

Inteligencia Artificial y servicio de referencia: ¿amenaza u oportunidad?

Nicolás Petrosini
Universidad de Palermo
npetro@palermo.edu

Resumen

Los Grandes Modelos de Lenguaje de Inteligencia Artificial como ChatGPT pueden ser considerados, desde un enfoque estratégico, como una amenaza para los servicios de referencia. Entre otras fortalezas, se destaca su capacidad para dialogar sobre distintas áreas del conocimiento, la inmediatez de sus respuestas y la disponibilidad sin límites de horarios. Sin embargo, no pueden interpretar las necesidades de información de los usuarios, ni evaluar la veracidad de las respuestas que generan, ni sentir empatía hacia ellos. Por estas razones, en tanto sean utilizados como fuentes de información, se requerirá de mediadores que garanticen la satisfacción de los usuarios. Los referencistas, por su formación en recuperación de la información, capacidad para evaluar las respuestas y vocación de servicio, son profesionales idóneos para interactuar con los modelos de lenguaje. Para ello, deberán aprender Ingeniería de Prompts, una competencia que consiste en optimizar las instrucciones para que los modelos generen respuestas pertinentes. En síntesis, esta ponencia aborda la necesidad de adaptación y actualización, para que los bibliotecarios y los servicios de referencia sean relevantes en la Era de la IA.

Palabras clave: Servicios de referencia, Inteligencia Artificial Generativa, Ingeniería de Prompts

Introducción

En el año 2013, dos investigadores de la Universidad de Oxford publicaron un estudio que causó un *shock* en la comunidad científica y en el mundo empresarial. Según sus conclusiones, el 47% de los empleos se encontraba en alto riesgo de ser automatizado veinte años más tarde (Benedikt Frey y Osborne: 2013). Para demostrarlo, diseñaron un algoritmo que ponderó la probabilidad de cada trabajo para ser informatizado. Y entre los profesionales con un valor superior al 50% de las probabilidades se encontraban los bibliotecarios¹.

Con el tiempo, los autores relativizaron sus afirmaciones (Merino, 2019). Sin embargo, en la actualidad estas se han resignificado a partir de los últimos avances en el desarrollo de la Inteligencia Artificial Generativa; en particular de los Grandes Modelos de Lenguaje (LLM) como ChatGPT.

En el 2023, Bill Gates declaró que la Inteligencia Artificial (IA) “tendrá un impacto abrumador en todas las esferas de la sociedad.” El fundador de Microsoft recomendó “estar listos para las oportunidades que surgirán en este entorno en constante evolución”, para lo cual es “esencial adquirir conocimientos en IA y desarrollar competencias digitales” (*Bill Gates asegura que solo tres tipos de trabajo tendrán futuro con la inteligencia artificial*, 2023).

Los bibliotecarios, en su función de custodios del conocimiento y organizadores de colecciones, han existido durante miles de años, desde las civilizaciones sumerias y egipcias. Ante este nuevo escenario, ¿cómo pueden adaptarse para que los servicios de referencia sean relevantes en la Era de la Inteligencia Artificial?

El objetivo de esta ponencia es explorar el rol del referencista como mediador entre la IA Generativa y los usuarios, para lo cual deberá aplicar la Ingeniería de Prompts.

¹ Bibliotecarios (*Librarians*): 360ª posición en el ranking; probabilidad de ser reemplazado 0.65, siendo el rango de valores entre 0 y 1. Para la investigación, los autores se basaron en un listado de 702 empleos del Departamento de Trabajo de los Estados Unidos.

Discusión

Fortalezas y debilidades de los LLM

Desde un enfoque estratégico, cabe considerar a los modelos de lenguaje como una amenaza para los servicios de referencia.

Por un lado, pueden dialogar sobre todas las áreas del conocimiento, ya que han sido entrenados con enormes conjuntos de datos (Adetayo, 2023: 19) de diversas fuentes: entre ellas, las disponibles libremente en internet, como libros, artículos, páginas web y foros (en el caso de ChatGPT, se estima que la versión 3.5 procesó el equivalente a 300.000 millones de palabras) (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2023: 5). Pero no responden con la frialdad de un motor de búsqueda, sino que dialogan² simulando una conversación con otro ser humano; lo cual promueve que la experiencia sea “cautivante y dinámica” (Panda y Kaur, 2023: 23).

