

Carlos García Torres

Espacios de pensamiento y libertad frente a sistemas de IA conversacional

RESUMEN: La aparición de sistemas de Inteligencia Artificial conversacionales que aspiran a ser sistemas generales de IA ha traído diversas perplejidades a la ciencia, a la filosofía y al derecho. De todos esos campos posibles de enfoque pretendemos abordar en este artículo solo aquellas que se refieren a la escritura y a su relación con el pensamiento humano. Para este efecto partiremos de diversas reivindicaciones de la escritura, tanto, como acto fundador del pensamiento, y como acto liberador. Nos referiremos fundamentalmente a las ideas de Kant sobre la naturaleza del saber y de Jacques Derrida y sus amplios trabajos sobre la escritura y su relación con el discurso.

PALABRAS CLAVE: Inteligencia artificial; Modelos amplios de lenguaje; LLM; Libertad; Derrida.

Spaces of thought and freedom in front of conversational AI systems

ABSTRACT: The emergence of conversational Artificial Intelligence systems that aspire to be general AI systems has brought various perplexities to science, philosophy and law. Of all these possible fields of focus, we intend to address in this article only those that refer to writing and its relationship with human thought. For this purpose, we will start from various vindications of writing, both as a founding act of thought, and as a liberating act. We will basically refer to Kant's ideas on the nature of knowledge and Jacques Derrida and his extensive works on writing and its relationship with discourse.

KEYWORDS: Artificial Intelligence; Large Language Models; LLM; Freedom; Derrida.

Pensamiento y relaciones

La afirmación de Kant sobre saber y relacionar constituye parte muy importante de su argumentación en la *Crítica de la Razón Pura*. Como es bien sabido su esfuerzo por unir las formas exactas de la razón con las formas variables de la experiencia, es decir, por encontrar los modos mediante los cuales se encuentre

► **Carlos García-Torres**, Departamento de Ciencias Jurídicas. Cátedra UNESCO de Ética y Sociedad en la Educación Superior. Universidad Técnica Particular de Loja, Ecuador. **Autor de correspondencia:**  cegarcia@utpl.edu.ec –  <http://orcid.org/0000-0003-1170-6765>.

el conocimiento que la humanidad necesita, lo llevó a la vasta creación de varias nociones entre las cuales tienen especial importancia las «categorías» y los «principios sintéticos». Entre las categorías se encuentra la de relación que se refieren a «inherencia y subsistencia (*substantia et accidens*, causalidad y dependencia (causa y efecto), comunidad (acción recíproca entre el agente y el paciente)» (Kant I. 1928: I, 212), de forma que en la visión kantiana el conocimiento se construye sobre relaciones entre los fenómenos y de los fenómenos con las formas puras de la razón. Tal es (dicho muy de prisa) el camino por el cual obtenemos conocimientos humanos válidos. También aquellos principios sintéticos llevan a la idea de relación sobre todo en la parte que habla de las analogías de la experiencia, aquí Kant señalaba en la primera edición de la *Crítica* que «Todos los fenómenos están, según su existencia, *a priori* bajo reglas de las determinaciones de su relación unos con otros en un tiempo» lo cual cambió en la segunda edición diciendo: «la experiencia es posible solo mediante la representación de un enlace necesario de las percepciones» (Kant I., 1928, II,39). De cualquiera de las dos formas se entiende la importancia de establecer las relaciones entre los fenómenos sensibles en el pensamiento de este filósofo.

Partimos entonces de esta afirmación esencial, saber es relacionar. De forma que relacionar informaciones es una actividad intelectual esencial que, hasta hace poco, estaba reservada a los seres humanos. Los avances informáticos han derivado estas acciones, al menos de una forma básica, a aplicaciones informáticas que pueden establecer relaciones entre datos que se encuentran en la internet o en determinadas bases de datos y que nos permiten encontrar respuestas a diversas inquietudes. Las aplicaciones de inteligencia artificial referidas a modelos amplios de lenguaje «*Large Language Model*» (en adelante LLM) permiten que exista una interacción conversacional entre el usuario y el sistema de IA desde la cual se pueden obtener diversos textos escritos en los cuales se puede observar relaciones entre ideas y hechos que dan la impresión de textos realizados por seres humanos. Textos de este tipo se vienen utilizando ampliamente en ámbitos laborales y académicos. Como es natural la posibilidad de deferir la escritura de textos universitarios a la IA ha generado gran inquietud alrededor del mundo, más aún cuando diversos investigadores han admitido abiertamente que utilizaron esta herramienta para producir artículos publicados en revistas indexadas. Ahora bien, las relaciones que realizan las máquinas no

pueden entenderse como relaciones originales porque se basan siempre en constructos intelectuales que ya han sido hechos previamente y que se encuentran algún lugar de la internet, esto quiere decir que se está utilizando pensamiento humano ajeno para presentarlo como propio.

