

La importancia de la ética en la inteligencia artificial

No se trata de ciencia ficción. Desde hace décadas, son muchos los científicos que vaticinan una explosión de inteligencia artificial en algún momento del siglo XXI. Esta inteligencia artificial podría llegar a ser algo único, algo enormemente poderoso que podría imponerse a la inteligencia humana. ¿Es el momento de establecer criterios éticos sólidos?

En 1985, Judith Jarvis Johnson puso de manifiesto en la revista *The Yale Law Journal* (la publicación de la Facultad de Derecho de Yale) algunas preguntas que llevaban rondando desde principios del siglo XX en el conocimiento humano. Este reconocido jurista expuso la complejidad que supondría para una inteligencia artificial encontrar la solución de un dilema humano complejo. Imaginemos que una persona conduce un coche sin frenos; ante él se bifurca el camino en dos: en el camino de la izquierda hay cinco personas, en el de la derecha sólo hay una. ¿Es lícito matar a una persona para salvar a otras cinco? Vamos con la segunda parte del dilema: un joven acude a una clínica para un chequeo rutinario; en esa clínica hay cinco pacientes esperando trasplantes de órganos. Para poder vivir, dos de ellos necesitan un pulmón, otros dos necesitan sendos riñones y el quinto requiere un corazón. Curiosamente el joven que ha ido a hacerse el chequeo tiene el mismo grupo sanguíneo que ellos, lo que le convierte en el donante idóneo. Repetimos la pregunta: ¿Es lícito matar a una persona para salvar a otras cinco?

Un ser humano tendría que despejar las incógnitas de la ecuación atendiendo criterios éticos. En cualquiera de los casos, cuantitativamente hablamos de matar a una persona para salvar a otras cinco. No obstante, casi todo el mundo estaría de acuerdo en que es preferible atropellar a esa persona para salvar a las otras cinco. Hay muchos factores puramente humanos que podrían justificar esta decisión, como la ausencia de contacto físico (no es lo mismo matar a un desconocido que a alguien que tratamos y conocemos personalmente). Sin embargo, es difícil que una máquina pueda calibrar este tipo de decisiones, por eso es tan importante que la inteligencia artificial pueda estar dotada de una ética, o código de valores, que condicione, a nuestra imagen y semejanza, la esencia de sus actos.

Siguiendo este razonamiento, hay teóricos que se postulan en creencias catastrofistas. Es el caso del filósofo Nick Bostrom, que sostiene que una inteligencia artificial avanzada podría tener la capacidad de provocar la extinción humana, ya que sus planes pueden no implicar tendencias motivacionales humanas. No obstante, Bostrom también plantea la situación contraria en la que una super inteligencia podría ayudarnos a resolver problemas tediosos y constantes en la humanidad, tales como la pobreza, la enfermedad o la destrucción del planeta.

Una ética empresarial

La infinita complejidad de un sistema de valores humanos hace que la inteligencia artificial no encuentre motivaciones amigables en formas de proceder humanas. No obstante, es la ética la que sostiene la mayoría de organizaciones. El común entender y la aceptación de nuestros esquemas culturales, implica el funcionamiento de complejos recovecos y motivaciones de la psique humana. Al incorporar la Inteligencia Artificial a los procesos de las organizaciones, es importante dotar a esta nueva tecnología de valores y principios. Y, dentro de las organizaciones, son los desarrolladores de esta tecnología las

personas que realmente han de trabajar siendo conscientes de las implicaciones morales y éticas que conlleva su trabajo.

La problemática planteada se está materializando en la Asociación sobre Inteligencia Artificial, creada por Elon Musk y Sam Altman, en la que se invita a los principales líderes tecnológicos para poder identificar dilemas éticos y prejuicios. Su objetivo primordial es establecer unas reglas de juego, basadas en un marco de comportamiento moral, donde la Inteligencia Artificial pueda desarrollarse en representación de la humanidad.

Según un informe realizado por SAS en el año 2017, el 92% de las empresas considera prioritario capacitar a sus tecnólogos en ética, y el 63% cuenta con comités en este tipo de asuntos para revisar el buen uso de la inteligencia artificial. Se trata, pues, de un tema necesario y de solución posible. Por ello, algunas compañías están comenzando a formar a las personas que a su vez están formando a las máquinas. Es, al fin y al cabo, una inclusión de la ética en los algoritmos que rigen la inteligencia artificial. Por ejemplo, la concesión de un préstamo no debe realizarse atendiendo criterios de sexo, edad o religión. Es importante que el machine learning se nutra de principios universales de respeto, libertad e igualdad. La cultura es uno de los vehículos de nuestra supervivencia: es básico que continuemos formando e informando atendiendo a unos principios básicos.

Roboética: principios para el siglo XXI

Hoy en día, la humanidad está inmersa en una revolución tecnológica sin precedentes. La preocupación ética por la creación de nuevos tipos de inteligencia requiere de un exquisito criterio moral de las personas que diseñan estas nuevas formas de tecnología. El algoritmo ha de ser capaz de discernir y reconocer fallos cuando se centran en acciones sociales con repercusión que antes realizaba un ser humano. El concepto es claro: el código no puede dañar a personas o a empresas.

Fruto de esta preocupación, el Parlamento Europeo realizó un informe sobre robótica en 2017 llamado Código Ético de Conducta y, recientemente (diciembre de 2018), ha publicado el primer borrador de la Guía Ética para el uso responsable de la Inteligencia Artificial. 52 expertos han escudriñado y exprimido los rincones de la problemática, centrándose en el ser humano siempre bajo la luz de la defensa de los derechos fundamentales.

Son estándares morales dirigidos a humanos, a los creadores de tecnología. Los principios son los siguientes:

Se debe asegurar que la IA está centrada en el ser humano.

Se debe prestar atención a los grupos vulnerables, como los menores de edad o las personas con discapacidades.

Debe respetar los derechos fundamentales y la regulación aplicable.

Debe ser técnicamente robusta y fiable.

Debe funcionar con transparencia.

No debe restringir la libertad humana.

Las grandes organizaciones, empresas y gobiernos están centrándose en los problemas que pueden surgir en el tema de la ética de la inteligencia artificial para trazar consideraciones, prácticas y marcos comunes de cara al futuro. Es importante alcanzar un acuerdo donde se pueda conceptualizar y, sobre todo, regular las prácticas derivadas. Al fin y al cabo, la tecnología es un paso más de nuestra evolución... Y su código de ceros y unos debe ser un reflejo de nuestros genes.