Por otro lado, pueden responder inmediatamente, incluso ante preguntas complejas. Si se compara con los tiempos de una consulta de referencia digital, el usuario no debe esperar por la respuesta, que en ocasiones necesita de una repregunta para aclarar alguna cuestión y, posiblemente, de una espera adicional. Desde el punto de vista del usuario, ya sea para investigar, completar una tarea o buscar orientación sobre un tema, los LLM pueden proveer “un mundo de conocimiento” en solo unos segundos (Panda y Kaur, 2023: 23).

Por último, están disponibles sin límites de horarios. Los usuarios pueden chatear en cualquier momento del día, los siete días de la semana. Esta particularidad brinda una solución conveniente para ellos, que les permite ahorrar tiempo (Chen, 2023: 125).

Sin embargo, los LLM poseen ciertas limitaciones.

En primer lugar, no pueden interpretar las necesidades de información. El funcionamiento de un modelo de lenguaje se basa en predecir palabras y frases en un contexto determinado (*How ChatGPT and our language models are developed*, 2024); en otros términos, un LLM no comprende lo que el usuario le solicita, sino que simula una conversación, en la cual selecciona las palabras más probables a partir de la instrucción dada, llamada *prompt*. Sin embargo, el usuario no siempre demanda lo que necesita: hay matices que intervienen en los procesos de comunicación. Por ejemplo, el

² Además, se trata de una interfaz basada en el lenguaje natural, con el cual los usuarios están familiarizados; al menos en el uso básico, que no necesariamente implica saber recuperar información de manera efectiva.

hecho de que frecuentemente el usuario “oculta” su verdadera pregunta, a veces por miedo o vergüenza de admitir algo que no sabe; y, para remediarlo, puede hacer consultas más generales (Mulac, 2012).

En segundo lugar, los modelos no evalúan las respuestas que generan: estas deben ser creíbles y coherentes, lo cual no siempre coincide con lo verdadero. De hecho, cuando no saben una respuesta pueden inventarla, fenómeno conocido como “alucinación” (Zucon et al., 2023). Otro aspecto que incide en la veracidad de las respuestas es que los modelos de lenguaje no se rigen por principios éticos, y los datos con los cuales se entrenan son elaborados por humanos, con lo “correcto y lo incorrecto, lo verdadero y lo falso” (UNESCO, 2023: 11).

En tercer lugar, no sienten empatía. Si bien pueden responder de una forma que parece empática, los LLM no son capaces de “leer” las emociones que subyacen en un *prompt* (Adetayo, 2023b: 10) como lo haría un ser humano, porque no tienen emociones.

En síntesis, en contraste con el servicio de referencia, pueden observarse grandes fortalezas en los LLM, que los convierte en una amenaza; pero también ciertas debilidades que son críticas para la satisfacción de las necesidades informativas de los usuarios. Hasta que estos utilicen los modelos de lenguaje con frecuencia, y comprendan la importancia de los aspectos mencionados, es posible que descarten otras opciones para buscar información, como los servicios de referencia. Entonces, ¿es conveniente competir con la IA Generativa?

Como sostiene Bilinkis (2022), “cuando se habla de la relación entre los humanos y las máquinas el énfasis suele ponerse en la rivalidad”. Pero hay algo más “poderoso que un humano o una máquina: la combinación de ambos” (p. 59).

Referencistas como mediadores entre los LLM y los usuarios

El proceso de diseño de un *prompt* para una tarea específica involucra conceptos similares que los referencistas han aplicado durante mucho tiempo: cómo las consultas de búsqueda ingresadas impactan en la información recuperada y la importancia de evaluar la calidad de las respuestas obtenidas (Lund, 2023). Por lo tanto, pueden ser considerados profesionales capacitados para mediar entre los LLM y los usuarios. En particular, cabe mencionar algunas aptitudes que conforman su perfil, que los posiciona para ocupar ese rol.