Si volvemos al punto de vista kantiano y a la idea general de que saber es relacionar encontraremos que no se crea conocimiento nuevo a partir de relaciones viejas y que sólo se trata de mero reciclaje de ideas que se presentan como originales. Es decir un texto de un programa de LLM sólo da la apariencia de pensamiento y no se trata, por tanto, de ninguna elaboración intelectual propia sino de la expresión variada de algo que ya se dijo con anterioridad. En definitiva, se trata de mero parafraseo. Por tanto, en sentido riguroso no puede hablarse de escritura cuándo se obtiene un texto por estos métodos.

Pensamiento y escritura

Conviene ahora referirse a la posibilidad de los sistemas de LLM de reemplazar las funciones humanas de lectura y escritura. En diversos espacios de las redes sociales se pueden ver reseñas de personas vinculadas al sector empresarial y al sector académico en las que se ponderan las capacidades de estos sistemas para resumir grandes cantidades de información y ofrecer productos coherentes e instructivos, así como para redactar textos con calidad académica suficiente como para merecer publicación. Sobre estos puntos tan delicados que atañen al acto mismo de leer y escribir (ejercicios intelectuales por excelencia) se han escrito numerosos textos dado que constituyen experiencias intensamente personales. Por esta razón los análisis que en su momento hicieron Jacques Derrida, Maurice Merleau-Ponty, Jean Paul Sartre, entre otros muchos, son relevantes para contrastar la posibilidad, la necesidad, la urgencia de ceder tales facultades a un sistema de IA. Por ahora nos limitaremos a una revisión somera de las ideas de Derrida. Como es sabido Derrida nos habla de diversas oposiciones con las que construimos nuestra interpretación del mundo. En estas oposiciones hay siempre un ganador y un excluido, por ejemplo, en la oposición luz y oscuridad tiene preeminencia la luz y la oscuridad se considera como algo no deseable; lo propio sucede con todas las oposiciones que nuestro bagaje cultural nos permite imaginar. Desenterrar estas oposiciones y dar el lugar que le corresponde al elemento excluido es una forma de deconstrucción. En la

oposición entre habla y escritura resulta claro que, a través de la historia, la escritura siempre ha sido un elemento marginal que sirve mayormente para registrar el pensamiento formulado oralmente. El ejemplo de Sócrates sirve para ilustrar este punto de vista, también la preeminencia de los contratos verbales en la cultura jurídica romana con la pronunciación de palabras rituales y solemnes. Llegados a este punto Derrida se decanta por la escritura y se lanza a una cruzada en su defensa buscando rescatarla de la posición secundaria en la cual la ha colocado la cultura occidental. Sobre esto nuestro filósofo demuestra que para explicar la relevancia del habla sobre la escritura se utilizan metáforas y ejemplos que surgen de la propia escritura. Según Derrida debe quedar establecido que la escritura precede al habla demostrando que ninguno de los dos términos tiene preeminencia.

Puede pensarse entonces que los sistemas de LLM, en tanto elementos conversacionales recuerdan más al habla dado que la interacción entre la máquina y los seres humanos se realiza de forma dialógica (socrática tal vez) en la forma de preguntas y respuestas. De estas interacciones pueden aparecer determinados textos y, en este sentido, puede considerarse un nuevo ejemplo de la construcción occidental de la dupla habla-escritura en la cual la escritura es un producto subordinado tal como parecen creer quienes utilizan estas herramientas como auxiliares de la investigación para una labor que consideran como secundaria y meramente derivada de su labor investigativa. Si consideramos que de acuerdo con Derrida el significado de una palabra solo cobra su valor real en el juego infinito con los significados de las otras palabras a los cuales responde (es decir que el significado se encuentra interminablemente «deferido» a otros significados) entonces las respuestas que dan los sistemas de LLM están siempre limitados a los significados que el usuario le otorga en sus preguntas, y su valor último, así como su fortaleza ética dependen del contexto que aporte el usuario.

