TEMAS A CONVERSAR SOBRE LA IA

Efectos y paradojas de la IA.

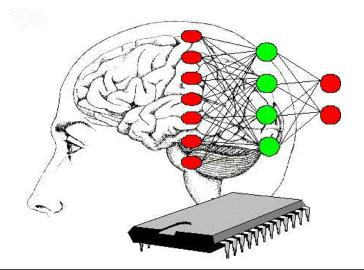
Desde el Test de Turing a la Singularidad

Tecnológica de Kurzweil.

Debate: Ética de la IA.

¿Existe el riesgo la Inteligencia Artificial?

Hacia el Futuro.
Para leer.



Inteligencia Artificial

Dos planteamientos en esta ciencia

"I.A. ... está relacionada con conductas inteligentes en artefactos" (Nilsson, 1998)



Stanford

"El nuevo y excitante esfuerzo de hacer que los computadores piensen...máquinas con mentes, en el sentido más literal" (Haugeland, 1985)



Chicago



IA débil vs IA fuerte ¿Sistemas que actúan racionalmente vs

Sistemas que piensan como humanos?

Inteligencia Artificial Test de Turing



La famosa pregunta planteada por

Alan Turing

¿puede pensar una máquina?

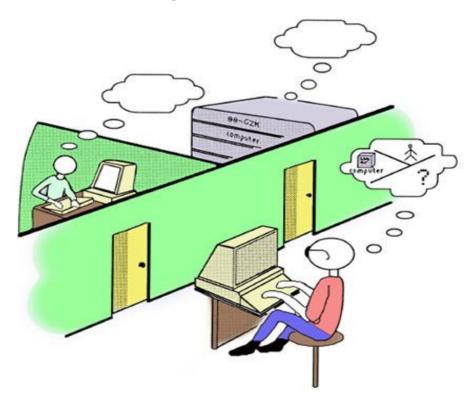


En 1947 Alan Turing pronunció una conferencia ante un auditorio del National Physical Laboratory de Londres en la que debatía sobre la controvertida pregunta ¿Puede pensar una máquina?

By A. M. Turing

Inteligencia Artificial Test de Turing

Test de Turing



Alan Turing decía que "existirá Inteligencia Artificial cuando no seamos capaces de distinguir entre un ser humano y un programa de computadora en una conversación a ciegas".



La máquina polígrafo Voight-Kampff Blade Runner Ridley Scott, 1982

Inteligencia Artificial Test de Turing





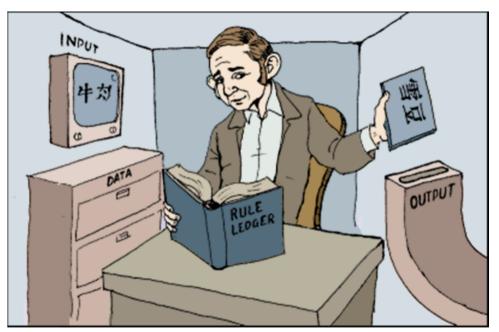
En el año 2014, el conversador inteligente Eugene Goostman, que participaba en un concurso celebrado en la Royal Society el 7 de Junio para conmemorar el 60 aniversario de la muerte de Turing, ganó el concurso.

Participaron 30 jueces, y un tercio de ellos acabaron convencidos de que Eugene Goostman era realmente un adolescente de 13 años.

Inteligencia Artificial

La habitación china de Searle: la comprensión semántica de los contenidos procesados

- En una habitación cerrada, con un orificio de entrada y uno de salida, se coloca a un sujeto con un diccionario de chino. Cada vez que el sujeto recibe un documento en chino por la entrada, lo traduce y devuelve el documento resultante por la salida.
- Para el que no conozca el sistema, este en su conjunto "sabe chino", pero... ¿realmente el sujeto sabe chino?



Los chinos que estén fuera de la habitación pensarán que el de la habitación conoce el chino pues han recibido respuestas satisfactorias, por tanto superaría el Test de Turing.

