

MICHAEL KEARNS
y AARON ROTH

El algoritmo ético

La ciencia del diseño de algoritmos
socialmente responsables

Traducción de la obra original

«**The ethical algorithm**
The Science of Socially Aware Algorithm Design»

■ LA LEY

 Wolters Kluwer

MICHAEL KEARNS
y AARON ROTH

El algoritmo ético

La ciencia del diseño de algoritmos
socialmente responsables

Traducción de la obra original

«The ethical algorithm

The Science of Socially Aware Algorithm Design»

■ LA LEY

 Wolters Kluwer

© **Michael Kearns y Aaron Roth**

Traducción a cargo de:

GEA Textos, S.L.

© **Wolters Kluwer España, S.A.**

© **Oxford University Press**, 1.ª edición en inglés, 2020

© **Wolters Kluwer España, S.A.**, .ª edición en castellano de la 1ª edición en inglés, 2020

Wolters Kluwer

C/ Collado Mediano, 9

28231 Las Rozas (Madrid)

Tel: 902 250 500 — Fax: 902 250 502

e-mail: clientes@wolterskluwer.com

<http://www.wolterskluwer.es>

Primera edición en castellano: septiembre 2020 (traducción de la obra original *The ethical algorithm*, 1ª edición de 2020)

Depósito Legal: M-23688-2020

ISBN versión impresa: 978-84-18349-44-7

ISBN versión electrónica: 978-84-18349-45-4

Diseño, Preimpresión e Impresión: Wolters Kluwer España, S.A.

Printed in Spain

© **Wolters Kluwer España, S.A.** Todos los derechos reservados. A los efectos del art. 32 del Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba la Ley de Propiedad Intelectual, Wolters Kluwer España, S.A., se opone expresamente a cualquier utilización del contenido de esta publicación sin su expresa autorización, lo cual incluye especialmente cualquier reproducción, modificación, registro, copia, explotación, distribución, comunicación, transmisión, envío, reutilización, publicación, tratamiento o cualquier otra utilización total o parcial en cualquier modo, medio o formato de esta publicación.

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la Ley. Diríjase a **Cedro** (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra. El editor y los autores no aceptarán responsabilidades por las posibles consecuencias ocasionadas a las personas naturales o jurídicas que actúen o dejen de actuar como resultado de alguna información contenida en esta publicación.

INTRODUCCIÓN

Ansiedad algorítmica

Se ha dicho que vivimos en una edad de oro de los datos. Prácticamente para cualquier asunto sobre las personas o sobre la sociedad que despierte nuestra curiosidad existen conjuntos de datos inmensos que pueden explorarse y analizarse para ofrecer respuestas con certidumbre estadística. ¿De qué modo influyen nuestros amigos en lo que vemos en la televisión o en nuestra decisión de voto? Preguntas como esta pueden abordarse con datos tomados de Facebook, que registra la actividad en las redes sociales de miles de millones de personas en el mundo. ¿Son las que menos ejercicio realizan las que más pendientes están de su correo electrónico? Quienes utilicen un reloj Apple Watch o un teléfono Android con la aplicación Google Fit tienen acceso a datos como estos. Por otra parte, si un comercio minorista quiere publicitar sus productos ante clientes potenciales sin saber dónde están ni cómo viven su día a día, docenas de empresas compiten entre sí para venderle tan valiosa información.

Tendencias como estas nos sumen en una situación enigmática. Un acceso a los datos como nunca antes en la historia abre perspectivas extraordinarias y aporta grandes beneficios: permi-

te entender cómo funciona la sociedad y mejorar la salud pública, los servicios municipales y los productos para el consumidor. Ahora bien, como individuos, no actuamos como meros receptores de los frutos de estos análisis de datos: al contrario, *somos* los datos que se utilizan para tomar decisiones *sobre nosotros*, a veces con profundas consecuencias.

En diciembre de 2018, el *New York Times* obtuvo un conjunto de datos comerciales que contenían informaciones de localización recabadas de aplicaciones de móviles cuyo objetivo primario se centraba en cuestiones mundanas, como partes meteorológicas o recomendaciones sobre restaurantes. Tales conjuntos de datos contienen la localización exacta de centenares de millones de individuos, actualizados cientos (o incluso miles) de veces al día. Por lo general, los agentes comerciales que compran estos datos están interesados en información agregada (p. ej., un fondo de inversión podría querer informaciones de rastreo del número de personas que compran en una determinada cadena de venta al por menor para predecir la cifra de negocios trimestral). Sin embargo, los datos quedan registrados con sus teléfonos individuales. Superficialmente son anónimos, sin nombres asociados, pero en esta débil promesa de anonimato se graban todos los movimientos de las personas.

