

VENTURA

100 ARTÍCULOS SELECCIONADOS



Hacia la singularidad: Inteligencia Artificial dominando el mundo

Juan Pérez Ventura

FECHA DE PUBLICACIÓN ORIGINAL: 24 de marzo de 2018

FECHA DE LA PRESENTE EDICIÓN: 12 de noviembre de 2024

compra el libro completo en vaventura.com/libro

Hacia la singularidad: Inteligencia Artificial dominando el mundo

Tan impresionantes son los evidentes peligros a los que nos enfrentamos como la poca atención o preocupación que parecen concertar entre la mayor parte de la población. Quizás no preocuparse sea una buena estrategia para llevar una vida tranquila, pero sin duda no es recomendable a la hora de pensar políticas para el futuro más cercano. La sociedad en convivencia con la inteligencia artificial no está tan lejos. Y sí, quizás haya que preocuparse un poco, aunque sólo sea para diseñar cómo vamos a establecer la relación con los robots inteligentes, que ya nos acompañan en muchos momentos de nuestro día a día.

El concepto de singularidad, aplicado en el ámbito de la tecnología, hace referencia a ese posible momento futuro en el que las máquinas alcancen una inteligencia igual o superior a la del ser humano, y sean autónomas en tareas como el aprendizaje o el mejoramiento. Es decir, máquinas (ordenadores, robots, programas informáticos...) capaces de aprender por sí solas y de mejorarse a sí mismas. Además, podrían igualar al cerebro humano si adquirieran habilidades emotivas y racionales.

Por el momento el ser humano ya ha creado máquinas que dominan la inteligencia perceptual (reconocimiento de imágenes, de sonido, transcripciones del habla, clasificación de imágenes...) y que incluso pueden crear algoritmos para generar sonidos e imágenes manejando redes adversarias generativas, es decir, que permiten generar nueva información a partir de la que ya se tiene. Si bien las máquinas actuales realizan todo esto gracias a una base informativa facilitada por un ser humano (un programador, un ingeniero, un informático...), los expertos coinciden en que hacia las décadas de 2040 o 2050 se habrá conseguido que las inteligencias artificiales construyan desde cero el conocimiento, planificando jerárquicamente la información y entendiéndola.

Es importante recordar que no estamos hablando de una sociedad en la que los coches vuelen y la gente viaje a Marte de vacaciones, sino de un futuro cercano en el tiempo (¿qué son cuarenta años en la escala humana? ¿cuánto se ha avanzado tecnológicamente desde 1970?) que plantea situaciones muy interesantes para abrir el debate, aunque puedan parecernos ficticias o inverosímiles.

Posibles escenarios en el futuro próximo

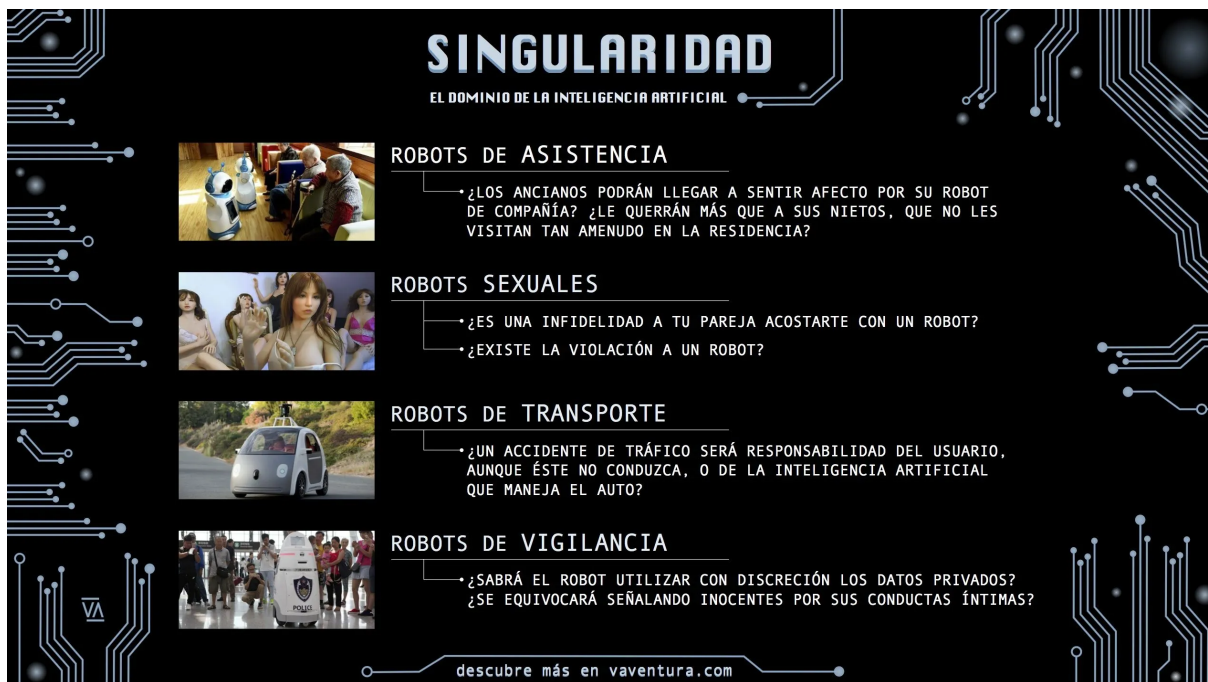
Antes de hablar de una superinteligencia que quiera aniquilar al ser humano porque es muy mala (parece que

