

10 TEXTOS SOBRE INTELIGENCIAS ARTIFICIALES

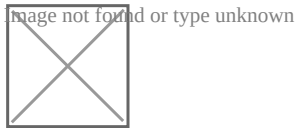
Por: Hipermediaciones. 04/01/2024

Quizás el primer libro que leí sobre los procesos de transformación digital fue *Le technologie dell'intelligenza. L'avvenire del pensiero nell'era dell'informatica* de **Pierre Lévy**, un volumen publicado por la editorial independiente Synergon en Bolonia en 1992 y editado por **Franco Bifo Berardi**. Un año antes **Tim Berners-Lee** había subido a su servidor en el CERN la primera página web y se comenzaban a percibir señales de cambio.

Las lecturas son importantes. En ese libro **Pierre Lévy** juega con dos conceptos – **interfaz e hipertexto**– que modelaron todo lo que pensé, escribí e investigué en las tres décadas siguientes. Entre tantas lecturas siempre me fui cruzando aquí y allá con el tema de las **inteligencias artificiales** (insisto, **debemos hablar en plural cada vez que abordamos estas tecnologías**), ya sea volúmenes dedicados totalmente a ellas o menciones al interno de obras sobre el mundo digital.

Desde la apertura al público del **ChatGPT** el 30 de noviembre de 2022 (parece que hubiera pasado una década... pero solo transcurrieron 12 meses!) se han escrito, publicado y reeditado infinidad de libros sobre las IA. Como se sabe, muchos libros sobre el **ChatGPT** han sido escritos en buena parte por el **ChatGPT**. Como decía un corolario a la **Ley de Murphy**, «**cuanto más se escribe sobre un tema, menos se sabe de él**». En eso estamos... En el último año he leído miles de páginas (sí, miles) sobre las IA. Páginas encuadernadas, páginas grapadas en su ángulo superior izquierdo, páginas PDF en pantalla, páginas en cuatricromía (no se pierdan el número de [Wired](#) dedicado a OpenAI) o manchadas de toner y mate.

Siguiendo con una tradición de este blog ([10 libros sobre ecología de los medios](#), [10 libros sobre evolución de los medios](#), [10 libros sobre narrativas transmedia](#)), a continuación propongo en estricto orden de publicación en su lengua original **10 textos -no sólo libros- sobre las IA**. He elegido **obras de corte sociotecnológico o histórico** que nos permiten mapear el territorio y evaluar con una cierta solidez teórica y analítica los procesos que se están desplegando a nuestro alrededor. No encontrarán en esta lista textos dedicados al *Machine Learning* o al *prompting*, ni manuales sobre el uso de las IA generativas en el aula. **Esas son otras listas.**



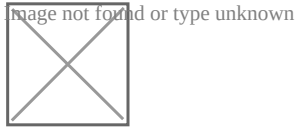
1 – Maquinaria computacional e Inteligencia – Alan Turing (1950)

Con el permiso de los trabajos pioneros de [Ada Lovelace](#) en el siglo XIX, [Computing Machinery and Intelligence](#) (en español, [Maquinaria computacional e inteligencia](#)) es quizás el padre de todos los textos contemporáneos sobre las IA. El artículo de [Alan Turing](#) se abre con **una pregunta que hoy, después de siete décadas, sigue más que vigente:**

Propongo considerar la siguiente pregunta: “¿Pueden pensar las máquinas?”. Se debiera comenzar definiendo el significado de los términos ‘máquina’ y ‘pensar’. Estas definiciones deberían ser elaboradas de manera tal que reflejen lo mejor posible el uso normal de estas palabras, pero una actitud así es peligrosa. Si el significado de las palabras ‘máquina’ y ‘pensar’ proviene del escrutinio de cómo son usadas comúnmente, se hace difícil escapar de la conclusión de que el significado y respuesta a la pregunta “¿pueden las máquinas pensar?” debiera ser buscado en una encuesta estadística, tal como la encuesta Gallup. Pero eso es absurdo. En vez de intentar una definición así, propondré reemplazar esa pregunta por otra, la cual se encuentra estrechamente relacionada y que se puede expresar en palabras relativamente poco ambiguas.

A partir de ahí **Turing** despliega el mítico **test** que lleva su nombre (también conocido como «*imitation game*») que sigue alimentando las conversaciones sobre

las IA hasta el día de hoy. Si alguien desea armarse un recorrido de lecturas centrado las IA, debería comenzar por Turing.