En los términos de Derrida puede considerarse una nueva oposición que hay que deconstruir, la idea de que la información es más importante que el conocimiento. Los sistemas de LLM proveen información incesante, muchas veces inexacta, pero de fácil acceso y presentación, de aquí que se considere más importante presentar mucha información antes que convertir esa información en auténtico conocimiento. Esto se logra, volviendo a Kant, a través de la creación personal de relaciones originales y auténticas. Más aún si consideramos

que el propio acto físico de la escritura ayuda a construir el pensamiento. En este sentido Derrida señala «Escribir es saber que lo que no se ha producido todavía en la letra no tiene otra morada, no: nos espera como prescripción en algún *topos ouránios* o algún entendimiento divino. El sentido debe esperar a ser dicho o escrito para habitarse él mismo y llegar a ser lo que es al diferir de sí: el sentido» (Derrida, 1989:19) y agrega: «Es porque es inaugural, en el sentido nuevo de esta palabra, por lo que la escritura es peligrosa y angustiante. No sabe adónde va, ninguna sabiduría la resguarda de esta precipitación esencial hacia el sentido que ella constituye, y que es, en primer lugar, su futuro. Sin embargo no es caprichosa más que por cobardía. No hay, pues, seguridad contra ese riesgo. La escritura es para el escritor, incluso si no es ateo, pero si es escritor, una primera navegación y sin gracia» (Derrida, 1989:21). Más adelante Derrida cita a Merleau-Ponty en una frase que, me parece, condensa el argumento que se trata de sostener en este artículo: «Mis palabras me sorprenden a mí mismo y me enseñan mi pensamiento» (Derrida, 1989: 21). Vistas estas ideas podemos apreciar el real valor de la escritura propia en la construcción de pensamiento nuevo.

En la dupla de información versus conocimiento podemos agregar que la información obtenida de los sistemas de LLM, por su rapidez y su comodidad puede asimilarse a un producto de consumo rápido que puede sacar a un estudiante o un académico de un apuro pero que, en realidad, no tiene la calidad de conocimiento verdadero si es que consideramos que tal clase de conocimiento debería tener dos características:

1. Tiene un efecto duradero en la mente de las personas
2. A partir de ese efecto duradero es un agente de cambio personal y social.

De forma que existe una diferencia esencial entre mera información y conocimiento. Según se ha dicho la información se encuentra dispersa y no puede ofrecer las relaciones esenciales que permiten una mejor comprensión de las realidades del mundo complejo en el que vivimos y más aún, pueden evitar esa comprensión lo cual equivale a decir que la información masiva a la que estamos sometidos es un agente perjudicial para la construcción de pensamiento. Por otra parte, la información rápida (que simula un largo análisis) no colabora

con los mecanismos de cambio que son esenciales al verdadero conocimiento y al acto educativo en sí.

A todo esto, podemos agregar el valor de la lectura dado el uso cada vez mayor de resúmenes elaborados por estos sistemas. Resumir exige hacer una abstracción de las cosas más relevantes, importantes y urgentes. Esto requiere capacidades de criterio que pocas personas poseen y que, incluso cuando se hace bien, muy rara vez se alcanzan resultados satisfactorios para todos. Por esta razón el reemplazo de la función lectora implica numerosos peligros éticos entre los cuales el reduccionismo cultural parece tener preeminencia.

Es necesario tener en cuenta que las mismas limitaciones tecnológicas de estos sistemas llevan a que puedan surgir respuestas erróneas para cuya detección es indispensable un conocimiento previo de la materia sobre la cual se está preguntando. Es claro que los sistemas de LLM operan sobre bases de datos muy amplias (cuya compilación requiere un enorme trabajo humano) pero, por muy grande que sea el caudal de información disponible, siempre será limitado en el amplísimo marco del conocimiento que la humanidad ha acumulado hasta ahora y por eso, de forma fatal y necesaria, un sistema de este tipo cometerá errores que pueden autorreproducirse y que requieren de la indispensable supervisión humana. Esta posibilidad de información errónea conlleva también peligros para su uso indiscriminado y sin control docente en los ámbitos educativos. Evidentemente el principal peligro es la difusión de contenidos erróneos que pueden llevar a visiones distorsionadas de realidades históricas o políticas. Un segundo peligro está dado por la visión comercial que resulta ser una de las finalidades principales de esta clase de sistemas y que puede dejar de lado las cuestiones éticas esenciales que se han reseñado cuando su atención afecte las ganancias esperadas.

