



¡AI, AI! ¡Qué vienen las Amenazas Inteligentes!

La popularización del término Inteligencia Artificial no lleva asociado su comprensión por parte de la población. Ahora aparece como algo nuevo, pero en realidad tiene profundas raíces en sueños antiguos de la humanidad. Antes se llamaban robots, pero las nuevas campañas no quieren llamarlo así. Como las consecuencias de esta tecnología pueden llegar a lo más alto en las escalas de riesgo, presentes y futuros, de la humanidad, puede ser interesante echar un vistazo, aunque incompleto, a lo que se esconde detrás de esas dos letras, A.I.

Pocos se acuerdan del ilustre rabino Judah Loew ben Bezalel, conocido como el Maharal de Praga en el siglo XVI. La tradición le imputa haber creado un Golem para defender de los antisemitas el gueto de Praga a la vez que, de paso, se encargaba del mantenimiento de la Sinagoga.

El primer Golem se remontaría a los orígenes de la humanidad, y según la leyenda fue creado por un creyente cercano a Dios. Al igual que Adán, el Golem habría sido creado a partir del barro al que después una chispa divina le habría dado la vida. Desde este punto de vista, toda persona con un cierto grado de santidad y acercamiento a Dios podría adquirir algo de la sabiduría y el poder divinos. Siempre ha sido muy deseado ese importante poder que insufla vida a la materia inerte.

Sin embargo, la tradición está de acuerdo, más allá del grado de santidad que tuviera el creador de un Golem, de que el ser creado no sería más que una sombra de aquel creado por Dios, ya que, significativamente y por definición, todo Golem carece de alma.

En realidad el Golem es fuerte, pero no inteligente. Ordenada una tarea, la llevará a cabo de un modo lento, sistemático, y ejecutando las instrucciones de un modo literal, sin cuestionamiento alguno. Como ejemplo de ello, cuenta la leyenda que la esposa del rabino pidió al

Golem que fuera al río a sacar agua, y éste comenzó a hacerlo sin parar, hasta que terminó inundando toda la ciudad.

Para hacerlo funcionar (*boot*), el maestro del Golem habría de meterle un papel con una orden (*software*) por la boca u otro orificio. Otra manera de hacerlo funcionar era inscribir en su frente alguno de los Nombres de Dios (*password*) o bien la palabra *Emet* (אמת, "verdad") (*password* por defecto). Al borrar la primera letra de "Emet" de su frente y quedar sólo "met" (מת, "muerte"), el Golem se desactivaba (*shutdown*) y volvía a ser una masa de barro inerte (*hardware*).

anuncios¹ en sí, han sido creados con supuestas Intelligencias Artificiales², y hablan de coches que, a su vez, son también flamantes inteligencias artificiales. A pesar de lo mucho que se menciona, no hay una definición clara de qué es eso de la AI.

El batiburrillo de términos

Lo que recibe la población, los periodistas, políticos y tertulianos en general, es un batiburrillo de términos que engloban, el *Machine Learning*, *Deep Learning*, *Artificial Intelligence*, *Big Data*, etc., pero nadie parece estar interesado en saber lo

cas sin utilizar o recibir instrucciones explícitas para ello, sino basándose sólo en patrones e inferencias que se le proporcionan. Dichos algoritmos terminan construyendo un modelo, su modelo, basándose en un conjunto concreto de datos ("datos de entrenamiento"), para luego poder hacer predicciones o tomar decisiones sin haber sido explícitamente programados para ello. Lo que está claro es que el sistema no aprenderá más allá de lo que esté contenido en el conjunto de datos que se le han ofrecido como alpiste, luego el sesgo, la parcialidad y la subjetividad de su creador están servidas.



Los algoritmos que se obtienen con estas tecnologías en boga no son inteligentes pero, dado el lugar en el que se instalan y en los que se desea instalar, son tecnologías que mandan, actúan y no son responsables de las consecuencias de sus actos.

Se habla mucho de la Inteligencia Artificial; "AI" para los amigos. De un tiempo a esta parte todos los anuncios, los de casi cualquier cosa, la mencionan como si de un pequeño espíritu maravilloso se tratase. Algunos de esos

que realmente hay detrás de ellos. Pasemos a explicarlos y valorarlos someramente.

