

¿Algoritmos que tomen decisiones éticas y morales?

Crear máquinas que tomen decisiones basadas no solo en datos sino también en criterios éticos y morales es un reto no solo tecnológico sino filosófico del que se habla a partir de las nuevas realidades. El dilema está en si ese tipo de opciones se le pueden dejar a un aparato programado o si es una función indelegable del ser humano.

SEBASTIÁN AGUIRRE EASTMAN

Colaborador

La idea le surgió a Amazon en 2014: construir un sistema de inteligencia artificial que revisara las hojas de vida de quienes se postulaban a las vacantes de empleo que ofrecía la compañía y tomara decisiones respecto a su contratación. Unos meses después, se habrían de sorprender con algunos resultados del experimento.

El proceso era simple: personal de recursos humanos ingresaba los datos de 100 currículos y el sistema se encargaba de seleccionar los cinco mejores con base en algunos criterios que le fueron programados.

Un año después de estar funcionando, es decir en 2015, la compañía se percató de que el sistema tenía un sesgo contra las mujeres y sugería contratar solo hombres. Aunque el personal técnico intentó enmendar el error, en 2018 el proyecto fue suspendido, según reportó el diario español *El País*, citando un reporte de la agencia de noticias Reuters.

Este es solo un ejemplo de una serie de desafíos éticos y morales que ya empieza a enfrentar la industria 4.0 con los avances tecnológicos que han puesto al servicio de los seres humanos herramientas como la inteligencia artificial, el internet de las cosas, el *Blockchain*, entre otras, que si bien serán muy útiles y bien utilizados potenciarán numerosos campos del conocimiento y de la vida cotidiana, también plantean una discusión sobre sus peligros.



Foto Shutterstock.

“VENDEN LAS VENTAJAS, NO LAS DESVENTAJAS”

Patrici Calvo es autor de publicaciones sobre ética de las cosas y profesor de Filosofía Moral en la Universidad Jaume I, en Castellón de la Plana, una ciudad española ubicada frente al mar Mediterráneo, en la Comunidad Valenciana.

Desde hace algunos años publica artículos y ensayos respecto a los problemas que las personas y las instituciones están enfrentando con la industria 4.0 y cómo esta idea “revolucionaria” se ha ido introduciendo casi que en cualquier ámbito de la actividad humana -política, economía, salud, entre otros-, pero cuyos avances, en su concepto, “nos han emborrachado” y han hecho que “se les abrace sin ningún tipo de crítica”.

Ese, señala, es el punto más importante y el que más problemas ocasiona, pues en su criterio se habla en abundancia de las ventajas de la industria 4.0 -que son muchas- pero de sus desventajas es poco lo que se menciona. Y allí hace hincapié en dos asuntos: la brecha de desigualdad que esta genera y los desafíos desde la gobernanza para implantar marcos jurídicos y éticos a su desarrollo.

Calvo plantea que hay problemas que se vienen presentando desde hace algún tiempo que invitan a pensar no en lo que va a ocurrir, sino en lo que ya está sucediendo.

“Hoy en España, por ejemplo, una empresa de telecomunicaciones va a poner en circulación las primeras líneas 5G [una nueva generación de la red inalámbrica que aumentará la velocidad de transferencia de los datos, entre otros desempeños mejorados]. Quien tenga acceso a ellas, tendrá

mayor capacidad de influir en los algoritmos. Hoy todo se está conectado con el internet de las cosas, tenemos millones de cámaras para controlar y conectar. La ideología de las 4.0 es que los algoritmos tengan el poder de toda actividad humana y eso es muy peligroso”.

Calvo cuestiona que algunos defensores de los beneficios de estos sistemas plantean que, al contrario, estas tecnologías reducirán las brechas de desigualdad. En cambio, el profesor considera que habrá grupos vulnerables de la sociedad que corren con desventajas frente a los algoritmos, al no tener acceso a internet o a equipos de telecomunicaciones, ya sea por su costo elevado o porque en las zonas donde viven no cuentan con servicios cuya velocidad los ponga al mismo nivel de las grandes ciudades, no tendrán la misma capacidad de influir con sus opiniones, hábitos o comportamientos y, por tanto, serán sesgados por las programaciones.

Se refiere, por ejemplo, a habitantes de zonas rurales con conectividad limitada o, incluso, a mujeres que en algunos países -porque se los prohíben sus familias o hasta el mismo Estado- no pueden acceder a un celular o a internet.