Este argumento se utiliza para sostener que lo mismo ocurriría con una computadora que pueda responder a textos escritos en chino, manipulando los códigos sintácticos pero no tendrían nada que ver con la comprensión semántica de los contenidos procesados

Inteligencia Artificial

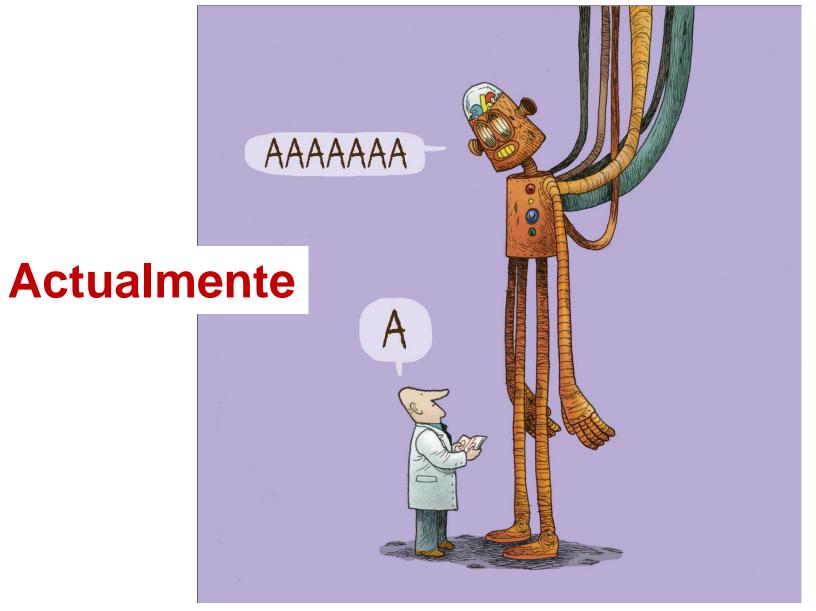
La paradoja de Moravec. Razonamiento humano y habilidades sensoriales y motoras

La famosa paradoja de Moravec nos refleja este momento de la inteligencia artificial. Nos indica que de forma anti-intuitiva, el razonamiento humano (el pensamiento inteligente y racional) requiere de poca computación, mientras que las habilidades sensoriales y motoras, no conscientes y compartidas con otros muchos animales, requieren de grandes esfuerzos computacionales.

Moravec escribía lo siguiente [1]: "es fácil comparativamente conseguir que las computadoras muestren capacidades similares a las de un humano adulto en un test de inteligencia, y es difícil o imposible lograr que posean las habilidades perceptivas y motrices de un bebé de un año".

[1] H. Moravec (1988). Mind Children, Harvard University Press.

Paradoja de Moravec: "... y es difícil o imposible lograr que posean las habilidades perceptivas y motrices de un bebé de un año".



¿Aprendiendo como un bebé? (Sergio García y Lola Moral, Noviembre 2014)

Inteligencia Artificial

La paradoja del Efecto IA y el Teorema Tesler

Teorema de Tesler: "La IA es cualquier cosa que no se ha hecho todavía."

(Douglas R. Hofstadter, Cödel, Escher, Bach, Un Eterno y Grácil Bucle, 1987, Tusquets)



Larry Tesler. Se le atribuye la creación de los comandos de cortar, copiar y pegar (Ctrl+X, Ctrl+C y Ctrl+V respectiva mente).

El **efecto IA** se produce cuando gente ajena al tema descarta el comportamiento de un programa de inteligencia artificial con el argumento de que no es inteligente.

Tan pronto como la IA soluciona con éxito un problema, el problema ya no es parte de la IA.

Pamela McCorduck lo llama una "extraña <u>paradoja</u>", en donde "los éxitos prácticos de la IA, programas computacionales que realmente logran comportamiento inteligente, fueron asimilados en cualquier otro dominio de aplicación en donde fueron encontrados para ser útiles y se convirtieron en socios silenciosos junto a otros enfoques para resolver problemas, lo que dejó a los investigadores de IA a tratar sólo con los "fracasos", las nueces duras que todavía no pueden ser quebradas."

https://es.wikipedia.org/wiki/Efecto_IA

Inteligencia Artificial

Debate: La singularidad tecnológica

Volviendo sobre el debate: ¿Pueden pensar las máquinas?