Machine learning (ML) son algoritmos³ y modelos estadísticos a ser ejecutados en ordenadores de todo tipo, que hagan tareas especifi-

Deep Learning (DL) es un caso particular de *Machine Learning* que utiliza redes neuronales artificiales (ANNs) sometidas a entrenamientos supervisados o incluso sin supervisión. Las redes neuronales⁴ están inspiradas

¹ Ver <https://youtu.be/6qEbgOKXpLg>

² Ver <https://ipmark.com/anuncio-de-lexus-inteligencia-artificial/>

³ "A computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P* if its performance at tasks in *T*, as measured by *P*, improves with experience *E*" definición de Tom Michael Mitchell, profesor de la Universidad de Carnegie Mellon.

⁴ W. McCulloch y W. crearon en 1943 un modelo computacional denominado "threshold logic". Este modelo abrió el camino para la investigación en redes neuronales. Este trabajo llevó al estudio de redes de nervios y su relación con los autómatas finitos. Por su parte, el algoritmo bautizado como Perceptron se inventó en 1958 el Cornell Aeronautical Laboratory por Frank Rosenblatt y con dinero de la United States Office of Naval Research. El Perceptron se inventó pensando en ser una máquina en lugar de un programa de ordenador.

en el procesado distribuido de la información que se da en los sistemas biológicos, sin embargo, unas y otros se diferencian considerablemente entre sí. En general, *Deep Learning* es un tipo de algoritmo de aprendizaje que consta de múltiples capas que progresivamente van extrayendo características más abstractas de los datos de entrada.

Por su parte, **Big Data** es el campo de actividad humana que busca modos de analizar y **extraer sistemáticamente información de un conjunto de datos**, o bien tratar conjuntos de datos que son tan grandes y tan complejos que resultan intratables con el software habitual de tratamiento

En realidad los algoritmos que se obtienen con estas tecnologías en boga **no son inteligentes** pero, dado el lugar en el que se instalan y en los que se desea instalar, **son tecnologías que man-**



El problema empieza cuando el trabajo de los robots y de los seres humanos se vuelve indistinguible; un empobrecimiento en las capacidades del hombre favorece el crecimiento de la robotización. La AI y la nueva robótica quieren quedarse con el trabajo intelectual de los seres humanos, y decidir, reconocer, diagnosticar, actuar libremente y juzgar, como los humanos y a los humanos, entre otras cosas.

dan, actúan y no son responsables de las consecuencias de sus actos.

Ejemplos de la aplicación de AI en nuestra vida real los tenemos en el coche autó-

coreano, en drones civiles¹⁴ y militares¹⁵, en herramientas de auditoría, en el marketing digital, etc. Cualquier error o mero “descontrol” en cualquiera de estos sistemas tendrá causas, muchas de ellas

escritor checo Karel Capek en su obra *“Los Robots Universales de Rossum”* (1920) que es donde aparece por primera vez la palabra robot.

En ruso, el verbo работать (*rabotatch*) es “trabajar”, y la

letales, sobre la población humana de este planeta y, en algunos casos, los errores pueden tener dimensiones considerables.

raíz RBT también se usa para designar el trabajo en alemán (*arbeit*) y en noruego (*arbeite*). En realidad todas estas palabras derivan de una: la raíz indoeuropea usada para referirse a “alejar”, “cambiar de estatus” o “lealtad”. La idea original de esta raíz está en una reducción de estatus similar a la que uno sufre cuando se convierte en esclavo.

Los robots fueron extremadamente populares en la filmografía¹⁷ del Siglo XX y, con formas absolutamente imprevistas, terminaron colonizando todo el tejido industrial en el mundo; empezando por la industria del automóvil, y llegando al montaje de todos los circuitos electrónicos, pasando por la recolección de cosechas y productos agrícolas, la mecanización digital y automática de cualquier pieza física, y la gestión de la paquetería planetaria.

Hace ya tiempo que los robots hacen nuestro trabajo físico y, como el Golem, lo hacen muy bien y sin rechistar (no hay sindicatos robóticos), pero ahora estamos entrando en otro ámbito mucho más peligroso. El problema empieza cuando el trabajo de los robots y de los seres humanos se vuelve indistinguible. **Un empobrecimiento en las capacidades del hombre favorece el crecimiento de la robotización.**

La Inteligencia Artificial y la nueva robótica quieren



La AI pretende inocentemente ser el resultado del mero entrenar estructuras lógicas para repetir con tenacidad y precisión lo mismo que se les enseña, pero realmente lo que hacen es que no queden huellas de ello. Las relaciones causales y la responsabilidad desaparecen. ¿Por qué un algoritmo da una determinada respuesta? Porque ha sido entrenado para ello, porque lo dice el ordenador, pero no por alguna razón auditable y conocida de ello.

de datos. Los objetivos del Big Data son la obtención de los datos, su almacenamiento, análisis, consulta (*querying*), compartición, transferencia, visualización, actualización, su privacidad y la autenticidad de sus fuentes, siendo estas dos últimas sus asignaturas pendientes.

