

Pensamiento humano e IA desde una perspectiva filosófica. Exposición en Fundación Ética Visionaria (FEV) 24/01/2026

Esta exposición se origina en un artículo publicado por el periódico Le Monde el 13 de noviembre de 2025, y que me pareció interesante porque presenta un enfoque diferente del habitual, ya que la mayoría se ocupa de analizar la cuestión financiera (si las inversiones en la IA son un “burbuja” o no). O del impacto ambiental que implican los equipos necesarios para su implementación (la administración Trump declaró emergencia energética en EU por este motivo). O del ingente consumo de agua para enfriamiento de los equipos. O del evidente impacto en el empleo (ya que las compañías sustituirán personal en muchas tareas que pueda realizar la IA, cuestión que ya está sucediendo, en varios campos, por ej. en empresas como Amazon, call centers, etc.). O por la creciente influencia en la gobernanza del mundo de las “Siete Magníficas” (Alphabet, Amazon, Apple, Meta, Microsoft, Nvidia y Tesla). En otras publicaciones se habla de Gafam, acrónimo para las cinco tecnológicas estadounidenses: Google (Alphabet), Apple, Facebook (Meta), Amazon y Microsoft, que, según dice el buscador de Google, “dominan la tecnología, el comercio electrónico, las redes sociales y el software, destacando por su recopilación de datos y su gran capitalización en el mercado. Este término agrupa a las empresas que tienen un enorme poder e influencia global, moldeando la vida moderna a través de sus omnipresentes servicios y productos.” Evidentemente es un amplio campo que va desde la política, hasta la ética. Sin embargo, en esta oportunidad, trataré de aportar alguna luz desde el punto de vista filosófico.

El título del artículo al que he hecho referencia al principio plantea, de por sí, un cuestionamiento de carácter filosófico: ***“Al delegar la escritura a la IA, corremos el riesgo de impedir nuestra reflexión”*** (Corentin Lamy¹, Le Monde, 13/11/2025).

El análisis filosófico de una temática no se queda en el plano social, ni técnico, ni histórico, sino que trata de escudriñar los fundamentos últimos del problema. Porque, como veremos, o trataré de mostrar, la IA cuestiona la propia naturaleza del ser humano. ¿Por qué?

¹ Chef de service adjoint chez Le Monde

Desde la antigüedad se consideraba que una de las características más esenciales del ser humano era su razón, su inteligencia, su pensamiento; y esto nos distinguía de los demás animales.

Sin embargo, desde hace un buen tiempo se habla, persistentemente, desde la terminología de la mercadotecnia, de objetos, de útiles, inteligentes: existen hoy, lavarropas, televisores, teléfonos, coches, casas, ciudades, rutas, herramientas inteligentes... Pareciera haber un curioso desplazamiento de la inteligencia humana, que, por supuesto, maneja una cantidad relativamente limitada de datos en comparación con las Ias; pero que los procesa de un modo absolutamente diferente.

Se hace necesario, sin embargo, distinguir, entre varios tipos de IA, no todas tienen el mismo acceso a fuentes de datos específicos y, por otra parte ya desde hace mucho, todos los usuarios de internet, somos también usuarios de algún tipo de inteligencia artificial, todos los motores de búsqueda lo son. Pero lo que ha producido esta supuesta revolución, son los chatbot (ChatGpt, Luzia, chatbot de Bing, Tidio, Zendesk, etc.) porque ha puesto en las manos del público en general un artilugio que puede generar textos sencillos como los que corresponden a informes rutinarios, y respuesta a ciertas órdenes por reconocimiento de voz, manifestando una especie de magia electrónica que aviva nuestro espíritu infantil. Pero también de reconocimiento facial y otros usados por policías y ejércitos para control social y la guerra. No habría inversiones tan multimillonarias en este campo, simplemente por el interés de poner a disposición de usuarios particulares “gadgets” de diversión o trabajo de emprendedores o mipymes. Está en juego mucho más, el dominio geopolítico del mundo, mediante el uso de este tipo de tecnología. También desde un punto de vista mental, en el manejo de grandes masas como la Pandemia Covid-19.

Nuestro tema es más bien otro, y es el de establecer diferencias entre la inteligencia humana y las Ias.

En la síntesis inicial de su artículo, Corentin Lamy dice:

*Escribir no nos permite solamente comunicarnos, sino que **nos ayuda a organizar nuestras ideas**. Recurrir a generadores de texto basados en inteligencia artificial equivale, por lo tanto, a dejar que los multimillonarios de Silicon Valley **piensen por nosotros**. [lo destacado es mío]. [...]*

Luego comentaré la referencia, que no es poco importante, sobre los “multimillonarios del Silicon Valley”, por ahora me remito a la esencia de lo que quiero analizar.

