos y máquinas y cuáles son las implicaciones filosóficas de este proceso relacional entre humanos, máquinas y robots?

**MARK COECKELBERGH**

Volviendo al monstruo de Frankenstein, podemos ver las máquinas como una amenaza o como algo que no es externo a los humanos: las máquinas, la tecnología, son creación nuestra y las utilizamos, integrándolas en nuestra vida y en la sociedad. Hay que partir del hecho de que la tecnología y los humanos siempre hemos estado juntos; en ese sentido, somos cíborgs. Podemos criticar la tecnología desde el punto de vista humanista, con el objeto de pensar qué configuraciones son necesarias para que podamos convivir bien con ella, pero no podemos hacerlo como si fuera algo que nos es ajeno. Tenemos que pensar cómo dar forma a esa relación entre humanos y tecnología de manera que sea positiva para llevar una buena vida y construir una buena sociedad.

**JUAN IGNACIO CRIADO**

Otra cuestión muy interesante que recoges en el libro es la del estatuto moral de la IA y de los robots. Presentas ideas como la de agencia moral e introduces, entre otras cuestiones, reflexiones acerca de si las máquinas han de tener derechos o si merecen algún tipo de consideración moral. ¿Apoyas la idea de una moralidad funcional de la IA?

**MARK COECKELBERGH**

Es una pregunta muy técnica. Para ponerla en su contexto, hay que plantearse cuál es su origen. Hay personas que «maltratan» a los robots. Se podría decir que no hay nada de malo en ello, pues no son más que máquinas. Sin embargo, no está bien, porque es perjudicial para el carácter moral de los propios humanos. Respecto a la cuestión del estatuto moral, es importante, desde el punto de vista filosófico, pensar a qué o a quién se lo adjudicamos: a las lenguas, por ejemplo; a los animales, cuando les ponemos un nombre propio... Sin embargo, si pasamos a un contexto industrial, con objetos inanimados, entonces pensamos de otra manera. He intentado reflexionar sobre ello y sobre lo que debemos hacer normativamente. En mi opinión, no se trata tanto del trato que reciben los robots, sino de cómo nos comportamos y nos relacionamos los humanos con ellos.

**JUAN IGNACIO CRIADO**

En países como Japón, donde los robots se dedican a los cuidados de personas mayores. ¿se les asigna un estatuto moral?

**MARK COECKELBERGH**

No creo que los robots tengan estatuto moral en ningún sitio, al menos, entendido este como conciencia. Lo que sucede es que las personas construyen una relación con los robots, como en el caso de los robots canguro. Y precisamente por ese contexto que se crea, el robot adquiere un valor para el ser humano. En ese sentido, sería éticamente malo que el humano lo destruyera o dañara.

**JUAN IGNACIO CRIADO**

Dos capítulos del libro tratan sobre la tecnología de los datos. Ahora mismo, entre la audiencia hay gente que está interactuando con personas fuera de la sala a través de sus teléfonos móviles, utilizando diferentes plataformas y por medio de redes sociales basadas en algoritmos. ¿Cómo entiendes este proceso de intercambio y las implicaciones en nuestra vida cotidiana desde el punto de vista físico, de la recuperación del tiempo y el espacio físicos?

**MARK COECKELBERGH**

Hemos creado otro mundo, un ciberespacio. Es una idea de la década de 1990. En ese momento tenía sentido, pues no había tantas personas que trabajaran con ordenadores, no teníamos teléfonos móviles, los ordenadores estaban en un lugar fijo: encendías el ordenador y entrabas en otro mundo. Hoy tenemos que pensar de manera diferente, también desde un punto de vista filosófico. La visión dualista de la ideología ciborg, que distingue el mundo real del

digital, ya no es apropiada. Necesitamos metáforas y narrativas distintas. Podemos utilizar la metáfora del ciborg, pero debemos superar el pensamiento dualista. Por ejemplo, ahora mismo, mientras hablo, no estoy haciendo algo solo físico, en un plano ajeno a lo tecnológico, ya que estoy siendo grabado. Las tecnologías intervienen en nuestra vida. Tenemos que reconocerlo y, en lugar de situarnos en su contra, pensar qué tipo de ciborg queremos ser.

**JUAN IGNACIO CRIADO**

En el libro también hablas de problemas ético-prácticos. Estableces tres categorías principales: la privacidad, la responsabilidad y la explicabilidad de las decisiones, los prejuicios y el significado de la vida. Desde el punto de vista de la privacidad, ¿qué cabe esperar de este proceso en las sociedades democráticas?

