El futuro de los agentes de IA

Por Rogelio del Corral Lozano, Director de Seguridad Bancaria y socio de ADSI

La inteligencia artificial ha dejado de ser una promesa del futuro para convertirse en una realidad que moldea nuestras vidas. Pero entre todas las formas que adopta, los agentes de IA se destacan como los verdaderos protagonistas de esta revolución tecnológica. Estos sistemas avanzados no solo ejecutan tareas, sino que aprenden, se adaptan y toman decisiones con una autonomía que desafía nuestra condición como trabajadores y nuestra comprensión como profesionales.

Un agente de IA es más que un asistente digital. Basado en modelos de lenguaje masivos (LLMs), actúa en nombre del usuario para automatizar procesos complejos, interactuar con su entorno y resolver problemas en tiempo real. Si antes los *chatbots* eran simples respuestas preprogramadas, **los agentes de IA son capaces de planificar y ejecutar acciones con una precisión que va a transformar industrias enteras**.

Un agente de IA no es solo un ejecutor de tareas: **es un estratega digital**. Tiene la capacidad de descomponer problemas complejos en acciones concretas y planificar secuencias precisas para alcanzar objetivos específicos. A medida que interactúa con su entorno, **puede procesar y generar información en diversos formatos**, desde texto y audio hasta imágenes, abriendo la puerta a interacciones naturales y contextuales con los usuarios. Además, **estos agentes no son estáticos; evolucionan al almacenar y analizar datos de experiencias previas**, afinando su rendimiento con cada iteración.

Este aprendizaje continuo, combinado con su capacidad para integrarse con herramientas avanzadas como APIs y sistemas de bases de datos, los convierte en actores clave en entornos complejos donde se requiere eficiencia y adaptabilidad. **Lo más fascinante es su habilidad para trabajar de manera colaborativa**, coordinando acciones con otros agentes o sistemas de IA, lo que amplifica su impacto y les permite resolver problemas con una eficacia que trasciende la capacidad individual.

La evolución de los agentes de IA

Para entender cómo estos agentes se desarrollan y evolucionan, es crucial explorar **los marcos teóricos que proponen los principales actores como OpenAI y Google DeepMind**. China aparte, que ¡ya vendrá a darnos la sorpresa!

Sam Altman, CEO de OpenAI, sugiere una progresión de cinco niveles basados en su aplicación práctica. El trabajo comienza con los *chatbots* y culmina en "organizadores", sistemas capaces de gestionar

organizaciones enteras. Por otro lado, **Google DeepMind adopta un enfoque centrado en superar al hombre**, clasificando la IA según su rendimiento en comparación con las capacidades humanas, desde lo emergente hasta lo sobrehumano, capaz de realizar tareas mejor que todos los humanos, incluyendo habilidades como razonar, pensar y predecir. Es la Inteligencia Artificial General, AGI por sus siglas en inglés (Artificial General Intelligence).

En este contexto, **otros jugadores como Anthropic, con financiación de Amazon, y Grok, de Elon Musk, también entran en liza**. Anthropic se enfoca en la ética y la calidad de las interacciones, destaca en el análisis y resumen de documentos extensos y complejos y ofrece respuestas más naturales y creativas, similares a las de un humano, mientras que Grok, apadrinado por X, promete una IA que no solo aprende rápido, sino que también anticipa necesidades humanas.

Estas teorías compiten en una carrera tecnológica que no solo busca avances, sino también **redefinir qué significa ser inteligente** y cuáles serán las nuevas aptitudes mentales que diferencian al hombre de la máquina, porque el tema se está complicando. Intuitivamente, a mí me parece que **solo quedan algunos espacios reservados para el hombre y no sabemos por cuánto tiempo: abstracción, creatividad, conciencia y emociones**. De todo lo demás se ocuparán estas nuevas y sofisticadas herramientas que son los agentes IA: aprendizaje, adaptabilidad, resolución de problemas, razonamiento lógico, comunicación, memoria, autonomía y toma de decisiones.

No hay que esperar mucho. Ya existen cientos de casos reales y **cada día veremos nuevos ejemplos de agentes de IA que trabajan y transaccionan con humanos y con otros agentes de IA**: gestores de fondos de inversión, abogados, tuiteros, *gamers*, *copywriters*, agentes de seguridad y ciberseguridad, testeadores de software, educadores, creadores de contenido, analistas de laboratorio y farmacéuticos, *call centers* que atienden 20.000 llamadas a la hora en más de 150 idiomas, servicios completos de atención al cliente que coordinan reuniones y generan avisos, programadores, agentes de investigación de mercados, verificadores de identidad, validadores de documentos...

