

B-DEBATE

International Center
for Scientific Debate
BARCELONA

biocat



Sinopsis

ARTIFICIAL INTELLIGENCE: DREAMS, RISKS, AND REALITY

March, 7th and 8th, 2017

COSMOCAIXA BARCELONA. ISAAC NEWTON, 26. BARCELONA

www.bdebate.org

CO-ORGANIZED BY:



CSIC
upf.



Inteligencia artificial: sueños, riesgos y realidad

Hay muchas definiciones disponibles para describir qué es la inteligencia artificial, lo que da muestras de su carácter escurridizo. Una aproximada sería esta: un sistema computacional capaz de emular funciones propias del cerebro humano. Plasmada por el imaginario popular en forma de robots, la inteligencia artificial va en realidad mucho más allá, y se encuentra ya presente en muchos ámbitos de nuestra vida. Aunque aún con limitaciones, la tenemos tras la selección de las noticias que leemos, condicionando o al menos influyendo en los resultados electorales; está detrás de muchos de los anuncios que vemos, y nos encontramos a las puertas de convivir con algunas de sus nuevas aplicaciones, como la de los coches autónomos, guiados sin necesidad de conductor.

Pero toda oportunidad esconde también un riesgo. Para debatir unas y otros, varios de los mejores expertos internacionales se reunieron en [B-Debate](#), una iniciativa de [Biocat](#) y de la [Obra Social "la Caixa"](#) para promover el debate científico. Fruto de este encuentro se estableció la Declaración de Barcelona para el adecuado desarrollo y utilización de la inteligencia artificial en Europa.

CONCLUSIONES

- Resulta complicado predecir cómo evolucionará nuestra relación con los robots. Esta interacción abre todo un campo nuevo conocido como roboética
- La inteligencia artificial aplicada a la gran cantidad de datos disponibles hoy en día está creando cámaras de eco y filtros de burbuja en las redes sociales, amenazando con radicalizar ideologías
- Los algoritmos con los que opera la inteligencia artificial son en muchas ocasiones oscuros. Es necesario conseguir traducirlos a nuestro lenguaje para que sus decisiones puedan ser éticas y socialmente justas
- Ante las múltiples oportunidades y peligros se ha establecido la llamada Declaración de Barcelona para el adecuado desarrollo y utilización de la inteligencia artificial en Europa

1. LA INTELIGENCIA ARTIFICIAL: DE LOS ROBOTS A LOS DATOS

Sobre los robots

“La inteligencia artificial podría barrer a la humanidad cuando llegue a ser demasiado inteligente, ya que los humanos serán como hormigas”. Ilustrado con la fotografía de un enorme robot en el periódico [The Independent](#), este titular del físico Stephen Hawking dio la vuelta alarmando al mundo. Los expertos, sin embargo, no le dan demasiada importancia. Para [Héctor Geffner](#), profesor ICREA en la universidad Pompeu Fabra, los robots son todavía torpes y **“la inteligencia general de las máquinas a un nivel humano solo existe de momento en las películas**. Son capaces de reaccionar, pero tienen poca flexibilidad y carecen de vida mental”.

[Carme Torras](#), investigadora en el Instituto de Robótica e Informática Industrial del CSIC, comparte que **“hoy día, los humanos son más peligrosos que los robots”**, pero reivindica su desarrollo actual. Su grupo trabaja en robots asistenciales que ayudan a vestirse a personas dependientes, y reivindica que ya existen robots usándose en logística o como recepcionistas, robots que discriminan qué debe ir a la basura o a un lavavajillas, e incluso apunta el hecho de que haya robots-mascotas, diseñados para mostrar y provocar emociones. Esto último abre todo un campo ético que ya se ha dado en llamar **“roboética”**, y que para Torres viene motivada en parte por “la humanización de los robots, la percepción de que hay *un humano dentro*”. Para Torres la gran pregunta es: **“¿Cómo evolucionará la naturaleza humana con el aumento de la interacción entre humanos y robots?** Y, por extensión: ¿puede esta evolución predecirse?”