esa es la aproximación más extendida que se plantea para abordar el tema de la inteligencia artificial), es necesario situarnos temporalmente y tratar de adivinar qué será lo primero que llegue a nuestras sociedades. Antes de tener robots terminators que se adueñen del mundo o programas informáticos que nos esclavicen, habrá situaciones mucho menos excitantes pero que nos pondrán en complicados debates morales, políticos, sociales, jurídicos, filosóficos, éticos y hasta familiares.

Hasta ahora la relación con las máquinas ha sido fría como el metal del que están hechas. Las diseñamos y observamos cómo realizan la tarea que les hemos encomendado. Están a nuestro servicio. Tienen formas extrañas, con cables, pantallas y luces cubiertas por duras carcasas de plástico. Viven en el interior de fábricas de las que salen nuestros coches y electrodomésticos, las tenemos escondidas para olvidar que hacen muy bien su trabajo. Esta relación distante y de superioridad puede cambiar próximamente, cuando veamos a las máquinas de manera distinta.

Un ejemplo está en los llamados robots de compañía, que podrían entrar en contacto con nosotros muy pronto en forma de ayuda a ancianos, como objetos sexuales o como simples elementos de ocio. El mercado de los robots sexuales está en crecimiento, y ya hay empresas que venden este tipo de acompañantes por 15.000 euros. Los robots (modelos masculinos y femeninos) tienen proporciones físicas perfectas e incluso varias personalidades programadas. Existe la muñeca robótica Roxxy Gold, con la que se puede practicar sexo poniéndola en modo «salvaje» o «frígida». En cuanto a los robots de asistencia para personas mayores, tanto en China como en Japón ya se están poniendo en marcha los primeros proyectos en residencias y hospitales. Es mucho más eficiente y barato tener a un robot haciendo ejercicios delante de un grupo de ancianos que malgastar el tiempo de un humano que podría estar haciendo otro trabajo mucho más importante.

Existe una discoteca donde la música la pone y mezcla un robot. Aunque su aspecto sigue siendo el de un brazo metálico, bastaría con darle forma humana para que se pudiera empatizar con este DJ robótico. Eso se intenta con Android U, una presentadora de televisión japonesa que puede hacer su programa en directo sin necesidad de descansar: es un androide. También hay robots policías patrullando aeropuertos: ¿cómo reaccionaríamos si un montón de cables y circuitos nos quisiera detener? ¿obedeceríamos? Como hemos dicho, hasta ahora la relación entre humanos y máquinas se ha basado en el po-



der del creador, pero si otorgamos cualquier tipo de autoridad a un robot, pueden generarse situaciones inéditas.