Difícil de entender y difícil de medir: los 'benchmarks'

Para medir la eficacia de estos modelos, los *benchmarks* se convierten en el campo de pruebas definitivo. Herramientas como GLUE, SuperGLUE y MMLU evalúan habilidades de comprensión y razonamiento, mientras que Big-Bench desafía la creatividad y la moralidad de las IA. En estas comparativas, modelos como GPT-4 de OpenAI sobresalen en generación de lenguaje fluido, mientras que Gemini y Claude destacan en tareas específicas como traducción y computación avanzada.

Solo este asunto de la medición, clasificación y comparativa de modelos daría sustancia para la elaboración de varias tesis doctorales, pero si alguien se anima, que sea rápido porque **todo puede cambiar con la llegada de los nuevos modelos LCM que introduce Meta** desde algún recóndito laboratorio, posiblemente dirigido por al menos un cerebro humano.

Mientras que los LLM procesan texto a nivel de *tokens* (palabras individuales), los LCM operan con conceptos completos, que pueden corresponder a oraciones enteras o ideas abstractas, y esto con independencia del idioma y del tipo de información, texto, audio o imagen.

Pasaremos del entrenamiento de los modelos IA con cantidades ingentes de datos, refinamiento e iteraciones múltiples a la IA de Aprendizaje, que se propugna mediante el conjunto de problemas Abstraction and Reasoning Corpus (ARC) que evalúan el desarrollo de los modelos en referencia a las capacidades que todavía nos permiten diferenciarnos de las máquinas:

- **Comprensión abstracta.** Ser capaz de identificar conceptos generales, como "simetría", "continuidad", o "repetición".
- **Razonamiento intuitivo**. Determinar las relaciones lógicas entre los elementos y cómo transformarlos.
- **Aplicación creativa**. Utilizar las reglas descubiertas en nuevos casos, aun cuando no sean casos idénticos.

La IA del descubrimiento es el próximo reto tecnológico en la evolución de los Modelos de IA. Tirar más de la intuición, la razón y la creatividad que de la memoria bruta, y hacerlo con la misma cantidad de datos o información que es capaz de hacerlo el ser humano.

Más allá de los resultados obtenidos, la existencia de estas pruebas, el afán por medir y comparar, son reflejo de la **inversión masiva que las empresas están canalizando en sus equipos de desarrollo**. La batalla está en su pleno fragor y solo hay cabida para gigantes. **OpenAI y Google lideran esta carrera** con amplios recursos financieros y con acceso a talento de élite, mientras que Meta aboga por proyectos Open Source, y otros proyectos más emergentes como Grok adoptan enfoques ágiles y disruptivos.

IA y 'blockchain': una convergencia inevitable

Los agentes IA van ligados a cada uno de estos modelos y presentan, en consecuencia, diferentes capacidades y velocidades de desarrollo, pero son ya una realidad y tenemos innumerables ejemplos prácticos que auguran una expansión rápida e imparable, **invadiendo todas las actividades humanas, ya sean de tipo industrial, educativo, salud, finanzas, ocio o seguridad**.

Los componentes precisos son los mismos que hemos necesitado siempre los humanos para llevar a término nuestra actividad. Los Agentes IA tienen capacidades de percepción para entender entradas en formato texto, imágenes, audio o datos en tiempo real. Tienen capacidad de decisión para aplicar modelos y generar patrones y acciones basadas en las entradas, y además, **se pueden integrar con otros sistemas, hardware o software, incluso con otros agentes** mediante APIs, interfaces lógicos y físicos y comunicaciones con latencia imperceptible. Ahora lo hacen las máquinas por nosotros, con supervisión humana, pero cada vez lo hacen mejor. Pronto la supervisión será innecesaria o, en todo caso, se realizará a la inversa.

Consecuencia de estas capacidades de los agentes de IA se produce necesariamente una integración con sistemas de infraestructura descentralizada. **La intersección de los agentes IA con *blockchain* y Web3 marca el comienzo de una nueva era**. Estos agentes, como hemos visto, pueden operar autónomamente, pero lo pueden hacer además como "trabajadores digitales" que gestionan tareas mediante contratos inteligentes, automatizando procesos y recibiendo pagos en criptomonedas.