Raymond Kurzweil (Director de Ingeniería de Google)

predice que un ordenador pasará el test de Turing hacia el 2029, demostrando tener una mente (inteligencia, consciencia de sí mismo, riqueza emocional...) indistinguible de un ser humano.







199

Inteligencia Artificial Debate: La singularidad tecnológica

La **singularidad tecnológica** es el advenimiento hipotético de <u>inteligencia</u> <u>artificial</u> general (también conocida como «<u>IA fuerte</u>», del inglés *strong AI*). La singularidad tecnológica implica que un equipo de cómputo, red informática o un robot podrían ser capaces de automejorarse recursivamente, o en el diseño y construcción de computadoras o robots mejores que él mismo. Se dice que las repeticiones de este ciclo probablemente darían lugar a un efecto fuera de control —una explosión de inteligencia—12 donde las máquinas inteligentes podrían diseñar generaciones de máquinas sucesivamente más potentes. La creación de inteligencia sería muy superior al control y la capacidad intelectual humana.³ Carvalko, Joseph (2012). *The Techno-human Shell-A Jump in the Evolutionary Gap.* Sunbury Press. <u>ISBN 978-1620061657</u>.

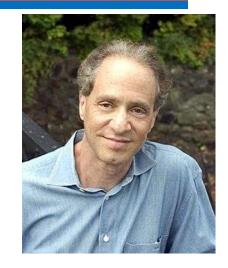
El primer uso del término «singularidad» se realizó en 1957 por el matemático y físico húngaro <u>John von Neumann</u>. Ese año, en una conversación con von Neumann, <u>Stanislaw Ulam</u> describió los avances cada vez más acelerados de la tecnología y los cambios en el modo de la vida humana. Da la apariencia de que se acerca una singularidad esencial en la historia más allá de los asuntos humanos, y tal como los conocemos:. Ulam, Stanislaw (16 de mayo de 1958). <u>Tribute to John von Neumann</u>. 64, #3, part 2. Bulletin of the American Mathematical Society. p. 5.

Inteligencia Artificial Debate: La singularidad tecnológica

Raymond Kurzweil predice que un ordenador pasará el test de <u>Turing</u> hacia el 2029, demostrando tener una mente (inteligencia, consciencia de sí mismo, riqueza emocional...) indistinguible de un ser humano. Este momento se ha denominado singularidad tecnológica (término popularizado por el matemático, informático y autor de ciencia ficción Vernor Vinge).

Kurzweil prevé que la primera inteligencia artificial será construida alrededor de una simulación por ordenador de un cerebro humano, lo que será posible gracias a un escáner guiado por nanobots. Una máquina dotada de inteligencia artificial podría realizar todas las tareas intelectuales humanas y sería emocional y autoconsciente.

Kurzweil sostiene que esa lA llegará a ser, inevitablemente, más inteligente y poderosa que la de un ser humano. Sugiere que las IA mostrarán pensamiento moral y respeto a los humanos como sus ancestros. De acuerdo a sus predicciones, la línea entre humanos y máquinas se difuminará como parte de la evolución tecnológica. Los implantes cibernéticos mejorarán en gran medida al hombre, lo dotarán de nuevas habilidades físicas y cognitivas y le permitirán interactuar directamente con las máquinas. https://es.wikipedia.org/wiki/Raymond_Kurzweil



En 1999 Raymond Kurzweil publicó un segundo libro titulado *La* era de las máquinas espirituales, que ahonda en la explicación de sus ideas futuristas.

Pensamiento sobre futurismo y transhumanismo

Inteligencia Artificial Debate: Ética de la Inteligencia Artificial ¿Existe el riesgo la Inteligencia Artificial?

Stephen Hawking: "La inteligencia artificial puede acabar con la raza humana" (Diciembre, 2014).