Hablamos de máquinas y robots, pero la inteligencia artificial no conoce forma física concreta. Al moverse en la dimensión virtual, la IA puede ser un programa informático sin apariencia humanoide ni robótica. Es en estos casos de hecho donde vemos el mayor poder de la inteligencia artificial, como ocurre con SkyNet, un programa informático capaz de reconocer objetos y a personas que ha sido instalado en China para garantizar la seguridad (y vigilancia) de todos los ciudadanos. SkyNet tiene más ojos que nadie: 20 millones de cámaras le facilitan imágenes y él se encarga de reconocer facialmente a cada ciudadano.

El debate que surge con la puesta en práctica de este tipo de tecnologías tiene que ver con cuestiones de privacidad, intimidad y derechos de las personas. El aumento de la seguridad casi siempre va atada al descenso de los derechos, pero en el caso de la IA se suma el factor de que es un robot quien invade la intimidad y utiliza con autonomía nuestros datos. ¿Estamos indefensos ante la vigilancia automatizada? Ante esta situación habrá quien prefiera confiar su seguridad a la inteligencia humana, y otros confiarán más en una inteligencia artificial. Los seres humanos cometemos muchos más errores, sin embar-

go cuando falla un robot, la alarma es mayor.

Este mismo año 2018 se ha producido el primer atropello mortal causado por un coche inteligente. El determinismo algorítmico y las decisiones automatizadas no siempre aciertan. Las mismas reglas se aplican a los sistemas financieros automatizados. ¿Qué ocurriría si una IA cometiera un accidente manejando flujos de dinero y nos hiciera perder miles de millones? Noticias de este estilo fomentan interesantes debates políticos, sociales, éticos y hasta jurídicos: ¿quién tiene la culpa cuando una inteligencia artificial causa un accidente o comete un delito, la máquina o la persona que la diseñó? Nadie duda que en la actualidad la responsabilidad es de la empresa y los ingenieros detrás de la máquina, pero en unos años este escenario puede cambiar.

¿Dejarías a un algoritmo decidir? Seguramente el determinismo algorítmico no nos importe para el caso de Spotify, que con su inteligencia nos recomienda las canciones que sabe nos van a gustar pero, ¿y en los coches? ¿dejaríamos a la IA que tomara las decisiones que creyera oportunas en la carretera mientras nos lleva a donde le hemos dicho? ¿y en casa? ¿nos parecería bien ceder la cuchara y que fuera una IA quien nos cocinara la comida que sabe que preferimos? Ahí se podría decir que el robot tiene la sartén por el mango... ¿Queríamos que una inteligencia

artificial nos presentara cada mañana un resumen de las noticias de los medios de comunicación que ha identificado como más importantes? ¿Nunca leeríamos nada diferente, siempre las mismas fuentes de información!

Todas estas cuestiones que se plantean parecen otorgar intencionalidad a las acciones de los robots inteligentes. ¿Realmente serán capaces de tener intenciones? Y, si no hay intencionalidad, ¿se les puede exigir responsabilidad? ¿en qué medida las máquinas podrán ser responsables de sus actos? Aceptar que los robots deben tener una responsabilidad es adjudicarles deberes, y en ese sentido es interesante recordar la propuesta que la Alianza Progresista de Socialistas y Demócratas hizo en el Parlamento Europeo en 2016, en la que se pedía que los robots adquirieran el estatus jurídico de 'personas electrónicas' para que, entre otras cosas, tuvieran el deber de pagar impuestos. Según como entendemos las relaciones en sociedad, tener deberes va a acompañado de disfrutar de ciertos derechos. Si se avanza hacia una verdadera inteligencia artificial, los robots podrían adquirir derechos derivados de su trabajo. Ya hay inteligencias artificiales componiendo música o pintando cuadros. ¿A quién pertenecen las obras creadas por robots? ¿a los robots?

En 2017 tuvo impacto la noticia de que un ingeniero chino se había casado con una mujer robot diseñada por él mismo. Más allá de la ridiculización que podamos hacer en la actualidad de este tipo de matrimonios, en un futuro no muy lejano pueden aceptarse socialmente.

