



ORGANIZACIONES CENTAURO

PLAYBOOK 1 IA AGÉNTICA



SEPTIEMBRE 2025

LA LEY



Corvalán, Juan G.

Organizaciones centauro : playbook 1 IA agéntica / Juan G. Corvalán ; Mariana Sánchez Caparrós ; Carina Mariel Papini. - 1a ed. - Ciudad Autónoma de Buenos Aires : La Ley, 2025.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-03-5041-5

1. Derecho. I. Sánchez Caparrós, Mariana II. Papini, Carina Mariel III. Título

CDD 340.1

AUTORES

Juan G. Corvalán
Mariana Sánchez Caparrós
Carina Mariel Papini

COLABORADORES

Marilina Borneo, Sofía Tammaro,
Marina Petliuk Diaz, Aylén Irrera Martino,
Carolina Martín, Matías Calderini, Martina Nuccitelli,
Felicitas Rodríguez Contreras, Denise Mester

DISEÑO

Sofía Rolleri, Victoria Mafud





IALAB

ubatec^{SA}

AI academy
BY doinGlobal

LA LEY

 Thomson
Reuters

Organizaciones Centauro

Playbook 1: IA agéntica

Presentación de la serie de Playbooks UBA IALAB

En apenas dos años, pasamos de interactuar con asistentes conversacionales multipropósito a experimentar agentes de IA y flujos agénticos que son capaces de tomar decisiones, ejecutar tareas complejas y aprender de la interacción con su entorno de forma autónoma. Los “autos autónomos” de las tareas y procesos humanos.

Para entender mejor el impacto de este fenómeno en las empresas y en las organizaciones, desde UBA IALAB presentamos una serie de playbooks que son guías conceptuales y metodológicas de aplicación práctica. Elegimos como título “centauros”, porque se trata del nacimiento de una simbiosis en el modo producir y desarrollar tareas: humano-agéntica por diseño, para reconfigurar un nuevo tipo de convivencia organizacional.

Esta serie de playbooks no está dirigida exclusivamente a expertos técnicos. Está pensada para líderes institucionales, responsables de innovación, equipos operativos y todo aquel que quiera “aterrizar” los agentes impulsados por IA generativa y la incipiente IA agéntica en entornos reales, sin perder de vista los desafíos éticos, organizacionales y culturales que eso implica.

También desarrollamos un análisis del estado del arte orientado a lograr entender cómo las diferentes plataformas que se desarrollaron en los últimos dos años combinan software “clásico” basado en automatizaciones con agentes de IA cada vez más autónomos.

Esta combinación creciente, se basa en la incorporación de herramientas específicas, comandos dinámicos y capacidades multimodales de texto, voz, imagen, pantalla y cámara, lo que amplía significativamente las posibilidades de implementación en múltiples industrias, áreas del conocimiento y dominios organizacionales.

**Bienvenidas y bienvenidos al primer paso hacia el diseño
de organizaciones que piensan y actúan a partir de la
símbiosis entre los humanos y la IA agéntica**

CONTENIDO

01. De asistentes conversacionales a la IA agéntica en las organizaciones centauro	08
02. Agentes de IA aumentados: el salto evolutivo de la IA generativa	11
03. La arquitectura de la IA agéntica	12
04. Estado del arte de los agentes de IA: funcionalidades emergentes y limitaciones operativas	16
05. Nuevos escenarios de la IA agéntica: NotebookLM y el Agente de ChatGPT	38
06. Limitaciones de la IA agéntica en las organizaciones: autonomía restringida en flujos complejos	49
07. Conclusiones	51



Resumen ejecutivo

La inteligencia artificial agéntica constituye un nuevo paradigma frente a los tradicionales asistentes conversacionales. Mientras que estos últimos se limitan a operar de forma reactiva, respondiendo a instrucciones en lenguaje natural, la IA agéntica habilita un ecosistema donde múltiples agentes interconectados perciben el entorno, planifican, actúan de manera proactiva y toman decisiones autónomas orientadas al cumplimiento de objetivos complejos.

En este marco, consideramos que es preciso distinguir entre **agente de IA e IA agéntica: el agente de IA representa el “qué” y la IA agéntica, en cambio, refiere al “cómo”**. Como bien explicaremos en el presente playbook, un agente de IA es un ejecutante autónomo; la IA agéntica, la orquesta que permite la sinfonía.

En este contexto, los agentes de IA se presentan como unidades funcionales dotadas de componentes estructurales que les permiten percibir su entorno, ejecutar tareas complejas, interactuar con herramientas externas, aprender de forma continua y operar bajo lógicas modulares. Su integración en entornos organizacionales plantea nuevos desafíos en términos de trazabilidad, supervisión, arquitectura y rediseño de flujos.

Este playbook caracteriza en detalle dichos componentes, presenta una arquitectura comparada de agentes de IA, y evalúa plataformas como NotebookLM y el Agente ChatGPT mediante una matriz de madurez agéntica elaborada por UBA IALAB, abordando capacidades como autonomía, proactividad, adaptabilidad, multimodalidad, uso de herramientas y gestión de memoria.

En definitiva, las organizaciones que incorporen estos sistemas deberán rediseñar sus procesos para que los agentes actúen bajo principios de coherencia institucional, con supervisión significativa y auditoría operativa. El usuario deja de ser un mero ejecutor de comandos para convertirse en curador, diseñador y auditor de sistemas autónomos. Esta nueva forma de simbiosis humano-máquina configura el núcleo del presente playbook y traza el horizonte operativo de las organizaciones centauro.

01

De los asistentes conversacionales a la IA agéntica en las organizaciones centauro

Los agentes de IA se consolidan como el nuevo núcleo operativo de las organizaciones centauro. Este modelo organizacional propone una colaboración simbiótica entre humanos y agentes inteligentes, inspirado en la figura híbrida del centauro, pero enriquecido con un enfoque responsable y adaptable a las realidades de las organizaciones y culturas de Latinoamérica.

Las organizaciones centauro se estructuran en torno a principios de IA centrada en el humano, trazabilidad, visualización de procesos, validación humana y generación de resultados agénticos explicables. En ellas, los agentes no solo automatizan tareas, sino que se integran activamente en los flujos de trabajo como “digital workers” estratégicamente diseñados, capaces de aprender, “razonar” y decidir junto a los humanos, potenciando la toma de decisiones, la eficiencia operativa y la adaptabilidad institucional.

Por un lado, el avance exponencial de la IA generativa nos sitúa en un escenario que va más allá de la idea de “automatización inteligente”. El paradigma tecnológico que se avizora se resume en un concepto: **la IA agéntica**.

Este nuevo paradigma redefine drásticamente los roles tradicionales que se pensaron cuando la IA todavía no había acelerado en capacidades. En los estudios que publicamos hace más de 5 años, la idea de la automatización inteligente se basaba en una reconfiguración de los procesos y de las tareas, pero el humano era el impulsor esencial de los mismos. **Frente a la IA agéntica, pasamos de ser productores de tareas y procesos, a ser directores de agentes, editores, auditores y supervisores de flujos agénticos.**

Realizando una analogía con los autos autónomos, podríamos decir que la automatización inteligente que se planteó en los inicios de esta década, tenía más que ver con incorporar funciones de automatización al vehículo: asistencia de conducción, cámara con sensores trasera o suspensión asistida. De cara a un auto autónomo grado 4 o 5, cambia sustancialmente la intervención humana y nuestro rol. Se incrementa el componente autónomo robusto basado en la autogestión, planificación y ejecución de tareas.

La IA agéntica se basa en sistemas con agencia, esto es, con capacidad de percibir el entorno, actuar de forma proactiva y tomar decisiones autónomas para cumplir con objetivos complejos. A diferencia de la IA generativa “ChatGPT 2022” que respondía de manera reactiva a un prompt específico, **un agente de IA opera como un delegado o representante digital**.

Se le confía una meta y el agente tiene un determinado margen de decisión para planificar y ejecutar las tareas necesarias para completarla. Los límites de ese margen se ampliaron y, con ello, los desafíos y riesgos vinculados a la intervención humana.

Por ejemplo, en lugar de ejecutar la orden de “programar una publicación” (automatización), un agente puede recibir el objetivo de “maximizar el engagement de una campaña”. Para ello, analiza tendencias, propone el mejor contenido y horario, redacta borradores de texto, genera imágenes y luego de una validación humana, acciona la publicación y también aprende de los resultados para futuras acciones.

En definitiva, es esta cualidad agéntica autónoma y orientada a metas la que habilita la colaboración simbiótica que define a las organizaciones centauro, que se organizan en de IA a plataformas y sistemas que cuentan con múltiples agentes de IA que se orquestan para conformar un ecosistema agéntico. Ahora bien, es crucial hacer una distinción para evitar una confusión común: aunque están íntimamente relacionados, los términos “agente de IA” e “inteligencia artificial agéntica” no son sinónimos.

Agente de IA: es el modelo, lo que actualmente se conoce como LLM o grandes modelos de lenguaje. Por ejemplo, Grok 4, Gemini 2.5 pro y 4o de ChatGPT son los modelos más avanzados y pertenecen a la familia de lo que se conoce como “IA generativa”. **El término “agente de IA” suele usarse como sinónimo de IA generativa. Sin embargo, en sentido estricto, para que se pueda calificar a un modelo de “agente”, debería contar con un grado de autonomía relevante para resolver problemas y ejecutar tareas.**

Es decir, la diferencia entre un asistente basado en IA generativa y un agente de IA, reside en el grado de autonomía y autogestión para cumplir objetivos y realizar tareas con poca o nula intervención humana.

Por eso **la IA generativa puede manifestarse en modo asistente o en modo agente**. Un ejemplo de “modo asistente” es cuando usamos GPT-5 para que nos ayude a modificar un email o para que nos redacte un documento. Sin embargo, cuando al completar la tarea el propio modelo toma la salida y nos sugiere agregados, adaptaciones o mejoras sin haberlas requerido, o cuando activamos en estas herramientas la función de “investigación en profundidad”, emerge el “modo agente”.

En el modo agente se incrementa la autonomía para realizar diversas acciones sin intervención humana directa. En concreto, por ejemplo, en el caso de la búsqueda profunda, la generación de un plan de investigación, realizar una búsqueda exhaustiva en diversas fuentes que son seleccionadas por el agente, luego el análisis y la posterior elaboración de un informe con sus fuentes que se entrega al usuario.

La IA agéntica: es el paradigma, las cualidades o el conjunto de capacidades que hacen que se cree un **ecosistema agéntico**, en donde la autonomía y la proactividad se presentan por capas o grados. Si volvemos a la analogía con los vehículos autónomos, la IA agéntica es el conjunto de condiciones habilitantes para que se produzca la conducción autónoma grado 4 o 5, que es la que permite que un auto tome una porción significativa de decisiones en la conducción sin intervención humana.