*Pero escribir no es solamente **comunicar**, es también **pensar**. En una reciente entrevista en la revista *Usbek & Rica*, el filósofo Eric Sadin lamenta que «mil*

millones de individuos» encuentran en estas tecnologías «la ocasión de no ejercer más sus facultades fundamentales, entre las que destacan las de hablar y escribir en primera persona». Y continúa: « ¿Entendemos que una vida privada de la expresión de nuestras facultades y de vínculos activos con nuestros semejantes solo puede ser caldo de cultivo para la tristeza, el rencor y la locura?» [el destacado es mío]

*«Un escritor no escribe sólo palabras, agrega Ed Zitron, (autor norteamericano especialista en inteligencia artificial, en su Acusación contra la IA generativa); hace **entrechocar las ideas, ideales, emociones, reflexiones, hechos y sentimientos** (...) La buena escritura es una tensión (...) un proceso atravesado por la emoción...una emoción que una IA no sabría replicar» [el destacado es mío][...]*

***Más generalmente, poner orden en un texto, es poner el orden de su pensamiento.** [el destacado es mío] [...]*

Comencemos por el **pensar**:

Escribir, sobre todo, significa pensar. Y éste ha sido, desde el principio de su historia el tema fundamental de la filosofía: ¿Qué significa pensar? Es, asimismo, el título de un libro de Martin Heidegger (*Was heisst Denken*). Libro laborioso y profundo, resultado de clases impartidas por Heidegger en 1951-1952 en la Universidad de Friburgo de Brisgovia.

¿Qué significa pensar?, repite una y otra vez Heidegger en sus clases magistrales de la Universidad de Friburgo, porque según él, históricamente, se ha pensado muy poco.

La memoria es la aglutinante del pensamiento.(Heidegger, 2005), (p. 22)

En esta única cita a la que he recurrido del texto de Heidegger, el autor, relaciona *memoria* con *pensamiento*. Aprovecho esta relación para hacer notar que mientras la memoria digital es un sistema electrónico de almacenamiento de datos, la humana no es un reservorio pasivo, sino una estructura dinámica que, como dice Heidegger, permite *aglutinar* el pensamiento. Y, por lo tanto, allí convergen cuerpo y “alma”.

Evidentemente, en estos últimos años nos han invadido una serie de metáforas que son confusas y que además, falsean la naturaleza de las cosas. Pensamos que el cerebro funciona como un ordenador que recibe información (input) y que expelle ciertos resultados de diverso tipo (output); pero no lo concebimos como parte esencial de un sistema complejo; pero vivo.

Por eso dice muy claramente Corentine Lamy:

¿Y en qué piensa el ChatGPT? Felizmente, en nada: ya se ha dicho, el programa se contenta con regurgitar los textos que figuran en su corpus de entrenamiento, desarrollando variantes según un modelo probabilístico. Los generadores de texto producen frases como la montaña torrentes, sin tener conciencia a dónde van, de lo que acarrearán. Ni siquiera de que ellos existen.

Con la memoria humana entramos directamente a la intersección “*cuerpo-alma*” para decirlo en los términos de Aristóteles, en el Tratado del alma:

”el alma es la primera entelequia del cuerpo físico que posee la vida en potencia”. Aristóteles pensaba el mundo en general en una forma dinámica, como *potencia y acto*. Potencia significa que algo no está todavía realizado (en la semilla está en potencia el árbol), y el acto es la realización, en este caso lo expresa como entelequia, como perfección, ya que para él la realización de la potencia es una forma de perfección.

Desde la época de Aristóteles (384-322 AC) hasta hoy se ha ido abandonando paulatinamente este concepto de alma (en griego, ψυχή (psujé)). Sin embargo todavía perdura en palabras como psicología y otras relacionadas.

Dejemos por ahora el concepto de alma para hacer la distinción entre *recuerdo y memoria*. Ya Platón (428-347 AC) hacía la distinción entre μνήμη (mneme) (memoria) y ἀνάμνησις (anánmnesis) (recuerdo), la memoria sería la facultad de atesorar recuerdos.