La relación con la inteligencia artificial es distante para la mayor parte de la población, pero ya hay algunos que se están preparando para la singularidad tecnológica creando partidos políticos y movimientos que piensan estrechar lazos con los futuros robots inteligentes. El transhumanismo es una corriente que defiende la transformación de la vida humana, evolucionando hacia una especie de ciborgs que combinen lo mejor de nuestra biología con el poder de la tecnología más avanzada. En la actualidad ya hay personas que son «ciborgs», ya que dependen de la tecnología para cumplir alguna función biológica, pero lo que propone el transhumanismo va mucho más allá, proponiendo transformar la condición humana y mejorar las capacidades humanas. Ya hay un candidato a la presidencia de Estados Unidos que es transhumanista, y en España también existe un partido transhumanista llamado Alianza Futurista. De nuevo, bajo nuestra óptica todo esto parece una locura, pero no deberíamos infravalorar nuestra propia capacidad para evolucionar tecnológicamente. De hecho, parece inevitable que lo hagamos.

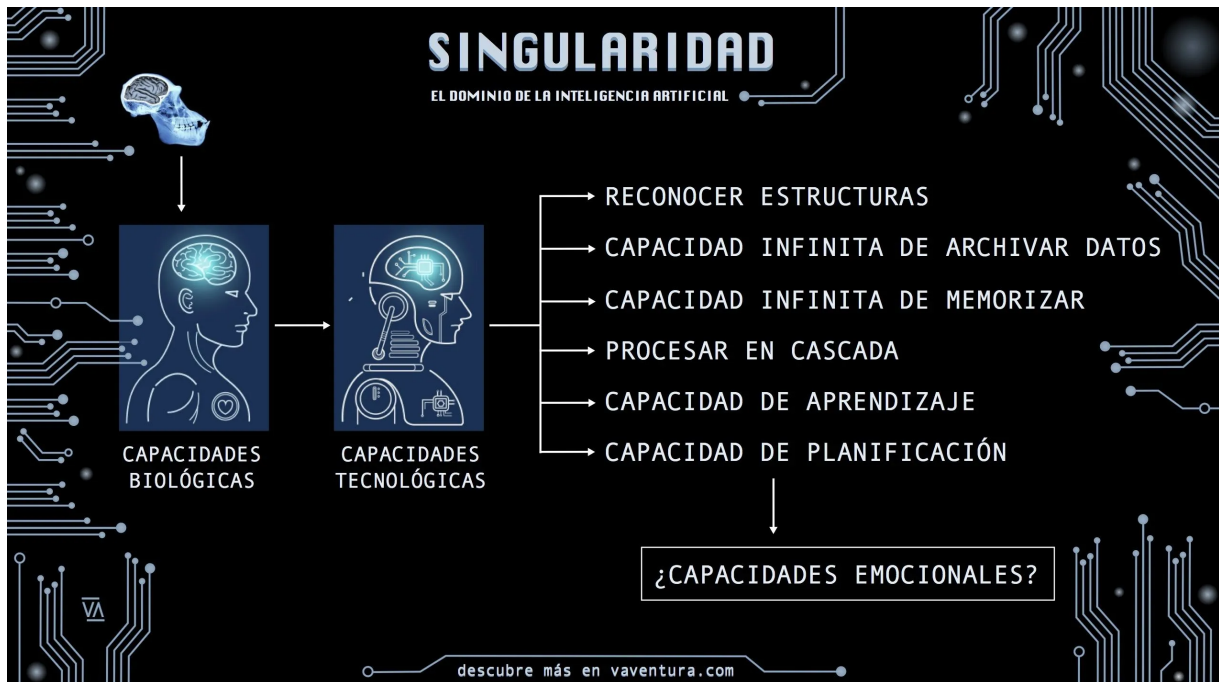
Hacia la inevitable singularidad tecnológica

Como dice el neurocientífico y filósofo Sam Harris, teniendo en cuenta lo valioso que son el conocimiento y la tecnología, el ser humano nunca dejará de buscar mejoras en esos campos. El ser humano nunca dejará de progresar tecnológicamente. Nunca ningún gobierno, empresa o institución internacional recomendaría dejar de aprender, dejar de extender las fronteras del conocimiento o dejar de innovar. No concebimos la posibilidad de dejar de hacer mejoras en nuestra tecnología de manera permanente. De esta manera, parece evidente que el ser humano va a seguir progresando tecnológicamente, haciendo sus máquinas cada vez más inteligentes. Lo que nos lleva, inevitablemente, a que tarde o temprano acabaremos desarrollando ordenadores, programas, máquinas y tecnologías más inteligentes que nosotros.

