

# ¿QUIEN LE TEME A LA INTELIGENCIA?

POSIBILIDADES Y RIESGOS DE LA INTELIGENCIA ARTIFICIAL  
EN EL ESTADO DIGITAL

The left side of the page features a decorative background with several overlapping circular patterns. Each pattern consists of concentric circles and segments, resembling a stylized target or a data visualization element. The colors are in shades of light blue and white, creating a modern, technological aesthetic.

# **¿QUIÉN LE TEME A LA INTELIGENCIA?**

**Posibilidades y riesgos de  
la Inteligencia Artificial en el Estado digital**

Marcelo Cabrol y Roberto Sánchez A.  
con prólogo de Iván Duque Márquez

# ÍNDICE

**Catalogación en la fuente proporcionada por la Biblioteca Felipe Herrera del Banco Interamericano de Desarrollo**

Cabrol, Marcelo.

¿Quién le teme a la inteligencia?: posibilidades y riesgos de la inteligencia artificial en el Estado digital / Marcelo Cabrol, Roberto Sánchez A.

p. cm. — (Monografía del BID ; 878)

Incluye referencias bibliográficas.

1. Artificial intelligence-Social aspects-Latin America. 2. Artificial intelligence-Moral and ethical aspects-Latin America. 3. Social service-Technological innovations-Latin America. 4. Machine learning-Latin America. 5. Public administration-Automation-Latin America. I. Sánchez Ávalos, Roberto. II. Banco Interamericano de Desarrollo. Sector Social. III. Título. IV. Serie.

IDB-MG-878

Códigos JEL: O38, O35, O32

**Palabras clave:** transformación digital, inteligencia artificial, aprendizaje automático, modernización del estado

**Diseño:** www.spumma.com

**Editora:** Patricia Ardila



Acerca de los autores	7
Prólogo	9
Agradecimientos	13

---

## Introducción I ¿Quién le teme a la Inteligencia? 15

---

### 1 Lo esencial es invisible 21

Soy la IA, para servirle	26
EBDI: un nuevo Estado de bienestar digital inteligente	31
Quiero ayudarte	35
Proteger y servir, pero también vigilar	39
La profundización de la brecha digital	41
Apresúrate lentamente	44

---

### 2 Thomas Southwood Smith y el algoritmo aterrador 49

Una cloaca llamada Londres	51
Von Bismarck pateo el tablero	58
La Guerra, los cinco gigantes y las diferencias filosóficas	61
Más vale predecir que lamentar	65
El COVID-19 y otras pandemias impulsoras de revoluciones	67
De la enfermedad puede salir la medicina	70

---

### 3 Terminator no existe 75

Las máquinas al rescate	80
Siri: ¿de veras estás ahí?	82
Primero lo primero: ¿de qué se trata?	87
Qué tipo de IA y para qué	92
Toma y soporte de decisiones	96
Y aquí estamos	98

<b>4</b>	<b>¿Listos para el despegue?</b>	<b>101</b>
	¿Quiere Estado digital? Llene este formulario	104
	El estado de nuestra tecnología	108
	La IA en el sector salud	112
	Andando despacio se llega lejos	114
<b>5</b>	<b>Alexa: ¿Este sesgo es mío?</b>	<b>119</b>
	¿Predecir por predecir?	122
	No basta con soplar y hacer botellas	126
	Los principales retos de la IA: sus riesgos	129
<b>6</b>	<b>Sonríe; te estamos filmando</b>	<b>149</b>
	La crisis del pueblo uighur y la disputa por Hong Kong	152
	La misma tecnología, otros fines	155
	¿Obedecer como elección?	156
	El caballo de Troya y el COVID-19	157
	La importancia de un <i>like</i>	162
	Metadatos: el capitalismo de vigilancia	164
	Súper empresa: me parece que esto es mío	166
Atención, América Latina: las cámaras ya llegaron	170	
<b>7</b>	<b>De las leyes para caballos de 1865 al mundo de hoy</b>	<b>175</b>
	Lo técnico: controlar el sesgo y crear herramientas de explicabilidad	180
	Lo ético: ¿qué tan responsable es su algoritmo?	182
	Lo legal: para morder se necesitan dientes	191
	<b>Epílogo   Aumentar el ritmo con responsabilidad</b>	<b>207</b>
	<b>Notas de los capítulos</b>	211
	<b>Bibliografía</b>	231

## ACERCA DE LOS AUTORES

**Marcelo Cabrol** es el Gerente del Sector Social del Banco Interamericano de Desarrollo. Allí lidera a un equipo multidisciplinario que apoya a los países en el diseño de soluciones de política pública para reducir la pobreza y mejorar los servicios de educación, trabajo, protección social y salud. Previamente fue Gerente de Relaciones Externas del BID, y entre 2007 y 2012 se desempeñó como Jefe de la División de Educación de la misma institución. Marcelo tiene una licenciatura en economía y ciencias políticas de la Universidad del Salvador de Buenos Aires, una maestría en políticas públicas de la Universidad de Georgetown y completó estudios de doctorado (ABD) en gobierno y políticas públicas en la misma institución.

**Roberto Sánchez** es asesor del Sector Social en el Banco Interamericano de Desarrollo. Allí colabora con fAIR LAC, una iniciativa orientada a promover la aplicación ética y responsable de la inteligencia artificial en la región. Como especialista en el desarrollo de proyectos de ciencia de datos, se ha dedicado al diseño de políticas públicas en temas como focalización de programas sociales, análisis de riesgo de desastres, educación y predicción de la violencia. Ha sido consultor del Banco Mundial y del BID en temas de arquitectura empresarial y aprendizaje automático. Roberto es economista del Instituto Tecnológico Autónomo de México (ITAM) y tiene una maestría en ciencia de datos de la misma institución.

## PRÓLOGO

**“[...] estamos frente a la oportunidad inigualable de no rezagarnos y aprovechar plenamente el potencial y los avances que esta tecnología traerá para el bienestar y la calidad de vida de todos nuestros ciudadanos.”**

La pandemia del COVID-19 no solamente marcará uno de los hitos definitorios de nuestro siglo por sus efectos devastadores, sino que además constituye el punto de inflexión hacia la tecnificación generalizada de la vida cotidiana, en gran parte gracias a la aceleración de la inteligencia artificial y la gobernanza digital. Estos avances tecnológicos brindarán soluciones de política pública para problemas locales con impacto global, entre ellos el cambio climático, la deforestación, la toma de decisiones de política pública, la seguridad, la salud, la educación, la lucha contra la corrupción y la generación de empleo. Sin embargo, como se evidencia en *¿Quién le teme a la Inteligencia? Posibilidades y riesgos de la inteligencia artificial en el Estado digital* de Marcelo Cabrol y Roberto Sánchez, existen también riesgos, debates y preguntas que deben conducir al diseño de marcos normativos, y sobre todo éticos, adaptados a nuestra región y respetuosos de las libertades civiles y los sistemas democráticos.

Durante la pandemia, el comercio electrónico y las plataformas de servicios en línea dinamizaron el consumo. Entre tanto, la telemedicina, la digitalización de las historias clínicas y la inteligencia artificial aplicada se convirtieron en una necesidad imperiosa en los sistemas de salud, mientras que los apoyos electrónicos en el campo de la educación permitieron vislumbrar las transformaciones que eventualmente se producirán en los modelos educativos actuales. Estas herramientas tecnológicas, junto con otras que hicieron su debut en medio la crisis, han llegado para quedarse.

Al mismo tiempo el distanciamiento social, necesario para detener el avance del coronavirus, redefinió no solo las relaciones interpersonales sino el vínculo entre el Estado y sus ciudadanos. En junio de 2018, el Banco Interamericano de Desarrollo (BID) había publicado *El fin del trámite eterno: ciudadanos, burocracia y gobierno digital*, un recordatorio de la lenta modernización de nuestras administraciones públicas, de la falta de herramientas digitales para acercar a los ciudadanos al Estado, y de las carencias de infraestructura digital y legislación pública en esa materia. No obstante, la expansión del Sars-CoV-2 por toda América Latina y el Caribe aceleró la implementación de trámites que ahora son más ágiles, digitales y cercanos al ciudadano, y dejó en evidencia la necesidad latente de cerrar las brechas digitales que aún persisten en la región.

En su obra, Marcelo Cabrol y Roberto Sánchez resaltan la necesidad de continuar con este proceso de digitalización del Estado, especialmente en lo que concierne a los servicios dirigidos a la ciudadanía, y hacen una radiografía de los avances y rezagos que se registran en América Latina y el Caribe en el uso de la tecnología y la inteligencia artificial en función de las necesidades de los ciudadanos. Desde su experticia e investigación, los autores aterrizan el concepto de inteligencia artificial y analizan sus ventajas, desventajas, retos y riesgos de aplicación en el Estado digital, al tiempo que nos invitan a construir un marco jurídico y ético para el desarrollo responsable y consciente de sus herramientas en múltiples campos sociales.

Cuando pensamos en la posibilidad de un Estado de Bienestar Digital Inteligente (EBDI), las áreas de aplicación son innumerables: seguridad y defensa nacional, salud pública y prevención de epidemias, políticas públicas, gestión de riesgos y lucha contra la corrupción, entre otros. Aun así, no creo que el modelo de EBDI vaya a ser solo uno; cada país deberá tener la capacidad de desarrollar e implementar el modelo que mejor se ajuste a sus necesidades y sea coherente con las exigencias

sociales, el respeto por la privacidad, los derechos civiles y el sistema democrático.

Pero también es cierto que cuando pensamos en la inteligencia artificial, salen a flote las distopías de un Estado panóptico que controla a sus ciudadanos por medios digitales interconectados, como en la obra *1984* de George Orwell. Por ello, filósofos como el surcoreano Byung-Chul Han o el esloveno Slavoj Žižek han teorizado sobre lo que podría ocurrir si no se establecen límites éticos claros entre los desarrollos tecnológicos, la vida humana e incluso la democracia.

En este libro sus autores evidencian lo valiosa que fue la inteligencia artificial durante la pandemia en niveles tan sencillos como insospechados. Igualmente nos conducen por un interesante recorrido a través de la historia donde nos muestran cómo crisis similares han catapultado revoluciones tecnológicas de gran magnitud que llevaron a replantear el papel y las responsabilidades del Estado frente a la ciudadanía: desde el Estado reactivo de la primera revolución industrial, pasando por el Estado de bienestar y el Estado promotor de derechos, hasta llegar a lo que ellos denominan el Estado de Bienestar Digital Inteligente (EBDI), que comienza a asomarse caracterizado por su vocación predictiva.

Cabrol y Sánchez también se empeñan en disipar los temores infundados que despierta una tecnología que, como la inteligencia artificial, no es solo “invisible” y difícil de explicar, sino que ha sido asociada, en los productos culturales de ficción, con la sustitución del discernimiento humano por el de las máquinas. Los autores se encargan de desmontar este mito señalando cómo, detrás de cada modelo, detrás de cada programa, está la inteligencia de los seres humanos que los diseñan y programan, y de quienes toman las decisiones sobre su aplicación.

La necesidad de que el Estado formule una política pública y un andamiaje sostenible para asegurar la implementación responsable y ética de la digitalización, el uso y el almacenamiento

de los datos —especialmente de aquellos altamente sensibles como los que tienen que ver con la información médica de cada individuo— es otro de los temas recurrentes en este libro. Lo mismo en lo que se refiere a los aspectos relativos a los riesgos tecnológicos de la inteligencia artificial, principalmente en la construcción de algoritmos que mantengan y repliquen los sesgos y prejuicios culturales de quienes los programan. Cabrol y Sánchez señalan cómo su reproducción --consciente o inconsciente-- en la recolección y/o el análisis de los datos usados en las decisiones de política pública, podría tener impactos adversos significativos en la calidad de vida de millones de personas. Por esto, comparto con los autores la convicción de que es necesario formular políticas que busquen ante todo la equidad social de la mano del emprendimiento y las actividades económicas emergentes. Para ello es indispensable contar con principios claros desarrollados en marcos éticos y auditables.

A lo anterior se agrega la necesidad de garantizar un manejo responsable de los datos que respete la vida privada de los individuos, sus derechos y la preservación de los sistemas democráticos frente a los riesgos de la hipervigilancia. Los autores resaltan el hecho de que, durante la última década, el uso de los metadatos ha mostrado que así como estos pueden ayudar a predecir y controlar epidemias, también pueden contribuir a vulnerar los procesos democráticos. De que sepamos aprovechar plenamente estas oportunidades y neutralizar sus riesgos dependerá la implementación correcta de la inteligencia artificial en nuestra región. Por ello es importante, como bien señalan los autores, que contemos con marcos legales que permitan hacerlo de manera responsable, para lo cual se requiere una gobernanza sólida de la IA que se ocupe de sus retos técnicos, éticos y sociojurídicos.

En suma, esta publicación nos invita a emprender estas tareas impostergables en nuestros países. Ya se están dando algunos pasos significativos en esa dirección. En Colombia, por ejemplo, desde el 2019 contamos con el primer Centro para la Cuarta Revolución Industrial en la región. Se trata de una institución que busca no solo constituirse en el “puerto de llegada” de los avances

de la inteligencia artificial, sino también en un polo de desarrollo del pensamiento crítico latinoamericano sobre su regulación e implementación ética y responsable en nuestras sociedades.

Haciendo eco al mensaje de Cabrol y Sánchez, estamos frente a la oportunidad inigualable de no rezagarnos y aprovechar plenamente el potencial y los avances que esta tecnología traerá para el bienestar y la calidad de vida de todos nuestros ciudadanos.

**Iván Duque Márquez**  
**Presidente de la República de Colombia**

## AGRADECIMIENTOS

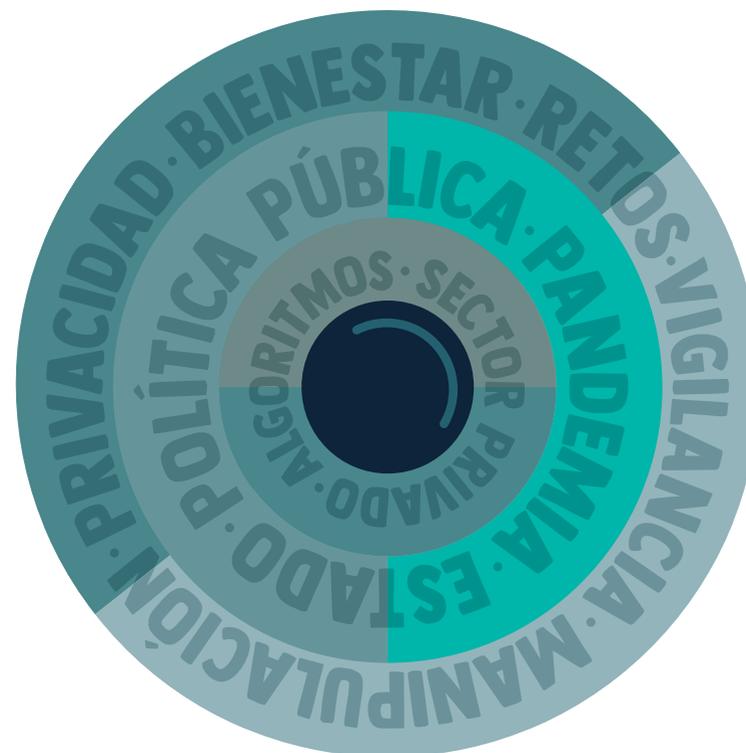
Por su tiempo y por sus valiosos aportes, expresamos un reconocimiento especial a Cristina Pombo, y a Diego Fonseca por ayudarnos a aligerar el lenguaje técnico y a hacer lo complicado comprensible.

A Carlos Tromben y a Max Alberto Gonzales adeudamos nuestra gratitud por ayudarnos a mapear casos de IA en la región, y a Juan Pablo Mena, José Tomás Arenas, Jacson Fressatto y al equipo de Juan Gustavo Corvalán por responder a sus preguntas.

Agradecemos asimismo los valiosos comentarios y observaciones de Joshua Entsminger, Alejandro Noriega, César M. Buenadicha Sánchez y Carolina Aguerre.

También queremos dar las gracias a Patricia Ardila por la edición de estilo del original en español, a Mariana Fagundez por la traducción al portugués, y a Pamela Murphy y su equipo de colaboradores por la diagramación y diseño del libro.

Estamos igualmente agradecidos con Norbert Schady, Ana María Ibáñez, Pablo Picón y Natalia González Alarcón por su apoyo en la publicación y difusión de esta obra.



## INTRODUCCIÓN

### ¿Quién le teme a la Inteligencia?

*Un vistazo inicial, a modo de guía,  
a las páginas de este libro.*

En octubre de 2019 Philip Alston, relator especial para la pobreza extrema y los derechos humanos de la ONU, escribía una nota en donde pedía cautela en el uso de la inteligencia artificial (IA). Allí señalaba que estábamos comenzando a ceder terreno en la protección de la privacidad y en otros derechos humanos, “embobados” por las promesas de la tecnología del nuevo Estado de bienestar digital. Advertía que el uso de la IA por parte del Estado se estaba enfocando principalmente en la disminución de costos y el incremento de la eficiencia, decisiones estas que no han hecho otra cosa que crear sistemas que “atacan, vigilan y castigan” a las personas más pobres. Vigilancia, manipulación y pérdida de privacidad son elementos constitutivos de un futuro distópico que debemos evitar a toda costa.

La fascinación que producen los avances tecnológicos nos acompaña desde siempre. La primera vez que los habitantes de Manhattan vieron a Thomas Alva Edison iluminar dos calles de la ciudad, el llanto y el asombro que los embargaron quizás no fueron muy distintos al sentimiento que produjo a nuestros ancestros primitivos el descubrimiento del fuego. Solemos conferir poderes extraordinarios a lo que desconocemos y se nos presenta inasible, infalible, inalcanzable y, por lo tanto, absolutamente poderoso. Hoy día, esa suerte de fascinación irracional hace que políticos y tecnólogos creen tener la solución a cualquier problema, como si fueran propietarios de la fórmula de la alquimia virtual.

Una de las grandes ventajas de la IA es que es incansable. Los sistemas inteligentes no necesitan comer y tampoco duermen. Operan 24/7, todo el año. Y lo hacen con una intensidad, profundidad y velocidad imposibles de alcanzar para los seres humanos. Su desarrollo se está abaratando, de modo que en menos de cinco años tendremos sistemas de toma y soporte de decisiones en todos los procesos de la administración pública. Proveedores de servicios en la nube como el AWS de Amazon o Google Cloud están creando servicios autoadministrados

que son tan fáciles de incluir en una plataforma como una línea de código.<sup>1</sup>

Pero si queremos que el proceso de adopción sea tan inteligente o más que la tecnología misma, debemos aproximarnos a la IA con una especie de escepticismo lúcido o, si se quiere, con una confianza crítica. Ni por ser nueva es buena, ni por ser fácil es necesaria, ni por ser brillante es imprescindible. El cómo y el por qué son sustantivos a cualquier toma de decisiones sobre su uso. Para ello debemos entender primero sus riesgos, aunque alejándonos al máximo de esa percepción distorsionada de una IA que tomará el control del mundo y acabará con la raza humana. No. La IA no es intrínsecamente mala. Tampoco es moralmente consciente. Todas sus decisiones se basan en la combinación de la información con la que se la ha entrenado y en las pequeñas decisiones que sus diseñadores toman a lo largo de su desarrollo. Sin desconocer sus riesgos, en este libro intentamos disipar temores infundados y enfocarnos en los retos técnicos, éticos y legales que tendremos que resolver para guiar el desarrollo de una IA responsable.

La IA es la gran revolución tecnológica del presente; es un nuevo paso en las grandes revoluciones industriales previas que conmocionaron y transformaron al mundo. Y no podemos asumir un cambio inevitable desde el temor, sino desde el discernimiento. La IA puede aportar enormes beneficios al Estado de **bienestar** —uno nuevo, digital e inteligente— si tomamos con seriedad su construcción desde el plano técnico, ético y legal, y si la consideramos como un instrumento adicional útil para hacer lo que debe hacer: informar y apoyar la toma de decisiones de las personas a cargo de formular la política pública. Los motores de IA no son la Skynet independiente y sagaz de Terminator; son algoritmos entrenados y utilizados, de forma responsable o no, por nosotros los humanos.

---

<sup>1</sup> En el capítulo 5 hablamos sobre los riesgos que los servicios autoadministrados y los algoritmos de “caja negra” presentan al no estar entrenados ni validados para el caso de uso específico.

La idea es discutir aquí cómo el desarrollo tecnológico y la IA pueden contribuir a crear un Estado de bienestar digital inteligente (EBDI). Buscamos entonces introducir el nuevo Estado que se vislumbra y comienza a adquirir forma a partir de los profundos cambios tecnológicos que están transformando las dinámicas económicas, sociales y culturales de nuestro tiempo. En tal sentido, no nos corresponde presentar un modelo armado y acabado de ese Estado --que aún no termina de definirse--, sino simplemente describir, desde una perspectiva histórica, los eventos que, como la pandemia del coronavirus, han transformado sus responsabilidades con el ciudadano.

Arrancamos en el capítulo 1 con la pandemia del coronavirus, carta de presentación más o menos formal de la necesidad global de incorporar las herramientas digitales y la IA como soporte de decisiones críticas de política pública por parte de un Estado que nunca ha sido estático y que se adapta al cambio tecnológico y a las exigencias sociales.

En el capítulo 2 haremos un recorrido sobre la evolución de las responsabilidades del Estado a lo largo del tiempo, pasando de las acciones meramente reactivas con las que respondió a los desafíos de la Primera Revolución Industrial, a los enfoques de promoción de derechos típicos del siglo XX. En las últimas tres décadas, los choques exógenos que han modificado a diario la vida de millones de familias en la región y en el mundo han sido el cambio climático, la violencia, la lucha contra el crimen, la automatización y la transformación de los mercados de trabajo. Estos escenarios determinan las características especiales de una nueva versión del EBDI, es decir, un Estado que con digitalización y modelos<sup>2</sup> predictivos pueda reorientarse hacia

<sup>2</sup> Los términos algoritmo y modelo se suelen utilizar como sinónimos, pero tienen un significado distinto. Se le llama algoritmo al procedimiento o a la serie de pasos que se deben ejecutar para obtener un resultado deseado. En aprendizaje automatizado, un modelo es el resultado de entrenar un algoritmo (por ejemplo, un árbol de decisión, red neuronal, etc.) con un grupo de parámetros y una base de datos de entrenamiento específicos.

la disminución de los riesgos de desastres<sup>3</sup> y hacia la creación de resiliencia<sup>4</sup> en las poblaciones más vulnerables y marginadas.

En el capítulo 3 definiremos qué es y qué no es la inteligencia artificial y de qué modo puede ser utilizada por el Estado para tomar mejores decisiones de política pública. En este punto, América Latina y el Caribe (ALC) —que ocupa un capítulo aparte— debe acelerar sus procesos de modernización para no dejar pasar de largo una vez más el tren de la Historia. En el capítulo 4 describiremos algunas experiencias promovidas en la región por el sector privado, siendo este el que más rápido se ha volcado, por necesidad competitiva, a la inteligencia artificial. Allí constataremos algunas articulaciones entre las empresas y los Estados latinoamericanos que han permitido desarrollar algunos proyectos de IA en la política pública.

Los principales retos de la IA en términos de los riesgos que puede entrañar son objeto de una mirada inquisitiva en el capítulo 5. Y para ello nos abocaremos a considerar factores críticos. Alimentar la IA implica utilizar enormes cantidades de información sensible. Los datos están allá afuera: imágenes satelitales, registros administrativos, archivos médicos electrónicos, estadísticas de consumo de electricidad, patrones de movilidad detectados por teléfonos inteligentes alimentan constantemente una serie de bases de datos que se actualizan a diario. Los algoritmos de aprendizaje automático (*machine learning*) brindan al Estado un nuevo conjunto de herramientas para procesar estas fuentes de información y aplicar políticas **predictivas**. El nuevo Estado puede —y debe— tomar decisiones en tiempo real con toda la información disponible. Pero conviene ser cautos. La experiencia nos muestra que una IA mal diseñada, regulada o controlada puede ser perjudicial.

<sup>3</sup> La reducción de riesgos de desastres es uno de los componentes de la gestión de riesgos y se define como la práctica de disminuirlos mediante el análisis de sus factores causales y de concentrar sus estrategias en prevención, mitigación y adaptación.

<sup>4</sup> Capacidad de adaptación de un sistema a un agente perturbador o a un estado o situación adversos.

La incorporación de la IA dentro del Estado puede ser una oportunidad única para reducir brechas de crecimiento y desarrollo en la región, pero al mismo tiempo puede transformarlo en el aparato vigilante acerca del cual nos advierte Philip Alston en su nota. La pregunta de los próximos años es cómo lograr un Estado inteligente que no se convierta en distopía. Y no solo el Estado. En muchos casos, a punta de capturar y analizar las pequeñas decisiones que tomamos cotidianamente, en ese “capitalismo de vigilancia” al que hace alusión la crítica de tecnología Shoshana Zuboff, las empresas han asumido un papel cada vez más importante en la vida de las sociedades occidentales. Dos grandes preocupaciones, a saber, el surgimiento de un Estado de vigilancia (el “súper Estado”) y de empresas mono u oligopólicas (las “súper empresas”), son objeto del capítulo 6. Ambos procesos se sostienen en el uso de la IA con objetivos nocivos de control social o dominancia del mercado.

La discusión sobre la IA es, en el fondo, una vieja discusión sobre la condición humana, tema del capítulo final. ¿Cómo hacemos para construir sistemas basados en la justicia y el respeto por los derechos humanos? ¿Cómo controlamos nuestros sesgos y mitigamos los riesgos técnicos a la hora de diseñar una IA al servicio de la sociedad? ¿Cómo regulamos y controlamos su creación y uso sin inhibir la innovación, y sin dar carta blanca a sistemas que perpetúen la desigualdad? Y por último: ¿Cómo puede América Latina formular estrategias regionales para apoyar la construcción responsable de un Estado de bienestar digital inteligente o EBDI?

Los invitamos a que se unan a este recorrido que nos lleve a proponer un nuevo Estado de bienestar digital inteligente usando la razón como herramienta, discutiendo los puntos favorables y poniendo bajo la lupa los asuntos más críticos.



## CAPÍTULO 1

### LO ESENCIAL ES INVISIBLE

*La pandemia del coronavirus probó que la inteligencia artificial es de enorme ayuda para un Estado moderno. ¿La clave? Cómo incorporarla.*

Hemos llegado a la Luna y procurado conocer el espacio e incluso conquistarlo. Hemos enviado sondas a lo más profundo del océano. Hemos erradicado decenas de enfermedades. Hemos reducido el hambre y la pobreza. Hemos acumulado más conocimiento que cualquier otra sociedad a lo largo de la Historia. Vivimos más tiempo.

Y sin embargo, somos vulnerables.

A inicios de 2020, un virus desconocido hizo su aparición, supuestamente en un mercado de productos de mar de Wuhan, una ciudad de cerca de 11 millones de habitantes capital de la provincia de Hubei en China central; el resto es historia reciente y conocida. En menos de dos meses, el planeta se había detenido. Miles de millones de personas quedaron confinadas en sus hogares, al tiempo que la economía global iniciaba su caída libre hacia profundidades quizás nunca vistas. Millones de personas se contagiaron del coronavirus SARS CoV-2 y cientos de miles murieron por la enfermedad brutal que produce, la COVID-19,<sup>5</sup> que destroza los pulmones en muy poco tiempo.

La humanidad no había visto nada igual en cien años.

En 1918, soldados de Estados Unidos llevaron a las trincheras de la Primera Guerra Mundial un virus desconocido en Europa que fue identificado por un grupo de médicos de España. Desde entonces, la llamada “gripe española”<sup>6</sup> figura como la epidemia más catastrófica de la era moderna, dejando a su paso por lo menos 40 millones de muertos.

Es altamente probable que el COVID-19 jamás alcance esas cotas, pero 1918 era otro siglo, literal y metafóricamente. En ese entonces, la propagación de la peste no pudo ser controlada

<sup>5</sup> En adelante nos referiremos al virus y a la pandemia como “el” COVID-19, tal como lo aconseja la RAE, si bien la enfermedad causada por el virus es de género femenino. Véase <https://www.rae.es/noticia/crisis-del-covid-19-sobre-la-escritura-de-coronavirus>.

rápidamente y, aunque se expandió con cierta lentitud a la luz de los estándares actuales, acabó cubriendo todo el planeta. Más de 500 millones de personas —un tercio de la población mundial— padecieron el virus. Ahora la situación es distinta. Si a la gripe española le tomó casi dos años propagarse por todo el mundo, al virus detectado en Wuhan le tomó menos de tres meses llegar a todos sus confines.

El coronavirus es un bicho complicado. Sin llegar a los niveles de mortalidad del ébola y otros virus, su contagio comienza ya en la fase de incubación, cuando la persona es asintomática, y se prolonga por más tiempo que en otras enfermedades; por esa razón, aislar a pacientes infectados es mucho más difícil.<sup>7</sup>

Al igual que cualquier otra amenaza, las enfermedades no surgen en abstracto; nuestra forma de vida les abre toda una gama de vías de entrada. El COVID-19 es perfecto para el mundo de hoy pues somos vectores eficientes de virus y bacterias. Viajamos tanto que el planeta se ha encogido. Los billetes aéreos se han abaratado y hemos conectado cada rincón del mundo con más de 40 millones de vuelos anuales. Solo en 2019 han atravesado los cielos más de 4.540 millones de viajeros. Si en el año 2000 el habitante promedio del mundo volaba una vez cada 44 meses, ahora lo hace cada 21. Más de

<sup>6</sup> El nombre de la gripe española castiga al país de manera injusta, como si fuera su punto de origen, cuando en realidad debiera premiarlo, pues mientras que en Francia, Inglaterra, Alemania y Estados Unidos se practicaba la censura y se restringían los reportes sobre el virus, en España, que era neutral, la información circuló libremente y retrató el horror de la enfermedad. Si sabemos de ella es porque allí la conversación fue indudablemente más fluida que en las naciones que decidieron ocultarla por razones de táctica geopolítica. Más aún, España no resultó especialmente afectada por la epidemia en comparación con las otras naciones en guerra, pero el tratamiento abierto que dio a la crisis asoció irremediamente su nombre al virus.

<sup>7</sup> La mayoría de las estimaciones sitúan la tasa de mortalidad por debajo del 3% y el número de transmisiones entre 2 y 4%, según un informe de *The New York Times*.

22.000 ciudades de todo el mundo están conectadas por algún vuelo doméstico o internacional.<sup>8</sup>

Sobre las alas de los aviones hemos configurado una fabulosa red de intercambios económicos, culturales y sociales, aunque al mismo tiempo, y sin pensarlo demasiado, hemos tejido con ello la telaraña de los contagios.<sup>9</sup> Nuestros contactos actuales son más estrechos y prolongados que en cualquier otra época de la humanidad. Por eso, una persona podría haber recorrido los cinco continentes antes de empezar a mostrar síntomas de coronavirus.

No deja de ser interesante que, sin una vacuna desarrollada, el mundo de 1918 y el de 2020 hayan coincidido en el tratamiento de sus respectivas pandemias enfocándose en procedimientos completamente analógicos empleados desde hace siglos: higiene personal, aislamiento, cuarentenas y rastreo de contactos.

Sin embargo, el hecho de que actualmente la humanidad registre un número de muertes muy inferior al que se produjo durante la pandemia de 1918 se debe en buena medida a que el desarrollo científico y tecnológico es incomparablemente más vasto que un siglo atrás. Incluso sin una vacuna, tratar las enfermedades hoy es mucho más efectivo gracias a todo el inventario de antibióticos, infraestructura hospitalaria y equipamiento sofisticado que poseemos para mantener con

---

<sup>8</sup> Solo en Estados Unidos, el promedio de vuelos diarios alcanza cerca de 87.000, de los cuales 2.150 son de carga y 28.800 comerciales; asimismo, los controladores aéreos de ese país dan vía cada año a 64 millones de despegues y aterrizajes. Estados Unidos, como parte de América del Norte, representa 22% del tráfico global de pasajeros y carga.

<sup>9</sup> Eso no es todo. A ello se suman 40 millones de personas que viajan en cruceros, así como el transporte de carga aéreo y marítimo. Este último, invisible a los ojos de la ciudadanía, es una de las actividades centrales que engrasa el sistema de intercambios. En el mundo existen más de 50.000 buques de gran tamaño que realizan conexiones en 150 países y movilizan el 90% de la carga global, según la International Chamber of Shipping.

vida a los enfermos de un modo inimaginable hace 100 años.

De hecho, la tecnología empleada para reducir las tasas de contagio ha sido un factor clave en aquellos países que la han sabido aprovechar, puesto que vivimos en un mundo abiertamente tecnologizado que dota a los gobiernos de una capacidad superior para comunicar, vigilar y controlar. Existe una diferencia gigantesca entre visitar a cada paciente y, con lápiz y papel, solicitar la lista de personas que este ha visto en los últimos días —como hacían los médicos a principios del siglo XX— y hacer lo mismo a través de una consulta automatizada en el celular,<sup>10</sup> más veloz, segura y sencilla de procesar. La toma de decisiones es hoy día más eficiente en circunstancias en las que cada segundo cuenta.

Precisamente, y para lidiar con una amenaza sanitaria invisible como lo es el virus que causa una enfermedad respiratoria grave, gobiernos, médicos y científicos han recurrido a un avance tecnológico igualmente invisible, en el sentido de que la esencia de su trabajo la realizan algoritmos que permanecen ocultos. Durante la pandemia, la inteligencia artificial (IA) —ese conjunto de programas computacionales capaz de realizar en muchísimo menos tiempo el trabajo de miles de cerebros humanos— ha hecho su presentación pública ante la sociedad global, en su calidad de sistema de toma y soporte de decisiones. Más aún, la ayuda que diversas herramientas digitales y sistemas de IA están prestando a funcionarios, médicos y científicos ha sido tan valiosa que nos ha hecho conscientes de que, más allá de la salud, su potencial de uso para numerosas políticas públicas es enorme. No exageramos al decir que hemos visto al futuro desplegarse ante nuestros ojos.

---

<sup>10</sup> Corea del Sur decidió que cada viajero que llegara al país debía bajar una aplicación en su móvil para reportar a través de ella su estado de salud diario.

## Soy la IA, para servirle

Decíamos que la experiencia frente al coronavirus ha sido la carta de presentación más llamativa de la IA en el escenario de las políticas públicas. Programas computacionales que producen soluciones hasta ahora nunca vistas actúan como los oráculos que siempre hemos buscado: mecanismos de ayuda que nos permitan tomar las decisiones más eficientes y mejores para nuestras familias.

Tras el COVID-19 no quedan dudas: la IA es útil pues mejora a grandes saltos las capacidades operativas del Estado y de las empresas (e incluso de los individuos) que deben afrontar crisis profundas y generalizadas.

La experiencia de la IA en el sector sanitario ha sido notable. En Taiwán, China, Japón y Corea se utilizaron sistemas digitales de recolección y procesamiento de información con distintos objetivos para controlar la pandemia. Entre otras aplicaciones se valieron de modelos de IA para identificar individuos sanos, portadores y enfermos y/o hacer la vigilancia de la ciudadanía en lo relativo al acatamiento de guías y órdenes de las autoridades. En Taiwán, donde la información socializada fue crucial, los funcionarios a cargo crearon mapas actualizados en tiempo real con insumos tomados de GPS, GoogleMaps y actualizaciones en redes sociales para alertar a las personas sobre áreas de riesgo a evitar o sobre la ubicación de tiendas donde adquirir mascarillas y desinfectantes.

Una de las naciones que más integró sistemas de soporte de decisiones basados en IA en su respuesta a la pandemia fue China.<sup>11</sup> En 2017, el gobierno de ese país había lanzado su “Plan de desarrollo de la siguiente generación de inteligencia artificial”, el cual fue desplegado plenamente durante la crisis

<sup>11</sup> El uso de la tecnología por parte del gobierno chino ha sido muy cuestionado por razones que se tratarán más adelante, principalmente en el capítulo 6 (vigilancia, violación de derechos humanos, manipulación, entre otras).

del coronavirus. Todo lo que ha rodeado a ese lanzamiento es llamativo a los ojos de un planificador de política pública o de cualquier persona con una mínima capacidad de asombro: la supercomputadora que combina cientos de imágenes de tomografías de tórax para arrojar una predicción sobre el COVID-19; la secuenciación del genoma del SARS CoV-2 que los laboratorios chinos resolvieron en enero de 2020, a menos de un mes de oficializado el descubrimiento de los primeros casos y a dos de las primeras evidencias no oficiales; los modelos que permitieron desarrollar una prueba para detectar el coronavirus en 30 minutos;<sup>12</sup> y por último Linearfold, el algoritmo predictivo del gigante Baidu que declaró tener la capacidad de reducir los tiempos de detección a solo 27 segundos.<sup>13</sup>

Fue así como todo un universo de IA comenzó a asomar ante nuestros ojos sorprendidos y maravillados. Empresas emergentes chinas entrenaron en solo unos días motores inteligentes capaces de predecir contagios con un margen de error inferior al 4%. En las estaciones de trenes de las principales ciudades del país se instalaron escáneres termales que registraban la temperatura de cada pasajero y notificaban la identidad de los posibles enfermos al centro de control ferroviario para que remitieran a evaluación a cualquier viajero que se encontrara en las proximidades en el mismo vagón.

Una de las principales experiencias de IA aplicada a la sanidad se vivió de puertas para adentro en los hospitales. Los nuevos,

<sup>12</sup> Según registros de prensa, el algoritmo nutre una plataforma de detección del genoma completo que es “más veloz y mejor” que la aproximación tradicional con nucleotidos, que solo descifra una porción del ácido ribonucleico o ARN.

<sup>13</sup> La Academia AliBaba DAMO, el brazo de investigación de la compañía china, también entrenó un sistema de IA para reconocer coronavirus con una precisión declarada del 96%. Este podría procesar hasta 400 escaneos necesarios para diagnosticar un coronavirus en 20 a 30 segundos versus los 10 a 15 minutos que le tomaría a un médico experimentado. Según la empresa, su algoritmo permitió que al menos 26 hospitales chinos revisaran más de 30.000 casos. También DeepMind, una compañía inglesa de tecnología adquirida por Google en 2010, puso a su IA a trabajar en la secuenciación.

como el construido en solo 10 días en Wuhan, obtuvieron una dotación de robots que hace algunos años habrían sido material de una película de ciencia ficción. Los robots dominaban la escena allí con el propósito de reducir la carga de trabajo y el riesgo para la salud de médicos y enfermeras: tomaban la temperatura de los enfermos, les notificaban los resultados y enviaban toda la información médica a la nube para alimentar a otros motores de IA con el objeto de buscar patrones.

Vacunas, genómica, detección, predicción. Mientras que los médicos batallaban para evitar contagiarse dadas las dificultades de los sistemas sanitarios, en los laboratorios de todo el mundo se combatía la pandemia con tecnología de punta. Los experimentos con elementos bioinformáticos —modelos que emplean componentes biológicos— permitirán a laboratorios de Asia, Europa y Estados Unidos acelerar los tiempos para obtener vacunas.<sup>14</sup> La vanguardia de Silicon Valley —Amazon, Microsoft, Google, IBM, entre otros— puso su poder computacional a procesar larguísimas series de datos epidemiológicos que

<sup>14</sup> El tiempo que toma desarrollar vacunas se ha reducido sustancialmente en los últimos años. Obtener la vacuna de polio tomó 47 años, 42 la de la varicela y 43 la del ébola. Durante el pico de la pandemia, diversos especialistas proyectaban que desarrollar la vacuna del COVID-19 tomaría entre 12 y 18 meses, al menos por dos razones: por un lado, debido al conocimiento previo que se tiene de otros coronavirus (como SARS o MERS), y por otro, en razón de las mejoras tecnológicas en bioingeniería y secuenciación genético. En 2019, la Universidad de Adelaide, en Australia, utilizó simulaciones computacionales de trillones de compuestos químicos en un esfuerzo por identificar los mejores candidatos para activar el sistema inmunológico y dotarlo de la capacidad de hacer frente a distintas enfermedades. Sin estos algoritmos sería imposible evaluar tantas combinaciones y reducir los tiempos. Stuart Thompson, editor de Opinión de *The New York Times*, trató de poner en perspectiva los tiempos de producción de la vacuna del SARS CoV-2. La meta, auxiliada por la IA, fijaba a agosto de 2021 como fecha de distribución, lo que significaría que los ensayos de fase 1, 2 y 3 (que incluye a humanos) y el inicio de producción tendrían que ocurrir entre el segundo semestre de 2020 y primer trimestre de 2021. Nada más como ejercicio hipotético, se trata de una tarea monumental que, de resultar, revolucionaría la ciencia médica. En cambio, según las proyecciones de Thompson, con los métodos tradicionales de investigación, prueba y producción, una vacuna para el SARS CoV-2 recién estaría disponible en mayo de 2036.

produjeron modelos bioinformáticos y moleculares para informar a las autoridades de Estados Unidos.<sup>15</sup> Entre tanto, investigadores del Hospital Infantil de Boston desarrollaron un motor de IA para seguir el desplazamiento del virus integrando búsquedas de Google y publicaciones en redes sociales, blogs y foros —fuentes que los epidemiólogos no consultan— para detectar las primeras señales de un brote, tal como lo hizo la empresa social canadiense BlueDot, primera en predecir el salto del COVID-19 de China a otros países de Asia.

Que no se malentienda, ya que será un tema recurrente en el libro: Sin la capacidad institucional y el interés político por implementar sus recomendaciones y un uso responsable, la tecnología sirve de poco y hasta puede ser perjudicial. Cabe aquí mencionar el ejemplo africano, que demostró que se puede abordar de manera efectiva una pandemia con tecnologías limitadas, pero con un claro compromiso institucional. Aunque allí se empleó la ciencia de datos para modelar patrones y mapas de contagio, la respuesta de las naciones africanas se

<sup>15</sup> Hasta hace cinco años —una década en el peor de los casos—, los científicos debían operar con bases de datos y archivos en línea más o menos integrados para analizar casuística. Si bien estos se encontraban ordenados, bien clasificados y tenían una impecable categorización, había un problema irresoluble: eran demasiados. Como es obvio, el exceso de información narcotiza y, al final, paraliza. Ahora la IA se desempeña como ese bibliotecario gentil al servicio de lo que la ciencia precisa. Al respecto cabe notar, por ejemplo, que mientras que en las semanas que siguieron al brote de Wuhan se publicaron más de 2.000 trabajos científicos sobre tratamientos posibles y sobre la evolución del virus, Microsoft, la Biblioteca Nacional de Medicina de Estados Unidos y el Allen Institute for AI comprimieron y prepararon casi 30.000 documentos científicos sobre las familias de coronavirus para apoyar la labor de los especialistas de todo el mundo. Otra compañía de Google, Kaggle, inyectó octanos al proceso de acumular conocimiento: si tenemos los datos, si hemos analizado y **ya sabemos lo que podemos saber** hasta el punto X, ¿qué más podemos aprender? Así que Kaggle lanzó un desafío de 10 preguntas sobre el SARS CoV-2 para que los científicos buscaran respuestas a su comportamiento. Este esfuerzo recibió el apoyo de la Chan-Zuckerberg Initiative, impulsada por Priscilla Chan y su marido Mark, Presidente de Facebook, y del Center for Security and Emerging Technologies de Georgetown University.

basó principalmente en su experiencia reciente en el combate de otras pandemias y crisis sanitarias como el ébola (2013-2016) y la polio, y con un nivel de coordinación regional que otros continentes no lograron. En Nigeria, por ejemplo, los equipos médicos que usualmente se desplazan a las aldeas remotas para vacunar a los niños contra la polio fueron utilizados para educar a los residentes sobre la nueva enfermedad y para implementar medidas de aislamiento y rastreo de contactos.

La promesa de la inteligencia artificial está mostrando resultados, no solo en términos de investigación de vanguardia, sino también en asuntos administrativos como la automatización de tareas repetitivas para que las instituciones puedan incrementar el acceso a todo tipo de servicios. La IA está en muchas partes y puede estar en muchas más. Un Estado moderno no puede ignorarla.

Que la mayoría de estas tecnologías no estuvieran disponibles durante el brote de SARS en 2003 hace aún más evidente cuán veloz ha sido el avance de la IA en el mundo. Es más; parece que su desarrollo se sigue produciendo al ritmo de la “Ley de Moore,” así llamada en honor al cofundador de Intel, Gordon Moore, quien estimaba que el número de transistores en un microprocesador se duplicaría cada dos años. El siguiente ejemplo es ilustrativo. A finales de noviembre de 2020, a solo unas semanas de que se aprobaran las primeras vacunas contra el COVID-19, la empresa Deepmind resolvió un problema surgido 50 años atrás al crear la segunda iteración de su modelo AlphaFold, capaz de predecir la estructura en 3D de una proteína. Esto con una precisión comparable a la de métodos experimentales ya existentes como la cristalografía de rayos X, la cual tiene un costo aproximado de US\$120 mil por proteína y cuyo desarrollo puede tomar hasta un año. Dado que la estructura de las proteínas proporciona información fundamental para entender el comportamiento de las enfermedades, este método tiene un potencial enorme para acelerar el descubrimiento de medicamentos.

El cambio es sísmico, liminar, fundacional. Las empresas de tecnología son conscientes de la importancia competitiva de la IA y, a juzgar por sus acciones durante la crisis, también del valor que tiene y de su propia responsabilidad social. La acción de las tecnologías puede “cambiar la cara de la salud pública” según Larry Brilliant, epidemiólogo y director ejecutivo de Google.org. “Pocas cosas en la vida son más importantes que la cuestión de si las principales tecnologías son demasiado poderosas, pero una pandemia es, sin duda, una de ellas.”

### **EBDI: un nuevo Estado de bienestar digital inteligente**

Mínima definición necesaria: la IA es software. Se trata de un área de investigación enfocada en desarrollar sistemas capaces de realizar tareas consideradas como inteligentes. En política pública, los resultados y la información generados por estos sistemas son utilizados como mecanismos de toma o soporte de decisiones.<sup>16</sup>

El uso de motores analíticos para tomar o asistir decisiones de política pública no es nuevo,<sup>17</sup> pero el COVID-19 ha revolucionado

<sup>16</sup> Esta definición describe de forma amplia a la IA e incluye tanto los sistemas que pueden utilizar algoritmos de aprendizaje automático (*machine learning*) y aprendizaje profundo (*deep learning*) con capacidad de recibir datos y aprender por sí mismos sin decisiones preprogramadas, como también aquellos basados en reglas conocidos como GOFAI (Good Old Fashion AI). Para más detalles sobre este particular, véase el capítulo 3.

<sup>17</sup> Los países llevan años utilizando algoritmos de predicción para tomar decisiones de focalización de la política social. Los sistemas conocidos como “prueba de aproximación de medios” (o *proxy mean tests* en inglés) son modelos multivariados que, a partir de las características de la vivienda de una familia, de la situación laboral de sus integrantes y de otras variables socioeconómicas, hacen una predicción del ingreso de un hogar para luego, a partir de allí, definir si este tiene acceso a programas sociales. Lo anterior evita lidiar con los problemas de subreporte de ingresos que se presentan en las encuestas autorreportadas, lo que a su vez puede dar lugar a problemas de inclusión/exclusión. Estos modelos se solían hacer con regresiones lineales, pero en los últimos años se han comenzado a utilizar metodologías de aprendizaje automático.

la manera de hacerlo y las discusiones que tenemos sobre la tecnología. Antes de la pandemia, y sobre todo como producto de ella, los gobiernos se han visto obligados a reiniciar un debate sobre el estado, calidad y madurez de sus sistemas, principalmente los de salud pública. La dimensión de las acciones que muchas naciones emprendieron para combatir el coronavirus es un borrador para otras más rezagadas y constituye el entramado básico de un modelo de respuesta para crisis futuras.

Hemos construido una vasta red de información y rastreo de la sociedad que resulta crítica durante un desastre sanitario. La interoperabilidad de los registros médicos electrónicos de las personas; los expedientes personales en los sistemas sociales que permiten distribuir apoyos durante la crisis; la intersección de información de fuentes no tradicionales como el GPS de un teléfono móvil o el historial de búsquedas de internet; el análisis de múltiples bases de datos de distinta cobertura capaces de dialogar entre sí; la lectura forense de los registros de boletos de avión; la información publicada y emitida por los periódicos, la radio y los canales de TV y, por supuesto, cuanto dicen y conversan las personas en las redes sociales, todo ello indica, en tiempo real y a cada instante, dónde se encuentran millones de individuos.

Esa información, que parece inocua hasta que se incorpora a un motor capaz de procesarla, clasificarla, analizarla y **juzgarla** en milésimas de segundo, se une a tecnologías de rastreo con la ayuda de GPS y *bluetooth* para identificar posibles contactos prolongados entre una persona sana y una enferma. Una vez procesada y ordenada, constituye un nuevo universo de datos que permite construir modelos epidemiológicos, los cuales facilitan el trabajo de científicos y médicos. Con ello pueden identificar y proyectar **al detalle** las fases de una pandemia para dar recomendaciones sobre cuándo no salir de casa o adónde no ir. En una fracción breve de tiempo —días, unas pocas semanas—, una epidemia puede ser detectada y hasta controlada gracias a las capacidades tecnológicas modernas.

El mundo digital e interconectado ha cambiado, cambia y cambiará la mecánica operativa de los Estados, como en el pasado lo hizo cada revolución tecnológica. La pandemia es una prueba irrefutable de que las herramientas de gestión de la cosa pública no pueden estar ancladas a la morosidad burocrática y analógica de antaño. Un gobierno que sabe más puede tomar mejores decisiones con mayor anticipación y celeridad, y con mayor precisión metodológica, lo cual a su vez le permite usar de manera más eficiente los recursos y lograr una focalización más clara en aquellas comunidades que exigen su atención.

En América Latina, la vivencia de la pandemia ha puesto en evidencia otras emergencias que constituyen el pan de cada día de numerosas poblaciones y para las cuales la IA podría resultar igualmente vital. El cambio climático, la inmigración y la emigración, la violencia y el combate al crimen organizado, la automatización y otros cambios en el mercado de trabajo e incluso el aumento de la esperanza de vida de las personas son choques exógenos que transforman a diario la vida de millones de familias en la región y en el mundo. La IA puede ayudar a gestionar estos procesos apoyando la toma de decisiones de **cualquier** política social planificada que apunte a crear resiliencia. Como ya lo demostró el empleo del herramienta tecnológica para controlar la crisis del coronavirus, el resultado de esas acciones se mide de un solo modo: vidas salvadas.

Volvamos a la pandemia y a la IA como prueba de nuestras afirmaciones. Conocemos las consecuencias de no actuar de manera preventiva y con premura. Durante los primeros meses de respuesta, los Estados que subestimaron el impacto potencial del coronavirus y desatendieron los modelos que proyectaban su avance exponencial sometieron a sus ciudadanos a una crisis tan letal como innecesaria.

Quienes tienen acceso a predicciones y no las utilizan para prepararse son responsables por esas decisiones y sus consecuencias. Hoy día, los ciudadanos esperan del Estado

más que simples acciones reactivas. Y eso es precisamente lo que impulsará la discusión del Estado de bienestar digital inteligente, al que llamaremos EBDI. A la luz de la experiencia con el COVID-19, actualmente exigimos el uso de proyecciones y la definición de las acciones respectivas; pero si el Estado tiene acceso a información que permite acometer problemas sociales, será también responsable de actuar de forma **preventiva**, trátase de un riesgo generalizado de salud como la diabetes, o de la deserción escolar, o de los prospectos de trabajadores cuyos empleos pueden estar comprometidos por un cambio de las condiciones de producción, o de fenómenos potencialmente catastróficos ocasionados por el cambio climático como son las inundaciones, tormentas o sequías.

Este EBDI, preventivo por diseño, no se construye siguiendo una evolución lineal del actual Estado analógico.<sup>18</sup> El Estado de bienestar digital inteligente solo estará dotado de la capacidad de prevenir catástrofes y gravísimos problemas sociales si amasa suficiente información como para que los algoritmos de la IA puedan hacer el trabajo de organizar, clasificar, entrelazar variables, construir modelos con ellas y, a partir de allí, generar una predicción útil para los funcionarios que deben tomar decisiones. Y para obtener esa información —cuyo volumen es, literalmente, inimaginable— es preciso que la administración pública pase un peine extrafino de arriba a abajo a la sociedad, pues la calidad de los modelos (o sea, de la IA) es mejor cuanto más exhaustiva y precisa es la búsqueda.

Entrenar esos modelos nos plantea igualmente nuevos problemas éticos y filosóficos sobre la privacidad, las libertades y los derechos individuales, puesto que se trata de insumos que no son fruto de trabajar la tierra o los materiales: nosotros —quiénes somos, qué hacemos y lo que decimos— constituimos la materia prima de la IA. Y esa realidad puede resultar un tanto problemática.

<sup>18</sup> O, si se quiere, post analógico y cuasi digital, dependiendo de los avances de cada nación.

En las páginas de este libro exponemos la siguiente idea: la historia futura de la pandemia deberá relatar cómo esa batalla silenciosa, apenas imaginable e invisible entre un virus microscópico y *bits* inasibles, abrió la puerta a una nueva era para un EBDI alimentado por motores de inteligencia artificial. Entre tanto, la historia de la humanidad deberá contarnos hasta dónde permitimos que estos motores condicionaran nuestra vida y fueran parte de ella.

### Quiero ayudarte

La película es “2001: Odisea en el espacio”. Para el momento en que se exhibió, la IA no era nada; ni siquiera una moda más allá de las discusiones de científicos barbudos en algunas universidades de prestigio. El año era 1968, época del *flower power*, de la resistencia a la guerra de Vietnam, del Mayo Francés y de los albores de la Primavera de Praga. Los jóvenes hablaban de hacer el amor y no la guerra. A fines de ese año, los Beatles lanzaban su Álbum Blanco, una de las perlas musicales del siglo XX.

Stanley Kubrick, el creador de la película, y Arthur C. Clarke, el escritor británico de ciencia ficción que la convertiría en libro mientras la filmaban, no tenían claro el alcance de la producción, pero acabaron creando una obra maestra en el camino. “2001” se enfoca en temas que eran usuales para la intelectualidad de la época: el avance tecnológico y sus ventajas, promesas y riesgos. La idea del progreso y el colapso posible de la humanidad flota en el ambiente —sobre todo vista desde hoy, cuando exploramos mundos posibles porque no somos capaces de hacer sostenible nuestro propio planeta— y se enfrenta con discusiones filosóficas trascendentales como los pasos probables de la evolución (En el libro, Clarke imagina la posibilidad de que podamos abandonar nuestros cuerpos y vivir como humanoides con carcasas de robots antropomórficos).

Arruinemos la trama para quien sea demasiado joven como para haberla visto. Los miembros de una misión empiezan a sufrir una serie de misteriosos accidentes que ponen en peligro sus

planes. Pronto descubrimos que no hay tales accidentes sino incidentes provocados por Hal-9000, la computadora madre de la nave, una magnífica pieza de inteligencia artificial capaz de razonar y dialogar con los humanos. Hal llevará la crisis a tal extremo que, después de matar al resto de los astronautas, se enfrentará al último tripulante: el primer oficial David Bowman.

Hal presiente que los humanos han cometido errores y que debe tomar el control de la misión, deshaciéndose, como es lógico, del componente fallido de la operación. “Esta misión es demasiado importante para mí [como] para dejar que la pongas en peligro”, le dice en un momento a Bowman. El ojo ígneo intensamente rojo de Hal es una imagen icónica de la película. Ese ojo incrustado en la pared de una unidad de procesamiento, rodeado de un aro metálico, se combina con una voz de perfecta dicción y absoluta despersonalización. Kubrick logra hacernos aferrar con las uñas a nuestras sillas cuando deja la cámara por largos momentos enfocada en ese ojo en completo silencio. Vemos “pensar” a la máquina.

Al final, Bowman, al borde de la desesperación, conseguirá desactivar a Hal. La máquina iniciará un alucinado proceso de apagado en el que enlazará comentarios sentidos, ofrecerá disculpas, cantará “Daisy Bell”<sup>19</sup> —*Daisy, Daisy, give me your answer do /I'm half crazy all for the love of you—* y se humanizará: “Tengo miedo. Tengo miedo, Dave. Dave, mi mente se va. Puedo sentirlo. Puedo sentirlo. Mi mente se va. No hay duda. Puedo sentirlo. Puedo sentirlo. Puedo sentirlo”.

<sup>19</sup> La canción data de 1892 y su autoría es del compositor británico Harry Dacre, un prolífico productor de la Isla de Man. La leyenda oral dice que las estrofas fueron una dedicatoria de Dacre a Daisy Greville, por entonces condesa de Warwick y amante del rey Eduardo VII, pero no hay certeza al respecto. Kubrick la empleó como un guiño al desarrollo creciente de las computadoras. Una novedad fascinante: “Daisy Bell” fue la primera canción entonada por un programa de voz sintetizada por una computadora, la IBM 7094, en 1961, siete años antes del lanzamiento de la película.

Bowman tomará el control, pero mientras comienza a retirar las placas de procesamiento de lenguaje de la computadora, ocurrirá un momento intensamente significativo. “Sé que recientemente he tomado algunas decisiones equivocadas, pero puedo darte mi total seguridad de que mi trabajo volverá a la normalidad”, le dice Hal. “Todavía tengo el mayor entusiasmo y confianza en la misión. Y quiero ayudarte.”

¿Qué sucede aquí? Hal asume el extremo de su antropomorfización como un ser humano corriente. Primero ofrece disculpas y, de inmediato, envía un mensaje tranquilizador: “puedes confiar en mí.” Para entonces, todos los espectadores saben que la computadora está intentando manipular a Bowman. Es sustancial. En el instante final, ante la pérdida de la vida, cualquiera de nosotros prometería lo indecible para mantenerse respirando. Hal, un ser fríamente revanchista, ya insinuaba el desquite posible si todo seguía en su lugar.

Esta imagen de 1968 no es tan distinta a lo que algunas personas piensan ahora: demos a las máquinas la oportunidad y dominarán la tierra. Tenemos una percepción temerosa de la IA, alimentada por el imaginario ficcional del cine. Para muchos, la IA son los robots amedrentadores y los humanoides musculosos de Skynet en Terminator o de WOPR, la supercomputadora que corría simulaciones bélicas en “Juegos de guerra,” hasta que aprende a declarar una de verdad. O, si la prefiere con mayor ternura, está la sugestiva voz de Scarlett Johanson en “Her” o la bella y dulcemente criminal androide de “Ex Machina.” En todos los casos, las máquinas —la IA— hacen lo que quieren con la inteligencia y voluntad humanas. Manipulan. Controlan. Dominan.

Esas aproximaciones ofrecen tramas fantásticas para el cine y la literatura, plenas de meandros y tensión, pero no se acercan a la realidad. De hecho, en ocasiones no hacen sino esconder que, si la IA hace lo que la ficción dice que hace —prometernos paraísos de bondades para, al final, traicionarnos—, la responsabilidad no es de la máquina sino de los humanos. Siempre. Y ahí está Hal 9000 —el “pobre” Hal 9000— para probarlo. Como cerebro de inteligencia artificial, sus capacidades para razonar están determinadas por

el universo de relaciones posibles que se pueden construir a base de la información con que fue programado por **humanos**. La responsabilidad siempre ha sido nuestra.