nomos y sus víctimas⁵, en la selección de personal en recursos humanos⁶, en el control de fronteras y aduanas⁷, en los negocios⁸ y el control de la población⁹, en todos los aspectos de la Sanidad¹⁰, en las Armas Letales Autónomas (LAW¹¹), en armas autónomas¹² como el SGR-A1¹³

El antecedente de los Robots

Antes de empezar a hablar de AI, lo que la población tenía en mente eran los “Robots”¹⁶. La palabra robot viene del checo *robota* (trabajo forzado) y *rabota* (servidumbre) y la popularizó el

⁵ Ver <https://soymotor.com/coches/noticias/el-coche-autonomo-de-uber-detecto-la-victima-pero-no-freno-tiempo-946937> y <https://www.autobild.es/noticias/accidentes-coche-autonomo-cuantos-van-ya-196922>

⁶ Ver <https://blog.acsendo.com/proceso-de-seleccion-de-personal-ia/>

⁷ Ver <https://ecija.com/detectores-de-mentiras-basados-en-inteligencia-artificial-para-control-migratorio-el-futuro-de-las-fronteras-europeas/>

⁸ Larry D. Wall: “Some financial regulatory implications of artificial intelligence”, Journal of Economics and Business. Vol 100, Pages 55-63. Nov-Dec 2018.

⁹ Ver <https://www.lavanguardia.com/tecnologia/20190202/46161322564/china-inteligencia-artificial-ciudadanos.html>

¹⁰ Ver <https://www.redaccionmedica.com/secciones/tecnologia/big-data-e-ia-van-por-delante-de-las-leyes-de-control-de-datos-sanitarios-5064>

¹¹ Ver https://en.wikipedia.org/wiki/Lethal_autonomous_weapon

¹² Ver <http://www.ensayos-filosofia.es/archivos/articulo/riesgos-asociados-al-desarrollo-de-robots-autonomos-dotados-de-inteligencia-artificial-avanzada-en-contexto-civil-y-militar>

¹³ Ver <https://en.wikipedia.org/wiki/SGR-A1>

¹⁴ Ver <https://theintercept.com/2019/05/11/drones-artificial-intelligence-europe-roborder/>

¹⁵ Ver <https://dronprofesional.com/blog/drones-robotica-e-inteligencia-artificial-en-las-nuevas-armas-de-guerra/>

¹⁶ Ver https://en.wikipedia.org/wiki/History_of_robots

¹⁷ Ver https://en.wikipedia.org/wiki/Category:Robot_films

quedarse con el trabajo intelectual de los seres humanos. La AI ahora quiere **decidir, reconocer, diagnosticar, actuar libremente y juzgar**, como los humanos y a los humanos, entre otras cosas.

Un ejemplo de lo que está pasando realmente, y que más de un lobby nacional e internacional están fomentando, lo tenemos en el caso de la empresa china **SenseTime**¹⁸; empezó siendo una empresa centrada en la Inteligencia artificial y el reconocimiento de imágenes, pero pronto se convirtió en un proveedor esencial de cualquier estado policial¹⁹. Sus servicios y productos de reconocimiento de personas llegan a unos niveles de precisión e inmediatez nunca antes vistos²⁰.

La AI pretende inocentemente ser el resultado del mero entrenar estructuras lógicas especialmente plásticas para repetir con tenacidad y precisión lo mismo que se les enseña, pero realmente **lo que hacen es que no queden huellas de ello**. Las relaciones causales y la responsabilidad desaparecen con la Inteligencia Artificial. ¿Por qué un algoritmo da una de-

zan son secretos celosamente guardados y codiciados por muchos. No son procesos o tecnologías auditables y sus propietarios quieren que siga siendo así, en cualquier caso.

De hecho, la Inteligencia Artificial empezó a gobernar la realidad humana cuando el primer directivo hizo caso a los resultados de simulación compleja hechos con una máquina. Se les suele llamar **Sistemas de Ayuda**



La AI y los nuevos robots, aun siendo los actuales seres inmateriales, al igual que los Golems medievales siguen estando hechos al capricho inconfesable de sus autores. Los filtros y algoritmos que esas compañías desarrollan y utilizan son secretos celosamente guardados y codiciados por muchos. No son procesos o tecnologías auditables y sus propietarios quieren que siga siendo así, en cualquier caso.

a **la Toma de Decisiones**²¹, y en ellos no se suele poner en duda ni la "calidad" de los datos que procesaban, ni la "corrección" de los tratamientos que aplican. Como en muchos otros casos, la Seguridad de estos sistemas no estuvo, y sigue ausente, en sus diseños, implementaciones, implantaciones y explotación.

y que realmente no se sepa qué es lo que está haciendo, y qué o quién está tomando las decisiones. El valor de los algoritmos AI es que sean secretos por lo que **nunca serán auditables** o correctamente auditados, **ni se podrá demostrar su integridad** mientras trabajan. Nunca se podrá probar que el algoritmo que corre es el que realmente se ha auditado/autorizado.

