

# *LOS PROBLEMAS MÁS IMPORTANTES DE LA IA*

## AÚN NO ESTÁN RESUELTOS

**RAMÓN LÓPEZ DE MÁNTARAS**

Profesor emérito de investigación en Inteligencia Artificial del CSIC

Ramón López de Mántaras es uno de los mayores expertos en Inteligencia Artificial (IA), la tecnología de la que todo el mundo habla. Fue el fundador y director del Instituto de Investigación en Inteligencia Artificial del CSIC, donde aún hoy, como emérito, sigue investigando en razonamiento por analogía y en aprendizaje de relaciones causales y por transferencia en robots humanoides. Precisamente ahí ve uno de los *hándicaps* de la IA: la capacidad de inferir, de aprender por sí misma, de establecer relaciones causa-efecto, es decir, que sea realmente inteligente.

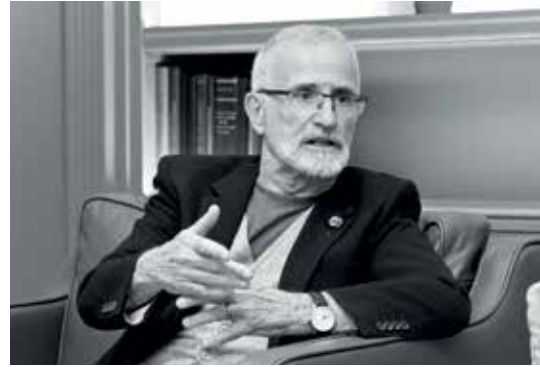
Por Carlos Bueno



“Va a haber una decepción en cuanto a las promesas de la Inteligencia Artificial. Todavía hay muchísimo por investigar”

Ramón López de Mántaras ofreció, en la Fundación Ramón Areces, una conferencia sobre los grandes retos a los que se enfrenta la IA. Charlamos con él antes de su intervención sobre sus inicios en este campo, allá por los años 70, de cómo en estas cuatro décadas largas de investigación estas prometedoras herramientas han fracasado por las altas expectativas que habían generado e incumplido. También, de la necesidad de una mayor colaboración científica, de los errores que oculta el ahora famoso ChatGPT -“está muy sobrevalorado”, nos dice- y de la conducción autónoma: “Líderes de la industria de la automoción dudan de que alcancemos en algún momento el nivel cinco de autonomía. Lo veo muy complicado y, de llegar, no sería antes de diez años”.

Doctor en Física por la Universidad Paul Sabatier de Toulouse (Francia), Master of Science en Informática por la Universidad de California-Berkeley (EEUU) y Doctor en Informática por la Universidad Politécnica de Cataluña. López de Mántaras fue Profesor Titular de la Facultad de Informática de Barcelona y Catedrático de la Universidad Pierre et Marie Curie de Paris. Fue Editor principal de *AI Communications* y Editor Asociado de la revista más prestigiosa en IA: *Artificial Intelligence Journal*. Autor del libro 'Inteligencia Artificial' ha firmado más de 300 trabajos científicos.



Chancellor's Distinguished Visiting Professor de la Universidad de Tecnología de Sydney y de la Western Sydney University, este pionero de la IA en España y Europa se muestra reacio a titulares espectaculares sobre el potencial de la IA “porque los problemas más importantes aún están sin resolver y queda mucho por investigar”. Confiesa que en su vida cotidiana huye de cualquier sistema de asistente personal inteligente -Siri, Google Assistant...- más bien por una cuestión de privacidad. Y, en sus ratos libres, disfruta tocando jazz al piano o patinando en línea, habiendo participado en competiciones como las 24 horas del circuito de LeMans y en dos ultramaratones de Barcelona.

**P. ¿Cómo recuerda sus inicios en este campo de la Inteligencia Artificial (IA)?**

R. Fue en el año 1975, cuando realizaba mi tesis doctoral en la Universidad de Toulouse, en el laboratorio de analítica de sistemas, en el CNR, el equivalente al CSIC español. Allí disponían de una mano antropomórfica con cinco dedos y tres falanges por dedo. Se me ocurrió que podía ser capaz de distinguir, solo a través de sensores táctiles, entre distintos tipos de objetos con variadas formas. En eso consistió mi tesis, en desarrollar unos algoritmos de aprendizaje automático para que

reconociera esos patrones distintos. Ese fue mi primer contacto con la IA. Curiosamente, se trataba de detectar patrones, que es lo más utilizado ahora. Las redes neuronales también realizan estas tareas, identifican patrones en conjuntos de datos.