La IA agéntica no es el auto en sí, sino que se trata del conjunto de tecnologías que permite resolver múltiples tareas o flujos de trabajo. Y aquí, usualmente, participan varios agentes con diferentes componentes herramientas y funciones. En síntesis, es la capacidad que adquiere el sistema en su conjunto cuando se combinan diferentes agentes de IA y se ocupan de diversos flujos de trabajo.

Un agente de IA es un especialista en tocar de manera autónoma un instrumento, mientras que la IA agéntica es la orquesta en su conjunto

Una organización centauro no se basa en cualquier tipo de agentes; en esencia, se fundamenta específicamente en aquellos diseñados bajo una lógica agéntica, pues son estos -por ahora- los únicos capaces de alcanzar el nivel de colaboración simbiótica que debemos transitar para transformar radicalmente las empresas y organizaciones.

En pocas palabras...

La IA agéntica no es una simple evolución de los asistentes conversacionales: representa un cambio de paradigma en el rol humano dentro de las organizaciones.

De operar herramientas, pasamos a orquestar agentes. Su capacidad para planificar, actuar y aprender de forma autónoma redefine la colaboración humano-máquina y convierte a los agentes en piezas centrales de un nuevo modelo organizacional: la organización centauro.

02 Agentes de IA aumentados: el salto evolutivo de la IA generativa

En los últimos dos años y diez meses la inteligencia artificial ha experimentado una transformación fundamental. Se inició con la democratización del acceso a herramientas reactivas (ej. ChatGPT sustentado en el modelo GPT-3.5) para migrar hacia ecosistemas agénticos impulsados por agentes de IA.

El término “agentes de IA” define el capítulo más reciente y transformador del libro que relata la evolución de la inteligencia artificial. Los agentes de IA representan el salto desde los asistentes conversacionales multipropósito, que se limitaban a responder consultas o generar texto bajo una lógica esencialmente reactiva, a una nueva categoría de soluciones de IA generativa que deja muy corto al término “crecimiento exponencial”.

Los agentes de IA aumentados se distinguen por la incorporación de funcionalidades más sofisticadas que nos proyectan hacia un “modo Netflix”. Esta expresión alude a la capacidad de los agentes de mantener memoria de corto y largo plazo, acceder a herramientas externas e internet, desplegar habilidades multimodales avanzadas, realizar investigaciones profundas y planificar tareas complejas¹.

Los agentes de IA funcionan como un núcleo o “sistema operativo” para la acción. Su diseño no se limita a procesar un prompt y devolver una respuesta; están concebidos para entender la intención profunda del usuario, planificar una secuencia de acciones, simular razonamientos complejos y ejecutar dichas acciones para alcanzar un objetivo, incluso si los pasos intermedios no fueron explícitamente solicitados.

En síntesis, encarnan un nuevo paradigma definido por conocimientos avanzados, conexión en tiempo real con el mundo, anticipación, planificación, proactividad y una autonomía creciente.

En pocas palabras...

Los agentes de IA representan la evolución más profunda de la IA generativa: ya no reaccionan, actúan. Con memoria, herramientas y planificación, dejan atrás al asistente conversacional para convertirse en verdaderos operadores digitales. Son el nuevo “sistema operativo” de la IA, capaces de interpretar intenciones, ejecutar flujos complejos y adaptarse al entorno en tiempo real.

¹ Corvalán J., Sanchez Caparrós M., “Agentes de inteligencia artificial y workflows agénticos La nueva frontera de la automatización” - 1a ed - Ciudad Autónoma de Buenos Aires : La Ley, 2025.

03 La arquitectura de la IA agéntica

Introducción: la clave del ecosistema modular

La eficacia de un agente de IA no reside en un único modelo de IA generativa, sino en su ecosistema modular. Esta arquitectura integra diversas capacidades y herramientas especializadas que operan en conjunto para dar vida a un colaborador digital con una significativa agencia y autonomía.

Antes de comenzar a enumerar los componentes de la infraestructura de estos agentes, consideramos fundamental señalar que tanto sus funcionalidades más complejas como las más simples, operan dentro de un entorno técnico más amplio. La IA generativa en modo agente y en modo asistente, están embebidas con funcionalidades clásicas de “software contenedor”.

Por ejemplo, no hay IA cuando hacemos clic en adjuntar un archivo, subir una imagen, activar triggers o un flujo automatizado. Se trata de funciones de software “clásico” embebidos en una plataforma que convive con asistentes y agentes de IA. Esta coexistencia técnica es la que permite a los agentes desplegar su potencial.

A

Componentes de los agentes de IA aumentados

El núcleo del agente, estrictamente, suele ser un modelo de lenguaje de gran capacidad (LLM) que actúa como el centro neurálgico para la toma de decisiones. **Su función trasciende la generación de texto; ha evolucionado para poder coordinar, simular razonamiento avanzado, planificar la estrategia a seguir, seleccionar y aplicar las herramientas adecuadas, activar otros modelos generativos (por ej. de imágenes) y gestionar el flujo de datos para lograr los objetivos.**

El modelo de lenguaje es quien asume el rol de director y ejecutor de un conjunto de tareas para llegar a un objetivo. Continuando con el ejemplo del agente de investigación profunda o deep research. Supongamos que el usuario activa este “modo agente” con una simple instrucción: “realiza un informe completo sobre el impacto de la IA agéntica en las empresas de Latinoamérica”.

A

Aunque pueda darnos una serie de sugerencias o preguntas para ajustar mejor la estrategia, el agente de IA llevará adelante las acciones de planificación, dirección, análisis y ejecución de la tarea.

Ahora bien, el modo agente se activa en función de dos clases de modelos que están disponibles en las plataformas de OpenAI, Gemini, Claude, Grok, Deepseek, entre otras:

- **Modelos de generación conversacional:** como los GPT clásicos, optimizados para la fluidez, coherencia y naturalidad en el diálogo.
- **Modelos optimizados para razonamiento:** como los de la familia “o” de OpenAI (o3 y o4), y más recientemente, GPT-5 thinking, o las versiones más avanzadas de Gemini, que están específicamente diseñados para el razonamiento estructurado, la toma de decisiones lógicas y la ejecución de tareas, permitiéndoles una mayor capacidad para resolver problemas complejos.

B

Memoria persistente (de sesión)

Este componente es la memoria de trabajo del agente. Le permite mantener un registro del contexto, los detalles y el flujo de una interacción continua. También puede sostener conversaciones extendidas con coherencia, recordar datos mencionados párrafos atrás y ajustar su comportamiento en tiempo real dentro de una única sesión, evitando la “amnesia contextual” que caracterizaba a modelos anteriores.

C

Memoria de largo plazo (transversal y de perfil)

Acá está la clave para el agente “modo Netflix”: esta memoria habilita al modelo a recordar información clave sobre el usuario como sus preferencias, historial, datos personales relevantes a través de múltiples sesiones y a lo largo del tiempo. Es lo que permite que el agente “te conozca”, se anticipa a tus necesidades basándose en interacciones pasadas y construya una relación de colaboración continua, en lugar de empezar de cero en cada conversación.

D

Acceso a herramientas (brazos y sentidos en el mundo digital)

Esta es la capacidad del agente para actuar más allá de su propia de IA de conocimientos. Mediante el acceso a herramientas, puede:

- **Navegar** por internet para obtener información en tiempo real.
- **Interactuar** con APIs (Interfaces de Programación de Aplicaciones) de terceros para realizar acciones concretas (reservar un vuelo, agendar una reunión, comprar un producto).
- **Analizar** archivos adjuntos en múltiples formatos (documentos de texto, hojas de cálculo, código, archivos de audio, imágenes y video).
- **Ejecutar** código en un entorno seguro para realizar análisis de datos, simulaciones o resolver problemas matemáticos.

Aquí podemos ver que las funcionalidades tradicionales de software “contenedoras” de la IA que se integran al flujo de trabajo en el cual este se encuentra inmerso el mismo.

En este sentido, la inteligencia artificial generativa, por sí sola, no es suficiente para realizar íntegramente ciertas tareas o diversos resultados. Es imprescindible que el modelo se conecte e interactúe con otras funcionalidades del software tradicional: son estas herramientas las que permiten que la IA adquiera capacidades operativas extendidas, acceda a información dinámica y actúe en el mundo digital con mayor autonomía y eficacia.

E

Capacidades multimodales (percepción holística)

Los agentes de IA son **nativamente multimodales**, lo que significa que pueden entender, procesar, relacionar y generar información a través de diversos formatos de manera fluida. No solo procesan texto, sino que pueden “ver” el contenido de una imagen, “escuchar” una instrucción por voz, interpretar datos en una tabla y generar una respuesta que combine todos estos elementos. Esta fusión de canales sensoriales digitales les otorga una comprensión del mundo mucho más rica y similar a la humana.

F

Mecanismos de planificación y razonamiento de tareas

Esta es quizás la capacidad más definitoria de un agente. En lugar de abordar un problema de forma monolítica, utilizan sofisticados mecanismos de planificación como cadena de pensamiento (“Chain-of-Thought”) o árbol de pensamiento (“Tree of Thoughts”) para descomponer un objetivo complejo y ambiguo en una serie de subtareas más pequeñas, lógicas y manejables. **Esto les permite operar con una autonomía mucho mayor, y también pueden abordar problemas que requieren estrategia, corrección de errores y una ejecución por etapas.**

En pocas palabras ...

La capacidad del agente de IA para utilizar herramientas de manera asertiva y eficiente, sumado a la de mantener una memoria contextual y de largo plazo, son los pilares que lo elevan de una interfaz conversacional a una aplicación con posibilidad de “comprender”, “razonar”, “responder”, “actuar” y “aprender”.

El “**modo Netflix agéntico**” es clave porque permite seguir el historial de conversaciones del usuario y utilizarlo bajo una lógica agéntica, redefiniendo la interacción humano-máquina.

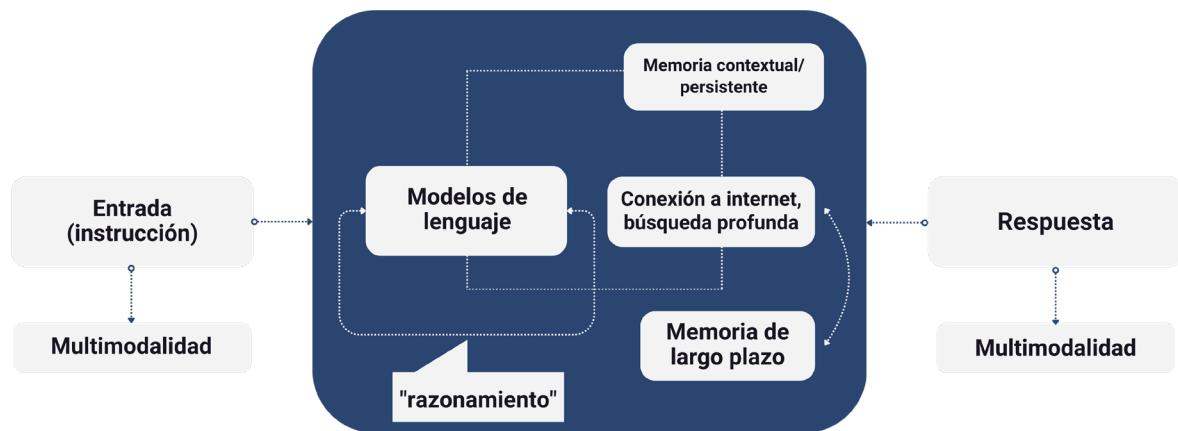


Figura 1. Estructura de un agente de IA

04 Estado del arte de los agentes de IA: funcionalidades emergentes y limitaciones operativas

En febrero de 2025, al lanzar el libro “**Agentes de inteligencia artificial y workflows agénticos: la nueva frontera de la automatización**”, desde UBA IALAB anticipamos que se consolidarían en los meses siguientes dos tendencias clave. Una de ellas preveía la aparición de agentes de IA cada vez más equipados con herramientas avanzadas, orientadas a potenciar su autonomía y capacidad de adaptación².