La AI es un monstruo hambriento con dos dependencias esenciales: la energía eléctrica y los datos. El proceso de entrenamiento de las distintas materializaciones de la AI necesitan correr largos procesos de entrenamiento para conseguir plasmar en sus algoritmos la esencia que el diseñador/docente está buscando, y

van los gastos que se dan cuando las AIs se ponen en producción. **La Inteligencia Artificial tiene una huella de CO₂ que no es despreciable**, y de ella deberíamos también hablar cuando se habla del calentamiento global y de la calentita extinción que preconiza.

Por otra parte, la Inteligencia Artificial no es nada sin datos, por lo que sus campañas de normalización

en realidad lo que ocultan es la **ansiedad casi enfermiza de los ingenieros de datos por que les dejen capturar y meter mano, a todos los datos privados** ya que, según ellos, necesitan saberlo todo para destilar ese perfume, ese aceite esencial, que llamamos intimidad humana y que otros llamarían alma.

No nos engañemos, aquí volvemos a ver nuestro viejo enemigo del totalitarismo social que gobiernan unos pocos. En este caso, los que deban mantenerlo en pie ya no necesitarían ser soldados y fanáticos seguidores adecuadamente descerebrados, sino "Inteligencias Artificiales" que hagan exactamente lo mismo, pero mejor, mucho más discretamente y sin descanso alguno. Alguien calificó a estas tecnologías de Big Data e Inteligencia Artificial como "armas



El valor de los algoritmos AI es que sean secretos por lo que nunca serán auditables o correctamente auditados, ni se podrá demostrar su integridad mientras trabajan. Nunca se podrá probar que el algoritmo que corre es el que realmente se ha auditado/autorizado.

terminada respuesta? Porque ha sido entrenado para ello, porque lo dice el ordenador, pero no por alguna razón auditable y conocida de ello.

La AI y los nuevos robots, aun siendo los actuales seres inmateriales, al igual que los Golems medievales, siguen estando hechos al capricho inconfesable de sus autores. El mejor ejemplo lo tenemos en los algoritmos de las redes sociales y los buscadores de Internet como Google. Los filtros y algoritmos que esas compañías desarrollan y utili-

El valor de los algoritmos AI

La inteligencia artificial productiva persigue trabajar en la sombra de todos

eso requiere ordenadores y mucha electricidad. Cualquier cambio del sistema AI requiere un re-entrenamiento, por lo que su consumo vital de electricidad es prácticamente continuo, aparte

¹⁸Ver <https://en.wikipedia.org/wiki/SenseTime>

¹⁹Ver <https://www.hrw.org/news/2018/12/26/chinas-bumblng-police-state>

²⁰Ver <https://www.efe.com/efe/espana/portada/los-ojos-ciberneticos-mas-avanzados-estan-en-china-seguridad-control/10010-3832751>

²¹Ver https://en.wikipedia.org/wiki/Decision_support_system

²²Cathy O'Neil: "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy" Random House LLC. September 2016. ISBN-13: 978-0553418811

²³Ver https://es.wikipedia.org/wiki/Imperativo_categorico

de destrucción masiva”²² y estoy completamente de acuerdo con ello. Todavía estoy a la espera de ver una propuesta de la pomposamente denominada Inteligencia Artificial que sea real y esencialmente buena para la humanidad; me gustaría ver si llegaría a satisfacer el Imperativo Categórico de Kant²³.

La AI como distopía robótica

La distopía robótica, ahora llamada Inteligencia Artificial²⁴, se puede evitar, pero para ello hay que actuar desde ahora mismo²⁵. Mientras se discuten sobre los derechos a la vida y a la muerte, el software para someter y probablemente destruir vidas, se está desarrollando rápidamente. Incluso se habla de una nueva “carrera armamentística”, esta vez construida alrededor de la Inteligencia Artificial²⁶.