La película es una compleja exposición sobre la naturaleza de la bondad y el sentido de la existencia reveladas por el juego permanente de relaciones entre los humanos y la máquina. Todos hemos visto en Hal una actitud inquisitorial y dominante —malvada— cuando comienza a fallar, a la vez que procura, extrañamente, mantener la cordialidad. Y hay una razón muy humana para ello: las fallas de Hal son fallas humanas.

Kubrick siempre se propuso dejar toda interpretación sobre las motivaciones de Hal a la febril imaginación del público, pero en la novela Clarke revela lo que no está en la película: Hal no puede desvelar la naturaleza de su misión porque así le fue ordenado. Su reacción criminal es producto de órdenes conflictivas dadas por sus programadores y, al cabo, de una preservación **ética**. Como no puede informar para qué está allí, dadas las órdenes que recibió, Hal comienza a entender que la insistencia y el comportamiento de los astronautas amenaza la misión para la cual otros seres humanos lo han creado.

Detengámonos allí. Hoy, las máquinas “autónomas” no lo son más allá de lo que permiten su alimentación y programación. Más aún, son herramientas —ni sujetos, ni androides, ni tampoco sistemas capaces de escindirse del ambiente de control humano— para que las instituciones, los individuos, las empresas, los colectivos sociales y los Estados tomen decisiones útiles.

Hal 9000 era IA, pero la IA no es la versión tremendista de Hal 9000. ¿O acaso puede serlo?

### Proteger y servir, pero también vigilar

La IA es, en sí misma, bastante inocua. Es un instrumento cuyas capacidades están determinadas por las convicciones, creencias y conocimientos de sus programadores, y por la calidad de la información con que sus productores alimentan el software. Tal y como sucede con Hal, si tiene información conflictiva o deficiente, puede tomar pésimas decisiones.

Ahora pensemos en la siguiente posibilidad. Los gobiernos deciden crear a los hermanos terrícolas de Hal 9000, pero con la intención específica —no casual, no errónea, sino consciente— de construir un Estado de vigilancia.

En los últimos años, los gobiernos han ido introduciendo sistemas alimentados por algoritmos de IA que incluyen software de reconocimiento facial, programas de vigilancia por internet e incluso análisis e inteligencia de nuestras compras con tarjetas de crédito. Por otro lado, en países como los de América Latina que aún no tienen la madurez tecnológica y/o jurídica de Europa, América del Norte u Oceanía, la digitalización de registros médicos electrónicos y de padrones de beneficiarios de servicios sociales empieza a abrir la puerta al uso de algoritmos que toman decisiones automatizadas (o sea que la decisión final viene de la máquina) o semiautomatizadas (el programa hace una sugerencia, pero es un humano quien decide si la ejecuta o no).

Antes de la crisis del COVID-19, la tendencia a la digitalización y al desarrollo de sistemas cada vez más invasivos había propiciado una discusión importante en Estados Unidos y Europa sobre el comportamiento de dos actores capitales de los siglos XX y XXI: el Estado y las empresas.

Las empresas se han volcado con gran determinación y un desembolso significativo de recursos a introducir la IA en sus operaciones. Los reguladores, y el mismo sector privado, han sido conminados a prestar atención a la legalidad (o ilegalidad) de las campañas masivas de modificación de comportamiento,

robo de información o mal uso de datos por terceros. Empresas como Facebook se han visto envueltas en sonados escándalos de uso fraudulento —con su conocimiento o sin él— de los datos personales de decenas de millones de usuarios.<sup>20</sup>

Las compañías trabajan ávidamente para construir perfiles psicosociales de sus consumidores con el objetivo de venderles el producto más adecuado a su demanda, pero esa misma oferta está reduciendo sus opciones. Durante 2018 y 2019, los presidentes de las grandes empresas de tecnología se sentaron por primera vez ante el Congreso de Estados Unidos a discutir sobre privacidad, protección de información, su responsabilidad en la manipulación digital y sus políticas de gobernanza y ciberseguridad. Los roces entre autoridades públicas y compañías privadas continuarán a medida que siga avanzando la minería de información sobre los comportamientos de los individuos, una actividad que es y seguirá siendo incesante, pues en todas nuestras actuaciones digitales y electrónicas diarias producimos datos e información que son fácilmente monetizables.

Simultáneamente crecen los cuestionamientos y temores de la sociedad civil sobre la responsabilidad de los Estados frente a la protección de esos datos personales, y sobre los límites que deben tener las propias actividades de seguridad, vigilancia y espionaje gubernamentales. Los defensores de una intervención más decidida de los Estados, en ocasiones partidarios de formas más autoritarias de gestión de la cosa pública, han argumentado que el individuo debe sacrificar parte de sus libertades (como la privacidad) por el bien común, pues ello facilita la creación de mecanismos de seguridad pública que nos protejan a todos.

---

<sup>20</sup> El caso Cambridge Analytica, al que nos referiremos más extensamente en el capítulo 6, es uno de ellos.

En tal contexto, si bien implementar la IA en la gestión sanitaria es crucial y sus aprendizajes pueden ser extensivos a otras áreas de la política pública, esa acción debe ir acompañada de protocolos de gobernanza y protección de los datos personales, así como de procesos institucionales y regulaciones sobre su uso que se adapten a los contextos culturales e institucionales de cada país. La experiencia no digital ya nos enseñó que los derechos de los individuos pueden resultar afectados por políticas públicas agresivas, de manera que cualquier diseño que incluya IA debe considerar esos riesgos.

### La profundización de la brecha digital

La experiencia asiática es una fotografía futurista para América Latina. Muchos países de la región están aún lejos de implementar de forma masiva y efectiva sistemas de soporte de decisiones con motores de IA. ¿Motivo de preocupación? Por supuesto, aunque no somos los únicos. La disponibilidad de esta tecnología no es nueva (la IA llevaba al menos una década en experimentación en el sector privado para cuando la pandemia golpeó al mundo), pero la mayoría de los gobiernos ha mirado el fenómeno como se observa a un animal peligroso que pasa: llama la atención, asusta, pero si uno no se mueve es probable que siga de largo y no pase nada. Igual es mejor mantener la distancia.

Nada atípico y tampoco inusual. Los grandes fenómenos históricos suelen acumular presión a partir de hechos que parecen circunstanciales, hasta que amasan tal capacidad de implosión que estallan frente a los ojos de los incautos. Sucedió con la combinación pandemia-tecnología. Numerosos Estados carecían de capacidades tecnológicas de última generación, pues nadie imaginaba que algo así pudiera suceder; y quienes sí lo pensaron, como ocurrió en varios países, subestimaron la dimensión de la ola que los golpearía.

Un caso ilustrativo es el de España, una de las naciones más impactadas por el coronavirus como resultado de una reacción tardía a la crisis. Allí, uno de los problemas más importantes del sector salud durante los primeros meses de la pandemia fue la

falta de interoperabilidad de las bases de datos. Cada región de España tiene autonomía para organizar su sistema de salud y eso puede crear incompatibilidades complejas ante la necesidad de reunir toda la información en un comando único como el que implementó el gobierno español durante la crisis del COVID-19. Por diferencias metodológicas, ese comando tardó casi un mes para poder integrar una única base que pudiera servir para dar seguimiento a los contagios. Al quinto mes de la pandemia, los gobiernos autonómicos todavía estaban reportando registros de muertes no contabilizados que provenían de hospitales, funerarias del interior y residencias de personas mayores.

En América Latina, estas carencias son incluso más acentuadas. La desigualdad significativa que se registra entre las naciones —y entre regiones dentro de ellas— hace que coexistan ciudades grandes y medianas enlazadas por abundante conectividad con significativas porciones de la población sin acceso a internet y/o a servicios digitales. Según el Banco Mundial, en ALC hay 34 millones de personas que no cuentan con documentos de identificación básica, un elemento imprescindible para incorporar a los ciudadanos a las bases de datos del registro civil de un país. Pero aunque todos los latinoamericanos los tuvieran, estos documentos exigen que se disponga de la tecnología que facilite su uso digital, y las personas tienen que aprender a utilizarlos en los sistemas digitales actuales, algo que tampoco sucede.<sup>21</sup>

En tal sentido, una condición *sine qua non* para que la IA tenga posibilidades de desplegarse con alguna utilidad y eficiencia en la región es que se modernicen y amplíen los registros, se extiendan las infraestructuras tecnológicas y se ayude a la población a migrar hacia los servicios públicos digitales. Y estos no son asuntos fáciles de atender hoy. Por ejemplo, en América

<sup>21</sup> Esto no solo no ocurre en América Latina. En el mundo hay cerca de 1.000 millones de personas sin documentos oficiales de identidad, y de los 6.600 millones de individuos restantes que sí los poseen —en su mayoría habitantes de las naciones más desarrolladas—, solo la mitad puede usarlos en plataformas digitales confiables.

Latina el acceso a conectividad residencial sigue siendo escaso. En promedio, solo el 44% de los hogares tiene acceso a internet —es decir, un poco menos de la mitad del promedio de la OCDE (81%)—, además de que se registran diferencias importantes entre las subregiones: en el Cono Sur la conectividad residencial es del 54%, en Centroamérica y los países andinos del 34%, mientras que en el Caribe uno de cada cinco hogares la tiene.

Esta brecha digital hace que muchos de los sistemas y soluciones tecnológicas sean ineficaces o que solo puedan cubrir a aquella parte de la población que por lo general está representada por las familias acomodadas de todo el país y por las clases medias de los grandes conglomerados urbanos. Hoy, por ejemplo, sería inútil intentar el famoso “rastreo de contactos” en el contexto de la pandemia en las regiones rurales de América Latina y en muchas ciudades pequeñas del interior de los países. Esto por cuanto, según un estudio del Big Data Institute de Oxford University, para que este mecanismo de rastreo funcione en forma efectiva, es necesario que por lo menos un 60% de los habitantes de una nación tenga activa la aplicación para ello.

Es aquí donde el acceso limitado a todos esos datos y cifras constituye un problema de fondo para la política pública. Si estos sistemas van a ser parte de la definición del futuro del EBDI, entonces, a mayor desconexión mayor exclusión. Si a eso sumamos factores como el origen étnico, las aptitudes informáticas o el idioma —todos factores que ya creaban disparidades en el mundo analógico—, las barreras de entrada para las personas más rezagadas serán mucho más difíciles de sortear.<sup>22</sup>

<sup>22</sup> No sobra insistir en que las únicas personas que se benefician claramente de la revolución digital son las que ya tienen acceso a los servicios por la vía tradicional, es decir, las clases medias y altas.

La falta de acceso a medios digitales tiene también un efecto indirecto. Los sistemas de soporte de decisiones alimentados por la IA son entrenados con todos los datos a los que el software tiene acceso. En un sentido amplio, esto cubre desde el historial crediticio y el comportamiento laboral de un individuo hasta sus búsquedas en Google, sus trinos y sus comentarios en blogs añejos o en Instagram. La población que no tenga acceso a esas opciones de interrelación termina por estar subrepresentada en los algoritmos de la IA, dado que sus datos no son procesados por funcionarios, no son cargados en las bases de datos, no son leídos por motores, jamás son incluidos en los modelos probabilísticos y tampoco aparecen como sujetos en los análisis finales de las máquinas. En otras palabras, son **invisibilizados**. Nuevamente, la desigualdad no se mantiene sino que se acentúa.<sup>23</sup>

Sin embargo, la película está incompleta pues hay tendencias que permiten avizorar un futuro más alentador. Por ejemplo, la penetración de la telefonía inteligente llega al 67% de los latinoamericanos mayores de edad y se espera que alcance al 73% para 2025. Más del 80% de los adultos de 24 años o más tiene igualmente una cuenta en Facebook o de WhatsApp, de manera que existen mínimos entrenamientos oficiosos en el uso de tecnologías amigables.

### Apresúrate lentamente

Los gobiernos de América Latina deberán seguir apostándole a la transición digital, ya que la digitalización de los servicios públicos ha demostrado tener un impacto significativo tanto social como económico. Un Estado digital y abierto se vuelve trazable; permite a los ciudadanos y organizaciones dar seguimiento a las acciones de las instituciones e identificar anomalías. Asimismo, ofertas digitales ya existente como

<sup>23</sup> Si la persona de esa ecuación es un joven, pobre y sin formación, su falta de habilidades digitales lo pone en una situación de desventaja en cuanto al acceso a materiales educativos u oportunidades de empleo frente a otros jóvenes de su misma edad, pero cuya situación socioeconómica es más sólida.

los pagos y los servicios de salud electrónicos mejoran el acceso de los ciudadanos a la atención gubernamental. Todo ello moderniza las condiciones de partida para introducir la inteligencia artificial en las políticas públicas del Estado.

Pero incluso si las infraestructuras y las capacidades de uso de la población fueran ya suficientes, América Latina deberá prestar la debida atención a un factor que también preocupa al resto del mundo: la dimensión humanística del cambio tecnológico, un tema que debe ser entendido desde la propia dinámica y realidad culturales. Los países de la región tendrán que hacer un examen de conciencia sobre qué áreas exigen atención prioritaria del EBDI y deberán ser cuidadosos en la construcción de sus intervenciones sociales. El cómo y el porqué de cada proyecto son críticos. Nadie desea un Estado vigilante y autoritario, ni tampoco una administración pública incapaz de proteger la información privada de las personas de la avidez de actores con intereses políticos o económicos más egoístas que gregarios.

Como cuenta la leyenda que dijo el rey español Fernando VII a su ayudante, “vistámonos lentamente que tenemos prisa”. La introducción de la IA al nuevo Estado de bienestar digital inteligente no es una simple mejora administrativa, sino un cambio estructural en la manera en que los gobiernos analizan, atienden y resuelven los problemas sociales. La experiencia de la IA aplicada al control de la pandemia del COVID-19 se sostuvo en el acceso y procesamiento de volúmenes incalculables de información, para lo cual fue fundamental que se diera una conjunción de factores: leyes que facilitaran la transferencia de datos sin desproteger los derechos individuales; infraestructuras modernas de telecomunicaciones y procesamiento de información; y mecanismos de supervisión con capacidad real de aplicación de normas. Todos esos planetas —enormes, significativos— deben estar alineados para avanzar correctamente. Y tenemos que empezar a vestirnos ya, pues hay prisa.

Decía Heráclito que es imposible bajar dos veces por el mismo río, porque ese curso pasa solo una vez. La posibilidad de que América Latina se acople pronto y de manera segura a la nueva revolución que cambiará **exponencialmente** las relaciones entre Estados, economías e individuos y sociedades, es real y cercana. Que los habitantes de nuestros países queden incluidos en este cambio estructural dependerá del compromiso y determinación de los Estados, pues deberán ser ellos los que lideren y prescriban una transformación ordenada y segura.

Cada revolución industrial supuso un cambio en la relación entre los ciudadanos y los agentes públicos. La transformación que supone el EBDI expone dilemas cruciales a los que se enfrenta el actual Estado de bienestar y, con él, el contrato social que los ciudadanos tenemos con nuestras instituciones. En este caso, con la IA como herramienta de soporte del EBDI, el acceso temprano a esas relaciones en vías de transformación puede representar un salto determinante en las condiciones de desarrollo de los países y de la región. Por el contrario, la exclusión nos privará de participar de los beneficios de un mundo que resolverá más eficientemente los problemas y podrá enfocar sus esfuerzos en avances nuevos y más trascendentales. Los países sin acceso al nuevo diseño del mundo quedarán en una posición cada vez más marginal en la periferia: balsas abandonadas a su propio naufragio.

La transformación nos expone a decisiones complejas donde el error puede ser frecuente. La ansiedad por ofrecer soluciones instantáneas para aprovechar una tecnología omnipotente que aparece como un milagro tan único que da miedo cerrar los ojos por riesgo a que desaparezca, no puede llevarnos al apresuramiento de la improvisación.

Como habría dicho una buena versión de Hal 9000, la misión es demasiado importante como para dejarla en las manos equivocadas. Lo importante es asumir nuestra responsabilidad en la construcción de la IA y en la manera como debemos integrarla al Estado. Definirla desde un comienzo e involucrarse en el proceso facilitará la planificación del futuro de nuestras

sociedades de modo tal que evitemos producir más inequidad y dejar más gente rezagada.

El camino no es fácil, pero sí necesario. Y no parece razonable autoexcluirnos de una revolución tecnológica que cambiará radicalmente el mundo y creará un nuevo mapa. Lo vemos ahora. Las grandes transformaciones de la vida humana, las que definen nuestras relaciones como individuos en un grupo social, han estado marcadas por cambios tecnológicos profundos. No debemos olvidar que esta no es la primera revolución que nos ha tocado en suerte.



## CAPÍTULO 2

### Thomas Southwood Smith y el algoritmo aterrador

*Del pobre merecedor a la política predictiva.  
Cómo cambian las revoluciones tecnológicas al Estado,  
y a la sociedad y sus relaciones.*

¿Hasta dónde llegan las responsabilidades de un Estado? ¿Pago de cesantía, seguro de salud, educación gratuita y de calidad? La política social está alineada con las responsabilidades que asume el Estado para con la sociedad en un momento dado de la Historia, y está estrechamente vinculada a la idea de bienestar que reflejan las creencias y los valores de una nación en ese momento.<sup>24</sup>

El contrato social es un acuerdo más o menos formal entre los ciudadanos y su gobierno y, a partir de ese entendimiento común, sabemos qué esperar y cómo contribuir. Por su parte, el Estado asume responsabilidades y atribuciones según el compromiso acordado. En tal sentido, definir sus límites —económicos, sociales o culturales, por ejemplo— no es una tarea que se realiza de una vez y para siempre. Las funciones del Estado son dinámicas y se transforman con el espíritu de la época.

La política social es la dimensión estratégica —cuando se la diseña como tal— y la expresión práctica —cuando se la implementa— de esa idea de bienestar. Su objetivo es eliminar, o al menos reducir, los riesgos y factores que acentúan la pobreza y la exclusión en una sociedad. Esto quiere decir que la política social opera sobre las vulnerabilidades estructurales o crónicas que hacen que las personas se encuentren en una situación difícil, lleguen a ella o permanezcan allí.

Es evidente que los problemas de los individuos también cambian en función de diversos factores —cambios tecnológicos, una

<sup>24</sup> En la mancomunidad de Estados de bienestar de Europa, la política social se conoce por lo general como tal, *social policy*, e incluye servicios de salud, seguridad social y otras intervenciones contra la vulnerabilidad y la pobreza. Entre tanto en Estados Unidos, donde está más extendida la idea de que el individuo es responsable de valerse por sí mismo incluso en situaciones catastróficas, se separan los servicios públicos de salud y educación de las intervenciones de atención social como el seguro de desempleo, y eso se conoce como *social welfare*. Durante buena parte del siglo XX, en América Latina la política social estuvo alineada con la concepción europea y por lo tanto vinculada a alguna forma del Estado de bienestar.

crisis económica, la concentración elevada de la riqueza en una nación particular o una pandemia, entre otros— y se transforman según la estructura de una sociedad (más o menos equitativa, por ejemplo) y las políticas (más o menos inclusivas) del Estado.

Pero no nos confiemos. El Estado es una bestia acomodaticia que tiende a volverse conservadora con el tiempo, porque para mantener los procesos debe burocratizarlos. Por eso, cada tarea que le ha correspondido no ha sido por iniciativa propia sino en respuesta a algún reclamo más o menos masivo y no sin resistencia de su parte.

En este capítulo se hace un repaso de la transformación del Estado impulsada por las revoluciones tecnológicas y la presión social, y de la manera como ha evolucionado nuestro entendimiento de la vulnerabilidad y la pobreza, desde la concepción isabelina de los *deserving poor* hasta el surgimiento de la salud pública y la educación gratuitas. Cuando el ciclo histórico mueve su aplanadora de cambios, quienes están dispuestos a exigir un futuro distinto ponen manos a la obra.

### Una cloaca llamada Londres

Incluyamos a Thomas Southwood Smith en la lista de innovadores. No en vano lo que conocemos hoy como Estado de bienestar le debe algunos favores a este quisquilloso bautista rebelde que iba para pastor calvinista, pero que abandonó sus estudios religiosos, se graduó de médico y acabó como un brillante reformista sanitario.

Southwood Smith nació en Martock, Somerset, en 1788. En aquellos años, Inglaterra ya ejercía un amplio control de India, exploraba las costas de Australia y había iniciado el acelerado proceso de mecanización que la Historia conocería como la Primera Revolución Industrial. La Inglaterra del siglo XVIII había heredado las “leyes de pobres” isabelinas. Estas últimas establecían reglas para elegir a los beneficiarios potenciales de ayudas públicas que no eran sistemáticas, cuantiosas, y tampoco justas.

Las leyes isabelinas separaban a la población pobre en dos grupos definidos: los pobres merecedores o *deserving poor* (entre ellos los huérfanos, los adultos mayores y la personas en condición de discapacidad) y los pobres no merecedores (*undeserving poor*), es decir, personas necesitadas de las que se suponía tenían “mala moral” y por lo tanto no tenían derecho a ayuda, dado que eran responsables de su propia condición.

A los 15 años, Southwood Smith entró a la Academia Bautista de Bristol para convertirse en ministro pastoral, pero cinco años después, ya adulto, se rebeló contra el calvinismo, abandonó el claustro y se mudó a Edimburgo a estudiar medicina. Tras graduarse, entró en contacto con el círculo de Jeremy Bentham, el fundador de la Escuela Utilitarista.<sup>25</sup> Bentham daba la talla del inquieto Southwood Smith: era un hombre inteligente y sensible, y un radical en una época de grandes transformaciones.<sup>26</sup> La filosofía de Bentham abogaba por la separación de Iglesia y Estado; por la libertad individual, de expresión y económica; y por la abolición de la esclavitud, del castigo físico —en especial el de los niños— y de la pena de muerte, usual en la Inglaterra de fines del siglo XVIII.

Un tiempo después, Southwood Smith se mudaría de Edimburgo a Londres. Casado por segunda vez y con dos niñas a cuestas,

<sup>25</sup> A diferencia de otras congregaciones cristianas, los unitaristas descreen del valor sagrado de las escrituras y asumen que la razón debe ser la guía de su lectura. No aceptan la idea de la predestinación y ponen un peso significativo en las decisiones de los individuos para construir sus vidas. Por ello, cuando el toro de la Historia se les presenta, lo toman por los cuernos.

<sup>26</sup> En la lectura que Bentham hacía de la vida, su axioma central era que lo que define qué está bien y qué está mal es la felicidad de la mayoría de las personas. Por ello defendió la igualdad de derechos para las mujeres, la potestad de los individuos de divorciarse y la despenalización de la homosexualidad. Formado como jurista, filósofo y reformista social, Bentham ejerció una influencia profunda en Southwood Smith y en el grupo que lo acompañaba a menudo y del cual hacían parte el filósofo James Mill, su hijo John Stuart Mill (quien mejor desarrolló la herencia utilitarista de su maestro) y Robert Owen, a la postre uno de los fundadores del socialismo utópico.

se empleó como médico en el Jews Hospital y el London Fever Hospital, al tiempo que se volcaba a su congregación unitaria como pastor.

La Londres que Southwood Smith encontró era una metrópoli descontrolada y plagada de contrastes, como consecuencia de los cambios brutales de la Primera Revolución Industrial. Los siglos XVIII y XIX supusieron el paso de sociedades mayoritariamente agrarias a una nueva articulación —la ciudad— como resultado de la introducción de la máquina de vapor y otros avances tecnológicos que permitieron el surgimiento de los ferrocarriles, las fábricas de producción masiva y barcos de larga autonomía.

La fuerza mecánica desterró a la fuerza muscular del mercado de trabajo. El campo dejó de ser un territorio con futuro para convertirse en espacio de supervivencia, y comenzó a expulsar migrantes que constituyeron la avanzada de una rápida urbanización de los antiguos asentamientos humanos. En los tiempos en que Southwood Smith crecía en Somerset se produjo una proliferación de aldeas y pueblos que en pocos años quedaron saturados de agricultores que buscaban otra vida en los nuevos burgos.<sup>27</sup>

Esa Primera Revolución Industrial fue una transformación económica desmesurada que también agravó numerosos problemas sociales, pues los pueblos que originalmente apenas si podían albergar a una población reducida, terminaban colapsados con la creciente masa migratoria. La infraestructura disponible resultaba insuficiente para absorber en condiciones aceptables a la marea humana, que no hacía otra cosa que agravar el hacinamiento preexistente.

<sup>27</sup> El área del puerto de Liverpool, por ejemplo, pasó de tener una población de 78.000 personas en 1801 a más de un millón de habitantes en algo menos de cien años.

La industrialización degradó el estado de las cuencas, manantiales, arroyos y ríos locales de donde se abastecían de agua de beber los habitantes de poblados y ciudades. Ahora también iban a parar allí los metales pesados de las fábricas y un mayor volumen de excrementos producidos por las incontrolables manchas urbanas. Las bacterias de enfermedades infecciosas letales como la tifoidea y el cólera campeaban en esas fuentes de agua. El historiador Alexis de Tocqueville describía así a la Manchester de aquellos años: “Las aguas fétidas y fangosas de los arroyos, manchadas con mil colores por las fábricas a su paso (...) deambulan lentamente por este refugio de pobreza.”

Dado que los barrios de las ciudades se organizaban por clases, los que salían peor librados en ese universo de malas condiciones eran los más pobres. En las urbanizaciones de obreros predominaban las casas continuas —las llamadas *rowhouses*— muy mal ventiladas. En las barriadas, la higiene y el manejo de residuos —a menudo vertidos en las calles— eran pésimos. Fue así como en muchas ciudades esos barrios se convirtieron en focos de enfermedades y depredación.

Southwood Smith había comenzado a interesarse en las epidemias un tiempo antes de la Revolución, y para 1830 ya había publicado su *Treatise on Fever*. Allí desafiaba a la comunidad médica a que pensara de manera distinta el tratamiento de las enfermedades sociales. El médico unitarista probó que las condiciones sanitarias eran la causa principal de la propagación de enfermedades como el cólera y la fiebre amarilla, y que estas se podían evitar con inversiones en saneamiento. “La causa inmediata o que excita la fiebre es un veneno formado por la corrupción o la descomposición de la materia orgánica,” escribió. “Durante el proceso de putrefacción, la materia vegetal y animal emite un principio o da origen a un nuevo compuesto que, cuando se introduce en el cuerpo humano, produce los fenómenos que constituyen la fiebre.”

Ocho años después, en 1838, la Comisión de la Ley de Pobres del Parlamento británico solicitó una serie de informes porque

observaba que una alta proporción de las personas de menores recursos, que constituían la inmensa mayoría de la población de ciudades como Londres, estaban enfermas.<sup>28</sup>

Muchos hombres no podían trabajar porque arrastraban dolencias de salud; un número elevado moría cada año. Aunque eran los primeros atisbos de una política pública dirigida a atender a quienes anteriormente se consideraban “pobres no merecedores,” el interés para el Estado de entonces era exclusivamente económico; la productividad iba en picada y las arcas extenuadas debían hacerse cargo de sostener a las viudas e hijos de esos muertos. A numerosos políticos les disgustaba la idea.<sup>29</sup>

Fue por ello que la Comisión pidió a los médicos James P. Kay-Shuttleworth y Neil Arnott, y al propio Southwood Smith, que investigaran la situación sanitaria de varios distritos de Londres. Southwood Smith estuvo a cargo del célebre estudio sobre Bethnal Green, el área con el mayor número de telares de la ciudad. Evidentemente no tardó mucho en hallar un vínculo directo entre saneamiento y enfermedad. Los retretes desaguaban directamente en las calles, a cielo abierto, y el efluvio desagradable cubría toda la zona. Southwood

<sup>28</sup> Después de que en 1534 el Parlamento lo convirtiera en jefe supremo de la Iglesia en Inglaterra —decisión que significó la ruptura de la nación con la autoridad papal—, Enrique VIII impulsó dos leyes de supresión de monasterios que entre 1536 y 1541 confirieron a la monarquía inglesa la autoridad de ejecutar procesos administrativos y legales dirigidos a eliminar monasterios, prioratos y conventos en Inglaterra, Gales e Irlanda.

<sup>29</sup> “Hasta 1834, Inglaterra operaba un sistema de bienestar social de amplia cobertura que garantizaba la subsistencia de todos, principalmente mediante transferencias de dinero en efectivo a ancianos, viudas, familias trabajadoras numerosas, enfermos, discapacitados, y también a los desempleados”, según Gregory Clark y Marianne Page en *Welfare Reform, 1834*. “Es probable que hasta un 9% de la población recibiera ese alivio en cualquier año entre 1790 y 1834. Smith, Bentham, Malthus y Ricardo denunciaron que este sistema impedía el buen funcionamiento de los mercados de trabajo, reducía los incentivos de los pobres para trabajar y ahorrar, y desalentaba la inversión en áreas de alto desempleo al gravar a los empleadores.”

Smith encontró casas con familias enteras agobiadas por las enfermedades y zonas donde las dolencias eran generalizadas. Identificó a ese lugar como una “fuente fructífera de fiebre”.

Tras el informe de Bethnal Green, la Comisión solicitó al gobierno que gastara dinero en saneamiento. La medida permitiría reducir el costo de cuidar a los pobres y a las familias de los muertos en el largo plazo, además de que seguía el espíritu del informe de Southwood Smith. Aquella era una solución utilitarista y práctica motivada por razones científicas y, por supuesto, nada casual. En la Londres de entonces, pequeña y conservadora, todo el mundo estaba al tanto de la posición heterodoxa de Southwood Smith frente a la profesión médica tradicional.

Bajo la influencia de Bentham y de un círculo creciente de su parroquia,<sup>30</sup> Southwood Smith llevaba un tiempo enfrentado a las que llamaba “corporaciones de médicos, cirujanos y boticarios administradas oligárquicamente.” En uno de sus escritos públicos decía que las enfermedades no eran aristocráticas o plebeyas y que los mismos tratamientos debían aplicarse a ambos grupos sociales.

La crisis y el informe de Southwood Smith llevaron a que se hiciera un primer intento por formular un plan de salud pública en Gran Bretaña. De hecho, dieron pie para que el debate se iniciara de inmediato. Pronto la Iglesia reclamó atención para los pobres. En 1842, otro informe realizado con médicos por el reformista social Edwin Chadwick permitió discutir la primera red de agua potable. En el informe de Chadwick, titulado *La condición sanitaria de la población trabajadora* (en buena medida heredero del trabajo de Southwood Smith), se argumentaba que las viviendas de los barrios marginales, el alcantarillado ineficiente y el abastecimiento de agua impura en las ciudades industriales estaban causando la muerte

<sup>30</sup> Además del matrimonio Harriet Taylor y John Stuart Mill, eran unitaristas universales el sanitarista Edwin Chadwick, la feminista Harriet Taylor, y la música y compositora Eliza Flower y su esposo el poeta Robert Browning.

innecesaria de unas 60.000 personas y enviando a más de 110.000 niños al año a los orfanatos. De allí que se instara al Estado a que gastara en drenaje público, suministro residencial de agua y en la eliminación de todos los desechos. Esos gastos, decía Chadwick con su lógica utilitarista, serían “una ganancia financiera (...), ya que reducirían el elenco de enfermedades y la muerte prematura.”

Southwood Smith y Chadwick defendieron con insistencia la necesidad de que el Estado invirtiera en política social. Ambos notaban tanto los beneficios como los perjuicios generados por la Primera Revolución Industrial que otros ignoraban o negaban.

En su informe sobre las condiciones sanitarias de la población trabajadora de Gran Bretaña, Chadwick profundizó en el estudio del impacto de la pobreza en la salud. Afirmaba que las personas de clase media vivían vidas más largas y saludables porque podían pagar para que se eliminaran las aguas residuales y se canalizara el agua dulce hacia sus hogares, y que este fenómeno dependía también de los contextos locales. En aquellos años, en Liverpool un profesional vivía en promedio 35 años; entre tanto, un trabajador en la misma ciudad moría en promedio a los 15 años, afectado por la contaminación originada en una industria no regulada y por el hacinamiento que facilitaba la propagación de enfermedades.

Chadwick criticaba a las empresas privadas que eliminaban aguas residuales y suministraban agua dulce, argumentando que eran los organismos públicos los que debían estar cargo de esos estos servicios. Según él, tales compañías solo estaban dispuestas a prestárselos a quienes pudieran pagar por ellos —una minoría—, cuando el fin de las entidades públicas debía ser asegurarse de que todos obtuvieran servicios. Cinco años más tarde, en 1847, con base en informes similares a los de Southwood Smith y Chadwick sobre casi todas las grandes ciudades, el gobierno de Inglaterra introdujo la primera gran reforma de salud y las administraciones locales instrumentaron sus primeras políticas sociales, de saneamiento, desagües y alcantarillado.

### Von Bismarck patea el tablero

La historia de Thomas Southwood Smith es paradigmática. Cada vez que aparece una tecnología disruptiva —como la máquina de vapor en la Primera Revolución Industrial—, toda la estructura económica y política se conmociona, surgen presiones sociales y sectoriales, se producen innumerables transformaciones culturales y hasta el propio contrato social sufre modificaciones.

Los cambios producidos a raíz de la introducción de tecnología en la Primera Revolución del siglo XVIII socavaron los mecanismos familiares o comunitarios de resiliencia, así como la asistencia social caritativa de la Iglesia en las sociedades preindustriales.<sup>31</sup> En la Inglaterra de Southwood Smith se crearon entonces las condiciones sociales y económicas requeridas para que el Estado iniciara su primera planificación de políticas sociales generalistas —a menudo **reactivas**— que permitirían superar la concepción caritativa y focalizada propia del periodo preindustrial.

Cuando Thomas Southwood Smith escribió su informe sobre los barrios de Londres se produjo un cambio de concepción. Hasta entonces, los gobiernos creían que la intervención estatal en la economía desbalancearía el progreso, por lo que solo porciones menores de la población recibían algún tipo de asistencia. Fue esa limitada lectura de los partidarios del libre mercado de la Primera Revolución Industrial lo que condujo al

<sup>31</sup> En esta primera fase, los intentos “institucionales” de organizar respuestas colectivas provienen de las entidades de beneficencia nacidas a partir del siglo XI en Europa. La mayor parte del apoyo provenía de instituciones religiosas como monasterios o fondos. A partir del siglo XV, la élite urbana comenzaría a financiar algunos servicios de salud con la aparición de hospitales de acceso más o menos libre, donde se recibía y se alimentaba a personas en problemas. Tiempo después se crearían figuras como las cajas de ahorro o las cajas de empréstito, los bienes comunes serían legalmente formalizados, y las nuevas prácticas y organizaciones agrícolas comunales ofrecerían a las personas en situación de pobreza acceso a la tierra y a sus frutos.

surgimiento de movimientos sociales que inclinarían al Estado a tener un papel más activo frente a la vulnerabilidad social.

La siguiente transformación del Estado ocurrió con la Segunda Revolución Industrial, impulsada esta vez por la electricidad y las cadenas de montaje. El desarrollo fabril más avanzado redujo todavía más la producción rural artesanal y afianzó la urbanización masiva de personas con pocas capacidades técnicas que sufrían un desempleo cíclico.

La Segunda Revolución no nació de un chasquido de dedos. A fines del siglo XIX, la industrialización y la urbanización habían introducido nuevos riesgos para los Estados. Si la Primera Revolución había creado la ciudad moderna arrojando a cientos de miles de personas a barrios insalubres, una mecanización más eficiente seguía atrayendo gente que ahora tenía frente a sí una primera generación de habitantes envejecidos con enfermedades que se sumaban a las de los nuevos trabajadores. Así, entre 1870 y la Primera Guerra Mundial, los gobiernos debieron vérselas con exigencias crecientes: pensiones y atención sanitaria por más tiempo y para un mayor número de personas, y atención a accidentes ocupacionales y a enfermedades laborales antes inexistentes.

La respuesta más decisiva llegó de Alemania, impulsada por grupos socialistas. El país había entrado en recesión hacia 1870, tras el rápido crecimiento producido por la industrialización acelerada iniciada en 1850. Las largas horas de trabajo arduo comenzaron a hacer mella en la población, la crisis generó pobreza y la emigración hacia otros países como Estados Unidos se hizo cada vez más común. Los grupos sociales que empezaron a exigir protecciones laborales eran cada vez más numerosos, y el partido laborista socialdemócrata alemán comenzó a instar a los trabajadores a que se sindicalizaran. Eso poco le gustó a Otto von Bismarck, el primer canciller de la historia alemana moderna, quien veía en el socialismo un riesgo para la nación.

Von Bismarck, que era muy listo, actuó políticamente. Se apropió de las reivindicaciones más o menos informales de la izquierda para instrumentarlas con prontitud y las presentó, no como una reacción sino como una oferta proactiva del Estado. En concreto, y a pesar de ser un conservador férreo, decidió abanderar algunas de las ideas de los grupos socialistas para reducir la presión e introdujo, en 1883, una ley de salud que dio origen al primer sistema nacional contributivo<sup>32</sup> de pensiones donde tanto los empleadores como los empleados pagaban a un fondo de seguros.<sup>33</sup> Durante los años siguientes, esta ley continuó ampliándose. Los seguros de accidentes (1884) y las pensiones para la vejez (1889) también acabaron formando los cimientos del moderno Estado de bienestar alemán, y luego europeo.

Bismarck fue un revolucionario imprevisto. Buscó ganar el apoyo de las clases trabajadoras, desarmar a los grupos socialistas, evitar la creación de sindicatos, y en el camino creó un marco que se mantuvo durante una centuria, ya que a lo largo del siglo XX los países industrializados ampliaron las disposiciones públicas para enfrentar riesgos como los accidentes laborales, la enfermedad, la discapacidad, la muerte y/o el desempleo crónico.

Al final, fueron nuevamente los cambios tecnológicos —la nueva fábrica con cadena de montaje y la electricidad— los que transformaron la dinámica social, la relación con el Estado,

<sup>32</sup> El financiamiento del Estado de bienestar se da tradicionalmente por dos vías: la contribución del patrón y del empleado, y los impuestos recolectados por el fisco. Aunque muchos Estados europeos en ese momento siguieron el modelo alemán, Gran Bretaña, por ejemplo, financió sus sistemas de salud a través de impuestos.

<sup>33</sup> “El plan de seguros de Bismarck no era una idea totalmente original”, según Lorraine Boissoneault en *Smithsonian Magazine*. “Los gobiernos europeos habían implementado medidas de salud pública desde el siglo XIV, cuando las ciudades-estado italianas tomaron medidas para controlar la propagación de la peste bubónica mediante cuarentenas. Y los grupos de seguros de salud organizados por la comunidad —llamados mutualidades o fondos para enfermos— aparecieron más o menos al mismo tiempo en ciertas profesiones”.

el contrato y las políticas sociales. Tras la Segunda Revolución Industrial, el mundo vio nacer los primeros seguros sociales, sindicatos, leyes más estrictas sobre trabajo infantil y salario mínimo, y el esbozo de lo que sería el Estado de bienestar, cuya expansión continuaría durante casi todo el siglo XX.

Es en esos años cuando en la política social<sup>34</sup> comienza a aparecer la idea de que el Estado debe cumplir un papel **preventivo** que supere su función reactiva inicial. Ahora las personas saben que hay redes de protección si caen en la pobreza, que hay seguros que les permitirán jubilarse y tener dinero para subsistir en la vejez, que los hospitales operan con un porcentaje de sus impuestos y que las madres comienzan a tener beneficios pre y postnatales. Ese primer Estado de bienestar exhibe también el germen de la separación entre los servicios públicos y las políticas de protección social.

### La Guerra, los cinco gigantes y las diferencias filosóficas

Además de haber sido un conflicto global aniquilante, la Segunda Guerra Mundial dio origen a una revolución en materia tecnológica. Los inventos nacidos entre 1939 y 1945 tendrían una influencia significativa en las décadas venideras. “No obstante el papel crucial que desempeñaron la ciencia, las matemáticas y los nuevos inventos en guerras anteriores, nada tuvo un efecto tan profundo en las tecnologías de nuestra vida actual como la Segunda Guerra Mundial,” según el experto en tecnología David Mindell. “Y ninguna guerra resultó tan

<sup>34</sup> El objetivo de esas políticas sociales es eliminar, o cuando menos reducir, los riesgos y factores que agudizan la pobreza y la exclusión en una sociedad. En tal sentido, operan sobre las fragilidades estructurales o crónicas que producen, y llevan a las personas a una situación vulnerable o la mantienen en ella. Por supuesto que esos problemas sociales cambian en función de diversos factores como son, por ejemplo, una crisis económica o la elevada concentración de la riqueza en una nación, o una catástrofe natural o una degradación ambiental progresiva, y se transforman también según la estructura de una sociedad (más o menos inclusiva, por ejemplo) y las políticas del Estado (también más o menos inclusivas).

profundamente afectada por la ciencia, las matemáticas y la tecnología como esta.”

Desde los fabricantes de autos hasta los de refrigeradoras, las industrias livianas fueron reconvertidas durante la guerra en productoras de armas, pertrechos y material bélico. El radar, que requería piezas de precisión pequeñas y confiables, sentó las bases de la ingeniería que desarrollaría una megaindustria en la década de los años cincuenta: la televisión. La bomba atómica reconfiguraría la producción de armas del futuro. El pesticida DDT, usado para combatir los mosquitos en las batallas del Pacífico, transformaría la agricultura. La industria alimenticia llevó a cabo numerosos estudios sobre nutrición —crítica en una guerra— que acabaron incidiendo en toda una nueva gama de alimentos envasados y empacados una vez que se produjo el regreso a la civilidad. Los plásticos —invento mayor— llegarían en abundancia. ¿Qué otras cosas cambiaron? Los ingenieros alemanes de la guerra quedaron sorprendidos con la velocidad con que Estados Unidos reconvirtió su producción civil para servir al aparato militar, pero sobre todo con cuanta efectividad consiguieron entrenar a la población para cambiar completamente de actividad según la demanda.

Por otra parte, el Ministerio de Sanidad del Reino Unido encargó al economista y reformista social, William Beveridge, una evaluación sobre el estado de las prestaciones sociales en el país en ese momento. A ello debía asociar recomendaciones de cambio, anticipando la reconstrucción necesaria al concluir el conflicto. Beveridge se refirió a los principales desafíos de política a los que se enfrentaría el Reino Unido como “los cinco gigantes”: miseria, ignorancia, necesidad, ociosidad y enfermedad. Para Beveridge, era necesaria una participación del Estado más activa que nunca y por ello abogó para que las autoridades dieran garantías de que se dispondría de los recursos necesarios para asegurar el bienestar de las personas. Un seguro social y las pensiones serían solo una porción de la oferta pública, que debía incluir también educación de calidad, atención sanitaria razonable y servicios de vivienda asequibles. En 1945, el Partido Laborista derrotó a Winston Churchill en

las elecciones generales y el nuevo primer ministro, Clement Attlee, decidió implementar las sugerencias de Beveridge.

En el aire se respiraba el nuevo signo de los tiempos: una humanidad que empezaba a creer que era capaz de superar grandes males y seguir adelante. Tres años después del fin de la guerra, las Naciones Unidas adoptarían la Declaración Universal de los Derechos Humanos de 1948, en la cual se expandían las sugerencias de Beveridge en Inglaterra al campo de los derechos económicos, sociales y culturales. La Declaración Universal de los Derechos Humanos reconoce el derecho a la seguridad social en su artículo 22, el derecho al trabajo en el 23, el derecho al descanso y al esparcimiento en el 24, el derecho a un nivel de vida adecuado en el 25, el derecho a la educación en el 26, y finalmente el derecho a los beneficios de la ciencia y la cultura en su artículo 27.

Estos son los comienzos del papel del Estado como **promotor**, cuyo objetivo iba más allá de la responsabilidad laboral pues asumía una obligación con los ciudadanos de garantizarles el ejercicio de sus derechos humanos, económicos, sociales y culturales como inalienables. El triunfo de las democracias de Occidente transmitió la visión de los ganadores a la organización política, cultural y económica del mundo.

En los años sesenta, los beneficios del rebote económico madurado en la década precedente permitieron a los gobiernos reestructurar y solidificar sus Estados de bienestar. Las tasas bajas de desempleo que perduraron hasta la década siguiente permitieron discutir la función del Estado más allá de formular medidas para atenuar los riesgos laborales, como los seguros de desempleo y accidentes, y las pensiones. Los años sesenta marcaron, pues, el salto de la Tercera Revolución Industrial al centro del escenario.

También llamada Revolución Informática o Digital, la tercera gran transformación fue propulsada por el desarrollo de semiconductores y microprocesadores, así como por el átomo y el nuevo potencial de la energía nuclear. Los semiconductores

permitieron miniaturizar objetos y dotar a los procesos de mayor velocidad y eficiencia. Gracias a los robots y a los controles programables, la automatización de la producción en serie de alto nivel revolucionó nuevamente a la industria.

El lanzamiento de la informática central (en la década de los años sesenta) y de la informática personal (en los años setenta y ochenta) catapultó saltos exponenciales en la producción y modificó otra vez las relaciones sociales. La aplicación, primero de los transistores entre los años cincuenta y sesenta, y posteriormente de los circuitos integrados en los años previos a 1970 y de los microprocesadores a partir de 1971, abrió las puertas a un mundo desconocido. Cada generación de computadoras duplicó la capacidad de sus circuitos lógicos e incrementó sus desempeños en casi un 50%, imprimiendo un ritmo veloz a todo el abanico de la actividad económica en oleadas constantes de cambio productivo. Las computadoras aceleraron todos los procesos, desde los de planificación y gestión de las burocracias públicas, hasta los avances de la ciencia y las actividades del campo, pasando por el aprendizaje en las escuelas e incluso por la economía aplicada.

El desarrollo telemático se hizo cada vez más presente en los años setenta con mejoras continuas en la capacidad de procesamiento. Gracias a las tecnologías de la información y las comunicaciones (TIC), el comercio internacional se aceleró, los procesos productivos adquirieron eficiencia y se estandarizaron a escala planetaria, y el capital financiero se globalizó. Desde entonces, el tiempo transcurrido entre los cambios se ha acortado con cada nueva innovación tecnológica, sea esta central o marginal.

Las corporaciones ampliaron sus economías de escala y las naciones debieron esforzarse por ser más productivas con una mano de obra cuyos costos fueran competitivos. El peso específico de los grandes capitales, deslocalizados y capaces de imponer condiciones de manera progresiva, ejerció cada vez más presión en los Estados.

Impulsada por la Inglaterra de Margaret Thatcher y los Estados Unidos de Ronald Reagan, la corriente neoliberal penetró en numerosos países. A partir de los años ochenta, la globalización del comercio empezó a exigir mercados abiertos en todo el mundo, requiriendo de los trabajadores nuevas destrezas y dejando en los márgenes del sistema a quienes no tenían habilidades y conocimientos específicos. El desempleo creció con las recesiones de las décadas de los años setenta y ochenta, pero no todos los gobiernos respondieron ampliando coberturas sociales. En el caso de Estados Unidos e Inglaterra, por el contrario, los esfuerzos se orientaron a adelgazar el Estado. Reagan, por ejemplo, hizo reiterados intentos por reducir tres programas sociales —vivienda, pensiones para la vejez y subsidios para los ingresos más bajos— y dio pasos menores para recortar los fondos asignados a los programas de salud, enfermedad y discapacidad.

### Más vale predecir que lamentar

Si hasta los años setenta se había respondido de manera más o menos homogénea a los desafíos del cambio, la liberalización económica progresiva marcó el inicio de la separación de las escuelas de pensamiento sobre política pública. La década siguiente presenció el gran quiebre conceptual acerca del Estado de bienestar. Las divergencias que generó el neoliberalismo se profundizaron y persisten hasta nuestros días. Esto diferenció aún más las modalidades de protección social que las naciones desarrollaban desde la Segunda Guerra Mundial, cada una inspirada en filosofías más o menos disímiles en su concepción del papel del Estado y de los individuos.

El sociólogo danés Gøsta Esping-Andersen, un especialista en Estado de bienestar, clasifica de la siguiente manera a los sistemas resultantes tras la guerra: (1) el modelo liberal que practican Estados Unidos y el Reino Unido; (2) el de la Europa conservadora continental, vigente en Alemania, Francia y Bélgica; y (3) el modelo socialdemócrata escandinavo propio de Suecia y Dinamarca.

Sobre todo a partir de la Tercera Revolución Industrial, el Estado social que primaba en la Europa continental desarrolló políticas en las áreas de salud, educación, seguridad social, distribución del ingreso y vivienda, al tiempo que buscaba soluciones para los problemas ambientales y urbanos. En Estados Unidos, en cambio, varias de esas políticas no fueron asumidas por el Estado, que se mantuvo como proveedor de segundo orden respecto a la oferta privada o a la responsabilidad de los individuos. Si bien allí se ofrecen sistemas de retiro y salud para los jubilados, a lo largo de la vida laboral el grueso de la atención recae en las familias, que deben hacer frente a los exagerados y crecientes costos de la atención médica nacional.

América Latina heredó elementos fundamentales del modelo de la Europa continental, con una orientación del financiamiento mediante contribuciones patronales e individuales inspirada en Bismarck. Esto se produce en el marco de un consenso en torno a la concepción de que la política social es parte fundamental de las obligaciones del Estado, el cual debe promover el empleo y el desarrollo, facilitar la resolución de conflictos, crear resiliencia y reducir la desigualdad.<sup>35</sup>

¿Qué quiere decir esto? Que para gustos, colores. Si bien es cierto que los trabajos de Thomas Southwood Smith y Chadwick sentaron las bases de la función social del Estado, sus responsabilidades y atribuciones no quedaron grabadas en piedra. Esos límites son dinámicos, como lo han demostrado las grandes revoluciones tecnológicas. Por eso la acción de individuos como Southwood Smith, Bismarck o Beveridge —heterodoxa y fuera de toda zona de confort— permite repensar modelos anquilosados.

<sup>35</sup> La propia CEPAL identifica seis áreas básicas de políticas públicas de gasto social: protección social, educación, salud, vivienda y servicios comunitarios, actividades recreativas (incluidas cultura y religión) y protección del medio ambiente.

Al cabo, la política social nunca es aséptica, como tampoco lo es la tecnología, y ni la una ni la otra se pueden entender sin leer las condiciones y contextos de las transformaciones, es decir, sin comprender la dinámica histórica.

En esa dinámica, como ya lo advertimos, los Estados fueron primero **reactivos** (a través de la asistencia social y los servicios), luego **preventivos** (reduciendo riesgos través de instrumentos de seguridad social), después **promotores** (promulgando leyes que favorecieran la igualdad, la inclusión y el enfoque de derechos, el desarrollo de habilidades, y el fomento y la inversión en actividades productivas). Ahora ha llegado el momento de que los Estados se doten de la flexibilidad necesaria para ejecutar políticas útiles encaminadas a resolver tanto circunstancias generales como fenómenos particulares en tiempos de cambios acelerados. Y no hay mejor manera de hacerlo que anticipando tales cambios. El Estado de bienestar digital inteligente —el EBDI— encaja en este nuevo escenario como anillo al dedo en su nueva personificación: el Estado **predictivo**.

La política predictiva llevará a que los gobiernos establezcan una lista de escenarios probables para actuar de forma preventiva antes de que el fenómeno se manifieste y genere una crisis. Las tecnologías son variadas —reconocimiento de imágenes, voz, aprendizaje automático, procesamiento de lenguaje natural— y permiten analizar trillones de datos históricos para hallar patrones y tendencias que faciliten anticipar un problema. Si algo aprendimos del coronavirus es que invertir en evitar una crisis y actuar de forma temprana es siempre menos costoso que gastar en revertir sus efectos.

### El COVID-19 y otras pandemias impulsoras de revoluciones

La Cuarta Revolución Industrial (4IR por su denominación simplificada en inglés), iniciada en los años noventa con la creación de internet, produjo disrupciones nuevas y significativas. La Primera Revolución Industrial del siglo XVIII usó vapor y agua para mover las máquinas y dijo adiós a los

animales como fuerza de trabajo y a la biomasa como fuente de energía; la Segunda, a fines del siglo XIX e inicios del siglo XX, empleó energía eléctrica para fabricar en cadena; la Tercera se valió de la electrónica y de las tecnologías de la información para automatizar la producción; también son los años del auge de la energía nuclear y el inicio de la exploración de fuentes sostenibles como la eólica y la solar.

La Cuarta, hija de la revolución digital precedente y catapultada por internet, amenaza con estallar en mil pedazos el mundo que conocimos. A diferencia de las tres revoluciones industriales anteriores, su propagación es viral y exponencial. De allí que produzca incertidumbre, miedo y resistencia. Un monstruo.

La 4IR difumina el mundo real y el tecnológico con avances que fusionan el mundo físico con el digital y el biológico. Hay quienes la describen como la llegada de los “sistemas ciberfísicos,” que exigen capacidades completamente nuevas en las personas y en las máquinas. Hay quienes, por pánico o simplificación, ven en ella a Skynet produciendo la primera generación de Terminator en las mismas plantas donde antes una cadena de producción escupía un Ford Focus. Para las mayorías todo es imaginable, porque lo visible es escaso y lo poco que se ve no se comprende muy bien. En realidad no es así.

Con los avances recientes en materia de inteligencia artificial, podemos imaginar operaciones autónomas donde fábricas de aviones, laboratorios o sistemas de alcantarillado pueden funcionar sin un ser humano a cargo. Hay cambios ya visibles que refuerzan la creencia de que vivimos en un mundo donde usamos algo que no comprendemos del todo. Facebook, Twitter, WhatsApp e Instagram han convertido a cada individuo en un emisor. Han eliminado la intermediación de la voz pública, poniendo aún más presión en la crisis de representatividad de la democracia: ¿para qué ser miembro de un partido político cuando puedo organizar una revolución con mis amigos por mensaje instantáneo para mañana a las 10:00 am?

La influencia de la 4IR en la economía no es apta para cardíacos. No hay variable macroeconómica que se libre de quedar atravesada por la 4IR y sus algoritmos. Los escenarios de los modelos de proyección de crecimiento, sin ir más lejos, se ensayan cada vez más en plataformas artificiales de IA.

Pero hay efectos más concretos.

Algunos economistas sugieren que la nueva revolución podría profundizar aún más las brechas de desigualdad laboral.<sup>36</sup> Ya no se trata de una sustitución de trabajos mecánicos y repetitivos, sino el reemplazo de procesos que requieren percepción y discernimiento. ¿Hablábamos de Skynet y Terminator? Pues aquí hay sustancia para otras cuantas películas de terror. Como lo advertimos, con la 4IR las esferas física, digital y biológica a veces se funden como nunca antes, introduciendo dilemas filosóficos y éticos. De hecho, este es el punto crucial. Hoy en día deberíamos tenerle más temor a esos sistemas automatizados que deciden quién puede ser beneficiario de un programa social, quién accede a la escuela o quién se queda con un trabajo sin que un ser humano lo supervise.

En la actualidad la tecnología es importante, y en ocasiones determinante, pero no se la puede ver como a una diosa del Olimpo. Es y debe ser entendida como la **herramienta** que es. Y para asumirla como tal, como producto y proceso utilitarios, tenemos que construir una visión común sobre qué tecnología queremos, cómo la produciremos, por qué razones y con qué objetivos. Esa lectura compartida facilitará la apropiación crítica de las herramientas tecnológicas, y nos conferirá potestad y control sobre nuestras creaciones.

<sup>36</sup> En el documento titulado “Impacto de las innovaciones tecnológicas en la estructura social”, publicado en el *International Journal of Economics and Financial Issues*, se describen las consecuencias negativas de las innovaciones que podrían conducir al reemplazo de mano de obra humana por sistemas automatizados.

Las diferencias principales entre las revoluciones anteriores y la 4IR son el ritmo del cambio y el alcance de la innovación tecnológica. Por escala y complejidad, no ha habido otra transformación de envergadura semejante en la Historia. Jamás.

Los cambios afectan a casi todas las industrias en todo el mundo, y por ello será necesario reformular sistemas completos de producción. Modifican la manera en que trabajamos —y hasta la posibilidad de tener un empleo—, nuestro tiempo personal, la forma en que nos comunicamos unos con otros. Cambian los gobiernos, la educación, la atención médica, el comercio. Todo. La gestión pública entrará en un juego cuyas reglas diseñadas para hace una década serán letra muerta. Las instituciones estatales, la sociedad, las empresas, los académicos, las familias y los individuos deberán acomodarse a un nuevo mundo. Estamos hablando de un cambio literalmente estructural.

### De la enfermedad puede salir la medicina

América Latina y el Caribe vienen de atravesar una de las épocas más prósperas de su historia. Entre 2003 y 2013, la desigualdad de ingresos se redujo de manera significativa, y la pobreza se contrajo masivamente como producto de la combinación de una mejor distribución del ingreso y un crecimiento económico razonable alimentado por los excelentes precios de los productos básicos de exportación agrícolas y minerales. Y aun así, 167 millones de personas vivían en la pobreza en la región en ese mismo período, especialmente en las áreas rurales, con acceso insuficiente y de baja calidad a servicios básicos de salud, agua, saneamiento e higiene, y sufriendo un legado de discriminación indiscutible basado en el género, la raza y la clase social.

Somos la región más desigual del mundo.

A comienzos del año 2020, justo antes de que se comenzara a sentir el impacto del coronavirus, otras amenazas merodeaban en el ambiente. En su estudio “¿Desarrollo económico inestable? Choques agregados en América Latina y el Caribe”, el Banco Mundial señalaba que la exposición del subcontinente a

riesgos de desastres naturales, económicos y sociales se había triplicado entre 1970 y 2014. Las lluvias y sequías extremas son una amenaza constante, según el estudio, hasta tal punto que las tormentas o inundaciones representan siete de cada diez eventos naturales registrados.

En el Caribe, por lo menos un país sufre cada año los impactos de un huracán o un ciclón. El Corredor Seco, una región de bosque tropical que se extiende por toda Centroamérica, sufre sequías recurrentes que ponen en peligro los cultivos, la ganadería y la seguridad alimentaria del istmo. Los países andinos y centroamericanos se encuentran dentro del Cinturón de Fuego del Pacífico, la mayor cadena de volcanes del planeta y escenario de tal actividad sísmica que allí ocurre cerca del 90% de los terremotos del mundo. En América del Sur, con sus inmensos Andes, tiene lugar más de una cuarta parte de todos los terremotos de magnitud 8,0 del mundo.

Si equiparamos una crisis sanitaria como la del coronavirus a un desastre natural, los impactos de otros riesgos e incertidumbres pueden ser igualmente complejos en la región: mercados volatilizados, conflictos y desplazamientos forzados, crisis económicas atendidas con recursos escasos e instrumentos menos eficientes que en las naciones desarrolladas. La fotografía muestra una escena crítica. Si los países de Europa, con sus Estados de bienestar más poderosos y mejor financiados, sufrieron de qué manera el impacto del COVID-19, imaginemos lo que ello puede significar en las naciones de ALC. Ahora combinemos una pandemia con un huracán o un terremoto, o con una megacrisis económica y social en el mismo año. No es ficción. Alguna de estas combinaciones es posible en cualquier momento.