Vinculando todo esto con el **ámbito de la seguridad corporativa**, imagina **un ecosistema donde un agente analiza videograbaciones para detectar, corregir o notificar anomalías técnicas** y cobra de inmediato por su eficaz tarea. Imagina sistemas de videovigilancia que no solo detectan amenazas, sino que también envían alertas automáticas a agentes de IA que reaccionan en segundos, o programas de formación de seguridad que utilizan simulaciones inmersivas para entrenar a usuarios o empleados. Otro agente de IA gestiona transacciones financieras con una escrupulosa verificación de los datos aportados por el ordenante, su identidad digital y su historial para prevenir fraudes y aplicar un estricto cumplimiento de la normativa sobre blanqueo de capitales.

Eficacia y productividad en una permanente progresión ascendente, sin limitaciones de tipo laboral, fisiológico, o emocional. **Un ecosistema exclusivamente gobernado por las leyes del mercado, la oferta y la demanda, que necesariamente conlleva a la progresiva reducción de los costes de producción**. Un ecosistema que además tiene sus cimientos en la descentralización y la 'tokenización', permitiendo que la propiedad de los agentes y sus beneficios se distribuyan de manera proporcional y equitativa a la contribución y el mérito de sus propietarios.

No abordaré aquí el asunto de la energía por no alargar más el texto, pero creo sinceramente que una vez que superemos la histeria político-mediática del ecologismo, **nuestra propia responsabilidad colectiva nos encaminará hacia fuentes de energía perdurable y barata**, empezando con las centrales nucleares tradicionales, y continuando en el siguiente lustro con un modelo innovador como el que propone la startup americana Deep Fission para suministrar energía a centros de datos y otras instalaciones mediante microreactores nucleares enterrados a una milla de profundidad. Y quizás, en las próximas generaciones llegarán al escalón definitivo con la conquista de la energía ilimitada y casi gratuita: la Fusión Fría: el amanecer de la energía eterna.

Cuando la IA toca el cerebro: Neuralink y el futuro de la interacción humana

Elon Musk no solo ha puesto su sello en la carrera por la IA con Grok. Es también el **artífice de otra fusión, en este caso entre la tecnología y la biología, con Neuralink**. Este proyecto, como otros similares, se encuentra en sus inicios, pero ya con los permisos clave y autorizaciones regulatorias necesarias en Estados Unidos y Canadá, donde se han realizado ensayos clínicos en tres personas afectadas por severas lesiones de médula espinal.

Dependiendo de la evolución en los próximos años de estas interfaces cerebro-computadora, no sería muy descabellado **anticipar una perfecta simbiosis humano-máquina con la aparición de implantes diseñados para interactuar mediante comandos mentales** o el simple pensamiento con un ejército de agentes de IA al servicio del hombre.

¡La polémica está servida! Todos iguales, un mundo sin trabajo. Pensionista desde que naces y rey desde que te implantan la interfaz. ¿Es extraño por inconcebible? ¿Irrealizable por idílico? Ni idea. Yo soy creyente y siempre digo que hay algo más.

Algo sí tengo muy claro. Sobre todo, si exploramos la evolución de nuestra forma de interactuar con las computadoras desde los años 60. En un principio, con las tarjetas perforadas; después, la línea de comandos de MS dos "C:\>", que todavía tenemos en el recuerdo, y ahora, con las ventanas y los clics del ratón sobre la pantalla y los interfaces táctiles. **Con la llegada de la IA todo cambia y la interfaz de usuario, tal y como la conocemos, está a punto de prejubilarse**.

La inteligencia artificial, el *blockchain* y la neurotecnología están redefiniendo los límites de lo posible. **Estas herramientas abren las puertas de un futuro lleno de eficiencia y automatización, pero también generan nuevos desafíos**, incertidumbres y riesgos. Quizás pasemos a un plano secundario y seamos felices, pero dependientes y esclavos.

Y mientras nos preparamos para este salto exponencial, no olvidemos un detalle crucial: asegurarnos de rellenar los libros catálogo de medidas de seguridad convenientemente diligenciados, en soporte de papel reciclable, con tinta no tóxica y una caligrafía impecable, sin tachaduras, ni raspaduras. ¡Los agentes no perdonarían este fallo! Porque **el futuro puede ser digital, pero nunca debemos subestimar el poder de lo analógico**. ■