Primero, para realizar su trabajo tienen que desarrollar habilidades de comunicación que les permiten comprender las necesidades informativas de los usuarios. “Un buen referencista sabe captar las suposiciones no expresadas, los sesgos o las omisiones en una consulta de referencia, e indagar por la pregunta detrás de la pregunta” (Bates, 2024: 35).

Segundo, están entrenados para evaluar la información. Los referencistas deben “responder verazmente a las consultas sin omisiones ni errores deliberados” (Merlo Vega, 2000: 99), con lo cual desarrollan una capacidad para validar la información que brindan.

Tercero, poseen vocación de servicio. En palabras de Thomsen (1999), “el referencista debe tener un deseo genuino e inquebrantable de ayudar a otros a alcanzar sus propios objetivos, sin importar cuáles sean estos” (p. 25). Es decir, un referencista no juzga la necesidad del usuario, sino que se compromete a ayudarlo. Y, en esa actitud, muchas veces debe entrar en juego la empatía, la cual “permite entender las intenciones de los demás” (Baron-Cohen y Wheelwright, 2004: 163).

Sin embargo, para poder interactuar eficazmente con los LLM no basta con ejercer las aptitudes mencionadas: los referencistas también deben aprender a dialogar con estos sistemas.

Una nueva competencia: Ingeniería de Prompts

Los usuarios tienden a buscar información en un LLM como lo harían en un motor de búsqueda, ingresando una consulta básica y aceptando la respuesta proporcionada (Lund, 2023). Pero, sin el *prompt* adecuado, el modelo falla en proveer una respuesta relevante (Solomon, 2023). Esencialmente, esto ocurre porque los LLM responden al lenguaje natural, que es impreciso. “La misma frase puede tener diferentes significados dependiendo del contexto, lo que dificulta que el modelo entienda lo que el usuario quiere que genere” (Smith, 2023).

La Ingeniería de Prompts (*Prompt Engineering*) es el proceso de optimizar las entradas (*inputs*) para que los modelos generen salidas (*outputs*) relevantes (Lund, 2023). Korzynski y otros (2023) consideran que es una competencia centrada en el lenguaje humano y relacionada con la alfabetización digital (p. 26). Más allá de estas definiciones, existen tres estrategias básicas de la Ingeniería de Prompts que los referencistas pueden aplicar (*Effective prompts for AI*, 2024):

- Ser específico: un *prompt* bien diseñado debe ser claro y específico para guiar al modelo hacia la respuesta deseada.

- Iterar: consiste en ingresar nuevos *prompts* para seguir “construyendo sobre la conversación”. Este proceso permite “desbloquear más potencial” de la IA.

- Proveer contexto: para dirigir al modelo hacia una respuesta relevante es crucial aportar suficiente información que le permita predecir las palabras y frases más coherentes.

Como puede observarse, en estas estrategias hay algunos principios que los referencistas utilizan en su trabajo cotidiano, pero adaptadas a las características y funcionamiento de los modelos de lenguaje. Para ejemplificarlas se explicará un caso elaborado para esta investigación. El objetivo de la tarea era que una persona, sin conocimientos profundos de computación, pudiera comprender la magnitud del conjunto de datos (*dataset*) que se usó para entrenar a ChatGPT.

Primero se le preguntó al chat: “¿De cuántas palabras es el dataset que utilizaron para entrenarte?” (nótese que la consulta es por “palabras” y no, por ejemplo, por “información”, es decir, se usó un término específico). Sin embargo, con este *prompt* no pudo obtenerse una salida precisa: el chat respondió que OpenAI, la empresa que lo desarrolló, no ha revelado esos datos.

Para refinar la salida, se iteró, aportando el dato estimado ya mencionado en este trabajo: “Suponé que la versión de GPT-3.5 fue entrenada con 300.000 millones de palabras”. Ante esta instrucción, el chat respondió sobre distintos aspectos e hizo una comparación con un libro “típico” de 100.000 palabras: por medio de un cálculo, estimó que fue entrenado con un volumen de datos “equivalente de 3 millones de libros”.