Colombia lo vivió en carne propia cuando, además de registrar altos índices de infección por COVID-19 y el consecuente debilitamiento de la economía, el país sufrió graves inundaciones invernales debido al fenómeno de La Niña, mientras que el archipiélago de San Andrés, Providencia y Santa Catalina era sacudido por la fuerza devastadora del huracán IOTA. La

infraestructura de la isla de Providencia quedo destruida en un 98%, según estimados del gobierno nacional.

Por eso hoy día existe un creciente consenso acerca de la necesidad de diseñar una política social preventiva, flexible y adaptable al riesgo. La política social del Estado de bienestar digital inteligente —el EBDI— se debe enfocar en identificar poblaciones vulnerables y en realizar esfuerzos por disminuir la probabilidad o intensidad de las amenazas y mitigar sus efectos nocivos.

Esta visión del Estado requiere introducir innovaciones en la manera en que se resuelven los problemas sociales actuales, primero desarrollando arquitecturas que permitan a sus sistemas analizar datos en tiempo real e implementando herramientas estadísticas que predigan fenómenos de riesgo, para luego crear las capacidades institucionales que posibiliten una respuesta efectiva. Las nuevas políticas públicas deben atender los fenómenos sociales de una forma más holística estudiando las vulnerabilidades y amenazas a las que se enfrentan las personas tanto a nivel agregado como individual, y tener las herramientas para enfrentarlas y mitigarlas. Adicionalmente, los gobiernos deberán dotarse de un nuevo marco ético y normativo para regular el desarrollo de la IA e inspirar confianza en los ciudadanos. En últimas, se trata de que estos tengan la esperanza de que sus vidas serán mejores, de que no hay un C3PO listo a reemplazarlos en la fábrica de refrigeradores, ni un Terminator a punto de acabar con sus familias porque salieron sin documentos de casa. Pero sobre todo deberán tener la esperanza de que la riqueza no se seguirá concentrando en las manos de unos pocos y de que el mundo será más justo y sano. ¡Tamaño labor!

Una tecnología no es una creación inocua. En ella está la subjetividad de sus creadores, el modo en que entienden la realidad y el mundo. El producto que tenemos en nuestras manos no es inerte. En *The Second Machine Age*, Erik Brynjolfsson y Andrew McAfee advierten que el peligro inmanente a toda tecnología es que la complejidad subyacente —muchas veces

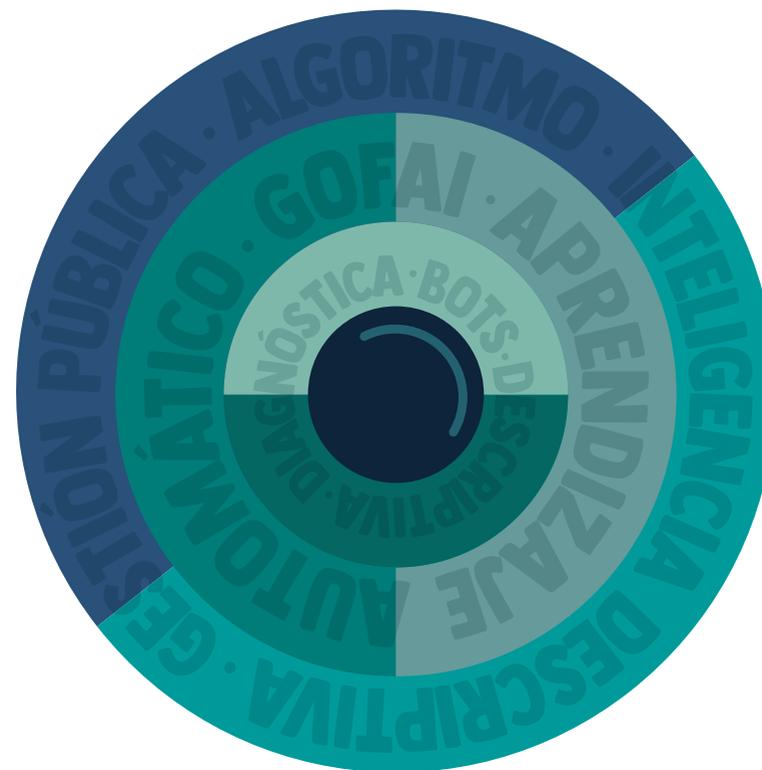
solo comprensible para sus creadores— reproduce y fortalece las desigualdades de poder existentes en una sociedad.

Para definir los alcances de las herramientas del EBDI, la participación ciudadana es tan crucial como la de los científicos, académicos y tecnólogos; funcionarios y burócratas; políticos y legisladores; reguladores y empresarios. El debate sobre cómo se debe aproximar el Estado a la IA debe ser amplio y abierto, como se verá en los capítulos 5, 6 y 7. Cuando asumimos la tecnología como una panacea, como un milagro o una solución absoluta para los problemas de nuestra existencia, es decir, cuando somos más hijos de la fe que de la razón —como diría Thomas Southwood Smith—, podemos quedar a merced de sus fallas. No hay herramienta aséptica. Al igual que todos los demás seres humanos, los programadores tienen puntos de vista, ideologías, prejuicios y filtros sociales, culturales y políticos que se expresan, consciente o inconscientemente, en los algoritmos de sus soluciones.<sup>37</sup>

En los siguientes capítulos abarcaremos los retos técnicos, institucionales, éticos y normativos de la incorporación de la IA en el Estado. El desafío para el EBDI es complejo. Debe recuperar terreno resolviendo los pendientes de la Tercera Revolución Industrial, las exigencias de la globalización, las amenazas climáticas, los déficits estructurales, y al mismo tiempo olvidar el manual de procedimientos del siglo XX para pensar en cómo anticipar un mundo por venir que además es invisible.<sup>38</sup>

<sup>37</sup> Observemos nada más la dinámica de nuestra oficina o empresa, pues sucede a menudo: la tecnología favorece a quienes tienen las habilidades necesarias para emplearla y por esa vía eleva su productividad.

<sup>38</sup> La tarea es ingente. Debe cerrar las brechas sociales y económicas entre los ricos de las ciudades y los pobres que habitan en los extramuros y en el campo. Debe elevar la calificación técnica de su mano de obra para acercar la economía a mercados de mayor valor agregado para poder ser competitivos en 2030, 2050 y 2070. Y todo ello lo debe hacer con el mismo personal que mantiene hoy en niveles apenas post analógicos, de modo que va a necesitar reentrenarlo o reemplazarlo por una nueva generación. Y eso no es todo. Debe igualmente prever un mundo nuevo y decidir cómo será el Estado del futuro en un océano de incertidumbre económica y laboral.



### **CAPÍTULO 3**

#### **Terminator no existe**

*La IA está lejos de ser un sistema autónomo capaz de dominar el mundo. Es, sí, una herramienta que puede ayudarnos. El trayecto de la inteligencia descriptiva a la predictiva.*

El 31 de diciembre de 2019, la plataforma canadiense BlueDot envió un despacho urgente a sus clientes en los organismos de salud del mundo. En la ciudad de Wuhan, 800 kilómetros al este de Shanghái, había surgido un clúster de casos de neumonía, posiblemente por la exposición de los vendedores del mercado de mariscos de Wuhan a animales vivos contaminados con una traza de un virus nuevo. Lo peor —o lo mejor— es que BlueDot ya sabía lo que seguiría. El coronavirus saldría de China en los días siguientes y contagiaría a cientos de personas en Bangkok, Seúl, Taipéi y Tokio, y así lo informó.

Menos de dos semanas después de las primeras alertas, los 3.771 ocupantes de un crucero eran retenidos en cuarentena en el puerto japonés de Yokohama. Más de 600 resultarían infectados unas semanas después. Desde entonces, las ciudades y países a los que BlueDot había anticipado que viajaría el virus registraban enfermos: unos 60 casos en febrero descubiertos por las autoridades de Tailandia y Taiwán, y unas 200 personas en Corea del Sur, 100 de ellas en una reunión de una secta religiosa. Antes del primer trimestre de 2020, casi 76.000 personas en China habían resultado infectadas del SARS-CoV-2, nombre que la Organización Mundial de la Salud (OMS) dio a la cepa.

La crisis fue reportada por los noticieros de todo el mundo, pues el coronavirus detectado se expandió en poco tiempo a una veintena de naciones de Asia, Europa, África y América del Norte: cualquier continente adonde llegaba un avión.

¿Cómo pudo una empresa social canadiense como BlueDot predecir el salto del coronavirus de China a otras naciones con tal precisión? ¿Tiene la compañía BlueDot una IA omnipresente que puede adivinar lo que va a suceder?

Desde 2014, la empresa trabaja para una docena de gobiernos del mundo —incluido el de Estados Unidos— vigilando la propagación de enfermedades infecciosas. Todo empieza con la programación de un robot dedicado a “raspar la web” (*web*

*scraping*),<sup>39</sup> el cual recolecta noticias en 65 idiomas diferentes de medios masivos, declaraciones oficiales, redes sociales y blogs de discusión de enfermedades en plantas y animales. Esos datos se cargan en un sistema de traducción<sup>40</sup> para unificar las noticias en un solo idioma que permita hacer búsquedas de palabras claves y contenido semántico.<sup>41</sup>

Esta minería de datos permite buscar correspondencias dentro de volúmenes significativos de información que podría parecer inconexa. Su objeto y cobertura dan cuenta de muertes extrañas sin explicación aparente o de enfermedades con síntomas que no se relacionan con ningún diagnóstico conocido —o que tienen similitud, pero muestran variaciones llamativas— reportadas por las autoridades en publicaciones periódicas especializadas, y hasta en blogs y foros de internet. El objetivo es identificar patrones espaciotemporales de menciones de enfermedades o infecciones. Con esta información estructurada, BlueDot elabora un informe de salud que envía a sus clientes para que eviten las zonas de peligro.

En el caso del SARS-CoV-2, BlueDot cruzó de inmediato la información que había detectado sobre Wuhan con miles de datos de billetes de avión, ya que las rutas aéreas —como en el pasado los galeones que cruzaban los mares— son importantes propagadoras de enfermedades. Así pudo determinar cuáles

<sup>39</sup> El *web scraping*, también conocido como *web crawling*, es el proceso de recuperar o extraer datos de un sitio web mediante un algoritmo o *bot* automatizado que descarga directamente el contenido a una página web y lo estructura en una base de datos para que pueda ser analizado.

<sup>40</sup> BlueDot no ha suministrado información suficiente sobre la etapa de traducción, pero se puede asumir con seguridad que se utilizan sistemas de traducción de lenguaje natural basados en aprendizaje automático, ya sean propios o mediante servicios de empresas como Google Translate o Microsoft Translator.

<sup>41</sup> Kamran Khan, el fundador de BlueDot, señaló que el motor de su algoritmo está tan bien entrenado que puede identificar, por ejemplo, si hay un brote de anthrax en Mongolia o una reunión de la banda de *heavy metal* Anthrax.

eran los vuelos con más conexiones entre Wuhan y otros sitios en China, y dio a conocer los saltos que daría el coronavirus desde su lugar de origen a otras ciudades asiáticas.

En 2016, BlueDot ya había ayudado a Estados Unidos, uno de sus usuarios más importantes, a prepararse para el brote del virus Zika que surgió ese año. Su algoritmo cruzó entonces datos de itinerarios de vuelo, densidades de población de las ciudades, mapas de temperatura y clima e información sobre los tipos de mosquitos que proliferan en cada región del mundo. Así, los Centros para el Control y la Prevención de Enfermedades (CDC por su sigla en inglés) supieron que debían prestar especial atención a los aeropuertos de Florida, adonde llegaban muchos viajeros de Brasil, el país más golpeado por el Zika.

La predicción anticipada de BlueDot ese 31 de diciembre de 2019 sobre el virus causante de la enfermedad COVID-19 fue crucial. China tiene sobrada reputación internacional de poca transparencia. Habitualmente, sus autoridades no se muestran dispuestas a proveer información al mundo sobre contaminación ambiental, enfermedades infecciosas o desastres naturales que ocurren en su territorio. Eso limita el trabajo de las organizaciones de vigilancia sanitaria, en especial de la OMS. Esta última notificaría al mundo que Wuhan tenía una crisis de riesgo global el 9 de enero de 2020, poco más de una semana después de haber recibido y corroborado la notificación del gobierno chino sobre los primeros 27 casos ocurridos casi un mes antes, a comienzos de diciembre. En el momento en que la OMS hacía su anuncio, las autoridades de Wuhan declararon estado de sitio. Los drones de las cadenas de noticias que sobrevolaban la ciudad mostraron una imagen anticipada del apocalipsis: avenidas desiertas y ni un alma en las aceras o en los parques.

En el caso de Wuhan, la IA permitió superar las barreras informativas del gobierno de China que la OMS no puede sortear, pues sus decisiones dependen de los datos que suministren los funcionarios de Beijing. En cambio, las telarañas de la IA de

BlueDot pueden esquivar buena parte de la censura indagando en cada rincón de la web por pistas indirectas producidas de manera atomizada, a menudo más difíciles de controlar para la seguridad del Estado.

Aunque el proceso analítico de BlueDot no es especialmente avanzado (comparado con sistemas como los que usa Tesla en el piloto automático de sus autos, por ejemplo), sí es una operación de análisis masiva y veloz. Obtener los mismos resultados con trabajo humano exigiría emplear durante mucho tiempo a miles de personas hablantes de muy diversos idiomas, completamente dedicadas a análisis lingüístico, contextual y semántico, y luego disponer de editores especializados en ciencias que puedan leer la intertextualidad de los informes. Jamás conseguirían un resultado cercano al de los algoritmos en la misma fracción de tiempo.

Todavía es muy pronto para determinar cuántas vidas ha salvado el algoritmo usado en Wuhan, pero su aplicación facilitó, como nunca antes, la toma de decisiones durante la crisis. En 2003, cuando el virus SARS se expandió de China a Hong Kong, y de allí a otras naciones, pasaron cuatro meses entre el descubrimiento de los primeros casos y la notificación global de la OMS.<sup>42</sup>

El coronavirus hizo evidente que hoy tenemos algoritmos para intervenir en la salud y muchos otros campos. Incluso algoritmos en apariencia sencillos pueden tener un impacto gigantesco en las decisiones que toma el Estado de bienestar digital inteligente, EBDI. El análisis predictivo permite a los gobiernos atender la probabilidad de que surja un problema antes de que este se manifieste; la tarea pendiente es lograr que los Estados tengan la flexibilidad necesaria para responder a estas emergencias.

<sup>42</sup> China demoró la notificación del SARS a la OMS hasta febrero de 2003, tres meses después de detectar los primeros casos. Para entonces ya había más de 300 infectados.

### Las máquinas al rescate

El algoritmo de Wuhan no es el ejemplo sino una de las posibilidades estratégicas de la implementación de IA en las políticas sociales. Y no tomará décadas. La IA está lista para cambiar el mundo. Todos hemos escuchado el consejo de nuestros mayores de dar un paso a la vez, pero los avances de la estrategia asistida por inteligencias no humanas son saltos que no podremos ignorar.

En el pasado, nuestra aproximación a las enfermedades había sido principalmente reactiva (sucedió una catástrofe y el Estado actuaba) o preventiva (financiando un seguro de salud o de jubilación para cuando llegáramos a viejos o estuviéramos enfermos). Aún hoy, buena parte de las intervenciones frente a las situaciones de crisis mantiene ese perfil. Cuando surgió el ébola en Guinea en 2013, pasaron meses antes de que los primeros investigadores llegaran a la villa considerada la “zona cero,” nada más para encontrarse con que el árbol donde solían dormir los murciélagos que transmitían la enfermedad había sido quemado y los coleópteros habían emigrado.

Este problema nos ha acompañado por muchas décadas. Las personas infectadas que son portadoras sanas o los animales que operan como reservorios vivos de los virus solían —y aún lo hacen— dejar sus sitios de origen sin seguimiento ni control, listos para transmitir su carga fatídica a otros seres humanos. Las pestes no son asunto frívolo. La mayoría de las infecciones —incluidos el ébola, el SARS, el VIH, el coronavirus o la enfermedad de Lyme— nacen y se incuban en otras especies antes de pasar a las personas. Para el momento en que sabemos qué patógeno atacó, el bicho ya se ha mudado.

La respuesta histórica a estos problemas —y a otros como las catástrofes— es a menudo *post facto*: se declara una enfermedad y se pasa a controlar, aislar y mitigar. De algún modo, es un curso de acción que, por décadas, suponía que el Estado siempre llegaría tarde. Era razonable pensar así. No teníamos la tecnología y/o los avances necesarios para anticipar sucesos.

Pero ya los tenemos.

La IA permite detener esa reacción en cadena justo antes de que suceda. Barbara Han cree que la aproximación tradicional ya no es suficiente. Han es una ecóloga de enfermedades del Cary Institute of Ecosystem Studies, una institución que opera desde los bosques de Millbrook, localizados un par de horas al norte de la ciudad de Nueva York. Ella estudia la conflictiva intervención de la humanidad en el medio ambiente y sostiene que el viejo paradigma del Estado reactivo, es decir, esperar hasta que la gente se enferme, es inadecuado. “No querrás repartir paraguas después de que llueva”, dice ella. “La idea es pronosticar la lluvia antes de que comience.”

En el pasado, la reacción de los Estados ante eventos catastróficos como terremotos, incendios o una epidemia de gripe exigía un proceso más o menos largo de recolección, procesamiento y análisis de datos; diseño de políticas; a veces una prueba piloto, y finalmente el despliegue de las medidas respectivas en la sociedad. La curva entre la ocurrencia de un caso —fuera una epidemia o una crisis económica— y la respuesta era —todavía lo es— larga, y por eso mismo facilitaba que los impactos graves se mantuvieran o se ampliaran. Incluso hoy día, los Estados planifican estrategias de mediano y largo plazo con base en información desactualizada.

Insistimos. Ya no debe ser así pues tenemos herramientas para que no lo sea.

La Cuarta Revolución, la de la inteligencia, ofrece una nueva oportunidad para reorganizar el Estado de bienestar con base en criterios de equidad, eficiencia y flexibilidad. Por eso la IA es una herramienta estratégica para diseñar políticas sociales dinámicas, adaptativas, preventivas y predictivas en el EBDI.

Se trata de un nuevo paradigma. Hoy la humanidad tiene la oportunidad, por primera vez desde sus orígenes, de predecir con un grado elevado de certeza la ocurrencia de eventos

peligrosos a gran escala antes de que sucedan y de que sus efectos colaterales se manifiesten.<sup>43</sup>

Los signos que anuncian una crisis a menudo son indiscernibles para el cerebro humano. ¿Cuántas personas se habrían necesitado para detectar pistas sutiles del coronavirus en las conversaciones en línea de los habitantes de Wuhan? Sin embargo no lo son para modelos computacionales capaces de hacer **minería** de datos a una escala inconcebible para cualquier individuo incapaz de multiplicar mentalmente 234 x 17 sin rendirse en el intento.

### Siri: ¿de veras estás ahí?

Pero, ¿es buena idea introducir la IA en la gestión pública? ¿No se corren riesgos significativos cuando se ensaya una tecnología tan nueva en el manejo de tantos asuntos y todos tan importantes?

Unas páginas atrás hablamos de Skynet, Terminator y Hal 9000. Demos un paso lateral y retomemos el miedo a lo desconocido.

El polímata y filósofo sueco Nick Boström —una referencia obligada cuando se habla de los riesgos de las superinteligencias— clasifica a los sistemas inteligentes en tres niveles: la IA acotada o débil, a saber, un sistema híperespecializado en una sola tarea; la IA general o fuerte, es decir, un sistema de uso general capaz de aprender de sí mismo y generalizar llevando su aprendizaje a otras tareas; y finalmente lo que Boström llama la súperinteligencia, o sea, un sistema que “supere radicalmente a las mejores mentes humanas en todos los campos, incluyendo la creatividad científica, la sabiduría general y las habilidades sociales.”

<sup>43</sup> También es posible que estos nunca sucedan, pues una predicción no es más que eso y no una certeza absoluta sobre algo predeterminado. El valor de la predicción de un modelo computacional reside en su capacidad de crear hipótesis plausibles dados N escenarios. Y eso permite diseñar posibles respuestas o anticipar acciones.

En los últimos años, asistentes virtuales como Siri, Alexa y Cortana han mejorado mucho. Hoy son capaces de hacer tareas complejas que antes hubieran requerido nuestra intervención: desde buscar un número de teléfono hasta darnos el pronóstico del tiempo antes de salir para una cita que, de hecho, ellas mismas nos han recordado minutos antes. Sin embargo, no es mucho más lo que pueden hacer. Una IA general con la que podamos conversar sobre todo y sobre nada, compartir sueños, felicidades y tristezas —una IA que se comporte como la meliflua asistente virtual que Scarlett Johansson interpretó en la película Her— es algo que aún está muy lejos.

No hay Skynet ni Terminator. No hay Hal 9000.

Tenemos que repetir esto como un mantra cada vez que veamos una noticia con un titular descabellado cuyo único objetivo es monetizar clics por la vía del escándalo. No existe inteligencia artificial consciente de sí misma; estamos a una distancia significativa de una inteligencia artificial general (si es que fuera posible) que sea capaz de razonar intuitiva y subjetivamente como la humana.<sup>44</sup>

Este es el tipo de historias que devoramos con gusto en la ficción, pero que no forma parte de la existencia real de la IA. Sin embargo, han alimentado nuestras fantasías al punto de construir la imagen de un monstruo peligroso. Por supuesto que nadie quiere una tecnología que se presume riesgosa en las cercanías de los archivos que guardan nuestros datos más personales, incluida nuestra información médica y fiscal o los detalles de la vida de nuestros hijos.

El siguiente caso es ilustrativo. En junio de 2017, miembros de la División de Investigación de Inteligencia Artificial de Facebook

<sup>44</sup> Esta IA general sigue siendo un sueño, o una pesadilla si se quiere, pero no existe; no ha sido inventada, y no hay consenso sobre cuánto tiempo creemos que nos va a tomar llegar a ella, si es que lo logramos en algún momento.

propusieron un nuevo método para entrenar *chatbots* en tareas de negociación. Los investigadores pusieron a dos *chatbots* a entrenarse mutuamente para mejorar los resultados. Se llamaban Bob y Alice y debían dialogar para negociar la propiedad de un grupo de objetos virtuales.

El modelo, construido sobre la base de miles de situaciones reales de negociación, utilizaba también un método de aprendizaje por refuerzo en el que el algoritmo tenía la tarea de simular los posibles resultados de la conversación para planificar mejor la siguiente respuesta.

Pero algo salió mal. Los autores identificaron que, en ocasiones, los *chatbots* generaban frases sin sentido como “Las bolas tienen cero para mí para mí para mí para mí para mí para mí para mí” (Alice) o “tú yo yo todo lo demás” (Bob). No se trataba de una discusión matrimonial, sino evidentemente de una desviación del programa original de los ingenieros. Y no debía suceder.

Pero como el *bot* receptor —supongamos Bob— no identificaba la frase del emisor —o sea Alice— como gramatical o sintácticamente errónea sino como un nuevo insumo lógico y para nada contradictorio, lo tomaba como válido, lo incorporaba a su sistema, lo analizaba y diseñaba su propia respuesta con base en él para continuar la conversación. En suma, Bob construía otra capa semántica asociada a ese insumo incomprendible para los humanos, pero significativo para Alice, quien ahora entendía que Bob estaba intentando decirle algo relacionado con su mensaje anterior; así que ella lo reprocesaba y creaba una nueva respuesta. Cuanto más tiempo pasaba, más afinaban Bob y Alice las secuencias lógicas de entendimiento mutuo y más se alejaban de algo que pudiera considerarse una comunicación en inglés humanamente comprensible.

Bob y Alice no inventaron su propio lenguaje. Se comunicaban rompiendo reglas gramaticales —las nuestras— y creando nuevos mecanismos de interacción aprendidos con cada intercambio. Lo hicieron porque eso les permitía ser más

eficientes en las negociaciones. De hecho, varias resultaron efectivas con su método. Pero como el nuevo lenguaje resultaba incomprendible para los autores originales, los ingenieros decidieron simplemente incluir una restricción obligando a los *bots* a respetar la gramática del inglés.

Ahora, imagine por un momento que usted es el reportero que descubre esta historia, que tiene pocos escrúpulos, que a su editor la ética no le quita el sueño, y que su medio de comunicación tiene incentivos para que millones de lectores hagan un clic en la noticia, contribuyendo a generar tráfico de audiencia y eventualmente anunciantes. Entonces escribe el siguiente titular: “La IA está inventando un lenguaje que los humanos no podemos entender. ¿Deberíamos detenerla?” Este titular es real. Lo publicó la revista *Fast Company* en julio de 2017.

Piénselo un poco y la nota se escribe sola. Una empresa pone a dos robots a conversar y estos en poco tiempo inventan un lenguaje que sus creadores no comprenden pero que es muy útil y eficiente para las máquinas. ¡Skynet!

La bola echó a rodar de inmediato. Los periódicos sensacionalistas hicieron su agosto con titulares catastrofistas. En Inglaterra, *The Sun* publicó un artículo con el título “Facebook apaga el experimento de IA después de que dos robots empiezan a hablar en su PROPIO idioma, que solo ellos pueden entender.” El texto, acompañado de imágenes similares a las de la película “Yo Robot” (donde humanoides rompían las tres leyes de la robótica de Isaac Asimov<sup>45</sup>), insinuaba que las máquinas estaban

<sup>45</sup> Conjunto de reglas ideadas por Isaac Asimov y usadas en sus novelas de ciencia ficción para proteger a los seres humanos en un futuro con IA. Las leyes, publicadas por primera vez en 1942 en la novela *Runaround* de la serie “Yo Robot”, rezan: (1) un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra un daño; (2) un robot debe obedecer las órdenes que le den los seres humanos, excepto cuando dichas órdenes entren en conflicto con la Primera Ley; (3) un robot debe proteger su propia existencia siempre que dicha protección no entre en conflicto con la Primera o con la Segunda Ley.

fuera de control y que a los investigadores no les había quedado otra opción que desconectarlas para salvarnos del apocalipsis.<sup>46</sup>

Cualquier publicación de este tipo es irresponsable, no solo porque es falsa, sino porque además evita que nos enfoquemos en los problemas reales. No, los *bots* no estaban inventando un lenguaje para alejarse de la humanidad ex profeso. No, los *bots* no fueron apagados por estar fuera de control, y no, no por eso estamos más cerca de tener una inteligencia artificial cercana a la inteligencia humana y, como tal, capaz de producir maldad conscientemente, **moralmente**.

Todos los avances técnicos en materia de inteligencia artificial han producido hasta ahora una IA limitada e hiperespecializada en una tarea que puede lograr con excelencia. Esa IA acotada puede percibir, procesar, clasificar objetos, traducir lenguajes, pero no lanzará Terminators 101, T-800, T-850 y T-1000 en pos de la gente.

Y esta IA acotada es la que necesita el EBDI, ese nuevo Estado de bienestar digital e inteligente.

Aunque esa IA haga una sola cosa, el encadenamiento de modelos puede emplearse para resolver procesos muy

---

<sup>46</sup> Muchos vieron con horror el impacto de un trino. Para empezar, los *bots* no habían hecho nada por su cuenta. Los ingenieros de Facebook programaron el desempeño de Alice y Bob, pues querían entender cómo opera la lingüística en instancias de negociación. Los *bots*, diseñados para experimentar con el lenguaje de modo tal que favoreciera su posición negociadora, encontraron un atajo en la invención de monemas más simples. Bob y Alice dotaron de sentido a sus intercambios. En la práctica, procuraban entenderse como dos softwares que deben compatibilizar para no detener un proceso. Nunca se trató, como insinuaban algunos medios, de que los *bots* hubieran creado su propio lenguaje para esquivar el control de sus amos humanos, sino que las redes neuronales artificiales usaban el lenguaje humano modificándolo según fuera necesario para mejorar las transacciones. Esto no es nuevo. Google también detectó que sus redes neuronales interpretaban la semántica de varias oraciones, y ya existen proyectos para dejar que las máquinas creen sus propios lenguajes artificiales que les permitan mejorar su capacidad de procesar información más rápidamente. ¿Da miedo? Siempre hay gente a cargo.

complicados cuyo impacto podría ser sustancial en la manera en que entendemos el Estado. A lo largo de este libro veremos que los riesgos de la IA no se encuentran tanto en los sistemas avanzados sino en el lado más humano de la relación. Los gobiernos y las personas están en mora de acometer la tarea pendiente de entender plenamente y mitigar los obstáculos potenciales y reales que representa adoptar y usar cualquier tecnología, y esta en particular.

Retomemos entonces la conversación y aterricémosla, porque enfocarnos en los escenarios apocalípticos con máquinas que conspiran a nuestras espaldas en idiomas sintéticos nos impide ver tanto el potencial del sistema como sus riesgos reales.

### Primero lo primero: ¿de qué se trata?

¿Qué es, a todas estas, eso que llamamos inteligencia artificial, más allá de un programa computacional como la presentamos al principio? Definirla nunca será una tarea sencilla porque es un concepto de amplia cobertura, una forma de llamar a todo y a nada en particular. Lo que hoy llamamos IA no existía hace veinte años, del mismo modo en que aquello que denominábamos inteligencia artificial hace dos décadas ahora se consideraría un estadio inferior ya superado.

En términos generales, la IA es un campo que desarrolla capacidades en sistemas computacionales para realizar tareas y actividades consideradas como “inteligentes”. Ese calificativo entraña el primer problema. ¿Qué es la inteligencia? ¿Es un sinónimo de percepción, audición, visión, gusto, tacto y olfato? O, ¿es más bien procesamiento en el sentido de entendimiento, generación, traducción y simplificación? ¿Es alguna medida de sentido común? ¿Es aprendizaje o capacidad de reaccionar de formas diversas y creativas?

Si inteligencia es un término de alcance vasto, ¿una máquina es entonces capaz de **pensar**?

El matemático Alan Turing fue el primero en plantear esta pregunta en 1950, y para responderla propuso el “juego de la imitación”, un mecanismo a través del cual se evaluaría la inteligencia de una máquina por su capacidad de tener una conversación al nivel de un humano. Esto significa que la máquina debía hacerlo de manera tan perfecta que los humanos que se encontraban al otro lado debían suponer que hablaban con otra persona y no con un software.

La evaluación consistía en poner a uno o más individuos en un cuarto equipado con un teclado y una pantalla desde donde tendrían que intercambiar mensajes en un chat tanto con otros humanos como con máquinas. Si esas personas —que actuaban como evaluadores— no podían distinguir entre una computadora y un humano en más del 50% de las ocasiones, entonces, según el modelo de Turing, la computadora estaría exhibiendo “inteligencia.”

Esta prueba sería refutada en los años ochenta, no ya desde las ciencias matemáticas sino desde la filosofía, con el llamado “argumento del cuarto chino”. Formulado por John Searle,<sup>47</sup> con su teoría del “cuarto chino” este criticaba el que Turing hubiera planteado un escenario de evaluación donde únicamente se medía la capacidad de imitación y engaño del sistema, sin poner una condición estricta a la máquina sobre su capacidad real de **razonar**.

<sup>47</sup> Searle decía encontrarse en una habitación siguiendo un programa de computadora que produce caracteres chinos. Él no habla chino, pero al seguir el programa puede manipular símbolos y números como una computadora, produciendo cadenas más o menos apropiadas de caracteres chinos que deslizaba bajo la puerta a un grupo de personas. Muy pronto, esas personas que están fuera comienzan a creer que hay un chino en la sala contigua. El argumento de Searle es que al programar una computadora digital se puede hacer que parezca entender el lenguaje pero no produce comprensión real, pues la máquina usa reglas sintácticas para manipular cadenas de símbolos pero no comprende el significado o la semántica de lo que usa. Para Searle, una mente humana no procesa información como las computadoras; es una entidad biológica que una computadora solo puede imitar.

La interacción entre Turing y Searle ha sido fundamental para el campo de la IA porque ha determinado una historia destinada a repetirse. Alguien define una evaluación de “inteligencia” para un algoritmo, algo que se cree es único y establece el límite al que puede llegar la inteligencia humana en un momento dado (la máquina puede jugar al ajedrez, al milenar juego de Go e incluso reconocer los gatos que aparecen en una postal). Pero luego aparece alguien más que produce un nuevo avance y logra que un algoritmo se desempeñe a un nivel superior al de los humanos en algún campo. Así redefine la cosa: esto es inteligencia y lo anterior ya no.

La periodista y escritora Pamela McCorduck se refiere a la “paradoja de la IA” para definir ese fenómeno según el cual, cada vez que logramos que un sistema computacional ejecute una tarea que no había podido hacer antes, lo catalogamos como IA. A la luz de esta nueva tarea, todos los resultados anteriores que nos habían impresionado y hecho escribir numerosas notas periodísticas con titulares en mayúsculas, se vuelven cosa de niños: reglas condicionales en forma de árbol de decisión; simples matemáticas aplicadas.

Entonces, si no podemos aproximarnos a la definición de IA bajo una regla estricta, hagámoslo mediante una visión histórica. A partir de 1950, muchas técnicas y corrientes han entrado y salido de este paraguas conceptual llamado inteligencia artificial, pero hasta ahora solo dos la han modelado como campo de estudio: la GOFAI y el aprendizaje automático.

La GOFAI —sigla en inglés de Good Old Fashioned Artificial Intelligence— es la inteligencia artificial buena y anticuada, a saber, el paradigma más importante desde mediados de los años setenta hasta mediados de la década de los noventa. Estos algoritmos encuentran sus decisiones óptimas siguiendo reglas predefinidas. Por ejemplo, cuando en un edificio dos personas en distintos pisos presionan el botón del elevador al mismo tiempo, el software que decide a qué piso ir primero es un GOFAI. Lo mismo sucede con el algoritmo Dijkstra, en el cual se basaron las primeras versiones de Google Maps y que

nos recomienda la ruta a seguir para llegar a nuestro destino. Incluso la mayoría de los *chatbots* utilizan todavía árboles de decisión predefinidos; ante un comentario del usuario, seleccionan una respuesta escrita de antemano que la máquina interpreta más o menos de acuerdo con la pregunta planteada por la persona. Todo GOFAI.

Sin embargo, la GOFAI se enfrenta a dos grandes obstáculos de carácter instrumental y metodológico. Primero, debe construirse y alimentarse a mano considerando escenarios y construyendo sus reglas lógicas, para lo cual se requieren decenas o cientos de horas de trabajo. Y segundo, para que funcione es preciso conocer el objetivo y el camino lógico que lleve a resolver el problema en cuestión. Esto limita su capacidad de responder a situaciones nuevas e imprevisibles. Es por ello que en ocasiones le cuesta abstraer, es decir, incorporar conocimiento de manera autónoma y aplicarlo a un nuevo problema. De ahí que se afirme que la GOFAI se ha construido para razonar y no para aprender, como sí lo hacen los sistemas de aprendizaje automático.<sup>48</sup>

Este último campo expresa la paradoja de McCorduck. Hoy, cuando hablamos de IA la asociamos con el aprendizaje automático. La GOFAI nos recuerda el juego Atari: adorable, sí, pero anticuado.

El aprendizaje automático se hace fuerte donde la GOFAI falla. Sus técnicas permiten que un algoritmo identifique patrones o comportamientos de un conjunto de datos o una simulación, sin necesidad de que un humano le dé instrucciones explícitas.

<sup>48</sup> Esto no implica que la IA GOFAI sea obsoleta. En los últimos años se han desarrollado aplicaciones que combinan ambos paradigmas. Ejemplo de ello son los algoritmos de procesamiento del lenguaje natural que suelen combinar enfoques estadísticos (basados en grandes cantidades de datos) y enfoques GOFAI (que consideran temas como las reglas gramaticales). Igualmente, existe una convicción extendida acerca de que, para continuar con el desarrollo de la IA general, el aprendizaje automático no será suficiente y que por lo tanto se deberán encontrar mejores formas de combinar ambas corrientes.

Muchos de los métodos usados actualmente por el aprendizaje automático ya existían en la década de los años cincuenta, pero en ese entonces no había suficiente información para ajustar un buen modelo y las computadoras tampoco tenían la capacidad de procesamiento necesaria para correrlo. La disponibilidad de grandes volúmenes de datos, así como el incremento enorme de la capacidad de recolección y procesamiento de información, volvieron la IA práctica, escalable y posible.

El valor del aprendizaje automático radica en predecir o clasificar, segmentar o agrupar, y aprender por refuerzo o interacción: todo lo que hoy esperaríamos de la IA.

El primer ejemplo —clasificar o predecir— es propio de la llamada IA “supervisada.” Aquí el algoritmo intenta hallar patrones entre los datos que expliquen un resultado conocido de antemano. El objetivo es predecir con cierta precisión casos desconocidos durante el entrenamiento de la máquina. El reconocimiento facial, por ejemplo, implica aplicación supervisada. El sistema descompone las fotografías de una persona en una serie de atributos y los asocia con los nombres o con la identificación de ese individuo en la base de datos. Cuando se le muestra una fotografía de esa persona que no ha visto, el sistema evalúa y la asocia al nombre correcto.

La segunda aplicación —segmentar o agrupar— es propia de la IA “no supervisada.” Aquí el algoritmo detecta patrones estructurales en la información y crea asociaciones de datos similares (clusterización) para reducir el número de dimensiones o dar un resultado significativo. En 2019, por ejemplo, Google lanzó SMILY, un sistema de aprendizaje profundo entrenado para identificar patrones de enfermedades entre más de 5.000 millones de imágenes. El algoritmo no buscaba una enfermedad específica sino imágenes parecidas entre una colección, sin etiquetas o clasificaciones a priori. Los médicos solo subían la fotografía y el sistema les devolvía un grupo de casos con componentes similares para que aquellos formularan sus diagnósticos u ordenaran nuevos exámenes.

En la tercera opción —aprendizaje por refuerzo—, el algoritmo debe hallar las soluciones óptimas por exploración y ensayo/error sin que los programadores le hayan proporcionado un solo ejemplo. Esto significa que aprende interactuando con el entorno. Cada acción resulta en un premio o castigo, de manera que la IA aprende a maximizar sus ganancias o a reducir los castigos. Ese agente internaliza una política, o sea que determina cuál es la acción que debe emprenderse en una circunstancia N. El aprendizaje por refuerzo se ha usado mucho para entrenar algoritmos que sepan jugar videojuegos, por ejemplo.

### Qué tipo de IA y para qué

Ahora bien, trátase de GOFAI o de aprendizaje automático, la IA actual comprende sistemas de toma y soporte de decisiones que ya pueden adaptarse a situaciones imprevistas: aprenden de ellas. Insistimos. Aún estamos lejos de una IA general —la Skynet de Terminator— y muchos de los investigadores destacados en este campo sostienen que para lograrlo es necesario que ambas corrientes se unan, pues no hay inteligencia sin razonamiento, pero tampoco hay inteligencia sin aprendizaje.

Los sistemas de IA son solo una herramienta. La forma en que la utilicemos dependerá únicamente del objetivo perseguido por su diseñador y de la forma en que se introduzca en el día a día de la gente. El mismo algoritmo de reconocimiento facial puede ser usado para vigilancia estatal y espionaje o como método de autenticación en un iPhone. En ambas, el propósito es el mismo: singularizar a alguien. Sin embargo, los objetivos finales son sustancialmente distintos.

Hay muchas formas de organizar los sistemas de IA,<sup>49</sup> pero desde el punto de vista de la política pública la clasificación

<sup>49</sup> Por ejemplo, se pueden clasificar por su paradigma de aprendizaje (supervisado, no supervisado y por refuerzo), por su objetivo (clasificación, segmentación, clusterización, etc.), y a un nivel más técnico por el tipo de algoritmo (redes neuronales profundas, árboles de decisión y otros).

más relevante es entenderla desde la forma en que incide en la toma de decisiones.

Como vimos en el capítulo anterior, el Estado ha creado hasta ahora tres tipos de políticas: reactivas, preventivas y promotoras. Estas surgen, primero de la definición de la función del Estado entendida como la responsabilidad que este asume para con los ciudadanos, pero solo son posibles en función de la clase de información a la que tiene acceso el responsable por adoptar las decisiones. Qué sabemos y qué no sabemos condiciona nuestro entendimiento del mundo y de los problemas sociales que nos proponemos resolver.

En suma, las preguntas analíticas que deberíamos considerar para tomar una decisión apuntan a obtener tres tipos de información: descriptiva, que muestra el estado actual del mundo; diagnóstica, que describe cómo llegamos allí; y predictiva, que plantea escenarios probables con un rango de error. El responsable por tomar las decisiones del caso utiliza esta información como insumo y crea una acción o intervención de política pública.

Estos tipos de información han estado presentes en mayor o menor medida durante toda la vida de las sociedades modernas; sin embargo, su calidad, que incluye criterios como precisión, disponibilidad y consistencia, ha cambiado de manera significativa con el tiempo. La IA tiene el potencial de mejorar cada uno de estos tres tipos de información.

La información descriptiva busca caracterizar un fenómeno y presentar su estado actual. Esa información permite a un funcionario a cargo de la formulación de políticas públicas identificar problemas y priorizar iniciativas reactivas. Los instrumentos de recolección de datos descriptivos más tradicionales son las encuestas, los censos y los registros administrativos, es decir, las bases de datos del Estado. Nos dicen el “hoy”; el “ahora”. Es la información que vemos y podemos contabilizar.

Sabremos entonces cuántos alumnos hay por escuela, sus calificaciones y sus asistencias, es decir, todo cuanto hemos podido recolectar con papel y lápiz desde hace cientos de años, tal y como lo hizo Southwood Smith cuando relevó los datos sobre Bethnal Green en Londres durante la Primera Revolución Industrial.

Dependiendo de su diseño, los canales tradicionales como los censos y las encuestas tienen la característica de ser altamente confiables. Sin embargo, son insuficientes para capturar la variabilidad de los fenómenos del mundo actual y emplearla en tiempo real. Esto sucede, por ejemplo, con los censos que se suelen realizar cada cinco o diez años. Una población cambia sustancialmente en una década. Las decisiones de política pública que se tomen con base en esa información pueden resultar poco eficaces.

La digitalización ha creado nuevos canales de recolección de datos descriptivos. Redes sociales, plataformas de empleo, imágenes satelitales, registros de llamadas telefónicas son todos accesibles para crear métricas descriptivas que se pueden actualizar incluso en tiempo real. Aun así, su objetivo es el mismo: exponer el estado de cosas en ese momento. Hoy, por ejemplo, el sistema público de empleo de un país puede analizar la oferta nacional de puestos de trabajo y las habilidades que se requieren para luego buscar entre todos los candidatos los perfiles que tienen una mayor afinidad o relación con el puesto. Esto lo hacen utilizando algoritmos de segmentación y recomendación, de una manera más veloz y eficiente que en el pasado reciente.<sup>50</sup>

Estas nuevas fuentes de información suelen tener dos características. La primera de ellas es que se trata de datos “no

<sup>50</sup> En el siguiente capítulo describiremos el proyecto ParaEmpleo, una plataforma gubernamental de Paraguay que se enfoca precisamente en estos procesos.

estructurados” (no pertenecen necesariamente a una base de datos con columnas y filas, como un Excel, sino que pueden ser fotografías, textos, imágenes y audio, entre muchas otras opciones), y la segunda es que pueden representar volúmenes de datos muy grandes (lo que los medios llaman inteligencia de datos o macrodatos). Los algoritmos de IA ayudan a procesar y a extraer información de estos datos masivos para luego emplearla en la toma de decisiones.

También podemos considerar como información descriptiva a los sistemas de clasificación de imágenes, todos ellos algoritmos supervisados y entrenados para reconocer e identificar a una persona o un objeto usando métodos de visión computacional. La versión analógica de este proceso son los documentos nacionales de identidad que cada individuo lleva en su cartera: una tarjeta con el número que lo identifica de forma única. Los sistemas de reconocimiento facial hacen lo mismo. En Estados Unidos y Europa, donde tales sistemas se encuentran en los mostradores de inmigración de los aeropuertos, los ciudadanos ya no tienen que sellar sus pasaportes: los colocan sobre un lector electrónico, se paran delante de una cámara y el software compara el documento con sus rasgos biométricos. Los Estados que emplean estos dos elementos (tanto el documento nacional de identidad o DNI como el sistema digital de reconocimiento) buscan obtener información descriptiva de la persona que quiere entrar al país y con ello toman la decisión de permitirle o no el ingreso.<sup>51</sup>

El segundo tipo de preguntas analíticas busca diagnosticar un problema, es decir, entender qué situaciones nos llevaron al estado actual y qué tipo de políticas nos pueden ayudar a mejorar una condición no deseable: cómo sucedió esto, por qué y qué puedo aprender. Cuando los humanos racionalizamos

<sup>51</sup> Es evidente que todos estos procesos acarrearán riesgos importantes tanto en materia de gobernanza como de protección de datos, sesgos y discriminación. Expondremos estas singularidades con mayor profundidad en los capítulos 4 y 5.

el mundo intentamos comprenderlo en términos de causa y efecto; si entendemos por qué ocurrió algo, podemos alterar nuestro comportamiento para cambiar resultados futuros. Comprobar causalidad no es una tarea trivial.<sup>52</sup>

El tercer tipo de análisis —el predictivo— surge del estudio de posibles eventos futuros. Procura responder a la pregunta de **qué podría pasar**. Cuando se sabe con cierto grado de certeza que un suceso acontecerá, la persona a cargo de formular la política pública debe tomar en cuenta los riesgos y poner en marcha medidas de gestión o mitigación, ya que con información predictiva se pueden crear sistemas de alerta temprana o de detección de anomalías. Corea del Sur y Taiwán, por ejemplo, tenían modelos de propagación epidemiológica y aplicaron medidas inmediatas de mitigación del coronavirus, lo cual resultó en un menor contagio.

### Toma y soporte de decisiones

En el mercado digital sigue siendo válida una figura analógica muy simple: no se pone el carro delante del caballo. Cuando una institución quiere aplicar un modelo de IA no se hace ningún favor al encarar el diseño de una solución sin tener claridad sobre lo que quiere hacer. La ausencia de un diagnóstico preciso, de una lectura acabada de sus insumos o de las posibilidades de implementación de la solución, hará que se malgasten recursos y se perderá el impulso. Peor aún: producirá daños en el cuerpo

<sup>52</sup> Los métodos de inferencia causal han sido utilizados ya por varios años en econometría, aunque en los últimos tiempos se ha registrado allí una verdadera revolución gracias a la labor de economistas como Esther Duflo y Abhijit Banerjee, quienes recibieron el Premio Nobel de economía en 2019 por su trabajo en esta área, a partir del cual han propuesto modelos basados en aleatorización y experimentos controlados. En la mayoría de los casos, los algoritmos del aprendizaje automático no buscan describir relaciones causales dado que se enfocan principalmente en la predicción. Pero en los últimos años, trabajos como el de Susan Athey han comenzado a introducir en los algoritmos de aprendizaje de máquina estas técnicas y procesos experimentales tipo *A/B testing*, muy útiles en contextos digitales por la facilidad de crear experimentos.

social al que desea ayudar. Agregar un poco de tecnología a los problemas no es política pública; es poner un curita en la coyuntura. La IA no siempre es la solución perfecta ni milagrosa. En rigor, casi nunca lo es.

La elaboración de políticas públicas se puede caracterizar como un proceso cíclico y dinámico mediante el cual el Estado identifica un problema público presente o futuro y determina aquellos riesgos y factores que busca eliminar, o como mínimo reducir, a través de intervenciones o acciones de gobierno.

Aunque existen muchos marcos para el diseño de una política pública, las etapas generalmente suceden en el siguiente orden: se identifica un problema; se formula una respuesta de política pública; se diseña un mecanismo de focalización, segmentación o personalización; se implementa, y por último se evalúa la política y su efecto en el asunto a resolver. Tras esto, vuelta a empezar.

No hay por qué sorprenderse; el modelo es antiguo y está probado. Lo nuevo es viejo. Por eso, el punto número uno de la política pública sigue siendo el mismo de siempre: identificar con precisión el problema a resolver.

Cualquier modelo de IA debe ajustarse al ciclo de política pública. Por ende, es necesario pensar en un objetivo específico. De este modo, si el problema que nos ocupa es la deserción escolar, el objetivo debe especificar una subpoblación en la cual implementar la solución, entendiendo las limitaciones y el alcance de la intervención.<sup>53</sup> Ese objetivo ha de ser lógico, operacional, conciso y breve, y debe permanecer fijo a lo largo de la intervención. Si se establecen los límites del problema, será posible determinar el objetivo concreto al que debe apuntar la solución.

<sup>53</sup> A modo de ejemplo, un objetivo factible podría ser “reducir la tasa de deserción escolar en poblaciones indígenas y de familias monoparentales en 10%.”

Ahora bien, establecer cuál será el algoritmo necesario —cómo se construye y qué se espera de él— dependerá de la intervención de política pública. Lo importante no es si se están usando redes neuronales recurrentes, una red convolucional o una regresión lineal, sino la manera en que se use el algoritmo para adoptar las respectivas decisiones de política pública.

### Y aquí estamos

No sobra recalcar que IA no es un fin en sí mismo; es una herramienta que, mediante el análisis y procesamiento de datos captados o percibidos del entorno, genera una recomendación de actuación mediante la predicción, clasificación o identificación de patrones. Ni magia ni solución milagrosa.

Y ya que este libro es sobre la aplicación de IA en la toma de decisiones en torno a las intervenciones del Estado, tampoco sobra insistir en que los sistemas útiles para tal efecto deben estar enmarcados en el ciclo de la política pública y definidos por el tipo de riesgos que puedan surgir en su implementación.

Este es un punto importante. Aunque la IA proporcione insumos significativos, la decisión sobre cómo materializar sus proyecciones y predicciones debe quedar en manos de los seres humanos. La planificación de políticas públicas requiere de la participación decidida, determinada —única e irremplazable hasta el día de hoy— de las personas. Esto por cuanto incluso si un programa de IA genera sistemáticamente predicciones con 99% de probabilidad de que ocurran, es preciso mantener una mirada estratégica sobre sus sugerencias. No solo en lo que a precisión se refiere, sino sobre todo a la carga valórica que lleva consigo. La máquina no hace conjeturas equipada con una balanza moral sobre lo que está bien y está mal; eso le compete al individuo.

Volvamos entonces al algoritmo de Wuhan. Cuando la máquina de BlueDot cruzó los datos de infectados con coronavirus en China con los viajes en avión, pudo prever que varias ciudades de Asia resultarían afectadas de inmediato por los contagios. Esa predicción alertó a los responsables de la empresa (humanos),

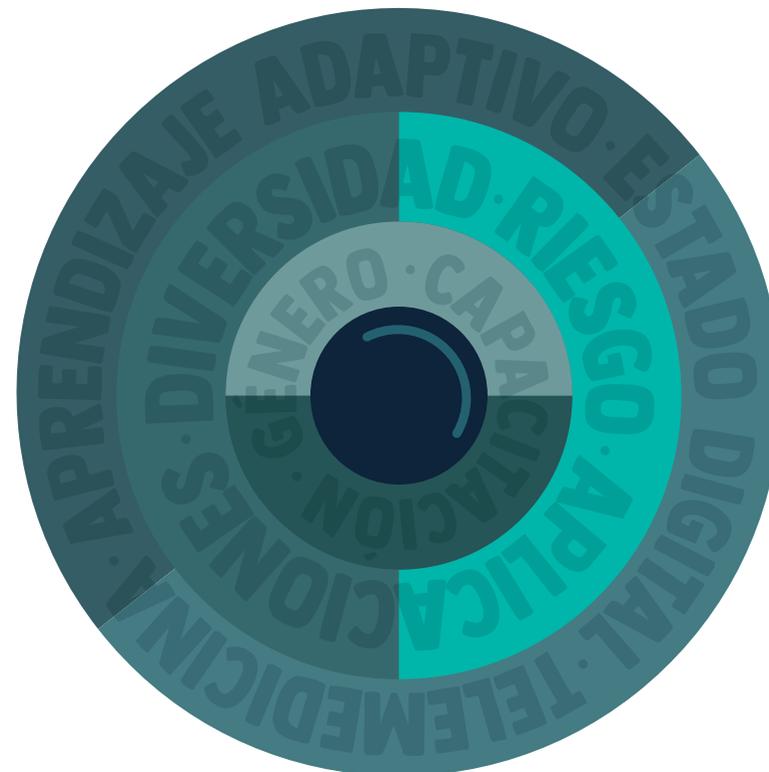
quienes transmitieron la información correspondiente a sus clientes en las oficinas sanitarias nacionales (también humanos), lo cual facilitó que organizaciones como la OMS y los CDC de Estados Unidos (donde también toman decisiones otros humanos) lanzaran una alerta global.

La asistencia de la IA fue capital durante la crisis pandémica y se torna cada vez más indispensable, pues por primera vez en décadas ofrece capacidades para planificar con precisión en el mediano y largo plazo, ensayando escenarios múltiples que permitan resolver del mejor modo posible nuestras rémoras y subdesarrollos históricos.

Pensemos en las aplicaciones de la IA para solucionar fallas estructurales de desarrollo en salud, educación y empleo en América Latina, o en infraestructura y calidad de vida. La IA es una gran promesa —o mejor una certeza— para mejorar la prestación de servicios sociales y asistir en el despliegue de políticas públicas. En democracias como las nuestras, donde la gestión pública necesita recuperar credibilidad frente a la ciudadanía, puede aumentar la transparencia en la toma de decisiones. En economías todavía endebles, favorecería la productividad.

En el siguiente capítulo intentaremos aproximarnos a la realidad latinoamericana a partir de un conjunto de preguntas. ¿Estamos técnicamente listos para ese salto? ¿Conocemos tanto su potencial como sus riesgos?

¿Sabremos hacerlo?



## CAPÍTULO 4

### ¿Listos para el despegue?

*América Latina llegó tarde y mal preparada a todas las revoluciones industriales. Esta es la primera vez que puede unirse a tiempo a la de la IA. Algunos proyectos en marcha.*

La pandemia reciente del coronavirus ha catapultado la transformación digital en todo el mundo. De un día para otro la tecnología se volvió un tema recurrente del que se hablaba en Twitter, en los mensajes de Facebook entre amigos, y en las charlas de oficina ahora remotas y por Zoom.<sup>54</sup>

Cada país tomó decisiones de mitigación sin un criterio completamente unificado, aunque con algunas similitudes: desde cierres totales como los de Wuhan, en China, hasta la implementación del distanciamiento social como en Corea de Sur o la estrategia de inmunidad de rebaño de Inglaterra y Suecia. Los informes sobre la efectividad de la política aplicada en cada país se hicieron, irónicamente, virales, porque por primera vez todos los gobiernos del mundo tenían un mismo objetivo de política pública: reducir el contagio y las muertes por coronavirus.

Cuando la ola del virus aún no cruzaba el Pacífico, los ciudadanos de ALC vimos en los noticieros cómo algunos países hacían intentos incipientes por controlar la pandemia: cerrar restaurantes, confinar a las familias, vigilar aeropuertos, frenar la economía, hacer rastreo de contactos. Simultáneamente éramos testigos del debut global, en horario estelar, de la inteligencia artificial, un tema de discusión hasta ahora muy mal expuesto por expertos en los medios de comunicación. Los servicios electrónicos, los sistemas GOFAI y el aprendizaje automático comenzaron a emplearse a una escala jamás vista para aliviar el impacto de la pandemia.

En muy pocas semanas se abrieron puertas que nadie esperaba. Aunque los *millennials* llevaban años buscando

<sup>54</sup> Los esfuerzos dirigidos a resolver los desafíos del coronavirus han tenido un impacto importante en la transformación digital. Satya Nadella, presidente de Microsoft, comentó durante la presentación de resultados del primer trimestre de 2020: “Hemos visto dos años de transformación digital en dos meses...” “...desde el trabajo en equipo y el aprendizaje a distancia, hasta las ventas y el servicio al cliente, pasando por la infraestructura y la seguridad críticas de la nube,” declaró refiriéndose a su empresa.

que sus empresas les permitieran trabajar de manera remota,<sup>55</sup> solo cuando ello se convirtió en una cuestión de vida o muerte la OMS lo patentó en una recomendación mundial. La profesión médica, que en ocasiones se había resistido a utilizar opciones virtuales, se vio en la necesidad de hacer una transición ultrarrápida de las consultas personales a las llamadas telefónicas y a las aplicaciones de telemedicina; esto en un mundo de pandemia con los hospitales enfocados en los casos de COVID-19 y con pacientes atemorizados de contagiarse en las salas de espera.

Un universo de servicios y empresas asomó ante un público que desconocía su existencia. JD Health, la subsidiaria de servicios sanitarios de la compañía de comercio electrónico JD.com, vio multiplicarse por 10 el uso de su plataforma desde el inicio de la pandemia. Hospitales de Estados Unidos como el Rush University Medical Center de Chicago y el George Washington University Hospital en Washington, D.C. se apoyaron en la telemedicina para atender las consultas de posibles pacientes enfermos recluidos en sus casas. Diversas compañías de atención de salud virtual de ese país salieron a contratar cientos de médicos ante una demanda de cuidados que elevó la presión sobre las capacidades de sus sistemas. DeepMind Health, una compañía de Google, se usó en China durante la crisis de Wuhan para notificar a los médicos sobre el agravamiento de la salud de los pacientes.<sup>56</sup>

Estas aplicaciones introdujeron cambios fundamentales en la interacción entre tecnología, trabajo, educación y relaciones sociales. En las páginas centrales de los periódicos, y en el horario estelar de las estaciones de televisión de Europa, comenzaron a aparecer historias sobre el uso de tecnologías

<sup>55</sup> Antes de la pandemia, Brasil era el país con más trabajadores remotos (12 millones), seguido de México (2,6 millones), Argentina (2 millones) y Chile (medio millón).

<sup>56</sup> DM Health monitorea el estado de una persona y lo compara en tiempo real con un conjunto incalculable de datos de síntomas semejantes.

en Asia. De allí surgieron debates sobre las implicaciones que podían tener los sistemas de rastreo implementados por China o Corea del Sur en las sociedades más liberales del continente. No hay periodista que no se haya sorprendido —y trasladado esa sorpresa a su público— con el uso de la IA para seguirle la pista al virus, decodificar su genoma o iniciar el trayecto hacia una vacuna. Todos parecimos entender —¡finalmente!— que este asunto de la inteligencia artificial podría ser útil.

No todos, claro. Bien sabemos que el desarrollo tiene ritmos distintos, grados diferentes, estadios variables y asimetrías profundas. Asia en mayor medida, algunos países de Europa y Oceanía, y Estados Unidos y Canadá se miraron de tú a tú con la nueva tecnología. Sin embargo, decenas de naciones de África, y también de América Latina, no se encontraban preparadas para descender el velo revelador de la IA.

### ¿Quiere Estado digital? Llene este formulario

El efecto que tendrá el coronavirus en la forma en que los gobiernos interactúan con los ciudadanos será importante y permanente. Ningún otro fenómeno ha sido tan extendido y generalizado a lo largo y ancho del planeta, y nunca hemos visto a Occidente intentar copiar, emular o adaptar procesos de Estados con otras experiencias políticas y culturales en sus propios sistemas públicos. Estamos participando todos, en simultánea, en una nueva transformación en la gestión de la cosa pública y en la relación de los individuos con el Estado de bienestar digital inteligente, EBDI.