Escritura y libertad

Desde diversas aproximaciones literarias y filosóficas se ha considerado que la escritura es un espacio de libertad. En la famosa novela «1984» el protagonista busca un último rincón en el que puede expresar sus pensamientos de forma libre alejado del enorme flujo de información falsa al que está sometido en el exterior, así como de la constante vigilancia gubernamental. Providencialmente el lugar en el que puede ejercer su libertad es un modesto cuaderno en donde ejerce la

escritura como último acto liberador, Orwell nos muestra esta escena y poco a poco va demostrando que su personaje, a través del acto de escribir va alcanzando el pensamiento:

Durante algún tiempo permaneció contemplando estúpidamente el papel. La telepantalla transmitía ahora estridente música militar. Es curioso: Winston no sólo parecía haber perdido la facultad de expresarse, sino haber olvidado de qué iba a ocuparse (...) El hecho mismo de expresarse por escrito creía él, le sería muy fácil. Sólo tenía que trasladar al papel el interminable e inquieto monólogo que desde hacía muchos años venía corriéndose por la cabeza (...) De repente, empezó a escribir con gran rapidez, como si lo impulsara el pánico, dándose apenas cuenta de lo que escribía. Con su letrita infantil iba trazando líneas torcidas.» (...) Pero lo curioso era que mientras lo hacía se le había aclarado otra faceta de su memoria hasta el punto de que ya se creía en condiciones de escribir lo que realmente había querido poner en su libro (Orwell, 2022: 7–8).

Cuando se difiere la capacidad humana de concretar físicamente las ideas, a más de limitar el propio pensamiento, se limita también un espacio de libertad personal que hasta ahora había estado reservado a los seres humanos. El amplio campo de expresión que provee la Internet constituye un ejemplo único en la historia humana sobre las enormes posibilidades de comunicación y de interacción que puede ofrecer la tecnología. Es necesario recordar que la riqueza de la internet, el enorme atractivo que tiene, se basa en la noción de que detrás de cada opinión, de cada texto que se comparte existe un ser humano que está haciendo uso de su libertad y que busca, en último término, contactarse con otros seres humanos que se encuentran realizando la misma tarea. Este deseo de encuentro de libertades se había limitado, hasta hace poco, a los métodos primarios de la imprenta y de los medios de comunicación masivos. La red global ha potenciado estas oportunidades, pero el uso de elementos tecnológicos sustitutos de la propia escritura claramente corta de raíz aquellas posibilidades y, en cierta forma, si se utiliza de forma exclusiva, se constituye en un elemento que disminuye la riqueza de lo humano.

Conflicto de intereses: El autor declara que no tiene ningún posible conflicto de intereses. **Aprobación del comité de ética y consentimiento informado:** No es aplicable a este estudio: el autor no realizó estudios en animales o humanos. **Contribución de cada autor:** C.G.T. confirma que ha conceptualizado, desarrollado

las ideas y escrito el trabajo como único autor y ha leído y aprobado el manuscrito final para su publicación. Para consultas sobre este artículo debe dirigirse a: (✉) cegarcia@utpl.edu.ec.

Referencias

- Britannica, T. (2022) Editors of Encyclopaedia. «Jacques Derrida». *Encyclopedia Britannica*. <https://www.britannica.com/biography/Jacques-Derrida>.
- Derrida, J. (1989). *La escritura y la diferencia*. Editorial Anthropos. Barcelona.
- Jankowiak, Tim. «Immanuel Kant». *Internet Encyclopedia of Philosophy*.
- Kant, I. (1788) *Crítica de la razón pura*. Traducción de Manuel G. Morente. Madrid: Librería general de Victoriano Suárez.
- Lawlor, Leonard (2022). «Jacques Derrida», *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/fall2022/entries/derrida/>>.
- Orwell, G. (2022) 1984. Edición electrónica de la Escuela de Filosofía de la Universidad ARCIS. <https://www.philosophia.cl/biblioteca/orwell/1984.pdf>

Información sobre los autores

► **Carlos García Torres** es Profesor Titular en el Departamento de Derecho y Coordinador de la Cátedra UNESCO de Ética y Sociedad en la Educación Superior de la Universidad Técnica Particular de Loja, Ecuador. Es Doctor en Derecho y Ciencias Sociales por la Universidad Nacional de Educación a Distancia, España. Sus intereses de investigación se concentran en la bioética, la ética, la filosofía del derecho y el derecho romano. Es autor de obras como: *Derecho romano: una revisión sumaria* (Dykinson, 2011); *Derecho Romano* (UTPL, 2020) o *Sociología Jurídica* (UTPL, 2020). **Contacto:** Cátedra UNESCO de Ética y Sociedad en la Educación Superior, Universidad Técnica Particular de Loja, Calle Marcelino Champagnat s/n, 110107 San Cayetano Alto, Loja, Ecuador. — (✉): cegarcia@utpl.edu.ec — iD <https://orcid.org/0000-0003-1170-6765>.

Como citar este artículo

García Torres, Carlos (2022). «Espacios de pensamiento y libertad frente a sistemas de IA conversacional». *Analysis* 35, no. 10: pp. 1–8.