



Palabras clave: inteligencia artificial, información, crisis, expertos humanos, máquina y humano.



MANUEL CEBRIÁN
Investigador distinguido y catedrático de Excelencia Permanente en la UC3M

CÓMO LA INTELIGENCIA ARTIFICIAL PUEDE AYUDARNOS A NAVEGAR POR LAS CRISIS

La IA como salvavidas

La IA puede ser un salvavidas en tiempos de crisis, un socio que nos ayuda a mantener la cabeza fuera del agua. Puede ayudarnos a procesar la avalancha de información, a encontrar claves que de otro modo pasaríamos por alto, a coordinar una respuesta rápida y efectiva. La IA por sí sola no es suficiente; requiere la guía de expertos humanos.

*How artificial intelligence can help us navigate crises
AI AS A LIFESAVER*

AI can be a lifesaver in times of crisis, a partner that helps us keep our heads above water. It can help us process the flood of information, find clues we would otherwise miss, coordinate a quick and effective response. AI alone is not enough; it requires the guidance of human experts.

Keywords: artificial intelligence, information, crisis, human experts, machine and human.

Imagina que se desata una crisis. La información fluye desde todas partes, un torrente de datos abrumador en su volumen y complejidad. En medio del caos, la inteligencia artificial entra en acción. No es una IA cualquiera, sino un Modelo de Lenguaje Grande (MGL), una maravilla de la tecnología moderna que puede comprender y generar lenguaje humano. Comienza a recopilar datos, filtrando el ruido para encontrar la señal. Resume la información, destilando el caos en claridad. Interactúa con las personas, proporcionando actualizaciones, respondiendo preguntas, ofreciendo tranquilidad. Es un faro de calma en la tormenta, una mano firme en medio de la turbulencia.

Pero el MGL no trabaja solo. Detrás de escena, hay un equipo de expertos, trabajando con algoritmos complejos y herramientas de optimización. Son los arquitectos de la IA, los titiriteros que

mueven los hilos. Toman la información recopilada por el MGL, los planes que ha hecho, y los refinan. Utilizan su experiencia, su comprensión de la situación, para ajustar la salida de la IA. Son el toque humano en la máquina, el corazón en la IA.

En una crisis, esta combinación de IA y experiencia humana puede marcar un cambio radical. La IA puede rápidamente unificar todas las informaciones de las personas, dar a las personas una comprensión de la situación y rápidamente hacer planes. Los expertos humanos pueden luego mejorar esos planes con algoritmos más detallados y consejos. Es una asociación, una simbiosis entre máquina y humano, cada uno complementando los puntos fuertes del otro, compensando las debilidades del otro.

La IA puede ser un salvavidas en tiempos de crisis, un socio que

nos ayuda a mantener la cabeza fuera del agua. Puede ayudarnos a procesar la avalancha de información, a encontrar claves que de otro modo pasaríamos por alto, a coordinar una respuesta rápida y efectiva. La IA por sí sola no es suficiente; requiere la guía de expertos humanos. Pero juntos, la IA y el experto humano pueden ser una combinación poderosa para ayudar a navegar incluso la peor de las tormentas. La IA es el salvavidas, y nosotros somos sus navegantes. Juntos, podemos llegar a puerto seguro.

En el corazón de Getafe, una vibrante ciudad al sur de Madrid, el constante zumbido del aire acondicionado es mi fiel compañero. Es una nana mecánica que ahoga el sofocante verano español del exterior, un fuerte contraste con el calor de otro tipo que se gesta dentro de los confines de mi oficina en la Universidad Carlos III de Madrid. Me encuentro en el epicentro de un equipo global de científicos. Nuestros miembros están dispersos por todo el mundo, desde los sagrados pasillos del MIT y Georgia Tech en Estados Unidos hasta los verdes paisajes estivales de la Universidad de Aalto en Helsinki, un lugar más comúnmente asociado con la nieve, pero que actualmente se deleita con el resplandor del sol de medianoche.