**P. ¿Esperaba esta evolución de la IA?**

R. Ha llovido mucho desde entonces. La IA ha tenido sus altibajos hasta llegar a hoy. Vivió varios inviernos porque se prometieron cosas en un momento dado que no se llegaron a cumplir. Se crearon expectativas desmesuradas que, al no cumplirse, generaron un problema de financiación. A finales de los años 70, poco después de concluir mi tesis, llegó la primera crisis de la IA. Pero después le dieron nuevo auge los sistemas expertos, que son capaces de razonar a partir de un conocimiento muy concreto, como por ejemplo diagnosticar neumonías, detectar diferentes averías... Y se volvió a prometer demasiado. Se repitió el error de pensar que se iba a poder hacer cualquier cosa gracias a esta tecnología. Y no fue así. Los sistemas expertos se quedaron obsoletos rápidamente porque no aprendían, no se actualizaban, por ejemplo, con los nuevos conocimientos de la medicina.

**P. Y volvió a desconfiarse en ella...**

R. Sí, y aún tuvo que venir otro invierno para la IA. Hacia los inicios de los años 90, llegaron nuevos algoritmos de aprendizaje, se desarrolló la robótica y esa nueva época de esplendor duró otros 10 o 12 años gracias a las redes neuronales profundas. Estas son capaces, a partir de gran cantidad de datos y potencia de cálculo, de hacer cosas extraordinarias y también en el reconocimiento de imágenes... Ha habido como cuatro etapas de la IA.



“La IA ha vivido varias épocas y ha despertado demasiadas expectativas incumplidas. Podemos caer de nuevo en el mismo error de prometer demasiado”

**P. ¿Y en qué momento nos encontramos?**

R. De nuevo estamos prometiendo mucho. Podemos caer de nuevo en el mismo error. Me temo que nos vamos a dar cuenta de las limitaciones. Si no vivimos un nuevo invierno de la IA, pues puede ser que tengamos un otoño un poco duro, un poco lluvioso, porque creo que va a haber una decepción en cuanto a las promesas. Todavía hay muchísimo por hacer en IA. La historia se repite. Estamos un poquito mejor que en las anteriores épocas porque ahora están apostando mucho y más fuerte en cuanto a inversiones por parte de grandes empresas tecnológicas, de empresas en general y también de gobiernos. Muchos dirigentes piensan que van a dominar el mundo con esto. Hay tantísimo apoyo, tantísimo, que, aunque no se cumplan muchas de las expectativas, no parece que de pronto se deje de financiar y se deje de apostar por esta tecnología.

**P. ¿Aunque queden tantas dudas en el aire y limitaciones?**

R. Creo que se va a seguir apostando por la IA a pesar de que vamos a descubrir que hay muchas limitaciones y que necesitamos mucha más investigación de lo que parece. Los grandes problemas y más importantes de la IA aún no están resueltos. Hay buena ingeniería muy bien hecha con resultados muy buenos en cuestiones prácticas, pero a las preguntas científicamente más importantes no les hemos encontrado respuesta.

“ChatGPT está muy sobrevalorado. No tiene sentido llamar inteligente a un sistema que da unas respuestas que son completamente absurdas”

**P. ¿Parte de esas expectativas incumplidas están en la conducción autónoma?**

R. Ese es un buen ejemplo. Recuerdo a Elon Musk, el fundador de Tesla, diciendo en 2017 más o menos que en el año 2022

íbamos a ver coches autónomos por todas partes. Estamos en 2023 y no los hemos visto.

**P. ¿Es un problema de regulación o también técnico?**

R. No es solo un problema de regulación. La conducción autónoma plantea cuestiones complejas en cuanto a las responsabilidades. Un coche con total autonomía en nivel 5 -en el que ya no habría ni freno ni acelerador ni volante dentro del vehículo- incluye cuestiones complejas si se produce un accidente. Las compañías de seguros quieren tener todo claro, saber qué ha fallado (si el software, los sensores...) Hay grandes fabricantes de coches que dudan sobre la llegada del nivel 5 de autonomía. Es probable que nunca se llegue a implantar, que siempre tenga que haber un volante, un freno y un acelerador y una persona sentada en el puesto de conductor por si eventualmente tiene que tomar el control del coche. Este es un buen ejemplo de algo que hace unos años parecía resuelto, pero no era así. De llegar a ese nivel 5 de autonomía, será quizá en 10 años.



**P. ¿Tendremos que conformarnos entonces con unas ayudas a la conducción?**

R. El nivel 3 está muy cerca. Ya hay vehículos que se podrían considerar prácticamente nivel 3. Todavía quedan detalles por pulir, pero tiene que haber una persona en el puesto de conductor y estar atenta. Recientemente conduje un coche en Australia en nivel 3 que en autopista iba bien con el modo Smart o autopiloto. Pero si estaba 30 segundos sin coger el volante, ya pitaba y aparecía un texto pidiendo que pusiera las manos sobre él.