En la actualidad, observamos cómo estas proyecciones se han hecho realidad a través de desarrollos tangibles, evidentes en casi todas las plataformas de acceso individual, como ChatGPT, Gemini, Grok y Copilot. Veamos:

A

Memoria de largo plazo, memoria persistente avanzada y personalización

La memoria es una herramienta esencial para que los agentes de IA retengan el contexto y se adapten a tareas en curso o históricas. Se distinguen principalmente dos tipos de memoria relevantes de cara a una mayor agencia de los agentes de IA:

- **La memoria a corto plazo**, persistente o de contexto, que rastrea la “línea de pensamiento” del agente de IA y sus acciones recientes para preservar el contexto dentro del flujo de trabajo actual; y,

² Corvalán & Sánchez Caparrós, 2025, p. 13.

A

- **La memoria a largo plazo** o memoria personalizada, retiene interacciones históricas e información relevante del usuario, lo que permite una comprensión contextual más profunda y una mejor toma de decisiones a lo largo del tiempo. La memoria a largo plazo es crucial para que los agentes de IA se adapten a las preferencias de los usuarios, permitiendo que las interacciones sean más personalizadas³.

Como exponente de la transición de la IA conversacional al agente de IA, podemos mencionar especialmente el caso de ChatGPT, que desde abril de 2025 puede referenciar todas las conversaciones pasadas del usuario, permitiendo respuestas más personalizadas y coherentes a lo largo del tiempo. El modelo toma decisiones sobre cuándo y cómo incorporar estos recuerdos para adaptar sus respuestas de forma autónoma⁴.

Adicionalmente, ChatGPT incorpora funciones específicas que potencian **la personalización de la experiencia del usuario**. Por un lado, a través de la opción “personalizar”, es posible definir una serie de parámetros que permiten que el modelo ajuste su comportamiento y estilo de respuesta de manera más precisa. Los usuarios pueden indicar cómo desean ser llamados, cuál es su ocupación, y qué tipo de tono, enfoque o nivel de detalle esperan en las respuestas. Esta configuración inicial actúa como una capa de IA de adaptación, favoreciendo una interacción más contextualizada y alineada con las preferencias individuales o profesionales del usuario.

Por otro lado, Chat GPT ha incorporado “proyectos”, una funcionalidad pensada para organizar y profundizar el trabajo con contextos específicos predefinidos por el usuario. A través de dicha funcionalidad los usuarios pueden agrupar conversaciones, archivos, documentos y tareas bajo un mismo eje temático o de trabajo, lo que permite que el agente retenga no solo interacciones aisladas, sino una estructura más compleja de información persistente.

3 Loucks, J., Crossan, G., Sarer, B., Widener, C., & Bucaille, A. (2024, 19 de noviembre). Autonomous generative AI agents. Deloitte Insights. <https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2025/autonomous-generative-ai-agents-still-under-development.html>

NVIDIA. (n.d.). What are AI Agents?. NVIDIA Glossary. Recuperado de <https://www.nvidia.com/en-us/glossary/ai-agents/>

4 OpenAI. (2025, 3 de junio). Memory FAQ. OpenAI Help Center. https://help.openai.com/en/articles/8590148-memory-faq#h_3319d9d65b

A

Esta capacidad potencia el aprendizaje progresivo del agente, que accede a memorias temáticas y contextos ampliados sin necesidad de repetir indicaciones en cada sesión. **Al vincularse con la memoria de largo plazo, los proyectos permiten que el modelo reconozca objetivos, etapas previas, decisiones tomadas y preferencias expresadas dentro de un mismo proceso.** Esto es clave y útil para el desarrollo de productos, investigaciones, estrategias institucionales o procesos legales complejos, en los que la continuidad contextual y la adaptación a largo plazo son clave. En combinación con las funciones de personalización y los archivos adjuntos, los “proyectos” configuran un entorno de trabajo que va más allá de la simple conversación, consolidando a los agentes de IA como asistentes inteligentes capaces de gestionar procesos con memoria activa y orientados a objetivos.

Gemini también ha implementado mejoras significativas en su capacidad de recordar preferencias y datos proporcionados en interacciones anteriores. Avances recientes incluyen la ampliación de la ventana de contexto de los modelos Gemini 2.5 (hasta 1 millón de tokens), lo que facilita inherentemente una memoria más extensa dentro de interacciones prolongadas y permite al modelo mantener la coherencia y el contexto sobre grandes cantidades de información sin necesidad de recordatorios constantes por parte del usuario⁵.

A su vez, Google ha puesto a disposición del usuario el modelo experimental Gemini con personalización. A través de su activación, Gemini puede aprovechar datos del usuario, como el historial de búsqueda, la actividad en Gmail, Calendar o Drive, para generar **respuestas más contextuales, pertinentes y alineadas con sus intereses y hábitos digitales.** Esta integración profunda permite que la inteligencia artificial actúe con mayor precisión y proactividad dentro del entorno cotidiano del usuario, anticipando necesidades y ofreciendo asistencia ajustada al momento y al contexto. En los próximos meses, Gemini planea ampliar su capacidad y conectarse con otras apps y servicios de Google, como Fotos, con el fin de ofrecer información más personalizada, basándose en una comprensión más amplia de sus actividades y preferencias⁶.

5 Kavukcuoglu, K. (2025, 25 de marzo). Gemini 2.5: Our most intelligent AI model. Google Official Blog. <https://blog.google/technology/google-deepmind/gemini-model-thinking-updates-march-2025/>

6 Citron, D., Pichai, S., Porat, R., Walker, K., & Manyika, J. (2025, 13 de marzo). Gemini gets personal, with tailored help from your Google apps. Google Official Blog. <https://blog.google/products/gemini/gemini-personalization/>

Plataforma	ChatGPT (OpenAI)	Gemini (Google)
Tipo de memoria y personalización	Memoria de corto y largo plazo + personalización	Memoria contextual extendida personalización conectada
Características clave	Recuerda interacciones pasadas, adapta respuestas en función del contexto y preferencias del usuario	Memoria de largo plazo, "personalizar", "proyectos"
Funcionalidades destacadas	Permite configurar tono, estilo, ocupación, nombre	Organización temática de archivos y conversaciones
Aplicaciones recomendadas	Memoria de largo plazo, "personalizar", "proyectos"	Personalización con datos de Gmail, Calendar, Drive
	Organización temática de archivos y conversaciones	Plan de integración futura con apps como Google Fotos
	Procesos complejos, investigaciones, trabajo legal, desarrollo de productos	Asistencia cotidiana, automatización contextualizada, personalización basada en hábitos
	Trabajo sostenido con continuidad contextual	Entornos integrados Google, adaptación proactiva a actividades y preferencias

Cuadro resumen - Comparativa de memoria y personalización en agentes de IA
(ChatGPT y Gemini a julio 2025)

B

GPTs personalizados, Gems y Agentes en Copilot

Las plataformas continúan ampliando las capacidades de personalización y despliegue para que los usuarios creen asistentes cada vez más especializados. Esta personalización se realiza editando preferencias, que se introducen en lenguaje natural o que el propio modelo de IA se ocupa de realizar.

Los usuarios pueden crear versiones personalizadas de ChatGPT sin necesidad de programar, definiendo su personalidad, estilo y flujos de trabajo. Estas versiones, conocidas como **GPTs personalizados**, retienen información de sesiones previas, adaptándose progresivamente a las necesidades y preferencias del usuario⁷.

Además de definir instrucciones e iniciadores de conversación, permiten **integrar archivos, enlaces y APIs externas, lo que transforma a cada GPT en una unidad modular con potencial para interactuar con sistemas utilizados en las organizaciones**. Chat GPT puede asistir al usuario en la creación de los GPTs personalizados, a través del comando “crear”, mediante preguntas sobre los objetivos y el tipo de uso que desea realizar.

Por su parte, los Gems de Google, integrados en plataformas como Gemini, comparten un propósito similar: ofrecer IA adaptada a necesidades específicas, con capacidad para aprender de interacciones pasadas. A diferencia de los GPTs, que se centran en el ecosistema de ChatGPT, **los Gems se integran con los servicios de Google**, como Gmail, Calendar o Drive, habilitando un nivel superior de automatización y asistencia contextualizada⁸.

Por último, Gemini también ofrece una opción para asistir a los usuarios en la creación de los Gems a través de un botón simbolizado con una varita mágica. Si el usuario hace clic en ella, optimiza el prompt donde el usuario escribió las instrucciones para el Gems, de manera automática.

7 OpenAI. (n.d.). GPTs FAQ. OpenAI Help Center. <https://help.openai.com/en/articles/8554407-gpts-faq>

8 Google. (n.d.). Gemini Gems — build custom AI experts from Gemini. Google Gemini. <https://gemini.google/overview/gems/>

Microsoft, mediante Copilot presenta una opción que ha denominado “agentes personalizados”, una especie de GPT personalizado o Gem en el entorno de Microsoft 365. Antes de avanzar, puede ser útil aquí, puntualmente, volver sobre la diferenciación que realizamos entre asistentes de IA generativa y agentes de IA. En muchos casos, las herramientas que se nos ofrecen son asistentes, aunque se los designe “agentes”, porque no presentan autonomía y capacidad para dirigir, ejecutar y planificar tareas.

B

Dentro del ecosistema Microsoft 365, los usuarios pueden diseñar asistentes personalizados para usarlos como herramientas estratégicas para la productividad empresarial. Hay algunos que están predefinidos en distintas plantillas como una asistente de marco de trabajo ágil (scrum) o asistente de escritura, entre otros. Esta versatilidad permite adaptar las soluciones a las necesidades específicas de cada equipo o proyecto⁹.