Quiero citar el final de un cuento de Jorge Luis Borges, que se llama Funes el memorioso, del libro Artificios (1944). (Borges, 1974)

Se trata de un hombre de campo, un paisano, un gaucho, que tiene un accidente, se cae de su caballo; y este accidente tiene una consecuencia extraordinaria y a la vez terrible: surge en él una memoria absoluta, recuerda todo. Y todo es todo, hasta el más mínimo detalle de cada pliegue de una cortina. Al final del cuento, a modo de epílogo, dice Borges:

*Había aprendido sin esfuerzo el inglés, el francés, el portugués, el latín. Sospecho, sin embargo, que no era muy capaz de **pensar**. Pensar es olvidar diferencias, es generalizar, abstraer. En el abarrotado mundo de Funes no había sino detalles, casi inmediatos.*

Borges señala aquí con toda claridad que una cosa es memorizar, en este caso extremo, **memorizar todo**, y otra muy distinta es **pensar**; pero sin embargo, siguiendo ahora a Heidegger, el pensamiento necesita de la memoria; pero ésta es una memoria humana, transida de recuerdos, de pasión, de olvidos, que, como dice Freud es una forma de recuerdo. Para él el olvido, y el recuerdo, tienen que ver con el subconsciente, también con el *ello*; que, en definitiva son pulsiones corporales como la sexualidad tan puesta de manifiesto en la obra freudiana. Esto no tiene nada que ver con una memoria digital, ni con el “aprendizaje profundo” de las Ias. El pensamiento humano se da en un cuerpo vivo, el cuerpo no es algo aleatorio o circunstancial y no hay unión física del cuerpo y el alma como pensaba Descartes.

Para Henri Bergson (1859-1941) hay una memoria psicofisiológica y otra que es la memoria pura, ésta constituye la propia esencia de la **conciencia**. La entenderíamos aquí, en forma general, como autoconocimiento, como experiencia vivida.

En cuanto a que el sistema nervioso es un sistema corporal: “el perder de vista las raíces orgánicas del sistema nervioso es una de las fuentes mayores de confusión respecto a su operar efectivo” (Maturana, 1997) [falta cita]

Lo vivo (autopoietico, autónomo), lo mecánico (la máquina, heterónoma)

En su libro *El fenómeno de la vida*, Francisco Varela, define con gran precisión el fenómeno autopoietico, que no es más que definir una de las características fundamentales del fenómeno de la vida:

[...] la célula es capaz de automantenimiento gracias a un patrón constitutivo o proceso de generación circular que reemplaza continuamente los componentes que están siendo destruidos y recrea las condiciones para discriminar entre el sí mismo y el no-sí mismo. Claramente, esta pauta de organización básica podría presentarse en diferentes tipos de unidades, no exclusivamente en la célula básica tal como la conocemos. Esta organización mínima es lo que con Maturana denominamos una unidad autopoética, la que se define más precisamente del siguiente modo: un sistema autopoietico está organizado (esto es, se define como una unidad) como una red de procesos de producción (de síntesis y destrucción) de componentes, en forma tal que estos componentes: (i) se regeneran continuamente e integran la red de transformaciones que los produjo, y (ii) constituyen al sistema como una unidad distinguible en su dominio de existencia. (Varela, 2002)(p30)

La lengua

Para Humberto Maturana² las “palabras son acciones, no son cosas que se pasan de aquí para allá. Es nuestra historia de interacciones recurrentes la que nos permite acoplamiento interpersonal efectivo, y encontrar un mundo que estamos especificando en conjunto a través de nuestras acciones” (Maturana H. V., 1997)(p.154)

Maturana está considerando la evolución de los homínidos desde hace miles de años, en su evolución filogenética y en la interacción en la que emerge el ser humano.

[...] la aparición del ser humano y de todo el contexto social en el que aparece, genera este fenómeno inédito –hasta donde sabemos– de lo mental y de la conciencia de sí como la experiencia más íntima del ser humano. (Maturana H. V., 1997)(p.154)

*Nos realizamos en un mutuo acoplamiento lingüístico, no porque el lenguaje nos permita decir lo que somos, sino porque **somos** en el lenguaje, en un continuo ser en los mundos lingüísticos y semánticos que traemos a la mano con otros. (Maturana H. V., 1997) (p155)*

² Doctor en biología (PhD) de la Universidad de Harvard y profesor titular de la Universidad de Chile.

En otras palabras, el lenguaje es verdaderamente esencial a la condición humana y además, según la concepción evolutiva de Maturana, el lenguaje ha sido condición de su evolución, mental y social. Por supuesto que con esto no clausuro de ninguna manera toda la complejidad del lenguaje que no podría ser expresada ni remotamente en un espacio tan breve de tiempo. Sólo quería exponer un pensamiento bastante peculiar del lenguaje, aunque emparentado, sobre todo por Varela³ con el pensamiento fenomenológico de Heidegger y Merlau-Ponty.