“Al final será la clase media alta la que va a influir en las políticas públicas, pues los algoritmos solo tendrán en cuenta sus opiniones, hábitos o comportamientos”, expresa.

Calvo menciona casos recientes en países como Nueva Zelanda, Japón y Estados Unidos en los que hay quienes proponen que se deje en manos de los algoritmos el diseño de políticas públicas, lo cual puede tener consecuencias inesperadas y lo peor, dice, “es que nadie se quiere dar cuenta”.

Una de las propuestas frente a estos dilemas pide aumentar la conciencia entre los líderes de las organizaciones para que dediquen más recursos a identificar las preocupaciones de ética de la inteligencia artificial. Foto Shutterstock.



PATRICI CALVO

Licenciado en Humanidades, máster en Responsabilidad Social Corporativa y en Nuevas Tendencias y Procesos de Innovación en Comunicación. En los últimos años ha compaginado su labor como investigador predoctoral del Ministerio de Ciencia e Innovación (2008-2012), investigador posdoctoral de la Conselleria d'Educació, Formació y Ocupació de la Generalitat Valenciana (2013-2015) y Profesor Ayudante Doctor del Departamento de Filosofía y Sociología de la Universitat Jaume I (2016 y presente).

Patrici Calvo, profesor e investigador de Filosofía Moral en la Universitat Jaume I (España). Foto Ximo Górriz.

¿QUÉ HACER DESDE LA GOBERNANZA?

Esa es la pregunta del millón, según Patrici Calvo. El profesor valora el hecho de que en algunas regiones, como en Europa, se hayan registrado avances para crear un marco jurídico, legal y ético, y que algunos gobiernos nacionales hayan tomado en cuenta opiniones como las que él ha expresado en torno a que lo jurídico y lo ético vayan de la mano.

Entre los avances destaca que se proponga una ética de la inteligencia artificial, que es un camino. Pero, de otro lado, cree que podría haber más desarrollo en otra vía: la llamada "ética algorítmica". Este campo estudia cómo la inteligencia artificial, el aprendizaje de las máquinas y los robots interpretan los datos (este concepto, de acuerdo con la propuesta de Luciano Floridi y Mariarosaria Taddeo en su artículo del año 2016 titulado *What is data ethics?*).

"Es decir -plantea Patrici Calvo- incluir ciertos códigos en los códigos algorítmicos para que estos piensen y tomen decisiones éticas. Un camino lo llevan los técnicos, quienes intentan que las máquinas piensen éticamente; por tanto, cómo se puede hacer que las máquinas aprendan y sean morales. Y, por el otro, crear un marco que oriente no solo a las máquinas sino a los legisladores para que creen leyes y estas, a su vez, orienten a los técnicos para que cuando hagan esos códigos tengan en cuenta ciertas cuestiones éticas".

"Quienes desarrollan tecnologías nos llevan mucha ventaja: cuando pensamos en alguna postura ética, ellos ya trabajan en algo más avanzado".

Patrici Calvo, profesor de Filosofía Moral



Un informe reciente encontró que dos tercios de los trabajadores tecnológicos desea más oportunidades para evaluar los posibles impactos de sus productos, pero que esa es la prioridad más baja en su trabajo. Foto Shutterstock.

UN TEMA POLÉMICO

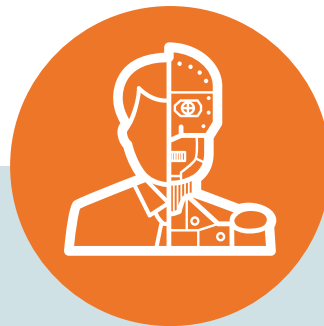
En 2015, algunas organizaciones como Amazon, Apple, BBC, Facebook y Google, entre otras, se unieron y crearon Partnership on AI, con la misión de diseñar las mejores prácticas, investigaciones y diálogos públicos sobre los beneficios de la inteligencia artificial para las personas y la sociedad.

En junio pasado, la investigadora Alice Xiang participó en el evento AI for Good Summit, en San Francisco (Estados Unidos). Ella trabaja en temas como la equidad, la transparencia y la responsabilidad con la inteligencia artificial. Habló sobre la importancia de que la ética de la inteligencia artificial tenga un puente entre la teoría y la práctica.

Dijo: "Como los actos de la inteligencia artificial son un intermedio entre el desarrollo humano y los resultados finales, puede ser difícil de identificar y entender las consecuencias éticas de varios de los pasos en el proceso de desarrollo".