**MARK COECKELBERGH**

Existe una confusión entre lo público y lo privado. Por un lado, nuestras vidas privadas son cada vez más públicas; ya no estamos solos en la cocina, ni en el dormitorio... No es el Gran hermano, pero quizá sí la «pequeña hermana». Vivimos intervenidos por tecnologías. A través de ellas compartimos cosas, nos relacionamos, pero también sufrimos la censura de las grandes compañías tecnológicas, que ni siquiera están en suelo europeo. Esto supone una privatización de la esfera pública, algo extremadamente peligroso para la democracia. No podemos prescindir de estas tecnologías, pero sí podemos transformar las redes sociales, por ejemplo, para evitar esa privatización. También los filtros burbuja y las cámaras de eco representan un peligro para la democracia, al constituir espacios en los que las personas de la misma opinión hablan entre sí en pequeños grupos; la democracia implica polémicas, discusiones, debates emocionales, y nada de esto se da dentro de una burbuja. De ahí que tengamos que pensar cómo transformar las redes sociales para recrear una esfera pública y algún tipo de bien común. Necesitamos construir una idea nueva de lo colectivo. Es muy difícil, pero estamos en un momento que nos brinda esa oportunidad.

Tenemos que hacernos la pregunta filosófica de qué es *nosotros* y *nuestra*. La búsqueda de la innovación en Occidente se ha debido a la importancia que le han otorgado a lo colectivo la mayoría de las sociedades del pasado. Sin embargo, la ética moderna se ha centrado excesivamente en el individuo. Ahora debemos encontrar el equilibrio entre la individualidad del ser humano y el hecho de que somos interdependientes y dependemos de la colectividad. Es importante tener presentes

Teatro de autómatas de Gonzalo Cañas, principios del siglo XX. © Diario de Madrid CC 4.0



los derechos colectivos, el bien común del que dependemos, que no tiene nada que ver con mi privacidad personal, por ejemplo, sino con la dependencia ecológica. Tenemos que caminar hacia ese punto de vista relacional sin perder los beneficios de la sociedad moderna.

#### JUAN IGNACIO CRIADO

Al hablar de decisiones responsables tomadas por máquinas, abordamos el aspecto político de la IA. ¿Quién debería ser responsable del *machine learning* y de las decisiones sobre cuestiones como los tratamientos médicos, los seguros, la concesión de visados o de becas? ¿Qué responsabilidad tienen los gobiernos sobre estas decisiones?

#### MARK COECKELBERGH

En una democracia siempre debe existir la rendición de cuentas y la asunción de responsabilidad frente a las decisiones que se toman. Respecto a la energía, por ejemplo, los humanos toman decisiones de manera no democrática, mientras que las máquinas podrían adoptar decisiones democráticas. Para que eso ocurra no podemos seguir viendo la tecnología como algo puramente técnico, sino como algo que forma parte de la política.

Por otro lado, cuando estas tecnologías se utilizan para implementar en la sociedad políticas predictivas, podemos predecir la delincuencia, por ejemplo, pero tenemos la desventaja de que nuestras vidas estén sometidas a la vigilancia permanente de las máquinas. Tendríamos que tener cuidado a la hora de usar estas tecnologías, de tal manera que nuestras libertades y nuestros derechos fundamentales estén protegidos. Esto no me preocupa como filósofo, sino como ciudadano. Con la pandemia hemos visto lo fácil que les resulta a los gobiernos o la policía intervenir en nuestras vidas personales. Si estas tecnologías se utilizan en esa dirección, peligran nuestras libertades.

#### JUAN IGNACIO CRIADO

En distintos ámbitos, incluido el sector público, los ciudadanos utilizan cada vez más servicios gestionados por tecnologías, que hacen muy necesarias la transparencia y la explicabilidad, también con el objeto de generar confianza entre la ciudadanía. ¿Llegaremos a ser capaces de desarrollar sistemas abiertos de IA transparentes y explicables?

#### MARK COECKELBERGH

Es un desafío técnico. Se puede comprender el funcionamiento de estas tecnologías técnicamente, pero no es posible describir los pasos concretos y, por otra parte, los modelos estadísticos de los algoritmos no permiten la transparencia. Se trata, pues, de un reto técnico, pero no solo. Aunque se pueda explicar técnicamente cómo funciona una máquina, es importante para la democracia y para la rendición de cuentas poder explicar a la ciudadanía el porqué de las decisiones que toman las máquinas. Por lo tanto, debemos establecer qué tipo de conocimiento necesitan los ciudadanos del siglo XXI para que la democracia funcione. La cuestión va más allá de la explicabilidad técnica.

#### JUAN IGNACIO CRIADO

Desde un punto de vista filosófico y político, ¿cómo ves el futuro de nuestras sociedades respecto al

## **El pensamiento occidental hace una distinción muy estricta entre los humanos y el resto del mundo. Si queremos avanzar, necesitamos una ontología relacional.**

MARK COECKELBERGH

sentido de la vida? ¿Y en lo que atañe al trabajo, que hasta ahora ha dado forma a nuestras vidas y a nuestras sociedades?