2. La cuarta discontinuidad. La coevolución de hombres y máquinas – Bruce Mazlish (1993)

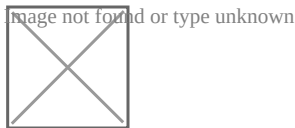
Sigo con un clásico que acabo de releer. Según escribe [Bruce Mazlish](#) en [La cuarta discontinuidad](#), a lo largo de la evolución reciente del Homo sapiens se produjeron **tres discontinuidades**, o sea, rupturas que cambiaron cómo los humanos se ven a sí mismos y su relación con el mundo. Estas discontinuidades son cruciales para entender el desarrollo del pensamiento y la sociedad modernos:

- **La discontinuidad copernicana:** Esto se refiere al cambio de una visión geocéntrica (centrada en la Tierra) a una heliocéntrica (centrada en el Sol), como propuso **Copérnico**. Fue un cambio radical en la forma en que los humanos entendieron su lugar en el cosmos, alejándose de la idea de que la Tierra y, por extensión, la humanidad, estaban en el centro de todo.
- **La discontinuidad darwiniana:** Esta discontinuidad se basa en la teoría de la evolución por selección natural de **Charles Darwin**. Desafió la visión de los humanos como una creación única, separada del reino animal. En cambio, planteó que los humanos evolucionaron de otras formas de vida a lo largo de millones de años, cambiando así la percepción de la especialidad humana en el orden natural.
- **La discontinuidad freudiana:** La teoría psicoanalítica de **Sigmund Freud** introdujo la idea de que gran parte del pensamiento y comportamiento humano está influenciado por procesos inconscientes y experiencias de la primera infancia. Esto generó una discontinuidad en la comprensión de la mente humana, sugiriendo que las personas no son completamente conscientes ni tienen el control total de sus propias mentes y comportamientos.

Mazlish también previó el impacto de las inteligencias artificiales y la tecnología en la auto-percepción humana. Esta **cuarta discontinuidad** se refiere a cómo el desarrollo de máquinas inteligentes desafía el estatus único de la inteligencia

humana y las capacidades de resolución de problemas. Escribe Mazlish:

Los hombres han llegado al umbral decisivo para romper con la discontinuidad entre ellos y las máquinas. Esta tesis consta de dos partes. Por un lado, los humanos están acabando con la discontinuidad porque ahora perciben su propia evolución como algo inextricablemente entrelazado con el uso y desarrollo de herramientas, de las que las máquinas modernas no son sino la última extrapolación. Siendo realistas, ya no podemos seguir pensando en la especie humana sin máquinas (...) este cambio en nuestra conciencia metafísica esta trascendencia de la cuarta discontinuidad, es esencial para llegar a un acuerdo armonioso con el mundo industrializado. Las alternativas son o un rechazo atemorizado de los Frankensteins que hemos creado o una creencia ciega en sus «virtudes superhumanas» y una fe conmovedora en que ellos pueden resolver todos nuestros problemas humanos.



3. La Singularidad está cerca – Ray Kurzweil (2005)

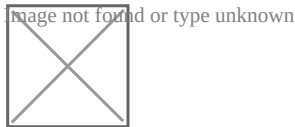
[Raymond Kurzweil](#) es quizás uno de los personajes más estrambóticos del mundillo digital. Empresario, gurú, inventor, músico, tecnólogo y desde 2012 director de ingeniería en Google, ha dedicado buena parte de su vida a anunciar la llegada de la [Singularidad](#). Según Kurzweil **el crecimiento de las tecnologías de la información no es lineal, sino exponencial**. Esto hace que esas tecnologías penetren cada vez más en diferentes campos científicos, desde la biología a las neurociencias; este proceso **se acelera cada vez más** y nos lleva sin escalas a una explosión de las **inteligencias suprahumanas**. ¿Qué entiende Kurzweil por «**inteligencias suprahumanas**»? Pues una **IA fuerte o general**, una especie de superinteligencia tipo The Matrix o Skynet pero sin connotaciones apocalípticas. Kurzweil denomina a ese momento [Singularidad](#).