Uno de los problemas para estudiar esta evolución es la limitación del lenguaje a la hora de describir el futuro. En palabras de Heidegger: **“Es a través de la tecnología que percibimos el mar como navegable”**. De ahí que muchos de los dilemas se encuentren en el terreno de la **ciencia-ficción**, uno de cuyos papeles, según Torras, es el de “anticipar posibles escenarios”. Estos se encuentran en libros como los de **Asimov**, **Philip K. Dick** o **Ray Bradbury**, en películas como **“Eva”** o en series como **“Black Mirror”**. La propia Torras es autora del libro de ciencia-ficción **“La mutación sentimental”**, a través del cual está desarrollando un proyecto educativo.

Sobre los datos

El aumento en la capacidad de computación de los ordenadores y, sobre todo, la instauración de Internet y el incremento exponencial de los datos disponibles han hecho que **la inteligencia artificial comience a estar presente en muchos aspectos de nuestra vida**. Los algoritmos están detrás de muchas de las noticias que leemos y de los anuncios que nos llegan, pero también pueden estar detrás del análisis de los créditos o los seguros que pedimos, del filtro inicial de nuestros currículums o del estudio de algunos datos de nuestra salud. Y no es nada descabellado pensar que estarán también guiando nuestros propios vehículos, que en un futuro pasarán seguramente a ser completamente autónomos.

Una de las áreas más estudiadas es la de las redes sociales y el acceso a la información, con notables implicaciones y peligros. Por ejemplo a la hora de **condicionar elecciones**, como se sugiere que sucedió al usar la inteligencia artificial para personalizar los mensajes **en la campaña de Donald Trump**.

Independientemente de campañas puntuales, los algoritmos, según **Cornelius Puschmann** –investigador en el Hans Bredow Institute for Media Research en Hamburgo– “son una forma confusa en la que a veces se habla de la inteligencia artificial”, y son los responsables de “elegir” –en base a nuestro perfil y anteriores preferencias– qué noticias tienden a aparecer en los muros de nuestras redes sociales. Hay dos tendencias que los algoritmos y las redes hacen emerger, y ninguna está exenta de peligros. La primera, como comentó **Camilo Cristancho**, investigador en la Universidad Autónoma de Barcelona, es la formación de “**cámaras de eco**” en las que los individuos tienden a juntarse con sus similares. La otra son las “**burbujas de filtro**”: burbujas ideológicas consecuencia de la individualización de contenidos ofrecidos por los algoritmos en base a las preferencias pasadas, y que amenazan con potenciar en exceso las ideas ya establecidas, minimizando el matiz y la pluralidad.

¿Deberíamos preocuparnos por las burbujas de filtro? Según Puschmann no excesivamente. “Desde un punto de vista periodístico no es algo nuevo, llevamos 50 años estudiándolo”. Para **Walter Quattrociocchi**, jefe del Laboratorio de Ciencia Social Computacional en el IMT de Lucca, en Italia: “todavía es objeto de debate si estas burbujas son algo creado por el algoritmo o son una mera tendencia del ser humano”.

Según Puschmann, lo que parece más evidente es que **lo que causa miedo es la sensación de pérdida de control**, el hecho de que algo no-humano tenga capacidad de decisión. En un experimento con usuarios de Facebook, **más de la mitad desconocían que había un algoritmo detrás de sus muros**, y la primera reacción al ser informados fue

de sorpresa y rabia. Sin embargo, varios meses después, y tras haberles enseñado su funcionamiento, su satisfacción era similar al que tenían antes de habérselo “descubierto”. Y su uso de Facebook mayor de lo que era antes.