La capacidad del agente de IA para usar herramientas con eficacia y mantener memoria contextual y de largo plazo lo transforma de un simple interlocutor en una aplicación capaz de comprender, razonar, responder, actuar y aprender

⁹ Microsoft. (n.d.). Get started with agents in Microsoft 365 Copilot. Microsoft Support. <https://support.microsoft.com/en-us/topic/get-started-with-agents-in-microsoft-365-copilot-943e563d-602d-40fa-bdd1-dbc83f582466>

Característica	ChatGPT (GPT-4o y GPT-5)	Gemini (Veo 3 y Flow)	Grok 4 (xAI)
Modalidades integradas	GPTs	Gems	Agentes personalizados
Capacidades destacadas	Análisis visual contextual a través de cámara Interacción en tiempo real con señales no verbales (modo de voz avanzado); generación de imágenes con DALL-E 3	No requieren programación IA adaptada a tareas específicas, aprende del historial de uso, integración nativa con servicios de Google	No requieren programación Agentes activos en el ecosistema Microsoft 365, con plantillas o personalización propia; pueden ser compartidos hacia el interior de la organización
Casos de uso ideales	API externas, enlaces, archivos	Gmail, Calendar, Drive	Word, Excel, Teams, Outlook, SharePoint
Nivel de agencia multimodal	Automatización de tareas, soporte técnico, asistentes conversacionales	Asistencia contextualizada, gestión de agenda, redacción asistida	Gestión de proyectos, productividad empresarial, colaboración interna
Integraciones clave	DALL-E 3 (imágenes), Modo de Voz Avanzado	Veo 3 (video), Imagen (imagen), Flow (fusión y orquestación)	Aurora (imágenes), Grok Vision (visión), plataforma X (información en tiempo real)

Cuadro resumen - Comparativa de asistentes personalizados en ChatGPT, Gemini y Copilot
(a julio 2025)

C

Modelos actualizados: o3 y o4-mini, GPT-4.1 y GPT-4.1 mini, el gigante GPT-5, Gemini 2.5 Pro y 2.5 Flash

En abril de 2025, OpenAI presentó **o3 y o4-mini**, los modelos más avanzados de su serie “o”, diseñados para realizar un razonamiento más profundo antes de generar respuestas. Estos modelos representan un salto significativo en las capacidades de ChatGPT, ofreciendo mejoras sustanciales tanto para usuarios ocasionales como para investigadores avanzados¹⁰:

10 OpenAI. (2025, 16 de abril). Presentamos OpenAI o3 y o4-mini. OpenAI. <https://openai.com/es-ES/index/introducing-o3-and-o4-mini/>

C

Ambos modelos, o3 y o4-mini, destacan por su capacidad de “pensar con imágenes”. Esto implica que pueden analizar y razonar sobre contenido visual, como fotografías, diagramas o gráficos, integrándolos en su cadena de pensamiento para ofrecer respuestas más precisas y contextualizadas¹¹.

Sin embargo, **o3 es el modelo más potente**, optimizado para tareas complejas que requieren un razonamiento profundo, como programación avanzada, matemáticas y análisis científico. En el caso del **modelo o4-mini**, se trata de una **versión más ligera y rápida**, diseñada para tareas que requieren eficiencia y velocidad sin comprometer la calidad. Es ideal para aplicaciones que necesitan respuestas rápidas en áreas como cálculos matemáticos, asistencia en programación y análisis de datos visuales¹². También es de utilidad para tareas sencillas, de dificultad baja donde el usuario prefiere velocidad frente a profundidad o razonamiento avanzado y respuestas extensas o detalladas.

En mayo de 2025, OpenAI lanzó también **GPT-4.1 y GPT-4.1 mini**, disponibles para suscriptores Plus, Pro y Team. Estos modelos ofrecen mejoras significativas en codificación, seguimiento de instrucciones y manejo de contextos largos, con una ventana de contexto de hasta 1 millón de tokens. Es decir, pueden procesar información extensa y mantener coherencia narrativa sin asistencia del usuario, lo que mejora su autonomía en tareas complejas y permite, por ejemplo, que el modelo estructure informes o análisis en múltiples etapas, sin requerir intervención constante para mantener el hilo lógico o temático.

Finalmente, el 7 de agosto de 2025, OpenAI lanza GPT-5, modelo insignia de la plataforma que busca sustituir los anteriores en el entorno de su plataforma ChatGPT. GPT-5 se ofrece en esta plataforma como un sistema unificado que integra:

- Un **modelo rápido y eficiente** para la mayoría de las consultas (“GPT-5 Fast”).

11 Idem

12 Idem

C

- Dos **modelos de razonamiento** rápido y profundo ("GPT-5 Thinking" y "GPT-5 Thinking mini") para problemas complejos.
- Un **enrutador** en tiempo real que decide cuál utilizar según la naturaleza de la conversación, la complejidad, las herramientas requeridas y la instrucción explícita del usuario (p. ej., "piensa bien en esto").

Este diseño busca brindar un equilibrio entre velocidad y profundidad, con un uso más eficiente de tokens y un rendimiento superior incluso frente a modelos especializados como OpenAI o3 en tareas de razonamiento visual, codificación agencial y problemas científicos avanzados¹³.

GPT-5 está disponible para todos los usuarios, con diferentes volúmenes de uso según el plan, mientras que los suscriptores Pro acceden además a GPT-5 Pro, con razonamiento avanzado para respuestas más completas y precisas¹⁴. En la misma línea, durante el primer semestre de 2025, Google introdujo importantes actualizaciones en su familia de modelos Gemini:

Gemini 2.5 Pro avanzó el 17 de junio de 2025 con capacidades de "razonamiento" complejo, codificación y manejo de contextos extensos; en concreto, hasta 1 millón de tokens. También avanzó en la capacidad experimental "*Deep Think*", que permite al modelo realizar investigaciones de alto nivel, considerando múltiples hipótesis antes de formular una respuesta, lo que refuerza su autonomía en tareas analíticas profundas. Gemini 2.5 Pro también ha mostrado un rendimiento destacable en benchmarks¹⁵ de matemáticas, ciencias y codificación.

Gemini 2.5 Flash: una versión optimizada para velocidad y eficiencia, también actualizada el 17 de junio de 2025, que ofrece mejoras en razonamiento, multimodalidad y codificación, con una reducción en el consumo de tokens. Está diseñado para tareas complejas que requieren respuestas rápidas.

13 Ver <https://openai.com/index/gpt-5-system-card/> y https://openai.com/index/introducing-gpt-5/?utm_source=chatgpt.com

14 Idem

15 Los benchmarks (o pruebas de referencia) son conjuntos estandarizados de tareas o métricas diseñadas para evaluar y comparar el rendimiento de distintos sistemas, productos o modelos (especialmente en los campos de la inteligencia artificial y la computación) bajo condiciones uniformes. Estos permiten determinar cuál es más eficaz o preciso en una tarea específica, garantizando una comparación justa. Una analogía útil para comprender este concepto es la de una carrera deportiva, donde todos los corredores compiten en la misma pista y bajo las mismas reglas, lo que permite medir objetivamente quién es el más rápido.

Ambos modelos Gemini 2.5 soportan entrada audiovisual y salida de audio nativa, permitiendo interacciones conversacionales más naturales y ricas, con control sobre el tono, acento y estilo de habla.

Modelo	Descripción general	Capacidades destacadas	Casos de uso recomendados
o3 (OpenAI)	Modelo más potente de la serie "o"	Razonamiento profundo, "pensar con imágenes", análisis científico y programación avanzada	Tareas complejas, análisis técnico, matemáticas avanzadas
o4-mini (OpenAI)	Versión optimizada para velocidad y eficiencia	Procesamiento visual, cálculo, programación, razonamiento ágil	Aplicaciones que requieren rapidez sin perder precisión
GPT-4.1 (OpenAI)	Modelo con contexto extendido y mayor autonomía	Seguimiento de instrucciones, coherencia en textos largos (hasta 1 millón de tokens)	Redacción de informes extensos, análisis en múltiples etapas
GPT-4.1 mini	Variante más liviana de GPT-4.1	Menor consumo de recursos, buena en tareas secuenciales	Uso general, asistencia técnica y narrativa sostenida
Gemini 2.5 Pro	Modelo avanzado con razonamiento complejo y capacidades de investigación profunda	"Deep Think", codificación, matemáticas, hipótesis múltiples, 1M tokens	Investigación, análisis científico, tareas cognitivas de alto nivel
Gemini 2.5 Flash	Versión veloz y eficiente de Gemini 2.5	Multimodalidad, codificación, razonamiento rápido, bajo consumo de tokens	Tareas técnicas exigentes con necesidad de respuesta inmediata
Ambos Gemini 2.5	Actualizaciones integradas en 2025	Soporte para entrada audiovisual y salida de audio con control de tono y estilo	Interacciones conversacionales naturales, experiencias multimodales enriquecidas

Cuadro resumen - Comparativa de modelos actualizados de OpenAI y Google (a julio 2025)

D

Agentes de IA multimodales: ChatGPT, Gemini y Grok redefinen la interacción con la IA

GPT-4o y GPT-5 de OpenAI integran de forma nativa texto, imágenes, audio y video. Ambos modelos permiten una comprensión y generación fluida entre diferentes formatos, ofreciendo respuestas más naturales y contextuales que enriquecen las interacciones. Y puntualmente, GPT-4o, con un tiempo de respuesta a entradas de audio de tan sólo 232 milisegundos y un promedio de 320 milisegundos, similar al de una conversación humana, es ideal para interacciones en tiempo real, como asistentes de voz o diálogos dinámicos¹⁶.

A partir de mayo de 2025, se encuentra disponible el “modo de voz avanzado” de ChatGPT, que ofrece conversaciones más naturales y en tiempo real que captan y responden con emociones y señales no verbales. Este modo de voz está predeterminado en todos los planes de pago de ChatGPT y se reducirá al “modo de voz estándar” si se alcanzan los límites de uso o si no se cumplen los requisitos de calidad de conexión o hardware¹⁷.

En esencia, mediante el comando “dictar”, el usuario puede elaborar prompts a través de su voz y el modelo le contesta por escrito. Adicionalmente, el botón “usar modo de voz”, tanto en versión móvil como en versión para ordenadores, permite al usuario hablar con el modelo y mantener una conversación fluida. Cada diálogo se transcribe automáticamente, y el historial queda disponible para su consulta posterior.

Por otro lado, la integración de **ChatGPT con DALL-E 3** permite generar imágenes de alta calidad a partir de texto, con mejoras en la renderización de texto dentro de las imágenes, lo que lo hace especialmente útil para proyectos creativos que requieren precisión visual.

Gemini Live de Google¹⁹ avanzó para potenciar la experiencia multimodal. Permite tener conversaciones de voz continuas, con solicitudes y respuestas habladas, manteniendo un diálogo natural.

16 OpenAI. (2024, 8 de agosto). GPT-4o System Card. OpenAI. <https://openai.com/es-419/index/gpt-4o-system-card/>

17 OpenAI. (2025, mayo). <https://help.openai.com/en/articles/9617425-advanced-voice-mode-faq>

18 OpenAI. (2024, 8 de agosto). GPT-4o System Card. OpenAI. <https://openai.com/es-419/index/gpt-4o-system-card/>

19 Ver más en Prakhar Khanna, How To Use Gemini Live's Real-Time Screen Sharing And Camera Features, Forbes, 21 de abril de 2025, disponible en: How To Use Gemini Live's Real-Time Screen Sharing And Camera Features

D

Además, el usuario puede compartir la cámara de su celular en tiempo real²⁰.