Y precisamente en este punto surgen varias preguntas cruciales: ¿los desarrollos de las IAs, por ahora embrionarios, podrán cruzar la barrera de la complejidad del cerebro humano y el lenguaje? [según Maturana y Varela en el libro citado hay más de “10¹⁰ o quizás más de 10¹¹ neuronas (decenas de miles de millones), y que cada una de ellas reciben múltiples contactos de otras neuronas y se conecta a su vez con muchas células, la combinatoria de posibles interacciones es más que astronómica”(p.105)] ¿Se trata de una “cuestión de tiempo” o de imposibilidad ontológica, es decir, una imposibilidad intrínseca de la no-vivo para acceder al *pensamiento*? ¿Podrían las “redes neuronales artificiales” (otra de las metáforas engañosas del lenguaje cibernético) tener la complejidad plástica del sistema nervioso humano, o sólo simulan esa complejidad, o sentimientos, por puro entrenamiento imitativo con output estadístico?

Aprendizaje profundo (Deep learning) y sesgos

Comienzo con una pregunta: ¿Podrá compararse este “aprendizaje” (quizás sería mejor hablar de *entrenamiento*) con el aprendizaje que desarrolla un bebé desde cero a cien años (ya que el aprendizaje es permanente, aunque los primeros años de vida sean decisivos)?

Veamos sólo algunas cosas importantes que dice Harari en *Nexus*, cuyo subtítulo es: *Una breve historia de las redes de información desde la Edad de Piedra a la IA*:

La IA no es un autómatas tonto que repite los mismos movimientos una y otra vez independientemente de los resultados. Más bien, está equipada con potentes mecanismos de autocorrección que le permiten aprender de sus propios errores [diríamos que aprende por “ensayo y error”, pero este “ensayo-error no es tampoco del mundo “real” sino el de los algoritmos que confrontan millones de datos] (Harari, 2024) (p.324-325)

Y continúa:

³ Francisco J. Varela García. Doctor en Biología (PhD) de la Universidad de Harvard. Profesor titular del CNRS, París.

*Esto significa que la IA comienza su vida como un "algoritmo bebé" con mucho potencial y capacidad de cálculo, pero que en realidad no sabe mucho. Sus padres humanos solo le otorgan la capacidad de aprender y acceder a un mundo de datos. Luego, dejan que el algoritmo bebé **explore el mundo [?] [¿cómo podría hacerlo?]. Como recién nacidos orgánicos [?]**, los algoritmos bebés aprenden **detectando patrones en los datos a los que tienen acceso**. Si toco fuego, me duele. Si lloro, mamá viene. Si sacrifico una reina por un peón, probablemente pierda la partida. Al encontrar patrones en los datos, el algoritmo bebé aprende más, incluyendo muchas cosas que sus padres humanos desconocen (Harari, 2024) (p.325)*

Con todo respeto por el esfuerzo de Harari y su prestigioso nombre como creador de voluminosos best sellers, diría yo, que el aprendizaje humano difiere absolutamente del entrenamiento de una IA. Traigo aquí a Jean Piaget, el epistemólogo y psicólogo genético suizo, quien estableció, a partir de la evolución de sus propios hijos y la observación clínica distintas etapas del aprendizaje de un niño.

El período que se extiende entre el nacimiento y la adquisición del lenguaje está marcado por un extraordinario desarrollo mental.[...]este período consiste en una conquista, mediante las percepciones y los movimientos, de todo el universo práctico que rodea al niño. Pero esta «asimilación sensorio-motriz» del mundo exterior inmediato lleva a cabo, de hecho, toda una revolución copernicana en miniatura: mientras que en el punto de partida de este desarrollo el recién nacido lo refiere todo a sí mismo, o , más concretamente a su propio cuerpo, en la meta, o sea cuando se inicia el lenguaje y el pensamiento, el niño se sitúa ya prácticamente, como elemento o cuerpo entre los demás, en el universo que él ha construido paulatinamente y que siente exterior a sí mismo. (Piaget, 1978)(ps.17-18)

Me parece que con este testimonio de Piaget contestamos al planteo de Harari en cuanto a esa comparación excesiva entre el aprendizaje de un bebé y de una IA.

Paso ahora a considerar algo que me ha parecido muy importante en el planteo de Harari en Nexus, y es el de los sesgos. Veamos:

Sin embargo, las bases de datos presentan sesgos. Los algoritmos de clasificación facial estudiados por Joy Buolamwini se entrenaron con conjuntos de datos de fotos etiquetadas en línea, como la base de datos

Labeled Faces in the Wild. Las fotos de esa base de datos se tomaron principalmente de artículos de noticias en línea. Dado que los hombres blancos dominan las noticias, el 78 % de las fotos de la base de datos eran de hombres y el 84 % de personas blancas. George W. Bush apareció 530 veces, más del doble que todas las mujeres negras juntas. (Harari, 2024)(p325)[...]