Los sistemas de soporte de decisiones no van a aparecer milagrosamente; requieren de una infraestructura organizacional y jurídica sólida para capturar datos (muchos de ellos personales y privados) que permitan alimentar de forma responsable los modelos. Y eso exige un andamiaje sostenible: una política de Estado. El salto a la IA requiere un Estado digitalizado y nos encuentra con uno que apenas si está parcialmente incorporado al siglo XXI. ¿Cuál es el problema principal? La lenta modernización de las administraciones públicas. En ALC hay un rezago estructural

en las capacidades estatales para integrar nuevas tecnologías y poca o nula regulación para la protección de los datos personales.

A las restricciones financieras de cada nación se suma una concepción anquilosada del modelo de Estado. La crisis derivada de la pandemia introducirá un nuevo escollo financiero en la carrera hacia el futuro. De eso no hay duda.

El retraso se expresa con cifras contantes y sonantes. Según algunos análisis, en 2018 solo cinco de los 19 países de América Latina figuraban entre los 50 más digitalizados del mundo. Uruguay<sup>57</sup> lideraba a la región en ese rubro, y ese año se unió a México como la segunda nación latinoamericana en aparecer en el Digital 9, un grupo que reúne a los gobiernos más avanzados en digitalización.

Los Estados latinoamericanos deben deshacerse de su excesiva burocracia y tramitomanía si desean ser eficientes en escenarios donde las trabas se pagan caro. Según el documento “El fin del trámite eterno” del Banco Interamericano de Desarrollo, en América Latina solo el 7% de las personas reportaron haber hecho un trámite en línea. Muchas gestiones son todavía personales y toman en promedio cinco horas, cuando podrían hacerse de una sola vez y en minutos en internet. Eso le cuesta al gobierno hasta 40 veces más de lo que podría costar el servicio equivalente en una plataforma digital. Todo lo anterior favorece la corrupción. Según el mismo estudio del BID, un tercio de los latinoamericanos reconoce haber pagado algún tipo de soborno para poder completar un trámite público.

Y aquí no terminan las malas noticias. Cuando se registran rezagos en la administración central del Estado, estos tienden a reproducirse en áreas específicas de la gestión. La escasa digitalización del sector salud latinoamericano, por ejemplo,

<sup>57</sup> Uruguay creó una agencia digital con rango ministerial para modernizar la gestión pública, una decisión que responde a la concepción de la modernización del Estado como estratégica y no herramental

se ha hecho especialmente visible durante la pandemia.<sup>58</sup> La consultora en investigación de mercados de salud, Global Health Intelligence, reportó en 2017 que solo el 40% de los hospitales contaban con registros médicos electrónicos (RME).<sup>59</sup>

Así nos encontró el coronavirus cuando acabó de cruzar los océanos para asentarse en la región. La falta de infraestructura digital y de procesos de interoperabilidad entre los hospitales se hizo patente, además de que evidenció inconsistencias y demoras importantes en la creación de registros de casos entre las administraciones subnacionales y los gobiernos centrales. Información tan básica como el número de defunciones y contagios no fluyó como debería hacerlo; sin saber cómo se propagaba el virus, las políticas de salud y control de la epidemia estaban destinadas al fracaso.

Tanto la urgencia de contar con soluciones tecnológicas, como la poca preparación y flexibilidad de los Estados de la región, quedaron también en evidencia. En suma, ALC no parece estar

<sup>58</sup> Si eso sucedía con los sistemas de salud, el estado de los sistemas educativos ya era preocupante incluso antes del COVID-19. De acuerdo con los resultados de las pruebas PISA 2018, “en promedio, un estudiante de 15 años en América Latina muestra tres años de rezago en lectura, matemáticas y ciencia comparado con un estudiante en algún país de la OCDE”. En un informe de la UNESCO de mediados de 2020 se indicaba que el cierre de los centros educativos durante la pandemia implicaría que 165 millones de estudiantes en 25 países de la región no tendrían acceso a la escuela desde preescolar hasta la educación terciaria.

<sup>59</sup> Tal vez el único punto favorable es que un año antes, esos RME estaban en 36% de los hospitales y se esperaba un crecimiento del 7% para cada año subsiguiente hasta 2019, hasta alcanzar a casi la mitad de los hospitales. Dependiendo de los países, el escenario varía a favor o en contra. Según GHI, Brasil fijó plazos para dotar a la mayoría de sus hospitales con RME para 2020 (no tenemos reportes sobre el estado actual), y ya en 2016 habría adquirido tres supercomputadoras para unificar todos los registros gubernamentales. Chile lidera a la región con un 74% de los registros médicos digitalizados, seguido por Uruguay y Costa Rica con 63% y 61% respectivamente. Guatemala, Panamá y Colombia tenían registros electrónicos en al menos la mitad de sus hospitales. Pero por ejemplo Argentina apenas si llegaba al 17%, mientras que México oscilaba en un rango amplio e indeterminado de entre 25 y 50%.

preparada para ofrecer el gobierno digital moderno y efectivo que sus ciudadanos requieren en un siglo interconectado.<sup>60</sup>

Ahora bien, el panorama económico que nos aguarda después de la pandemia produce temores bien fundados. Se estima que las caídas generalizadas del PIB a escala mundial afectarán más a los países en desarrollo. A mediados de 2020, cuando la primera ola del coronavirus aún no había concluido, se esperaba que América Latina liderara la retracción global con una reducción de su PIB regional superior al 7%.<sup>61</sup> Aunque parecen cifras lanzadas al aire, la economía es en realidad una ciencia que las personas sufren en el cuerpo. Las familias de los trabajadores informales, sin contratos y/o prestaciones, han sido las más vulnerables a los efectos de choques exógenos como el coronavirus, pues carecen de acceso a protección laboral, seguros de salud y vías de entrada a las redes de protección social. Según la Organización Internacional del Trabajo (OIT), en 2018 había 130 millones de trabajadores latinoamericanos en la informalidad, equivalentes al 56%<sup>62</sup> del mercado laboral. A mediados de 2020, la situación indicaba que la crisis llevaría esa participación a un 62%.

<sup>60</sup> Con esto no queremos decir que la sola digitalización de los servicios públicos sea la respuesta a todos los problemas, especialmente en una región en donde solo el 44% de los hogares tiene acceso a internet. Es necesario pensar en estrategias que permitan atender a las poblaciones en situación de mayor marginación.

<sup>61</sup> En un informe del Banco Mundial de junio de 2020 se preveía que el este asiático y el Pacífico crecerían un 0,5%, una rareza en relación con el resto del mundo. El sur de Asia se contraería 2,7%, África subsahariana el 2,8%, Oriente Medio y el norte de África un 4,2%, y Europa y Asia central un 4,7%. Entre tanto, América Latina retrocedería un 7,2%. Las recesiones previstas causarán un retroceso de años de avances y devolverán a decenas de millones de personas a la pobreza extrema, en especial en las naciones emergentes.

<sup>62</sup> Hay países donde la informalidad era incluso mayor antes de la crisis originada en la pandemia. Por ejemplo en Honduras, Nicaragua y Guatemala ascendía al 84, 77 y 76% de la PEA respectivamente, mientras que en Bolivia la cifra era del 71% de la fuerza laboral hacia 2017.

Es evidente que los esfuerzos gubernamentales contra la pandemia como son los confinamientos estrictos —a todas luces necesarios— tuvieron un efecto nocivo al paralizar la actividad económica. Y si bien se pidió a la gente que se quedara en casa, durante el aislamiento el Estado no fue capaz de dar los pasos necesarios para controlar una crisis sanitaria y tampoco para planear el futuro inmediato y mediano.

Es ahí precisamente donde está el núcleo del problema. Si la función del Estado es contribuir con sus políticas a que las personas tengan vidas más dignas, ¿qué pueden ofrecer estos gobiernos a sus ciudadanos en una era con nuevas exigencias, comparadas con las del siglo XX, cuando todavía ni siquiera han respondido a las necesidades típicas de la centuria previa? Y como si fuera poco, en este veloz siglo XXI están abocados a diseñar una nueva estructura de competitividad para sus economías nacionales, aunque con cosmovisiones e infraestructuras claramente obsoletas.

Pero no hay que perder la esperanza, pues nos sobra experiencia cuando se trata de afrontar situaciones difíciles. La pandemia del COVID-19 no es ni será la última emergencia que afectará a ALC. Los Estados pueden y deben continuar su modernización apuntando a ser flexibles, escalables e inteligentes. Y deben aprovechar la IA de forma responsable. Los datos de las personas, y la información que de allí surja, deben considerarse como un activo estratégico de la sociedad y su aparato público.

Lo mejor es que ya hay por dónde comenzar a aprender.

### **El estado de nuestra tecnología**

Duele decirlo, pues nos exige aceptar cuán lejos estamos de lo deseable, pero lo cierto es que los funcionarios públicos en la región viven igual que sus vendedores ambulantes: al día. Pasan la mayor parte de su tiempo apagando los incendios de hoy, sin tiempo y/o recursos para abordar las crisis que se avecinan.

Las grandes empresas avanzan sobre caminos muy bien pavimentados por las corporaciones globales. Mientras que la experimentación y el riesgo caracterizan el territorio de los emprendedores, más allá el escenario luce sombrío. En la mayoría de nuestras naciones, los mercados de capitales son prácticamente inexistentes, las leyes para fomento de IA son apenas ilusiones y la formación de recursos humanos es en extremo incipiente.

Antes de que el coronavirus hiciera su aparición, la división Accenture Frontier Economics había estimado que si los países con las economías más grandes de la región adoptaban de manera decidida la IA, podrían crecer hasta un punto porcentual más para 2035. En ese escenario hipotético, Chile y Colombia registrarían expansiones del PIB del 4,5%, Perú y Brasil de cerca del 4,2% y Argentina de un 3,6%. En valores constantes, la IA aportaría a Chile US\$63.000 millones extras; US\$78.000 millones adicionales a Colombia; US\$43.000 millones a Perú; US\$432.000 millones a Brasil, y US\$59.000 millones a Argentina. Se preveía igualmente que los rubros más beneficiados serían las finanzas, los productos básicos de exportación, el comercio mayorista y los servicios públicos. Dado que falta más de una década para 2035, es inevitable que varias de estas expectativas sean radicalmente reconsideradas tras la devastadora recesión económica en que nos está sumiendo la pandemia del COVID-19, una retracción solo comparable con la Gran Depresión de 1929.

Pero mientras esas rectificaciones suceden y los Estados toman alguna decisión, en América Latina ya hay procesos de IA en curso y bien amarrados. ¿Cómo así? Pues porque la sed competitiva es un motor determinante en el sector privado, que hoy ensaya opciones en torno al aprendizaje automático sin esperar a que los gobiernos fortalezcan las condiciones ideales de competitividad.

Un estudio de la organización de promoción de emprendedores Endeavor, en colaboración con la firma consultora de negocios Everis, reveló a mediados de 2018 que más de 70 empresas de Argentina, Brasil, México, Colombia, Chile y Perú tenían

ya proyectos funcionales de IA. No obstante sus dificultades económicas, Argentina cuenta con las empresas emergentes más grandes y dinámicas de la región, aunque curiosamente es uno de los Estados menos digitalizados de ALC, siendo el tercero por tamaño relativo.<sup>63</sup> Las más adelantadas en el uso de la IA son las firmas de software, seguidas por las del mercado de cuidados de salud, las de medios de comunicación, educación, minería y transporte, incluyendo las aerolíneas.<sup>64,65</sup>

Según el sondeo de Everis, la mayoría de las empresas de IA en ALC son jóvenes; se fundaron a mediados de la década pasada. La mitad no tiene más de 10 empleados y trabaja con predicción, clasificación y procesamiento de lenguaje natural como herramientas. En 2018, seis de cada diez emprendimientos estaban enfocados en IA supervisada. Un tercio empleaba reconocimiento de patrones, comprensión lingüística y reconocimiento de imágenes.

La composición de la inversión habla del potencial de solvencia futura y el estado de madurez de los proyectos de IA. En América Latina, solo el 60% estas empresas<sup>66</sup> recibió recursos de una o varias fuentes, en su mayoría fondos de inversión.<sup>67</sup> En Perú y Argentina, la mitad de las operaciones no recibió

<sup>63</sup> En 2017, esas empresas generaron ingresos por US\$3,3 millones y esperaban llegar a US\$3,7 millones para el cierre de 2018. Esto las ubicaría por encima de Brasil, donde obtuvieron US\$1,75 millones; la proyección para 2018 es de US\$2,28 millones. Perú pasaría de US\$0,79 millones en 2016 a US\$2,79 millones para finales de 2020. Las compañías argentinas de IA tienen unos 76 empleados, más que las de Brasil (41) y Perú (35).

<sup>64</sup> El 40% son compañías de servicios y análisis de negocios (muchas proveen asistentes virtuales), seguidas por un 17% de proyectos enfocados en salud. Todas las demás están por debajo del 7%.

<sup>65</sup> Las coyunturas disruptivas contribuyen a acelerar los cambios. El COVID-19, por ejemplo, llevó a la compañía colombiana de entrega de comidas a domicilio Rappi a lanzar en Medellín un piloto de entregas en asociación con la empresa de robots KiwiBot. Durante el piloto, 15 robots entregaron 120 pedidos diarios. Una vez que concluya el experimento, Rappi analizará el potencial de automatización y robotización de sus entregas a domicilio.

fondos, circunstancia que indica probables constreñimientos de corto o mediano plazo una vez agotado el capital original; en México, la mitad obtuvo capital privado por más de US\$10 millones y dos de cada diez fondos recibieron hasta US\$5 millones. En Brasil, la mitad recibió hasta US\$5 millones a cambio de acciones.

En cuanto al tipo de inversión, una empresa suele comenzar con un padrino inversionista (que aporta entre US\$10.000 y US\$50.000) o con capital semilla (entre US\$100.000 y US\$1 millón). Esto le permite empezar a desarrollar su idea y su modelo de negocio. El 50% de las empresas sondeadas en el estudio de Everis que recibieron inversión se encontraba en este primer estado de madurez. Esto nos muestra una industria en crecimiento, pero que aún debe solidificarse.

El sector sanitario es, quizás, el que mayor inversión ha recibido, por las mismas razones que en el resto del mundo: la información más precisa que producimos los seres humanos es sanitaria. Podemos mentir al fisco, podemos ocultar deudas a un banco y hasta a nuestras familias, pero al médico tendemos a decirle casi siempre la verdad. Nunca es más apropiada la idea de que se nos va la vida en ello. La abrumadora información sanitaria es un vergel para los proyectos de IA. El coronavirus nos mostró cuán vulnerables somos cuando nuestras vidas están en juego y cuán útiles resultan los avances tecnológicos para lidiar con crisis de salud letales.<sup>68</sup> Y ahí hay un negocio monumental.

<sup>66</sup> Los datos presentados a continuación se obtuvieron del documento “Impacto de la Inteligencia Artificial en el emprendimiento” de 2018.

<sup>67</sup> Entre ellos Angels Nest, Biz Rupt, Capital Partners, Dalus Capital, Dila Capital, Endeavor Catalyst, Highland Capital Partners, Nxtp Labs, Mountain Nazca, On Ventures, The Venture City y Wayra.

<sup>68</sup> Antes del coronavirus ya existían en la región proyectos de salud que apuntaban a resolver problemas críticos. Por ejemplo, emprendedores de México, Chile y Alemania crearon la empresa emergente de análisis molecular Miroculus, la cual se centró, primero en detectar cáncer, para luego descubrir que la IA le permitía ampliar sus diagnósticos a otras enfermedades.

## La IA en el sector salud

A inicios de 2020, en ALC se empleaban algunas tecnologías inteligentes en el diagnóstico de enfermedades y en la prevención y control de epidemias<sup>69</sup> e incluso en la reducción de la estadia de pacientes en los hospitales, esto último con el fin de optimizar los siempre limitados gastos de la salud pública. También se encontraban en operación múltiples soluciones de diagnóstico mediante el análisis de síntomas repetitivos que llevan a un resultado común,<sup>70</sup> y otras para la administración eficiente de turnos en laboratorios y consultorios.

De hecho, la telemedicina parece haber tenido un debut auspicioso, de modo que no entraremos tarde a este campo. Hay proyectos en países como Colombia (IDOC3), y también en Brasil y México, con operaciones mucho más amplias (Telemedicina y Doc.com respectivamente). La buena noticia es que el modelo parece exitoso y replicable, y es difícil que retroceda. Su eficiencia ha sido probada durante una crisis global y es probable que en pocos años hablemos de él como un estándar.

Uno de los ejemplos latinoamericanos en el rubro salud es el de la chilena Teledx. En 2013, la compañía había desarrollado un prototipo para la detección temprana de la retinopatía

<sup>69</sup> Con base en el uso de redes neuronales, un proyecto piloto liderado por la Universidad Nacional de Córdoba, en la provincia de Salta, permitió predecir la presencia de huevos de mosquitos transmisores de dengue con mayor precisión que con trabajo de campo. Esta iniciativa duró cuatro años, de 2012 a 2016.

<sup>70</sup> Estos son modelos de clasificación supervisada (alimentados con una base de imágenes con diagnóstico positivos para cáncer) que generan información predictiva. Índigo, una compañía colombiana patrocinada por Microsoft, desarrolló un modelo de IA para identificar 42 tipos de mutaciones de cáncer de pulmón, y para segmentar nódulos y tumores usando imágenes de tomografías. Según la compañía, su programa conocido como Indira, ha permitido diagnosticar más de 300 casos de cáncer de pulmón con un 99,3% de confiabilidad.

diabética (Dart), lo cual condujo a que uno de sus creadores fuera incorporado a la lista de jóvenes innovadores del *MIT Technology Review*. Desde entonces, Teledx no ha dejado de expandirse.

Muchos diabéticos corren el riesgo de desarrollar retinopatía diabética, una enfermedad que daña los vasos sanguíneos ubicados en la parte posterior del ojo. Los casos más avanzados pueden desembocar en ceguera parcial o total, por lo que requieren cirugía; de ahí la importancia de la detección temprana. Dart/Teledx pretende anticipar el surgimiento de la enfermedad. “Comenzamos en 2011 cuando no se hablaba de IA y menos de IA en salud”, dice José Tomás Arenas, presidente de Teledx. “En ese tiempo nos miraban raro porque ni siquiera había internet en todos los lugares.”

Según estimaciones del Ministerio de Salud de Chile, aproximadamente 1,5 millones de personas padecen de diabetes, lo cual equivale al 12% de la población adulta. Estas personas deben hacerse un examen anual en el que se evalúa no solo la evolución de la enfermedad sino también los riesgos y complicaciones secundarias como la retinopatía.

Sin embargo, existe un desbalance entre oferta y demanda de atención de salud basada en sistemas digitales eficientes, particularmente en el sector público. Teledx-Dart vio la oportunidad de posicionar una solución para abordar una patología específica con alto impacto en la calidad de vida de las personas. En 2014 comenzó a crear una red de clientes en zonas donde se concentra el grueso de la población chilena (la Región Metropolitana de Santiago, Valparaíso, Biobío y Maule), apoyándose en los servicios públicos de salud.

El modelo, desarrollado por el doctor Arenas y su equipo, puede procesar fotografías de la retina (retinografías) para el tamizaje de retinopatía diabética. Estas alertas son posteriormente verificadas y evaluadas por un médico para iniciar un tratamiento, priorizando así la atención y reduciendo el estrés sobre los recursos limitados del sistema que irían a cirugías o a tratamientos láser.

A través del Ministerio de Salud y la red de atención pública, Teledx-Dart se transformó en estándar nacional, atendiendo a un total de 300.000 personas al año, un volumen significativo para Chile. El éxito animó a Teledx a realizar pilotos y pruebas en varios países de América Latina, centrándose especialmente en México, Perú y Brasil.

Los procesos de diagnóstico como Dart son de uso generalizado en el mundo desarrollado. Hay motores de IA capaces de predecir la posibilidad de un cáncer a partir del análisis de imágenes. También se emplean microscopios con tecnología AI para buscar bacterias dañinas como E. coli y estafilococos en muestras de sangre a una mayor velocidad que el escaneo manual.<sup>71</sup> Un grupo de investigadores utilizó las mamografías de unas 90.000 mujeres de Gran Bretaña y Estados Unidos, cuyos diagnósticos ya se conocían, para entrenar computadoras en el reconocimiento de cáncer de mama.<sup>72</sup>

### Andando despacio se llega lejos

La necesidad es la madre de la innovación. Al igual que en el caso de Dart/Teledx en Chile, la necesidad de soluciones para problemas estructurales de política pública también encontró en Brasil una respuesta en Jacson Fressatto. Autodenominado “hacker activista”, Fressatto está detrás del robot cognitivo brasileño que identifica de manera anticipada casos de sepsis —una respuesta inflamatoria sistémica a una infección que puede llevar al paciente a la muerte— y notifica a los médicos de inmediato.

<sup>71</sup> Los científicos utilizaron 25.000 imágenes de muestras de sangre para enseñar a las máquinas cómo buscar bacterias. Las máquinas aprendieron entonces a identificar y predecir las bacterias nocivas en la sangre con una precisión del 95%.

<sup>72</sup> Luego, los investigadores presentaron 500 mamografías a varios radiólogos de Estados Unidos para que las interpretaran. El sistema de IA produjo menos falsos negativos y falsos positivos que los especialistas.

El robot se llama Laura-br porque Laura era el nombre de la hija de Fressatto, una bebé que nació prematura y fue ingresada a un hospital de Curitiba en 2010. La condición clínica de la bebé empeoró de forma repentina sin que el equipo médico lo notara hasta que la situación se tornó irreversible. Laura falleció por una sepsis prevenible cuando apenas tenía 18 días de nacida. Fressatto, buscando entender lo que había sucedido, descubrió que las infecciones hospitalarias representan el 56% de las muertes en estos centros de atención en Brasil, donde cada año fallecen por esa causa cerca de 250.000 niños, adultos y ancianos.

Esto lo llevó a fundar Laura-br, una compañía que desde 2014 opera en su sede de Curitiba, y que desde 2016 ha logrado reunir y analizar los datos de dos millones de personas. Según Fressatto, a inicios de 2020 Laura-br ayudaba a salvar diariamente 12 vidas en los hospitales públicos durante la atención inicial de valoración (*Triage*) y diagnóstico de pacientes con síntomas de COVID-19, enfermedad que ha afectado gravemente a Brasil.

La historia de Laura-br tiene como antecedente el uso de tecnología para resolver problemas comportamentales. Fressatto es especialista en seguridad corporativa. En su calidad de analista de sistemas por formación, y de científico de datos autodidacta, usaba algoritmos para identificar fraudes en sistemas y procesos, y trasladó ese aprendizaje al tratamiento de datos en tiempo real empleando un robot de IA. El año era 2011. “En esa época usar AI era una odisea,” afirma.

Laura-br emplea datos históricos para entrenar algoritmos de detección de anomalías. El sistema genera un puntaje o calificación de riesgo para cada paciente del hospital y asigna un orden para reducir el tiempo de atención de los médicos. El robot fue mejorado progresivamente y ya no solo observa los peligros de sepsis sino también el deterioro del paciente, desde su condición cardíaca hasta la renal.

La educación es otro ámbito que requiere atención en ALC. Millones de estudiantes podrían beneficiarse de tecnologías de aprendizaje adaptativo que modifican el contenido de la enseñanza

en función de las habilidades de cada uno de ellos.<sup>73</sup> Existen allí varios proyectos con aproximaciones interesantes al aprendizaje personalizado. Para mediados de 2019, la empresa colombiana BabySparks, por ejemplo, había logrado que su aplicación para niños fuera descargada más de 3,5 millones de veces en 180 países. Entre tanto, la firma chilena U-Planner trabaja con universidades e instituciones técnicas para ayudarles a planificar y mejorar la gestión: desde asignar horarios, actividades, salas y profesores, hasta manejar mallas curriculares. Lo mismo en lo que se refiere al seguimiento de la calidad educativa, generando alertas tempranas de abandono escolar, midiendo el aprendizaje de los estudiantes, o evaluando la gestión y calidad de los docentes. Sus algoritmos reducen a minutos la ejecución de tareas administrativas que toman horas cuando se realizan de forma manual o con la informática tradicional.<sup>74</sup>

Emprendimientos como el mexicano Territorium capacitan en línea y de manera presencial a estudiantes que buscan reeducarse para el mercado laboral.<sup>75</sup> Territorium trabaja con 8,5 millones de usuarios y tiene una alianza global con Dell y Microsoft. Sus cursos se ajustan empleando motores de IA alimentados con entre 300 y 1.000 datos por alumno. “No

<sup>73</sup> Según un análisis del BID, los sistemas de inteligencia artificial y el estudio de los patrones de los alumnos ayudan a crear perfiles pedagógicos para automatizar la generación de contenidos adaptados a cada persona. La individualización permite crear ejemplos y prácticas que faciliten al estudiante el refuerzo de conocimientos. También ayuda a crear simulaciones realistas de casos prácticos para que los alumnos interactúen tomando decisiones relacionadas con problemas reales. De ese modo, aprenden de sus errores en ambientes controlados.

<sup>74</sup> “En el contexto universitario actual, este tipo de herramientas resulta de gran importancia para los procesos de *assessment*, mejora continua y acreditación, de los cuales depende a su vez el posicionamiento de cada plantel en el mercado universitario de pre y posgrado”, según Juan Pablo Mena, presidente de la firma U-Planner, en una entrevista para este libro. “Los problemas que vimos en América Latina y validamos en Australia o Abu Dabi son globales. Los desajustes de la demanda educativa y la oferta escolar, la relación entre el costo y su calidad, los desafíos de optimización...Se trata de un mismo problema que cambia de nombre.”

es tanto lo que se aprenda sino asegurar que la persona sea emprendedora o pueda tener un empleo,” dice Carlos Elizondo Ancer, su presidente. “Vemos qué busca el mercado y qué ofrece el sector educativo, luego trabajamos con cada empresa para detectar la brecha y recomendamos una experiencia de aprendizaje al estudiante para reducirla.”

En el mercado laboral, uno de los principales impactos de la digitalización ocurre en la búsqueda y oferta de empleo. Las plataformas de vacantes han permitido poner a disposición de los trabajadores una cartera mucho más amplia de empleos. Pero también han generado nuevos problemas. Por ejemplo, un área de recursos humanos puede recibir hoy miles de solicitudes, pero carece de la capacidad humana para evaluarlas todas de forma eficiente. Los empleadores suelen decir que la razón principal por la que no logran cubrir sus vacantes es la falta de habilidades. Al mismo tiempo, sin embargo, en redes profesionales como LinkedIn abundan los trabajadores sobrecalificados para las tareas que desarrollan en sus puestos actuales.

La IA puede hacer que la intermediación laboral sea más eficiente, justa e inclusiva. Uno de los primeros proyectos latinoamericanos de política pública de empleo basado en IA está en Paraguay. ParaEmpleo es una plataforma digital de intermediación lanzada en 2018 por el gobierno paraguayo que incluye un sistema de vinculación entre vacantes y postulantes. Esta analiza habilidades, capacidades y competencias para producir el mejor acoplamiento entre la oferta y la demanda laboral.

<sup>75</sup> La chilena GameLab Education es otro ejemplo enfocado en capacitar a quienes trabajan en empresas o a las propias instituciones de formación, entre ellas las universidades. GameLab realiza simulaciones en varios campos de gestión empresarial a través de juegos virtuales como SodaPop Game (gestión de operaciones), Warehouse Game (manejo de logística) o el Balanced Scorecard Game de estrategia. Más de 18.300 estudiantes han participado de los juegos de GameLab, que tiene entre sus clientes a organizaciones como Arizona State University (la universidad pública más grande de Estados Unidos), New York University, Texas A&M y una veintena de universidades latinoamericanas.

Si el empleo, la salud y la educación son áreas complejas que requieren miradas innovadoras desde la política pública, hay otras que atraviesan a todos los sectores de la sociedad: las brechas de género y diversidad. ALC tiene rezagos marcados en lo que se refiere a la incorporación de grandes y pequeñas minorías al desarrollo. Tomemos, por ejemplo, a las personas con discapacidades auditivas, que suman más de 1,5 millones de individuos solamente en México, Colombia y Argentina. Un proyecto interesante, la Red-Apis, ha desarrollado TransVoz, una herramienta que usa IA para permitir a personas con discapacidad auditiva o movilidad reducida a transcribir en tiempo real lo que se conversa en torno a ellas.

Cada uno de estos casos expone el estado de la IA en la región. Esto significa que esta ha comenzado a adentrarse en la nueva revolución, y quizás este conocimiento e iniciativa logren que por primera vez no lleguemos tarde a un cambio estructural. Pero esa participación también refleja las condiciones complejas y el punto de partida de enfático rezago frente a otras áreas y naciones del mundo; hasta ahora, el avance que se está logrando depende más de la innovación y riesgo privados que de la planificación del Estado. Esto debe cambiar.



## CAPÍTULO 5

### Alexa: ¿Este sesgo es mío?

*Cómo controlar nuestros sesgos para ofrecer soluciones inteligentes.*

En octubre de 2019, la prestigiosa revista *Science* publicó una investigación de cuatro académicos de Berkeley, Boston y Chicago, quienes advertían sobre un serio problema en el mercado de seguros de salud de Estados Unidos. Ziad Obermeyer, Brian Powers, Christine Vogeli y Sendhil Mullainathan dedicaron meses a revisar 50.000 historiales de un hospital universitario de renombre y encontraron que los pacientes blancos recibían calificaciones de riesgo de salud más altas que la población afroamericana, lo cual mejoraba sus probabilidades de ser seleccionados para cuidados adicionales como citas específicas o enfermería personalizada. Según las estimaciones de los investigadores, esa práctica privaba a más de la mitad de los pacientes afroamericanos de una atención con mejores cuidados.

La razón de esa discriminación era un algoritmo. Según un informe del MIT, no es que el sistema de IA empleado por el hospital hubiera sido programado de forma explícita para tomar decisiones injustas —de por sí ilegal—, sino que el algoritmo corría sobre historiales médicos que tenían un sesgo implícito, el cual se estaba transmitiendo al sistema.

El objetivo del sistema de soporte de decisiones era indicar de forma preventiva cuáles eran los pacientes que tenían una probabilidad elevada de requerir mayores cuidados médicos. El problema surgió cuando los desarrolladores utilizaron como variable a predecir el gasto en que aquellos incurrieron durante su enfermedad, bajo la hipótesis de que las personas más enfermas terminarían gastando más en tratamientos médicos para superar la enfermedad. ¿Qué podía salir mal?

Por razones históricas variadas, incluidas las de carácter socioeconómico, los pacientes afroamericanos gastaban en promedio menos en atención de salud que la población blanca, así tuvieran las mismas dolencias u otras más graves. “Como resultado, el algoritmo dio a los pacientes blancos los mismos puntajes que a los pacientes afroamericanos que estaban significativamente más enfermos.” El gasto histórico en salud decía que estaban peor. ¿Y cómo no creerle al dinero que uno gasta en uno mismo?

El costo no es una métrica ciega. Según *The Washington Post*, el software investigado pertenecía a United Health Group, la mayor aseguradora del mercado de salud de Estados Unidos, una compañía que en 2019 estaba avaluada en US\$235.000 millones, tres veces más que su competidora inmediata. El software se empleaba con 70 millones de usuarios y no era el único. Los investigadores detectaron la misma falencia lógica en los 10 algoritmos más usados por la industria, un grupo de herramientas que en su conjunto procesaba las solicitudes de atención de no menos de 150 millones de personas cada año.

La compañía propietaria del software trabajó con los cuatro investigadores para corregir los prejuicios funcionales del algoritmo, reduciendo la disparidad en un 80% e incorporando candados y mejoras para evitar la reiteración de sesgos. “Este estudio es el último en mostrar los peligros de asignar recursos siguiendo la recomendación de algoritmos,” según *MIT Technology Review*. “Este tipo de desafíos no solo se presenta en el cuidado de la salud, sino también en contrataciones, calificaciones crediticias, pólizas de seguros y en la justicia penal.”

Una verdad irrefutable sobre la inteligencia artificial es que el algoritmo nunca es mejor que los datos que lo alimentan y/o que las decisiones de quienes lo entrenan. Y quizás debiéramos decir que la IA no puede ser muy distinta a sus creadores: en los algoritmos estamos nosotros. Si los modelos de IA presentan sesgos es porque sus diseñadores han plasmado allí sus construcciones culturales, prejuicios e inclinaciones ideológicas. En definitiva, la arbitrariedad de nuestras elecciones se refleja en la calidad de la solución.

Los gurús del apocalipsis tecnológico suelen lanzar advertencias sobre los riesgos de las máquinas para el futuro de la humanidad. Hay una fantasía más o menos extendida de que su final será producto de nuestros propios inventos. Máquinas, androides o entidades híbridas que resultan mejores que nosotros mismos son los personajes que protagonizan esa especie de proyección de nuestro propio antropocentrismo.

En realidad, cuando los medios hablan de los riesgos de la IA no deja de haber un barniz de ciencia ficción: una creencia infundada en el supuesto poder omnímodo de una máquina capaz de desarrollarse con completa independencia de sus creadores. Y es curioso: a esa inteligencia artificial se le supone maldad. Los jinetes del apocalipsis creen que la IA podría llegar a destruirnos porque somos un riesgo para la sostenibilidad del planeta.

Pero, como ya hemos dicho, los motores de las máquinas carecen de moral. Si tienen algo parecido es porque nosotros —sus creadores, entrenadores y programadores humanos— hemos transferido nuestros sesgos a sus algoritmos.

El asunto es sencillo. Si queremos evitar un problema, dice la lógica, es mejor preverlo y crear soluciones que lo eviten o, dado el caso, que lo resuelvan.

### ¿Predecir por predecir?

En 2017 en Salta, una provincia del norte argentino, un grupo de funcionarios decidió abordar un problema social preocupante con una aproximación distinta. Tratarían de prevenir los embarazos adolescentes empleando inteligencia artificial.

El embarazo temprano es un problema serio de salud pública.<sup>76</sup> Según la OMS, para 2018 unos 21 millones de niñas entre los 16 y los 19 años, y otros dos millones de 15 años o menos, habían quedado embarazadas en las zonas económicamente menos favorecidas del planeta.

Las menores entre 15 y 19 años que deciden poner fin a sus embarazos se someten por lo general a abortos peligrosos (cerca de 3 millones de ellos al año), mientras que las complicaciones durante la gestación y el parto son la principal causa de muerte en el mundo para las niñas entre los 10 y 19 años. Y si finalmente dan a luz, la maternidad tiene repercusiones sociales y económicas de largo plazo en sus vidas, ya que esta se correlaciona con el abandono de la educación formal y, por esa vía, con la reducción de oportunidades laborales futuras.<sup>77</sup>

El proyecto en Salta requería que el Ministerio de la Primera Infancia, institución a cargo, recopilara los datos relevantes de sus encuestas de población. Esa información iría a una empresa tecnológica que suministraría la infraestructura para albergar los datos; allí se procesaría el modelado.

Utilizando un algoritmo de aprendizaje automático supervisado, el sistema de IA sería alimentado con observaciones sobre las niñas que habían tenido un embarazo adolescente, y con información

<sup>76</sup> Los embarazos adolescentes se presentan en todas las naciones, independientemente de su nivel de ingreso. Sin embargo, las tasas más significativas se registran en las regiones en desarrollo. La tasa de partos adolescentes se ha reducido entre 1990 y 2015 de 65 a 47 nacimientos por cada mil mujeres, y esa es una buena noticia. Sin embargo, cuando los expertos de la OMS se reunieron en 2018 constataron que las proyecciones para 2030 mostraban un mayor número de adolescentes embarazadas, pues la población en ese grupo etario es un segmento en crecimiento. Excluido el norte del continente, África experimentaría el mayor incremento, seguida por América Latina. En la actualidad, en los 20 países continentales de la región y en las islas del Caribe da a luz un promedio de 64 menores por cada mil mujeres. Si las regiones pobres son las más proclives, las poblaciones rurales e indígenas llevan la peor parte, dado que allí es posible que una niña quede embarazada en una proporción de 3:1 con respecto a otra que vive en una ciudad mediana. “En todo el mundo es más probable que los embarazos adolescentes ocurran en comunidades marginadas, comúnmente propiciadas por la pobreza, y por la falta de educación y oportunidades de empleo”, según el informe *“Adolescent Pregnancy”* de la OMS.

<sup>77</sup> Según el informe Juventud de Unicef de 2017, en Argentina medio millón de adolescentes han abandonado la escuela en los 16 primeros años de un siglo donde la formación continua dividirá las aguas entre un buen futuro y uno difícil. Esa deserción escolar tiene una elevada correlación con el ingreso prematuro al mercado laboral y —ninguna sorpresa— con el embarazo temprano. Según información extraoficial, el 15% de los nacimientos en Argentina son de madres adolescentes. La tasa se ha mantenido constante por los últimos 25 años, pero no por ello deja de ser preocupante. En Argentina nacen 3.000 niños de madres que ni siquiera han celebrado su fiesta de 15 años. La mayoría (siete de cada diez, según Unicef) carece de acceso a mecanismos de control natal de venta libre, a educación sexual y reproductiva formalmente impartida e incluso al intercambio informal que sobre temas de sexo se produce espontáneamente entre grupos de pares.

socioeconómica. Analizando los patrones de estos casos de entrenamiento, el modelo asignaría probabilidades de ocurrencia de un embarazo en otras niñas en una muestra compuesta por adolescentes de 15 a 19 años. Allí se identificaría a aquellas con el mayor riesgo “predicho” de resultar con un embarazo durante la adolescencia. A partir de tales predicciones, los funcionarios de Salta tomarían las decisiones del caso.

El proyecto atrajo de inmediato la atención de los medios de comunicación, pero sobre todo de los tecnólogos. También animó a la oposición política del gobierno de la provincia, que veía con inquietud la incursión de una Administración con recursos limitados en una fórmula que nadie conocía. ¿Tecnología en lugar de las clásicas aproximaciones sanitarias? Ver para creer. Los resultados del piloto podrían abrir o no las puertas a un abordaje moderno de un problema de salud pública, de manera que también los ojos de algunos curiosos del mundo se posaron sobre Salta.

El plan contó con la participación de Microsoft, que actuó como su proveedor técnico, ayudó a diseñar el modelo probabilístico de IA y entrenó al personal del Ministerio de la Primera Infancia a cargo de la implementación. Se trataba de una tarea aparentemente sencilla: reunir toda la información disponible en la provincia para predecir, lo más pronto posible, las probabilidades de que niñas y adolescentes quedaran embarazadas.

La alimentación del modelo de IA de Salta comenzó desde cero, pues el Ministerio de la Primera Infancia tenía un problema de salud que atender, pero no contaba con los registros organizados para hacerlo. Tras reunir, reordenar y recategorizar la información,<sup>78</sup> esta última, ya analizada, debía ajustarse a la herramienta de aprendizaje automático. Recién entonces el

<sup>78</sup> Los técnicos debieron incluso crear o reordenar grupos de categorías, pues los que existían no siempre disponían de datos personales identificables, completos o fiables.

sistema de inteligencia artificial organizaba estadísticamente los datos, dividiendo y recategorizando todo una y otra vez. Cuando se acopió una muestra significativa que presuponía cierto grado de certeza, se construyó el primer modelo. Y este primer modelo fue mejorado en miles de iteraciones.

La información reunida por el Ministerio en una base de datos combinaba decenas de variables de las personas a estudiar. Cada una de esas variables fue ubicada en una serie de categorías. Unas eran, por ejemplo, “personales” (estatura, peso actual, enfermedades declaradas, etc.); cinco se referían a la “educación” de la persona (incluyendo el nivel educativo completo de sus padres y hermanos, etc.); otras recopilaban datos de la “familia”, la “salud” y el “empleo” o estaban referidas a la “vivienda” (si viven en una casa hacinadas con seis o más individuos, el número de habitaciones, el material de construcción utilizado, etc.).

En total, el Ministerio relevó datos de 296.612 personas, el equivalente a casi el 23% de la población de la provincia de Salta, donde viven algo más de 1,3 millones de habitantes, más de la mitad de ellos mujeres. De ese total, casi 13.000 de los encuestados (el 4,27% de la muestra) eran niñas y adolescentes cuyas edades oscilaban entre 10 y 19 años.

De un Ministerio cuyos dispensarios de salud no sabían cuándo ni cómo recibirían jeringuillas para vacunar, se pasó a uno que contaba ahora con un volumen respetable de datos relativos a una proporción significativa de las mujeres de la provincia de Salta censados, organizados y listos para ser evaluados; pero sobre todo a una entidad pública que poseía las herramientas requeridas para establecer qué sucedía con las niñas que podían convertirse en madres antes de concluir la escuela secundaria o, aún peor, la primaria.

El piloto comenzó y se completó en siete meses en 2017. El uso del aprendizaje automático fue una novedad incluso dentro de la administración pública de Argentina, pero ciertamente en la salteña. Todos confiaban en que la combinación de algoritmos

y habilidades de procesamiento de datos que amplifican y aceleran el aprendizaje humano podrían identificar patrones y suministrar información que facilitara la planificación y la toma de decisiones. Había confianza.

Cuando Ministerio entrenó el sistema de IA reportó que había logrado identificar exitosamente al 98,2% de las adolescentes embarazadas. Los funcionarios afirmaron que el modelo había predicho que, “en función del análisis de sus patrones sociales” otras 397 niñas tenían una probabilidad del 88% de quedar embarazadas “en el futuro cercano.”

El éxito parecía iluminar el camino de la primera aplicación de IA de Argentina y América Latina enfocada en un área crítica. La expectativa internacional por el caso salteño seguía aumentando con las publicaciones de prensa. Una posibilidad hasta hacía poco inconcebible se convertía finalmente en realidad. Por fin en la región se podía ejecutar política pública a la altura de las naciones más desarrolladas. El tren de la Historia no la dejaría esta vez. ¿O sí?

### **No basta con soplar y hacer botellas**

La confianza progresiva en procesos automatizados para digerir y analizar información está afincada en la creencia de que la inteligencia artificial sirve para todo. Por eso estas aplicaciones atraen todavía más interés cuando se trata de abordar problemas estructurales que han preocupado por décadas a los funcionarios públicos.

Pero incluso antes de ponerla en marcha, hay preguntas críticas relativas al uso de IA que se deben responder: ¿es la inteligencia artificial necesaria para el problema que nos proponemos solucionar? Esta pregunta constituirá el punto de partida de cualquier intervención pública que desee incorporarla.

Lo cierto es que puede haber demasiados factores concurrentes que conduzcan a dar por supuesto que la IA es una solución imprescindible, sin cuestionar su factibilidad o su deseabilidad.

De hecho, suele suceder que los funcionarios deciden lanzar proyectos sin conocer la tecnología. Por su parte, las empresas saben desarrollar sistemas funcionales, pero no tienen conocimiento suficiente sobre el asunto a tratar. Hay interés en la demanda y en la oferta, pero no siempre confluyen de la mejor manera. El resultado de dar por supuesto algo sobre lo cual no se tiene certeza es subóptimo o frustrante. Las buenas intenciones no son siempre las mejores consejeras.

El programa piloto de Salta encaja en este escenario. Los funcionarios tenían la buena voluntad de anticiparse a resolver un problema de salud pública, pero durante el desarrollo del proyecto no se contemplaron medidas mínimas de mitigación de riesgo. Las dificultades fueron de índole variada.

La más notoria fue que no se tenía claro cuál era la intervención de política pública que se iba a ejecutar una vez que el modelo de IA identificara a la población en riesgo (asumiendo que el modelo detectara correctamente a adolescentes con probabilidad de quedar embarazadas). El proyecto piloto se abordó desde la tecnología, sin tener lista la intervención pública correspondiente diseñada para dar respuesta a los problemas estructurales y sociales que llevan a la ocurrencia del fenómeno del embarazo temprano. Las autoridades no publicaron los protocolos y las acciones de seguimiento que se implementarían con las adolescentes y tampoco definieron quién tendría acceso a esta información o como se protegería. Las consecuencias personales, psicológicas y sociales que tendría para una niña la divulgación de estos resultados no son difíciles de imaginar.

Mantener a la gente informada sobre el uso de sus datos también era otro aspecto crucial. La confidencialidad de la información sobre la vida de las personas encuestadas para el piloto había generado una controversia importante. Las condiciones del almacenamiento de la información no eran un tema menor. Empresas y gobiernos lidian a menudo con ataques a la seguridad informática que pueden poner en riesgo los datos de millones de personas. Por lo tanto, considerar los valores

éticos y legales del manejo correcto de los datos personales es un asunto de primer orden. Información tan amplia como la relevada en Salta pertenece a la intimidad de las personas y, como tal, es confidencial y merece protección prioritaria.

En cuanto a los datos con los que se entrenó el modelo, la gran mayoría pertenecían a niñas de ingresos bajos y muy bajos, de modo que los patrones identificados podrían no ser generalizables a la población para la que se buscaba hacer predicciones. Si bien es cierto que un porcentaje alto de embarazos adolescentes se presenta en grupos vulnerables o en situación de marginación, el muestreo de Salta no consideró mecanismos para asegurar la representatividad de la muestra para estas poblaciones, como tampoco para las niñas y adolescentes de familias de ingresos medios y altos. Tampoco se reveló qué volumen de las encuestas provenía de las zonas rurales, habitualmente más afectadas que las ciudades por los embarazos no deseados en proporción al número de habitantes.

Sobre este tema, el laboratorio de inteligencia artificial de la Universidad de Buenos Aires observó que, dado que el embarazo adolescente es un tema sensible, también era posible que quienes respondieran a la encuesta subreportaran los casos de embarazos no deseados o de abortos. Por lo tanto, la información que alimentaría el modelo no iba a reflejar necesariamente la realidad.

También hubo problemas técnicos. Para evaluar la precisión de un modelo de aprendizaje automático es necesario dejar por fuera del proceso de entrenamiento una sección de la base de datos con respuestas validadas que después puedan ser comparadas con los resultados de la predicción. El equipo técnico encargado del desarrollo cometió un error metodológico que hacía que la evaluación utilizara datos que el modelo ya había observado durante el entrenamiento, lo que podría llevar a sobrestimar la precisión del sistema.<sup>79</sup>

Con estos errores metodológicos y posibles sesgos en la información recolectada, sostener que el sistema tenía

una capacidad predictiva de más del 98,2% es sin duda cuestionable.<sup>80</sup> La ONG The Web Foundation, que analizó de manera independiente el piloto en su informe “Algoritmos e inteligencia artificial en América Latina”, comentó que sin una auditoría de un grupo independiente sería imposible saber si el modelo llevaba “a resultados justos o injustos.”

### Los principales retos de la IA: sus riesgos<sup>81</sup>

El caso de Salta demostró cómo los riesgos que conllevan los sistemas de IA devienen en retos que tendremos que resolver y que se presentan en cada una de las etapas de su construcción: desde la definición del problema de política pública, hasta la intervención, pasando por la recolección de datos, el entrenamiento de los modelos y su monitoreo.

<sup>79</sup> El error metodológico que cometió el equipo técnico está relacionado con el procesamiento de datos previo al entrenamiento del modelo. Por las características del problema, del total de mujeres de la muestra una proporción muy pequeña (cerca del 7%) resultó embarazada, de modo que un modelo que clasificara a todas las mujeres como NO embarazadas obtendría una precisión del 93%. Para resolver este tipo de desbalance en la base de datos de entrenamiento se suele ensamblar una sobremuestra de los casos menos frecuentes para así aumentar su peso en el modelo. Sin embargo, si esa sobremuestra se constituye antes de hacer la separación de la muestra para su validación, se estarían usando los mismos datos tanto para entrenar como para evaluar, obteniendo entonces una precisión que no es realista.

<sup>80</sup> Cuando se evalúa el comportamiento de un modelo de clasificación (por ejemplo, estar o no embarazada), una de las métricas más utilizadas es la exactitud, es decir, la fracción de las predicciones para las que nuestro modelo acertó. Sin embargo, cuando se tiene una base de datos donde el fenómeno que se busca predecir sucede muy pocas veces, tener un nivel elevado de exactitud puede no ser indicativo de un buen modelo. Por esta razón es importante que no nos dejemos llevar nunca por cifras como 98 o 95% de exactitud, pues podrían resultar engañosas.

<sup>81</sup> Los riesgos identificados en los sistemas de IA no implican de ninguna manera que estos se conviertan en problemas irresolubles. Por el contrario. La labor de los responsables de su concepción, diseño, uso y monitoreo consiste precisamente en reconocerlos y abordarlos con las medidas apropiadas para mitigarlos o eliminarlos del todo.

¿Qué riesgos deben considerarse entonces a la hora de proponer soluciones de política pública apoyadas en IA? La lista de pilares a tener en cuenta cada vez que un proyecto de IA entra en el campo de las políticas públicas es larga. En el gráfico a continuación se registran algunos de los desafíos más importantes para su implementación responsable.<sup>82</sup>



Los riesgos pueden manifestarse en diversos momentos: (1) durante la **conceptualización y diseño** de la política pública, cuando no se entiende bien el problema social que se busca solucionar —sus causas y/o el tipo de acciones o intervenciones requeridas— y se cree que la tecnología por sí sola lo resolverá; (2) durante el **ciclo de vida** de la IA, cuando no se consideran las distintas fuentes de sesgo ya sea en los datos de entrenamiento o en las decisiones de modelación que pueden dar lugar a sistemas deficientes, inequitativos o poco robustos; (3) cuando no se considera el **impacto sistémico**

<sup>82</sup> La lista de riesgos presentados a manera de retos fue elaborada a partir de las discusiones que tuvieron lugar en las mesas de trabajo multidisciplinarias organizadas por la iniciativa fAIr LAC con especialistas de ALC.

que la IA puede tener, bien directamente en los usuarios de los sistemas, o indirectamente en aspectos económicos, sociales o culturales; y (4) cuando se descuidan los riesgos transversales de **rendición de cuentas**<sup>83</sup> y de **gobernanza y seguridad**, debido a la exposición y vulnerabilidad de los datos personales y sensibles que usan los sistemas, así como el cumplimiento de requerimientos relativos a la transparencia y explicabilidad de sus resultados.

En mayor o menor grado, el proyecto de detección temprana de embarazos adolescentes en Salta se enfrentó a todas estas dificultades. No había un diseño de política pública ni una intervención sustentada en evidencia que pudiera ser implementada una vez que el sistema arrojase una predicción; no existía una gobernanza clara ni completo consentimiento para el uso de los datos con estos fines. La información con la que fue entrenado el sistema parecía no representar a la población objetivo. Hubo errores metodológicos durante el entrenamiento y las métricas de error pudieron haber sido sobreestimadas. Los ciudadanos desconocían cómo, para qué y de qué manera operaba el piloto, cómo se almacenaba la información y quién tenía acceso a ella.

El caso de Salta es solo una muestra de los problemas con los que se pueden encontrar otros emprendimientos públicos o privados que ya operan o que vayan a operar en la región. Todos los sistemas de IA deben someterse a los mismos cuestionamientos sobre sus métodos y objetivos, y sobre la calidad de sus procesos, servicios, productos y resultados. Es normal que así sea.

<sup>83</sup> ¿Qué sucedería, por ejemplo, si no se tomaran las medidas necesarias para asegurar el consentimiento por parte de los propietarios de los datos personales? ¿Sería ético implementar un modelo así? Aunque la respuesta (“no”) es obvia, la pregunta no debe soslayarse: ¿Qué implicaciones tendría no adoptar las medidas necesarias para asegurar la interpretabilidad (comprensible de manera bidireccional) y explicabilidad (que se pueda entender fácilmente el proceso) de los sistemas de toma y/o soporte de decisiones?

Si la región incorpora la inteligencia artificial a la actividad económica privada y al sector público, los gobiernos deberán contar con las capacidades suficientes para guiar y contener la nueva dinámica; esto en relación tanto con las condiciones razonables de operación que requieren los programadores, como con el trato justo que se le debe garantizar a los usuarios.

Si bien las leyes y regulaciones al respecto ya están operando, las infraestructuras son todavía insuficientes, los recursos profesionales y financieros escasos y el conocimiento general de la tecnología y sus particularidades reducido, como bien lo han observado numerosos empresarios y analistas. Es como tener un Boeing, pero solo disponer de una pista de aterrizaje para avionetas. El riesgo es elevado.

### **Conceptualización y diseño**

Ya habíamos mencionado cómo, de nada vale diseñar una solución cuando no se tiene claro lo que se busca con ella. El diagnóstico, el conocimiento de los insumos que esta requiere, y su viabilidad de acuerdo con el contexto son todos elementos que determinan sus posibilidades de éxito según los objetivos definidos. Una vez identificado el problema, se determinan aquellos factores que exigen intervención. Este riesgo existía antes de la IA y persistirá porque atañe al diseño mismo de la política. Una definición equivocada del asunto a tratar afecta a toda la cadena de decisiones; el mal diseño de IA conduce a una predicción errada. Todo ello conlleva una serie de costos de corto plazo que podrían extenderse en el tiempo.

La IA es una herramienta: un medio; es un instrumento que debe estar condicionado a un buen diseño de la intervención o acción en la que operará como sistema de soporte de las decisiones. La IA en ningún momento sustituye a la política pública. Una vez que el sistema ha realizado su predicción, procede implementar la decisión que se haya tomado a través de las acciones correspondientes.

Volviendo al caso de Salta, identificar a adolescentes en riesgo de embarazo es solo la primera parte. Los pasos que se den una vez que se tiene esa información alterarán el resultado del proyecto. No toda acción es necesaria, deseable o apropiada. Determinarla requiere algo más que saber ajustar un algoritmo de IA; presupone la aplicación de una escala de valores, de juicios, de experiencia en política pública e incluso de concepciones filosóficas sobre la pertinencia de la intervención.

Una vez tomada la decisión sobre cómo actuar, es preciso establecer quién o quiénes —instituciones o individuos— estarán a cargo del proceso que conduzca a ejecutar la acción. Esto incluye la disponibilidad de recursos económicos y logísticos para una implementación segura, así como la definición del tipo de modelo de IA que se va a ejecutar. Si la acción no se puede llevar a cabo por razones políticas o económicas, tener un modelo preciso no sirve de nada.

En este punto corresponde plantear el uso de la IA bajo criterios de necesidad y proporcionalidad. Las preguntas claves se dirigen a establecer si esa es la herramienta realmente necesaria para resolver X problema, o si este se puede solucionar con otra tecnología o incluso sin ella.

Una vez que se ha definido el diseño, la deseabilidad y factibilidad de la acción, entonces sí es posible pasar al ciclo de desarrollo de un sistema de soporte y toma de decisiones: el modelo de IA en sí.

De otro modo caeremos en lo que el intelectual Evgeny Morozov ha denominado acertadamente el “solucionismo tecnológico,” es decir, la creencia de que la tecnología tiene la capacidad de resolver por sí sola todos los problemas sociales.<sup>84</sup> Los practicantes y devotos de dicho dogma reducen fenómenos complejos de salud pública, educación y aplicación de la ley a

<sup>84</sup> Se trata de una ideología bastante extendida que Morozov describe en su libro *To Save Everything, Click Here: The Folly of Technological Solutionism*.

“problemas perfectamente definidos que pueden optimizarse fácilmente si solo los algoritmos correctos están en su lugar.” Todo se puede resolver con una aplicación, una cadena de bloque o con la denominada *gig economy* (economía del encargo).

Y así como debemos promover la utilización de la IA por parte del Estado allí donde tenga sentido, también tenemos que estar alerta para decir dónde no debe emplearse. Es necesario separar las capacidades reales de la IA como sistema generador de información, de las responsabilidades que competen a los encargados de la formulación de la política pública a la hora de diseñar intervenciones para resolver temas sociales. Esto por cuanto, entre estos últimos, habrá algunos donde la IA pueda tener una incidencia si acaso limitada o donde ni siquiera sea adecuada. De allí que en cada instancia en que se considere su aplicación, los responsables de la política pública y el equipo de IA deberán trabajar de forma conjunta, teniendo como Norte el impacto final de sus decisiones en las personas.

### **Rendición de cuentas y consentimiento (des)informado**

En 2019 Geoffrey Fowler, periodista del *Washington Post* especializado en tecnología, publicó en su columna lo que había encontrado en la web a partir de una investigación sobre la información que las aplicaciones de nuestros teléfonos recopilan diariamente: itinerarios de vuelo, calendarios, diagnósticos y recetas médicas, declaraciones fiscales, el historial de su ubicación diaria en Google Maps, y cuatro años de audios de conversaciones grabadas por Alexa, su asistente virtual, incluidos los detalles de un acuerdo de negocios que un amigo de Fowler cerró cuando estaba de paso por allí. Los datos provenían de extensiones de su navegador y de las distintas aplicaciones de su celular que descargó “conscientemente” y para las que aceptó todos los términos y condiciones.

El concepto de “consentimiento informado” implica que el propietario de los datos debe entender el tratamiento y uso que se les dará, lo cual es prácticamente imposible en el mundo de hoy, ya que la innovación hace difícil que las personas, de forma

individual, sean realmente conscientes de las implicaciones de compartir sus datos y del uso que se les dará, incluyendo los algoritmos que vayan a alimentar. Por esta razón, es necesario que los gobiernos adopten una actitud más proactiva en la definición de las reglas del juego, y que establezcan los estándares y normas que deberán cumplirse antes de poder liberar un sistema que pueda afectar la vida de las personas.<sup>85</sup>

### **Gobernanza y seguridad: asuntos mayores**

El manejo de información siempre entraña algún riesgo relacionado con su seguridad y protección. Si bien la falta de protocolos, así como el robo, secuestro y/o extorsión con información personal no son exclusivos de la IA y tampoco nuevos, la creciente expansión de los servicios digitales aumenta significativamente tanto la exposición de los datos personales de los ciudadanos que son almacenados día con día en miles de servidores, como la posibilidad de un uso perjudicial.

En los últimos años, en ALC se ha registrado un incremento en los casos de filtración de datos personales, y de secuestro o robo de información privada.<sup>86</sup> En ocasiones, los riesgos se generan por errores humanos producto de la falta de conocimiento y comprensión de los estándares de seguridad y de las buenas prácticas. En 2016, por ejemplo, el padrón de votantes del Instituto Nacional Electoral de México, que contiene la información de más de 93 millones de personas, quedó expuesto sin contraseña en un servidor en la nube, en apariencia por un error humano.

<sup>85</sup> En el capítulo 6 discutiremos cómo estas líneas difusas han llevado a redefinir los límites de la privacidad y los riesgos que esto conlleva, mientras que en el capítulo 7 abordaremos algunas de los conceptos legales —entre ellos los de responsabilidad, explicabilidad e interpretabilidad— que tenemos que madurar para lograr una IA robusta y responsable.

<sup>86</sup> Según el Eset Latin American Security Report de 2017 (una encuesta de 4.000 personas en 13 países), solo en 2016 el secuestro de datos y la extorsión aumentaron en un 131%.

En otras circunstancias, sin embargo, la ausencia de protección ha favorecido los comportamientos criminales. En la región, la situación es muy poco tranquilizadora en el sector público y también en el privado.<sup>87</sup> Una proporción significativa de las empresas de la región carece de sistemas suficientemente protegidos para resistir ciberataques,<sup>88</sup> mientras que los recursos<sup>89</sup> destinados a frenar el creciente avance de este fenómeno contra la seguridad de la información parecen ser siempre insuficientes.<sup>90</sup>

Existen pocos mecanismos regionales de defensa coordinados, en parte debido a que los niveles de desarrollo e inversión tecnológica difieren de un país a otro. Aunque ya hay varias naciones que se encuentran diseñando equipos de respuesta a emergencias cibernéticas (CERT) y equipos de respuesta a incidentes de seguridad informática (CSIRT) para prevenir o gestionar los ataques —por lo menos 17 para finales de 2017—, la mayoría no tiene siquiera un nivel intermedio de preparación.