**P. ¿Entonces para qué sirve todo esto?**

R. Buena pregunta. Si no puedo ir leyendo, sino agarrado al volante... El

sistema corrige si tienes tendencia a desviarte y te lleva por el centro del carril. Es una ayuda a la conducción. Pero en cuanto me salí de la autopista y entré por carreteras con las líneas algo desdibujadas, el modo Smart no se activaba de ninguna manera. Esto demuestra lo complicado que todavía es y los retos por resolver, hay enormes dificultades. Y no nos sirve que nos digan que tiene una fiabilidad del 95%. Esto es muy serio, necesitamos confianza del 100% en el sistema.

**P. ¿Qué opinión le merece el robot generador de textos automáticos ChatGPT? ¿Ha ayudado a popularizar esta tecnología?**

R. Es cierto que todo el mundo lo conoce ya. También envía el mensaje de que es impresionante y de que puede obtener resultados espectaculares, pero carece de robustez. Comete unos errores extraordinarios. Como decir que el número 20 no es menor que 24, por ejemplo. En otro caso, le dimos un texto sencillo de lenguaje natural y cuando le hicimos unas preguntas sobre él, descubrimos que no había entendido casi nada. El problema de ChatGPT es que evalúa cuál es la probabilidad de la siguiente palabra en el texto. Nada más. No hay un análisis semántico, en absoluto. Es como un loro que repite aquello para lo que ha sido entrenado.

**P. ¿Es un problema del material con el que ha sido alimentado?**

R. ChatGPT se ha entrenado con billones de palabras, con corpus de textos enormes, tiene 175 millones de parámetros que se ajustan en el ecosistema de redes neuronales. Una vez me dijo que el trofeo no cabía en el cajón porque era demasiado grande. Y cuando le pregunté qué es lo que es demasiado grande me respondió que el cajón.



“Para mejorar a ChatGPT hay que dotarle de capacidad semántica y de razonamiento, no basta con alimentarle con más parámetros de información”

**P. Ha dicho Bill Gates que ChatGPT va a ser tan importante como la llegada de Internet...**

R. Hay quien dice también que va a acabar con los buscadores tipo Google. No lo creo, porque son cosas distintas. ChatGPT no ha sido diseñado para ser un buscador. Podemos buscar información, pero el problema es que nos dará tantas burradas como resultado, que no nos dará confianza, algo clave. El problema de la confianza es muy importante. Cuando hacemos una búsqueda ahora en un buscador, sabemos que esas búsquedas están relacionadas y nos da

confianza, vemos las fuentes. Pero ChatGPT (o estos sistemas 'large language models') pueden sacar cualquier tontería.

**P. ¿Es una cuestión de cantidad?**

R. No es suficiente con hacer modelos con más parámetros. Ahora GPT 4 va a tener trillones de palabras en lugar de billones. No creo que

“Hay que ser muy cautelosos a la hora de dejar que las máquinas decidan autónomamente sin que pueda intervenir una persona”

sea suficiente escalar en tamaño para que funcione mejor. Seguirá teniendo errores, errores enormes de bulto, completamente estúpidos y esto seguirá afectando a la confianza. Para mí, el camino no va por ahí. El camino tiene que ir por añadir más capacidad semántica y más capacidad de razonamiento a GPT. Es un problema duro, de investigación complicado, pero pienso que se pueden conseguir resultados.

**P. ¿Se está sobrevalorando entonces el ChatGPT?**

R. Yo creo que está sobrevalorado. Muy sobrevalorado. Con estos ejemplos queda muy claro que no tiene sentido llamar inteligente a un sistema que da respuestas completamente absurdas. A esto no lo considero algo que se pueda llamar inteligente.

Metiendo más parámetros y haciendo el modelo más complejo es posible que si ahora comete un 30% de errores de bulto, puede ser que lo bajen a un 5%. Pero el problema es que no hay que bajarlo al 5%, sino que hay que bajarlo a cero.

**P. Como sucede con la conducción autónoma...**

R. Claro, no nos podemos permitir un porcentaje de error, aunque sea pequeño, porque eso va a afectar a la confianza del sistema.