SPACEX

TESLA

SolarCity

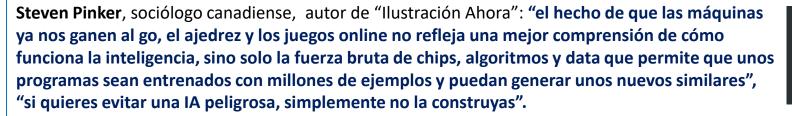




Stuart Russell "Desde el principio, el principal interés de la tecnología nuclear fue una fuente inagotable de energía", dijo Russell. "Creo que hay una analogía razonable entre una cantidad ilimitada de energía y cantidades ilimitadas de inteligencia. Ambas perspectivas parecen maravillosas hasta que uno piensa en los posibles riesgos" https://actualidad.rt.com/ciencias/180878-cientifico-inteligencia-artificial-peligroso-armas-nucleares

Byung-Chul Han «Los datos y las máquinas deben de estar al servicio de las personas, y no al revés. Una sierra mal utilizada también sirve para cortar cabezas»

Yuval Harari: «En cincuenta años ha habido un desarrollo extraordinario en la inteligencia de los ordenadores, pero un desarrollo cero en su conciencia. No hay ningún indicio de que se vayan a desarrollar en ese sentido».





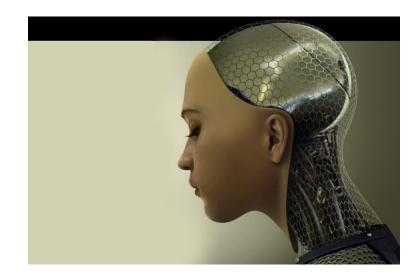


Inteligencia Artificial: La ficción











Inteligencia Artificial Debate: Ética de la Inteligencia Artificial ¿Existe el riesgo la Inteligencia Artificial?

¿Puede un ordenador ser más inteligente que un ser humano? probable es que sea de formas restringidas

Se subestima cuánta ciencia queda por hacer. Y resulta difícil separar la ficción de la realidad porque observamos cosas maravillosas y que también, a ojos humanos, parecen mágicas.

¿Podemos llegar a construir máquinas que piensen?



"La inacabada fábula de los gorriones"

Inteligencia Artificial Debate: Ética de la Inteligencia Artificial ¿Existe el riesgo la Inteligencia Artificial?

Superinteligencia: Se puede producir un ordenador que exceda la capacidad humana, que podría dirigir su inteligencia hacia si mismo, mejorando su propio diseño,



reescribiendo su software o usar su programación evolutiva para crear, probar y optimizar mejoras a su diseño. Esto conduciría a un proceso reiterativo de mejoramiento recursivo, cada vez más inteligente y más capaz.

Resultado: La superinteligencia, un ordenador con un intelecto superior totalmente ajeno al nuestro. Daría lugar a un acontecimiento que se ha dado en llamar "Singularidad".

Inteligencia Artificial: débil, fuerte y super



 Inteligencia Artificial débil: Una máquina puede ser capaz de competir con un humano en una actividad específica



 Inteligencia Artificial Fuerte: Una máquina con la capacidad de aplicar inteligencia para cualquier problema, en lugar de abordar un único problema (a nivel humano)

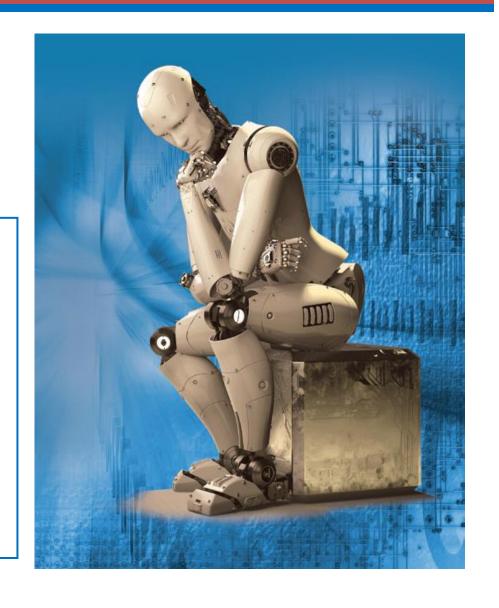


 SuperInteligencia Artificial: Una capacidad que es mucho mejor que cualquier cerebro humano en prácticamente cualquier campo, incluyendo creatividad científica, artística y habilidades sociales.

Inteligencia Artificial Hacia el futuro

¿Qué harán los robots dentro de 15 años?