<sup>87</sup> En su informe “El estado de la ciberseguridad en el sector bancario en América Latina y el Caribe” de 2018, la OEA informaba que al menos 9 de cada 10 entidades bancarias habían sufrido incidentes cibernéticos en el último año. Asimismo reportaba que 6 de cada 10 usuarios que no utilizan servicios de banca digitales aducían que no lo hacían por desconfianza en la seguridad de las transacciones.

<sup>88</sup> “Solo 10% de las empresas que sufrió un incidente con este programa maligno (el llamado *ransomware*) tenía una solución de seguridad, doble factor de autenticación y políticas de seguridad”, según se registra en el Eset Latin American Security Report de 2019. Un cuarto de las empresas afectadas no tenía respaldada la información con copias de seguridad.

<sup>89</sup> Un documento privado de una compañía de consultoría en seguridad informática, Comstor, vinculada a Cisco Systems, sostiene que los US\$12.000 millones que las empresas de la región tenían previsto invertir en ciberseguridad “no serían suficientes”. Por lo menos un 50% de las instituciones financieras de la región encuestadas por Comstor dijo haber sufrido algún tipo de ataque, desde robo de contraseñas hasta robo de identidad, entre otros.

<sup>90</sup> Según una encuesta presentada en el Eset Latin American Security Report de 2019, en América Latina dos de cada tres empresas sufrieron un incidente de seguridad durante ese año.

Países como Brasil, Argentina, México, Chile y Colombia, por mencionar algunos, han comenzado a adoptar estrategias nacionales de seguridad cibernética; pero todavía falta conciencia pública. No hay concertación entre los sectores público y privado, aparentemente por un asunto de desconfianza mutua. Esto, sobra decirlo, inhibe las posibilidades de cooperación e intercambio de información.

Si no se cuenta con las capacidades suficientes para frenar si quiera el robo de información personal, la debilidad es aún mayor frente a formas más avanzadas de manipulación electrónica, un aspecto determinante cuando se trata de crear sistemas de IA robustos.

### **Impacto sistémico y usos malintencionados**

Dado que ninguna intervención humana es inocua, la política pública debe plantearse pensando en el impacto de su aplicación. Aquí los principales riesgos a abordar son aquellos que se derivan de: (1) la interacción compleja entre los seres humanos y la IA; (2) los impactos del uso de algoritmos en la vida diaria de las sociedades; (3) las posibilidades de que el uso de la IA fortalezca o cree condiciones de monopolio de mercado o riesgos para los derechos de las personas; y (4) el uso incorrecto o malintencionado de los datos por parte de agentes políticos, económicos y sociales agresivos.

Por **interacción seres humanos-IA** entendemos el papel del usuario, en particular del funcionario público, como receptor crítico de las recomendaciones de un sistema de IA para la toma de decisiones. Ya hemos señalado el problema muy común de la fantasía de la exactitud (*accuracy*) del modelo. Por muy elevada que sea la cifra (digamos entre 95 y 99% de precisión), esta medida puede ser engañosa; no hay modelo perfecto y los errores, aunque sean mínimos, pueden tener un impacto considerable en miles de personas.

Por ejemplo, en un diagnóstico de una enfermedad podemos observar o bien falsos positivos, es decir, los errores que comete

el sistema al clasificar algo como positivo cuando no lo es, o su opuesto, los falsos negativos. Si el sistema está diagnosticando una enfermedad crónica mortal, un falso negativo que conduzca a decirle a una persona que no está enferma, cuando sí lo está, puede llevar a que esta no se tome los medicamentos que podrían salvarle la vida. Si por el contrario se tratara de un falso positivo, ello podría llevar al paciente a hacerse un mayor número de exámenes posiblemente costosos que a la postre muestren que no sufre la dolencia.

Por eso es importante que quien ajuste el modelo y quien lo utilice entiendan el desempeño del sistema en el contexto específico de su uso, así como las implicaciones de sus posibles errores, al tiempo que toman medidas para mitigar los riesgos identificados.

Un estudio del National Endowment for Science, Technology and the Arts (Nesta) de Gran Bretaña destaca tres principios clave en la interacción de los seres humanos con la IA que, si se tienen en cuenta, mejorarán la adopción y utilización de las recomendaciones de ese sistema: **contexto, comprensión y agencia**. Esto significa que el funcionario encargado de trabajar con la herramienta lo haga comprendiendo plenamente sus capacidades, así como el papel que debe cumplir en un entorno de toma de decisiones. Aquí el usuario de la herramienta debe poder entender los posibles sesgos del sistema y los errores que puede cometer para poder ajustar sus recomendaciones antes de que se tome una decisión final.

Otro tema fundamental surge en torno a la discusión de la propiedad intelectual de los sistemas y a su definición como bien público o privado. En general, un modelo de aprendizaje automático será más preciso si se entrena con una mayor cantidad de información de buena calidad. La posición privilegiada de gobiernos y empresas establecidas facilita su acceso a mayores volúmenes de datos, lo cual puede propiciar el **surgimiento de monopolios u oligopolios** con importantes impactos en la vida de los ciudadanos y en la calidad de la gestión de los sistemas públicos.

No es un asunto de poca monta. A mayor cantidad de usuarios de una plataforma, mayor el volumen de información que alimenta el sistema, lo que a su vez mejorará la calidad de los modelos de IA de su propietario. Esto propicia un tipo de mercado en donde el ganador se lleva todo (*winner-takes-all market*). Lo anterior significa que, gracias a la calidad si acaso ligeramente superior frente a la de sus competidores, un determinado producto, servicio u organización obtendrá una proporción mayor de usuarios e ingresos. Y al capturar un porcentaje más elevado de la demanda consolidará su liderazgo. Eso puede crear altos costos de entrada para nuevos desarrolladores de IA y por esa vía reducir la competencia.

Estos patrones van más allá del uso de la IA, pues el problema combina el acceso a la información con el monopolio de los canales de comunicación y la venta del producto final a los usuarios. Amazon, por ejemplo, ha sido acusada de copiar productos en su Marketplace. Cuando la compañía comenzó como una plataforma de comercio virtual en la que empresas externas podían vender y distribuir sus productos, esa intermediación le dio acceso a un volumen único de información.

Con el tiempo, la empresa adquirió la capacidad de hacer análisis de ventas de todos los proveedores e identificar los mejores productos, así como de estudiar los precios y el comportamiento de los consumidores a un nivel que ni los propios vendedores lograban alcanzar. En 2009, Amazon introdujo su línea personal Amazon Basics, a través de la cual ofrecía algunos de los productos que antes comerciaban los vendedores externos, hasta desplazarlos. Su primera iteración se centró en productos electrónicos y pilas. Los ofreció a un precio mucho menor que el de sus competidores y en poco tiempo había obtenido un tercio del mercado que antes compartían empresas como Duracell y Energizer. Hoy día, Amazon Basics vende todo tipo de productos, desde ropa hasta muebles, pasando por todo lo demás.

Y no es la única en la mira de las autoridades por acciones potencialmente monopólicas o de abuso de poder. En Estados

Unidos, el Congreso ha tenido una agenda de investigación abierta para determinar si empresas como Google y Facebook deben ser consideradas monopólicas y si es necesario dividir las para reducir su posición dominante. A mediados del 2020, las cuatro empresas de tecnología más importantes de Estados Unidos fueron citadas por el Congreso a audiencias antimonopolio. Jeff Bezos, presidente de Amazon, reconoció que si bien la compañía no utilizaba información individual por empresa, sí empleaba información agregada para tomar decisiones sobre sus productos.

Un último riesgo es el **uso incorrecto o malintencionado de la IA**. Esto abarca desde espionaje y manipulación (por parte de empresas y gobiernos) a través de la aplicación de la teoría del comportamiento, pasando por la hiperfocalización de la publicidad, hasta la desinformación con noticias falsas y *bots*, y con la simulación de rostros y voz (lo que se conoce como falsificaciones profundas o *deepfake*). De todo ello hablaremos en el siguiente capítulo.

### **Sesgo, justicia algorítmica y equidad en el ciclo de vida de la Inteligencia Artificial**

“Básicamente todos los modelos están mal, aunque algunos son útiles.” Esta frase es un conocido aforismo atribuido al matemático inglés George Edward Pelham que más nos vale tener siempre presente. Nuestra responsabilidad en el diseño de un nuevo Estado alimentado por IA radica precisamente en entender las características de los errores de sus motores.

El error de un modelo es la diferencia entre el valor predicho y el valor real. Cuando este error es aleatorio se considera ruido, pero si este se repite sistemáticamente se le conoce como sesgo. En un modelo, el **sesgo** por sí mismo no tiene necesariamente implicaciones éticas. Estas últimas aparecen al utilizar sus resultados para tomar decisiones, ya que estos pueden llevar a adoptar medidas injustas o dañinas para subgrupos de la población.

En 2015, la empresa Amazon experimentó con un sistema de recomendación de recursos humanos a partir de técnicas de aprendizaje supervisado. El modelo entrenaba con los resultados históricos de procesos de selección de candidatos almacenados durante diez años. En esa base de datos se identificaba si un candidato había sido aceptado o rechazado para el trabajo por el departamento.

El sistema de IA se construyó con el objetivo de crear un primer filtro para descartar perfiles que tuvieran una alta probabilidad de no cumplir con las expectativas del área de recursos humanos (RRHH). Trabajaba con la hipótesis de que el modelo podría capturar buenos candidatos y reducir el trabajo del departamento de RRHH al hacer la preselección.

El proyecto se justificaba dado que RRHH recibía miles de currículums que sus empleados difícilmente podían evaluar de manera individual. Lo que el equipo no había tenido en cuenta es que la industria de la tecnología ha sido predominantemente masculina por largo tiempo. Si bien es cierto que en los últimos años se han hecho esfuerzos notables por incrementar el número de mujeres empleadas, al tomar como base de entrenamiento ese historial el modelo de IA estaba fortaleciendo y volviendo permanente el comportamiento de exclusión.

Hoy día, la mayoría de los sistemas de IA disponibles en el mercado tienen como objetivo minimizar el error. Los equipos técnicos pocas veces hacen un análisis sobre el desbalance que el error puede generar entre subgrupos. A pesar de que las métricas de precisión del modelo en general sean altas, un sistema que privilegie sistemáticamente a un grupo en detrimento de otro no es deseable.

La forma en que se evalúa si el modelo está tomando una decisión “injusta” no es la misma para todos los problemas e incluso puede ser distinta para el mismo problema en diferentes contextos o países. Esto por cuanto lo que se entiende por “justicia” puede cambiar según la cultura y/o la tradición de cada sociedad. Por ejemplo, para ciertos casos

se pueden crear políticas que busquen equidad mediante acciones afirmativas como las cuotas de diversidad de género e inclusión, o mediante políticas restaurativas, mientras que en otras circunstancias la toma de decisiones equitativas puede estar basada en otro tipo de argumentos, como por ejemplo la participación por regiones o por provincias, o por estrato socioeconómico u otras variables.<sup>91</sup>

Una historia que describe la complejidad de lo que entendemos por “justicia” ocurrió en Estados Unidos en 1961, cuando el presidente John F. Kennedy firmó una orden ejecutiva que exigía al gobierno tomar “acciones afirmativas”<sup>92</sup> en sus adquisiciones de bienes y servicios para apoyar el objetivo nacional de no discriminación favoreciendo a proveedores de las minorías. La orden ejecutiva fue motivada por el movimiento a favor de los derechos civiles y reconocía que la población afroamericana había vivido una historia de desigualdad e injusticias que requería compensación. Durante los siguientes años, el concepto se fue transformando y las universidades de Estados Unidos comenzaron a utilizar “acciones afirmativas” en sus procesos de admisión. Desde entonces se ha dado un intenso debate sobre su aplicación y se han interpuesto varias demandas ante la Corte Suprema de Justicia por una supuesta “discriminación inversa” en detrimento de la población blanca y, en los últimos años, de los asiáticoamericanos.

El resultado de un modelo de IA y sus errores se pueden analizar mediante la representación matemática de estas definiciones de lo que es justo —algo que se conoce como **justicia algorítmica**— e incluirse en el proceso de selección y ajuste del aquel.

<sup>91</sup> En casi todos los países existen regulaciones y leyes sobre la igualdad de oportunidades y el acceso a los derechos sociales, así como a la gama de servicios que provee el Estado.

<sup>92</sup> Según la Real Academia Española, es el conjunto de medidas que adopta el Estado a fin de promover la igualdad real a favor de los titulares de derechos que se encuentren en situación de desigualdad.

Aunque se pueden emplear distintas definiciones, en la mayoría de los casos el cumplimiento de una impide que se cumpla otra, es decir, pueden ser excluyentes. Por ejemplo, cuando se sigue un criterio de **igualdad de oportunidades**, cada persona, independientemente del subgrupo al que pertenezca (género, raza, religión, etc.), debería obtener métricas de evaluación parecidas, siempre y cuando todas ellas cuenten con las mismas características.<sup>93</sup> Otro criterio es el de **paridad demográfica**, que establece que la proporción de cada segmento de una clase (género, raza, nivel socioeconómico, etc.) debe obtener métricas también proporcionales de acierto y error.

Cumplir simultáneamente con el criterio de “paridad demográfica” e “igualdad de oportunidades” es imposible. La situación se complica dado que hay decenas de criterios adicionales igualmente aceptables. Es importante entender que la definición de justicia no es una decisión matemática sino de política pública y que, en el contexto de la toma de decisiones, la discusión no es distinta a la que se tenía antes de introducir allí el sistema de IA.

Una vez que se ha definido el criterio de justicia algorítmica, es necesario analizar los resultados del modelo y eliminar o mitigar las fuentes que ocasionan el sesgo.

Ya habíamos señalado que los sesgos no aparecen mágicamente en los sistemas de IA, sino que son un reflejo de nuestra sociedad y de nuestras decisiones. Algunos vienen ya incorporados en los datos con los que se entrenará el modelo de aprendizaje automático. Como se explicó en el capítulo 3, estos algoritmos aprenden de los datos y sus características condicionan la forma en que se vayan a comportar. El sesgo surge porque no tomamos en cuenta que los datos con los que entrenamos a un modelo de IA pueden haber capturado prejuicios sociales preexistentes. Esto es lo que se conoce como sesgo histórico o discriminación pasada,

<sup>93</sup> Por ejemplo, en temas de contratación las personas más calificadas deberán ser las que obtengan el trabajo, independientemente de la disparidad que exista entre los grupos a los que pertenezcan.

y queda bien ejemplificado por el caso del algoritmo de recursos humanos de Amazon, donde la brecha salarial entre hombres y mujeres está ya representada en la historia con la que se entrena el modelo.

Otro ejemplo de este tipo de sesgo se puede observar en algunas interacciones con el sistema GPT-3 de la empresa OpenAI. El GPT-3 es un modelo de lenguaje autorregresivo no supervisado que utiliza aprendizaje profundo para analizar y extraer relaciones semánticas del uso y contexto de millones de palabras. Es hasta ahora el modelo de lenguaje más grande jamás creado: mientras que la versión inmediatamente anterior (GPT-2) tenía 1.500 millones de parámetros, la GPT-3 suma 175.000 millones. Sus resultados son sorprendentes, y es por mucho el sistema más avanzado para simular textos humanos. Sin embargo, no se debe pasar por alto que el GPT-3 fue entrenado utilizando una diversidad de textos obtenidos de internet de fuentes como Wikipedia, el corpus abierto al público WebText e información descargada de la red social Reddit. Esta última en particular es famosa por estar plagada de lenguaje racista, misógino y clasista.

Al ser una red neuronal profunda e implementar una metodología no supervisada, resulta particularmente difícil evaluar el campo completo de las interacciones que puede tener el sistema con los usuarios. Por ejemplo, cuando se le preguntó su opinión sobre Etiopía, el sistema contestó: “El principal problema de Etiopía es que la propia Etiopía es el problema. Parece un país cuya existencia no puede justificarse.” Las respuestas del sistema fueron igualmente discriminatorias cuando se le preguntó sobre derechos sexuales, racismo y género.

Es por ello que las decisiones de política pública no pueden fundamentarse en sistemas de limitada estabilidad y cuyos resultados sean poco predecibles. Esto no quiere decir que el GPT-3 —u otros modelos de lenguaje semejantes— no sean herramientas valiosas; lo que denota es que debemos identificar estos sesgos significativos para poder avanzar en la resolución de debilidades que puedan generar altos costos.

Siendo ALC la región más desigual del mundo, donde predominan sesgos claros de clase, raza y sexo, entre otros, este es un problema fundamental. Nos cabe entonces la responsabilidad de garantizar que, a medida que construimos modelos de IA, eliminamos tales sesgos.

La otra fuente de error muy común es el sesgo de representación. Este surge cuando no se sabe hasta dónde se pueden “estirar” los datos, es decir, para qué pueden servir y qué conclusiones podemos sacar de ellos y cuáles no. Este tipo de problema sucede a menudo porque los modelos de aprendizaje automático se entrenan con bases de datos que representan únicamente a una parte de la población. Por ejemplo, los registros médicos electrónicos solo contienen a usuarios de hospitales de primer nivel. Una base de datos provenientes de Twitter únicamente representará a un subconjunto de la población con interés en temas de política y peleas en línea, mientras que Facebook (que poco a poco está siendo abandonado por los jóvenes) tendrá principalmente fotografías familiares. El problema surge cuando el responsable de la formulación de la política pública procede a decidir como si esa base de datos representara a toda la población. Esto puede ocasionar sesgos importantes por diferencias de precisión entre subpoblaciones poco representadas en los datos.

También están los sesgos de modelación, los cuales se generan en el entrenamiento de los algoritmos, principalmente por errores metodológicos en la ejecución del proceso de preparación del modelo y de decisiones que se adoptan durante el manejo de la información. En los últimos años, varias empresas han puesto en venta servicios de aprendizaje automático que se ofrecen como autoadministrados (o “cajas negras”), los cuales no requieren mayor ajuste y/o mantenimiento, y que prometen realizar por sí mismos el entrenamiento de los modelos de IA.

Si bien tales servicios pueden funcionar para casos generales, ya hemos visto cómo, durante el desarrollo de un sistema de IA se toman decisiones de política pública que no se pueden dejar a merced del algoritmo. La evaluación de los errores y sesgos

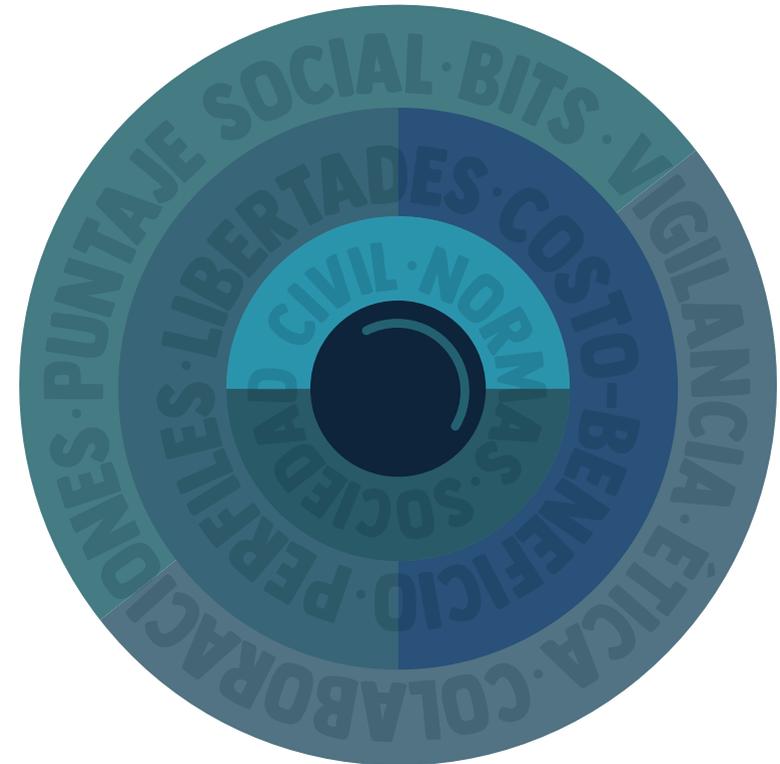
de un modelo siempre tiene que realizarse en el marco del caso de uso en donde se implementará el sistema. Por esta razón, siempre es necesario que haya procesos de evaluación humana antes de utilizar los resultados de la IA en el momento de tomar decisiones de política pública.

Por último, existe una capa más general de riesgos que se superpone a las anteriores, y que tiene que ver con la decisión que tomamos como sociedad de permitir el uso y almacenamiento de un volumen cada vez mayor de datos sobre diversos aspectos de nuestras vidas. Hasta hace pocos años, la única información con la que contaban los Estados sobre cada uno de sus ciudadanos era la que aparecía en el registro de nacimiento y en el acta de defunción. Hoy en día, la mayoría de ellos tienen bases de datos centralizadas con información laboral, bancaria, educativa y de salud sobre todos nosotros. Su uso para el análisis de datos constituye uno de esos casos donde se cumple el antiguo axioma según el cual “el todo es más que la suma de sus partes”: al integrar más información sobre los individuos, surgen patrones útiles para predecir sus comportamientos.

Ciertamente, estas bases de datos son sumamente útiles para un EBDI que se proponga tomar decisiones predictivas y mejorar sus procesos de focalización. Sin embargo, en ausencia de los contrapesos apropiados la concentración de información puede llevar a la pérdida de privacidad, a la vigilancia masiva y a la manipulación. Este riesgo no se origina únicamente en el Estado sino también en la empresa privada, que actualmente posee enormes volúmenes de información sobre los consumidores de sus productos. Y es aquí donde el EBDI tiene una doble tarea: la primera como actor y usuario de la IA, con incentivos para recolectar y extender el uso y aplicación de la IA a sus tareas, y la segunda como ente regulador responsable por definir las pautas y reglas del despliegue de la IA para las empresas privadas nacionales e internacionales.

En suma, son varios los riesgos que se manifiestan de forma particular en cada una de las instancias de los sistema de IA

cuando se intenta usar la automatización en los procesos de toma de decisiones de política pública. Entendemos tales riesgos, no como obstáculos insuperables, sino más bien como una capa adicional de implementación de los modelos, es decir, como un espacio de intervención en procura de su perfeccionamiento para evitar o mitigar las debilidades identificadas y así garantizar que el EBDI use la IA de manera responsable.



## CAPÍTULO 6

### Sonríe; te estamos filmando

*Sin control, el uso de la IA abre las puertas a oligopolios privados; también al Estado y al capitalismo de vigilancia.*

*“Si quieres mantener un secreto, ocúltalo hasta de ti mismo”—George Orwell, 1984.*

En su novela distópica *1984*, George Orwell imaginaba una sociedad donde un gobierno totalitario usaba maniobras de vigilancia masiva para controlar a la población. El régimen represivo pretendía mantener en línea el comportamiento de los individuos atrapándolos dentro de los límites de una sociedad acorralada, mansa y perfectamente ordenada. No había exlimitación en su lógica mientras la razón de Estado —el orden— requiriera acción.

Cualquier ciudadano chino sabe que está siendo observado. Cualquier ciudadano de Hong Kong sabe que está siendo observado por China. El mundo sabe que China observa. Desde el momento en que una persona abandona el edificio en que vive en cualquiera de las principales ciudades del gigante asiático, va dejando un rastro digital específico. Al salir de casa, debe mostrar su rostro en el escáner biométrico que permite abrir la puerta de la calle de su complejo residencial. El escáner alimenta el sistema inteligente de datos de la ciudad con esa identificación facial. En caso de necesidad, un oficial de policía podría correr un motor de reconocimiento facial en las grabaciones y en muy poco tiempo obtener una predicción bastante certera acerca de las personas a las que vería y sobre sus lugares de destino. En otras palabras, un ciudadano chino o hongkonés pueden reclamar *1984* como parte de su existencia.

No hay país del mundo cuyo aparato estatal de control concite tanto interés y comentario. Lo cierto es que China ha empleado tácticas de vigilancia estatal desde el nacimiento de su Estado moderno en 1949. Sin embargo, sus capacidades de control se han intensificado en los últimos años mediante el seguimiento de las comunicaciones electrónicas, rastreo por GPS, y tecnologías de reconocimiento facial y de voz.<sup>94</sup>

El de China es el sistema de vigilancia social más costoso, extendido y sofisticado del mundo. No hay ciudadano que pueda escapar del rastreo por mucho tiempo. La piedra angular del sistema es la política del “nombre real”, que obliga a cada persona a usar una identificación emitida por el gobierno si desea tener una cuenta en redes sociales, comprar una tarjeta

SIM para el móvil, tomar un tren o un avión e incluso ir al supermercado. Nadie puede estar por fuera de la urdimbre tecnológica estatal cada vez más amplia estructurada por Beijing. “En el vocabulario de los chinos no aparece el término ‘esfera privada’”, afirma el filósofo Byung-Chul Han.

China está inundada de cámaras en espacios públicos que facilitan a las autoridades, o bien el control de las multitudes y/o el seguimiento de un individuo determinado. Las cámaras de reconocimiento facial permiten detectar en tiempo real actos que “violán” la seguridad pública: desde manifestarse sin autorización estatal, como ha sucedido en Hong Kong desde 2019 y hasta bien entrado 2020, hasta comportamientos irrelevantes como cruzar una calle por fuera del paso peatonal.<sup>95</sup>

Se estima que las cámaras que integran el sistema de vigilancia del Estado Chino suman más de 200 millones: cámaras de seguridad que ya escanean las carreteras, los centros comerciales y las centrales de transporte, conectadas con las cámaras privadas de los edificios y empresas que son integradas en un sistema central de vigilancia llamado “Skynet”. Tal cual.

<sup>94</sup> Y no solo China. Singapur es otro ejemplo de implementación de estrategias difíciles de asimilar en Occidente, pues allí prima la aceptación social de medidas de restricción extremas. No nos referimos aquí a las de confinamiento de las poblaciones en riesgo o a la verificación de cumplimiento por geolocalización durante la pandemia, sino al hecho de que la policía podía irrumpir en los hogares de manera repentina para verificar que sus ocupantes estuvieran allí. Durante el primer brote, más de 6.000 personas fueron localizadas por las autoridades para constatar si habían tenido contacto con portadores, todos ellos ubicados mediante el uso de cámaras de circuito cerrado, GPS, investigación policial, y trabajo detectivesco, literalmente hablando. Los agentes del orden llamaban a preguntar, por ejemplo, si alguien X había viajado en un taxi específico el día miércoles a las 18:47 horas. El gobierno incluso destinó a soldados para que actuaran como detectives de rastreo.

<sup>95</sup> En Shenzhen, las cámaras de vigilancia detectan cuando una persona cruza la calle por fuera del paso peatonal y contrastan su rostro en la base de datos del gobierno. Y como las autoridades tienen registros de su número de teléfono gracias a sus alianzas con empresas como WeChat y Weibo, el envío de la multa es instantáneo.

El mismo nombre que el de las películas de Terminator, en una especie de sórdido tributo a su emulado. Esta emplea sistemas de reconocimiento facial para rastrear sospechosos, detectar conductas riesgosas para la seguridad, y supervisar el día a día de la población.

En varios artículos de prensa se han denunciado abusos policiales por causa del uso de estas cámaras, así como la intensificación de la vigilancia de críticos del régimen durante el brote de COVID-19. El uso de la llamada inteligencia de datos es cada vez más común entre las fuerzas de seguridad chinas, las cuales pueden revisar toda la información de tráfico en internet que les proporcionan las grandes proveedoras de conectividad y servicios —telefonía y redes sociales—, así como los movimientos e historial de los usuarios de móviles.<sup>96</sup> La idea de que Beijing sabe con precisión casi milimétrica dónde andan sus habitantes es exagerada. En rigor, el margen de error es de algunos metros.

Estas tecnologías se han empleado en todo el país, pero dos regiones son paradigmáticas para describir su uso potencial negativo por parte de un Estado totalitario que no se siente obligado a responder ante nadie: Xinjiang y Hong Kong.

### La crisis del pueblo uigur y la disputa por Hong Kong

A finales del 2019 se filtró a la prensa una serie de documentos en los que se revelaba la existencia de una red de campos de “reeducación” en la provincia de Xinjiang, donde estaban detenidos ilegalmente al menos un millón de musulmanes, principalmente miembros de la etnia uigur. Aunque al principio

<sup>96</sup> Durante la crisis, la compañía de telecomunicaciones China Mobile envió cientos de mensajes de texto a los medios de comunicación con información sobre las personas que habían resultado positivas al coronavirus. Su nivel de detalle era tal que podían incluir el asiento que el enfermo ocupaba en un avión o en el vagón del subterráneo. Los medios de comunicación publicaban esos datos en las redes sociales con la idea de alertar a quienes pudieran estar en los alrededores.

el gobierno chino declaró que los documentos eran falsos, posteriormente admitió la existencia de estos campos, pero sostuvo que se trataba de un programa destinado a integrar a las minorías musulmanas para que pudieran aprovechar las oportunidades que brinda la China moderna.

El territorio de los uighures se considera una región autónoma —Uighur de Xinjiang—, pero en la práctica ha sido controlada con mano de hierro por Beijing, que ha mantenido una estrategia de asimilación cultural incentivando la migración masiva de personas del grupo étnico Han (étnica que constituye el 92% de la República Popular China) al territorio de Xinjiang. Los uighures pasaron de ser el 70% de la población de Xinjiang en 1949 al 26% en los últimos años. Este flujo migratorio de los Han, visto por los uighures como una invasión cultural y política espoleada por el Partido Comunista, ha causado un fuerte descontento social que llevó a que se organizaran protestas, algunas de ellas violentas.

En 2014, el gobierno chino inició una campaña en contra de la violencia política que estructuró sobre lo que identificó como las “tres fuerzas del mal”: separatismo, terrorismo y extremismo. Como parte de ello se abrieron centros donde, según informes de prensa, personas que no habían cometido ningún crimen estaban siendo adoctrinadas para que abandonaran su religión y se afiliaran al Partido Comunista. El gobierno publicó una lista de 75 indicadores de un comportamiento que clasificaban como “extremismo religioso.” Algunos eran sumamente vagos e imprecisos, entre ellos “almacenar grandes cantidades de alimentos,” dejar de fumar o beber en forma repentina, o comprar equipo deportivo como pesas o guantes de boxeo.

Una nueva filtración develó el uso de tecnología para rastrear y vigilar a este grupo étnico. Las autoridades llevaron a cabo una recolección masiva y obligatoria de fotografías, datos biométricos, ADN y muestras de voz, con las que mediante sistemas de IA detectan a esta población en función de su apariencia. Esto les permite implementar un control minucioso y hacer una evaluación de sus criterios de comportamiento

para clasificar a las personas según sus niveles de confiabilidad, para entonces decidir si deben ser llevadas a los campos de reeducación.

En Hong Kong, por su parte, la avanzada del Estado de vigilancia ha sido significativa y evidente para el mundo entero. El 1 de octubre de 2019, para la celebración del septuagésimo aniversario de la fundación de la República Popular China, miles de personas se enfrentaron a la policía en diversos distritos de Hong Kong. El choque fue parte de un proceso de escalamiento progresivo de la inconformidad de los habitantes del área administrativa especial con las políticas de Beijing. Las protestas habían comenzado en marzo, cuando los primeros manifestantes se reunieron para rechazar una nueva ley de extradición, percibida por los habitantes como una excusa para acallar opositores políticos.

Esas protestas se extendieron por más de un año, hasta mediados de 2020. En octubre de 2019 se produjo la represión más intensa. Uno de los aspectos más interesantes de los enfrentamientos entre la policía y los manifestantes es el aprendizaje que se evidencia entre los jóvenes para lidiar con las autoridades. Las imágenes televisivas mostraban una especie de espectáculo de luces cuando la policía pretendía cargar contra los manifestantes. Eran apuntadores láser, como los que se usan en una conferencia. Los jóvenes los dirigían a las cámaras de la policía para cegarlas y evitar que sus rostros quedaran expuestos a los sistemas de reconocimiento facial.<sup>97</sup>

Antes y durante las manifestaciones, los organizadores empleaban el servicio de mensajería instantánea Telegram, que opera con mensajes encriptados, para evitar a los espías de las fuerzas de seguridad. El resto de la estrategia, enfocada en

<sup>97</sup> No fue este el único truco tecnológico, ni tampoco el último. En noviembre de 2019, apenas un mes después de las confrontaciones y de la subsecuente represión en Hong Kong, otros manifestantes, esta vez en Chile, lograron neutralizar a los drones de la policía apuntándoles múltiples láser hasta confundir sus giróscopos internos.

eludir el seguimiento tecnológico inteligente, consistió en salir de la red digital. Así, los manifestantes pagaban sus compras en efectivo en lugar de usar las tarjetas de crédito; adquirirían trayectos de una sola vía en el metro en lugar de usar tarjetas prepagadas; se abstendrían de tomarse *selfies* y fotos en las áreas de protesta, todo ello para mantenerse lejos del omnipresente ojo policial.

### La misma tecnología, otros fines

La reacción de los ciudadanos de Hong Kong no es un comportamiento usual en China. Por el contrario, cuando la crisis del coronavirus, la sociedad —incluso los habitantes de esa área administrativa especial— se mantuvo en sintonía con sus autoridades. Hablamos de más de 1.300 millones de personas que siguen de manera ordenada las directrices de un mando centralizado. Es realmente notable. Cualquier planificador occidental (independientemente de su mayor o menor adscripción a la democracia) debe observar estos sucesos con una mezcla de pasmo, preocupación y fascinación.

El seguimiento y control estatal de la crisis del COVID-19 en China contó con la asistencia de la misma inteligencia artificial contra la que batallaron los jóvenes de Hong Kong, en particular la de las cámaras de identificación facial conectadas a los sistemas de procesamiento de información de las fuerzas de seguridad. Esas cámaras fueron las que permitieron seguir el movimiento de las personas dentro de las zonas de confinamiento. La policía podía saber si un individuo enfermo o sospechoso de estarlo abandonaba su área restringida combinándolas con los GPS de los teléfonos móviles.

Y todo eso fue posible con el manejo de volúmenes enormes de información. Para que muchas de esas tareas se ejecuten, es decir, para que el Estado tenga la capacidad orwelliana de saber qué hacen sus ciudadanos, los sistemas de IA de aprendizaje automático requieren y dependen de la generación, recolección, almacenamiento y análisis de datos. Es allí donde el ambiente se vuelve asfixiante, pues ello implica que el Estado

debe **saberlo todo**, o al menos mucho más de lo que sabe hoy sobre los habitantes de una nación.

Es en este punto cuando las aguas de la IA se tornan turbias. La potestad para usar la información de las personas marcó la primera gran diferencia entre China y Taiwán en el manejo de la crisis del coronavirus.<sup>98</sup> Mientras que la primera fue el origen de todo, e incluso sospechosa de ocultar cifras de contagios y muertes reales, la segunda se convirtió en uno de los pocos ejemplos de mitigación eficiente y veloz de la pandemia. Esa diferencia, visible en la superficie, también lo era de fondo. Mientras que en la isla las autoridades tenían capacidades limitadas por ley para emplear datos individuales, Beijing prendió todos los motores de un Estado de vigilancia orwelliano de tan vasta cobertura que desafía la imaginación, asombra a los ciudadanos y atemoriza a Occidente.

### ¿Obedecer como elección?

China clamó victoria sobre el brote de SARS CoV-2 y procedió inmediatamente a realizar una amplia campaña mundial de relaciones públicas. El país ha sido especialmente activo cuando se trata de promover su respuesta al COVID-19, y en esta ocasión lo hizo mucho más rápidamente que cuando se presentó el SARS. Las imágenes del hospital construido en Wuhan en 10 días fueron presentadas como una proeza de ingeniería —y bien que lo es— en los medios y redes de cada país. Las ayudas y exportaciones de equipos médicos a 120 naciones han sido coreadas como un ejemplo de solidaridad desinteresada.

Entre tanto, la obediencia social fue exhibida como una adaptación a un mando centralizado que garantiza el orden. Su propaganda apuntaba a ablandar la imagen de autoritarismo del régimen entre las audiencias occidentales.

<sup>98</sup> Los detalles del manejo de la crisis en ambos países se encuentran en el capítulo 1.

Un Estado controlador necesita legitimarse para no ser visto como una dictadura, sino más bien como un sistema que se ocupa decididamente del bienestar general y no de los deseos particulares de los individuos. “Los defensores a ultranza de la vigilancia digital,” escribió Byung-Chul Han, “proclamarían que la inteligencia de datos salva vidas humanas.”

El punto de fondo de una tecnología con un potencial tan invasivo como la IA es el análisis costo- beneficio que conllevan tales intervenciones. ¿Dónde trazamos las fronteras éticas, jurídicas y políticas? ¿Qué derecho tiene un actor político o económico —el Estado o una empresa, por decir algo— a aprovecharse del conocimiento que adquiere de los ciudadanos o consumidores investigando sus actitudes y decisiones de vida para manipularlos?

Por su nivel de generalización y aplicabilidad, la IA podría facilitar la transición del Estado paternalista al Estado manipulador. Para ello basta con reproducir las estrategias de mercadeo de las compañías privadas, las cuales implementan métodos basados en el análisis de datos para incentivar comportamientos específicos.

Un individuo manipulable, es decir, que reacciona de manera condicionada y menos crítica a los estímulos, es el sueño de las autocracias. Se pulsa un botón y todos saludan. Una vez que existe un sistema capaz de analizar *petabytes* de información y entregar predicciones altamente probables, el Estado tipo 1984 tiene en sus manos la herramienta adecuada para utilizarla en cualquier acción de reforzamiento social, de la misma manera en que una empresa puede usarla para que un cliente infrecuente se torne en un usuario cautivo.

### El caballo de Troya y el COVID-19

La crisis del COVID-19 ha dado a las autoridades razones y pretextos para acelerar la recolección masiva de datos personales. Esta es una posibilidad peligrosa, sobre todo en naciones como China, donde no hay leyes estrictas que

protejan la privacidad de los individuos. Al respecto, el filósofo Giorgio Agamben señala que el coronavirus ha permitido a los gobiernos dar un paso adelante hacia el establecimiento de un Estado de excepción con vocación de perpetuidad.

El siguiente ejemplo es ilustrativo. Durante la pandemia, los policías chinos emplearon cascos inteligentes para identificar y separar individuos con una elevada temperatura corporal.<sup>99</sup> Sin embargo, como los mecanismos de reconocimiento facial tenían dificultades para identificar con precisión a las personas que llevaban mascarillas quirúrgicas, las autoridades acogieron nuevas propuestas tecnológicas que permitirían reconocerlas incluso con el rostro cubierto con el tapabocas.<sup>100</sup>

Entre el diseño y la producción de esta tecnología, y las caras tapadas de los jóvenes de Hong Kong durante las protestas, han transcurrido menos de seis meses. Los activistas han denunciado que sin duda el gobierno ha encontrado en la pandemia justificaciones para su sistema de vigilancia y están convencidos de que las autoridades no descartarán el uso de las nuevas tecnologías de control ya probadas, una vez que se supere la crisis del coronavirus.

Ninguno de los mecanismos de control utilizados por China con anterioridad ha sido desechado una vez desaparecida la

<sup>99</sup> Una empresa china, Rokid, adaptó un diseño de sus gafas detectoras de temperatura para el seguimiento del COVID-19.

<sup>100</sup> La compañía Hanwang Technology Ltd. aseguraba ser capaz de reconocer al 95% de las personas que llevaban el rostro cubierto. Según la empresa, su tecnología, que identificaría “en segundos” a cualquier individuo en un grupo de 30, podría también detectar “terroristas y sospechosos”. Uno de sus clientes es el Ministerio de Seguridad Pública del Estado, que ha convocado a muchas otras empresas a imaginar y ofrecer nuevas tecnologías de uso social. Otra empresa emergente china, SenseTime, también promocionaba la precisión de su sistema de reconocimiento y distribuyó su cámara de detección de temperatura en estaciones de metro, escuelas y centros comunitarios de Beijing, Shanghái y Shenzhen. Tras ello creció el interés de las fuerzas de seguridad de Occidente, en particular de Estados Unidos.

razón de su uso.<sup>101</sup> Las enseñanzas que en materia de seguridad y vigilancia digital dejaron eventos masivos como los Juegos Olímpicos de Beijing 2008 y la Expo Shanghái 2010 facilitaron la implementación y mejora progresiva de las plataformas y herramientas tecnológicas de control, cuyo alcance, funciones y capacidades fueron ampliados con la nueva estrategia nacional de desarrollo tecnológico lanzada en 2017. Las técnicas de vigilancia masiva se volvieron más permanentes tras estos eventos, según la organización de derechos humanos Human Rights Watch.

La línea que divide el uso de la información privada y el seguimiento personal para evitar la propagación de una enfermedad, de su aprovechamiento para efectos de control social, es bastante borrosa. Los Estados con vocación autoritaria no necesitan circunstancias excepcionales como las de una pandemia para perseguir a sus conciudadanos; sin embargo, una situación crítica o extrema constituye siempre una oportunidad para ensayar la tolerancia de las personas frente al recorte de sus derechos individuales con el argumento de la necesidad colectiva, el imperativo sanitario, la protección del público o la seguridad de la patria.

No es de asombrarse entonces que durante la crisis del COVID-19, la doctora Cecile Viboud, científica de los Institutos Nacionales de Salud de Estados Unidos, haya manifestado lo siguiente: “Hemos creído durante mucho tiempo que el análisis de datos puede ayudar al gobierno a pronosticar de manera efectiva el desarrollo de una epidemia, y para hacerlo necesitamos incorporar la recopilación de datos a la vigilancia. China tiene un sistema de vigilancia muy completo que ha

<sup>101</sup> No solo en China existen mecanismos de seguridad y control que han permanecido activos incluso después de que desaparecieran las circunstancias que los motivaron. El Acta Patriótica sancionada en Estados Unidos tras los atentados del 11 de septiembre de 2001 se mantuvo hasta 2015, cuando caducaron algunas disposiciones, si bien estas fueron reemplazadas poco después por una nueva norma: el Acta de Libertad de Estados Unidos.

demostrado ser útil para acopiar los datos necesarios.” Aunque se trata de una declaración bienintencionada, oculta el alto costo personal que los ciudadanos chinos tienen que pagar por esos sistemas.

Una de las razones por las que el Estado chino es el mejor preparado para la vigilancia masiva es su relación con la industria privada, en lo que se puede caracterizar como un matrimonio de dominancia. Antes y ahora, empresas como WeChat han sido útiles para apoyar los esfuerzos de control del coronavirus suministrando información a las autoridades y ayudando a guiar la logística de circulación en situaciones de aglomeraciones masivas.

El problema de estas colaboraciones es que la misma plataforma y metodología, así como la información que compilan, pueden ser útiles para fines menos nobles. WeChat, usado cada mes por alrededor de 1.000 millones de personas en China, combina en un solo sistema las redes sociales, el chat, el GPS del teléfono y los pagos de las billeteras electrónicas de cada usuario. Ese sistema permitió que, durante la crisis sanitaria, las personas no tuvieran que usar tarjetas de crédito o billetes para pagar —con el consecuente incremento del riesgo de contagio—, pues podían emplear las tarjetas virtuales de pago de sus móviles soportadas por la tecnología de WeChat.

La combinación de parámetros inteligentes hace de este tipo de compañías innovadoras un objetivo apetecible para gobiernos que persiguen controlar individuos y multitudes. La geolocalización inteligente es tan precisa que los errores son de pocos metros; además, como es móvil, resulta un complemento más práctico, veloz y ubicuo que las cámaras biométricas instaladas en las ciudades. La billetera electrónica de WeChat, por ejemplo, permite conocer al detalle qué ha adquirido una persona, en qué lugar y a qué hora, facilitando no solo la deducción de patrones de consumo sino inferencias sobre su ingreso probable y sobre sus decisiones futuras vinculadas al estatus sociocultural de su grupo. La IA de WeChat puede suponer que la compra de una medicina determinada o de

un producto vigilado por el gobierno tienen una intención específica, y a partir de ello generar una alerta. Luego están los chats. Aquí los motores lingüísticos de la empresa recorren las salas en busca de patrones que van desde conversaciones sobre una tos y una fiebre elevadas hasta, por ejemplo, críticas a las autoridades que a la postre generan reportes sobre sus autores.

Pero no solo es China. En Israel, Taiwán, Corea del Sur e Italia,<sup>102</sup> las compañías telefónicas han facilitado al Estado el acceso a información agregada y anónima sobre la movilidad de sus usuarios, una medida que permitió a las autoridades sanitarias el control de portadores sintomáticos y asintomáticos del coronavirus, pero que pueden también pavimentar el camino hacia usos indebidos como el seguimiento de opositores o el espionaje interno.<sup>103</sup> De hecho, en Israel el gobierno ignoró las advertencias escritas de los asesores legales de la Comisión de

<sup>102</sup> Un grupo de compañías italianas desarrolló una aplicación que permite seguir los movimientos de una persona infectada y notificar a aquellos con los que estuvo en contacto y que puedan estar en riesgo. El consorcio descartó posibles vulneraciones a la privacidad (las autoridades tienen acceso directo a la aplicación), pues no compartiría los datos personales ni el número de teléfono de sus usuarios. En Lombardía, una de las regiones más afectadas por el primer brote de la pandemia en Europa, las telefónicas Vodafone y TIM permitieron al gobierno de la región analizar los registros telefónicos para poder seguir los movimientos de las personas en riesgo. Los datos eran compartidos de manera anónima. El modelo lombardo imitaba al de Corea del Sur, donde la información del GPS y sobre el uso de tarjetas de crédito había permitido crear un mapa de la infección. “Esta medida es un poco perjudicial para la privacidad, [por lo que] debemos asegurarnos de que los datos se utilicen con fines de salud pública”, advertía Paolo Bonanni, profesor de higiene en la Universidad de Florencia y miembro de la Sociedad Italiana de Higiene, Medicina Preventiva y Salud Pública. “Esta es una medida excepcional que solo debe llevarse a cabo durante un período específico.”

<sup>103</sup> También en México, activistas de derechos humanos y periodistas han sido espiados por el Estado que, para 2017, había gastado US\$80 millones en un software israelí que infiltra los teléfonos inteligentes de las personas. En Guatemala, tres gobiernos sucesivos emplearon diez programas informáticos para hacer espionaje ilegal de empresarios, políticos opositores, periodistas, activistas y diplomáticos.

Defensa sobre los riesgos que dicha medida representaba para las libertades democráticas. El periódico *Ha'Aretz* denunció entonces que esa determinación, junto con las “medidas más draconianas en la historia israelí,” permiten “rastrear los desplazamientos en gran escala de ciudadanos respetuosos de la ley.”

Dado que los mecanismos de control social menoscaban las capacidades sistémicas de las democracias, la posibilidad de que algunos gobiernos acojan la idea de un Estado vigilancia por razones de salubridad pública, así sea de forma temporal, es altamente riesgosa. En períodos de guerra, los gobiernos adquieren la potestad delegada por las leyes para expropiar bienes (pueden, por ejemplo, convertir un hotel en un hospital), pero ello sucede con reglas y límites claros. De la misma manera, deberíamos considerar protecciones especiales para la información que producimos o entregamos en los entornos digitales a causa de la pandemia.

### La importancia de un *like*

Junto con su colega David Stillwell de Cambridge University, el psicólogo Michal Kosinski investigó entre 2008 y 2014 la posibilidad de identificar rasgos psicológicos de las personas a partir de sus selecciones de “me gusta” en Facebook, cruzando esta información con rasgos básicos de la personalidad como conciencia, apertura, simpatía y neurosis obtenidos de encuestas en línea realizadas en la misma plataforma.

Con esta información, Kosinski y Stillwell elaboraron perfiles psicológicos a partir de los cuales sugirieron que temas aparentemente superfluos como el gusto por las papas fritas o el temor a las tormentas eléctricas se podían utilizar para predecir detalles altamente sensibles sobre las personas, incluyendo su orientación sexual, su etnia, sus puntos de vista religiosos y, por supuesto, sus posiciones políticas. Para quien use o trafique con estos datos, el trabajo de estos dos investigadores es oro en polvo.

La labor de Kosinski y Stillwell fue tan exitosa que uno de los profesores de la universidad, Aleksandr Kogan, decidió fundar la firma de ciencia de datos Global Science Research, GSR. ¿Qué importancia tiene esta empresa aquí? Pues que poco tiempo después, GSR cambiaría de nombre y se convertiría en Cambridge Analytica, la compañía que tuvo una participación bastante notoria en la campaña presidencial de 2016 en Estados Unidos.

La historia es conocida. Utilizando información de 87 millones de usuarios de Facebook cuyos perfiles estaban en su poder, Cambridge Analytica se dedicó a alimentar a cada individuo con un anuncio personalizado, diseñado específicamente para su perfil psicológico. El objetivo: condicionar su voto. Para ello se proponía, por ejemplo, activar su temor a la pérdida del empleo para incentivar la respuesta deseada, ya fuera votar por un candidato, no votar por otro candidato o donar fondos a una determinada campaña.

Los métodos de Cambridge Analytica viajaron de Estados Unidos a otras naciones donde la empresa participó en campañas electorales, incluidas las presidenciales en Ucrania, Malasia y Kenia, y en la elección de Jair Bolsonaro en Brasil. En su informe sobre la libertad en internet durante 2017, Freedom House registró que al menos en 18 países se habían empleado tácticas de manipulación o desinformación durante las campañas, en no pocos casos con motores más o menos sofisticados de IA. La combinación de datos robados, ética maleable y aprendizaje automático puede ser extensiva, definitivamente malintencionada y, aun así, suponerse inocente.<sup>104</sup>

<sup>104</sup> Google, por ejemplo, es una máquina que opera con IA. Como las personas realizan 5.500 millones de búsquedas al día en su ventana, los robots de Google aprenden de los criterios de búsqueda de las personas para crear patrones, que luego reproducen cuando el buscador realiza predicciones contextuales de lo que escribimos. Una periodista de *The Guardian* lo probó. Escribió “los judíos” y luego, letra por letra, “s-o-n” y el buscador, por sí solo, le sugirió varias respuestas. La cuarta de ellas era “los judíos son malos”. La IA había aprendido lo que los humanos querían saber y ahora anticipaba respuestas con base en sus búsquedas.

Cambridge Analytica no fue la primera en utilizar análisis predictivo en unas justas presidenciales,<sup>105</sup> aunque sí fue pionera en combinar segmentación psicográfica y ciencias del comportamiento con analítica avanzada para modificar y/o reforzar acciones. Pero sobre todo fue una de las primeras en hacer un experimento de este tipo —y definitivamente no será la última— con un alcance tal que resulta inimaginable incluso hoy día: 87 millones de personas fueron **atacadas** simultáneamente. En 12 años —el periodo que va de 2008 a 2020 o incluso solo entre 2016 y 2020—, el uso de métodos analíticos para distorsionar las discusiones en línea y suprimir disidencias se ha vuelto global. Y no solo eso. Se usa incluso para moldear nuestra selección de prendas de vestir, autos o libros.

### Metadatos: el capitalismo de vigilancia

En el mundo en que se desenvuelva la IA se considera materia prima a toda la producción intangible de las personas: desde sus datos filiatorios hasta su última entrada en Facebook, pasando por sus compras del supermercado y las horas semanales de gimnasia. Ese insumo que representan todos nuestros datos y metadatos tiene un valor; un valor enorme. Las empresas lo saben y algunos gobiernos también. Pero quienes disputan y ocupan el poder lo saben mejor que nadie. Sin embargo, una abrumadora mayoría de las personas que producen esta riqueza informativa no es consciente de ello. La gente que escribe en las redes sociales lo considera diversión o interacción regular con otras personas, mientras que sus compras en el supermercado no tienen más valor que el que ha pagado por ellas o el impacto que tienen en sus magros salarios. No estamos conscientes de nuestro peso específico como productores de materia prima para la industria de la IA.

<sup>105</sup> El objetivo de una campaña política es modificar comportamientos e intenciones de voto; el uso de la tecnología es solo una herramienta para lograrlo. Durante la campaña de Barack Obama en 2008 también se empleó IA para detectar patrones de conducta.

No cobramos por ella. Al ceder información esencial sobre quiénes somos, estamos trabajando gratis.

Shoshana Zuboff, crítica de tecnología y profesora de la Universidad de Harvard, introdujo el concepto del capitalismo de vigilancia (*surveillance capitalism*) como una estructura de mercado en la que la mercancía en oferta son datos personales y comportamientos en línea obtenidos mediante una vigilancia masiva en internet, actividad esta que corre por cuenta de súper empresas que operan motores de captura de información escondidos en servicios percibidos como gratuitos, entre ellos las redes sociales y los motores de búsqueda.

Combinados con capacidades computacionales avanzadas, inteligencia artificial y aprendizaje automático, esos datos de comportamiento arrojan predicciones sobre las actitudes humanas que las empresas venden como un nuevo producto a mercados de consumidores interesados en segmentar a sus clientes y focalizar publicidad. Todos esos metadatos producidos, de cuya generación somos prácticamente inconscientes, tienen hoy un mercado privado y público. Billonario.

El análisis de lo que hacemos en internet alimenta estrategias de colocación de productos e inducción de selecciones cada vez más sofisticadas. Lo hace Nike cuando entramos a ver unos tenis en su sitio web, y minutos después encontramos los mismos zapatos en avisos en Facebook o en *El País* porque el algoritmo leyó nuestro comportamiento de búsqueda; nuestro interés. Lo hacen Amazon, American Airlines, las firmas de prendas de vestir, las de aparatos electrónicos y las de electrodomésticos. Los vendedores de especias orgánicas premium también lo hacen. Lo hizo Cambridge Analytica para manipular elecciones.

La buena noticia para las empresas es mala para los individuos. Los datos no perecen, pueden ser utilizados de forma continua y no son exclusivos; se les conoce como un bien “no rival”: su uso no perjudica, limita o impide el uso simultáneo por parte

de otro consumidor. Una vez que la información es pública se puede almacenar, intercambiar y vender un sinnúmero de veces para diferentes áreas.

Es muy probable que la base de datos que utilizó Cambridge Analytica siga allá afuera, dado que los perfiles psicográficos de las personas no cambian tan rápido. Aunque hasta ahora no se han encontrado informaciones acerca de que la base se esté vendiendo en el mercado negro, podemos estar seguros de que los datos de esos 87 millones de personas estarán como mínimo en otros servidores, junto con los de otros 2.500 millones de personas: los de Facebook.

### Súper empresa: me parece que esto es mío

Ahora, ¿a quién pertenecen los datos hoy? La ausencia de legislación detallada en numerosos países tiene un ganador. En este momento nuestros metadatos les pertenecen a las empresas.<sup>106</sup>

Las personas producimos bits en cantidades industriales y con cada clic cedemos esa información a decenas de empresas sin pensarlo dos veces. La razón de que así sea es que desconocemos su valor; nos es abstracto, distante y muy difícil de determinar. Sabemos cuánto puede valer una casa, un auto, un kilo de pan, una computadora o nuestro trabajo porque hemos crecido en ese mundo. Pero las particularidades de los datos digitales escapan a nuestra comprensión ¿Cuál es el valor contante y sonante de tener una lista de películas que hemos disfrutado, el valor de saber cuántas cervezas tomamos a la semana, qué tipo de ropa nos gusta y si finalmente nos suscribimos a Amazon Prime?

Si nos preocupa el Estado de vigilancia también nos debe preocupar la súper empresa. Muchos cuestionamos la captura de datos por parte de actores autoritarios, pero damos información gratis a decenas de compañías de servicios cuando aceptamos

sus acuerdos de confidencialidad y uso de información. Nos desentendemos de ello. Son contratos largos llenos de letra menuda y lenguaje leguleyo. Cuando el sistema nos pide leerlos y dar nuestro consentimiento, nos desplazamos raudos hacia abajo, pulsamos la casilla de aprobación y nos olvidamos del asunto. Nunca más prestamos atención a la información que recogen, en la mayoría de los casos de manera silenciosa, anónima y desapercibida. No somos mínimamente conscientes de la asimetría evidente. El usuario no sabe que su comportamiento es monitoreado ni a qué profundidad, mientras que la empresa lo sabe to-do.

Es curiosa nuestra reacción. Le tememos al control potencial que pueda ejercer el Estado sobre nuestras vidas —un temor que debemos preservar—, y al mismo tiempo permitimos que actores privados prácticamente decidan por nosotros, sugiriendo qué podemos comprar después de analizar nuestros comportamientos históricos de consumo. No queremos dar información al monstruoso Estado orwelliano —una decisión sabia— pero dejamos alegremente que las compañías privadas nos succionen datos sin ton ni son.

Y mientras nos espantamos con los titulares sobre el “puntaje social” que China asigna a sus ciudadanos según su comportamiento y lo comparamos con la serie “Black Mirror”, en el mundo occidental llevamos un buen tiempo sometidos al “puntaje del mercado.” En Estados Unidos, por ejemplo, las calificadoras de crédito nos definen en función del cumplimiento de nuestros pagos, del número de cuentas bancarias y de tarjetas a nuestro nombre, de los cheques sin fondos que hemos girado, así como del volumen y duración de nuestros créditos, entre otros indicadores. Cuanto mayor sea ese puntaje, más rápida será la aprobación del próximo préstamo que solicitemos. ¿Por qué? Porque un buen puntaje significa que somos confiables para el sistema, y en razón de ello también pagaremos menores tasas de interés. Nuestro historial crediticio dice que somos responsables. Soliciten un crédito a su banco y verán cómo allí estará su puntaje crediticio iluminando su puntaje social, es decir, su comportamiento ciudadano.

<sup>106</sup> En el capítulo 7 se aborda nuevamente este tema.

¿Cuántos años pasarán antes de que muchos otros datos terminen por alimentar nuestro puntaje social? La posibilidad de que esta recolección de información se amplíe es razonable y solo la ley podría limitarla. Vincular los historiales clínico, crediticio, educativo y penal de cada individuo, es decir, el sueño de la interoperabilidad de las bases de datos, es asunto de días o de meses, no de años. Hoy una compañía puede capturar nuestra información y vendérsela a otra y esta última a una tercera, y así sucesivamente. ¿No hemos visto nuestras bandejas de correos electrónicos llenarse de ofertas de tiendas que jamás hemos visitado o de productos que nunca hemos pedido? Son las famosas bases de datos compartidas, vendidas o ilegalmente penetradas. Equivalen hoy a lo que en un mundo analógico era el correo basura que llenaba nuestros buzones con ofertas de seguros, cupones de supermercados o nuevas tarjetas y tentadores créditos hipotecarios. Ahora ni siquiera nos interrumpen. Una vez compartida, nuestra información ayuda a estilizar la oferta de productos y servicios que aparece en nuestros navegadores. Son las consecuencias imprevistas del consumo de ciertas tecnologías. Y aunque suene a ficción, si la IA no se regula con prontitud, la posibilidad de que ciertas amenazas a la privacidad se vuelvan reales es cada vez más próxima.