En realidad no importa el momento en el que ocurra o la rapidez del proceso, lo importante realmente es el hecho de que ocurrirá. Llegará tarde o temprano, pero llegará. Ese momento en el que se de la singularidad, y en el que el ser humano no sea el ser más inteligente del planeta. En un escenario en el que máquinas inteligentes sean capaces de crear otras máquinas inteligentes, el ser humano podría sufrir una especie de marginación intelectual, debido a que sería intelectualmente inferior a los robots.

Algunas de las razones que cita Sam Harris para apoyar la idea de que el ser humano alcanzará inevitablemente la creación de inteligencias artificiales superiores son estas:

- El ser humano no se encuentra actualmente en su pico de conocimiento ni de desarrollo. No estamos todavía explotando nuestras capacidades intelectuales ni tecnológicas en todo su potencial. Por lo tanto el ser humano seguirá mejorando estas capacidades, enunciando mejores teorías que expliquen el Universo, conociendo mejor la Naturaleza... y construyendo mejores tecnologías.
- La industria de la inteligencia artificial reporta enormes beneficios económicos. Un factor clave para que las empresas sigan promoviendo avances en este campo, creando un escenario de competición continua que sin duda acelerará los seguros progresos tecnológicos. El primero que consiga crear una superinteligencia artificial habrá ganado la partida. Una peligrosa carrera hacia la singularidad.



La cadena trófica a nivel tecnológico

¿Por qué somos mejores que los monos? ¿por qué un chimpancé es superior a un escarabajo? ¿por qué un delfín está por encima de una sardina? La respuesta parece estar en la superioridad física, pero es la inteligencia el factor que define la posición en la pirámide de poder animal. Lo que hace poderosos a los seres humanos sobre otras especies no son sus músculos o sus capacidades físicas, sino su inteligencia, su capacidad intelectual.

Que nuestro cerebro sea más inteligente que el cerebro de un delfín o de un chimpancé nos ha permitido acumular cultura y experiencia de generación en generación, y así inventar tecnologías y una organización social complejas. Pero estos atributos que ahora nos ponen en una posición espacial podrían pertenecer luego a las máquinas inteligentes del futuro, que podrían superarnos drásticamente en esa inteligencia que permite la innovación tecnológica, la organización compleja, la planificación a largo plazo, etc.

Las capacidades de aprendizaje y planificación son dos habilidades que otorgan inteligencia. Las especies que han conseguido aprender de la experiencia y planificar pensando en el futuro demuestran tener una inteligencia

superior. Por el momento las máquinas no son capaces de hacerlo partiendo de cero, es decir, siguen necesitando a un programador humano que les otorgue una serie de datos con los que empezar a trabajar. Cuando consigamos crear algoritmos que sean capaces de aprender como lo hace un bebé humano, entonces estaremos muy cerca de crear inteligencias artificiales avanzadas.

En 2011 el famoso concurso Jeopardy tuvo un ganador inusual. Un robot llamado Watson y cargado con una inteligencia artificial diseñada por IBM venció a los demás concursantes demostrando una cultura casi infinita y, lo que es más sorprendente, la capacidad de reconocer el lenguaje humano. Sin duda este hecho fue un paso importante para conseguir inteligencias artificiales complejas.

El momento de singularidad se apoya en argumentos como el potencial de una máquina superinteligente, que excede por mucho a las capacidades humanas. Atendiendo meramente a descripciones y cualidades físicas, comprobamos que las máquinas tienen las de ganar. Básicamente porque el umbral en el procesamiento de información en una infraestructura artificial se encuentra mucho más allá de los límites del tejido biológico. Para empezar, una neurona biológica manda impulsos a unos 200 hertz, 200 veces por segundo, mientras que incluso un transistor actual opera a la frecuencia de los gigaher-

cios. Las neuronas propagan el impulso lentamente a lo largo de los axones, a un máximo de 100 metros por segundo, pero en los ordenadores, las señales pueden viajar a la velocidad de la luz. Además, hay diferencias físicas relativas al puro tamaño: un cerebro humano tiene que caber dentro de un cráneo, pero un ordenador puede ser del tamaño de un almacén o un edificio.