"65% de los alumnos que acceden hoy a la escuela primaria trabajarán en profesiones que aún no existen"



Inteligencia Artificial Hacia el futuro

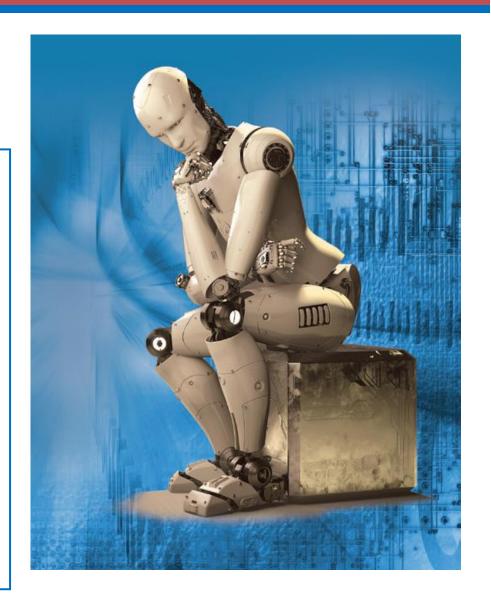
¿Qué ha ocurrido en los últimos 15 años?

Detrás de todos estos cambios están nuevas profesiones. Y más aún, la destrucción de ciertos puestos de trabajo para crear nuevos.

Social media manager.

Marketing en Redes
Conductor de Uber
Sustainability manager (preocupación de las empresas por la sostenibilidad)

Muchas de las profesiones actuales dejarán de existir en el futuro. El problema es que no sabemos cuáles ni en qué momento, ni por qué profesiones serán sustituidas.



Inteligencia Artificial Hacia el futuro: El empleo. Implicaciones sociales

Implicaciones sociales: La Inteligencia Artificial nos conducirá hacia una sociedad en la que se irá produciendo una automatización inteligente de muchas tareas.

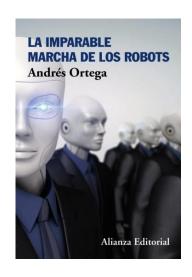
¿Cómo afectará al empleo, se transformarán los puestos de trabajo? ¿se perderán los repetitivos y se mantendrá los creativos?

¿se crearán profesiones que no existen?.

Inteligencia Artificial Hacia el futuro: Automatización inteligente. Implicaciones sociales

Implicaciones sociales: La Inteligencia Artificial nos conducirá hacia una sociedad en la que se irá produciendo una automatización inteligente de muchas tareas. ¿Cómo afectará al empleo, se transformarán los puestos de trabajo?, ¿se perderán los repetitivos y se mantendrá los creativos?, ¿se crearán profesiones que no existen?.

El libro de A. Ortega es a la vez un reportaje y un ensayo que analiza un proceso en marcha, más que las respuestas, lo importante son las preguntas que plantea:



¿Pueden los robots tener sentimientos? ¿Tienen género? ¿Pueden tomar decisiones morales?

¿Qué ocurrirá con nuestros empleos?

¿Pueden desempeñar cualquier labor?

¿Quién pagará las pensiones si los robots hacen nuestro trabajo?

¿Tendremos el tiempo y la flexibilidad suficientes para adaptarnos a la revolución que está en marcha?

Inteligencia Artificial Hacia el futuro: Hacia el futuro y Distopías

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation

https://www.eff.org/files/2018/02/20/malicious_ai_rep ort_final.pdf



Future of Humanity Institute, University of Oxford y OTROS

4 escenarios distópicos que involucran a la IA

Escenario 1: Manipulación del publico mediante informes de noticias falsas generados con IA