Otra base de datos preparada por una agencia del gobierno estadounidense contenía más del 75 % de hombres, casi el 80 % de piel clara y solo el 4,4 % de mujeres de piel oscura.[...]

*No es de extrañar que los algoritmos **entrenados** con dichos conjuntos de datos fueran excelentes para identificar a hombres blancos, pero pésimos para identificar a mujeres negras. Algo similar le ocurrió al chatbot Tay. Los ingenieros de Microsoft no le incorporaron ningún prejuicio intencional. Pero unas pocas horas de exposición a la información tóxica que circulaba por Twitter convirtieron a la IA en una racista furiosa [...]*

Y la cosa empeora. Para aprender, los algoritmos bebés necesitan algo más, además de acceso a datos. También necesitan un objetivo. Un bebé humano aprende a caminar porque quiere llegar a algún lugar. Un cachorro de león aprende a cazar porque quiere comer. Los algoritmos también deben tener un objetivo para aprender. (Harari, 2024)(p325)

Si no eliminamos el sesgo desde el principio, las computadoras podrían perpetuarlo y magnificarlo.[...]

Aquí viene lo más importante para nuestro análisis:

Pero deshacernos del sesgo algorítmico podría ser tan difícil como deshacernos de nuestros sesgos humanos. Podríamos decidir simplemente descartar el algoritmo sesgado y entrenar un algoritmo completamente nuevo con un nuevo conjunto de datos menos sesgados. Pero ¿dónde podemos encontrar un conjunto de datos totalmente imparciales? (Harari, 2024)(ps327-328)

Habría que ver la cantidad de sesgos acumulados en todos los campos y cómo esto puede incidir en la vida pública. Ya que desde hace tiempo hay una fuerte tendencia a concebir a los Estados y a la política como algo equivalente al gobierno de una empresa. Y en un programa español de análisis político se plantea la cuestión de si se puede privatizar la democracia. En verdad ésta es la idea, transida de transhumanismo del Sr. Tesuzo Matsumoto, ex Ceo de la compañía nipona Saufal, profesor universitario y experto en IA:

“El ser humano, simplemente, no es apto para la política, tiene su ego, tiene deseos, es inestable. Con la IA se puede lograr la “Razón pura”, esto lo sabemos por la filosofía del idealismo alemán. Se trata de saber cómo deberían ser las cosas, debemos programar a la IA para que “todo sea como debe ser”. (transcripción de un video de la DW (Límites éticos para la IA, Youtube) Por el contrario, el ser humano sólo puede ser, nunca alcanzará un estado ideal.

La política nació en Grecia íntimamente relacionada con la ética, así lo podemos ver claramente en Sócrates, Platón y Aristóteles. Sin embargo la democracia liberal pareciera estar tocando a su fin, sobre todo por la conformación de bloques muy poderosos que requieren economías de guerra y un sistema que no admite disensos. Dentro de estos poderes, se encuentran en muy primer plano, las llamadas Siete Magníficas que ya hemos mencionado. Las regulaciones legales son muy relativas, porque no pueden “impedir la innovación”, y las compañías prefieren pagar fuertes sumas en multas que revertir sus intromisiones en los datos de las personas haciendo una verdadera minería psicológica-mercadológica, haciendo perfiles de todo el mundo para control social. Cierro mi participación con una observación de Lamy en el artículo que origina esta presentación, repito la cita:

*Escribir no nos permite solamente comunicarnos, sino que **nos ayuda a organizar nuestras ideas**. Recurrir a generadores de texto basados en inteligencia artificial equivale, por lo tanto, a dejar que los multimillonarios de Silicon Valley **piensen por nosotros**. [lo destacado es mío]. [...]*

Gracias

Bibliografía

- Borges, J. L. (1974). *Obras Completas*. Buenos Aires: Emecé Editores.
- Harari, Y. N. (2024). *Nexus*. New York: Random House.
- Heidegger, M. (2002). *Was heisst Denken?* Frnkfurt am Main: Vittorio Klostermann.
- Heidegger, M. (2005). *¿Qué significa pensar?* Madrid: Trotta .
- Maturana, H. V. (1997). *El árbol del conocimiento*. Santiago: Editorial Universitaria.
- Piaget, J. (1978). *Seis estudios de psicología*. Barcelona: Barral Editores.
- Varela, F. (2002). *El fenómeno de la vida*. Santiago de Chile: Dolmen Ediciones.