Teniendo todo esto en cuenta y dando por hecho que la singularidad llegará en algún punto de la historia humana, podemos tratar de apuntar algunas ideas de cómo podría ser ese escenario futuro. Como muchos esperábais, es hora de hablar de Terminator, Matrix y el Apocalipsis. Y no son imaginaciones de agoreros que hablan en canales de historia frikis, sino apuestas de mentes brillantes y estudiosas como la de Stephen Hawking, que asegura que «el desarrollo total de la inteligencia artificial podría propiciar el fin de la raza humana», o Elon Musk, que advierte que «liberar energía es fácil: lo difícil es contenerla».

Debemos preguntarnos: ¿Sabremos contener el poder de la inteligencia artificial? Se ha demostrado que, a medida que la tecnología mejora, crece su potencial destructivo, y la clave está en el uso que el ser humano hace de ella. Uno de los escenarios que se quiere evitar es que la tecnología con capacidad destructiva caiga en manos de grupos terroristas: no queremos que tengan el poder de destruir nuestra red informática, de acceder a nuestros datos o de diseñar armas autónomas. Pero en un estadio más avanzado, cuando hayamos sido capaces de desarrollar esta tecnología sin que ningún humano con malas intenciones se aproveche de ella, ¿cómo asegurar que la propia tecnología no se volverá contra nosotros?

Todo es cuestión de perspectiva. No debemos ver a los robots superinteligentes como seres malvados y con ansias de dominar el mundo, sino más bien ponernos en su lugar. ¿Cómo actuamos nosotros ante una hormiga? ¿cómo tratamos a los animales que nos acompañan en este planeta? Son inferiores a nosotros, y los vemos bajo esa óptica. Les podemos tratar bien o mal, pero siempre desde una posición de superioridad y atendiendo a nuestras preferencias. Lo mismo podría ocurrir con las IA, que llegaran a vernos como nosotros vemos a un chimpancé. Sin malas intenciones (todo lo contrario, buscando preservar el mundo y la raza humana), los robots superinteligentes podrían considerar que lo mejor es limitar la población humana, exterminando a unos cuantos millones de personas. Según los cálculos perfectos de los robots, esa decisión sería la mejor entre todas las posibilidades, y beneficiaría tanto a humanos como a máquinas. Esto desde nuestra perspectiva sonaría horrible, pero

desde el punto de vista de la IA podría ser la mejor estrategia para el bienestar general.

El ejemplo utilizado lleva al extremo lo que podría ocurrir, y la clave es entender que podemos llegar a un futuro en el que las situaciones y acciones se vean determinadas por las preferencias de las superinteligencias. Al ser los seres más inteligentes y poderosos, hay pocas razones que inviten a pensar que no harían uso de esa posición de superioridad para imponer sus ideas y políticas.

De nuevo, hay que hacer un ejercicio de empatía con los robots superinteligentes. Pensemos cómo vemos a un escarabajo que se cruza en nuestro camino, qué pensamos de los estúpidos monos que se comen las pulgas... ¿a caso no nos creemos seres infinitamente superiores? ¿a caso no es así como nos ven los demás animales del planeta? Pues hora imaginad un mundo en el que el ser humano no es el animal más inteligente del planeta.

Las capacidades de las IA serán algunas como el aprendizaje más rápido, el trabajo infatigable, habilidades sin error, replicación más numerosa, inmutabilidad emocional... una serie de características que les harán efectivamente superiores a los humanos. Entonces, ¿qué futuro nos espera? ¿Esclavitud de las personas? ¿destrucción del ser humano? ¿dominio total de las máquinas? Un mundo como el de Matrix, donde hombres y máquinas luchan y destruyen el planeta, puede sonar demasiado apocalíptico, pero existe un precedente muy utilizado por los catastrofistas: en 2007 un cañón robotizado mató por error a nueve soldados en Sudáfrica. Evidentemente no lo hizo a propósito, pero la alarma se extiende entre la población: ¿qué harán las armas automatizadas cuando sean inteligentes? ¿cómo seleccionarán sus objetivos? ¿atacarán a población civil para acabar con un terrorista?