Como ha dicho el profesor de AI, Noel Sharkey²⁷, “Las guerras robóticas son una realidad. Los ejércitos quieren dar el poder de la

En este contexto, el pasado 22 de mayo los países que componen la OCDE²⁸ acordaron un conjunto de principios²⁹ que, según ellos, deberían guiar la gobernanza de la Inteligencia Artificial. La idea básica es que, en todo momento,



La AI no es nada sin datos, por lo que sus campañas de normalización en realidad lo que ocultan es la ansiedad casi enfermiza de los ingenieros de datos por que les dejen capturar y meter mano, a todos los datos privados, privados ya que, según ellos, necesitan saberlo todo para destilar ese perfume, ese aceite esencial, que llamamos intimidación humana y que otros llamarían alma.

tendrían que respetar los derechos humanos y los valores democráticos.

Los principios preconizados son prácticos y se han elegido también flexibles con la esperanza de poder resistir la rápida evolución que todo el mundo espera que tenga este ámbito. Esta declaración viene a complementar las posiciones anteriores de esa organización en temas como la defensa de la privacidad, la gestión

de mejorar el bienestar y la calidad de vida de la gente, de contribuir a un sostenimiento positivo de la actividad económica, de incrementar la innovación y productividad, y de ayudar a dar respuesta a retos globales clave”.

La OCDE arranca con un planteamiento optimista y positivo de la AI como “... una tecnología de propósito general que tiene la capaci-

dad de mejorar el bienestar y la calidad de vida de la gente, de contribuir a un sostenimiento positivo de la actividad económica, de incrementar la innovación y productividad, y de ayudar a dar respuesta a retos globales clave”.

La OCDE es la organización heredera del Plan Marshall³⁰ posterior a la II Guerra Mundial, y está netamente centrada en la actividad económica del mundo

los derechos económicos se enfrenten, y deba prevalecer sólo uno de ellos en detrimento del otro. Yo creo que serán siempre los derechos humanos los que pierdan.

No creo que multitudinarias declaraciones como la mencionada vayan realmen-

te a poner coto a las guerras robóticas, pero está claro que es necesario, desde ya mismo, desvelar claramente cuáles son los intereses reales, en cualquier escala de tiempo, de todos aquellos agentes que propugnan y trabajan a favor de la investigación, el desarrollo y la implantación silenciosa e isotrópica de la AI en todos los aspectos de la realidad planetaria.

Los promotores de la AI, en esencia, son de dos tipos, los ciegos de vidriosos y opalescentes ojos, al estilo de Jorge de Burgos³¹ en “El Nombre de la Rosa” de Umberto Eco, que desde su esencial inseguridad desconfían de la raza humana al más puro estilo *hobbiano*³², o los otros, los insensibles mercaderes inmediatoplacistas cuyo lema heráldico no es otro que el de “toma el dinero y corre”, y que en nada empatizan con los que habrán de sufrir las consecuencias de sus muy obscenos beneficios. ■



Es necesario, desde ya mismo, desvelar claramente cuáles son los intereses reales, en cualquier escala de tiempo, de todos aquellos agentes que propugnan y trabajan a favor de la investigación, el desarrollo y la implantación silenciosa e isotrópica de la AI en todos los aspectos de la realidad planetaria.

vida y de la muerte a máquinas sin razón ni conciencia”. ¿Cómo responderá la comunidad internacional, europea y/o española?

del riesgo digital y las conductas empresariales responsables. El 9 de junio de este 2019, los miembros del G20 adoptaron los princi-

occidental y del capitalismo global. Por su historia, en realidad no está claro de qué parte estará la OCDE cuando los derechos humanos y

²⁴Ver <https://www.bbc.com/news/technology-30290540>

²⁵Ver https://en.wikipedia.org/wiki/Existential_risk_from_artificial_general_intelligence

²⁶Ver <https://towardsdatascience.com/the-ai-arms-race-in-2019-fdca07a086a7>

²⁷Ver <https://www.theguardian.com/commentisfree/2007/aug/18/comment.military>

²⁸Además de los países miembros de la OCDE, también han firmado esa declaración Argentina, Brasil, Colombia, Costa Rica, Perú y Rumania.

²⁹Ver <https://legalinstruments.oecd.org/api/print?ids=648&lang=en>

³⁰Ver https://es.wikipedia.org/wiki/Plan_Marshall

³¹Ver <https://sites.google.com/site/pelliculaciei/jorgue-de-burgos>

³²“Lupus est homo homini, non homo, quom qualis sit non novit” locución creada por Plauto (254-184 a.C.) en su obra *Asinaria* (Comedia de los asnos), y popularizada por Thomas Hobbes.

JORGE DÁVILA
Consultor independiente
Director
Laboratorio de Criptografía
**LSIS – Facultad
de Informática – UPM**
jdavila@fi.upm.es