Nuestra exploración no es de paisajes físicos, sino del vasto e in-

trincado laberinto de la inteligencia artificial. Es un viaje que no exige resistencia física sino agilidad computacional, un agudo intelecto y una insaciable curiosidad. No estamos explorando la IA solo por el avance tecnológico. Nuestra misión es mucho más urgente, mucho más humana: estamos trazando los territorios inexplorados de la IA, buscando formas de aprovechar su inmenso potencial para ayudarnos a navegar las futuras crisis que la humanidad está destinada a enfrentar.

Desafíos humanos

Nuestra misión no está arraigada en la mera curiosidad académica. Es un clamor que resuena desde los desafíos urgentes y apremiantes con los que nuestro mundo está lidiando. El cambio climático, con su marcha implacable alterando paisajes y vidas; pandemias, barriendo naciones, dejando rastros de devastación; amenazas cibernéticas, invisibles pero insidiosas, socavando nuestra sensación de seguridad; disturbios sociales, el espíritu humano clamando contra la injusticia... La letanía de crisis que exigen una acción inmediata y coordinada está creciendo, cada una de ellas es un severo

recordatorio de la fragilidad e interconexión de nuestro mundo.

Nuestras instituciones de respuesta a emergencias, esos bastiones firmes diseñados para una era diferente se encuentran en terreno inestable. Son como antiguas fortalezas enfrentándose a la artillería moderna, sus altos muros y profundos fosos no son rival para el ritmo y la complejidad de los desafíos actuales. El mundo está cambiando a una velocidad asombrosa y un poco aterradora, y la complejidad de los problemas que enfrentamos es una red enmarañada que desafía a cualquier experto u organización a desentrañarla.

Nos encontramos al borde de este precipicio, mirando hacia el abismo de lo desconocido, y está claro: necesitamos ayuda. Necesitamos un nuevo tipo de aliado, uno que pueda mantener el ritmo del incansable avance del cambio, que pueda navegar la laberíntica complejidad de nuestros problemas, que pueda ayudarnos a trazar un rumbo a través de la tormenta. Necesitamos la IA.

En medio de una crisis, el tiempo es esencial. Cada segundo cuenta, cada decisión importa. Pero nuestra capacidad humana para procesar información, para dar sentido al caos, es limitada. Solo somos humanos, después de todo. Nos cansamos, tropezamos, cometemos errores. Y en una crisis, esas limitaciones pueden

tener consecuencias graves. Pero, ¿y si tuviéramos ayuda? ¿Qué pasa si tuviéramos un compañero que no se cansara, que pudiera procesar enormes cantidades de información en un abrir y cerrar de ojos, que nos ayudara a navegar la tormenta? Ahí es donde entra la IA.

Imagínate una crisis en desarrollo. La información llega desde todos los rincones, un torrente de datos que abruma por su volumen y complejidad. En medio del caos, la IA entra en acción. No es cualquier IA, sino un Modelo de Lenguaje Grande (MLG), una maravilla de la tecnología moderna que puede entender y generar lenguaje humano. Comienza a recopilar datos, cribando el ruido para encontrar la señal. Resume la información, destilando el caos en claridad. Interactúa con la gente, proporcionando actualizaciones, respondiendo preguntas, ofreciendo seguridad. Es un faro de calma en la tormenta, una mano firme en medio de la conmoción.

Pero el MLG no trabaja solo. Detrás de escena hay un equipo de expertos, trabajando con algoritmos complejos y herramientas de optimización. Son los arquitectos de la IA, los titiriteros que mueven los hilos. Toman la información recopilada por el MLG, los planes que ha hecho, y los refinan. Utilizan su experiencia, su comprensión de la situación, para afinar los re-

Nuestras instituciones de respuesta a emergencias, esos bastiones firmes diseñados para una era diferente, se encuentran en terreno inestable

sultados de la IA. Son el toque humano en la máquina, el corazón en la IA.