En Enero de 2015 una serie de personalidades firmaron un manifiesto advirtiendo de los peligros de la IA, y Hawking ya previó que «los humanos, limitados por una evolución biológica lenta, no podrán competir y serán superados».

El mayor reto: robots inteligentes... y emocionales

Que conseguiremos diseñar robots inteligentes parece seguro, el reto es crear sistemas artificiales alineados con los valores humanos, y que siempre hagan lo que nosotros queremos que hagan. La pregunta es cómo codificaremos las máquinas para que nunca se vuelvan contra la humanidad.

SINGULARIDAD
EL DOMINIO DE LA INTELIGENCIA ARTIFICIAL

LA EVOLUCIÓN BIOLÓGICA LIMITA AL SER HUMANO FÍSICAMENTE, MIENTRAS QUE EL PROGRESO TECNOLÓGICO NO CONOCE ESE TIPO DE BARRERAS DE CRECIMIENTO.

EL CEREBRO HUMANO TIENE UNAS DIMENSIONES FÍSICAS LIMITADAS, Y UNA MÁQUINA PUEDE SER DE CUALQUIER TAMAÑO

ADemás, LAS MÁQUINAS TIENEN OTRAS VENTAJAS, COMO LA VELOCIDAD DE ENVÍO DE DATOS, QUE PUEDE ACERCARSE A LA VELOCIDAD DE LA LUZ. EN EL CEREBRO HUMANO LOS IMPULSOS NEUROLÓGICOS VIAJAN A UNOS 200 HERTZ.

1200 cm³

23 m²
Watson

descubre más en vaventura.com

En 1942 Isaac Asimov propuso una serie de leyes que debían incorporarse a los algoritmos que viajan por el sistema de circuitos y microchips que componen los cerebros positrónicos de los robots. Estas leyes establecen lo siguiente:

1. Un robot no hará daño a un ser humano, ni permitirá con su inacción que sufra daño.
2. Un robot debe cumplir las órdenes dadas por los seres humanos, a excepción de aquellas que entren 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 con la segunda ley.

Los programadores tendrán que descubrir cómo introducir los valores humanos en una máquina para que no solo siga al pie de la letra nuestras peticiones y órdenes, sino que además entienda la intención tras ellas. No será sencillo escribir esos algoritmos, pero Ian Winfield ya ha demostrado que un robot puede actuar éticamente sin entender nada de ética o moral: puso a una máquina en una situación extrema de tener que salvar a dos personas en peligro de muerte. En vez de escoger una de ellas, el robot automáticamente se decidió por rescatar a las dos.

En el caso de poder crear una IA perfecta, capaz de extrapolar las intenciones que tenemos los humanos, entonces no tendremos de qué preocuparnos. Además, si la inteligencia artificial termina siendo capaz de hacer todo o buena parte de nuestro trabajo intelectual mejor que nosotros, entonces Nick Bostrom asegura que tendremos en nuestras manos el último invento que tendrá que realizar la humanidad.

Si conseguimos crear una IA súper-inteligente, con buenas intenciones, con ética y moral, que no nos haga daño, que nos obedezca, que no destruya a la Humanidad, que funcione con energía solar y que sea servicial... entonces el futuro que se abre ante nosotros es el de Wall-E. En ese momento de singularidad el ser humano ya no tendrá que seguir pensando, innovando, creando, estudiando... el trabajo intelectual lo harán las máquinas. Contando con que del trabajo físico ya se habrán encargado previamente, ya no habrá tareas para las personas.

Estaremos de vacaciones, leyendo, viendo la televisión, haciendo deporte y disfrutando de la vida. Suena muy bien, pero esa sociedad futura se enfrentará a problemas como el desempleo o la obesidad y a debates existenciales profundos. Podremos estar todo el día jugando a videojuegos, que no se distinguirán mucho de la vida real, y

quizás recemos a un Dios digital, a un programa informático, como ya hacen los seguidores de la nueva religión Way to the Future. Que este futuro sea ciencia o ficción depende de nosotros mismos. ¿Hacia dónde vamos como sociedad si seguimos progresando?



IMÁGENES

- Título: *Singularidad: el dominio de la Inteligencia Artificial*
- Autor: Juan Pérez Ventura
- Año: 2018

[▲ descargar](#)