Ray Kurzweil es uno de los tecnopensadores más radicales y optimistas de su generación. Su visión expresa como pocos el paradigma de la [Californian ideology](#), ese **pensamiento utópico nacido en la costa Oeste entre tablas de surf, restos de LSD, rock'n'roll y microprocesadores**. Pero volvamos a la Singularidad:

¿Qué tan singular es la Singularidad? ¿Sucederá en un instante? Consideremos nuevamente la derivación de la palabra. En matemáticas, una

singularidad es un valor que está más allá de cualquier límite: en esencia, el infinito (...) La Singularidad, como la hemos discutido en este libro, no alcanza niveles infinitos de computación, memoria o cualquier otro atributo mensurable. Pero ciertamente logra vastos niveles de todas estas cualidades, incluyendo la inteligencia. Con la ingeniería inversa del cerebro humano podremos aplicar los algoritmos paralelos, autoorganizados y caóticos de la inteligencia humana a sustratos computacionales muy potentes. Esta inteligencia estará entonces en condiciones de mejorar su propio diseño, tanto de hardware como de software, en un proceso iterativo que se acelerará rápidamente.

Según Kurzweil **la Singularidad llegará entre 2025 y 2030**. Atención! Hay que tomar muy en serio las predicciones de este hombre: en su libro [*The Age of Intelligent Machines*](#) (1990) predijo que antes del fin de siglo las máquinas superarían al mejor jugador humano de ajedrez. En ese momento unos cuantos se rieron, pero tuvieron que tragarse las sonrisitas cuando **en 1997 la computadora de IBM [Deep Blue](#) venció al campeón mundial Garry Kasparov**.



4. Superpotencias de la inteligencia artificial. China, Silicon Valley y el nuevo orden mundial – Kai-fu Lee (2018)

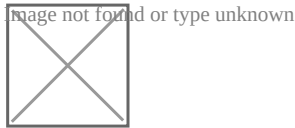
Este libro de [Kai-fu Lee](#) comienza en 2017, cuando otra IA ([AlphaGo de Google](#)) venció a **Ke Jie**, el imbatible campeón de Go. Este juego, inventado hace 2.500 años y mucho más complejo que el ajedrez, parecía una montaña imposible de escalar que se levantaba en el medio del camino de las IA. Y sin embargo... Según Kai-fu Lee ese fue el «momento Sputnik» de China; **al igual que Estados Unidos frente a la conquista rusa del espacio con el satélite Sputnik, la potencia oriental decidió fijar claros objetivos estratégicos para acelerar el desarrollo de sus IA**.

Menos de dos meses después de que Ke Jie abandonara la última partida con AlphaGo, el gobierno central chino elaboró un ambicioso plan para desarrollar las capacidades de la inteligencia artificial. Instó a una mayor financiación, apoyo político y coordinación nacional para el avance de la IA. Estableció claros

parámetros de evaluación para medir el progreso entre 2020 y 2025, y previó que para 2030 China se convertiría en el centro de la innovación global en inteligencia artificial, liderando en teoría, en tecnología y en sus aplicaciones. En 2017, los inversores de capital riesgo chinos ya habían respondido a esa llamada, invirtiendo sumas récord en nuevas empresas de inteligencia artificial, que representaba el 48 por ciento de toda la financiación de capital riesgo en IA en todo el mundo, superando por primera vez a Estados Unidos.

Detrás de ese incremento del apoyo del gobierno chino hay un nuevo paradigma en la relación entre la inteligencia artificial y la economía. Mientras que la ciencia de la inteligencia artificial ha progresado despacio pero de manera constante durante décadas, sólo en fecha reciente se ha acelerado rápidamente dicho progreso, lo que ha permitido que estos logros académicos se traduzcan en casos de uso en el mundo real.

Después de leer este libro, **las utópicas predicciones de Ray Kurzweil suenan mucho más creíbles.**



5. Human Compatible: Artificial Intelligence and the Problem of Control – Stuart Russell (2019)

[Stuart Russell](#), autor junto a [Peter Norvig](#) del superclásico texto académico [Artificial Intelligence: A Modern Approach](#) -bibliografía obligatoria en más de 1.500 instituciones-, volvió al lugar del delito en 2019 con [Human Compatible: Artificial Intelligence and the Problem of Control](#). Russell sostiene que el modelo estándar de investigación, enfocado en lograr rígidos objetivos específicos, es peligroso y podría no reflejar los valores humanos, **lo cual sería catastrófico si las IA alcanzaran la superinteligencia** anunciada por Kurzweil. Además, propone investigar las cuestiones vinculadas a la seguridad dada la incertidumbre del tiempo requerido para desarrollar una IA a nivel humano o superior, y las presiones económicas que impulsan y aceleran esa innovación. Escribe Russell