En una crisis, esta combinación de IA y experiencia humana puede cambiar el juego. La IA puede recopilar rápidamente toda la información de las personas, ofrecerles una comprensión de la situación y hacer planes rápidamente. Los expertos humanos pueden mejorar esos planes con algoritmos más detallados y asesoramiento. Es una asociación, una simbiosis de máquina y humano, cada uno complementando las fortalezas del otro, cada uno compensando las debilidades del otro.

En los anales de los desafíos tecnológicos, el Desafío de la Red DARPA 2009 ocupa un lugar especial. Fue una prueba que fue más allá del ámbito de la tecnología, adentrándose

en la esencia misma de la ingeniosidad, cooperación y habilidades de resolución de problemas humanos. El desafío era engañosamente simple en la superficie: localizar diez globos meteorológicos rojos dispersos por Estados Unidos continentales. Pero bajo esta simplicidad, se escondía una compleja red de desafíos que ponía a prueba los límites de la coordinación y colaboración humanas.

Los globos no estaban escondidos en lugares oscuros. Estaban a la vista, anclados en lugares públicos. Pero la escala misma de la tarea, la inmensidad del área dentro de la cual se dispersaron estos globos, hizo que el desafío fuera abrumador.

En muchos sentidos, el Desafío de la Red DARPA fue un micro- ▶▶▶

1 Subcontratación masiva voluntaria o colaboración masiva son alternativas en español al anglicismo *crowdsourcing*. <https://www.fundeu.es/recomendacion/crowdsourcing/#:~:text=Subcontrataci%C3%B3n%20masiva%20voluntaria%20o%20colaboraci%C3%B3n,en%20espa%C3%B1ol%20al%20anglicismo%20crowdsourcing.>

cosmos de las crisis que enfrentamos hoy. Al igual que los globos rojos, los problemas que enfrentamos (el cambio climático, las pandemias, las amenazas cibernéticas...) están esparcidos por todo el mundo. Son problemas complejos y multifacéticos que ningún individuo u organización puede resolver por sí solo. Requieren un esfuerzo coordinado y colaborativo. Nos exigen movilizar recursos, recopilar y analizar información, diseñar estrategias y adaptarnos a las circunstancias cambiantes. Nos exigen unirnos, combinar nuestros recursos y nuestro intelecto para resolver un problema que es más grande que cualquiera de nosotros.

En la actualidad, estamos realizando una simulación del Desafío de la Red DARPA para recrear este escenario en un entorno controlado. Los globos rojos son reemplazados por marcadores textuales, dispersos en un mar de datos. La tarea de la IA es generar estrategias para localizar estos marcadores, movilizar recursos y coordinar esfuerzos en una carrera contra el tiempo. Es una prueba de la capacidad de los Modelos de Lenguaje a Gran Escala (MGLs) para resolver problemas complejos del mundo real, trabajar en conjunto con los humanos y aumentar nuestras capacidades.

Los resultados de la simulación han sido a la vez iluminadores y sobrios, un testimonio del inmenso po-

tencial y las limitaciones inherentes de la IA. Resulta que los modelos son capaces de generar estrategias detalladas y matizadas. Pueden tamizar los datos, identificar patrones, anticipar posibles desafíos y proponer soluciones. Pueden pensar de manera estratégica, adaptarse a las circunstancias cambiantes y aprovechar el poder de la tecnología y el crowdsourcing¹. Pueden, en cierto sentido, imitar el proceso humano de resolución de problemas, pero a una velocidad y escala que supera con creces las capacidades humanas.

Herramientas

Pero la simulación también revela las limitaciones de estos modelos. Pueden procesar grandes cantidades de datos, pero carecen del toque humano, la capacidad de entender las sutilezas del comportamiento humano, la imprevisibilidad del mundo real. Pueden generar estrategias, pero necesitan la intervención humana para implementarlas, para adaptarlas a las realidades siempre cambiantes del mundo.