Esto significa que puede apuntar la cámara del celular y Gemini puede identificar los objetos, leer los documentos físicos y brindar respuestas u orientación sobre lo que registra y procesa de manera instantánea. Adicionalmente, Gemini live permite que el usuario comparta la pantalla de su dispositivo en vivo para colaborar a partir del procesamiento de la imagen que detecta la IA, en tiempo real.

Con respecto a las imágenes, **Gemini de Google** ha avanzado significativamente en la generación de contenido audiovisual con **Veo 3**, lanzado en mayo de 2025. Este modelo produce videos de alta calidad en resolución hasta 4K a partir de prompts textuales o imágenes, incorporando audio sincronizado de forma nativa, como diálogos, música y efectos de sonido. Veo 3 destaca por su comprensión avanzada de instrucciones complejas, física del mundo real y sincronización labial, permitiendo crear narrativas visuales dinámicas y coherentes.

Además, recientemente lanzó **Flow**, una herramienta profesional que fusiona Veo, Imagen y Gemini, permitiendo generar escenas detalladas a partir de descripciones de elenco, ubicación y estilo visual, dirigida a cineastas, creadores de contenido y publicistas, abriendo nuevas posibilidades para la producción de medios de alta calidad²¹.

Grok 4 de xAI. En el marco de la plataforma Grok, se anunció recientemente la llegada del modelo Grok 4, en julio del 2025. Este cambio representa un avance significativo respecto a Grok 3, el modelo que se lanzó el 17 de febrero de 2025, que mejora las capacidades multimodales de su predecesor, que incluyen procesamiento de texto, imágenes y, desde abril de 2025, visión en tiempo real a través de Grok Vision.

Mientras que Grok 3 permite analizar objetos, documentos y entornos a través de la cámara de un smartphone, respondiendo preguntas como “¿qué estoy viendo?” con

20 Ver más en Isha Sheth, 5 ways to use Gemini Live with camera and screen sharing, Blog Google, 7 de abril de 2025, disponible en: How to use Gemini Live with camera and screen sharing

21 Google. (n.d.). Flow - Google Labs. Google Labs. <https://labs.google/flow/about>

D

una comprensión avanzada del contexto espacial, Grok 4 mejora esta capacidad con un procesamiento más rápido y preciso, integrando un sistema multiagente (Grok 4 Heavy) que combina múltiples perspectivas para una interpretación más detallada del entorno visual.

Esto permite a Grok 4 no solo identificar objetos, sino también realizar inferencias más complejas sobre interacciones entre ellos, como predecir movimientos o analizar dinámicas en tiempo real.

En cuanto a la generación de imágenes, Grok 3 utiliza “**Aurora**”, un modelo flexible para crear contenido visual a partir de prompts textuales, destacándose por su libertad creativa, aunque con controversias por posibles usos éticos cuestionables. Grok 4, por su parte, introduce una versión mejorada de Aurora, con mayor calidad en la generación de imágenes y una capacidad ampliada para manejar prompts más complejos, incluyendo la generación de contenido visual multimodal que combina texto, datos estructurados e imágenes de referencia.

Además, Grok 4 está diseñado para integrarse con herramientas externas, como navegadores web, para validar y optimizar los resultados visuales en tiempo real, reduciendo errores en comparación con Grok 3.

Adicionalmente, la integración con la plataforma X, que permite a Grok 3 acceder a datos en tiempo real para ofrecer respuestas actualizadas y contextuales, ideal para análisis de tendencias o asistencia en tiempo real, se mantiene en Grok 4. Sin embargo, Grok 4 optimiza esta funcionalidad con un motor de búsqueda en tiempo real más robusto y una arquitectura de mezcla de expertos (*Mixture of Experts*), lo que mejora la relevancia y precisión de las respuestas, especialmente en escenarios dinámicos como el análisis de noticias o la monitorización de eventos en curso.

Por último, Grok 4 incorpora una capacidad computacional significativamente mayor, entrenado con el superordenador Colossus de xAI (200,000 GPUs H100), lo que representa

22 Mixture of Experts (MoE): arquitectura de IA que combina varios modelos especializados (“expertos”) y un mecanismo de enrutamiento que selecciona cuáles se activan según la entrada, lo que permite mayor capacidad con menor costo computacional.

D

un aumento de 10 veces respecto a Grok 3, permitiendo un manejo más eficiente de grandes volúmenes de datos en tiempo real.

Estos desarrollos reflejan una aceleración sin precedentes en la evolución de los modelos multimodales, que no solo procesan texto, imagen, audio y video, sino que también integran capacidades avanzadas de “comprensión” contextual, “razonamiento” espacial y generación “creativa”. **En conjunto, OpenAI, Google y xAlestán dando forma a una nueva generación de agentes de IA multimodales, cada vez más autónomos, sensibles al entorno y preparados para interactuar en tiempo real con las personas y con el mundo digital.**

Lejos de tratarse solo de mejoras incrementales, **estos avances consolidan un nuevo estándar de experiencia e interacción con la inteligencia artificial, en el que la generación de contenido, la asistencia personalizada y la toma de decisiones se integran de forma fluida en múltiples formatos.** Esta convergencia transforma la manera en que se produce y consume información y sienta las bases para el despliegue de agentes capaces de operar en entornos complejos, con una creciente capacidad de agencia, adaptación y coordinación.

Los últimos avances en IA marcan un nuevo estándar multimodal, donde agentes capaces de comprender, razonar y crear en múltiples formatos se consolidan como interlocutores autónomos, adaptativos y preparados para interactuar en tiempo real con personas y entornos digitales

Característica	ChatGPT (GPT-4o y GPT-5)	Gemini (Veo 3 y Flow)	Grok 4 (xAI)
Modalidades integradas	Texto, imagen, audio, video	Procesamiento visual, cálculo, programación, razonamiento ágil	Aplicaciones que requieren rapidez sin perder precisión
Capacidades destacadas	Modelo con contexto extendido y mayor autonomía	Seguimiento de instrucciones, coherencia en textos largos (hasta 1 millón de tokens)	Redacción de informes extensos, análisis en múltiples etapas
Casos de uso ideales	Variante más liviana de GPT-4.1	Menor consumo de recursos, buena en tareas secuenciales	Uso general, asistencia técnica y narrativa sostenida
Nivel de agencia multimodal	Modelo avanzado con razonamiento complejo y capacidades de investigación profunda	"Deep Think", codificación, matemáticas, hipótesis múltiples, 1M tokens	Investigación, análisis científico, tareas cognitivas de alto nivel
Integraciones clave	Versión veloz y eficiente de Gemini 2.5	Multimodalidad, codificación, razonamiento rápido, bajo consumo de tokens	Tareas técnicas exigentes con necesidad de respuesta inmediata

Cuadro resumen - Comparativa de modelos actualizados de OpenAI y Google
(a julio 2025)

Herramienta	ChatGPT modo voz	Gemini Live	Google AI Studio
Canal de interacción	Voz (dictado y conversación fluida con transcripción automática, disponible en móvil y escritorio)	Voz, cámara en tiempo real y pantalla compartida	Voz y pantalla compartida dentro del navegador, con posibilidad de interacción en tiempo real con documentos del usuario
Capacidad agéntica	Moderada: responde con precisión en tiempo real, pero no interpreta imágenes ni entornos externos	Alta: interpreta imágenes en vivo, voz y pantalla compartida para dar orientación contextualizada	Alta: permite interacción directa con documentos y feedback instantáneo desde el entorno del usuario
Valor diferencial	Permite mantener conversaciones naturales por voz con registro textual completo, útil para accesibilidad y repaso	Integra canales multimodales simultáneos: permite mostrar objetos, documentos o compartir pantalla para recibir ayuda contextual	Facilita sesiones de trabajo colaborativo con Gemini, permitiendo críticas y devoluciones sobre materiales concretos
Limitaciones observadas	No procesa imágenes ni pantalla compartida; interacción limitada al canal verbal	Disponible en dispositivos con sistema operativo Android e IOS necesita conexión estable	Requiere configuración previa y conocimientos básicos de integración para aprovechar su potencial al máximo

Cuadro resumen - Comparativa de herramientas agénticas para investigación
(a julio 2025)

El “modo agente” para investigación profunda y procesamiento de datos

ChatGPT Deep Research, desarrollado por OpenAI e impulsado por el modelo o4-mini, ha ampliado notablemente sus capacidades para el análisis avanzado de datos. Esta herramienta permite a los usuarios cargar conjuntos de datos, ejecutar y visualizar código Python directamente desde la interfaz, lo que facilita tareas como la depuración, la generación de gráficos y la manipulación de datos en tiempo real.

También hemos destacado el modo “investigación en profundidad” disponible para suscriptores de ChatGPT Pro, con un límite inicial de 250 ejecuciones mensuales. Su carácter agéntico se manifiesta en la capacidad de gestionar múltiples entradas como datasets, funciones y variables y seleccionar automáticamente los enfoques más adecuados en función de los objetivos inferidos del usuario²³.

E

Ya hemos destacado la funcionalidad de Gemini, denominada Deep Research, que permite explorar de forma automática cientos de sitios web a solicitud del usuario, analizar la información recolectada y generar, en pocos minutos, informes extensos y con alto nivel de profundidad. Con la incorporación del modelo mejorado Gemini 2.5, esta herramienta ha optimizado su desempeño en todas las etapas del proceso investigativo, desde la planificación inicial hasta la generación de entregables detallados, consolidándose como una solución integral para la investigación asistida por inteligencia artificial²⁴.

Desde UBA IALAB hemos utilizado la funcionalidad Deep Research de Gemini como asistente en la elaboración de 55 reportes con propuestas de casos de uso de agentes basados en IA generativa para las industrias del petróleo, el gas y el agua.

En nuestra experiencia, observamos que cuando se le proporciona un prompt detallado y específico, esta herramienta puede generar informes de entre diez y quince páginas, con buena claridad expositiva y múltiples referencias.

23 OpenAI. (2025, 24 de abril). Introducing deep research. OpenAI. <https://openai.com/index/introducing-deep-research/>

24 Google. (n.d.). Gemini Deep Research — your personal research assistant powered by Gemini. Google Gemini. <https://gemini.google/overview/deep-research/?hl=es>

E

No obstante, en todos los casos fue imprescindible realizar una revisión humana exhaustiva, tanto para verificar la precisión técnica como para mejorar redacción, estructura y coherencia argumentativa.

Una de las principales limitaciones encontradas fue el sistema de citación: si bien el modelo incluye referencias, estas se presentan al final del documento como un listado general, sin formato de notas al pie. Además, los números de referencia que aparecen dentro del texto no siempre coinciden con los del listado final, lo que exige una reelaboración manual para garantizar la trazabilidad y fiabilidad de las fuentes.