Por ejemplo, a partir de los datos mencionados, el *New York Times* pudo identificar a una profesora de matemáticas de 46 años llamada Lisa Magrin. Era la única persona que se desplazaba desde su domicilio, en el norte de Nueva York, a la escuela de enseñanza media en la que trabajaba, a 22 km de distancia. Una vez desvelada su identidad por estos medios, era posible descubrir mucho más sobre ella. El *Times* rastreó los datos de Lisa hasta

Weight Watchers (una clínica de adelgazamiento), a la consulta de un dermatólogo y a la casa de su exnovio. La mujer se sintió violentada y se explicó ante el *Times*: «Es embarazoso, se trata de esos detalles íntimos que no queremos que los demás sepan». Hace apenas un par de décadas, para un grado semejante de vigilancia intrusiva habría sido preciso contratar a un detective privado o recurrir a una instancia oficial; hoy no es sino un sub-producto de conjuntos de datos de amplia presencia comercial.

Claramente, nos adentramos en un mundo nuevo y salvaje.

A medida que prolifera la captura y el análisis de datos, no solo nos preocupa la privacidad, porque los algoritmos, esas pequeñas secuencias de código máquina que interfieren cada vez más en nuestros comportamientos a través de los teléfonos e Internet, no se limitan a analizar los datos que generamos con cada uno de nuestros desplazamientos. También los emplean activamente para tomar decisiones que influyen en nuestras vidas. Cuando pedimos una tarjeta de crédito, tal vez la solicitud nunca llegue a ser examinada por un ser humano. Es corriente que un algoritmo indague en los datos existentes sobre nosotros (y tal vez también sobre las personas que pulsan nuestros «me gusta») de numerosas fuentes distintas para aprobar o denegar automáticamente la petición. Aunque saber al instante que la solicitud ha sido aprobada tiene sus ventajas, sin que sea necesario esperar 5 o 10 días hábiles, esta dinámica reclama un poco de reflexión. En muchos estados, algoritmos basados en lo que se conoce como aprendizaje automático, o de máquina, se usan también para informar sobre decisiones relativas a avales bancarios, libertad condicional y sentencias penales. Los algoritmos sirven, asimismo, para organizar el despliegue de las fuerzas po-

liciales en las ciudades, así como para tomar decisiones en toda suerte de ámbitos con una repercusión directa y real en la vida de las personas. En torno a esta situación surgen dudas relativas no solo a la privacidad, sino también a la imparcialidad, junto con otros valores sociales básicos, como la seguridad, la transparencia, la responsabilidad e incluso el carácter moral.

De este modo, en la línea actual de seguir generando y utilizando inmensos conjuntos de datos para automatizar decisiones importantes (cuya reversión nos parece tan plausible como volver a una sociedad agraria), hemos de pensar seriamente sobre algunos aspectos críticos. Entre ellos se incluyen los límites en el uso de los datos y algoritmos, y las leyes, reglamentos y organizaciones que habrían de encargarse de fijar y aplicar tales límites. También debemos reflexionar a fondo sobre cómo abordar cuestiones como estas desde un enfoque científico, lo que habría de traducirse en codificar principios éticos directamente en el diseño de unos algoritmos cada vez más íntimamente entrelazados con nuestra vida cotidiana. Este libro trata sobre la ciencia emergente del diseño de algoritmos éticos, precisamente con esta pretensión.

Ordenar con algoritmos

En primer lugar, ¿qué es un algoritmo? En su nivel más fundamental, no es más que una serie de instrucciones especificadas con gran precisión para llevar a cabo una tarea concreta. Los algoritmos más simples, los que enseñamos a nuestros alumnos de informática de primer año, hacen cosas muy básicas, aunque a menudo importantes, como ordenar una lista de números de menor a mayor. Imaginemos que nos topamos con una colección de 1.000 millones de tarjetas, cada una de las cuales tiene

Este libro entrelaza tecnología, sociedad, ética y diseño de algoritmos y en él se analizan no pocos casos en los cuales los algoritmos revelan información personal sensible o discriminan a ciertos individuos por datos demográficos. Los más influyentes algoritmos de la sociedad, desde las búsquedas de Google y Facebook a la clasificación de los créditos y la valoración de los riesgos sanitarios, son desarrollados generalmente por científicos e ingenieros de alta cualificación que se esmeran por aplicar principios de diseño bien contrastados. En realidad, los problemas derivan del propio diseño algorítmico, sobre todo cuando se maneja aprendizaje automático.

En este libro veremos que es posible ampliar los principios sobre los que se basa el aprendizaje automático para solicitar que incluyan, de una forma cuantitativa, medible y verificable, muchos de los valores éticos que nos definen como individuos y como sociedad.

ISBN: 978-84-18349-44-7



9 788418 349447



3652K29059



ER-0280/2005



GA-2005/0110