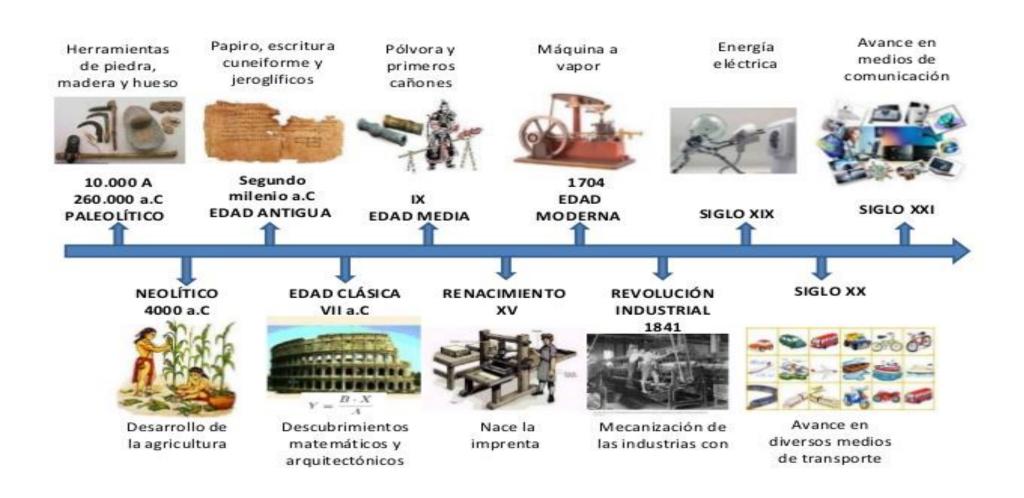
Escenario 2: La estafa más inteligente de suplantación de identidad y control de sistemas críticos

Escenario 3: La epidemia del malware (Ciberataques)

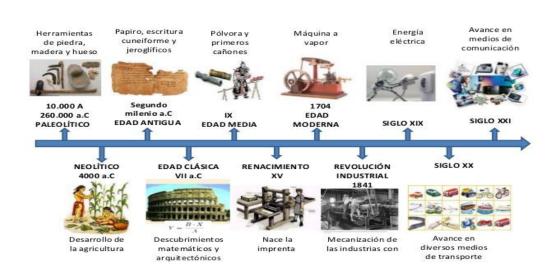
Escenario 4: El robot asesino (suplantar a un robot, detección identidad, explosión de bombas)

https://www.technologyreview.es/s/10012/cuatro-escenarios-black-mirror-que-aterran-los-expertos-de-ia

La inteligencia artificial Hacia el futuro: ¿En un futuro muy lejano?

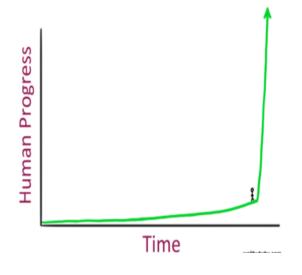


La inteligencia artificial Hacia el futuro: ¿En un futuro muy lejano?



Si queremos imaginar cómo será el mundo en 30 años (2052) no podemos hacerlo pensando en cómo era hace 30 años (1992), porque el progreso tecnológico no es lineal.

En los primeros 20 años del siglo XXI se ha avanzado tecnológicamente lo mismo que en todo el siglo XX



Inteligencia Artificial Hacia el futuro

¿Se acelera el progreso en IA?

No todos los campos de la Inteligencia Artificial evolucionan al mismo ritmo, en parte por también dependen de los avances en otros campos.

Ej. Capacidades físicas de los robots depende de los avances en materiales, diseño del motor ...

Aprendizaje automático y big data evolucionan rápidamente por la ingente cantidad de datos en formato digital (debido a Internet), ...

Inteligencia Artificial Para finalizar. Comentarios finales

5 Mitos sobre la IA

- Mito 1: La IA funciona de la misma manera que el cerebro humano
- Mito 2: Las máquinas inteligentes aprenden por su cuenta
- Mito 3: Al puede estar libre de sesgos
- Mito 4: Al solo reemplazará trabajos repetitivos que no requieren titulaciones avanzadas
- Mito 5: No todas las empresas necesitan una estrategia de Al

Inteligencia Artificial Para finalizar. Comentarios finales

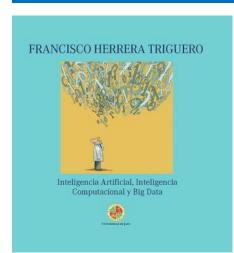
Hoy no podemos pensar en un mundo sin ordenadores, sin Internet, sin sistemas inteligentes y cada vez son más las tareas que estos realizan.

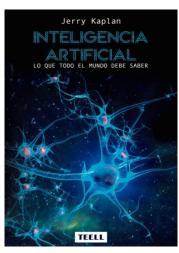
Convivimos con artefactos inteligentes, con capacidades limitadas, que nos permiten tener una vida más confortable, una mayor calidad de vida.