Si bien esta salida puede ser más concreta que hablar de millones de palabras, se intentó mejorarla para alcanzar el objetivo propuesto; para lo cual se ingresó el *prompt* “Redactá una analogía con esta información: si una persona puede leer 1 libro de 100.000 palabras por semana, ¿cuántos años tardaría en leer el *dataset*?” La salida del modelo fue la siguiente:

Imagina que decides emprender un viaje de lectura en el que cada semana lees un libro de 100.000 palabras. Si te embarcaras en la tarea de leer un dataset de 300.000 millones de palabras, te llevaría unos 57.692 años completarlo, dedicando una semana entera a cada libro. Esto es más tiempo del que ha existido la civilización humana, remontándose a tiempos anteriores de la historia registrada. (OpenAI, 2024).

La relación entre la cantidad de palabras del *dataset*, las palabras por libro y la cantidad de libros leídos por semana contiene suficiente contexto para que el modelo genere una salida relevante.

Por otro lado, es posible optimizar las respuestas de los LLM aplicando diferentes patrones (*prompt patterns*). De forma similar a la programación de computadoras, un *prompt*, además de ser una pregunta o solicitud, puede ser un conjunto de instrucciones para “programar” al modelo; es decir para personalizarlo o mejorar sus capacidades. De esta manera, los patrones brindan soluciones reutilizables para resolver problemas ante situaciones similares (White et al., 2023).

Por último, cabe aclarar que, en tanto las plataformas sobre las cuales funcionan los LLM están evolucionando cada vez más rápido, y se vuelven más sofisticadas, es posible que en el mediano o corto plazo algunas técnicas para conversar con los modelos se vuelvan más sencillas o, simplemente, ya no sean necesarias. Algunos autores como Acar (2023) se refieren a un enfoque basado en la resolución de problemas y no en la redacción de *prompts* adecuados. En todo caso, “es muy probable que aquellos que logren hablar con la IA de manera más efectiva y obtengan por tanto lo mejor de ellas, desempeñen la mayoría de los trabajos” (Sigman y Bilinkis, 2023: 131).

Conclusión

Los Grandes Modelos de Lenguaje pueden ser considerados como una amenaza o como una oportunidad para el servicio de referencia. Poseen grandes fortalezas, contra las cuales los bibliotecarios no pueden competir. Pero tienen debilidades, las cuales requieren de un mediador humano que adecúe las respuestas, para satisfacer las necesidades de información de los usuarios. En este nuevo escenario, los referencistas pueden desempeñar un papel fundamental, por las aptitudes que conforman su perfil profesional, entre ellas, la capacidad de interpretar las necesidades del usuario, la evaluación de la información y la vocación de servicio.

Pero no pueden confiar solamente en el saber adquirido hasta el presente. Deben avanzar en su carrera por mantenerse actualizados. El camino que propone esta ponencia es aprender a dialogar con los modelos de lenguaje por medio de la Ingeniería de Prompts. ¿Es la Inteligencia Artificial una amenaza o una oportunidad para el servicio de referencia? La respuesta depende de los bibliotecarios.