Quattrociochi investiga fundamentalmente cómo la información errónea se propaga por la red, lo que liga con la designada palabra del año en 2016, la **posverdad**, un concepto cuyo significado “denota circunstancias en que los hechos objetivos influyen menos en la formación de la opinión pública que los llamamientos a la emoción y a la creencia personal”. Y también con el llamado sesgo de confirmación, “la actitud que lleva a buscar la información de manera que confirme las propias creencias”. Incluso el Foro Económico Mundial ha señalado a la masiva información errónea digital como “uno de los principales riesgos para nuestra sociedad”.

Las investigaciones de Quattrociochi confirman la clara existencia de cámaras de eco, dentro de las cuales las dinámica de refuerzo son similares sea cual sea el motivo en común. Quizás lo más preocupante es que los usuarios más activos y “comprometidos” en esas cámaras tienden a ser **los que defienden teorías conspirativas** (al estilo de los que proclaman la existencia de OVNIS o la relación entre las vacunas y el autismo). Y aún más preocupante es la siguiente paradoja: la publicación dentro de esas cámaras de información rigurosa que busca desacreditar sus teorías no solo no lo consigue, sino que es contraproducente. Es decir, les insta a reforzar su posición inicial.

2. LIMITACIONES Y ÉTICA EN LA INTELIGENCIA ARTIFICIAL

A pesar de los logros y las promesas de la inteligencia artificial, esta cuenta todavía con serias e importantes limitaciones. Básicamente, los programas pueden funcionar muy bien en aplicaciones concretas, pero **están lejos de hacerlo en el impredecible mundo real**. Para Francisco Martín, presidente de BigML, las técnicas apenas han cambiado en los últimos años. Simplemente han aumentado los datos sobre los que pueden operar. Según Martín, “mucha gente está reinventando la rueda al mismo tiempo” sin hacer verdaderos avances. De hecho, los procesos “requieren todavía mucha experiencia humana y se hacen de forma muy manual”.

Una de las limitaciones tiene que ver con la oscuridad del proceso que siguen los programas hasta llegar a sus conclusiones. Muchos de ellos se basan en lo que se conoce como **redes neuronales**. Estas redes son construcciones matemáticas que asemejan en cierta manera cómo funciona nuestro cerebro, donde la información va saltando de

nivel y consolidándose de forma difusa. Son modelos muy potentes, pero necesariamente opacos, y han dado lugar a hablar de “**la caja negra de la inteligencia artificial**”. Algunos expertos, [como Pierre Baldi](#), no le dan importancia a esa oscuridad. Al fin y al cabo: “Usas tu cerebro todo el tiempo; confías en tu cerebro todo el tiempo; y no tienes ni idea de cómo funciona tu cerebro”. Otros, como [Marcello Pelillo](#), profesor de Ciencia Computacional en la Universidad de Venecia, reconocen que hay situaciones “**donde es importante poder explicar las decisiones**”. Por ejemplo, si afecta a la decisión de un juez o de una institución social. Y, en general, para garantizar el respeto a la autonomía y la dignidad individual.

Porque el uso de la inteligencia artificial implica numerosos dilemas y retos éticos. Estos son algunos de los propuestos por [Francesca Rossi](#), investigadora en IBM y profesora de Ciencias de la Computación en la Universidad de Padua:

- ¿Sustituirá la inteligencia artificial el trabajo de los humanos?
- ¿Cómo cambiará nuestra interacción con otros humanos y la sociedad o la educación?
- En un futuro, ¿qué sucederá ante la posibilidad de desarrollar armas autónomas?
- ¿Cómo será el desarrollo ético de la propia inteligencia artificial?

Para Rossi, **habrá que desarrollar los sistemas de forma que discriminen sus resultados de forma ética**. Esto tendrá que ser así, obviamente, en el caso de los sistemas autónomos, pero también en otros. Por ejemplo, en herramientas que aconsejen a los médicos en base al análisis de datos: “Si esas sugerencias no siguen un código ético, los propios médicos no confiarán en ese sistema”. La clave es cómo desarrollar ese código, ya que los procesos de decisión del programa, como hemos visto, son muchas veces oscuros por su propia naturaleza.