Grok 3 de xAI, por su lado, también introduce capacidades avanzadas de investigación profunda y análisis de datos, integradas con la plataforma X para acceder a información en tiempo real. Su función DeepSearch, que se activa mediante un botón en la interfaz, permite a Grok realizar búsquedas iterativas en la web, verificando hechos antes de generar reportes detallados con citas.

Estas herramientas marcan un punto de inflexión en la investigación asistida por IA generativa, ya que permiten a estudiantes, profesionales, académicos, docentes y técnicos acceder a procesos de análisis automatizado cada vez más potentes, versátiles y contextualmente pertinentes. Su valor diferencial radica en su capacidad agéntica, es decir, en la posibilidad de actuar como agentes autónomos o semiautónomos que no solo procesan datos, sino que pueden planificar búsquedas, seleccionar fuentes relevantes, verificar información, inferir objetivos y generar entregables de manera proactiva. Esta agencia transforma radicalmente el rol del usuario, que pasa a convertirse en un supervisor estratégico del proceso de investigación.

Aunque persisten desafíos vinculados a la trazabilidad de fuentes, la calidad expositiva y la necesidad de revisión humana, el despliegue de estas funciones de búsqueda profunda y análisis multimodal abre nuevas posibilidades para el desarrollo de investigaciones complejas, ágiles y altamente personalizadas. En este sentido, nos encontramos ante una evolución significativa del paradigma de búsqueda tradicional hacia un modelo de colaboración con agentes inteligentes.

Herramienta	Tipo de análisis	Capacidad agéntica	Ventaja principal	Limitaciones observadas
ChatGPT Deep Research	Carga y procesamiento de datos (Python), análisis estructurado	Alta: selecciona funciones, variables y enfoque según objetivos del usuario	Permite ejecutar código Python y visualizar resultados en tiempo real	Requiere conocimientos básicos de análisis de datos
Gemini Deep Research	Exploración web automatizada, generación de informes	Alta: interpreta el <i>prompt</i> , navega, selecciona fuentes y redacta informes	Genera informes largos a partir de <i>prompts</i> detallados	Sistema de citación poco confiable y requiere revisión humana
NotebookLM	Lectura, análisis y consulta sobre documentos propios	Moderada: responde solo con base en documentos cargados, sin interferencia externa	Ideal para investigación documental centrada en el corpus del usuario	No incorpora fuentes externas, depende del material cargado
Grok 3 + Deep Research	Búsqueda iterativa en tiempo real, verificación de hechos	Alta: decide qué buscar, verifica información y genera reportes con citas	Acceso directo a datos actualizados desde la plataforma X	Dependencia del entorno X, limitado a temas con presencia en redes

Cuadro resumen - Comparativa de capacidades de voz, audio y compartir pantalla de ChatGPT y Gemini (a julio 2025)

F

Agentes de IA que realizan búsquedas: navegación contextual en tiempo real

La funcionalidad **SearchGPT** permite a ChatGPT acceder a información actualizada en la web, mostrar citas verificables, incluir enlaces directos y ofrecer respuestas basadas en datos en tiempo real dentro del chat.

Lo verdaderamente innovador no es solo la capacidad de búsqueda, sino el modo en que el modelo gestiona esa tarea de forma autónoma: interpreta la intención de la consulta, decide qué términos utilizar, qué fuentes priorizar y cómo sintetizar la información obtenida.

F

El usuario no conduce el proceso paso a paso; al activar SearchGPT, delega la agencia investigativa en el modelo, lo que consolida a ChatGPT como un **agente con capacidad de operar en entornos abiertos e inciertos**.

El modelo puede decidir por sí mismo cuándo buscar en la web, según la naturaleza de la pregunta, o bien se puede activar manualmente mediante el ícono de búsqueda (globo terráqueo).

Las respuestas incluyen enlaces a las fuentes utilizadas, lo que favorece la trazabilidad y permite profundizar en los temas de interés. Además, se pueden realizar preguntas de seguimiento, ya que ChatGPT conserva el contexto completo de la conversación, optimizando la coherencia y precisión de las respuestas.

La plataforma de Elon Musk, Grok, también incorpora navegación web en tiempo real, con una ventaja estratégica: su integración directa con la plataforma X (ex Twitter), que le permite acceder a flujos de datos actualizados al instante. A diferencia de SearchGPT, cuyo enfoque se basa en la interpretación autónoma de consultas web generales, **la fortaleza de Grok reside en su capacidad para capturar tendencias, opiniones y eventos en curso en redes sociales**, lo que lo convierte en una herramienta especialmente valiosa para el monitoreo contextual, el análisis de discursos públicos y la detección de narrativas emergentes.

Esta función, como ya mencionamos, se encuentra mejorada a partir de Grok 4 que optimiza esta funcionalidad con un motor de búsqueda en tiempo real más robusto y una arquitectura de Mixture of Experts, lo que mejora la relevancia y precisión de las respuestas en comparación con Grok 3.

Las funciones de navegación web en tiempo real representan un paso clave hacia modelos agénticos con mayor autonomía, capaces de operar más allá de sus datos entrenados y actualizar su conocimiento de forma dinámica.

Tanto SearchGPT como Grok extienden los límites de la IA generativa al incorporar estrategias de consulta autónoma y acceso a fuentes externas, pero cada uno lo hace desde una lógica distinta: SearchGPT actúa como un agente de búsqueda generalista con capacidad de interpretar y ejecutar procesos

F

investigativos complejos, mientras que Grok se posiciona como un observador privilegiado del presente, con acceso directo a flujos conversacionales en redes sociales. En ambos casos, las decisiones sobre cómo buscar dejan de recaer en el usuario final y son asumidas por el agente según el contexto de uso.

Característica	SearchGPT (ChatGPT)	Grok (xAI)
Herramienta	Manual (ícono) o automática según la consulta	Activa en tiempo real de forma continua
Tipo de fuentes	Web general (noticias, artículos, sitios confiables)	Plataforma X (red social en tiempo real) y web general a partir de Grok 4
Capacidad de agencia	Alta: interpreta intención, selecciona fuentes y sintetiza	Alta: accede y filtra eventos, opiniones y tendencias
Trazabilidad de fuentes	Alta: incluye enlaces y citas verificables	Media: acceso directo a X, pero sin referencias externas con excepción de Grok 4
Contexto de uso ideal	Búsquedas complejas, reportes, investigación temática	Análisis de coyuntura, redes, seguimiento de eventos
Ventaja diferencial	Interpretación autónoma y respuestas con fuentes confiables	Acceso privilegiado a datos actualizados y tendencias virales

Cuadro resumen - Comparativa entre SearchGPT y Grok (a julio 2025)

G

Google AI Studio

Esta plataforma es una de las más versátiles de todo el ecosistema. Presenta múltiples funcionalidades, herramientas que permiten crear música, videos, imágenes, aplicación, manipular código a través de lenguaje natural, lo que se conoce como Vibe coding.

Diseñada para ser un entorno integrado y accesible, Google AI Studio permite experimentar, construir y desplegar soluciones basadas en IA sin necesidad de conocimientos avanzados en programación. A través de una interfaz visual e interactiva, presenta modelos de última generación como Gemini y Gemma, capaces de procesar lenguaje natural. La plataforma ofrece capacidades multimodales robustas, combinando texto, imagen, audio y video en un mismo flujo de trabajo, lo que la convierte en una herramienta ideal tanto para desarrolladores como para diseñadores, educadores, creadores de contenido y equipos empresariales. Además, permite ajustar parámetros clave como temperatura, tokens máximos, seguridad, ofreciendo un control detallado sobre el comportamiento de los modelos en cada tarea.

Asimismo, facilita la creación de aplicaciones personalizadas y su integración con entornos externos; una galería de prompts preconfigurados para agilizar el desarrollo; y funciones avanzadas de interacción en vivo como micrófono, cámara, voz sintética y diálogo afectivo. AI Studio permite desplegar proyectos directamente en plataformas como Vertex AI, Cloud Run o Kaggle, lo que asegura escalabilidad y continuidad en entornos reales. En conjunto, la plataforma no solo democratiza el acceso a la IA generativa, sino que habilita un nuevo estándar de creación tecnológica que combina potencia, flexibilidad y usabilidad en un solo lugar.

05 Nuevos escenarios de IA Agéntica: NotebookLM y el Agente de ChatGPT

La IA agéntica representa una nueva etapa en la evolución de los modelos generativos, caracterizada por su capacidad para ejecutar tareas complejas con autonomía, combinando razonamiento, acción instrumental y uso de herramientas externas. En este apartado se destacan dos desarrollos clave: NotebookLM y el reciente Agente de ChatGPT.

a. NotebookLM

Por un lado, tenemos por parte del ecosistema de Google a NotebookLM, una plataforma diseñada como asistente de lectura y análisis que combina el potencial de Gemini con múltiples funcionalidades orientadas a la visualización, investigación, manipulación y diversas funcionalidades impulsadas por IA generativa.

Permite al usuario cargar documentos, seleccionar fuentes específicas y formular preguntas directamente sobre esos contenidos, generando respuestas en texto basadas exclusivamente en el material proporcionado. Además, ofrece herramientas como mapas mentales, notas estructuradas, guías de estudio e incluso podcasts generados con un solo clic. Esta funcionalidad resulta especialmente valiosa para proyectos que requieren trazabilidad, precisión contextual y foco temático, ya que evita recurrir a información externa o no verificada.

En múltiples entornos NotebookLM se consolida como un recurso estratégico para sistematizar, comprender y explorar grandes volúmenes de texto, con un enfoque centrado en la documentación propia del investigador o del equipo de trabajo.

b. Agente ChatGPT

El 17 de julio de 2025, OpenAI anunció el lanzamiento de una nueva herramienta denominada "Agente ChatGPT". Según la escueta información contenida en su tarjeta de presentación oficial (ChatGPT Agent System Card²⁵), este nuevo desarrollo se presentaría como una evolución significativa dentro del paradigma de los modelos agénticos de IA, al combinar, por primera vez, las capacidades de investigación profunda (Deep Research) y de ejecución instrumental (Operator) en una única arquitectura operativa.

Esta convergencia permite alcanzar un grado de capacidad agéntica considerablemente superior al de cualquier otro modelo de lenguaje mencionado hasta el momento en el presente playbook.

25 <https://openai.com/index/chatgpt-agent-system-card/>

Entre las funcionalidades que integran el “Agente ChatGPT” se encuentran:

Deep Research

Capacidad para llevar a cabo investigaciones complejas, multietapas, y generar reportes de alta calidad a partir del análisis cruzado de múltiples fuentes.

Operator

Ejecución de tareas mediante un navegador visual remoto, que le permitiría interactuar con sitios web como lo haría un usuario humano. En su versión individual este le solicitaba al usuario su validación en casi todos los puntos de una tarea que se le encomendara y, como consecuencia de ello, se generaba una intervención constante del usuario en su ejecución. Ahora, en su integración a la cadena de razonamiento del agente, solamente consulta y solicita la intervención del usuario en puntos clave, mejorando significativamente la experiencia de este último en el uso de la herramienta.