Pero estas limitaciones no son una condena a estos modelos, sino un recordatorio de su naturaleza. Son herramientas, no reemplazos del juicio e intuición humanos. Son ayudas, no sustitutos de la comprensión y empa-

tía humanas. A pesar de estas limitaciones, veo no solo desafíos, sino potencial. Veo un atisbo de un futuro en el que los MGL y los sistemas de IA autónomos y agéntico, como por ejemplo la novedosa plataforma AgentGPT, podrían convertirse en aliados invaluable en nuestra búsqueda para navegar las complejidades de nuestro mundo.

AgentGPT no es solo otro modo de IA. Es una nueva categoría de IA diseñada para el comportamiento autónomo dirigido hacia metas. Es como un jugador de ajedrez, siempre pensando varios movimientos por adelantado, siempre adaptando su estrategia en función del estado del juego. Pero a diferencia de un jugador de ajedrez, AgentGPT no está confinado a las 64 casillas de un tablero de ajedrez. Puede operar en el vasto y complejo paisaje de datos del mundo real, generando estrategias, anticipando desafíos y proponiendo soluciones.

El uso de la IA, particularmente en las crisis, es una espada de doble filo. Por un lado, puede ayudarnos a responder rápidamente y de manera efectiva, a movilizar recursos, a coordinar esfuerzos. Por otro lado, también puede ser una fuente de daño, si sus resultados son poco éticos o perjudiciales. El potencial de mal uso o abuso de la IA es una preocupación real y apremiante, a la que debemos hacer frente.

Pero las salvaguardas son solo una pieza del rompecabezas. También necesitamos empujar los límites de nuestras simulaciones, probar nuestros modelos bajo condiciones cada vez más desafiantes. Necesitamos exponer nuestra IA a las complejidades e incertidumbres del mundo real, ver cómo responde, entender sus fortalezas y debilidades.

Y necesitamos llevar nuestra investigación fuera del laboratorio y al mundo.

Estamos comprometidos a garantizar que cualquier uso de la IA para manejar las crisis se mantenga centrado en el humano. Estamos comprometidos con una visión de la IA que no sea una amenaza, sino un socio, una herramienta que pueda aumentar nuestras capacidades,

que nos pueda ayudar a dar sentido al caos, que pueda ayudarnos a encontrar la señal en el ruido. Estamos comprometido con una visión de IA que respete nuestros valores, nuestra ética, nuestra humanidad.

Y así, mientras el sol del verano español se pone fuera de la ventana de mi oficina, proyectando largas sombras sobre el cuidado campus de la Universidad Carlos III de Madrid, me encuentro lleno de un sentido de propósito, de resolución, de esperanza. Los desafíos que tenemos por delante son formidables, pero también lo es nuestra resolución.

Para la redacción de este trabajo se ha utilizado como apoyo editorial la plataforma GPT4 de OpenAI y la plataforma Claude+ de Anthropic.

Bibliografía

- Cebrián M. (2021): "A time-critical crowdsourced computational search for the origins of COVID-19" en *Nature Electronics*.
- Crandall J.W., Oudah M., Tennom F., Ishowo-Oloko F., Abdallah S., Bonnefon J.-F., Cebrián M., Shariff A., Goodrich M.A. and Rahwan I. (2018): "Cooperating with machines" en *Nature Communications*.
- Kryvasheyeu Y., Chen H., Obradovich N., Moro E., Van Hentenryck P., Fowler J. and Cebrián M. (2016): "Rapid assessment of disaster damage using social media activity" en *Science Advances*.
- Pickard, G., Pan, W., Rahwan, I., Cebrián, M., Crane, R., Madan, A. and Pentland, A. (2011): "Time-critical social mobilization" en *Science*.
- Rahwan, I., Cebrián, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., 'Sandy', Roberts, M. E., Shariff, A., Tenenbaum, J. B. and Wellman, M. (2019): "Machine behaviour" en *Nature*.
- Rutherford, A., Cebrián, M., Dsouza, S., Moro, E., Pentland, A. and Rahwan, I. (2013): Limits of social mobilization, en *Proceedings of the National Academy of Sciences of the USA (PNAS)*.