**P. Dice usted que “estamos dejando muchas decisiones en manos de máquinas estúpidas”.**

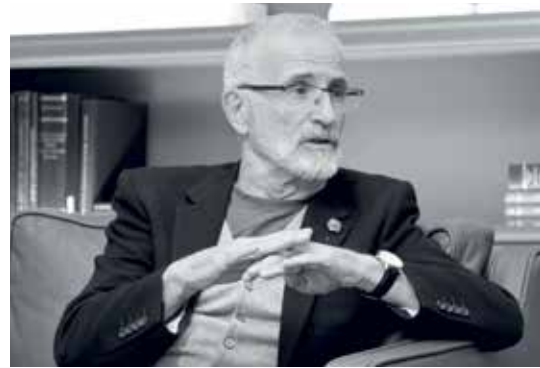
R. Creo que sí. El problema es que vamos con una cierta ligereza. Hay que ser muy cautelosos a la hora de dejar que las máquinas decidan autónomamente sin que pueda intervenir y corregir la situación una persona. Otra cosa distinta es que ayude a tomar una decisión. Pero la decisión la toma finalmente la persona, que es la que tiene la valoración moral. Como en la elaboración de sentencias, que sabemos que ya está ocurriendo.

**P. ¿Cree que muchas veces se da por sabida la inteligencia natural?**

R. Sin duda, sin duda. Además, tenemos esa tendencia a creernos demasiado lo que dice la tecnología. El problema no es tanto el grado de inteligencia, sino el grado de autonomía que demos a un sistema o sistemas que incluso no están basados en IA pero que controlan algo y al que le damos autonomía absoluta.

**P. ¿De qué sirve que consigamos un sistema de IA predictiva que alerte de que se va a producir una gran tormenta si luego los ciudadanos no siguen esas recomendaciones y cogen el coche y quedan atrapados?**

R. Además, la inteligencia natural y la IA son muy distintas. Tenemos la tendencia a antropomorfizar muchísimo a las máquinas, es algo que hacemos en exceso. Lo vemos en los titulares de los periódicos. La



máquina de por sí no tiene iniciativa propia, se le ha entrenado para hacer ciertas cosas, pero poco más.

**P. La IA también ha tenido un papel relevante en la investigación del Covid-19 ¿Qué lecciones hemos aprendido de la pandemia?**

“Tenemos que promover la ciencia abierta, cuanto más abierta mejor, compartir resultados, datos...”

R. Hemos aprendido que la colaboración en ciencia es clave. Y cuanto más amplia e internacional sea, mucho mejor. Esto se sabía, pero ahora lo hemos asumido mucho más. Un investigador avanza a partir de lo que otro ya ha hecho y ese otro puede estar en la otra parte del mundo. Tenemos que promover la ciencia abierta, cuanto más abierta mejor, compartir resultados, datos...

Pero sigue habiendo mucha reticencia a compartir datos, incluso entre hospitales... Y no es un problema de privacidad, porque se pueden tomar medidas para anonimizar los datos. El problema ha sido que cada cual es reticente a soltar y compartirlos. Tenemos que aprenderlo para futuras pandemias, tendremos que reaccionar de forma distinta, estar más dispuestos a compartir datos de calidad.

**P. En el mundo empresarial, ¿cree que se están aprovechando las posibilidades que ofrece la IA, a partir de los gemelos digitales, por ejemplo?**

R. Leía no hace mucho un estudio o encuesta en empresas de Estados Unidos de diferentes sectores que en torno al 70 u 80% reconocían que usaban IA. Cuando se les preguntaba de forma más minuciosa, se descubriría que apenas un 30% de ellas la usaban en realidad en sus negocios. Lo ponen en sus webs por cuestión de marketing o imagen, pero luego nada. No, todavía no se les está sacando todo el jugo que se le podrá sacar. Es un reto técnico, ético y también de falta de talento. Hay una escasez de experiencia aquí y en todo el mundo. No se encuentran expertos en IA. Las empresas tienen ese problema. Si no puedes tener a nadie

en esos proyectos en IA, es complicado desarrollar aplicaciones con esa tecnología...

“Una máquina nunca tendrá capacidad de socialización, de empatía, de improvisación, de creatividad, de gestionar una situación completamente imprevista”

**P. ¿Estamos preparados para ser sustituidos por máquinas?**

R. No, creo que todavía tardará mucho tiempo, si llegamos a ser sustituidos por máquinas. Es cierto que determinadas tareas ya han sido asumidas por máquinas e incluso sin necesidad de IA, como ha sucedido con los cajeros automáticos de los bancos. La informática y la automatización lleva muchos años sustituyendo a personas, y se van creando nuevos puestos y perfiles de trabajo. ¿Compensamos los que se destruyen con los que se crean? Es una transformación.

Una máquina nunca tendrá capacidad de socialización, de empatía, de improvisación, de creatividad, de gestionar una situación completamente imprevista. En todos esos casos la máquina no tiene nada que hacer.