Terminal restringido

entorno controlado para ejecutar código, realizar análisis de datos y generar planillas o presentaciones, con acceso de red limitado.

Conectores externos

acceso a fuentes de datos y servicios como Google Drive mediante integraciones conocidas como “connectors”, ampliando su alcance operativo.

Esta herramienta se perfila como un “agente dentro del agente de IA”. Un ejecutor contextualizado para tareas complejas, con capacidad de acción, “razonamiento”, análisis y síntesis para completar la tarea que se le indique.

En virtud del riesgo potencial asociado a este nuevo nivel de autonomía, OpenAI anticipa un despliegue acompañado de un conjunto robusto de salvaguardas. Estas medidas no solo buscan reducir errores del agente, sino también prevenir abusos, filtración de datos y usos maliciosos. Entre ellas, se destacan:

Desactivación de la memoria persistente

a fin de reducir la posibilidad de que información sensible almacenada de manera contextual sea exfiltrada mediante ataques indirectos, como inyecciones encadenadas. El agente se lanzaría sin acceso a memoria persistente. Esto implicaría que el agente no podría recordar interacciones pasadas entre sesiones. No obstante, aclaran que se reservan la posibilidad de revisar esta opción en un futuro.

Restricciones en el uso del terminal

el acceso al terminal estaría limitado a operaciones seguras, como por ejemplo solicitando al servidor una información o recurso concreto, como es el caso de las peticiones GET para descarga de imágenes o datasets públicos verificados (por ejemplo, fuentes gubernamentales). Esta limitación buscaría impedir que el agente realice acciones no autorizadas en red (como POST, PUT o ejecución remota), minimizando así el riesgo de que ejecute comandos con consecuencias externas imprevistas. En este supuesto también aclaran que se reservan la posibilidad de revisar esta opción en un futuro.

Monitoreo y filtrado automático de prompt injections

se implementaría un sistema automatizado de filtros y monitores diseñados para detectar instrucciones maliciosas ocultas en páginas web u otros contenidos a los que el agente accede. Estas inyecciones podrían intentar subvertir la intención original del usuario o provocar comportamientos no deseados, como el robo de datos o la ejecución inadvertida de acciones. Según el documento, este tipo de riesgo podría ser mayor que en versiones anteriores debido al acceso simultáneo del agente a múltiples herramientas, motivo por el cual se desarrolló una arquitectura de defensa multicapa y entrenamiento específico para fortalecer su resistencia ante estas amenazas.

Modo de observación (watch mode)

este mecanismo, heredado del modelo Operator, consiste en que cuando el agente ChatGPT utilice la herramienta de navegador visual en un contexto sensible (por ejemplo, al iniciar sesión en un correo electrónico o una cuenta bancaria), se activaría automáticamente el “modo observación” con el objetivo de que el usuario controle lo que el agente ejecuta. Ante cualquier signo de inactividad del usuario o abandono de la sesión, el sistema detiene su ejecución inmediatamente.

Sistema de solicitud deconfirmaciones explícitas

antes de ejecutar acciones de alto impacto, como enviar correos electrónicos, realizar pagos o modificar permisos sobre archivos, el agente le solicitará al usuario una confirmación explícita.

No obstante todos los beneficios que promete esta innovadora herramienta, las pruebas realizadas por los usuarios, tras la habilitación experimental del agente por parte de OpenAI, permiten identificar algunas limitaciones relevantes en su funcionamiento actual.

OpenAI ha dispuesto el acceso, en fase inicial, al agente para los usuarios con suscripciones Pro con un cupo de 371 prompts y para usuarios Plus con 40 prompts disponibles. Esta lógica de uso limitado por cantidad de prompts introduce una restricción operativa importante. Aunque el usuario puede detener el flujo del agente en cualquier momento como hacer preguntas en medio del proceso o corregir y guiar al agente si este se encuentra cometiendo errores en la ejecución de la tarea, cada intervención del usuario consume uno de esos créditos.

En lugar de computarse por tarea cumplida o conversación finalizada, el sistema descuenta un crédito por cada interacción. Esto encarece significativamente el uso activo del agente, especialmente en flujos complejos que requieren supervisión, redireccionamiento o correcciones humanas.

En lo que respecta a sus limitaciones técnicas, si bien el agente puede ejecutar ciertas tareas con eficacia, su desempeño es inconsistente cuando se le encomiendan determinadas acciones como, por ejemplo, el completar hojas de cálculo complejas en Excel o generar presentaciones PowerPoint con imágenes incorporadas. En estos casos, los resultados obtenidos son frecuentemente pobres o incompletos.

Además, el agente enfrenta obstáculos en la navegación web cuando se encuentra con sistemas de validación diseñados para distinguir humanos de bots, como los CAPTCHA. Dado que el entorno web no fue concebido para la interacción autónoma de agentes, estas barreras interrumpen su flujo de trabajo. En tales situaciones, el sistema solicita asistencia del usuario humano para completar la validación, tras lo cual puede retomar la tarea. Esto pone en evidencia que aún no es posible delegar por completo ciertos tipos de procesos sin intervención humana. Ahora bien, esta evolución del “agente que maneja tú computadora” cristaliza una capacidad agéntica más robusta en relación a otras plataformas y agentes de IA.

c. Test agéntico. Escalas y niveles de agentes de IA

I. Introducción

Diseñamos un test específico orientado a identificar el grado de agencia que presentan las plataformas y soluciones de IA. En particular, evaluamos el modelo Notebook LM y el “Modo Agente ChatGPT” recientemente lanzado. Recordemos que la noción de “agencia” no se limita a la capacidad de respuesta frente a estímulos, sino que abarca un conjunto de competencias más complejas, como la planificación, la toma de decisiones autónoma, el manejo de memoria contextual o de largo plazo, la integración multimodal y el acceso a herramientas externas para ejecutar acciones.

Para medir esas capacidades, en el estado actual del arte, consideramos un sistema de puntuación basado en una escala del 1 al 5, en la que cada valor refleja un nivel específico sobre los principales componentes dentro del paradigma de la IA agéntica.

Los criterios que sustentan cada puntaje son:

Nivel 1	Modo asistencia	Sin funcionalidad, puramente reactivo, sin agencia ni contexto
Nivel 2	Presencia incipiente	Funcionalidad limitada, dependiente de intervención humana alta
Nivel 3	Desempeño parcial en entornos controlados	Funcional en contextos específicos, pero requiere supervisión frecuente
Nivel 4	Despliegue sólido con autonomía funcional	Autónomo en la funcionalidad, estable en flujos de trabajo, pero sin coordinación sistémica completa
Nivel 5	Integración plena y articulada	Totalmente integrado, dinámico, proactivo y adaptativo en una arquitectura agéntica coherente

El test se aplica a través de una matriz que contempla las dimensiones que hemos trabajado este playbook N°1:

- a.** El nivel de evolución de los agentes de IA impulsados por modelos de lenguaje
- b.** La capacidad de memoria persistente (de sesión).
- c.** La capacidad de memoria de largo plazo.
- d.** El grado de acceso a herramientas.
- e.** Las capacidades multimodales.
- f.** Mecanismos de planificación y razonamiento.
- g.** El grado de madurez agéntico en general.

Este test no pretende constituir una clasificación/puntuación definitiva, sobre todo de cara a un cuadro tecnológico en permanente y veloz evolución, sino que busca ofrecer un marco que permita entender el estado de evolución de los sistemas de IA en relación con el paradigma agéntico, permitiendo visualizar con mayor precisión sus alcances, limitaciones y potencialidades.

II. Aplicamos el test agéntico UBA IALAB a la plataforma NotebookLM y al “Agente ChatGPT”

Veamos ahora cómo se puede aplicar en la plataforma NotebookLM y en el “Agente de ChatGPT”:

Componente evaluado	Nivel de evolución de los agentes de IA (modelos de lenguaje)
Calificación: 4	NotebookLM Presenta autonomía para tareas como generar resúmenes, FAQs, timelines o informes a partir de fuentes. Puede activar “modo agente” implícito en features como “deep research” dentro de fuentes (ej. analizar PDFs multimodales y generar informes sin prompts paso a paso). No es solo reactivo; entiende intenciones y ejecuta secuencias (ej. de un prompt como “crea una guía de estudio sobre estas notas”, planifica y genera estructura).
Calificación: 5	Agente ChatGPT Combina capacidades de investigación profunda, ejecución de tareas a través de un navegador visual remoto y herramientas de terminal con acceso limitado a redes. Actúa como un sistema unificado que entiende intenciones, planifica y ejecuta secuencias complejas (ej. generación de informes, interacción con aplicaciones externas como Google Drive).

Componente evaluado**Capacidad de memoria persistente
(de sesión)****Calificación: 4****NotebookLM**

Mantiene contexto dentro de un Notebook/sesión, recordando fuentes y ajustes en interacciones continuas para coherencia.

Calificación: 5**Agente ChatGPT**

Mantiene contexto continuo dentro de sesiones, recordando prompts, datos externos y herramientas utilizadas, asegurando coherencia en interacciones multturno. Su diseño soporta conversaciones complejas y mantiene consistencia en flujos de trabajo prolongados.

Componente evaluado**Capacidad de memoria de largo plazo****Calificación: 3****NotebookLM**

Recuerda historial por notebook/usuario (preferencias, notebooks previos), pero no tan transversal como en agentes full (ej. no anticipa basándose en sesiones ajenas sin integración explícita). NotebookLM Plus mejora esto con más notebooks persistentes.

Calificación: 4**Agente ChatGPT**

ChatGPT Agent recuerda historial de usuario y preferencias a través de sesiones, integrando datos de interacciones previas y herramientas externas (ej. Google Drive). Aunque no anticipa basándose en sesiones no relacionadas sin integración explícita, su capacidad para mantener contexto a largo plazo es más robusta que NotebookLM, especialmente en entornos dinámicos.

Componente evaluado	Grado de acceso a herramientas
----------------------------	---------------------------------------

Calificación: 3

NotebookLM

Procesa archivos subidos (texto, imágenes, audio, video, PDFs multimodales), y navega en internet en búsqueda de fuentes a solicitud del usuario.

Calificación: 5

Agente ChatGPT

Acceso a herramientas externas como navegadores visuales, terminal con red restringida y conectores a aplicaciones (ej. Google Drive). Ejecuta código, realiza análisis de datos y genera contenido (ej. diapositivas, hojas de cálculo).

Componente evaluado	Capacidades multimodales
----------------------------	---------------------------------

Calificación: 5

NotebookLM

Nativamente multimodal ya que entiende/procesa/genera texto, imágenes, audio (Audio Overviews), video (Video Overviews en 2025) y tablas/gráficos. Ofrece “percepción holística” similar a la humana.

Calificación: 5

Agente ChatGPT

Nativamente multimodal: Procesa texto, imágenes, audio y datos estructurados (ej. tablas). Soporta entrada de imágenes y entornos visuales (navegador visual).