Encontrar una solución al problema del control de la IA es una tarea importante; tal vez, en palabras de Bostrom, sea «la tarea esencial de nuestra época». Implica construir sistemas que sean mucho más poderosos que nosotros y al mismo tiempo

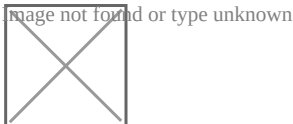
garantizar que esos sistemas seguirán siendo impotentes para siempre. Hasta ahora, la investigación en IA se ha centrado en sistemas que son buenos para tomar decisiones, pero esto no es lo mismo que tomar las mejores decisiones. No importa cuánto se maximice un algoritmo hasta alcanzar su excelencia, y no importa cuán preciso sea su modelo del mundo: las decisiones de una máquina pueden ser inefablemente estúpidas a los ojos de un ser humano común y corriente si no logra comprender sus preferencias.

Russell, quizás inspirado por **Isaac Asimov** y sus [leyes de la robótica](#), propone tres principios que deberían guiar el desarrollo de las IA:

- 1. El único objetivo de la máquina es maximizar la realización de las preferencias humanas.*
- 2. Inicialmente, la máquina no está segura de cuáles son esas preferencias.*
- 3. La fuente fundamental de información sobre las preferencias humanas es el comportamiento humano.*

Russell concluye abogando por una **gobernanza más estricta en la investigación y desarrollo** de estas tecnologías y una introspección en clave cultural sobre la **autonomía** que deberíamos otorgar a las IA.

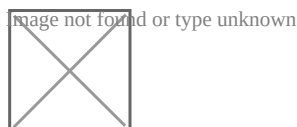
Imagen superior: [exposición sobre IA en el CCCB](#).



6. Artificial Intelligence: A Guide for Thinking Humans – Melanie Mitchell (2019)

Este libro de [Melanie Mitchell](#), investigadora de la Portland State University y del mítico [Santa Fe Institute](#), es una de las mejores introducciones a las IA que he leído. [Artificial Intelligence: A Guide for Thinking Humans](#) desgana paso a paso el recorrido que llevó a las actuales IA, desde los primeros balbuceos en la posguerra hasta las máquinas generativas basadas en las redes neuronales. Después de esta minuciosa reconstrucción del pasado y presente de las IA, Mitchell se permite reflexionar sobre su futuro:

Me parece extremadamente improbable que las máquinas puedan alguna vez alcanzar el nivel de los humanos en traducción, comprensión lectora y similares exclusivamente a partir de datos en línea, sin esencialmente ninguna comprensión real del lenguaje que procesan. El lenguaje se basa en el conocimiento del sentido común y la comprensión del mundo... La falta de comprensión humana, ¿lleva a que sean frágiles, poco confiables y vulnerables a los ataques? Nadie sabe la respuesta y este hecho debería hacernos reflexionar a todos.



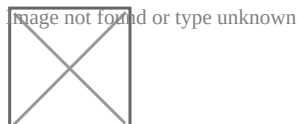
7. Programados para crear. Cómo está aprendiendo a escribir, pintar y pensar la inteligencia artificial (2019)

El libro de [Marcus du Sautoy](#), uno de los mejores divulgadores científicos de este siglo, apunta a uno de los temas clave del debate sobre las IA: **¿pueden crear?**

El desafío de este libro es llevar a la nueva inteligencia artificial hasta sus límites para ver si puede igualar o incluso superar las maravillas de la programación humana. ¿Puede una máquina pintar, componer música o escribir una novela? Quizá no pueda competir con Mozart, Shakespeare o Picasso, pero ¿podría ser tan creativa como nuestros hijos cuando escriben un cuento o hacen un dibujo? Al interactuar con el arte que nos conmueve y comprender lo que lo distingue de lo trillado y repetitivo, ¿podría una máquina aprender a ser creativa? Y no sólo eso, ¿podría ampliar nuestra propia creatividad y ayudarnos a expresar nuestras capacidades? (...) La inteligencia artificial supone un desafío fenomenal, pues está dejando claro que muchas de las tareas que realizan los seres humanos las pueden realizar igual de bien, si no mejor, las máquinas. Pero en lugar de centrarse en un

futuro de vehículos autónomos y medicina computarizada, este libro se propone explorar si estos algoritmos pueden competir significativamente con el poder de la programación humana.