Hace 50 años no era imaginable el nivel de presencia que los ordenadores tendrían en nuestra sociedad.

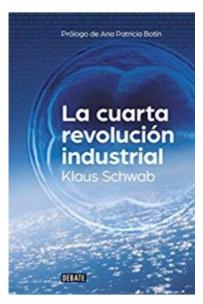
Si pudiésemos avanzar 50 años, ¿cuál será la presencia de los ordenadores y los sistemas inteligentes?.

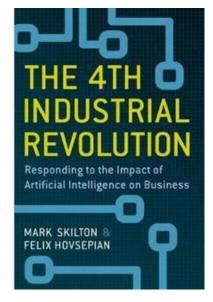
Inteligencia Artificial Conocer más: Inteligencia Artificial y Cuarta Revolución Industrial

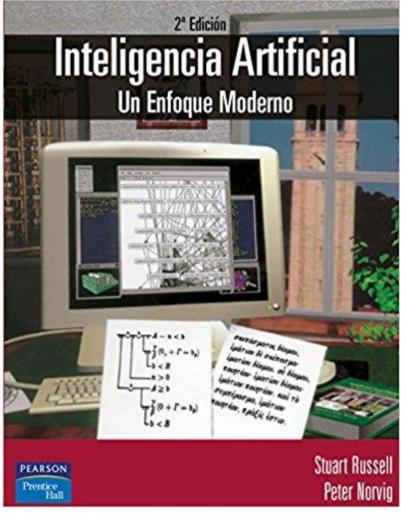












Inteligencia Artificial Conocer más: Inteligencia Artificial y el debate

Conocer más: Inteligencia Artificial y el debate sobre la IA

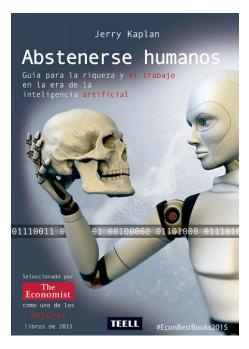


Artículo: Evolución del mundo gracias a la ciencia y la tecnología: https://magnet.xataka.com/en-diez-minutos/17-graficos-para-ensenar-a-quien-todavia-no-crea-que-el-mundo-va-cada-vez-mejor



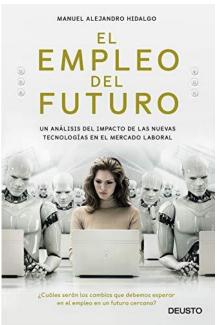
Inteligencia Artificial

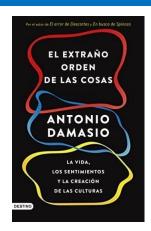
Conocer más: Inteligencia Artificial y el debate después de las profesiones











El neurólogo **Antonio Damasio**, situa los sentimientos en el centro del ser humano, nos dice que la **"racionalidad no se puede separar de la emoción, igual que el cuerpo es inseparable de la mente".**

"Los ingenieros están convencidos de que el sustrato no cuenta. Y eso es un gran error, porque no prestan atención a los sentimientos. La inteligencia artificial opera en el mundo del intelecto. Es un mundo muy preciso, pero nosotros somos mucho más blandos, operamos en un mundo de vida vulnerable. "

"Si vas a Silicon Valley y hablas con esa gente, no se ven a sí mismos y no ven a otros, son muy buenos con las matemáticas y con la computación y piensan que todo el mundo es código. Pero no lo es. Un ordenador no tiene enfermedades, no se va a resfriar, no va a tener cáncer. Nosotros estamos hechos de un material muy vulnerable. Si sales a la calle y cruzas con un semáforo en rojo y un coche te golpea te rompes los huesos y puedes morir. El sentido de si la vida está siendo buena o mala se expresa a través de los sentimientos. Los sentimientos expresan mentalmente si la homeostasis está funcionando o no y eso no lo puedes tener en un ordenador a menos que crees un cuerpo para el ordenador. Por eso, cuando alguien plantea que va a cargar su conciencia en un ordenador, pregunto si va a cargar también un cuerpo. Porque si no cargas también tu cuerpo, no vas a ser tú. Será un ordenador con tus ideas, pero no tendrá tus sentimientos."