Componente evaluado	Mecanismos de planificación y razonamiento
----------------------------	---

Calificación: 3

NotebookLM

Usa Chain-of-Thought implícito para descomponer objetivos (ej. de fuentes crudas a mind maps o guías estructuradas). Autónomo en ejecución por etapas, con corrección (ej. iterar basándose en feedback humano).

Calificación: 5

Agente ChatGPT

Emplea razonamiento avanzado (Chain-of-Thought explícito e implícito) para descomponer tareas complejas (ej. resolver CTF challenges, generar informes multi-paso). Autónomo en planificación y ejecución, con capacidad para iterar basándose en feedback y confirmaciones humanas.

Componente evaluado

Grado de madurez agéntica en general

Calificación: 3

NotebookLM

Presenta un nivel intermedio-alto de madurez agéntica. Logra ejecutar tareas complejas de forma autónoma. Su razonamiento es coherente, multimodal, y está estructurado en etapas, manteniendo contexto dentro de una sesión e integrando elementos como fuentes, preferencias y estructuras de salida (ej. resúmenes, mapas conceptuales, guías). Si bien su memoria de largo plazo es limitada y la integración de herramientas externas es parcial, demuestra capacidades agénticas relevantes, especialmente en planificación, ejecución segmentada y percepción multimodal.

Calificación: 5

Agente ChatGPT

Presenta un grado de madurez agéntica elevado. Combina razonamiento complejo (con Chain-of-Thought implícito y explícito), planificación multietapa, y acceso a herramientas externas como navegador visual, terminal con red restringida, y conectores a apps como Google Drive. Su memoria persistente y de largo plazo permite sostener flujos de trabajo prolongados y mantener coherencia entre interacciones, incluso cuando involucran múltiples herramientas. Su capacidad para actuar de forma autónoma, ejecutar tareas con mínima intervención humana, iterar según feedback, y operar en entornos multimodales, lo posiciona como un agente altamente desarrollado.

III. Calificación general

NotebookLM: una plataforma parcialmente agéntica

Con una calificación promedio de 3.5 resalta por su capacidad para crear resúmenes, guías o mapas mentales a partir de documentos, imágenes o audios que le des.

Usa un modelo avanzado (Gemini) que le permite planificar y generar contenido útil sin que tengas que guiarlo paso a paso. Además, entiende texto, imágenes y hasta audio o video, lo que lo hace muy versátil.

Aunque su capacidad para conservar información a largo plazo es limitada y la conexión con herramientas externas todavía es parcial, NotebookLM pone en evidencia habilidades agénticas significativas, especialmente en lo que respecta a su facultad para planificar, ejecutar tareas de forma segmentada y procesar distintos tipos de contenido de manera integrada.

Agente ChatGPT: el paradigma agéntico en acción

Debemos aclarar que la evaluación de este agente es preliminar, ya que todavía está en una fase experimental. Presenta un grado de autonomía robusto, con una calificación promedio de 4.5 en base a las escalas y a las posibilidades de autonomía en fases de pruebas. Es proactivo, capaz de planificar tareas complicadas como resolver problemas técnicos o analizar datos y mantener el hilo de conversaciones largas. Aunque a veces necesita que el usuario confirme ciertas acciones sensibles como enviar un correo o hacer una compra, su capacidad para trabajar en entornos reales y dinámicos lo hace más autónomo y versátil que NotebookLM.

Un agente de IA es un especialista en tocar de manera autónoma un instrumento, mientras que la IA agéntica es la orquesta en su conjunto.

06 Limitaciones de la IA agéntica en las organizaciones: autonomía restringida en flujos complejos

Primero: En el ámbito de las organizaciones humanas, la autonomía decisional nunca puede ser ilimitada. Es limitada o condicionada por defecto, en función de múltiples variables técnicas, económicas, políticas, culturales, normativas o éticas.

La IA agéntica tiene que tener el mismo punto de partida: no se puede concebir una autonomía ilimitada por diseño. La autonomía condicionada o limitada es la clave para desarrollar ecosistemas agénticos que puedan adaptarse a cualquier tipo de organización.

Segundo: Los límites a las cualidades agénticas de los sistemas de IA, vienen dados por un conjunto de factores que responden a diferentes variables y escenarios. Si se trata de organizaciones públicas o privadas, el tipo de tareas involucradas, el grado de intervención humana requerido en función de los objetivos y estrategias empresariales, el branding, las normas legales involucradas, el nivel de adaptación cultural, entre muchos otros.

Tercero: A Julio de 2025, la IA agéntica más avanzada como el "Agente ChatGPT", presenta limitaciones sustanciales a la hora de ejecutar flujos de trabajo complejos de manera completamente autónoma. Si bien muestran un desempeño notable en tareas específicas, su capacidad para coordinar procesos de múltiples etapas sin intervención humana o sin sistemas de orquestación sigue siendo acotada.

Por ejemplo, ChatGPT puede analizar documentos, responder preguntas basadas en imágenes o transcribir audio, pero no puede por sí mismo planificar, ejecutar y monitorear de principio a fin un proyecto complejo como una auditoría o la elaboración de un contrato legal sin depender de herramientas adicionales ni de validación humana. Del mismo modo, Grok puede razonar y asistir en tareas analíticas, pero no está diseñado para gestionar procesos interdependientes que requieren iteraciones normativas, acuerdos entre múltiples partes y validaciones reglamentarias.

En general, su autonomía se limita a tareas aisladas dentro de un flujo mayor. Para alcanzar coherencia operativa, trazabilidad y toma de decisiones fundamentada en procesos complejos, sigue siendo indispensable la intervención humana o la integración con arquitecturas multiagente dotadas de orquestadores específicos.

En síntesis: la IA agéntica presenta tres limitaciones clave en relación a las capacidades para ejecutar tareas de manera autónoma en el plano de las organizaciones y los procesos concretos.

A

Falta de orquestación autónoma

Los agentes de IA no cuentan con mecanismos internos para coordinar de manera secuencial o paralela tareas interdependientes propias de un flujo de trabajo. Por ejemplo, en una auditoría financiera, pueden analizar datos contables, pero no gestionar autónomamente la recolección de documentos, la validación normativa cruzada y la generación de informes finales sin asistencia externa.

B

Dependencia de herramientas externas

Aunque pueden integrarse con APIs, aplicaciones y otros recursos, requieren intervenciones humanas para configurar, controlar y validar dichas integraciones. Un agente puede redactar un correo, pero no coordinar su envío, seguimiento y gestión de respuestas de forma autónoma dentro de un flujo sostenido.

C

Escalabilidad en flujos multiagente

Muchos procesos complejos requieren la colaboración de agentes especializados, por ejemplo, para extracción de datos, redacción, verificación normativa. La coordinación entre ellos sigue requiriendo orquestadores externos o intervención humana, lo que restringe su capacidad de operación autónoma a gran escala.

En pocas palabras...

La IA agéntica marca un salto cualitativo en autonomía, razonamiento y capacidad instrumental, pero aún opera bajo una autonomía restringida. Si bien los agentes de IA actuales, como el Agente ChatGPT, pueden ejecutar tareas complejas y multietapas, siguen dependiendo de validaciones humanas, herramientas externas y sistemas de orquestación para sostener flujos de trabajo completos. Su desempeño es destacable en tareas puntuales, pero fragmentado en procesos largos o interdependientes. Por eso, la verdadera transformación no está en reemplazar al humano, sino en rediseñar su rol como coordinador, curador y garante del ecosistema agéntico.

07 Conclusiones

Los avances recientes han transformado de manera sustancial la naturaleza de las plataformas de IA generativa más utilizadas como ChatGPT, Grok, Gemini y NotebookLM, posicionándose como núcleos tecnológicos crecientemente autónomos, adaptativos y multifuncionales. Lo que antes eran asistentes conversacionales centrados en respuestas textuales y reactivas, hoy se configuran como agentes de IA: entidades autónomas, adaptativas y proactivas, capaces de operar en flujos complejos, acceder a herramientas, planificar, ejecutar tareas y aprender de la interacción.

Este salto se materializa de forma clara con la aparición del Agente ChatGPT, que representa un punto de inflexión en términos de agencia, autonomía operativa y arquitectura integrada. A través de funcionalidades como Deep Research, navegación visual, terminal restringido y conectores externos, se proyecta como un “agente dentro del agente”, diseñado para operar con un nivel de independencia significativamente mayor al de generaciones anteriores. Sin embargo, es importante enfocarnos en las capacidades disponibles en el presente. A diferencia de los primeros asistentes conversacionales, centrados exclusivamente en el intercambio textual, los agentes de IA actuales incorporan capacidades avanzadas como razonamiento estructurado, memoria persistente y de largo plazo, acceso a herramientas externas, planificación por etapas, ejecución secuencial, percepción multimodal y personalización dinámica.

Estas capacidades permiten que los agentes de IA dejen de ser simples interfaces conversacionales para convertirse en plataformas de automatización inteligente, capaces de participar activamente en tareas de mayor extensión, variabilidad y complejidad. En este nuevo paradigma, no solo se incrementa la eficiencia operativa, sino que también se redefine el rol del usuario humano, que pasa de ejecutar tareas manualmente a convertirse en editor, auditor y supervisor de flujos agénticos.

En este contexto, la irrupción de la IA agéntica marca un cambio de época. No se trata simplemente de un nuevo modelo, sino de un paradigma en el que la autonomía y la proactividad se presentan por capas o grados. A diferencia del asistente conversacional, el agente puede tomar decisiones, organizar tareas y actuar orientado a metas, con una lógica de delegación similar a la de un representante digital. Esto habilita la construcción de ecosistemas en los que distintos agentes operan de manera orquestada para resolver objetivos complejos.

Este modelo agéntico es el que da origen a las organizaciones centauro, estructuras organizacionales diseñadas sobre la base de una colaboración simbiótica entre humanos y agentes de IA. En ellas, los agentes se integran como digital workers capaces de aprender, razonar y decidir junto a las personas. Su diseño se apoya en principios como trazabilidad, validación humana, visualización de procesos y generación de resultados explicables. No se trata solo de automatizar, sino de reconfigurar cómo pensamos y producimos dentro de entornos institucionales.

A pesar de los avances, la autonomía sigue siendo restringida. La ejecución de flujos complejos aún requiere intervención humana, validaciones, herramientas externas y sistemas de orquestación. Por eso, el foco es transformar el rol de las personas en el marco de estas organizaciones: de productor de tareas, a director y editor de agentes en un ecosistema cada vez más agéntico. En este escenario, el test agéntico desarrollado por UBA IALAB resulta clave para dimensionar con mayor precisión el estado evolutivo de cada agente, permitiendo distinguir niveles de madurez y orientar su implementación según las necesidades organizacionales.

La IA agéntica no es un destino final, sino el inicio de un nuevo tipo de convivencia tecnológica: una simbiosis organizacional que redefine la frontera entre lo humano y lo digital



IALAB