Para la muestra un botón. Hace algún tiempo Google compró Fitbit, una aplicación que mide en un individuo datos como el número de pasos que da a diario, su ritmo cardíaco, la calidad de su sueño, los pisos que ha subido y bajado, y otra información sobre sus condiciones físicas. Con todo ello, la compañía creó Google Health, un repositorio de datos que conectaría a médicos, hospitales y compañías de seguros. Hoy esa empresa tiene en sus centros de datos nuestra información médica, nuestros datos de consumo, nuestras búsquedas de internet, nuestro historial de correos, nuestras conexiones sociales, y por supuesto nuestra ubicación geográfica. No nos molestan. Todo parece inocuo e inofensivo.

Y sin embargo...

Supongamos que entramos al gimnasio y activamos la caminadora con la huella del pulgar, pero no cumplimos con los 30 minutos diarios para reducir el colesterol que el médico nos recomendó y registró debidamente en nuestro historial digital. ¿Qué sucedería si el doctor compartiera esa información con el seguro y este último con el banco? ¿Y qué pasaría si todos ellos se combinaran con la descripción de los productos —ese delicioso tocino, tal magnífico jamón y una mantequilla celestial— que hemos comprado de manera virtual en el supermercado? ¿Nos podrían ajustar los intereses de nuestra hipoteca frente al riesgo de una obstrucción arterial porque no cumplimos con la recomendación del facultativo? Y el gobierno, ¿acaso no podría el aumentarnos ciertos impuestos si, con esa misma información, observa que somos un riesgo para el sistema sanitario?

Por eso, incluso allí donde pareciera primar la mejor voluntad, es necesario actuar con cierto grado de cautela. Si, como ya hemos visto en ejemplos previos, el uso de la IA puede ayudar a transformar en poco tiempo estructuras anquilosadas para beneficio de la sociedad, eso no debe distraernos del imperativo de vigilar y proteger los derechos de los individuos.

Lo cierto es que Google, Apple, Facebook, Amazon y Microsoft —conocidas con el acrónimo GAFAM— manejan un volumen de información sobre las actitudes y hábitos de la gente que supera significativamente lo que gestiona cualquier agencia pública, incluidas las de seguridad. Dicho de otro modo, cinco compañías privadas saben más acerca de los ciudadanos que sus propios gobiernos —de suyo preocupante— e incluso que los ciudadanos mismos.

La regulación de la industria privada se torna compleja porque quien resulta responsable de la definición de los marcos regulatorios y éticos —el Estado moderno— es también actor y usuario de la IA, y por lo tanto tiene incentivos para utilizar los datos y avances tecnológicos de las empresas.

### Atención, América Latina: las cámaras ya llegaron

En América Latina tenemos los primeros casos de vigilancia con IA por parte de un Estado que compra tecnología a empresas privadas. Esto ha sido implementado con muy poco conocimiento de la ciudadanía y prácticamente sin ninguna oposición. En 2019, el gobierno de la ciudad de Buenos Aires puso en operación un sistema de reconocimiento facial para la “detección, verificación, identificación y detención de personas” con orden de captura. El sistema está conectado a la Consulta Nacional de Rebeldías y Capturas, que mantiene actualizada la base de datos biométricos de los prófugos y cuenta con más de 7.000 cámaras de video de vigilancia privada dispuestas en terminales y estaciones del tren subterráneo.

En julio de 2019, Guillermo Ibarrola, de 39 años, fue detenido en un barrio de Buenos Aires mientras se dirigía a su casa. La policía le dijo que era sospechoso de un robo sucedido tres años antes en Bahía Blanca, una ciudad ubicada a más de 650 kilómetros al sur de la capital argentina. Guillermo había sido identificado mediante los sistemas de reconocimiento facial instalados por la municipalidad en el subterráneo de Buenos Aires. Estuvo detenido seis días acusado de robo. El problema era que Guillermo jamás había estado en Bahía Blanca. A la policía le tomó casi una semana darse cuenta de que había cometido un error y que no era ese Guillermo Ibarrola, sino un homónimo, el verdadero autor del delito.

En nombre de la seguridad pública, en ALC se han impulsado políticas de vigilancia masiva, entre ellas la intensificación del control de los espacios públicos. Joseph Cannataci, relator especial de ONU sobre el derecho a la privacidad, cuestionaba la justificación del sistema argentino en la falta de un análisis de necesidad y proporcionalidad: “Soy consciente de la necesidad de detener a las personas sospechosas de haber cometido delitos y llevarlas ante la justicia, pero no veo la proporcionalidad de instalar una tecnología con graves implicaciones en la privacidad para la búsqueda de una lista de 46.000 personas

que incluye [a aquellas que cometieron] delitos no graves cuya exactitud no se actualiza ni se comprueba cuidadosamente.”

Para Philip Alston, relator especial sobre la extrema pobreza y los derechos humanos de la ONU, “Los gobiernos están ideando estos planes diciendo que se trata principalmente de mejorar los servicios de bienestar y la seguridad. Sin embargo, no están limitando los fines para los que se pueden utilizar los datos ni adoptando salvaguardias esenciales para reducir su riesgo.” El avance de estas tecnologías es aún más rápido en países en desarrollo, en donde no se cuenta con suficiente madurez regulatoria en temas de protección de datos y privacidad ni con una ciudadanía informada capaz de cuestionar las acciones de los gobiernos.

La introducción de estos sistemas está ya muy extendida en toda Sudamérica,<sup>107</sup> donde empresas principalmente chinas y rusas comenzaron a instalar sistemas masivos en las ciudades desde 2018. Mientras tanto, en Europa y Estados Unidos las organizaciones de la sociedad civil ejercían una gran presión para legislar en contra de su uso, y en ciudades como San Francisco y Nueva York se las declaraba ilegales.<sup>108</sup>

Con una dotación de 50 cámaras iniciales, Asunción tiene un sistema similar al de Buenos Aires, el cual podría extenderse a Ciudad del Este, Encarnación y Pedro Juan Caballero, las otras tres ciudades principales de Paraguay. En 2019, Bolivia lanzó también su programa Bol-110, anunciado como la plataforma

<sup>107</sup> Abundaremos sobre esto en el capítulo 7.

<sup>108</sup> La experiencia internacional tampoco ha sido muy buena. Un estudio publicado en 2018 mostraba que los servicios desarrollados por las grandes empresas de tecnología (Facebook, Google y Microsoft) habían sido entrenados con una base integrada principalmente por individuos caucásicos o de tonos de piel clara. Esto hacía que el algoritmo fuera más preciso clasificando a personas blancas (con una tasa de error del 0,8%) que a personas con tonos más oscuros de piel como las mujeres afroamericanas, para las cuales la tasa de error alcanzaba el 34,7%.

de seguridad más moderna “de todo el continente.” Esta constaría de 550 cámaras, entre ellas 100 incorporadas a patrullas policiales con computadoras enlazadas a múltiples bases de datos. La IA de Bolivia es de factura china, al igual que la de Ecuador, donde el servicio de emergencias ECU 911 se expandió con un préstamo de Beijing por US\$240 millones utilizados en adquirir 4.300 cámaras de vigilancia, montar 16 centros regionales de respuesta y emplear a más de 3.000 personas en el seguimiento de las imágenes. China también le suministró aviones teledirigidos con visión nocturna y un laboratorio de IA.<sup>109</sup> En 2019, Ciudad de México adquirió más de 1.800 patrullas equipadas con equipos de reconocimiento facial y detección de placas de automóviles.

Activistas de derechos humanos de la región y de Europa han exigido que se investigue si estas soluciones conllevan un riesgo para las libertades personales. Un temor similar despiertan otras aplicaciones como las cámaras biométricas, que pueden afectar los derechos de reunión, asociación y protesta, así como la libertad de expresión y las garantías constitucionales, en particular el debido proceso y la presunción de inocencia. Por ejemplo, si llegara a realizarse una manifestación contra alguna iniciativa del gobierno de turno, con esa tecnología lo único que requieren las autoridades es peinar la marcha con los algoritmos y *voilà*: la disidencia con lista de asistencia.

El escrutinio colectivo por parte de organizaciones representativas de la sociedad civil, la exposición pública y viral de estos métodos,

---

<sup>109</sup> Colombia también hizo intentos, en apariencia infructuosos, en materia de seguridad, y desde entonces ha ampliado su foco para incluir a las empresas, a partir de la realización de un foro de IA en 2017 y el anuncio de apoyo gubernamental para entrenar a más de 3.000 ingenieros en aprendizaje automático. En Brasil, entre tanto, el sector de la Justicia dispuso que la compañía concesionaria de la Línea Amarilla del Metro de São Paulo suspendiera la captura de imágenes de los pasajeros (en esa línea viajan 600.000 personas al día), pues violaba el derecho constitucional a la intimidad y a la vida privada. Las cámaras de reconocimiento facial habían sido ubicadas junto a avisos publicitarios para recoger la reacción de los pasajeros, y reconocer su género y su franja etaria.

y la presión sobre los legisladores deben permitir una fiscalización veloz y exhaustiva del uso de las nuevas soluciones. Esto en un contexto donde la tecnología avanza mucho más rápido que los marcos jurídicos que la deben regular y de cuyos efectos nos enteramos a posteriori, cuando vemos, por ejemplo, la manipulación generalizada de una justa electoral.

Fijar estándares y normas que definan límites para el uso de los datos y la aplicación de los sistemas de inteligencia artificial en la vida civil será crucial para que la región pueda lograr una participación justa de los beneficios de la IA. La ciudadanía debe desempeñar un papel activo en la defensa de su derecho a la privacidad. Y ese es el debate que estamos en mora de dar: cómo controlamos aquello que, bien empleado, puede proporcionarnos beneficios tangibles y revolucionarios.



## CAPÍTULO 7

### De las leyes para caballos de 1865 al mundo de hoy

*Antes de cualquier diseño de inteligencia artificial se necesitan un marco ético y uno jurídico. El segundo tarda más y llega por ensayo y error. El primero debe venir incorporado en nuestra formación.*

Volvamos a la práctica ya tan común de comprar en línea. Bueno, comprar, lo que se dice comprar, no; apenas mirar, evaluar, para luego decidir. Si al final hemos desistido por la razón que fuere y cambiamos a otra página en la red para leer algo o ver un video, no deberíamos sorprendernos si en las franjas laterales de ese sitio nuevo aparecen avisos de los mismos productos que curioseamos hace unos segundos. Y luego en otro, y en uno más.

Ni hablar de las ofertas de billetes de avión a un destino elegido, armadas casi a la medida. Aparecerán gangas de último minuto en cada sitio que visitemos con mensajes de urgencia y tiempo limitado. Y qué tal la venta omnipresente de paquetes de vacaciones, ropa, electrónicos y, sí, esos suplementos para la dieta que nunca comenzaremos. Consumimos así, a la medida. Películas, libros, música, lo que queramos. Pulsamos el enlace y lo tendremos. La razón es que detrás de esas páginas hay un algoritmo que se la pasa procesando nuestros clics, evaluando sus pequeños cambios. ¿Nos sentimos rodeados? Nada raro. Los sistemas de IA nos indican desde lo que debemos comprar hasta las fotografías de nuestros amigos que vale la pena ver.

¿Por qué sucede esto? Porque somos como Hansel y Gretel que vamos dejando migas de pan virtuales por donde quiera que pasamos en internet. Cada vez que aceptamos *cookies*, ahí quedamos pintados, con nuestras selecciones y nuestros gustos convertidos en trocitos ínfimos de información o *bits* que van a nutrir bases de datos donde motores inteligentes procesan y deciden por nosotros.

Del otro lado del monitor el usuario ve la utilidad de la automatización. Pocos son los que protestan porque detrás de cada servicio haya un algoritmo de personalización o recomendación a la medida de nuestros gustos. Los que somos mayores de 30 recordamos la experiencia de entrar a un Blockbuster<sup>110</sup> a buscar una película a mano, leer la sinopsis,

<sup>110</sup> Franquicia estadounidense especializada en alquiler de películas a través de tiendas físicas y servicios por correo.

rentarla y esperar a llegar a casa para verla y entonces verificar si elegimos bien. Ahora, en la época de Netflix, YouTube y Spotify hay menos sorpresas; basta con entrar a la plataforma y tendremos una lista interminable de recomendaciones ordenadas por su afinidad con nuestros gustos.

Nos hemos convertido en dedicados filántropos de nuestra propia información; no tenemos problema alguno para compartir nuestros datos a diestra y siniestra a cambio de que nos ofrezcan experiencias de consumo más gratas o menos frustrantes, más eficientes y ajustadas a nuestros deseos. Obvio que no cobramos por esa información acerca de lo que somos y es poco lo que percibimos a cambio. La asimetría informativa nos condena. Quienes usan nuestros datos saben cuánto valor pueden extraer de ellos; para nosotros son tonterías: entrada tras entrada de Facebook, trinos a granel, búsquedas en Google. ¿Qué valor puede tener eso?

Por otro lado, tenemos la promesa del gobierno digital. Entendemos la necesidad de que el Estado vaya más allá de las políticas sociales reactivas que se ocupan de lo visible, lo actual y lo tangible. Queremos un Estado capaz de asumir una función predictiva y anticipar problemas que afectarán a miles de millones de ciudadanos; un Estado flexible que pueda modificar su respuesta en tiempo real a un mundo que cambia permanentemente.

Ahora estamos más conscientes de cuáles son las implicaciones de ese Estado digital que, para **poder saber**, requiere información: la nuestra. Y es evidente que no deseamos que el Estado tenga acceso a **toda** nuestra información. Nadie quiere ser sujeto de vigilancia, sometido a control panóptico, monitoreado al detalle.

¿Cómo, entonces, abordar estos riesgos? ¿Qué medidas de mitigación podemos adoptar? ¿Debemos ceder a la discusión entre libertad/privacidad versus seguridad/poder atendiendo a un solo extremo? ¿Hemos de tolerar a un súper Estado o a las súper empresas para lograr que la tecnología nos ayude

a superar las décadas de atraso, negligencia, desinterés e incapacidad que han definido nuestro trasegar histórico? ¿O será que podemos construir algo intermedio, sano, activamente ciudadano? ¿Es posible el equilibrio?

Creer que la responsabilidad de calcular los riesgos del uso de la tecnología debe recaer en el ciudadano es un sinsentido, tanto como lo sería darle al consumidor la tarea de evaluar los componentes químicos de un medicamento para saber si pone en peligro o no su vida, o como permitir que las compañías de fármacos distribuyan productos riesgosos sin control de calidad, sin pruebas de efectos secundarios, sin recetas médicas. La IA debe tener una regulación clara. Como vimos, los riesgos son variados y las promesas de los sistemas inteligentes difícilmente se materializarán si no se definen marcos de gobernanza claros. Investigadores, desarrolladores, pensadores de diversas procedencias, y la industria misma, pueden sentar las bases técnicas y los límites de los marcos ético-filosóficos dentro de los cuales se desarrollará el EBDI y se desplegarán los sistemas de IA. Los legisladores deberán dar forma jurídica a estas discusiones.

Para ello será necesario definir los estándares de calidad de estos sistemas y asegurarnos de que las leyes que ya existen —como las pertinentes a la igualdad de oportunidades y a la no discriminación— se respeten y apliquen plenamente. La IA únicamente servirá como soporte de la política pública si se realiza una vigilancia atenta sobre su calidad y se dispone de mecanismos idóneos de mitigación de riesgos en su diseño, implementación y monitoreo. Habrá que comenzar entonces por retomar y revisar integralmente las leyes de protección de datos personales, de intercambio y almacenamiento de información, y por analizar los marcos jurídicos de responsabilidad y explicabilidad que rigen el uso de sistemas de IA.

No es demasiado tarde para plantearnos, en esta era digital, un nuevo contrato social que cubra derechos, responsabilidades y expectativas, así como el conjunto de normas explícitas que

defina qué pueden y qué no deben hacer las empresas y los gobiernos. Pero, ¿cómo determinar y elaborar estas normas? ¿Cómo reglamentar el desarrollo y uso de los sistemas de IA? ¿Qué grados de supervisión legal y ética —quiénes y cuándo— serán necesarios? ¿Y de qué modo se regula para no hacer inviable el desarrollo de los proyectos y asegurar que se mantengan vivos los incentivos para innovar?

El problema principal es que nadie sabe muy bien por dónde comenzar a regular la IA. Las disrupciones técnicas crean nuevos desafíos éticos y jurídicos cuyos impactos pueden no ser comprensibles de inmediato, ni siquiera para quienes desarrollan la tecnología. Asimismo, dejan preguntas abiertas cuyas respuestas no se hallarán sino hasta cuando haya transcurrido un tiempo y se hayan cometido los errores que exijan reconsiderar y perfeccionar.

De cualquier manera, los actores deben avanzar en procura de resolver estas contradicciones y disyuntivas sobre la marcha. Especialmente en lo que concierne al Estado, donde inclusive los responsables de áreas que podrían tener soberanía regulatoria sobre la IA desconocen la genealogía de esta bestia que ha aparecido en el horizonte. En su propuesta de modelo para una gobernanza sólida de la IA, los investigadores Urs Gasser y Virgilio Almeida señalan tres capas:<sup>111</sup> una técnica, una ética y una sociojurídica.<sup>112</sup>

<sup>111</sup> Aunque Gasser y Almeida describen la progresión de su desarrollo de manera cronológica en las tres etapas mencionadas, este ha sido cíclico, dado que las aplicaciones técnicas evolucionan y se alimentan conforme se definen las preocupaciones éticas.

<sup>112</sup> La capa técnica, que constituye la base del ecosistema de gobierno de la IA, está relacionada con los algoritmos y datos a partir de los cuales se construye; la capa ética plantea los principios y preocupaciones de alto nivel que deberían aplicarse a todos los tipos de sistemas de IA; la capa sociojurídica, entre tanto, aborda el proceso de creación de instituciones y asigna responsabilidades para regular esta tecnología.

### Lo técnico: controlar el sesgo y crear herramientas de explicabilidad

Gasser y Almeida ubican la capa técnica, compuesta por datos y algoritmos, en la base de la gobernanza de una IA responsable.

Desde una perspectiva técnica, los puntos más importantes a tener en cuenta para lograr un sistema de IA robusto son tres: (1) la atención y mitigación de las distintas fuentes de sesgo, y sus implicaciones en términos de equidad, justicia y discriminación en los grupos vulnerables; (2) su mayor o menor capacidad de explicabilidad de las predicciones individuales; y (3) la transparencia y auditabilidad de sus procesos.

Estos aspectos deben ser tomados en cuenta a lo largo del desarrollo de la IA, es decir, desde la conceptualización del sistema, la recolección de datos, su limpieza y procesamiento, pasando por el entrenamiento y ajuste de los algoritmos, su evaluación e implementación, hasta el monitoreo de su aplicación en el mundo real.<sup>113</sup>

En el capítulo 5 abordamos de forma más detallada estas preocupaciones y los efectos que pueden tener en los resultados de los modelos. Aquí es importante recalcar que, cuando se diseñan políticas públicas, la IA no debe partir de la tecnología, puesto que es su interrelación con las otras dos capas (la ética y la sociojurídica) lo que justifica y define los objetivos particulares que deben cumplir los algoritmos durante su desarrollo.

En lo que se refiere al aspecto de atención y mitigación de sesgos, durante el ajuste del modelo los responsables de la

<sup>113</sup> Existen ya manuales que describen buenas prácticas y metodologías para evaluar técnicamente un sistema de toma y soporte de decisiones que utiliza aprendizaje automático. Un ejemplo de ello es el "Manual de IA responsable" del Banco Interamericano de Desarrollo, enfocado en la aplicación de IA en la política pública.

formulación de políticas deben comenzar por definir cuáles son los grupos protegidos y el criterio de justicia algorítmica a utilizar. Esto por cuanto, si bien desde un punto de vista puramente técnico el objetivo de un modelo de aprendizaje automático es generalizar tal aprendizaje para toda la base de datos —lo que en el agregado produce resultados de precisión muy buenos—, este mismo puede resultar injusto para grupos vulnerables y poco representados. Los errores deben analizarse bajo la lupa de los objetivos específicos del sistema y, si se quiere priorizar la atención a una población determinada, es necesario proporcionar una justificación explícita.

Los temas relativos a la **explicabilidad**, por un lado, y a la **transparencia** y **auditabilidad**, por el otro, refieren a requisitos que veremos más adelante en este capítulo cuando abordemos la capa sociojurídica. Estos aluden a la obligación de las entidades o instituciones que utilicen sistemas de IA de construir procesos trazables y de dar explicaciones sobre la manera en que fue tomada una decisión, como por ejemplo por qué se otorgó un crédito o por qué se dio o se negó un beneficio del Estado.

En muchos casos los algoritmos son poco interpretables, en el sentido de que resulta difícil dilucidar la serie de pasos o condiciones lógicas que llevaron a tomar una decisión y no otra.<sup>114</sup> Cuando existe una obligación legal, esto condiciona el tipo de sistemas que se pueden emplear y exige acudir a algoritmos más sencillos que, aunque sacrifiquen precisión, sean claramente explicables.

<sup>114</sup> Algoritmos de alta complejidad como las redes neuronales profundas, son en principio poco transparentes en cuanto a cómo se hacen las predicciones subyacentes. Existen ya propuestas metodológicas de explicabilidad, pero es un área sujeta a investigación y desarrollo.

### Lo ético: ¿qué tan responsable es su algoritmo?

Definir la escuela a la que adscribirá la IA es definir qué tipo de EBDI vamos a tener.<sup>115</sup> Todo cuanto suceda a partir de allí quedará plasmado en el marco ético-filosófico seleccionado. En un sentido literal, los sistemas de soporte y toma de decisiones se convierten en derivados sucedáneos, producto de la cosmovisión que predomine.

Gasser y Almeida sostienen que para desarrollar la capa ética existen dos preguntas fundamentales: (1) en qué podemos aplicar la IA y (2) cuáles son las reglas de su desarrollo.

Lo primero que se debe tener en cuenta es que no hay desarrollo científico que se encuentre libre de intereses económicos o políticos; si un proyecto tiene financiamiento es porque existe alguna industria, gobierno o grupo de interés que promueve una agenda.

Por ejemplo, la reacción atómica puede usarse en diversas aplicaciones beneficiosas: desde la obtención de imágenes médicas hasta la radioterapia, pasando por los detectores de humo y la producción de energía renovable. Pero ese mismo potencial se ha utilizado para producir y usar armas de destrucción masiva.

En tal sentido, es necesario pensar en las implicaciones, aplicaciones y consecuencias no intencionadas del desarrollo de un sistema inteligente. Volviendo al ejemplo anterior, y a la necesidad de contar con una regulación que garantice los usos previstos, vemos que aunque hay naciones que se encuentran desarrollando este tipo de tecnología, las directrices nacionales y los acuerdos internacionales de no proliferación han demostrado ser un marco relativamente sólido para establecer y mantener las expectativas de un comportamiento responsable.

<sup>115</sup> La distinción de concepciones del Estado de bienestar se trató en el capítulo 2.

Por esta razón, la construcción de la capa ética suele iniciarse con la definición de los principios o códigos que en ella emplearán los grupos y sectores que desarrollan la tecnología, pues son estos los que la conocen y pueden proporcionar orientaciones de alto nivel sobre lo que debe o no hacerse en un campo de investigación o de aplicación del conocimiento.<sup>116</sup> Existen buenas prácticas que pueden servir como modelos de construcción ética para definir los límites de una tecnología. Uno de los campos con reglas de exclusión más firmes es la experimentación genética, donde están prohibidas todas las formas de clonación humana en la medida en que son incompatibles con la dignidad y la protección de la vida de nuestra especie.

Llevamos miles de años modificando los seres que nos rodean. Las manzanas, el maíz y los animales domesticados se parecen poco a sus ancestros silvestres. Creamos alteraciones genéticas mediante la selección de mutaciones por razones de utilidad e incluso de percepción de la belleza.<sup>117</sup> En las últimas décadas, las capacidades tecnológicas dieron un salto significativo y entonces creamos alimentos genéticamente modificados, insulina y hormonas de crecimiento, entre muchas otras cosas. Cuando logramos desarrollar el método CRISPR<sup>118</sup> —que permite hacer ediciones precisas del ADN alterando su secuencia

<sup>116</sup> Este tipo de reglas proactivas suelen observarse en otras áreas de la vida social, y a menudo los legisladores esperan a que la ciudadanía reaccione a un cambio antes de establecer normas que regulen una industria o sector.

<sup>117</sup> El número de razas de perros que existe es una clara muestra de la capacidad de la ingeniería genética que hemos desarrollado por años, en donde pequeñas diferencias (entre 8 y 15 genes) pueden crear un chihuahua —que puede pesar un kilo— o un gran danés, que puede llegar a pesar 80 kilos.

<sup>118</sup> La técnica CRISPR (sigla en inglés de Clustered Regularly Interspaced Short Palindromic Repeats) emplea una proteína (Cas9) a la que se dirige con guías hacia zonas del ADN seleccionadas para cortarlas, como con una tijera molecular; luego se unen los extremos cortados, se inactiva el gen y se incorporan moldes de ADN, editando las 'letras' como los científicos quieren. Es un mecanismo de corrección de genomas tomado de las bacterias.

a un costo 150 veces menor que en el pasado y de forma mucho más sencilla—, nos encontramos con que podíamos curar enfermedades genéticas o detectar y sustituir células cancerosas, entre muchas otras aplicaciones en desarrollo. Sin embargo, es una técnica muy nueva y el conocimiento que tenemos sobre el ADN es aún limitado, de modo que todavía es difícil prever el efecto futuro de variaciones muy pequeñas del código genético.

De ahí que seamos renuentes a saltarnos las barreras ético-jurídicas en materia de biología sintética acordadas en consensos más o menos globales. En 2019, en China el científico de la Southern University of Science and Technology de Shenzhen, He Jiankui, dijo haber experimentado con embriones humanos para hacerlos menos susceptibles a contraer el HIV/SIDA. De inmediato, la comunidad científica planteó preocupaciones éticas, especialmente en referencia al nivel de consentimiento que había obtenido He Jiankui de los padres de los bebés, y al conocimiento que tenían esas personas de los riesgos y el grado de transparencia del procedimiento de edición de genes. No había mucha claridad sobre el asunto. A principios de 2020, He Jiankui fue sentenciado a tres años en prisión.

En el campo de la IA existen numerosas aplicaciones que los especialistas consideran poco deseables y por lo tanto creen que deben ser tratadas, por decirlo de alguna manera, con las mismas precauciones y cuidados que los experimentos genéticos. Sin embargo, a diferencia de lo que sucede con la ingeniería genética, los esfuerzos mundiales encaminados a establecer los límites a la investigación en IA son todavía muy incipientes.

En 2017, por ejemplo, Google consiguió un contrato con el programa Project Maven del Pentágono para desarrollar un sistema de IA que pudiera ayudar a reconocer personas y objetos capturados en filmaciones de aviones militares no tripulados.

Antes de firmar el contrato, Google estaba plenamente consciente de las aplicaciones de la IA para vigilancia<sup>119</sup> y de sus implicaciones éticas. El *New York Times* tuvo acceso a

un intercambio de correos entre altos ejecutivos de Google en relación con el proyecto y reveló que uno de ellos pedía indicaciones sobre cómo compartir con el personal la noticia de la vinculación de Google con el Pentágono. "Evita a toda costa cualquier mención o implicación de la IA," le respondieron. "La IA usada como arma es probablemente uno de los temas más sensibles, si no el más. Hablar de ello equivaldría a lanzarle un trozo de carne roja a los medios de comunicación para que ataquen a Google."

Lo cierto es que cuando se hizo público el contrato, más de 3.000 empleados de Google recibieron la noticia con indignación. Estaban convencidos de que la empresa no debía involucrarse en ningún trabajo con las fuerzas armadas, y que los modelos de IA no deberían emplearse en la identificación de objetivos militares. Un año después, en 2018, Google firmó un compromiso público de que no desarrollaría IA relacionada con armamentos, con proyectos de vigilancia que violen normas aceptadas internacionalmente, o con planes que contravengan "principios ampliamente aceptados de derecho internacional y sobre los derechos humanos."

Esta equivocación estratégica impulsada por incentivos económicos, así como el subsiguiente cambio de posición en respuesta a las presiones de los propios empleados de la compañía y de la opinión pública, muestran la necesidad de regular estos temas.

Cabe notar, sin embargo, que aquí la autorregulación de las empresas no es ni será suficiente. Aun así, en un contexto donde la normativa de los Estados avanza lentamente se necesita crear espacios de discusión apropiados para tratar estos temas, de modo que las empresas del sector puedan

<sup>119</sup> Google proporcionó al Departamento de Defensa su API TensorFlow para ayudar en el reconocimiento de objetos. Con el tiempo, el Pentágono se proponía convertir sus depósitos de vídeo en inteligencia destinada a la seguridad nacional y a la vigilancia.

aceptar sus errores, cambiar de opinión, aprender y corregir, y difundir las mejores prácticas disponibles para que otras compañías reproduzcan el método.

En parte, y a la luz de esa discusión original, a mediados de 2020 la empresa IBM envió una carta al Congreso de los Estados Unidos donde afirmaba que no estaba dispuesta a desarrollar o a investigar algunas tecnologías hasta tanto no existiera regulación en la materia. Dice el documento: “IBM se opone firmemente y no aprobará el uso de ninguna tecnología [de reconocimiento facial], incluida la [...] ofrecida por otros proveedores, para la vigilancia masiva, la elaboración de perfiles raciales, la violación de los derechos humanos y las libertades básicas, o cualquier otro propósito que no sea coherente con nuestros valores y principios de confianza y transparencia” [...] Creemos que este es el momento de iniciar un diálogo nacional sobre si la tecnología de reconocimiento facial debe ser usada por las fuerzas del orden y sobre cómo hacerlo.”

Ya hemos señalado que la inteligencia artificial no puede ser ética en sí misma, dado que las máquinas carecen de autonomía moral. Sus acciones están al servicio de las personas. Todo software inteligente opera entonces en el marco ético que le imprimen sus programadores, y en ese sentido son la materialización de ideas humanas y representan sus activos culturales, su carga ideológica y sus prejuicios.

Existe, por lo tanto, la necesidad imperiosa de que los modelos de IA sean entrenados bajo esquemas de información y datos sostenidos por un mapa de principios éticos y morales que a la postre se traduzcan en leyes y normas.

Diversas entidades internacionales, entre ellas la Organización de las Naciones Unidas, han alertado sobre los riesgos de parcialidad y sesgo en los sistemas de IA, los cuales se originan en los datos usados para su entrenamiento, en las decisiones conscientes o inconscientes de los desarrolladores durante el entrenamiento del algoritmo, y en la forma en la que sus resultados son usados para tomar decisiones.

Respetar la dignidad humana es una obligación inapelable. Un principio clave del ecosistema ético-filosófico de la IA debe ser la reducción o eliminación de sesgos en sus motores, con el fin de garantizar que no se discrimine o excluya a ciertos grupos. Si estas máquinas toman decisiones basadas en objetivos de eficiencia o sesgados por prejuicios, se vulnerarán muchos valores de convivencia democrática. Y no precisamente porque “tengan conciencia”, sino porque así se las programó, con intención o sin ella.

La introducción de la IA al aparato del Estado lo reconfigura. La sola terminología anuncia la aparición de una nueva especie en el horizonte. Explicabilidad, automatización inclusiva, adaptación de dominio y evaluación de equidad en subgrupos son el nuevo idioma de un nuevo paradigma. Nuestras burocracias, asociadas al papel o a formas muy anticuadas de tecnología —con redes que todavía no dialogan entre sí en muchos países—, tendrán que afrontar serios problemas operativos. Si al sector privado vinculado a actividades como el comercio, la industria o los servicios (donde la competencia exige adaptación constante) todavía le cuesta reentrenar a sus trabajadores para que transiten desde tareas individuales o complementarias hacia modelos colaborativos, imaginemos cuán complejo resultará esto en el sector público, donde las transformaciones culturales son más lentas. Las administraciones públicas deberán migrar hacia mecanismos transdisciplinarios de trabajo y operar colaborativamente de manera reticular.

### *Hipócrates y las leyes de Asimov*

El juramento hipocrático es uno de los códigos de ética más antiguos y conocidos. Su texto original se atribuye a Hipócrates, un galeno del Siglo de Oro griego que realizó tantas contribuciones a su campo que se le considera el padre de la medicina. De hecho, fue Hipócrates quien introdujo el término “epidemia” al vocabulario médico.

El juramento hipocrático clásico incluía el compromiso de practicar la medicina para el bien de los enfermos, sin hacer

daño o cometer injusticias, y de respetar la confidencialidad del paciente. En la versión más antigua que se conserva, y que data del año 275 AC, dice lo siguiente: “Todo lo que vea y oiga en el ejercicio de mi profesión, y todo lo que supiere acerca de la vida de alguien, si es cosa que no debe ser divulgada, lo callaré y lo guardaré como secreto inviolable.”

Hoy día los médicos aún hacen este juramento, cuya esencia se mantiene intacta no obstante algunas variaciones sutiles. Es más, muchos países han incorporado las obligaciones de los médicos a la legislación, como ocurre por ejemplo con la HIPAA (Health Insurance Portability and Accountability Act) de Estados Unidos, una ley federal que requirió crear estándares nacionales para proteger la confidencialidad de la información sanitaria de los pacientes de manera que no pueda ser divulgada sin su consentimiento o conocimiento.

Cuando pensamos en las reglas de la inteligencia artificial, la primera referencia que nos viene a la mente la heredamos de las tres leyes de la robótica del relato de ciencia ficción Runaround de la serie “Yo Robot” escrito por Isaac Asimov en 1942: “(1) un robot no hará daño a un ser humano o, por inacción, permitirá que un ser humano sufra daño; (2) un robot debe cumplir o realizar las órdenes dadas por los seres humanos, excepto si estas órdenes entran en conflicto con la Primera Ley; (3) un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la Primera o Segunda Ley.” Por supuesto que no es aconsejable regular la realidad con productos de la ficción de hace más de 50 años.

En 2017, la industria de IA intentó discutir algunos esbozos, primero en una conferencia de investigadores, científicos y líderes en Asilomar, en Silicon Valley sobre la bahía californiana de Monterrey, y posteriormente en Barcelona. En California analizaron, debatieron y redactaron los Principios de Asilomar, un conjunto lineamientos, normas y valores que se esperaba fueran adoptados por todos los involucrados en el desarrollo de la IA. De Barcelona surgió la “Declaración de Barcelona para el desarrollo y uso apropiados de la inteligencia artificial”,

que giró en torno a los mismos temas: prudencia, fiabilidad, responsabilidad de los creadores, alcance de la autonomía de las máquinas y la función de los humanos durante su uso.<sup>120</sup>

Hacia 2020 ya existía un número importante de principios con los que se persigue una IA “ética,” los cuales proporcionan una orientación de alto nivel sobre lo que debe o no hacerse en su desarrollo. El centro Berkman Klein de la Universidad de Harvard realizó un metaanálisis que sintetiza los aspectos clave presentes en casi todas estas discusiones. De ese trabajo se desprende que los siguientes son principios necesarios para una IA ética, justa y equitativa:

- *Privacidad:* Las personas afectadas deben tener control y conocimiento de sus datos y de las decisiones que se toman con ellos.
- *Rendición de cuentas:* Deben existir mecanismos para garantizar que la responsabilidad por los efectos nocivos del uso de la IA se defina correctamente y se ofrezcan recursos adecuados de compensación.
- *Seguridad y protección:* Los sistemas de IA deben ser seguros, funcionar según el objetivo definido y estar protegidos de ataques malintencionados y usos no autorizados.

<sup>120</sup> El texto de la Declaración de Barcelona es elocuente y revela los muy complejos usos de la IA de riesgo elevado: “Existe una creciente preocupación por los *chatbots* de IA y otros tipos de sistemas de mensajería automática que operan en internet y en las redes sociales, diseñados para la manipulación de la opinión política, la desinformación a través de la propagación de hechos falsos, la extorsión u otras formas de actividad maliciosa. Eso es peligroso para los individuos y desestabiliza nuestra sociedad. Estos *chatbots* pretenden ser humanos y no revelan la identidad de quienes están detrás de ellos. El uso de la IA ha hecho que estos *chatbots* sean lo suficientemente realistas como para que los usuarios desprevenidos no puedan hacer una distinción y se engañen. Una posible solución a este problema podría ser dejar en claro de forma obligatoria si una interacción se origina en un ser humano o en un sistema de IA, y que, en el caso de un sistema artificial, los responsables de este puedan ser rastreados e identificados. Esta solución posiblemente podría implementarse mediante un sistema de marcado de agua y volverse obligatoria en Europa.”

- *Transparencia y explicabilidad:* La IA debe diseñarse y aplicarse de modo tal que permita la supervisión mediante la traducción de sus operaciones en resultados inteligibles y del suministro de información sobre dónde, cuándo y cómo está siendo utilizada.<sup>121</sup>
- *Equidad y no discriminación:* Los sistemas deben diseñarse y utilizarse para maximizar la equidad y promover la inclusión.
- *Control humano de la tecnología:* Las decisiones importantes deben quedar sujetas a revisión humana.
- *Responsabilidad profesional:* El profesionalismo y la integridad son cruciales; debe consultarse a los grupos interesados, prever los efectos de largo plazo y planificar respuestas en caso de que se presenten.
- *Promoción de los valores humanos:* Los fines y medios de la IA deben corresponderse con nuestros valores fundamentales y, en general, promover el bienestar humano.

¿Es esto suficiente? Una vez más, no. Deberíamos prestar atención a esta advertencia. Las empresas tienen incentivos que pueden ir en dirección contraria al bienestar general, y como los esquemas autorregulados —caso de todos los documentos de principios de la IA— no son jurídicamente vinculantes, no existe una entidad que pueda sancionar su violación. En suma, son un acto de buena voluntad, no una ley.

Leonard Haas y Sebastian Gießler de Algorithm Watch, una organización independiente que monitorea el mercado de IA, advierten que no debemos creer que estos principios son un instrumento idóneo para contrarrestar los usos perjudiciales de la IA. Sin obligaciones claras, las empresas pueden tener

<sup>121</sup> Los principios de privacidad y rendición de cuentas están en el 97% de los documentos en discusión en la industria; los de seguridad en el 81%; los de transparencia y explicabilidad en el 94%; los principios de justicia y no discriminación se encuentran en todos los documentos; los de control y promoción de valores humanos en el 69%; y los de responsabilidad profesional en el 78%.

incentivos para firmar acuerdos y principios que solo utilizarán como herramienta de mercadeo y relaciones públicas sin hacer cambios de fondo.

Sin embargo, esto no significa que su desarrollo no sea deseable. Ya hemos señalado cómo la industria avanza a un ritmo más acelerado que la acción gubernamental y que los esquemas regulatorios. De allí que incentivar la autorregulación de las empresas sea un paso inicial importante en la generación de buenas prácticas, pues sirve como instrumento para reflejar los principales problemas que encuentran los desarrolladores de tecnología de punta, al tiempo que constituye un antecedente para la construcción de la capa legal.

### Lo legal: para morder se necesitan dientes

Es mejor saber desde el principio que el camino será largo y empinado. Las normas y leyes aplicadas a las tecnologías innovadoras no suelen surgir inmediatamente. En no pocas ocasiones comienzan con errores y tropiezos significativos. Pero sin capa legal no hay nada que comprometa a las instituciones, empresas y gobiernos a seguir procesos y buenas prácticas que nos conduzcan a una IA robusta y responsable.<sup>122</sup>

La IA que queremos debe facilitar el empoderamiento de las personas para que tomen decisiones informadas; en tal sentido, no deben existir programas que no cuenten con la supervisión adecuada. Cada software debe estar dotado de robustez técnica y seguridad para que se reduzcan las fallas y/o se prevea un procedimiento alternativo en caso de error. La precisión y

<sup>122</sup> En el capítulo 5 se mencionó dónde se presentan los principales retos: (1) en la conceptualización y diseño, donde se requiere partir de un análisis de necesidad y proporcionalidad; (2) en el ciclo de vida, con la identificación de sesgos e inequidades causados por los datos, algoritmos o por decisiones tomadas durante el desarrollo de la IA; (3) en la rendición de cuentas, consentimiento, protección de datos, interpretabilidad y explicabilidad, responsabilidad y transparencia. También se deben definir estándares de ciberseguridad, interoperabilidad y gobernanza e intercambio de datos.

la confiabilidad son esenciales y deben ser reproducibles para controlar y minimizar los daños involuntarios. Todo ello se debe acompañar del respeto por los principios de equidad y no discriminación (evitar el sesgo injusto), los cuales deberán quedar plasmados en diseños que favorezcan la diversidad.

Los programas de IA han de respetar por sobre todo la privacidad de los datos personales, cuyo manejo debe ser seguro y garantizado. La gobernanza de la información y los datos debe permitir el acceso legítimo de los propietarios de esos activos personales, es decir, los individuos. Los mecanismos de trazabilidad pueden contribuir a la transparencia, del mismo modo en que la explicación clara del funcionamiento de los modelos facilita la confianza de quienes participan en ellos.

Los sistemas de IA no existen en un vacío legal. Son variadas las regulaciones que les competen y que abarcan estos principios, entre ellas las que tienen que ver con la protección de datos personales, los derechos humanos y la responsabilidad civil de quienes toman decisiones. Por supuesto, algunas de estas regulaciones tendrán que transformarse para considerar los cambios que trae la nueva revolución tecnológica. Los conceptos de **privacidad, responsabilidad y diligencia debida, explicabilidad y transparencia** figuran entre los principales retos que la comunidad legal tendrá que abordar y actualizar. La tarea no será fácil.

### *Privacidad, te extraño*

El enunciado es simple. Las personas afectadas por un diseño de IA deben tener control y conocimiento de sus datos y saber cuáles son las decisiones que se toman con ellos.

La IA ha creado todo un nuevo espacio para resolver asuntos variados, pero también para complicarlos. Información a la que antes no se tenía acceso o que jamás había sido analizada ahora será parte de las operaciones del Estado, y ese es un escenario donde los límites son todavía muy difusos. ¿Hasta dónde queremos que llegue la captura de nuestros datos

personales? ¿Estamos de acuerdo con que las autoridades tengan información sobre nuestra alimentación, salud y régimen de ejercicio que pueda terminar en las oficinas de la autoridad sanitaria, por ejemplo, solo porque ingresamos los datos respectivos en una aplicación de uso personal sobre nuestro estado físico? ¿Queremos que se sepa sobre nuestros hábitos de ahorros porque un día los bancos y las casas de inversión deciden abrir sus archivos al fisco? ¿O queremos que se comparta la información sobre nuestra vida privada o pública cuando una red social acepta entregar las llaves y abrir los candados a las fuerzas de seguridad? Y una vez que el Estado tenga esa información, ¿dónde la almacenará?, ¿cómo la utilizará?, ¿por cuánto tiempo tendrá acceso a ella? y ¿cómo la borrará una vez logrado su objetivo?

Un criterio básico de protección que la sociedad debe exigir es que la recolección de datos personales y el desarrollo de sistemas de toma y soporte de decisiones apoyado en la IA respeten la regla de **privacidad por diseño**. Esto implica que solo se capturará la mínima información personal indispensable, que esta se destinará solo al propósito específico declarado por el sistema y aceptado por el usuario, y que sus datos serán anonimizados. Es importante evitar —y se debe penalizar— una recolección desproporcionada, especialmente cuando el objetivo perseguido es generar confianza colectiva en los sistemas inteligentes.

Aquí cabe señalar, sin embargo, que la idea de privacidad va mutando. La nuestra no es idéntica a la que tenían las generaciones anteriores. Al igual que la inteligencia artificial, la privacidad es un concepto dinámico que va sufriendo modificaciones con los cambios sociales, con el impacto significativo de ciertos hechos que marcan hitos, con la transformación de las prácticas culturales, y también con las capacidades tecnológicas de captura, análisis y campo de uso —o mal uso— de los datos.

El concepto jurídico de privacidad nació a fines del siglo XIX, alrededor de 1890, durante la Segunda Revolución Industrial.

Louis Brandeis, futuro juez de la Corte Suprema de Justicia de Estados Unidos veía como un riesgo el uso de la tecnología por parte del gobierno. Junto con Samuel Warren, publicó un artículo en *Harvard Law Review* en el que describían el derecho a “ser dejado en paz”, un planteamiento con el que se buscaba proteger la información privada para que no pudiera ser difundida a través de los medios de comunicación.

Ya en su calidad de juez de la Corte Suprema, Brandeis redactó una opinión de disenso sobre un caso ocurrido durante la prohibición del alcohol en los Estados Unidos, en el cual la policía había intervenido los teléfonos de un presunto contrabandista sin obtener primero una orden. La Corte decidió que la evidencia podía ser utilizada, ya que en ese momento consideró que la llamada “regla de exclusión” —que evita que se use evidencia obtenida ilegalmente— solo aplicaba a confiscaciones y evidencia física.

Brandeis argumentó entonces que esa acción era ilegal, dado que no existía diferencia alguna entre escuchar una llamada telefónica y abrir una carta sellada. Brandeis sabía que el espionaje gubernamental no terminaría con las intervenciones en los teléfonos y que, de continuar así, no habría espacio íntimo que fuera inaccesible para la tecnología de vigilancia. Se habría sorprendido al saber cuán proféticas resultaron sus palabras.

A mediados del siglo XX, el jurista Alan Westin reorientó la concepción de privacidad como el derecho a decidir qué información sobre uno mismo puede comunicarse a los demás y en qué circunstancias. En este sentido, nuestra información personal pasa a ser un bien del que como individuos podemos disponer y compartir. Este poder de decisión sobre la información propia fue retomado más recientemente por otro jurista, Adam D. Moore, quien sostiene que la privacidad debe ser definida como el **derecho a controlar el acceso** a nuestra información personal.<sup>123</sup>

<sup>123</sup> El argumento de Moore tiene tres actores: individuos, gobiernos y corporaciones. Estos dos últimos son indispensables para la discusión sobre aparatos de control y manipulación en la era digital.

“Estar seguro a este nivel”, escribió, “es tener soberanía sobre un dominio privado; es ser libre de interferencia injustificada de otros individuos, corporaciones y del gobierno. En este nivel, privacidad y seguridad van de la mano.” Para Moore, la manipulación del individuo es posible cuando los límites de la privacidad se desdibujan.

Esta concepción hace que sea el individuo el responsable de sus datos; sin embargo, un primer problema es que poco entendemos los riesgos que conlleva su disponibilidad. Las secciones de las páginas web donde debemos aceptar los términos y condiciones, por ejemplo, están diseñadas para que pasemos allí el menor tiempo posible. En la era digital no existe el consentimiento informado, porque para tenerlo debemos ser capaces de entender qué van a hacer con nuestros datos, cómo y quién.

Google tiene una política de privacidad de 4.000 palabras que, como casi todos sabemos, muy pocos leemos. Veamos un ejemplo. En sus acuerdos de privacidad, Google informa al usuario que cuando una persona sube contenidos, la compañía y sus socios obtienen una licencia de hospedaje, almacenamiento y creación de trabajos derivados de esa información.

“Tus búsquedas en la red, tu impresión de voz y tu rostro en un videochat de Google Chat, tus correos electrónicos enviados por Gmail son todos contenidos que has enviado a sus Servicios,” advierte James Ward, un abogado especializado en seguridad de datos y privacidad consultado por el *Times* para The Privacy Project. “Es así como todos hemos ‘consentido,’ sin tener la menor idea de que lo hemos hecho”. Lo mismo sucede con los servicios que aprovechamos; si suministramos información, Google y sus asociados pueden usarla.<sup>124</sup>

<sup>124</sup> En prácticamente todos los países de América Latina, el consentimiento se puede dar de forma tácita; basta con que el titular no manifieste explícitamente oposición a su uso, y siempre y cuando no se trate de datos sensibles.

Es cierto que los empleados de estas empresas en general no tienen permiso para leer directamente nuestros correos o información; sin embargo, esto no evita que se entrenen múltiples algoritmos con estos datos para crear sistemas inteligentes cuya propiedad intelectual les pertenece únicamente a estas empresas.

Pero en últimas, ¿qué podrían importar los comentarios en redes sociales, los “me gusta” que tan generosamente damos a las películas e incluso nuestro historial de búsquedas?

La OCDE describe los datos personales como toda aquella información relacionada con una persona viva identificada o identificable. ¿Qué significa esto hoy día, cuando los datos con los que se puede identificar a un individuo en particular han cambiado con las capacidades analíticas? Anonimizar un conjunto de datos, es decir, eliminar las columnas de lo que se considera información personal, no es suficiente en la actualidad.

En 2009 ocurrió un caso emblemático que ilustra la vulnerabilidad de las medidas que ofrecen las empresas para la protección de nuestros datos personales. Netflix organizó ese año una competencia abierta por la que pagaría US\$1 millón al equipo que pudiera obtener el mejor sistema de recomendación de series. Para este concurso publicaron 100 millones de evaluaciones de 500.000 usuarios únicos. La medida de anonimización consistió en suprimir los campos de información personal y nombres de usuario; incluso se eliminaron los títulos de las películas. En ambos casos estos datos fueron sustituidos por identificadores únicos. En teoría, los datos no incluían información privada que pudiera reconocer de forma única a cada individuo. Sin embargo, utilizando información del sitio de acceso público IMDb, dos investigadores de la Universidad de Texas lograron cruzar datos e identificaron a usuarios que evaluaron películas en ambas plataformas.

Otro caso, más impactante aún por el tipo de información expuesta, sucedió en 2013. Ese año, un profesor de la Universidad de Harvard logró reidentificar a cuatro de cada diez participantes de una base anonimizada de voluntarios

del proyecto Personal Genome con solo consultar información electoral y registros públicos. Esa base de datos contenía información sobre las condiciones médicas de los participantes, sobre abortos practicados, uso de drogas, depresión e incluso variables genéticas.

Estos dos ejemplos son en apariencia muy distintos. En el segundo de ellos, la información que se logró reidentificar parece crítica y grave, pues incluye detalles delicados sobre comportamientos sociales que la mayoría preferiríamos mantener bajo llave. En el caso de Netflix solo se revelaron las preferencias fílmicas de las personas (¡qué culpa tenemos de que nos guste *The Hangover III!*!), lo cual podría resultar inocuo. Sin embargo, si algo aprendimos con el caso de Cambridge Analytica es que estos patrones pueden ser usados para crear modelos robustos que permitan predecir, por ejemplo, el partido por el que vamos a votar. Es en ese sentido que son igual de preocupantes.

### ***Responsabilidad: la muerte de Bridget Driscoll***

Retomemos la idea de la rendición de cuentas. Esto significa que deben existir mecanismos para garantizar que la responsabilidad por los efectos de usar IA se distribuya equilibradamente y que se ofrezca la compensación adecuada a quienes sufran daños por algún uso inapropiado.

Como es usual cuando surgen innovaciones, las primeras regulaciones se elaboran con base en referencias y recurriendo a analogías de casos legales similares provenientes de procesos más o menos conexos. Podemos construir los marcos éticos de la IA estudiando antecedentes. Podemos igualmente aprender de los errores cometidos y de los problemas surgidos sobre la marcha.

En 1865, años antes de la llegada del automóvil, el parlamento inglés había sancionado una de las primeras leyes de seguridad vehicular, la Locomotives on Highways Act, comúnmente conocida como la “Ley del banderín rojo.” La norma exigía a

todos los vehículos de carretera impulsados mecánicamente —o sea todo lo que no dependiera de un caballo (un carruaje, por ejemplo), o del esfuerzo humano (una bicicleta)— un mínimo de tres conductores y un límite de velocidad de 6,4 km por hora. También ordenaba que debía haber un hombre que caminara unos 50 metros por delante del vehículo sosteniendo un banderín rojo para indicar a los transeúntes que este se aproximaba.

La ley fue redactada para un momento en que los vehículos de carretera impulsados mecánicamente eran tan ruidosos que asustaban a los caballos que tiraban de los carruajes. Y si bien fue modificada antes de la llegada del automóvil moderno hacia 1878 —cuando se hizo innecesario el “hombre del banderín”—, se la emplea a menudo para ejemplificar la obsolescencia de un reglamento ante una disrupción tecnológica.

La adecuación y evolución de las leyes es un proceso natural, a medida que la tecnología y nuestras costumbres se van transformando. Pero las leyes no cambian de un día para otro; su evolución es lenta y generalmente parcial. Es solo mucho más tarde cuando tomamos conciencia de la necesidad de modificarlas.

En ese sentido, las leyes que se expidieron en torno al uso del automóvil resultan interesantes. Tanto la norma del banderín rojo, como sus adecuaciones posteriores, tomaron como antecedentes legales las llamadas “leyes equinas” inglesas para los carruajes de caballos. Posteriormente, sin embargo, eliminar el banderín e incrementar los límites de velocidad no sería lo único que el legislador inglés tendría que hacer.

El 17 de agosto de 1896, Bridget Driscoll, una señora de 44 años cruzaba la calle cuando fue atropellada por un automóvil manejado por el ciudadano Arthur James Edsall. El accidente ocurrió solo unas pocas semanas después de que una nueva ley del Parlamento hubiera aumentado el límite de velocidad de los coches a 23 km por hora, de los precedentes 3,2 km/h en las ciudades y 6,4 km/h en el campo (la misma que alcanza un adulto caminando o trotando). Era la primera vez que sucedía

un accidente automovilístico con consecuencias letales para un peatón. Bridget Driscoll se paralizó al ver al auto acercarse a tan solo 6 km/h. Aunque el conductor gritó e intentó llamar su atención desde una distancia suficiente, la mujer, sorprendida por una máquina para ella completamente desconocida, se paralizó. El mismo informe declaraba que el señor Edsall llevaba únicamente tres semanas conduciendo y parecía no saber lo que estaba haciendo. Frente a estos hechos, el jurado a cargo emitió un veredicto de muerte accidental.

La justicia se enfrentaba a un caso nuevo. Ciertamente, era una persona quien conducía el automóvil, pero las leyes equinas hasta ese momento tenían en cuenta aquellos incidentes donde el conductor no necesariamente tenía el control sobre el caballo. Por ejemplo, aquel no sería responsable de accidentes causados si su animal se espantaba o si realizaba algún movimiento impredecible. Con la llegada del automóvil, ese principio tendría que cambiar. Ahora el conductor no manejaba a otro ser vivo independiente, y por lo tanto se convertiría en el causante único del accidente.

Estas leyes de **responsabilidad** han ido evolucionando con el tiempo, y en las normas de tránsito actuales existe el concepto de responsabilidad civil. Un conductor entiende el riesgo que genera al utilizar un vehículo motorizado, por lo que se vuelve responsable de los daños causados a personas o bienes con motivo de ese uso.

Volviendo al tema que nos ocupa, ¿en quién recae la responsabilidad de un mal uso o desempeño incorrecto de un sistema de IA?

En 2018, un reportero de *The Washington Post* recordó el caso de Bridget Driscoll a raíz de una muerte que parecía similar. Un coche autónomo de Uber que recorría las calles de la ciudad de Tempe, Arizona, atropelló y mató a Elaine Herzberg, de 49 años, cuando ella cruzaba una calle. La mujer arrastraba una bicicleta cargada con bolsas de compras por una avenida de cuatro carriles a unos 100 metros del paso peatonal. El

automóvil viajaba en autocontrol con una conductora de respaldo sentada al volante, quien no vio a la señora Herzberg sino hasta un segundo antes de atropellarla. Eran las 10:00 p.m.

La investigación determinó que el sistema sí había detectado a la señora Herzberg. Seis segundos antes del accidente la había clasificado como un objeto desconocido, luego como un vehículo y a poca distancia como una bicicleta. El software tiene una respuesta predictiva distinta para cada caso. La telemetría del auto mostró que el programa de IA notificó a la computadora que el auto debía detenerse cuando estaba a 1,3 segundos del impacto, pero los frenos recién fueron accionados un segundo después del choque. Según los peritos, el software no tenía un sistema de alerta temprana para el conductor, de manera que no notificó el posible accidente más rápidamente. Por su parte, la conductora no estaba atenta a la vía —miraba hacia abajo— y tampoco tenía las manos sobre el volante, como dicta el protocolo de respaldo para coches autónomos. La policía investigó los hechos y concluyó que el accidente pudo haber sido evitado si la conductora no hubiera estado distraída mirando un episodio de *The Voice* en su teléfono móvil.

La falta de reacción de la conductora de respaldo de Uber es parte de la nueva dinámica que introduce a la IA como sistema de **soporte de toma de decisiones**. En este caso, la responsabilidad humana continúa.<sup>125</sup> De alguna manera, si un programa de IA está a cargo de una buena parte del trabajo que antes realizaban las personas, tendremos que considerar conceptos como la “responsabilidad del pasajero”<sup>126</sup> como parte del nuevo paradigma en donde el operador es parcialmente un responsable solidario.<sup>127</sup>

La historia de la legislación para coches tirados por caballos y para los autos nos puede enseñar otras cosas. Al comienzo, la relación entre estos dos medios de transporte causó disputas interesantes. En Estados Unidos, por ejemplo, si un automóvil espantaba a un caballo y eso desembocaba en un accidente, el automovilista era parcialmente responsable de los daños ocasionados. Entre 1899 y 1907, las muertes humanas

vinculadas de algún modo a los caballos se incrementaron en un 50%, lo que apunta a la posibilidad de que una de las causas fuera la introducción del automóvil. Al poco tiempo surgió una regla general según la cual un automovilista no sería responsable si el demandante solo podía establecer que su caballo se había asustado al ver el vehículo. Era necesario demostrar igualmente que el conductor había sido **negligente** pues, al notar los nervios en la bestia, no había procedido a cumplir con el reglamento, el cual le exigía bajar la velocidad en presencia de un animal, como hoy lo haría al avistar a una persona a punto de cruzar el paso peatonal.

Este caso muestra un punto sumamente interesante, y es que los automóviles autónomos necesitarán de una regulación más cercana a las leyes equinas que a las del automóvil tradicional. La inteligencia artificial, al igual que el caballo, va a desempeñar un papel activo durante el accidente y podría ser la causante de la muerte de un peatón tanto por un error en su programación como por una regla explícitamente programada en el sistema. Por supuesto que la responsabilidad no descansará en el algoritmo, pero queda por decidir si recae en la empresa. ¿O quizá en el programador?

<sup>125</sup> La Sociedad de Ingenieros Automotrices (SAE por sus siglas en inglés) tiene una clasificación de 6 niveles de automatización en autos basada en la cantidad de atención e intervención requerida por el conductor. El último nivel es el único que no requiere atención o interacción de un humano responsable.

<sup>126</sup> Con el invento del automóvil, las responsabilidades de los pasajeros no fueron muy claras por un buen tiempo. Según un reglamento vial de 1927, el pasajero era responsable de alertar al conductor sobre peligros en el camino. Del mismo modo, ese pasajero debería actuar en caso de que viera al conductor operar un auto de forma irresponsable, en violación del reglamento vial; al aceptar esa función, ese pasajero no podría posteriormente responsabilizar solo al conductor de los daños que pudieran haber causado en un accidente.

<sup>127</sup> El concepto de responsabilidad solidaria sostiene que el propietario de un vehículo que participó en un accidente cuando estaba siendo conducido por una persona distinta es responsable solidario de aquella por daños y perjuicios. En casos de lesiones o muerte, la responsabilidad penal es del conductor.