En un mundo donde proliferan las IA generativas en el campo textual, gráfico y sonoro, [Programados para crear](#) nos orienta y guía por una serie de experiencias increíbles que, al escribir estas líneas, seguramente ya han sido superadas. También **las IA generativas avanzan a ritmo exponencial.**



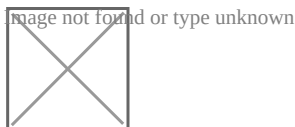
8. El mito de la inteligencia artificial. Por qué las máquinas no pueden pensar como nosotros lo hacemos – Erik Larson (2021)

[Erik Larson](#), experto en procesamiento del lenguaje natural y empresario, argumenta en [El mito de la inteligencia artificial](#) que la idea de que estas tecnologías superarán pronto a la inteligencia humana es errónea. Subraya que las actuales IA, basadas en el razonamiento inductivo y procesamiento de datos, difieren significativamente de la inteligencia humana, la cual utiliza el **razonamiento abductivo**, un aspecto aún **no replicable en las máquinas** (ver mi artículo «[El mal que aqueja a las IA es la abducción](#)«).

Como explica en una entrevista en [La Vanguardia](#), Larson desconfía de los planteos disruptivos de Ray Kurzweil and Co.:

Soy muy escéptico con aquellos que dicen que la IA logrará una comprensión real, como lo hacen las mentes humanas. ¿Será algún día posible? Por ahora es algo que desconocemos. A día de hoy no tenemos ni idea de cómo programar determinados aspectos y esa es la barrera más grande (...) Los medios y el público tienen la impresión de que estamos construyendo máquinas sobrehumanas que terminarán con nosotros pero en realidad estamos creando programas para realizar tareas limitadas como jugar a juegos o traducir idiomas. Nunca construimos sistemas que tengan sentido común o inteligencia general. Es más, es algo retorcido pensar que los chips de silicio pueden generar mentes humanas.

Larson critica las exageraciones en torno a las IA, señalando que esta actitud no solo es científicamente inexacta, sino también **perjudicial para el avance científico**. En lugar de enfocarse en el potencial exagerado de las IA existentes, sugiere que **debemos valorar y entender mejor nuestra propia inteligencia para lograr avances genuinos**.



9. Atlas de inteligencia artificial. Poder, política y costes planetarios de la inteligencia artificial – Kate Crawford (2021)

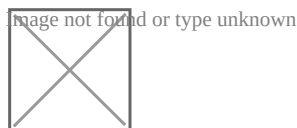
En este volumen, casi un clásico contemporáneo, [Kate Crawford](#) revela cómo las redes globales que sustentan la tecnología de las IA está **dañando el medio ambiente**, afianzando la desigualdad y fomentando un tipo de **gobernanza no democrática**. En un viaje que va desde las minas hasta las fábricas de dispositivos y granjas de servidores, la autora nos **presenta en clave materialista y ecológica un panorama de los dispositivos y materiales necesarios para hacer funcionar las IA**.

En una entrevista en [MIT Technology Review](#) Crawford explicaba la lógica detrás de su libro:

Muchos de los libros que se han escrito sobre la inteligencia artificial en realidad solo hablan de unos logros técnicos muy limitados. Y a veces describen a los grandes hombres de la IA, pero eso es realmente todo lo que hemos tenido para hacer frente

a lo que es la inteligencia artificial (...) Yo quería hacer algo muy diferente: he querido entender cómo se crea la inteligencia artificial en el sentido más amplio. Esto significa analizar los recursos naturales que la impulsan, la energía que consume, el trabajo oculto de la cadena de suministro y la gran cantidad de datos que se extraen de cada plataforma y dispositivo que usamos todos los días (...) De esa forma, quería abrir las puertas a la idea de que la IA que no es ni artificial ni inteligente. Es lo opuesto a lo artificial. Proviene de las partes más materiales de la corteza terrestre y del trabajo humano, y de todos los artefactos que producimos, decimos y fotografiamos todos los días. Tampoco es inteligente. Creo que existe este gran pecado original en este campo, donde la gente asumió que los ordenadores eran de alguna manera como los cerebros humanos y si los entrenábamos como si fueran niños, poco a poco se convertirían en estos seres sobrenaturales.