Referencias

- Acar, O. A. (2023, junio 6). AI Prompt Engineering isn't the future. *Harvard Business Review*. <https://hbr.org/2023/06/ai-prompt-engineering-isnt-the-future>
- Adetayo, A. J. (2023a). Artificial intelligence chatbots in academic libraries: the rise of ChatGPT. *Library Hi Tech News*, 40(3), 18-21. <https://doi.org/10.1108/LHTN-01-2023-0007>
- Adetayo, A. J. (2023b). ChatGPT and librarians for reference consultations. *Internet Reference Services Quarterly*, 27(3), 131-147. <https://doi.org/10.1080/10875301.2023.2203681>
- Baron-Cohen, S. y Wheelwright, S. (2004). The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2), 163-175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>
- Bates, M. E. (2024). Librarians as prompt engineers. *Computers in Libraries*, 44(2), 35.
- Benedikt Frey, C. y Osborne, M. (2013). *The future of employment: how susceptible are jobs to computerisation?* Oxford Martin School. University of Oxford. <https://www.oxfordmartin.ox.ac.uk/publications/the-future-of-employment>
- Bilinkis, S. (2022). Preparándose para los trabajos del futuro. *Revista Fasecolda*, 188, 58-61. <https://revista.fasecolda.com/index.php/revfasecolda/article/view/861>
- Bill Gates asegura que solo tres tipos de trabajo tendrán futuro con la inteligencia artificial. (2023, agosto 4). *La Gaceta*. <https://www.lagaceta.com.ar/nota/1001434/sociedad/bill-gates-asegura-solo-tres-tipos-trabajo-tendran-futuro-inteligencia-artificial.html>
- Chen, X. (2023). ChatGPT and its possible impact on library reference services. *Internet Reference Services Quarterly*, 27(2), 121-129. <https://doi.org/10.1080/10875301.2023.2181262>
- Effective prompts for AI: the essentials*. (2024). MIT Sloan Teaching & Learning Technologies. <https://mitsloanedtech.mit.edu/ai/basics/effective-prompts/>
- How ChatGPT and our language models are developed*. (2024). OpenAI. <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed>
- Korzynski, P., Mazurek, G., Krzypkowska, P. y Kurasinski, A. (2023). Artificial intelligence prompt engineering as a new digital competence: analysis of generative

- AI technologies such as ChatGPT. *Entrepreneurial Business and Economics Review*, 11(3), 25-37. <https://doi.org/10.15678/EBER.2023.110302>
- Lund, B. (2023). The prompt engineering librarian. *Library Hi Tech News*, 40(8), 6-8. <https://doi.org/10.1108/LHTN-10-2023-0189>
- Merino, M. (2019, julio 1). *El estudio que alarmó al mundo diciendo que en 2033 el 47% del empleo estaría en manos de robots ya no lo defienden ni sus autores*. Xataka. <https://www.xataka.com/inteligencia-artificial/estudio-que-alarma-al-mundo-diciendo-que-2033-47-empleo-estaria-manos-robots-no-defienden-sus-autores>
- Merlo Vega, J. A. (2000). El servicio bibliotecario de referencia. *Anales de documentación*, 3, 93-126. <https://digitum.um.es/digitum/handle/10201/3690>
- Mulac, C. M. (2012). The reference interview. En *Fundamentals of reference*. American Library Association.
- OpenAI. (2024). *Analogía de dataset para entrenamiento* (2 de junio) [Large language model]. <https://chatgpt.com/share/609d8669-46fb-4128-b15f-719e5c07ddaf>
- Panda, S. y Kaur, N. (2023). Exploring the viability of ChatGPT as an alternative to traditional chatbot systems in library and information centers. *Library Hi Tech News*, 40(3), 22-25. <https://doi.org/10.1108/LHTN-02-2023-0032>
- Sigman, M. y Bilinkis, S. (2023). *Artificial: la nueva inteligencia y el contorno de lo humano*. Debate.
- Smith, C. S. (2023, abril 5). *Mom, dad, I want to be a Prompt Engineer*. Forbes. <https://www.forbes.com/sites/craigsmith/2023/04/05/mom-dad-i-want-to-be-a-prompt-engineer>
- Solomon, L. (2023, febrero 10). *Are librarians the next prompt engineers? What does this mean to me, Laura?* <https://meanlaura.com/are-librarians-the-next-prompt-engineers>
- Thomsen, E. (1999). *Rethinking reference: the reference librarian's practical guide for surviving constant change*. Neal-Schuman.
- United Nations Educational, Scientific and Cultural Organization. (2023). *ChatGPT e inteligencia artificial en la educación superior: guía de inicio rápido*. https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. y Schmidt, D. C. (2023). *A prompt pattern catalog to enhance prompt engineering with ChatGPT*. arXiv. <https://doi.org/10.48550/arXiv.2302.11382>

Zuccon, G., Koopman, B. y Shaik, R. (2023). *ChatGPT hallucinates when attributing answers*. arXiv. <https://doi.org/10.48550/arXiv.2309.09401>