Al ser tan difícil dilucidar qué tan complicadas pueden ser esas consecuencias y con qué frecuencia el proceso de desarrollo implica interrogar decisiones técnicas aparentemente sencillas, Xiang expuso tres pasos para enfrentar este desafío:

- 1 Aumentar la conciencia entre los líderes de las organizaciones para alentarlos a dedicar más recursos a la identificación y abordaje de las preocupaciones de ética de la inteligencia artificial. Un informe reciente encontró que dos tercios de los trabajadores tecnológicos desearían más oportunidades para evaluar los posibles impactos de sus productos, pero que hacerlo es hoy la prioridad más baja en su trabajo.
- 2 Educar a los desarrolladores de inteligencia artificial sobre técnicas para abordar, detectar y mitigar preocupaciones éticas con sus productos. El mismo informe encontró que los trabajadores tecnológicos utilizan en la actualidad principalmente métodos informales, como su brújula moral personal o búsquedas en internet, para evaluar el impacto social de sus productos.
- 3 Educar a los formuladores de políticas y a los equipos legales o de políticas en las organizaciones sobre cómo construir estructuras de gobernanza que incentiven a las organizaciones a abordar los problemas de ética de la inteligencia artificial, y a revisar y refinar sus prácticas.



LA "INGENIERÍA INVERSA" DE LA REGULACIÓN EXISTENTE

La organización Eticas Foundation, en su publicación *Discriminación algorítmica. El reto de los datos en el siglo XXI*, se cuestiona si las decisiones algorítmicas "son una forma de amplificar, extender y hacer inescrutables los prejuicios y la discriminación que ya prevalecen en la sociedad".

En el documento, la entidad sugiere una taxonomía rigurosa de los impactos y las externalidades negativas y positivas de los algoritmos". Esto "permitiría entender los riesgos que conlleva la proliferación del uso de algoritmos en una muy amplia variedad de tareas con importantes consecuencias sociales. Solo a partir de este conocimiento se podrán diseñar las medidas a tomar por parte de los actores relevantes para mitigar sus consecuencias negativas".

En ese sentido, Eticas Foundation propone una "ingeniería inversa" de los marcos regulatorios existentes "para identificar cómo y por qué han surgido, qué nos dicen de las sociedades que las propusieron y qué impacto han tenido. Por otra parte, es necesario analizar las propuestas para traducir los ideales humanos (ética, equidad y justicia) a formas que puedan ser entendidas por los algoritmos junto con conceptos como la rendición de cuentas".

Un estudio encontró que quienes trabajan en campos tecnológicos utilizan básicamente métodos informales, como su brújula moral personal o búsquedas en internet, para evaluar el impacto social de sus productos.

PRINCIPIOS BÁSICOS

La conclusión de Xiang -con la que Calvo no está del todo de acuerdo- es que no se trata de una ética algorítmica sino una ética de la inteligencia artificial, con un marco regulatorio compuesto por cinco principios. Cuatro de ellos son fundamentos básicos desde la bioética:

- **Autonomía:** la capacidad diseñar normas o reglas para uno mismo sin influencia de presiones.
- **No maleficencia:** abstenerse, con intención, de realizar actos que puedan causar daño o perjudicar a otros.
- **Beneficencia:** actuar en beneficio de otros, promoviendo sus legítimos intereses y suprimiendo prejuicios.
- **Justicia:** tratar a cada uno como corresponda, con la finalidad de disminuir las situaciones de desigualdad.

Hay un quinto que no hace parte de la bioética y es quizá el original de este planteamiento, el cual se ha llamado "**explicabilidad**": consiste en que "aquellos que hacen el algoritmo se pueda explicar, y si esto no es posible, el desarrollo no puede avanzar".

El profesor argumenta que no le gusta porque "creo que se podrían buscar y aplicar principios mucho más importantes y claros. Utilizar principios propios de la salud tal cual sobre la inteligencia artificial, en algo que es tan diferente, me resulta chocante, pero eso es lo que están planteando".

Por último, Calvo lamenta que en la actualidad no existan leyes internacionales que regulen este tema y sugiere instaurar una suerte de comité mundial que recoja los avances de algunos países y que presente principios que orienten el desarrollo de los algoritmos para que luego cada nación los pueda aterrizar a su realidad y cree leyes nacionales. El tema es complejo y la polémica está servida. ■