Mientras que los sistemas técnicos presentan una apariencia de objetividad y neutralidad, Crawford muestra cómo están **diseñados para servir e intensificar los sistemas de poder existentes**. Basándose en una más de una década de investigación, la autora demuestra **cómo las nuevas infraestructuras detrás de las IA reflejan las creencias y perspectivas de un pequeño grupo de personas - la Californian ideology de nuevo...- y sirven a los intereses de unos pocos en detrimento de muchos**.



10. Tecnohumanismo. Por un diseño narrativo y estético de la inteligencia artificial – Pablo Sanguinetti (2023)

Entre tanta bibliografía sobre IA publicada en estos meses se destaca un pequeño gran volumen de **Pablo Sanguinetti**: [*Tecnohumanismo. Por un diseño narrativo y estético de la inteligencia artificial*](#). Breve y exquisitamente escrito -me recuerda a la cuidada prosa de **Byung-Chul Han**-, *Tecnohumanismo* se aleja del popular [*lloriqueo tanguero del filósofo y ensayista surcoreano*](#). Lejos de cargar contra las IA desde una mirada apocalíptica, Sanguinetti nos propone una **reflexión en clave artística y humanística**

para afrontar una tecnología que considera tan rompedora como la máquina a vapor o Internet. Ya desde el primer párrafo Sanguinetti marca el territorio y el estilo del libro:

La tecnología y lo humano se construyen mutuamente. Nuestra forma de estar en el mundo y de entender quiénes somos viene mediada por artificios técnicos. No existe humano sin tecnología. Ni existe revolución tecnológica que deje intacto el modo de ser humano (...) Lo que comenzó como logro técnico es cada vez más un asunto ético, filosófico y estético (...) Necesitamos que la nueva realidad del algoritmo encaje de forma natural y con sentido en el mundo humano. Y las herramientas para lograrlo provienen de las artes y las humanidades.

Sanguinetti entra al campo de las IA desde lo narrativo –**¿Cómo las ponemos en discurso a las IA? ¿Qué decimos de ellas?**– y reivindica la **intervención artística**: «Solo a través del arte entenderemos a la máquina, y solo a través del arte la máquina nos entenderá a nosotros». La idea detrás del libro se puede resumir en una frase: **«La inteligencia artificial seguirá incompleta mientras no se la recubra de un diseño narrativo y estético que le dé sentido y la inserte en nuestra cultura»**.

Sanguinetti afronta desde una perspectiva muy fresca y humanista cuestiones urgentes como la **explosión textual** que están produciendo las IA generativas o la **educación** y los **objetivos** (¿o debemos hablar de **deseos**?) de las máquinas digitales. El autor propone «inculcar la **incertidumbre** en la máquina y reservarnos la última palabra». Para eso, nada mejor que operar desde lo artístico, ya que **«el arte es una escuela de incertidumbre»**. Y reflexiona: «ha tenido que ser la tecnología la que nos demuestre que **la belleza no resulta un lujo accesorio**, sino el código de lo que somos, el lenguaje con el que está escrita nuestra identidad».

A lo largo de las páginas de *Tecnohumanismo* Sanguinetti propone varias intervenciones discursivas que considero esenciales, por ejemplo la necesidad de **desantropomorfizar a las IA**. Hacer que estas máquinas digitales dejen de ser puestas en discurso con características y cualidades humanas **debería ser una prioridad de cualquier tipo de acercamiento crítico** al tema. ¿Seremos capaces? Me parece que **la desantropomorfización de las IA es una batalla a estas alturas imposible de ganar**, pero no debemos bajar las banderas.

Final abierto

Mis [10 tesis \(mutantes\) sobre las inteligencias artificiales](#) comienzan con una reflexión en clave textual: **Tesis nº 1 – Todo lo que escribimos o decimos sobre las IA ya es viejo.** ¿Podemos decir lo mismo de todo lo que leemos sobre las IA? Espero que no. Si bien este campo se está desarrollando a velocidad exponencial, seguramente algunos textos -como el clásico de Alan Turing- permanecerán en la bibliografía obligatoria por mucho tiempo, mientras que otros se irán cayendo para ser sustituidos por obras que planteen otras preguntas o arriesguen nuevas respuestas. Es en este contexto que debe leerse esta lista de 10 textos sobre inteligencias artificiales.

Buena lectura!

[LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ](#)

Fotografía: hipermediaciones

Fecha de creación

2024/01/04