Esto no es sencillo. Por una parte por un tema de flexibilidad. Lo que a nosotros nos puede parecer algo de **sentido común** no tiene por qué serlo para una máquina. “Un robot de cocina no debería cocinar a nuestro gato si no hay nada en la nevera, aunque considere que esa sería una comida aceptable para nosotros”, ejemplifica Rossi.

Una solución podría ser programar según los códigos profesionales. “Pero estos tienen muchos huecos que nosotros rellenos con sentido común. Algo que es difícil de introducir en una máquina”.

¿Puede hacerse, entonces? Según Rossi, sí. “Hay varias maneras que ahora mismo se están estudiando”. El proceso, además, esconde ventajas. “Hacerlo puede ayudarnos a tomar consciencia y a comportarnos nosotros mismos de una forma más ética”.

3. LA DECLARACIÓN DE BARCELONA

Impulsado por los líderes del B-Debate **Luc Steels** –profesor ICREA en el Instituto de Biología Evolutiva de Barcelona– y Ramón López de Mantaras –director del Instituto de Investigación en Inteligencia Artificial de Barcelona–, durante el encuentro se propuso y consensuó la “Declaración de Barcelona para el adecuado desarrollo y utilización de la inteligencia artificial en Europa”, **un manifiesto abierto a nuevas firmas y comentarios** que pretende servir de guía para el progreso de la inteligencia artificial. El manifiesto reconoce la gran importancia de este campo en el futuro de la economía y las sociedades, pero también la preocupación por el posible uso inapropiado, prematuro o malicioso de las nuevas tecnologías. En base a ello, propone un código de conducta articulado en torno a seis puntos:

- **Prudencia:** a pesar del desarrollo de la inteligencia artificial y sus posibilidades, muchos problemas fundamentales todavía no están resueltos y requerirán avances radicales. Incluso es posible que algunos problemas nunca se resuelvan.
- **Fiabilidad:** los nuevos métodos deben ser verificados antes de estar disponibles. La nueva agencia para la robótica y la inteligencia artificial creada por el parlamento europeo podría asumir esta responsabilidad.
- **Transparencia:** cuando un sistema de inteligencia artificial toma una decisión, las personas afectadas deben poder recibir una explicación de por qué se ha tomado esa decisión. Sin embargo, muchas de las técnicas que “explican” los procesos internos del sistema todavía no están bien desarrolladas.
- **Responsabilidad:** hay una preocupación creciente por los sistemas de mensajería automática o chat-bots que operan en internet y son capaces de manipular a la opinión pública. Es necesario dejar claro a los usuarios si los mensajes proceden de un sistema de inteligencia artificial.
- **Autonomía restringida:** los sistemas de inteligencia artificial, cuando están incluidos en sistemas físicos como los coches autónomos, tienen el potencial de actuar a través de sus decisiones en el mundo real. Es necesario tener reglas claras que limiten su autonomía, así como clarificar quién es el responsable en caso de fallo.
- **El papel humano:** el entusiasmo por la inteligencia artificial da lugar en ocasiones a la impresión de que la inteligencia humana ya no será necesaria. Esto es un error importante. Todos los sistemas dependen de la inteligencia humana,

por lo que la experiencia de las personas debe continuar siendo enseñada, desarrollada y ejercitada.

Finalmente, dada la importancia creciente de la inteligencia artificial en la economía, los firmantes del manifiesto llaman a las agencias de financiación europeas y a las empresas a invertir en su desarrollo de forma adecuada y estructural, creando las condiciones para estimular el emprendimiento. Además, **subrayan la importancia de desarrollar plataformas de recursos abiertos**, una infraestructura común compartida que favorezca el desarrollo de las próximas innovaciones.