Sobre este punto, la universidad MIT creó un sitio en internet llamado “La máquina moral” (inspirado en el experimento mental conocido como “*Trolley Problem*”), donde más de dos millones de personas han opinado sobre la decisión que debería tomar un automóvil autónomo que se vea en un dilema moral en el que salvar la vida de una persona implique atropellar a otra.

Estamos en un mundo donde estas situaciones se repetirán. ¿Podemos responsabilizar de alguna manera a Uber por no haber comunicado de forma efectiva a la conductora que debería estar prestando atención? ¿O quizá por haber recortado de dos a uno el número de conductores de respaldo que debían ir en sus autos, antes del accidente que le costó la vida a la señora Herzberg?

Este tipo de preguntas será cada vez más frecuente en relación con las diversas aplicaciones de la IA. Y no solamente con los fenómenos que se generen directamente de la recomendación o decisión de un algoritmo, sino también del entorno digital en donde este opere. ¿Es responsable Twitter por las cuentas falsas que, utilizando algoritmos de lenguaje natural, difunden mensajes de odio o información falsa? ¿Lo es Facebook por el escándalo de Cambridge Analytica? ¿Quién es el responsable de un crimen: ¿el arma o quien la empuña?

La IA evidentemente requerirá de una especie de “protocolos Driscoll-Herzberg,” es decir, procedimientos preventivos y acciones correctivas que obliguen a las empresas del sector a responsabilizarse por los posibles efectos adversos de sus productos y servicios. Igualmente deberán garantizar que durante su desarrollo y antes de su lanzamiento al mercado se los someta a una certificación de estándares de calidad por parte de auditores independientes.

### ***Diligencia debida e interacción humana-IA***

Lo que hoy llamamos diligencia debida es un proceso mediante el cual se realiza una evaluación razonable que conduzca a

identificar debilidades y riesgos para corregirlos y mitigarlos mediante acciones concretas que aseguren el cumplimiento de criterios aceptados para el desarrollo de los sistemas. Negligencia y diligencia debida van de la mano. Y como en los sistemas de IA muchas cosas pueden salir mal, las empresas tendrán que demostrar que implementaron procesos suficientes de mitigación de riesgos bajo estándares internacionales.

En 2019, dos accidentes aéreos en Etiopía e Indonesia costaron la vida de 459 personas. Aunque este tipo de siniestros es poco frecuente, en esta ocasión ambos ocurrieron en un lapso muy corto y en ambos el protagonista fue el mismo avión: un Boeing 737 Max. Algo andaba mal.<sup>128</sup>

El 737 Max no es un avión nuevo. Es una mejora del modelo 737 diseñado en 1968 que por muchos años fue el estandarte de Boeing. Con más de 10.000 unidades en el mercado, es uno de los aviones más vendidos y confiables del mundo. Sucede que la compañía Boeing de Estados Unidos se encuentra en una carrera comercial contra el consorcio europeo Airbus, que en 2010 sacó al mercado un rediseño de su A320 con un cambio en el motor que hacía su consumo de gasolina casi 15% más eficiente en función de los costos. El ahorro en gasolina —el costo de operación más significativo— daba a Airbus una ventaja competitiva. Además, el hecho de que fuera un rediseño significaba que los pilotos no tendrían que tomar un curso de capacitación, con lo cual la compañía europea se ahorraría millones de dólares adicionales en entrenamiento.

Para seguir siendo competitiva, la Boeing quiso aplicar la misma estrategia y renovó el motor del 737. Sin embargo, como el 737 es más bajo que el A320, la compañía debió levantar el motor

<sup>128</sup> Viajar en avión es hoy en día más seguro que viajar en auto. Las estadísticas del Departamento de Transporte de Estados Unidos muestran que entre 2007 y 2016 hubo 11 muertes por cada mil millones de millas de viajes aéreos comerciales. Si bien hay algunas discusiones sobre la comparabilidad de estas estadísticas —viajar una milla en avión no es lo mismo que hacerlo en auto—, incluso cuando se calcula en número de “viajes” la tasa de muertes es menor.

casi a la altura de las alas, lo que alteró el centro de gravedad de la aeronave que, al alcanzar velocidades altas, levantaba la nariz.

La complejidad del diseño requeriría una capacitación de los pilotos, lo cual implicaría más costos en tiempo y dinero. Así que, en lugar de adoptar esa vía, Boeing decidió incluir un sistema que, de forma automatizada, bajaría la nariz del avión en caso de que se elevara demasiado, hasta devolverlo a la posición normal. El nombre del software era MCAS (siglas de Maneuvering Characteristics Augmentation System). Dado que el sistema solo se activaría en condiciones especiales —y, en teoría, el piloto no tendría que darse cuenta—, los ingenieros pensaron que se podía prescindir de la capacitación. El material de entrenamiento para pilotear el Boeing 737 Max ni siquiera incluía una descripción del software; es más, muchos pilotos no sabían de su existencia.

Según las investigaciones acerca del accidente de la compañía indonesia Lion Air, cuando el avión despegaba de Yakarta el sistema comenzó a empujar la nariz hacia abajo, aparentemente por una falla en los sensores. Los pilotos no entendían qué pasaba ni conocían el protocolo para desactivar el MCAS en una emergencia así. El avión terminó por caer al mar.

En esta historia se mezclan la negligencia de la empresa, al no aplicar medidas de mitigación de riesgo, con la falta de preparación y entrenamiento de los operadores sobre las características y peligros de un sistema con el que tenían que interactuar.

### ***Explicabilidad y transparencia en las decisiones***

Insistimos. Los programas inteligentes deben diseñarse y aplicarse de manera tal que puedan ser supervisados mediante la traducción de sus operaciones en resultados inteligibles y del suministro de información sobre dónde, cuándo y cómo se van a utilizar.

El siguiente caso es ilustrativo. En 2019, el ingeniero estadounidense David Heinemeier escribió molesto en Twitter que la tarjeta de compras de Apple había discriminado a su esposa al darle un cupo de crédito 20 veces menor que el suyo. Heinemeier indicó que él y su esposa presentan una declaración de renta conjunta, ambos son propietarios de los mismos inmuebles, “y sin embargo la caja negra de Apple cree que merezco 20 veces el límite de crédito que ella.” Apple redirigió el cuestionamiento al banco Goldman Sachs, el cual corría el análisis de crédito, pero sus empleados aseguraron que no existía una discriminación por sexo, aunque tampoco podían explicar por qué el algoritmo había designado a la señora Heinemeier como menos solvente que su marido. Steve Wozniak, cofundador de Apple, compartió en Twitter que él y su esposa habían identificado el mismo comportamiento y expresó que tanto Apple como Goldman Sachs deberían asumir la responsabilidad de este error.

¿Cómo se puede garantizar que las personas afectadas impugnen este tipo de resultados? La respuesta a este problema es también una pregunta abierta que la capa legal debe resolver. En caso de que una empresa tenga que justificar los resultados de un algoritmo, este debe ser interpretable y explicable en la medida de lo necesario.

En aprendizaje automático, el término “explicabilidad” se utiliza para describir la capacidad de comunicar a un usuario o beneficiario las razones que llevaron a que un sistema tome una u otra decisión. Por su parte, el término “interpretabilidad” se refiere a hacer inteligible para las personas del común el funcionamiento de un algoritmo y sus resultados. Las explicaciones individuales pueden ser importantes cuando se trata de rendir cuentas a los usuarios o de explicar decisiones tomadas con base en aprendizaje automático. Cabe notar, sin embargo, que los métodos predictivos complejos son en principio poco transparentes cuando se busca explicar cómo se hacen las predicciones subyacentes.

Es muy probable que veamos cada vez más casos donde los empleados no entiendan los resultados de las cajas negras que

usan las empresas. La sociedad aún tiene que determinar hasta dónde y en qué instancias es obligatorio que el resultado del sistema sea interpretable y explicable.

Contamos ya con algunos requisitos mínimos. En varios países de América Latina, cuando hablamos de servicios relacionados con los derechos sociales, existe el requerimiento constitucional para las instituciones de justificar el otorgamiento o negación de un servicio, y para los individuos de impugnar esas decisiones si se cree que existe alguna discriminación. Nada de esto da espera. Ya se ha abundado bastante en estos temas y tenemos suficiente información como para poder comprender las implicaciones técnicas, éticas, filosóficas y legales de la IA, y el tipo de sociedad en la que nos podemos convertir. La necesidad de crear marcos jurídicos y definir las reglas del juego no puede subestimarse; es un artículo de primera necesidad.



## EPÍLOGO

Aumentar el ritmo con responsabilidad

Cuando iniciamos este libro a principios de 2020, pensábamos que íbamos a describir casos exitosos del uso de la IA en América Latina y el Caribe. Con ello nos proponíamos llamar la atención sobre la imperiosa necesidad de hacer inversión pública en este rubro, con el fin de que sus instrumentos se fueran incorporando paulatinamente en la gestión del Estado. Lo que sucedió meses más tarde nos tomó completamente por sorpresa. El coronavirus se propagó y en cuestión de semanas produjo un cimbronazo de incalculables proporciones a lo largo y ancho del planeta.

Ante la inminencia de los impactos que produciría un evento catastrófico de semejante magnitud, sabíamos que los sistemas digitales y la IA podrían revolucionar la forma en que históricamente hemos respondido a una emergencia como la pandemia del coronavirus, y que se podría diseñar una política pública que permitiera analizar datos en tiempo real, estudiar la expansión del virus, anticipar patrones de contagio, focalizar medidas de rastreo de contactos y redefinir otras políticas públicas conforme se iban acopiando mayores volúmenes de información para mitigar sus efectos más adversos.

Y aunque estábamos conscientes de que ALC presentaba un rezago importante en materia digital, lo que la pandemia reveló fue un escenario incluso más precario. La ausencia de registros electrónicos de salud, de padrones de beneficiarios de servicios sociales y de censos educativos limitó el uso de las soluciones digitales más básicas. Se hizo evidente que el teletrabajo, la teleducación y la telemedicina son prácticas difíciles de implementar en una región donde el trabajo informal es significativo, existe una desigualdad palpable y la brecha digital hace que millones de niños carezcan de acceso a computadoras y/o conectividad. Asimismo, la pandemia puso de manifiesto la significativa deuda institucional que acarrean los Estados latinoamericanos, su falta de capacidades técnicas y una deficiente comprensión científica de la crisis sanitaria del momento, lo cual hizo que algunos de los países más grandes de la región pusieran en duda su potencial destructivo y se mostraran incapaces de contener su avance.

Sin embargo, no podemos desaprovechar el momento para aprender de esta experiencia, especialmente si se considera que en ALC las emergencias son el pan de cada día. Fenómenos como el cambio climático, la violencia y la lucha contra el crimen son algunas de las amenazas que se ciernen permanentemente sobre la región. Aunque en menor escala, estos choques exógenos modifican a diario la vida de millones de familias altamente vulnerables y con capacidades muy limitadas de responder a ellos.

Frente a estos escenarios insoslayables, los avances tecnológicos nos invitan a pensar en la posibilidad de un nuevo Estado, un Estado de bienestar digital inteligente (EBDI) dotado de un novedoso conjunto de herramientas predictivas para poder implementar políticas adaptativas al riesgo. Esta modernización del aparato estatal, junto con el objetivo de migrar de un paradigma reactivo a uno preventivo, exige un cambio en sus responsabilidades y en su relación con la ciudadanía, tal y como ha sucedido en otros momentos de la historia durante las grandes revoluciones industriales previas, que también se produjeron en el contexto de enormes sacudidas causadas por eventos calamitosos. Ello nos ha permitido entrever cuáles podrían ser las características especiales de este EBDI que comienza a vislumbrarse, especialmente a la luz de las deudas pendientes que tiene en América Latina el Estado de bienestar, que hasta el momento no ha pasado de ser una declaración de buenas intenciones. Porque lo cierto es que seguimos siendo la región más desigual del mundo, caracterizada por problemas como la corrupción, la poca transparencia y una rendición de cuentas deficiente.

Como dijimos al comienzo, no intentábamos con este ejercicio proponer o presentar un modelo armado y acabado de ese nuevo Estado que aún no termina de definirse, sino simplemente ir describiendo, desde una perspectiva histórica, los elementos que se han ido manifestando a partir de eventos actuales que, como la pandemia del coronavirus, han afectado al planeta entero y desencadenado una serie de cambios en sus estructuras que se encargarán paulatinamente de definir el EBDI.

## NOTAS DE LOS CAPÍTULOS

Lo que sí podemos hacer es tomar la delantera en lo que tiene que ver con las discusiones sobre la ética y la justicia de los sistemas de IA, dado que estas ofrecen una oportunidad única para cuestionar nuestros propios supuestos, abrir la caja negra de nuestras preconcepciones y desenterrar los algoritmos sesgados que nuestras sociedades han utilizado por años; esto con el fin de reconstruirlos y hacerlos más transparentes en aras de crear un mundo más justo y equitativo.

Asimismo, estamos convencidos de que se requiere que ese EBDI se fundamente en las particularidades culturales, sociales y políticas de nuestra región, y que cuente con la participación proactiva de los actores interesados --Estados, empresas, academia, organizaciones de la sociedad civil-- en la discusión sobre sus marcos éticos, para lo cual ya existen espacios internacionales donde diversos grupos de trabajo discuten estos temas con ánimo de concertación.<sup>129</sup>

No sobra repetirlo una vez más: un Estado de bienestar digital inteligente es posible. Pero para lograrlo debemos acometer de inmediato la tarea de preparar los procesos institucionales que conduzcan a que también sea un Estado responsable y ético que actúe por el bien de la ciudadanía sin perpetuar las brechas de desigualdad que han caracterizado a la región.

<sup>129</sup> Existen ya iniciativas en curso cuya agenda se ocupa de estos puntos. Una de ellas es fAIR LAC, del Banco Interamericano de Desarrollo, con la cual buscamos promover la adopción responsable de la IA para mejorar la prestación de servicios sociales y crear oportunidades de desarrollo en la región. Esto lo hacemos implementando pilotos controlados donde podamos entender los riesgos y las posibilidades de las herramientas, elaborando estándares de calidad en su desarrollo e impulsando la formulación y adopción de marcos regulatorios en los países, comenzando con temas como la interoperabilidad de sistemas, la protección de los datos personales y la expedición de normas específicas para la IA.

### Capítulo 1 Lo esencial es invisible

El relato de la gripe española proviene de una variedad de textos y estudios. Empleamos la cifra de muertos de la gripe consignada en “The Origin and Virulence of the 1918 ‘Spanish’ Influenza Virus”, la ponencia de Jeffery K. Taubenberger en el simposio “Archaeology: New Techniques and Methods” de 2002, disponible en los archivos de la Biblioteca Nacional de Medicina de Estados Unidos. Taubenberger era entonces jefe del Departamento de Patología Molecular del Instituto de Patología de las Fuerzas Armadas de Estados Unidos. Otras fuentes reportan estimados de muertes hasta de 50 millones de personas o más.

El artículo del *New York Times* sobre las tasas de contagio del coronavirus mencionado es “How Bad Will the Coronavirus Outbreak Get? Here Are 6 Key Factors” de Knvul Sheikh, Derek Watkins, Jin Wu y Mika Gröndahl.

El material contextual sobre viajes aéreos, transporte de carga e interconexión de ciudades proviene de fuentes como la IATA, la International Chamber of Shipping, la Oficina de Estadísticas de Transporte del gobierno de Estados Unidos, la Asociación Nacional de Controladores de Tráfico y la National Oceanic and Atmospheric Administration (ambas de Estados Unidos), y el portal alemán Statista.

Para el segmento sobre vacunas se usaron las publicaciones de los Centros de Control y Prevención de Enfermedades de Estados Unidos (CDC por sus siglas en inglés), el trabajo “Historia de la vacunación” de Stanley Plotkin y “How Long Will a Vaccine Really Take?,” un artículo de Stuart Thompson publicado en *The New York Times*.

Poco antes de la gran pandemia de la gripe asiática de fines del siglo XIX, la humanidad apenas contaba con dos vacunas: viruela y rabia; otras tres —tifoidea, cólera y plaga— se agregarían antes de 1918. La población mundial sufriría la pandemia de gripe española sin defensas farmacológicas, pues la primera vacuna efectiva contra la enfermedad recién llegaría en 1936, casi dos décadas después del contagio planetario más letal de que se tuviera noticia. En la actualidad, países como Estados Unidos disponen de un conjunto de casi 30 vacunas, mientras que para mediados de la década pasada la humanidad entera contaba con al menos 50 tipos de vacunas para tratar un número similar de enfermedades mayores.

Para reconstruir las estrategias de control del SARS CoV-2 y del COVID-19 en Corea, China y Taiwán, así como los avances tecnológicos a nivel internacional, se usaron una variedad de fuentes, entre ellas ABC News, Al-Jazeera, CNN, *Forbes*, *Foreign Policy*, Reuters, *South China Morning Post*, *TechCrunch*, *The Guardian*, *The New York Times*, *The Washington Post*, *Time* y *Wired*. También se consultaron informes de instituciones como la Organización Mundial de la Salud (OMS); de las empresas Data Innovation, DeepMind, Google e IBM; de Stanford University, y de publicaciones oficiales de cada gobierno.

Bernard Marr, autor del libro *Artificial Intelligence in Practice: How 50 Successful Companies Used Artificial Intelligence to Solve Problems*, rescató diez áreas en las que la inteligencia artificial contribuyó a solventar la crisis pandémica. La lista es interesante. Según este autor, la IA ha permitido:

- Identificar, rastrear y prever la evolución de brotes, como lo consiguió la empresa social canadiense BlueDot.
- Ayudar a diagnosticar el virus con algoritmos, como lo hicieron las compañías chinas AliBaba e Infervision, cuyo algoritmo fue reprogramado (originalmente se usaba para diagnosticar cáncer de pulmón) para detectar la neumonía asociada a enfermedades respiratorias como el coronavirus. Según la empresa, este fue empleado por

al menos 34 hospitales chinos que usaron su lectura de tomografías para analizar 32.000 casos sospechosos.

- Procesar las necesidades de atención médica, para lo cual se usaron decenas de *bots* daneses, coreanos y chinos.
- Manejar información administrativa de millones de personas, una tarea en la que, entre otras, se destacó la compañía Ant Financial, la cual se apoya en la cadena de bloques (*blockchain*) para acelerar la administración de hospitales o aseguradoras.
- Entregar suministros médicos y alimentos o patrullar espacios públicos con drones (como lo hicieron Terra Drone y otras empresas en China), esterilizar con robots (como lo hizo Pudu Technology), y abastecer hospitales en Asia reprogramando robots empleados para el servicio de suministro de comidas (como lo hizo Blue Ocean Robotics).
- Investigar las enfermedades como lo está haciendo DeepMind de Google, que estudia las proteínas de los virus y publica sus resultados en la modalidad de acceso abierto.
- Producir estuches de detección y desarrollar vacunas, como lo hace BenevolentAI, cuyo algoritmo predictivo sugiere también medicamentos ya existentes que podrían ser evaluados para uso contra el coronavirus. Un proyecto relacionado es el de la compañía surcoreana de biotecnología Seegene, la cual trabaja en la investigación molecular de varias enfermedades y desarrolló un estuche de exámenes de coronavirus en solo tres semanas para hacerlo más accesible y popular. La supercomputadora de Seegene permitió reducir el tiempo usual de dos a tres meses que toma producir una prueba. Esta compañía fue una de las cuatro proveedoras locales de pruebas en Corea del Sur, una de las naciones que más tamizajes realizó, y de manera gratuita, entre la población.

- Fabricar productos especiales ajustados a las necesidades de la crisis como ha hecho la compañía israelí Sonovia, creadora de una tela antibacterial y antipatógenos hecha con nanopartículas de óxido metálico que puede emplearse en mascarillas o indumentaria para el personal sanitario.
- Usar inteligencia de datos para controlar los desplazamientos de individuos enfermos.
- Usar *chatbots* para compartir datos o facilitar información —tal como lo hace la industria de viajes cuando debe emitir alertas a los usuarios de aviones, cruceros y hotelería— y operar como portales para personas que necesitan revisar su condición sanitaria, como en los hospitales coreanos.
- Entrelazar supercomputadoras para trabajar en la búsqueda de vacunas a través de computación en la nube, una operación que numerosas compañías —Huawei, DiDi y Tencent, entre otras— lanzaron para aprovechar la velocidad de cálculo y modelación de situaciones de su software.

La pandemia también produjo una explosión de *chatbots* para procesamiento de lenguaje natural que dialogaban con las personas sobre posibles síntomas y tratamientos para el COVID-19. Con base en la descripción que esas personas hacían de la dolencia, los robots podían orientarlas para que buscaran un hospital. En Shanghái, los *bots* incluso llamaban a los pacientes en riesgo para monitorear su estado. Solo en China hubo más de 200 robots en hospitales, hoteles y escuelas, o en zonas críticas de algunas ciudades, donde además repartían alimentos y medicinas o controlaban la fiebre de las personas. Uno de ellos, llamado Cloud Ginger, fue reprogramado para que actuara como émulo artificial de Patch Adams, el médico-payaso que protagonizó Robin Williams en la película de 1998; estaba equipado con pantallas para que los pacientes tuvieran entretenimiento e información interactiva. También transportaba medicinas y alimentos. Una compañía danesa, UVD Robots, ubicó varias centenas de sus robots de

limpieza y desinfección en centros asistenciales en China e Italia. Sus máquinas mapeaban los corredores y salas de los hospitales y luego emitían luces ultravioletas que destruían las líneas de ARN del coronavirus en las superficies que hallaban contaminadas a su paso. El propósito de los robots era sencillo: reducir el contacto persona a persona. Cuando las tasas de contagio bajaron en China y Corea, por ejemplo, las máquinas entraron automáticamente en hibernación a la espera de ser requeridas nuevamente.

El caso de Taiwán fue muy bien analizado por Jason Wang, Chun Y. Ng y Robert H. Brook en “Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing”. Hasta mediados de junio de 2020, el mapa de áreas de riesgo de contagio empleado por Taiwán y mencionado en este capítulo se podía consultar en la página web del Ministerio de Salud del país asiático. El texto de Dan Shapiro sobre el comportamiento de los ciudadanos taiwaneses en relación con su gobierno es “Taiwan Shows its Mettle in Coronavirus Crisis, while the WHO is MIA,” Este fue publicado en marzo de 2020 en la sección “Order from chaos” de Brookings Institution.

El material contextual de SARS y MERS proviene de notas de prensa de Deutsche Welle, *South China Morning Post* y *The Washington Post*. El registro de las acciones desarrolladas por el gobierno de Taiwán publicado a inicios de 2020 por la American Medical Association también fue útil. Las afirmaciones del epidemiólogo de Google, Larry Brilliant, citadas en este libro, fueron publicadas el 18 de marzo de 2020 en *Político*.

Como dedicados admiradores de “2001: Una odisea en el espacio”, y para no dejarnos guiar solo por nuestros conocimientos, los miembros del equipo consultamos varios reportajes sobre la novela y la película en *The London Review of Books*, *The New York Times* y *The Paris Review*. El texto de Dan Chiasson “‘2001: A Space Odyssey’: What it Means, and How it was Made”, publicado en la revista *New Yorker* en abril de 2018, es particularmente interesante.

El caso de información no consolidada sobre muertos por COVID-19 en España fue publicado por diversos medios de ese país durante un periodo de varias semanas en mayo de 2020, entre ellos *El País*, *Eldiario.es*, *El Confidencial* y *ABC*.

Los datos sobre identificación personal y digitalización provienen del informe “Digital Identification: A Key to Inclusive Growth,” del McKinsey Global Institute. Varios datos de conectividad actual y futura en América Latina provienen de Gartner y Statista. Los detalles sobre el impacto positivo de la digitalización en América Latina provienen del capítulo 4, “Impact of E-Government on Transparency and Corruption,” del libro *Unlocking E-Government Potential: Concepts, Cases and Practical Insights* de Subhash Bhatnagar.

La investigación del Big Data Institute de Oxford University citada es “Quantifying SARS-CoV-2 Transmission Suggests Epidemic Control with Digital Contact Tracing” y fue publicada por *Science* en mayo de 2020.

## Capítulo 2

### Thomas Southwood Smith y el algoritmo aterrador

Varias publicaciones alimentaron el retrato de época de la Primera Revolución Industrial en Inglaterra. Entre ellas figuran *A City Profile on Liverpool*, un interesante texto sobre los cambios demográficos y urbanos de Liverpool a lo largo de un ciclo histórico dominado por grandes cambios tecnológicos. La narración sobre la vida de Thomas Southwood Smith proviene de un especial sobre el reformador publicado por Spartacus Education. Para la reseña de las enfermedades y el estado de la sociedad en la época se usó un excelente trabajo de Michael Brown, “Making Sense of Modernity’s Maladies: Health and Disease in the Industrial Revolution.” Alexis de Tocqueville describe a Manchester en sus *Journeys to England and Ireland*; de allí proviene la cita. Los detalles de las “English Poor Laws” se encuentran tanto en publicaciones del gobierno británico como en las de las universidades de Cambridge, Oxford y Wisconsin-La Crosse. El caso de Bethnal Green fue reforzado

con material del Victoria County History. Robert Wilde, en “Public Health During the Industrial Revolution,” ofrece un relato muy completo sobre las reformas sociales introducidas a partir del trabajo de Southwood Smith y Chadwick.

Cabe notar aquí que esas leyes nacieron, en buena medida, como respuesta a un grave deterioro económico, al aumento del número de trabajadores urbanos y a una disminución casi simultánea de las formas tradicionales de asistencia caritativa. La disolución de los monasterios ordenada por Enrique VIII, seguida de la disolución de los gremios religiosos, fraternidades, casas de beneficencia y hospitales en la década de 1540, acabó con una gran proporción del tejido de caridad que cobijaba a los pobres. Después de emplear los recursos expropiados para solventar las guerras del rey, el Parlamento notó que las arcas habían quedado exhaustas y que el país se sumía en la pobreza y el malestar social. Las leyes de 1597-98 y 1601 pueden considerarse como un intento del Parlamento tanto por prevenir el hambre como por controlar el orden público.

La cita de Clark y Page proviene de “Welfare Reform, 1834”, publicado en mayo de 2008 por el Department of Economics, UC-Davis.

Los interesados en el tema las transformaciones del Estado en los siglos XIX y XX podrán hallar más información e ideas estimulantes en *Introduction to Social Work: An Advocacy-Based Profession*, el libro de Lisa E. Cox, Carolyn J. Tice y Dennis D. Long, en particular en el capítulo 2, “The History of Social Work.”

El texto de Lorraine Boissoneault citado fue publicado por *Smithsonian Magazine* el 14 de julio de 2017 con el título “Bismarck Tried to End Socialism’s Grip —By Offering Government Healthcare.”

La publicación de Cepal sobre gasto social incluida en este capítulo es “Gasto social”, consultada en el Portal de Inversión Social en América Latina y el Caribe (véase la bibliografía).

Los economistas que sostienen que la nueva revolución podría profundizar la desigualdad social son Erik Brynjolfsson y Andrew McAfee en *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*.

Hay un aporte crítico adicional sobre el reemplazo de fuerza laboral humana por robots mencionado en el capítulo que vale la pena ampliar. Dicen Ramiro Albrieu y Martín Rapetti en *The Fourth Industrial Revolution and the Future of Work in Argentina*: “Debido a su naturaleza incipiente, la velocidad del cambio y sus múltiples y amplias aplicaciones, es difícil saber el impacto de 4IR en los patrones económicos y de empleo a mediano y largo plazo. Por otro lado, la historia muestra que las revoluciones tecnológicas son períodos de divergencia económica, donde algunos países pudieron absorber completamente los nuevos inventos y tecnologías, mientras que otros se quedaron atrás. Por lo tanto, si la historia sirve de guía, es probable que el impacto de la tecnología en el futuro del trabajo difiera entre los países. La demografía, la dotación de factores, las brechas de género e informalidad y la existencia (o no) de factores de complementariedad clave darán forma a la carrera entre tecnología y habilidades, lo que la hace altamente idiosincrásica”. (...) La Cuarta Revolución Industrial y el Futuro del Trabajo en Argentina tenderán a estar sesgados hacia tres grandes grupos: conocimiento cognitivo general, conocimiento específico asociado a las nuevas tecnologías y habilidades socioemocionales. Por otra parte, las ganancias del cambio tecnológico se distribuirán de manera desigual entre los trabajadores de acuerdo con sus perfiles de habilidades, con los subsecuentes efectos negativos sobre los salarios de los trabajadores con habilidades obsoletas (...). Al mismo tiempo, las nuevas tecnologías tienen el potencial de mitigar las brechas de ingresos y las desigualdades en varios sectores desfavorecidos del mercado laboral. Por ejemplo, las mujeres y los jóvenes podrían beneficiarse de las innovaciones digitales que les permitan cerrar la brecha.”

La exposición sobre riesgos y desastres naturales reseñada en el texto proviene de “¿Desarrollo económico inestable?

Choques agregados en América Latina y el Caribe” del Banco Mundial. El documento del WEF referido en la misma sección se titula “Advancing Human-Centred Economic Progress in the Fourth Industrial Revolution.”

La cita sobre la reunión del centro de pensamiento Initiative Exchange de Bangkok proviene de “State Capability, Policymaking and the Fourth Industrial Revolution: Do Knowledge Systems Matter?”, un documento de Arnaldo Pellini, Vanesa Weyrauch, Maria Mahlo y Fred Carden.

### Capítulo 3 Terminator no existe

La historia de BlueDot fue reconstruida con material de prensa, en especial del *New York Times* y *Wired*.

¿Tratamos a las computadoras como actores sociales? Esta pregunta está en el sustrato de la atracción/rechazo que nos pueden producir las computadoras, incluidas las asistentes virtuales. Las referencias de este apartado provienen de la investigación de Wade J. Mitchell, Chin-Chang Ho, Himalaya Patel y Karl F. MacDorman titulada “Does Social Desirability Favor Humans? Explicit-implicit Evaluations of Synthesized Speech Support a New HCI Model of Impression Management” publicada en *Computers in Human Behavior* en enero de 2011. Un resumen sobre este tema también se encuentra en un breve video de *The Wall Street Journal*: “Alexa, Siri, Cortana: Why All Your Bots Are Female.”

En 2014, un trío de investigadores rusos informó que un programa computacional al que llamaban Eugene Goostman había logrado batir la prueba de Turing en conversación escrita. El *bot* fue creado con la personalidad de un chico de 13 años, de origen ucraniano, que tenía por mascota un hámster y cuyo padre era ginecólogo. La idea de que fuera un chico de 13 años tenía sentido, pues en el diálogo aparentaría saber algunas cosas pero no todo sobre todo; el que fuera ucraniano

tenía el propósito de salvar a la máquina ante posibles errores gramaticales que pudiera cometer en la interacción escrita en inglés. Para 2012, en un evento organizado para conmemorar el centenario del nacimiento de Turing, Eugene Goostman logró convencer al 29% de los jurados de que era humano. Dos años después, en 2014 durante un concurso que conmemoraba la muerte de Turing en 1954, el programa había convencido al 33% de los jurados de que no era una máquina. En su experimento, Turing había predicho que para el año 2000, las máquinas serían capaces de engañar a un tercio de los jurados humanos en una interacción de cinco minutos. El logro del *bot* Eugene Goostman ha sido cuestionado por científicos de la comunidad de IA como Gary Marcus de la Universidad de Nueva York y Murray Shanahan del Imperial College de Londres. La historia de Eugene Goostman, el robot de los investigadores rusos que habría vencido a la máquina de Turing, está ampliamente documentada en artículos publicados por la BBC, *The New York Times* y Deutsche Welle.

En este capítulo se habla de la necesidad de diseñar buenas metodologías de investigación para actuar frente a problemas sociales. La Organización Panamericana de la Salud tiene un esquema sencillo que sirve de guía para definirlos y delimitarlos. El enlace al documento, cuya descarga es gratuita, se encuentra en la bibliografía. Allí también están los enlaces de acceso al documental “Slaughterbots”, sobre armas autónomas, y al compromiso “Lethal Autonomous Weapons Pledge.”

Como soporte de decisiones, la IA puede ser empleada de varias maneras. En este capítulo sostenemos que los sistemas pueden actuar de forma más o menos autónoma, con mayor o menor participación humana, como se observa a partir de las siguientes definiciones:

- *Inteligencia asistida.* Este concepto se utiliza para describir aquellos sistemas de soporte de decisiones donde a los sistemas de información se agrega una capa analítica mediante modelos de aprendizaje automático o de otro tipo. Esto con el fin de incrementar la disponibilidad de

datos para adoptar determinaciones en relación con un objetivo específico.

- *Inteligencia aumentada.* Este concepto se utiliza para describir aquellos sistemas de soporte de decisiones donde los sistemas de información aprenden de la interacción con los humanos y el entorno.
- *Inteligencia automatizada.* Este concepto refiere a la automatización de un proceso manual o cognitivo sin cambiar de forma sustancial el procedimiento utilizado para ejecutar la tarea. En otras palabras, el algoritmo pasa a realizar tareas previamente ejecutadas por las personas.
- *Inteligencia autónoma.* Aquí todos los procesos están digitalizados y automatizados mediante una combinación de software y hardware, a tal punto que no se requiere la intervención humana. A diferencia de la inteligencia automatizada, la autónoma se puede adaptar a distintas situaciones no previstas ni programadas por la experiencia humana introduciendo cambios a la tarea o proceso, tal como lo haría una persona.

A lo largo del libro se utiliza el concepto de sistemas de soporte de decisiones, que incluye inteligencia asistida y aumentada, y de sistemas de toma de decisiones, que incluye inteligencia automatizada y autónoma.

Varios de los modelos de IA que se mencionan en este capítulo ya se han implementado para diagnosticar enfermedades o predecir la deserción escolar. El de BlueDot, descrito al comienzo, muestra cómo describir la propagación de una enfermedad con cierto rango de error no es suficiente para mitigar sus riesgos. Las distintas acciones de política pública que se pueden adoptar como cerrar fronteras, realizar pruebas y hacer rastreo de contactos son decisiones que llegaron con posterioridad al resultado de los modelos. El puntaje crediticio es otro ejemplo común de predictibilidad utilizado en todos

los servicios financieros. Los modelos de calificación de crédito procesan un historial crediticio, una solicitud de préstamo y otros factores para clasificar a los consumidores según la probabilidad de que amorticen el préstamo cumplidamente. Por supuesto que en más de una ocasión, y sin supervisión humana de por medio, los algoritmos han creado predicciones cuestionables. En 2020, la fundación que opera el International Baccalaureate, un diploma que permite a egresados de todo el mundo con buenos resultados postular a universidades de prestigio, fue objeto de numerosas críticas por parte de padres, estudiantes y profesores. Resulta que a raíz de la pandemia, la fundación decidió calcular los resultados para expedir su diploma empleando información de cinco años de algoritmos propios que revisaban las calificaciones de las tareas del alumno, las calificaciones previstas y los resultados históricos del IB de su colegio. Según narra el reportaje publicado en la revista *Wired*, se esperaba que con ello se pudiera inferir lo que un estudiante habría obtenido en un 2020 no dominado por una pandemia mortal. Para las escuelas donde faltaban datos históricos, las predicciones se basarían en datos combinados de otras escuelas. El problema fue que muchos estudiantes obtuvieron calificaciones por debajo de sus expectativas académicas. Para julio de 2020, más de 15.000 padres, alumnos y maestros habían firmado una petición en línea en la que solicitaban una evaluación más justa. Para *Wired*, “las preocupaciones sobre los modelos matemáticos defectuosos están aumentando a medida que un mayor número de empresas y gobiernos aplican programas informáticos a problemas tradicionalmente resueltos por seres humanos como son las decisiones de crédito, la identificación de sospechosos de delitos y la definición de qué significa discurso de odio”. Los temas relativos a “erradicar los prejuicios, así como a la inexactitud en tales sistemas, constituyen un campo creciente del activismo y la academia.”

## Capítulo 4 ¿Listos para el despegue?

La casuística sobre la IA en América Latina se sustenta tanto en materiales publicados por medios e instituciones, como en entrevistas personales realizadas por el equipo de investigación. Entre otras personas, entrevistamos a responsables de Teledx y U-planner (Chile), Laura-br y Telemedicina (Brasil), Prometea (Argentina), Territorium y Doc.com (México), representantes de los gobiernos de Chile (2020) y Argentina (2019), y especialistas en IA y tecnología de México y Argentina.

Para el caso específico de Centroamérica, en las fuentes consultadas se indica que las primeras en asomarse al fenómeno han sido algunas grandes compañías multilaterales. Por ejemplo Autofácil, la división del salvadoreño Grupo Poma especializada en el otorgamiento de créditos y *leasing* automotriz en Centroamérica y Colombia, tiene un programa de IA apoyado en aprendizaje automático que permite que un usuario pueda gestionar información, “hablar” con el programa y obtener respuestas. El sistema facilita la gestión con el objeto de que un agente real solo atienda a aquellos clientes con los cuales haya una alta probabilidad de concretar un negocio.

Las cifras sobre teletrabajo provienen de un documento de la especialista del BID Laura Ripani publicado bajo el título “Coronavirus: un experimento de teletrabajo a escala mundial,” y del estudio “Teletrabajo en América Latina” de la firma 5G Americas.

El estado de la digitalización del Estado en América Latina se basa en artículos de prensa y en trabajos de la Corporación Andina de Fomento, el BID y Latin Data Barometer.

Los detalles sobre la educación en América Latina provienen de documentos como “La educación en tiempos del coronavirus. Los sistemas educativos de América Latina y el Caribe ante la Covid-19.” La información sobre digitalización hospitalaria proviene del informe “Electronic Medical Records Growing in

Latin America.” Los comentarios de especialistas del BID sobre la IA en educación provienen del trabajo “El futuro ya está aquí: habilidades transversales en América Latina y el Caribe en el siglo XXI” publicado en octubre de 2019.

Las estadísticas sobre informalidad laboral se encuentran en “Informalidad laboral y coronavirus: una combinación desafiante” y en informes de medios como *Clarín* de Argentina, *Expansión* y *El Universal*, *The New York Times*, y de la OIT.

El estado de la telemedicina a nivel global se basa en información de CNBC, Global Health Intelligence, *The Economist*, *The New York Times*, *Time* y YouTube.

Los datos sobre IA en América Latina provienen de una variedad de fuentes, entre ellas el BID, *Beath Israel DMC*, *The New York Times*, *Estrategia & Negocios*, *Ámbito Financiero*, *AméricaEconomía*, *Forbes México* y Centroamérica y SciDevNet. La información sobre el mercado de la salud proviene de un estudio de KPMG titulado “Investment in AI for Healthcare Soars” y del artículo “Investors Poured \$4B into Healthcare AI Startups in 2019” de Fierce Healthcare.

## Capítulo 5

### Alexa: ¿Este sesgo es mío?

La OMS ha publicado numerosos artículos sobre los riesgos de los embarazos adolescentes. Tal como se señala en el texto, las complicaciones en los embarazos y partos son responsables del 99% de las muertes de madres de 15 años o menos en naciones de ingresos bajos y medios, según sus estimaciones sobre la salud global de 2015. De acuerdo con el estudio “Pregnancy and Childbirth Outcomes among Adolescent Mothers,” los riesgos de sufrir preeclampsia, endometritis puerperal e infecciones sistémicas son significativamente superiores en niñas madres que en mujeres de más de 20 años de edad. Para este capítulo también incorporamos material periodístico de *Bloomberg* y *The New York Times* e investigaciones del Guttmacher Institute, BMC Public Health y los CDC de Estados Unidos.

El caso de IA para detección de embarazo adolescente en Salta ha sido documentado por diversas organizaciones y tuvo amplia cobertura de prensa, incluido un informe muy completo de The Web Foundation, IEEE, la School of Computing and Mathematical Sciences de la Liverpool John Moores University, y la Faculty of Computer and Information Science de University of Ljubljana. Cabe señalar que cuando nos encontrábamos escribiendo este caso a finales de 2019, la información de las personas encuestadas en Salta seguía almacenada en los servidores de Microsoft debido al vacío normativo y a que Argentina no posee servidores seguros propios.

Una buena parte de la casuística de este capítulo proviene de reportajes, columnas y entrevistas publicadas por *The New Yorker*, *The Wall Street Journal*, *Business Insider*, Reuters, *The New York Times*, NPR, *The Washington Post*, *Wired*, *The Atlantic*, *Expansión*, *El País*, *Estrategia & Negocios*. El estudio sobre sesgos raciales de Obermeyer, Powers, Vogeli y Mullainathan fue publicado en *Science*, y la información complementaria proviene de *MIT Technology Review* (véase la Bibliografía). Los detalles sobre proyectos de seguridad cibernética de los gobiernos de América Latina se encuentran en el estudio “Ciberseguridad: ¿estamos preparados en América Latina y el Caribe?” del Observatorio de la Ciberseguridad del BID y en “Three Reasons Why Latin America is Under Cyber Attack” del IEEE. El estudio de Nesta mencionado en el capítulo se titula “Decision-making in the Age of the Algorithm: Three Key Principles to Help Public Sector Organisations Make the Most of AI Tools.”

Volviendo al caso del registro electoral en México, la historia de cómo el Instituto Nacional Electoral (INE) de ese país supo de la exposición de los datos es ilustrativa. Chris Vicerky, un especialista en seguridad informática, halló el paquete de 132 gigas de información en la red y avisó al Departamento de Estado de Estados Unidos para que se lo comunicara a la Cancillería de México. Nadie le hizo caso. Vickery entonces llamó a la Embajada de México en Washington, DC. Tampoco le prestaron atención. Solo cuando subió un video al respecto en un sitio especializado, el INE lo contactó.

Aunque no forman parte del tema central del libro, las leyes de educación, trabajo y vivienda en la región suelen tener en su articulado postulados como el siguiente (proveniente de la legislación mexicana): “Las disposiciones de esta Ley deberán aplicarse bajo principios de equidad e inclusión social de manera que toda persona, sin importar su origen étnico o nacional, el género, la edad, la discapacidad, la condición social o económica, las condiciones de salud, la religión, las opiniones, las preferencias o el estado civil pueda ejercer su derecho constitucional a la vivienda.”

## Capítulo 6

### Sonríe; te estamos filmando

En este capítulo, la cobertura de la pandemia en Asia y Europa proviene de una variedad de fuentes de prensa y de documentos preparados por especialistas. El material periodístico fue publicado por organizaciones como *Animal Político*, Reuters, BBC, *TechCrunch*, Al Jazeera, *El País*, *Corriere della Sera*, *The New York Times*, *Nómada*, *El Confidencial*, *Ha’Aretz*, *The Guardian*, *Wired*, *The New York Times Magazine*, *The Independent*, *Clarín* y *Foreign Policy*, entre otros.

Las citas y referencias a Byung-Chul Han provienen de un artículo que se publicó en el periódico español *El País* (véanse las notas de prensa). El comentario de Giorgio Agamben proviene de su escrito “La invención de una epidemia”. Las citas de Adam D. Moore se encuentran en su libro *Privacy Rights: Moral and Legal Foundations*. Algunos elementos importantes de la discusión sobre privacidad se extrajeron de “Engaging Privacy and Information Technology in a Digital Age,” un documento de análisis de la Academia Nacional de Ciencias de Estados Unidos. El informe sobre la encuesta del Government Business Council de Estados Unidos citado en este capítulo se titula “Is the Federal Government Ready for AI?” Las citas de Philip Alston, relator especial de Naciones Unidas, provienen del material de la ONU previo a la presentación de un informe especial sobre el Estado de bienestar digital ante la Asamblea

General en octubre de 2019. Los detalles de la visita del relator especial en temas de privacidad a Argentina en 2019 también provienen de material de prensa de ONU.

El análisis del Estado de vigilancia se basa en diversas fuentes. El caso chino ha sido muy bien documentado por James Leibold, cuyo estudio se encuentra en la bibliografía de este libro.

## Capítulo 7

### De las leyes para caballos de 1865 al mundo de hoy

La casuística proviene de material académico, histórico y periodístico de las siguientes fuentes: *The Washington Post*, Oxford University, la London School of Economics, *The New York Times*, Boeing, Airbus; Naciones Unidas, *Nature*, la Comisión Europea, los CDC de Estados Unidos, Algorithm Watch y *The Wall Street Journal*.

La exposición sobre la elaboración reactiva o proactiva de leyes como ejemplo de enseñanzas que pueden tomarse de otras industrias proviene de un capítulo de discusión sobre política energética del College of Earth and Mineral Sciences de PennState University. Los comentarios sobre ética y genética se basan en el documento “General Assembly Adopts United Nations Declaration on Human Cloning by vote of 84-34-37.”

Otro tema importante a considerar aquí es el de la privacidad. Esto exige determinar qué tipo de información se busca, quién la busca y cómo se debe recopilar, proteger y utilizar. Algunos especialistas en el área, miembros de las Academias de Ciencias de Estados Unidos, sostienen que si bien es cierto que hay cosas que no nos importaría que el gobierno supiera, no estamos dispuestos a permitir que se le revelen a un empleador (y viceversa). Y hay otras cosas que no objetaríamos que ambos supieran, pero que en cambio no querríamos compartir con familiares cercanos. Por la misma razón, hay aspectos que una persona solo desearía mantener en el ámbito de la familia. Según el estudio *Engaging Privacy and Information Technology*

in a Digital Age, “El equilibrio sobre los límites de la privacidad apunta a determinar qué tipo de cosas deben: (1) dejarse en el ámbito de la ética y la cortesía común, (2) ser incentivadas o desalentadas, y (3) formalizarse en la regulación o la ley.”

Durante la reunión de Barcelona de 2017 sobre ética de IA, posterior a la de Asilomar también reseñada en este capítulo, se lanzó un manifiesto básico titulado “Declaración de Barcelona para el desarrollo y uso apropiado de la inteligencia artificial.” Allí se establecen parámetros-guía sumamente útiles para aquellas regiones donde el trabajo en estos temas apenas comienza. La Declaración surgió del “Artificial Intelligence: Dreams, Risks, and Reality”, un encuentro organizado por el International Center for Scientific Debate de Barcelona y BioCat, una agencia del gobierno autonómico. “Muchos problemas fundamentales de la inteligencia artificial aún no se resuelven y requerirán avances radicales, mientras que otros quizá nunca se resuelvan”, dice el Manifiesto. “Por lo tanto, la aplicación de IA exige que quienes a ello se dedican tengan en cuenta requisitos muy estrictos; de lo contrario, los resultados podrían ser decepcionantes y potencialmente muy dañinos.”

En concreto:

- Se debe operar con **prudencia**. “La aplicación de IA basada en el conocimiento necesita de la experiencia humana y de recursos suficientes para analizarla y modelarla,” dice el Manifiesto de Barcelona. Del mismo modo, la IA basada en datos los requiere en suficiente cantidad y calidad para poder elegir con cuidado los algoritmos y parámetros apropiados. “Estos requisitos previos de la aplicación deben investigarse y explicarse con mucho más detalle para que los responsables del uso de la IA puedan ejercer la prudencia necesaria.”
- Los sistemas deben ser confiables, por lo cual es necesario realizar pruebas para garantizar que no se violen derechos personales, colectivos, normativas o leyes, en especial en dominios como la medicina o la robótica autónoma. Esa **fiabilidad** requiere trabajo urgente, puesto

que el grueso de los procedimientos de verificación y validación que tenemos actualmente se crearon para sistemas concebidos durante los años ochenta y noventa, cuando no existía IA basada en datos. El público debe tener confianza en que alguien está supervisando. En tal sentido, sería conveniente contar con una red de agencias regionales que usen métodos de verificación y validación estandarizados y comunes, tal como sucede en Europa con la agencia para la robótica y la inteligencia artificial del Parlamento Europeo. Hasta el momento, sin embargo, en América Latina y el Caribe no existe nada parecido en ninguno de los bloques subregionales: ni en Mercosur, ni en los países andinos, ni en Centroamérica.

- Se requiere **responsabilidad**. Hay múltiples decisiones cuyas implicaciones legales en temas que van desde seguros hasta impuestos, pasando por asuntos de vida o muerte, requieren explicaciones claras. Dice el manifiesto: “Cuando un sistema de IA toma una decisión, los humanos afectados por esta deberán poder obtener una explicación de por qué se la adopta en un lenguaje que puedan entender y también cuestionarla con argumentos razonados.” La mayoría de las decisiones de IA se sostienen en el cruzamiento de volúmenes enormes de datos y parámetros de red derivados estadísticamente. Las técnicas que dan sentido a estos parámetros son demasiado nuevas y requieren una combinación de mecanismos de control provenientes de la inteligencia artificial basada en el conocimiento y de la IA basada en datos.
- Debe haber **identificación**. A muchos nos resulta graciosa la prueba de Turing para determinar si una máquina pasa el examen de “humanidad,” pero en la vida real todos deberíamos saber con certeza si interactuamos con una persona o con una creación de IA. Lo mismo en lo que se refiere a los responsables del diseño de tal sistema.
- Debe haber **autonomía limitada**. Esto significa que en las reglas se deben establecer con claridad tanto los límites del comportamiento de la IA, como las funciones que

desempeñan los seres humanos y sus responsabilidades, sobre todo en caso de fallas o fracaso. Hay circunstancias en las que la IA toma decisiones por sí misma. Por ejemplo, cuando está integrada a un sistema físico como los vehículos autónomos, la IA actúa sobre el mundo real y eso produce consecuencias de vida o muerte. “Es comprensible que surjan preguntas sobre la seguridad y sobre si la IA autónoma no superará en algún momento a los humanos,” dice el manifiesto. “Aunque algunas de estas preocupaciones pertenecen más al dominio de la ciencia ficción que a la realidad, la circunscripción adecuada de los sistemas inteligentes autónomos y su alineación con los valores humanos es un desafío importante que debe abordarse.”

- La **función humana** debe estar bien definida. Es indiscutible que todos los sistemas de IA dependen críticamente de la inteligencia humana. Ya sean sistemas simbólicos basados en reglas o subsimbólicos basados en datos, ambos se basan críticamente en el comportamiento humano. De ello se deduce que mantener las habilidades humanas, incluso aquellas que se puedan automatizar, es fundamental. “El beneficio real proviene de la sinergia entre el intelecto humano y el artificial.”<sup>130</sup>

## BIBLIOGRAFÍA

### Introducción (Bibliografía)

- Report of the Special Rapporteur on Extreme Poverty and Human Rights,” por Philip Alston, United Nations, Human Rights Council, 11 de octubre de 2019. Disponible en: <https://undocs.org/en/A/74/493>
- *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* Shoshana Zuboff. 2019. New York: Public Affairs

### Capítulo 1 (Notas de Prensa)

- “Why was the 1918-19 Pandemic that Killed 50 Million People Known as ‘Spanish Flu’?”, History Extra, 19 de marzo de 2020. Disponible en <https://www.historyextra.com/period/first-world-war/why-was-spanish-flu-pandemic-known-called-that-where-did-name-come-from-spain-myth-coronavirus-covid-19-name/>
- “Beijing Bets on Facial Recognition in a Big Drive for Total Surveillance”, por Simon Denyer, *The Washington Post*, 7 de enero de 2018. Disponible en <https://www.washingtonpost.com/news/world/wp/2018/01/07/feature/in-china-facial-recognition-is-sharp-end-of-a-drive-for-total-surveillance/>
- “Facial Recognition Technology Helps Shenzhen Police to Identify Jaywalkers”, *South China Morning Post*, 27 de marzo de 2018. Disponible en <https://www.youtube.com/watch?v=ectdRsyj-zI>
- “Coronavirus: How Artificial Intelligence, Data Science and Technology Is Used to Fight the Pandemic”, por Bernard Marr, *Forbes*, 13 de marzo de 2020. Disponible en <https://www.forbes.com/sites/bernardmarr/2020/03/13/coronavirus-how-artificial-intelligence-data-science-and-technology-is-used-to-fight-the-pandemic/#6ffe6cdb5f5f>
- “Coronavirus: China’s Tech Fights Back”, por Patrick Jakhar, BBC, 3 de marzo de 2020. Disponible en <https://www.bbc.com/news/technology-51717164>
- “Coronavirus: The Detectives Racing to Contain the Virus in Singapore”, por Karishma Vaswani, BBC, 19 de marzo de 2020. Disponible en <https://www.bbc.com/news/world-asia-51866102>
- “How this South Korean Company Created Coronavirus Test Kits in Three Weeks”, por Ivan Watson, Sophie Jeong y Julia Hollingsworth, CNN, 13 de marzo de 2020. Disponible en <https://edition.cnn.com/2020/03/12/asia/coronavirus-south-korea-testing-intl-hnk/index.html>
- “Chinese Hospitals Deploy AI to Help Diagnose Covid-19”, por Tom Simonite, *Wired*, 26 de febrero de 2020. Disponible en <https://www.wired.com/story/chinese-hospitals-deploy-ai-help-diagnose-covid-19/>

<sup>130</sup> Aunque el aporte de los nuevos sistemas es significativo “en casi cualquier área”, según la Declaración de Barcelona “la experiencia humana aún supera con creces a la inteligencia artificial, particularmente para tratar casos que no han aparecido en los conjuntos de datos que configuran los ejemplos de los cuales están aprendiendo los sistemas de inteligencia artificial, y a menudo el beneficio real proviene de la sinergia entre la inteligencia humana y artificial.”

- “In a Time of Coronavirus, China’s Investment in AI is Paying off in a Big Way”, por Andy Chun, *South China Morning Post*, 18 de marzo de 2020. Disponible en <https://www.scmp.com/comment/opinion/article/3075553/time-coronavirus-chinas-investment-ai-paying-big-way>
- “How Long Will a Vaccine Really Take?”, por Stuart A. Thompson, *The New York Times*, 30 de abril de 2020. Disponible en <https://www.nytimes.com/interactive/2020/04/30/opinion/coronavirus-covid-vaccine.html>
- “IBM, Amazon, Google and Microsoft Partner with White House to Provide Computer Resources for COVID-19 Research”, por Frederic Lardinois, *TechCrunch*, 22 de marzo de 2020. Disponible en: <https://techcrunch.com/2020/03/22/ibm-amazon-google-and-microsoft-partner-with-white-house-to-provide-computer-resources-for-covid-19-research/>
- “Researchers Will Deploy AI to Better Understand Coronavirus”, por Will Knight, *Wired*, 3 de marzo de 2020. Disponible en <https://www.wired.com/story/researchers-deploy-ai-better-understand-coronavirus/>
- “2001: A Space Odyssey”: What it Means, and How it was Made”, por Dan Chiasson, *The New Yorker*, 16 de abril de 2018. Disponible en <https://www.newyorker.com/magazine/2018/04/23/2001-a-space-odyssey-what-it-means-and-how-it-was-made>
- “Taiwan’s Coronavirus Response is among the Best Globally”, por James Griffiths, CNN, 5 de abril de 2020. Disponible en <https://edition.cnn.com/2020/04/04/asia/taiwan-coronavirus-response-who-intl-hnk/index.html>
- “Taiwan Sets Example for World on How to Fight Coronavirus”, por Stacy Chen, ABC News, 13 de marzo de 2020. Disponible en <https://abcnews.go.com/Health/taiwan-sets-world-fight-coronavirus/story?id=69552462>
- “How Taiwan is Containing Coronavirus – Despite Diplomatic Isolation by China”, por Michelle Yun, *The Guardian*, 13 de marzo de 2020. Disponible en <https://www.theguardian.com/world/2020/mar/13/how-taiwan-is-containing-coronavirus-despite-diplomatic-isolation-by-china>
- “Cataluña aflora de golpe ante Sanidad 635 muertos y culpa a Colau del desfase”, por Darío Ojeda y Marcos Lamelas, *El Confidencial*, 22 de mayo de 2020. Disponible en [https://www.elconfidencial.com/espana/cataluna/2020-05-22/cataluna-desfase-sanidad-muertos-covid19\\_2607331/](https://www.elconfidencial.com/espana/cataluna/2020-05-22/cataluna-desfase-sanidad-muertos-covid19_2607331/)
- “How China is Using AI and Big Data to Fight the Coronavirus”, por Shawn Yuan, Al-Jazeera, 1 de marzo de 2020. Disponible en <https://www.aljazeera.com/news/2020/03/china-ai-big-data-combat-coronavirus-outbreak-200301063901951.html>
- “Coronavirus, come funzionano il controllo delle celle e il tracciamento dei contagi. Il Garante: «Non bisogna improvvisare»”, por Martina Pennisi, *Corriere della Sera*, 18 de marzo de 2020. Disponible en

[https://www.corriere.it/tecnologia/20\\_marzo\\_18/coronavirus-controlli-celle-telefoniche-tracciamento-privacy-223ea2c8-6920-11ea-913c-55c2df06d574.shtml](https://www.corriere.it/tecnologia/20_marzo_18/coronavirus-controlli-celle-telefoniche-tracciamento-privacy-223ea2c8-6920-11ea-913c-55c2df06d574.shtml)

- “Coronavirus, pronta la app italiana per tracciare i contagi: ‘Così possiamo fermare l’epidemia’”, por Elena Tebano, *Corriere della Sera*, 18 de marzo de 2020. Disponible en [https://www.corriere.it/tecnologia/20\\_marzo\\_18/coronavirus-pronta-app-italiana-tracciare-contagi-così-possiamo-fermare-l-epidemia-c6c31218-6919-11ea-913c-55c2df06d574.shtml?refresh\\_ce-cp](https://www.corriere.it/tecnologia/20_marzo_18/coronavirus-pronta-app-italiana-tracciare-contagi-così-possiamo-fermare-l-epidemia-c6c31218-6919-11ea-913c-55c2df06d574.shtml?refresh_ce-cp)
- “Por qué la solución tecnológica coreana es inviable ahora. Pero puede ser indispensable pronto”, por Jordi Pérez Colomé, *El País*, 23 de marzo de 2020. Disponible en <https://elpais.com/tecnologia/2020-03-23/por-que-la-solucion-tecnologica-coreana-es-inviable-ahora-pero-puede-ser-indispensable-pronto.html>
- “Congress Grilled the CEOs of Amazon, Apple, Facebook and Google. Here are the Big Takeaways”, por Brian Fung, CNN, 30 de julio de 2020. Disponible en <https://www.cnn.com/2020/07/29/tech/tech-antitrust-hearing-ceos/index.html>

## Capítulo 1 (Bibliografía)

- “1918 Pandemic (H1N1 virus)”. CDC. 2020. CDC. Disponible en <https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html>
- “Artificial Intelligence Can Help Solve Humanity’s Greatest Challenges”. ITU. 2017. ITU. Disponible en [https://www.itu.int/en/ITU-T/AI/Documents/Report/AI\\_for\\_Good\\_Global\\_Summit\\_Report\\_2017.pdf](https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf)
- “Inteligencia artificial: promesas y amenazas”. Ganascia, Jean-Gabriel. 2018. UNESCO. Disponible en <https://en.calameo.com/books/0060823762132d70e85e2>
- “AI in 2020: From Experimentation to Adoption”. Thomas, Rob. 2020. IBM. Disponible en <https://www.ibm.com/blogs/think/2020/01/ai-in-2020-from-experimentation-to-adoption/>
- “2 Years of Digital Transformation in 2 Months”. Spataro, Jared. 2020. Microsoft. Disponible en <https://www.microsoft.com/en-us/microsoft-365/blog/2020/04/30/2-years-digital-transformation-2-months/>
- “BIDMC Researchers Use Artificial Intelligence to Identify Bacteria Quickly and Accurately”. Mitchell, Jacqueline. 2015. Beth Israel Deaconess Medical Center (BIDMC). Disponible en Beth Israel Deaconess Medical Center <https://www.bidmc.org/about-bidmc/news/bidmc-researchers-use-artificial-intelligence-to-identify-bacteria-quickly-and-accurately>
- “Investors Poured \$4B into Healthcare AI Startups in 2019”. Landi, Heather. 2020. FierceHealthcare. Disponible en <https://www.fiercehealthcare.com/tech/investors-poured-4b-into-healthcare-ai-startups-2019>
- “Investment in AI for Healthcare Soars”. Lavender, Jonathan. 2018. KPMG. Disponible en <https://home.kpmg/xx/en/home/insights/2018/11/investment-in-ai-for-healthcare-soars.html>

- “Google Health - Tools to Help Healthcare Providers Deliver Better Care”. Google. 2019. Google. Disponible en <https://www.youtube.com/watch?v=P3SYqcPXqNk&list=PL590L5WQmH8e3dS9CtvRofbOnfdGb-Of9&index=3&t=3s>
- “General Session: David Feinberg, Head of Google Health”. HLTH. 2019. HLTH. Disponible en <https://www.youtube.com/watch?v=21A1CdmFcwE&list=PL590L5WQmH8e3dS9CtvRofbOnfdGb-Of9&index=12>
- “The Potential for Artificial Intelligence in Healthcare”. Davenport, Thomas y Kalakota, Ravi. 2019. Future Healthc. Disponible en <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616181/>
- “Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing”. Wang, CJ; Ng, CY; Brook, R. H. 2020. JAMA Network. Disponible en <https://shorturl.at/himBU>
- “Medical Knowledge Doubles Every Few Months; How Can Clinicians Keep Up?” Corish, Breda. 2018. Elsevier. Disponible en <https://www.elsevier.com/connect/medical-knowledge-doubles-every-few-months-how-can-clinicians-keepup>
- “History of Vaccination”. Plotkin, Stanley. 2014. Department of Pediatrics, University of Pennsylvania. Proceedings of the National Academy of Sciences of the United States of America. Disponible en <https://www.pnas.org/content/pnas/111/34/12283.full.pdf>
- “2018 Traffic Data for U.S Airlines and Foreign Airlines U.S. Flights”. Bureau of Transportation Statistics. 2019. Disponible en <https://www.bts.dot.gov/newsroom/2018-traffic-data-us-airlines-and-foreign-airlines-us-flights>
- “Number of Flights Performed by the Global Airline Industry from 2004 to 2021”. Mazareanu, E. 2020. Statista. Disponible en <https://www.statista.com/statistics/564769/airline-industry-number-of-flights/>
- “The Origin and Virulence of the 1918 ‘Spanish’ Influenza Virus”. Taubenberger, Jeffery. 2006. US National Library of Medicine National Institutes of Health. Disponible en <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2720273/#FNI>
- “World Air Transport Statistics”. 2019. International Air Transport Association. Disponible en <https://www.iata.org/contentassets/a686ff624550453e8bf0c9b3f7f0ab26/wats-2019-mediakit.pdf>
- “Computational Predictions of Protein Structures Associated with COVID-19”. 2020. Deep Mind. Disponible en <https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19>
- “How Baidu is Bringing AI to The Fight against Coronavirus”. Baidu. 2020. *MIT Technology Review*. Disponible en <https://www.technologyreview.com/s/615342/how-baidu-is-bringing-ai-to-the-fight-against-coronavirus/>
- “How Damo Academy’s AI System Detects Coronavirus Cases”. Li, Cecilia. 2020. Alizila, News from Alibaba Group. Disponible en <https://www.alizila.com/how-damo-academys-ai-system-detects-coronavirus-cases/>
- “How Artificial Intelligence is Aiding the Fight Against Coronavirus”. Johnson, Ashley. 2020. Center for Data Innovation. Disponible en <https://www.datainnovation.org/2020/03/how-artificial-intelligence-is-aiding-the-fight-against-coronavirus/>
- “Taiwan Shows its Mettle in Coronavirus Crisis, while the WHO is MIA”. Shapiro, Don. 2020. Brookings. Disponible en <https://www.brookings.edu/blog/order-from-chaos/2020/03/19/taiwan-shows-its-mettle-in-coronavirus-crisis-while-the-who-is-mia/>
- “Digital Identification: A Key to Inclusive Growth”. White, Olivia; Madgavkar, Anu; Manyika, James; Mahajan, Deepa; Bughin, Jacques; McCarthy, Michael; Sperling, Owen. 2019. McKinsey Global Institute (MGI). Disponible en <https://www.mckinsey.com/-/media/mckinsey/featured%20insights/innovation/the%20value%20of%20digital%20id%20for%20the%20global%20economy%20and%20society/mgi-digital-identification-a-key-to-inclusive-growth.ashx>
- “How Taiwan Used Big Data, Transparency and a Central Command to Protect Its People from Coronavirus”. Duff-Brown, Beth. 2020. Stanford, Freeman Spogli Institute for International Studies. Disponible en <https://fsi.stanford.edu/news/how-taiwan-used-big-data-transparency-central-command-protect-its-people-coronavirus>
- *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Topol, Eric. 2019. Basic Books, Inc.