#### MARK COECKELBERGH

En un mundo en el que las máquinas desempeñan cada vez más trabajos, uno de los retos es enfocar el sentido de la vida hacia otras

direcciones. Si viviéramos en una sociedad en la que ya no tuviéramos que trabajar tanto o incluso en la que no trabajáramos nada, nos plantearíamos qué sentido queremos dar a nuestras vidas. Hay varias respuestas a esta cuestión. Personalmente, pienso que podemos hacer mucho trabajo social y creativo, medioambiental, etc. En cualquier caso, esta pregunta suscita un problema desde un punto de vista político, ya que la utopía de la sociedad del ocio, que surgió entre las décadas de 1960 y 1970, nunca tuvo lugar. Me temo que estamos creando una suerte de sociedad esclavista, jerárquica, donde unas personas viven en la sociedad del ocio y otras no. Las primeras son una especie de seres trágicos, ya que, en el sentido hegeliano, no pueden recibir el reconocimiento de las máquinas y están condenados a vivir en un consumo insensato. Si pasamos a una automatización inteligente, hay que detenerse a reflexionar sobre ello.

#### JUAN IGNACIO CRIADO

En otra parte del libro ofreces propuestas de políticas públicas, algunas respaldadas por varios gobiernos, e introduces ejemplos y desafíos para los diseñadores de estas políticas. ¿Qué opinas sobre las estrategias adoptadas por Estados Unidos, la Unión Europea o China, desde un punto de vista filosófico o práctico?

#### MARK COECKELBERGH

La tecnología tiene una relación muy estrecha con la cultura de cada país, también con las políticas. China trata estos asuntos desde un enfoque más autoritario, como hemos visto con la gestión de la pandemia, por ejemplo. En otros países hay un enfoque más de *laissez faire* que permite a las compañías tecnológicas tomar sus propias decisiones éticas sobre IA. En Europa tenemos la oportunidad de encontrar un equilibrio entre los valores políticos, la libertad y la cohesión social. Podemos dar ejemplo con nuestra regulación de la IA, que podría ser un referente para el resto del mundo. Tal vez no seamos líderes en innovación tecnológica, pero sí en innovación ética.

#### JUAN IGNACIO CRIADO

El diseño de la tecnología refleja relaciones de poder en la sociedad y en las instituciones; para el diseño y la aplicación de la IA tú sugieres un enfoque de abajo arriba. ¿Qué se puede hacer para que más ciudadanos se sientan concernidos, y cómo podemos dar voz a investigadores, a profesionales de la IA y, sobre todo, a las personas potencialmente discriminadas por este desarrollo opaco y vertical llevado a cabo por las compañías tecnológicas?

#### MARK COECKELBERGH

En ámbitos locales y nacionales, esta cuestión tiene que ver con cómo mejorar la democracia, cómo hacerla más participativa, también desde el punto de vista tecnológico, para que no decidan únicamente los expertos y los políticos. El problema es

## **Los ciudadanos utilizan cada vez más servicios gestionados por tecnologías y se hacen necesarias la transparencia y la explicabilidad.**

JUAN IGNACIO CRIADO



Versión cibernética del médico y poeta estadounidense Oliver Wendell Holmes realizada por una inteligencia artificial

que estamos en un contexto global en el que en Europa utilizamos *software* hecho en Silicon Valley, por ejemplo, por lo que el desequilibrio de poder es muy fuerte. Los europeos debemos asegurarnos de tener una voz al respecto. Es una gran pregunta la de cómo ordenar las sociedades en lo social y lo económico, qué hacer en un contexto capitalista, cómo movernos en las relaciones internacionales, pues hemos tomado medidas respecto a la política exterior que nos han perjudicado. En cualquier caso, los países no pueden hacer gran cosa contra estas grandes compañías. Debemos ser más estratégicos en el ámbito global como europeos.

#### **JUAN IGNACIO CRIADO**

En el último capítulo del libro, «Es el clima, estúpido», abres el debate sobre la IA y el cambio climático. ¿Crees que la IA ayudará a llevar a cabo grandes descubrimientos o innovaciones?

#### **MARK COECKELBERGH**

Estos años de pandemia nos han dejado claro que tenemos que establecer cuáles son nuestras prioridades. Cuestiones sociales como la crisis del clima están sobre la mesa y es fundamental la respuesta que le den la sociedad y los políticos. Aunque muchos piensen que estas y otras cuestiones semejantes son más importantes que la IA, se equivocan, pues la IA puede ser de mucha ayuda en la predicción de fenómenos climáticos extremos, por ejemplo, o a la hora de ahorrar energía. La IA y lo verde no han de estar en conflicto. Por otro lado, en escenarios futuros, el uso de la IA puede llevarnos por derroteros totalitarios y autoritarios, como ocurre cuando se considera que estas tecnologías son la única vía para resolver nuestros problemas. Tenemos que estar atentos para no llegar a situaciones en las que perdamos la libertad y los derechos fundamentales que nos corresponden como ciudadanos.