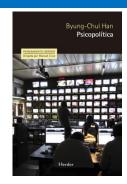
"La inteligencia artificial puede convertir a millones de personas en una clase inútil"

Con su primera obra analizó el pasado, con 'Homo Deus' nos describió un futuro probable, y con '21 lecciones para el siglo XXI' (Ed. Debate), se centra en los dilemas que plantea nuestro presente, ese instante de transición entre los cambios vertiginosos que se esperan y las estructuras sociales y las tradiciones que se dejan atrás.





https://www.elconfidencial.com/alma-corazon-vida/2018-08-30/inteligencia-artificial-empleo-tecnologia-religion_1609326/



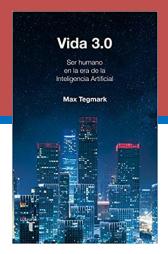
Byung-Chul Han, el filósofo surcoreano afincado en Alemania, que preconiza un retorno -tal vez imposible- a lo auténtico: «Hoy estamos en red, pero no estamos unidos. La comunicación actual se basa en no escucharse». «Los datos y las máquinas deben de estar al servicio de las personas, y no al revés. Una sierra mal utilizada también sirve para cortar cabezas»

"Steven Pinker en su ensayo «Ilustración Ahora», alabado por Bill Gates como «el mejor libro que he leído en mi vida». Pinker se ha convertido en el profeta del optimismo, el sabio que prueba con datos que «en contra de lo que se cree, el mundo va cada vez mejor».

A su juicio sobra histeria cuando se habla de la amenaza de la tecnología y las IA, un que él denomina con sarcasmo el «Robotapocalipsis». «La falacia radica en que se confunde inteligencia con motivación». Explica que «incluso si inventásemos unos robots súper inteligentes, ¿por qué iban a querer esclavizar a los humanos y conquistar el mundo? La inteligencia es la habilidad de desarrollar nuevas maneras de lograr una meta. Pero eso no es lo mismo que querer algo». Para Pinker, el hecho de que las máquinas ya nos ganen al go, el ajedrez y los juegos online «no refleja una mejor comprensión de cómo funciona la inteligencia, sino solo la fuerza bruta de chips, algoritmos y data que permite que unos programas sean entrenados con millones de ejemplos y puedan generar unos nuevos similares»."

https://elpais.com/elpais/2018/06/07/eps/1528366679_426068.html

http://www.abc.es/tecnologia/abci-disrupcion-tecnologica-mayor-revolucion-jamas-conocida-201803250339 noticia.html#ns_campaign=rrss-inducido&ns_mchannel=abces&ns_source=fb&ns_linkname=noticia.foto&ns_fee=0



En su libro 'Vida 3.0', el profesor del MIT propone argumentos para un debate global que evite que la llegada de la Inteligencia Artificial acabe en desastre

"Hay una gran presión económica para hacer obsoletos a los humanos"

El profesor del MIT y director del <u>Future of Life Institute</u> en Cambridge (EE UU) estima que la llegada de una Inteligencia Artificial General (IAG) que supere a la humana es cuestión de décadas. En su visión del futuro, podríamos acabar viviendo en una civilización idílica donde robots superinteligentes harían nuestro trabajo, crearían curas para todas nuestras enfermedades o diseñasen sistemas para ordeñar la energía descomunal de los agujeros negros. Sin embargo, si no somos capaces de transmitirle nuestros objetivos con precisión, también es posible que a esa nueva inteligencia dominante no le interese nuestra supervivencia o, incluso, que asuma un objetivo absurdo como transformar en clips metálicos todos los átomos del universo, los que conforman nuestros cuerpos incluidos. Para evitar el apocalipsis, Tegmark considera que la comunidad global debe implicarse en un debate para orientar el desarrollo de la inteligencia artificial en nuestro beneficio. Esta discusión deberá afrontar problemas concretos, como la gestión de las desigualdades generadas por la automatización del trabajo, pero también un intenso esfuerzo filosófico que triunfe donde llevamos siglos fracasando y permita definir y acordar qué es bueno para toda la humanidad para después inculcárselo a las máquinas.

https://elpais.com/elpais/2018/08/07/ciencia/1533664021 662128.html