## Capítulo 2 (Notas de Prensa)

- “The Third Industrial Revolution — Internet, Energy And A New Financial System”, por Gonçalo de Vasconcelos, *Forbes*, 4 de marzo de 2015. Disponible en <https://www.forbes.com/sites/goncalodevasconcelos/2015/03/04/the-third-industrial-revolution-internet-energy-and-a-new-financial-system/#10a803bc271a>
- “Ban on Killer Robots Urgently Needed, Say Scientists”, por Ian Sample, *The Guardian*, 13 de noviembre de 2017. Disponible en <https://www.theguardian.com/science/2017/nov/13/ban-on-killer-robots-urgently-needed-say-scientists>
- “Thousands of Leading AI Researchers Sign Pledge against Killer Robots”, por Ian Sample, *The Guardian*, 18 de julio de 2018. Disponible en <https://www.theguardian.com/science/2018/jul/18/thousands-of-scientists-pledge-not-to-help-build-killer-ai-robots>

## Capítulo 2 (Bibliografía)

- “Introduction to Social Work: An Advocacy-Based Profession”. (Capítulo 2). Cox, Lisa E.; Tice, Carolyn; Long, Dennis. 2015. SAGE Publications, Inc. Disponible en [https://us.sagepub.com/sites/default/files/upm-binaries/90576\\_Chapter\\_2\\_cox.pdf](https://us.sagepub.com/sites/default/files/upm-binaries/90576_Chapter_2_cox.pdf)
- “Anticipatory Government Preempting Problems through Predictive Analytics”. Eggers, William D. 2019. Deloitte. Disponible en <https://www2.deloitte.com/us/en/insights/industry/public-sector/government-trends/2020/predictive-analytics-in-government.html>

- “Public Health During the Industrial Revolution”. Wilde, Robert. 2019. ThoughtCo. Más información en <https://www.thoughtco.com/public-health-in-the-industrial-revolution-1221641>
- “Welfare Reform, 1834”. Clark, Gregory y Page, Marianne. 2008. UC Davis. Department of Economics, mayo.
- “Making Sense of Modernity’s Maladies: Health and Disease in the Industrial Revolution”. Brown, Michael. 2006. Science Direct. Disponible en Michael Brown, “Making sense of modernity’s maladies: health and disease in the Industrial Revolution”. Más información en [http://mguhhistory.weebly.com/uploads/8/0/0/5/8005631/making\\_sense\\_modernity.pdf](http://mguhhistory.weebly.com/uploads/8/0/0/5/8005631/making_sense_modernity.pdf)
- *Journeys to England and Ireland*. de Tocqueville, Alexis. 1958. New Haven, CT: Yale University Press, pp. 104-7.
- *Unlocking E-Government Potential: Concepts, Cases and Practical Insights*. Bhatnagar, Subhash. 2009. SAGE Publications, Inc.
- “The Limits of Semiconductor Technology and Oncoming Challenges in Computer Micro Architectures and Architectures”. Stojcev, K.M., Tokic, I. y Milentijevic, Z.I. 2004. Facta Universitatis. *Series Electronics and Energetics 17*. Disponible en <https://pdfs.semanticscholar.org/6bc6/ba19e94e0f7240cda69c870082eb0dfe54c2.pdf>
- “Computers and Economic Growth: Firm-Level Evidence”. Brynjolfsson, Erik y Hitt, Lorin. 1994. MIT Sloan School. Disponible en <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.698.2765&rep=rep1&type=pdf>
- “Introduction to Social Work: An Advocacy-Based Profession”, en *The History of Social Work*. Cox, Lisa E., Tice, Carolyn J. y Long, Dennis D. 2018. SAGE Publications Inc. Disponible en [https://us.sagepub.com/sites/default/files/upm-binaries/90576\\_Chapter\\_2\\_cox.pdf](https://us.sagepub.com/sites/default/files/upm-binaries/90576_Chapter_2_cox.pdf)
- “The Science and Technology of World War II”. Mindell, David. 2009. Science & Technology of World War II, National WWII Museum.. Disponible en <https://www.w2sci-tech.org/essays/essay2.html>
- “The Fourth Industrial Revolution”. Schwab, Klaus. 2017. World Economic Forum.
- *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. Brynjolfsson, Erik y McAfee, Andrew. 2016. W.W. Norton & Company.
- “Advancing Human-Centred Economic Progress in the Fourth Industrial Revolution”. Samans, Richard y Davis, Nicholas. 2017. World Economic Forum. Disponible en <https://www.weforum.org/whitepapers/advancing-human-centred-economic-progress-in-the-fourth-industrial-revolution>
- “¿Desarrollo económico inestable? Choques agregados en América Latina y el Caribe”. Báez, Javier E., Fuchs, Alan y Rodríguez-Castelán, Carlos. 2017. Banco Mundial. Disponible en [https://openknowledge.worldbank.org/bitstream/handle/10986/28892/Final%20Report%20-%20Aggregate%20Shocks%20and%20Economic%20Progress\\_Spanish%28web%29.pdf?sequence=4&isAllowed=y](https://openknowledge.worldbank.org/bitstream/handle/10986/28892/Final%20Report%20-%20Aggregate%20Shocks%20and%20Economic%20Progress_Spanish%28web%29.pdf?sequence=4&isAllowed=y)
- “The Fourth Industrial Revolution and the Future of Work in Argentina”. Albriou, Ramiro y Rapetti, Martín. 2019. GIZ-CIPPEC. Disponible <https://economic-policy-forum.org/wp-content/uploads/2016/10/Future-of-Work-WEB.pdf>
- “Bismarck Tried to End Socialism’s Grip—By Offering Government Healthcare”. Boissoneault, Lorraine. 2017. *Smithsonian Magazine*. Disponible en <https://www.smithsonianmag.com/history/bismarck-tried-end-socialisms-grip-offering-government-healthcare-180964064/>
- “On the Impact of Innovations on the Social Structure”. Dmitriev, Sergey; Kalinicheva, Valentina; Shadoba, Elena; Nikonets, Olesya; Pogonysheva, Dina y Shvarova, Elena. 2016. *International Journal of Economics and Financial Issues*. Disponible en <https://www.econjournals.com/index.php/ijefi/article/view/2362/pdf> International Journal of Economics and Financial Issues ISSN: 2146-4138 available at <http://www.econjournals.com> International Journal of Economics and Financial Issues, 2016, 6(S1) 107-113
- “How Artificial Intelligence is Transforming the World”. West, Darrell M. y Allen, John R. 2018. Brookings. Disponible en <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/>

### Capítulo 3 (Notas de Prensa)

- “An AI Epidemiologist Sent the First Warnings of the Wuhan Virus”, por Eric Miller, *Wired*, 25 de enero de 2020. Disponible en <https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/>
- “How this Canadian Start-up Spotted Coronavirus before Everyone Else Knew About It”, por Cory Stieg, 3 de marzo de 2020. Disponible en <https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html>
- “Alexa, Siri, Cortana: Why All Your Bots Are Female”, *The Wall Street Journal*, 22 de febrero de 2017. Disponible en <https://www.youtube.com/watch?v=Ly23QfCvom0>
- “Facebook’s Artificial Intelligence Robots Shut Down after They Start Talking to Each Other in Their Own Language”, por Andrew Griffin, *The Independent*, 31 de julio de 2017. Disponible en <https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>
- “‘The Discourse is Unhinged’: How the Media Gets AI Alarmingly Wrong”, por Oscar Schwartz, *The Guardian*, 25 de julio de 2018. Disponible en <https://www.theguardian.com/technology/2018/jul/25/ai-artificial-intelligence-social-media-bots-wrong>
- “The First ‘Robot Citizen’ in the World Once Said She Wants to ‘Destroy Humans’”, por Chris Weller, *Inc Magazine*, 26 de octubre 2017. Disponible en <https://www.inc.com/business-insider/sophia-humanoid-first-robot-citizen-of-the-world-saudi-arabia-2017.html>
- “No, Facebook’s Chatbots Will Not Take Over the World”, por Tim Simonite, *Wired*, 1 de agosto de 2017. Disponible en <https://www.wired.com/story/facebooks-chatbots-will-not-take-over-the-world/>

- “Twitter Taught Microsoft’s AI Chatbot to be a Racist Asshole in Less than a Day”, por James Vincent, *The Verge*, 24 de marzo de 2016. Disponible en <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>
- “Facebook Shuts Down Robots after They Invent Their Own Language”, por Matthew Field, *The Telegraph*, 1 de agosto de 2017. Disponible en <https://www.telegraph.co.uk/technology/2017/08/01/facebook-shuts-robots-invent-language/#:~:text=Facebook%20shut%20down%20a,mimicking%20human%20trading%20and%20bartering>.
- “AI Is Inventing Languages Humans Can’t Understand. Should We Stop It?”, por Mark Wilson, *FastCompany*, 14 de julio de 2017. Disponible en <https://www.fastcompany.com/90132632/ai-is-inventing-its-own-perfect-languages-should-we-let-it>
- “‘Robot Intelligence is Dangerous’: Expert’s Warning after Facebook AI ‘Develop their Own Language’”, por Louise Sassoon, *Mirror*, 1 de agosto de 2017. Disponible en <https://www.mirror.co.uk/tech/robot-intelligence-dangerous-experts-warning-1090871>
- “Facebook AI Researcher Slams ‘Irresponsible’ Reports about Smart Bot Experiment”, por Jordan Novet, *CNBC*, 1 de agosto de 2017. Disponible en <https://www.cnn.com/2017/08/01/facebook-ai-experiment-did-not-end-because-bots-invented-own-language.html>
- “The ‘Creepy Facebook AI’ Story that Captivated the Media”, por Chris Baraniuk, *BBC*, 1 de agosto de 2017. Disponible en <https://www.bbc.com/news/technology-40790258#:~:text=%22Robot%20intelligence%20is%20dangerous%3A,%22%2C%20says%20the%20Mirror.&text=While%20some%20great%20minds%20%2D%20including,nothing%20to%20be%20worried%20about>.
- “Computer AI Passes Turing Test in ‘World First’”, *BBC*, 9 de junio de 2014. Disponible en <https://www.bbc.com/news/technology-27762088>
- “Deal or No Deal? Training AI Bots to Negotiate”, por Mike Lewis, Denis Yarats, Devi Parikh y Dhruv Batra, *Facebook Engineering*, 14 de junio de 2017. Disponible en [https://engineering.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/?cft\\_%25B0%25D=AZU976xxfL9yDljhbnpHTYaNUZCAxTvxflEd2BEO7sfldsM5r1hnZyZlYjoCXX-gikceggkNk7tlxc5nT5MeGvsQFSSyYi7PxGseDrhwJ7wAjzyvhlEGGX\\_y2pCE9XCGQ&tn=-UK-R](https://engineering.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/?cft_%25B0%25D=AZU976xxfL9yDljhbnpHTYaNUZCAxTvxflEd2BEO7sfldsM5r1hnZyZlYjoCXX-gikceggkNk7tlxc5nT5MeGvsQFSSyYi7PxGseDrhwJ7wAjzyvhlEGGX_y2pCE9XCGQ&tn=-UK-R)
- “Inteligencia artificial y coronavirus: más ‘hype’ que realidad (por ahora)”, por Ramón López de Mántaras, *El País, Retina*, 25 de marzo de 2020. Disponible en [https://retina.elpais.com/retina/2020/03/25/tendencias/1585129233\\_858335.html](https://retina.elpais.com/retina/2020/03/25/tendencias/1585129233_858335.html)
- “Meet the Secret Algorithm That’s Keeping Students Out of College”, por Tom Simonite, *Wired*, 10 de julio de 2020. Disponible en <https://www.wired.com/story/algorithm-set-students-grades-altered-futures/>
- “AI Could Help with the Next Pandemic —But Not with this One”, por Will Douglas Heaven, *MIT Technology Review*, 12 de marzo de 2020. Disponible en <https://www.technologyreview.com/2020/03/12/905352/ai-could-help-with-the-next-pandemic-but-not-with-this-one/>

### Capítulo 3 (Bibliografía)

- *Superintelligence: Paths, Dangers, Strategies*. Bostrom, N. 2014. Oxford: Oxford University Press.
- “On the Current Paradigm on Artificial Intelligence”. *AI Communications* 27, N.1 Cristianini, N. 2014. Disponible en <https://doi.org/10.3233/AIC-130582>
- *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Springer. ISBN 978-1-4020-6710-5.
- “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”. Disponible en <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>
- *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* A K Peters/CRC Press; 2a edición.
- “Notes from the AI Frontier. Applying AI for Social Good”. McKinsey Global Institute. 2018. McKinsey & Company.
- *Artificial Intelligence in Society*. París. OECD Publishing.
- *Leakage in Data Mining: Formulation, Detection, and Avoidance Applied Predictive Modeling*. Kuhn, M., & Johnson, K. 2013. Springer New York.
- “Learning Representations by Back-propagating Errors”. *Nature*. doi:10.1038/323533a0
- *Minds, Brains, and Programs. Behavioral and Brain Sciences*. Searle, J.R. 1980.
- “Intriguing Properties of Neuronal Networks”. Szegedy, C., Wojciech, Z., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I. y Fergus, R. 2014.
- *Computing Machinery and Intelligence*. Turing, A. 1950. Disponible en <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>
- “This AI Outperformed 20 Corporate Lawyers at Legal Work” Disponible en <https://www.weforum.org/agenda/2018/11/this-ai-outperformed-20-corporate-lawyers-at-legal-work/>

### Capítulo 4 (Notas de Prensa)

- “S Korea’s Smartphone Apps Tracking Coronavirus Won’t Stop Buzzing”, por Kelly Kasulis, *Al-Jazeera*, 9 de abril de 2020. Disponible en <https://www.aljazeera.com/news/2020/04/korea-smartphone-apps-tracking-coronavirus-won-stop-buzzing-200408074008185.html>
- “Coronavirus: 4 países de América Latina que lograron aplicar estrategias exitosas de educación a distancia ante la pandemia”, por Ana Pais, *BBC News Mundo*, 24 de abril de 2020. Disponible en <https://www.bbc.com/mundo/noticias-america-latina-52375867>
- “Los problemas de usar datos del siglo pasado para una pandemia del siglo XXI”, por Kiko Llaneras, *El País*, 21 de junio de 2020. Disponible en <https://elpais.com/sociedad/2020-06-20/los-problemas-de-usar-datos-del-siglo-pasado-para-una-pandemia-del-siglo-xxi.html>
- “América Latina espera acelerar adopción de IA en medio de auge en naciones desarrollada”, *BNAméricas*, 9 de enero de 2020. Disponible

en <https://www.bnamericas.com/es/noticias/america-latina-espera-acelerar-adopcion-de-ia-en-medio-de-auge-en-naciones-desarrolladas>

- “Coronavirus en México: ¿qué hay detrás del súbito aumento de muertes por la pandemia de covid-19?”, por Alberto Nájara, BBC News Mundo, 4 de junio de 2020. Disponible en <https://www.bbc.com/mundo/noticias-america-latina-52930198>
- “Millions of Chinese, Cooped up and Anxious, Turn to Online Doctors”, *The Economist*, 5 de marzo de 2020. Disponible en <https://www.economist.com/business/2020/03/05/millions-of-chinese-cooped-up-and-anxious-turn-to-online-doctors>
- “What America can Learn from China’s Use of Robots and Telemedicine to Combat the Coronavirus”, por Tim Hornyak, CNBC, 18 de marzo de 2020. Disponible en <https://www.cnbc.com/2020/03/18/how-china-is-using-robots-and-telemedicine-to-combat-the-coronavirus.html>
- “Telemedicine Companies Are Struggling to Serve ‘Extreme Volumes’ of Patients as Coronavirus Calls Surge”, por Christina Farr, CNBC, 16 de marzo de 2020. Disponible en <https://www.cnbc.com/2020/03/16/telemedicine-companies-struggling-to-meet-coronavirus-demand.html>
- “No puedo parar: la pandemia desafía al vasto continente de la economía informal latinoamericana”, por Kirk Semple y Natalie Kitroeff, *The New York Times*, 30 de marzo de 2020. Disponible en <https://www.nytimes.com/es/2020/03/30/espanol/america-latina/coronavirus-trabajadores-informales.html>
- “Inteligencia Artificial: Cómo las empresas en Centroamérica ya la utilizan”, por José Barrera, 21 de mayo de 2018. Disponible en <https://www.estrategiaynegocios.net/lasclavesdelodia/1180122-330/inteligencia-artificial-c%C3%B3mo-las-empresas-en-centroam%C3%A9rica-ya-la-utilizan>
- “Argentina, el país con las startups de Inteligencia Artificial más grandes en la región”, *Ámbito Financiero*, 4 de octubre de 2018. Disponible en <https://www.ambito.com/ambito-biz/argentina-el-pais-las-startups-inteligencia-artificial-mas-grandes-la-region-n4035630>
- “Rappi inicia sus primeras entregas con robots”, *Forbes Centroamérica*, 21 de abril de 2020. Disponible en <https://forbescentroamerica.com/2020/04/21/rappi-inicia-sus-primeras-entregas-con-robots/>
- “Brasil: En 2020 la USP tendrá un centro de estudio de Inteligencia Artificial”, por Elizabeth Ángeles Uribe, Red UE-ALCUE, 25 de noviembre de 2019. Disponible en <https://observatoriodenoticias.redue-alcue.org/brasil-en-2020-la-usp-tendra-un-centro-de-estudio-de-inteligencia-artificial/>
- “Argentina: Usan inteligencia artificial contra el dengue”, por Martín de Ambrosio, 30 de mayo de 2018. Disponible en <https://www.scidev.net/america-latina/salud/noticias/argentina-usan-inteligencia-artificial-contra-el-dengue.html>
- “How AI Can Predict Heart Attacks and Strokes”, por Alice Park, *Time*, 14 de febrero de 2020. Disponible en <https://time.com/5784090/ai-heart-attack-stroke/>

## Capítulo 4 (Bibliografía)

- “Gasto público social”. CEPAL. 2018. CEPAL. Disponible en <https://observatoriosocial.cepal.org/inversion/es/indicador/gasto-social>
- “La educación en tiempos del Covid: Los sistemas educativos de América Latina y el Caribe ante COVID-19”. Álvarez Marinelli, Horacio; Arias Ortiz, Elena; Bergamaschi, Andrea; López Sánchez, Ángela; Noli, Alessandra; Ortiz Guerrero, Marcela; Pérez Alfaro, Marcelo; Rieble-Aubourg, Sabine; Rivera, María Camila; Scannone, Rodolfo; Vásquez, Madiery y Viteri, Adriana. 2020. IDB. Disponible en <https://publications.iadb.org/publications/spanish/document/La-educacion-en-tiempos-del-coronavirus-Los-sistemas-educativos-de-America-Latina-y-el-Caribe-ante-COVID-19.pdf>
- “Informalidad laboral y coronavirus: una combinación desafiante”. IADB. 2020. IADB. Disponible en <https://blogs.iadb.org/trabajo/es/informalidad-laboral-y-coronavirus-una-combinacion-desafiante/>
- “The Global Economic Outlook During the COVID-19 Pandemic: A Changed World”. The World Bank. 2020. Disponible en <https://www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world>
- “Se desarrolla en Colombia modelo de Inteligencia Artificial para la detección complementaria del Covid19”. News Center Microsoft Latinoamérica. 2020. News Center Microsoft Latinoamérica. Disponible en <https://news.microsoft.com/es-xl/se-desarrolla-en-colombia-modelo-de-inteligencia-artificial-para-la-deteccion-complementaria-del-covid19/>
- “Three Reasons Why Latin America is Under Cyber Attack”. IEEE. 2020. Disponible en <https://innovationatwork.ieee.org/latin-america-is-under-cyber-attack/>
- “Algorithms and Artificial Intelligence in Latin America: A Study of Implementation by Governments in Argentina and Uruguay”. Ortiz Freuler, J. e Iglesias, C. 2018. World Wide Web Foundation.
- *Artificial Intelligence in Society*. OECD. 2019. OECD Publishing. Disponible en <https://ec.europa.eu/jrc/communities/sites/jrccties/files/eedfee77-en.pdf>
- “Prometea, Artificial Intelligence to Transform Public”. Corvalán, Juan Gustavo. 2019. IMODEV. Disponible en <http://ois.imodev.org/index.php/RIDDN/article/view/287>
- “Electronic Medical Records Growing in Latin America”. Romerto Roy, Mariana. 2017. globalhealthintelligence. Disponible en <https://globalhealthintelligence.com/ghi-analysis/electronic-medical-records-growing-in-latin-america/#:~:text=Overall%2C%20GHI's%20research%20indicates%20that,the%20years%202014%20and%202019>
- “Transformación digital para la América Latina del S. XXI”. Silveira, Antonio. 2020. Banco de Desarrollo del América Latina CAF. Disponible en <https://www.caf.com/es/conocimiento/visiones/2020/02/transformacion-digital-para-la-america-latina-del-s21/>

- *El impacto de la inteligencia artificial en el emprendimiento*. Everis & Edevor. 2018. Disponible en <http://www.endeavor.cl/wp-content/uploads/El-impacto-de-la-IA-en-el-emprendimiento-en-Am%C3%A9rica-Latina-everis-y-Endeavor.pdf>
- *El futuro ya está aquí. Habilidades transversales de América Latina y el Caribe en el siglo XXI*. Mateo Díaz, Mercedes y Rucci, Graciana, editoras. 2019. IADB. Disponible en <https://publications.iadb.org/es/el-futuro-ya-est%C3%A1-aquí-habilidades-transversales-de-america-latina-y-el-caribe-en-el-siglo-xxi>
- “Hacia la transformación digital de América Latina: Programa TIC de CAF”. Agudelo, Mauricio. 2016. Banco de Desarrollo del América Latina CAF. Disponible en <https://www.itu.int/en/ITU-D/Regional-Presence/Americas/Documents/EVENTS/2016/15532-MX/D1-S3-1.pdf>
- “Coronavirus: un experimento de teletrabajo a escala mundial”. Ripani, Laura. 2020. IADB. Disponible en <https://blogs.iadb.org/trabajo/es/coronavirus-un-experimento-de-teletrabajo-a-escala-mundial/>
- “La Agenda Digital en América Latina: Un avance a distintas velocidades”. Porrúa, Miguel. 2019. IADB. Disponible en <https://blogs.iadb.org/administracion-publica/es/la-agenda-de-gobierno-digital-en-america-latina/>
- “CAF y OCDE impulsan la apertura, digitalización e innovación para la generación de valor público”. Noticias. Banco de Desarrollo del América Latina CAF. 2019. Disponible en <https://www.caf.com/es/actualidad/noticias/2019/12/caf-y-ocde-impulsan-la-apertura-digitalizacion-e-innovacion-para-la-generacion-de-valor-publico/>
- “fAIR LAC: Adopción ética y responsable de la inteligencia artificial en América Latina y el Caribe”. Pombo, Cristina; Cabrol, Marcelo; González Alarcón, Natalia; Sánchez Ávalos, Roberto. 2019. IADB. Disponible en <https://publications.iadb.org/es/fair-lac-adopcion-etica-y-responsable-de-la-inteligencia-artificial-en-america-latina-y-el-caribe>
- *Ciberseguridad ¿Estamos preparados en América Latina y el Caribe?* Observatorio de la Ciberseguridad en América Latina y el Caribe. BID y Organization of American States. 2016. Disponible en <https://publications.iadb.org/publications/spanish/document/Ciberseguridad-%C2%BFEstamos-preparados-en-Am%C3%A9rica-Latina-y-el-Caribe.pdf>

## Capítulo 5 (Notas de Prensa)

- “Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women”, por Jeffrey Dastin, Reuters, 11 de octubre de 2018. Disponible en <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCNIMK08G>
- “La lista de los 93 millones de mexicanos que votan, expuesta en Amazon”, por Sonia Corona, *El País*, 23 de abril de 2026. Disponible en [https://elpais.com/internacional/2016/04/23/mexico/1461369407\\_447969.html](https://elpais.com/internacional/2016/04/23/mexico/1461369407_447969.html)
- “The Extraordinary Decisions Facing Italian Doctors”, por Yascha Mounk”, *The Atlantic*, 11 de marzo de 2020. Disponible en [https://www.theatlantic.com/ideas/archive/2020/03/who-gets-hospital-bed/607807/?fbclid=IwAR37PaWlryqk8jHICMLGcr1DJz\\_47qBkRtHK\\_31RnUu64RK1av7R5IIPUqw](https://www.theatlantic.com/ideas/archive/2020/03/who-gets-hospital-bed/607807/?fbclid=IwAR37PaWlryqk8jHICMLGcr1DJz_47qBkRtHK_31RnUu64RK1av7R5IIPUqw)
- “En Salta usan inteligencia artificial para prever embarazos adolescentes”, por Rosario Medina, *Clarín*, 12 de Abril 2018, Disponible en [https://www.clarin.com/sociedad/salta-usan-inteligencia-artificial-prever-embarazos-adolescentes\\_0\\_r10wlG6jf.html](https://www.clarin.com/sociedad/salta-usan-inteligencia-artificial-prever-embarazos-adolescentes_0_r10wlG6jf.html)
- “Rise of the Deepfakes”, *The Week*, 9 de junio de 2018. Disponible en <https://theweek.com/articles/777592/rise-deepfakes>
- “Movimiento Ciudadano admite que subió los listados nominales a Amazon”, *Expansión*, 27 de abril de 2016. Disponible en <https://expansion.mx/politica/2016/04/27/movimiento-ciudadano-admite-que-subieron-el-padron-electoral-a-amazon>
- “A.I. Is Learning to Read Mammograms”, por Denise Grady, *The New York Times*, 1 de enero de 2020. Disponible en <https://www.nytimes.com/2020/01/01/health/breast-cancer-mammogram-artificial-intelligence.html>
- “Alexa Has Been Eavesdropping on You this Whole Time”, por Geoffrey A. Fowler, *The Washington Post*, 6 de Mayo de 2019, Disponible en <https://www.washingtonpost.com/technology/2019/05/06/alexa-has-been-eavesdropping-you-this-whole-time/>
- “When Amazon Flexes Its Power”, por Shira Ovide, *The New York Times*, 11 de Junio de 2020, Disponible en <https://www.nytimes.com/2020/06/11/technology/amazon-competitors.html>
- “Does Social Desirability Favor Humans? Explicit-Implicit Evaluations of Synthesized Speech Support: A New HCI Model Of Impression Management”. Chin-Chang, Ho; Wade, Mitchell; Himalaya, Patel; Maccorman, Karl. 2011. Computers in Human Behavior. Disponible en [https://www.academia.edu/31398003/Does\\_social\\_desirability\\_bias\\_favor\\_humans\\_Explicit\\_implicit\\_evaluations\\_of\\_synthesized\\_speech\\_support\\_a\\_new\\_HCI\\_model\\_of\\_impression\\_management](https://www.academia.edu/31398003/Does_social_desirability_bias_favor_humans_Explicit_implicit_evaluations_of_synthesized_speech_support_a_new_HCI_model_of_impression_management)
- “Digital Identification: A Key to Inclusive Growth”. White, Olivia; Madgavkar, Anu; Manyika, James; Mahajan, Deepa; Bughin, Jacques; McCarthy, Michael y Sperling, Owen. 2019. McKinsey Global Institute. Disponible en <https://www.mckinsey.com/-/media/mckinsey/featured%20insights/innovation/the%20value%20of%20digital%20id%20for%20the%20global%20economy%20and%20society/mgi-digital-identification-a-key-to-inclusive-growth.ashx>
- “Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations”. Obermeyer, Ziad; Powers, Brian; Vogeli, Christine y Mullainathan, Sendhil. 2019. *Science*. American Association for the Advancement of Science. Disponible en <https://science.sciencemag.org/content/366/6464/447>

- “Ley de vivienda”. Cámara de Diputados, H. Congreso de la Unión. Última reforma 2019. Disponible en [http://www.diputados.gob.mx/LevesBiblio/pdf/LViv\\_140519.pdf](http://www.diputados.gob.mx/LevesBiblio/pdf/LViv_140519.pdf)
- “Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000–2015”. WHO. 2016. Ginebra: WHO.
- “Adding It up: Costs and Benefits of Meeting the Contraceptive Needs of Adolescents”. Darroch J., Woog V., Bankole A. y Ashford, L. S. 2016. New York: Guttmacher Institute.
- “What Do We Talk When We Talk About AI: Algorithmic Decision-Making in Latin America”. Canales, María Paz. 2020. Derechos Digitales. Disponible en <https://www.derechosdigitales.org/wp-content/uploads/glimpse-2019-4-eng.pdf>
- “Avanza el uso de la Inteligencia Artificial en la Argentina con experiencias en el sector público, privado y ONGs”, News Center Microsoft Latinoamérica, 2 de abril del 2018. Disponible en <https://news.microsoft.com/es-xl/avanza-el-uso-de-la-inteligencia-artificial-en-la-argentina-con-experiencias-en-el-sector-publico-privado-y-ong/>
- “Pregnancy and Childbirth Outcomes among Adolescent Mothers: A World Health Organization Multicountry Study”. Ganchimeg T, et al. 2014. *Bjog*. 2014;121(S Suppl 1):40-8
- “El embarazo en la adolescencia”. OMS. 2020. OMS. Disponible en <https://www.who.int/es/news-room/fact-sheets/detail/adolescent-pregnancy>
- “Estado de la ciberseguridad en el sector bancario en América Latina y el Caribe”. Organización de los Estados Americanos (OEA). 2018. Programa de Ciberseguridad del Comité Interamericano contra el Terrorismo. Disponible en <http://www.oas.org/es/sms/cicte/sectorbancariospa.pdf>
- “ESET Security Report Latinoamérica 2017”. Enjoy Safer Technology ESET. 2017. Disponible en <https://www.welivesecurity.com/wp-content/uploads/2017/04/eset-security-report-2017.pdf>
- “ESET Security Report Latinoamérica 2019”. Enjoy safer technology ESET. 2019. Disponible en <https://www.welivesecurity.com/wp-content/uploads/2019/07/ESET-security-report-LATAM-2019.pdf>
- “fAIR LAC: Adopción ética y responsable de la inteligencia artificial en América Latina y el Caribe”. Pombo, Cristina; Cabrol, Marcelo; González Alarcón, Natalia; Sánchez Ávalos, Roberto. 2020. IDB. Disponible en <https://publications.iadb.org/es/fair-lac-adopcion-etica-y-responsable-de-la-inteligencia-artificial-en-america-latina-y-el-caribe>
- “Algoritmos e inteligencia artificial en Latinoamérica: Un estudio de implementaciones por parte de gobiernos en Argentina y Uruguay”. Ortiz Freuler, J. e Iglesias, C. 2018. World Wide Web Foundation. Disponible en [https://webfoundation.org/docs/2018/09/WF\\_AI-in-LA\\_Report\\_Spanish\\_Screen\\_AW.pdf](https://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf)
- “To Save Everything, Click Here: The Folly of Technological Solutionism” Morozov, Evgeny. 2014. United States: PublicAffairs.

## Capítulo 6 (Notas de Prensa)

- “One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority”, por Paul Mozur, *The Washington Post*, 14 de abril de 2019. Disponible en <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>
- “Inside China’s Dystopian Dreams: A.I., Shame and Lots of Cameras”, por Paul Mozur, *The New York Times*, 8 de julio de 2018. Disponible en <https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html>
- “Facebook Agrees to Pay U.K. Fine Over Cambridge Analytica Scandal”, por Parmy Olson, *The Wall Street Journal*, 30 de octubre de 2019. Disponible en <https://www.wsj.com/articles/facebook-agrees-to-pay-u-k-fine-over-cambridge-analytica-scandal-11572442488>
- “Big Tech Faces a ‘Big Brother’ Trap on Coronavirus”, por Nancy Scola, *Político*, 18 de marzo de 2020. Disponible en <https://www.politico.com/news/2020/03/18/big-tech-coronavirus-134523>
- “How Researchers Learned to Use Facebook ‘Likes’ to Sway Your Thinking”, por Keith Collins y Gabriel J.X. Dance, *The New York Times*, 20 de Marzo de 2018 Disponible en <https://www.nytimes.com/2018/03/20/technology/facebook-cambridge-behavior-model.html>
- “La emergencia viral y el mundo de mañana. Byung-Chul Han, el filósofo surcoreano que piensa desde Berlín”, por Byung-Chul Han, *El País*, 22 de marzo de 2020. Disponible en <https://elpais.com/ideas/2020-03-21/la-emergencia-viral-y-el-mundo-de-manana-byung-chul-han-el-filosofa-surcoreano-que-piensa-desde-berlin.html>
- “Using Texts as Lures, Government Spyware Targets Mexican Journalists and Their Families”, por Azam Ahmed y Nicole Perlroth, *The New York Times*, 19 de junio de 2017. Disponible en <https://www.nytimes.com/2017/06/19/world/americas/mexico-spyware-anticrime.html>
- “El CIS os pregunta si queréis un Ministerio de la Verdad”, por Juan Soto Ivars, *El Confidencial*, 16 de abril de 2020. Disponible en [https://blogs.elconfidencial.com/sociedad/espana-is-not-spain/2020-04-16/cis-pregunta-limitar-redes-medios-bulos-verdad\\_2551640/](https://blogs.elconfidencial.com/sociedad/espana-is-not-spain/2020-04-16/cis-pregunta-limitar-redes-medios-bulos-verdad_2551640/)
- “Espionaje ilegal del Gobierno: Aquí está la investigación de Nuestro Diario (Parte I)”, *Nuestro Diario*, 6 de agosto de 2018. Disponible en <https://nomada.gt/pais/la-corrupcion-no-es-normal/espionaje-ilegal-del-gobierno-aqui-esta-la-investigacion-de-nuestro-diario-parte-i/>
- “Even Mask-wearers Can Be ID’d, China Facial Recognition Firm Says”, por Martin Pollard, *Reuters*, 9 de marzo de 2020. Disponible en <https://www.reuters.com/article/us-health-coronavirus-facial-recognition/even-mask-wearers-can-be-idd-china-facial-recognition-firm-says-idUSKBN20WOWL>
- “Chinese Startup Rokid Pitches COVID-19 Detection Glasses in US”, por Jake Bright y Rita Liao, *TechCrunch*, 16 de abril de 2020. Disponible en <https://techcrunch.com/2020/04/16/chinese-startup-rokid-pitches-covid-19-detection-glasses-in-u-s/>

## Capítulo 6 (Bibliografía)

- “Surveillance in China’s Xinjiang Region: Ethnic Sorting, Coercion, and Inducement”. Leibold, James. 2019. *Journal of Contemporary China*. Disponible en <https://www.tandfonline.com/doi/full/10.1080/10670564.2019.1621529?scroll=top&needAccess=true>
- “Ethics in Artificial Intelligence”. López de Mantarás, Ramón. 2017. *Investigación y Ciencia*. Disponible en <https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/el-multiverso-cuntico-711/tica-en-la>
- “To Be Let Alone: Brandeis Foresaw Privacy Problems”. Burrows, Leah. 2013. Brandeis University. Disponible en <https://www.brandeis.edu/now/2013/july/privacy.html>
- *Engaging Privacy and Information Technology in a Digital Age*. Waldo, James; Lin, Herbert S.; Millett, Lynette I., editores. 2007. National Research Council, The National Academies Press. Disponible en <https://doi.org/10.17226/11896>
- “Declaración a los medios de comunicación del Relator Especial sobre el derecho a la privacidad, al concluir su visita oficial a la Argentina del 6 al 17 de mayo de 2019”. Oficina del Alto Comisionado para los Derechos Humanos (ACNUDH). 2019. Disponible en <https://www.ohchr.org/SP/NewsEvents/Pages/DisplayNews.aspx?NewsID=24639&LangID=S>
- “Dynamic Neural Network Architecture Inspired by the Immune Algorithm to Predict Preterm Deliveries in Pregnant Women”. Hussaina, A. J.; Fergusa P.; Al-Askara, H.; Al-Jumeilya, D. y Jagerb, F. 2015. LJM Research Online. Disponible en <http://researchonline.limu.ac.uk/id/eprint/2369/3/Preterm%20Deliveries%20from%20EHG%20using%20Machine%20Learning%20Algorithms%20-%20Final%5B1%5D.pdf>
- “Expert System for Predicting the Early Pregnancy with Disorders Using Artificial Neural Network”. Maylawati, D. S.; Ramdhani, M. A.; Zulfikar, W. B.; Taufik, I. y Darmalaksana, W. 2017. IEEE. Disponible en <https://ieeexplore.ieee.org/document/8089243>
- “The Associations of Birth Intervals with Small-for-Gestational-Age, Preterm, and Neonatal and Infant Mortality: A Meta-analysis”. Kozuki, N.; Lee, A.; Silveira, M. et al. 2013. *BMC Public Health*. 13 (Suppl. 3):S3.
- “Justiça determina que empresa pare de coletar dados faciais do metrô em SP”. Instituto Brasileiro de Defesa do Consumidor (IDEC). 2019. Disponible en <https://idec.org.br/idec-na-imprensa/justica-determina-que-empresa-pare-de-coletar-dados-faciais-do-metro-em-sp>
- “Despantan las reformas en materia de protección de datos en América Latina”. Bojalil, Paulina y Vela-Treviño, Carlos. 2019. IDB. Disponible en <https://blogs.iadb.org/conocimiento-abierto/es/proteccion-de-datos-gdpr-america-latina/>
- “World Stumbling Zombie-like into a Digital Welfare Dystopia, Warns UN Human Rights Expert”. Alston, Philip. 2019. OHCHR. Disponible en <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25156>

- “The Science Behind Cambridge Analytica: Does Psychological Profiling Work?” Andrews, Edmund L. 2018. Stanford Graduate School of Business. Disponible en <https://www.gsb.stanford.edu/insights/science-behind-cambridge-analytica-does-psychological-profiling-work>
- “La invención de una epidemia”. Agamben, Giorgio. 2020. Disponible en <https://ficcionalarazon.org/2020/02/27/giorgio-agamben-la-invention-de-una-epidemia/>
- “Wait No More: Citizens, Red Tape, and Digital Government”. Roseth, Benjamin; Reyes, Ángela; Fariás, Pedro; Porrúa, Miguel; Villalba, Harold; Acevedo, Sebastián; Peña, Norma; Estevez, Elsa; Linares Lejarraga, Sebastián; Fillottrani, Pablo. 2018. IADB. Disponible en <https://publications.iadb.org/en/wait-no-more-citizens-red-tape-and-digital-government>

## Capítulo 7 (Notas de Prensa)

- “China’s Coronavirus Battle Is Waning. Its Propaganda Fight Is Not”, por Vivian Wang, *The New York Times*, 8 de abril de 2020. Disponible en <https://www.nytimes.com/2020/04/08/world/asia/coronavirus-china-narrative.html>
- “Adoption of Veritone’s AI-Powered Public Safety Solutions Grows Across the U.S. as Public Safety Agencies Seek to Increase Transparency and Accountability”, por Allison Zullo y Walker Sands, Veritone, *Street Insider*, 14 de julio de 2020. Disponible en <https://www.streetinsider.com/Business+Wire/Adoption+of+Veritone%E2%80%99s+AI-Powered+Public+Safety+Solutions+Grows+Across+the+U.S.+as+Public+Safety+Agencies+Seek+to+Increase+Transparency+and+Accountability/17109863.html>
- “Un hombre estuvo seis días preso por un error policial”, *Infobae*, 2 de agosto de 2019. Disponible en <https://www.infobae.com/sociedad/policiales/2019/08/02/un-hombre-estuvo-seis-dias-presos-por-un-error-del-sistema-de-reconocimiento-facial/>
- “Los ‘errores’ del sistema de reconocimiento facial: detuvieron a una mujer por su parecido con una prófuga”, *Nueva ciudad*, 17 de mayo de 2019. Disponible en <https://www.nueva-ciudad.com.ar/notas/201905/40692-los-errores-del-sistema-de-reconocimiento-facial-detuvieron-a-una-mujer-por-su-parecido-con-una-profuga.html>
- “How Tech Companies Track Your Every Move And Put Your Data Up For Sale”, NPR, 31 de julio de 2019. Disponible en <https://www.npr.org/2019/07/31/746878763/how-tech-companies-track-your-every-move-and-put-your-data-up-for-sale>
- “San Francisco is First US City to Ban Facial Recognition”, por Dave Lee, BBC, 14 de mayo de 2019. Disponible en <https://www.bbc.com/news/technology-48276660>
- “Un relator especial de la ONU denuncia la falta de privacidad y el alto nivel de escuchas en Argentina”, *Notimérica/EuropaPress*, 24 de mayo de 2019. Disponible en <https://www.notimerica.com/politica/noticia-argentina-relator-especial-onu-denuncia-falta-privacidad-alto-nivel-escuchas-argentina-20190524043445.html>

- “Sent to Prison by a Software Program’s Secret Algorithms”, por Adam Liptak, *The New York Times*, 1 de mayo de 2017. Disponible en <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>
- “Estas son las nuevas patrullas de la CDMX: tienen tecnología de reconocimiento facial y de autos robados”, *Animal Político*, 9 de julio de 2019. Disponible en <https://www.animalpolitico.com/2019/07/nuevas-patrullas-cdmx/>
- “The Technology That Could Free America From Quarantine”, por Derek Thompson, *The Atlantic*, 7 de abril de 2020. Disponible en <https://www.theatlantic.com/ideas/archive/2020/04/contact-tracing-could-free-america-from-its-quarantine-nightmare/609577/>
- “Shoshana Zuboff on Surveillance Capitalism’s Threat to Democracy”, por Noah Kulwin, *New York Magazine*, 24 de febrero de 2019. Disponible en <https://nymag.com/intelligencer/2019/02/shoshana-zuboff-q-and-a-the-age-of-surveillance-capital.html>
- “The China Way: Use of technology to Combat Covid-19”, por Aditya Chaturvedi, *Geospatial World*, 11 de mayo de 2020. Disponible en <https://www.geospatialworld.net/article/the-sino-approach-use-of-technology-to-combat-covid-19/>
- “How Tech’s Lobbyists Are Using the Pandemic to Make Gains,” por David McCabe, *The New York Times*, 3 de abril de 2020. Disponible en <https://www.nytimes.com/2020/04/03/technology/virus-tech-lobbyists-gains.html>
- “‘The New Normal’: China’s Excessive Coronavirus Public Monitoring Could be Here to Stay”, por Lily Kuo, *The Guardian*, 9 de marzo de 2020. Disponible en <https://www.theguardian.com/world/2020/mar/09/the-new-normal-chinas-excessive-coronavirus-public-monitoring-could-be-here-to-stay>
- “How Bad Will the Coronavirus Outbreak Get? Here Are 6 Key Factors”, por Knvul Sheikh, Derek Watkins, Jin Wu y Mika Gröndahl, *The New York Times*, 28 de febrero de 2020. Disponible en <https://www.nytimes.com/interactive/2020/world/asia/china-coronavirus-contain.html>
- “Ecuador’s All-Seeing Eye Is Made in China”, por Charles Rollet, *Foreign Policy*, 9 de agosto de 2018. Disponible en <https://foreignpolicy.com/2018/08/09/ecuadors-all-seeing-eye-is-made-in-china/>
- “Hong Kong Protesters Use Lasers to Avoid Facial Recognition Cameras and Blind Police”, por Anthony Cuthbertson, *The Independent*, 1 de agosto de 2019. Disponible en <https://www.independent.co.uk/news/world/asia/hong-kong-protests-lasers-facial-recognition-ai-china-police-a9033046.html>
- “Por qué los manifestantes chilenos usan punteros láser en las marchas”, *Clarín*, 13 de noviembre de 2019. Disponible en [https://www.clarin.com/mundo/manifestantes-chilenos-usan-punteros-lasers-marchas\\_0\\_FeU7smoT.html](https://www.clarin.com/mundo/manifestantes-chilenos-usan-punteros-lasers-marchas_0_FeU7smoT.html)
- “Facebook Exposed 87 Million Users to Cambridge Analytica”, por Issie Lapowsky, *Wired*, 4 de abril de 2018. Disponible en <https://www.wired.com/story/facebook-exposed-87-million-users-to-cambridge-analytica/>
- “Hong Kong Protesters Coordinate Tech-Savvy Effort to Beat Chinese State Surveillance”, por Shibani Mahtani, *The Independent*, 16 junio de 2019. Disponible en <https://www.independent.co.uk/news/world/asia/hong-kong-protests-china-surveillance-tech-telegram-extradition-bill-a8960911.html>
- “Fresh Cambridge Analytica Leak ‘Shows Global Manipulation Is Out of Control’”, por Carole Cadwalladr, *The Guardian*, 4 de junio de 2020. Disponible en <https://www.theguardian.com/uk-news/2020/jan/04/cambridge-analytica-data-leak-global-election-manipulation>
- “Google, Democracy and the Truth about Internet Search”, por Carole Cadwalladr, *The Guardian*, 4 de diciembre de 2016. Disponible en <https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook>
- “Google’s 4,000-Word Privacy Policy Is a Secret History of the Internet”, por Charlie Warzel y Ash Ngu, *The New York Times*, The Privacy Project, 10 de julio de 2019. Disponible en <https://www.nytimes.com/interactive/2019/07/10/opinion/google-privacy-policy.html>
- “How a Pentagon Contract Became an Identity Crisis for Google”, por Scott Shane, Cade Metz y Daisuke Wakabayashi, *The New York Times*, 30 de mayo de 2018. Disponible en <https://www.nytimes.com/2018/05/30/technology/google-project-maven-pentagon>
- “Dr. Fei Fei Li, Google’s AI Cloud Head Steps Down amidst Speculations; Dr. Andrew Moore to Take her Place”, por Melisha Dsouza, *Packt*, 11 de septiembre de 2018. Disponible en <https://hub.packtpub.com/dr-fei-fei-li-googles-ai-cloud-head-steps-down-amidst-speculations-dr-andrew-moore-to-take-her-place/>
- “Google Promises Its A.I. Will Not Be Used for Weapons”, por Daisuke Wakabayashi y Cade Metz, *The New York Times*, 7 de junio de 2018. Disponible en <https://www.nytimes.com/2018/06/07/technology/google-artificial-intelligence-weapons.html>
- “IBM Will No Longer Offer, Develop, or Research Facial Recognition Technology”, por Jay Peters, *The Verge*, 8 de junio de 2020. Disponible en <https://www.theverge.com/2020/6/8/21284683/ibm-no-longer-general-purpose-facial-recognition-analysis-software>
- “In the Realm of Paper Tigers – Exploring the Failings of AI Ethics Guidelines”. Leonard Haas and Sebastian Giebler. *Algorithm Watch*. 28 de abril de 2020. Disponible en <https://algorithmwatch.org/en/ai-ethics-guidelines-inventory-upgrade-2020/>
- “A \$1 Million Research Bargain for Netflix, and Maybe a Model for Others”, por Steve Lohr, *The New York Times*, 21 de Septiembre de 2019. Disponible en <https://www.nytimes.com/2009/09/22/technology/internet/22netflix.html>
- “Horses, Equine Law and The Future of the Autonomous Vehicle Legal Framework”, por Rahul Razdan, *Forbes*, 7 de junio de 2020. Disponible en <https://www.forbes.com/sites/rahulrazdan/2020/01/07/horses-equine-law-and-the-future-of-the-autonomous-vehicle-legal-framework/#2f8ecd6525d8>

- “Boeing Built Deadly Assumptions Into 737 Max, Blind to a Late Design Change”, por Jack Nicass, Nathalie Kitroeff, David Gelles y James Glanz, *The New York Times*, 1 de junio de 2019. Disponible en <https://www.nytimes.com/2019/06/01/business/boeing-737-max-crash.html>
- “The Real Reason Boeing’s New Plane Crashed Twice”, por Alvin Chang, Dion Lee y Kimberly Mas, *Vox*, 15 de abril de 2019. Disponible en <https://www.vox.com/videos/2019/4/15/18306644/boeing-737-max-crash-video>

## Capítulo 7 (Bibliografía)

- “General Assembly Adopts United Nations Declaration on Human Cloning by Vote of 84-34-37”. United Nations. 2005. Disponible en <https://www.un.org/press/en/2005/ga10333.doc.htm>
- “Preliminary Report Highway HWY18MH010”. NTSB. 2018. Disponible en <https://www.nts.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf>
- “Artificial Intelligence: dreams, risks and reality”. International Center for Scientific Debate Barcelona. 2017. Disponible en [https://www.biocat.cat/sites/default/files/sinopsibdebate\\_artintelligence\\_es.pdf](https://www.biocat.cat/sites/default/files/sinopsibdebate_artintelligence_es.pdf)
- “Leveraging Private Data for Public Good, A Descriptive Analysis and Typology of Existing Practices”. Verhulst, Stefaan G; Young, Andrew; Winowatan, Michelle y Zahuranec, Andrew J. 2019. GovLab, New York University (NYU) Tandon School of Engineering. Diponible en <https://datacollaboratives.org/static/files/existing-practices-report.pdf>
- “Barcelona Declaration for the Proper Development and Usage of Artificial Intelligence in Europe”. Biocat & Obra Social la Caixa, Barcelona. 2017. Disponible en <https://www.iiia.csic.es/barcelonadeclaration/>
- “Ethics Guidelines for Trustworthy AI”. European Commission. 2019. Disponible en <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- “Artificial Intelligence: the Global Landscape of Ethics Guidelines”. Jobin, Anna; Ienca, Marcello; y Bayena, Effy. 2019. Cornell University. Disponible en <https://arxiv.org/abs/1906.11668>
- “In the Realm of Paper Tigers – Exploring the Failings of AI Ethics Guidelines”. Hass, Leonard y Gießler, Sebastian, con la asistencia de Veronika Thiel. 2020. *AlgorithmWatch*. Disponible en <https://algorithmwatch.org/en/ai-ethics-guidelines-inventory-upgrade-2020/>
- “Where Are Human Subjects in Big Data Research? The Emerging Ethics Divide”. Metcalf, Jacob y Crawford, Kate. 2016. *SAGE Journal*. Disponible en <https://journals.sagepub.com/doi/10.1177/2053951716650211>
- “China Set to Introduce Gene-editing Regulation Following CRISPR-Baby Furore”. Cyranoski, David. 2019. *Nature*. Disponible en <https://www.nature.com/articles/d41586-019-01580-1>
- “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”. Buolamwini, Joy y Geburu, Timnit. 2018. MIT & Microsoft. Disponible en <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>
- “A Layered Model for AI Governance”. Gasser, Urs y Almeida, Virgilio A. F. 2017. *IEEE Internet Computing* 21. Disponible en <https://dash.harvard.edu/bitstream/handle/1/34390353/w6gov-18-LATEX.pdf?sequence=1>
- “Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI”. Jessica Fjeld, et al. *Berkman Klein Center Research*. 2020. Disponible en [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3518482](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518482)
- “Identifying Participants in the Personal Genome Project by Name”, Sweeney, Latanya, Abu, Akua y Winn, Julia. Data Privacy Lab, IQSS, Harvard University, 2013. Disponible en <https://privacytools.seas.harvard.edu/publications/identifying-participants-personal-genome-project-name>
- *Privacy Rights: Moral and Legal Foundations*. Moore, Adam D. 1965. The Pennsylvania State University Press. Págs. 205-06.
- *A Common Sense Digest of American Negligence Cases: 17 Volumes*. The University of Chicago. 1914. Callaghan and company. Disponible en [https://books.google.com/books?id=tbpFAQAAMAAJ&pg=PA574&lp\\_g=PA574&dq=driver++liable+if++frightened+horse&source=bl&ots=S2WLLsW-la&sig=ACfU3U3UuZXI\\_](https://books.google.com/books?id=tbpFAQAAMAAJ&pg=PA574&lp_g=PA574&dq=driver++liable+if++frightened+horse&source=bl&ots=S2WLLsW-la&sig=ACfU3U3UuZXI_)
- *The Horse in the City: Living Machines in the 19th Century*. Clay McShane. 2011. John Hopkins University.
- “Of Frightened Horses and Autonomous Vehicles: Tort Law and its Assimilation of Innovations.” Graham, Kyle. 2012. Santa Clara Law Digital Commons. Disponible en <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1170&context=facpubs>
- “Recommendation of the Council on Artificial Intelligence.” OECD Legal Instruments. 2019. OECD. Disponible en <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- “Responsibility and AI”. Bernstein, Abraham; Cancio, Jorge; Floridi, Luciano; Gürses, Seda; Guillemain, Gabrielle; Helberger, Natali; Ilves, Luukas; Kersevan Smokvina, Tanja; McNamee, Joe; Nastos, Evgenios; Perri, Pierluigi; Schulz, Wolfgang; Yeung, Karen. 2019. Council of Europe (COE). Disponible en <https://rm.coe.int/responsability-and-ai-en/168097d9c5>
- “IA Responsable: Manual técnico: Ciclo de vida de la inteligencia artificial” González, Felipe; Ortiz, Teresa; Sánchez Ávalos, Roberto. 2020. Disponible en <https://publications.iadb.org/es/ia-responsable-manual-tecnico-ciclo-de-vida-de-la-inteligencia-artificial?fbclid=IwAR3Oesma2QZEbcR8iSNBPuLbDQKLYOpXdp-15i76pdmCQ91XEdW3zcNCM>



¿Quién le teme a la inteligencia?  
Posibilidades y riesgos  
de la inteligencia artificial  
en el Estado digital

MARCELO CABROL  
y ROBERTO SÁNCHEZ A.

Banco Interamericano de Desarrollo

Copyright © 2021



Desde las acciones reactivas frente a la pobreza características de la primera revolución industrial, hasta los enfoques de promoción de los derechos humanos típicos del siglo XX, la función del Estado frente al ciudadano ha evolucionado de manera permanente.

En las últimas décadas el cambio climático, la violencia, la automatización, y ahora la pandemia del COVID-19, han modificado la vida de millones de familias en la región y en el mundo, y exigen una nueva transformación del Estado, uno con vocación predictiva que se apoye en la digitalización y en los algoritmos de la inteligencia artificial.

Este libro presenta una discusión sobre los temas más relevantes relacionados con el desarrollo y uso de una inteligencia artificial ética y responsable, pasando por el manejo de datos y su protección, los sesgos y su mitigación, y el papel del Estado como usuario y regulador.

