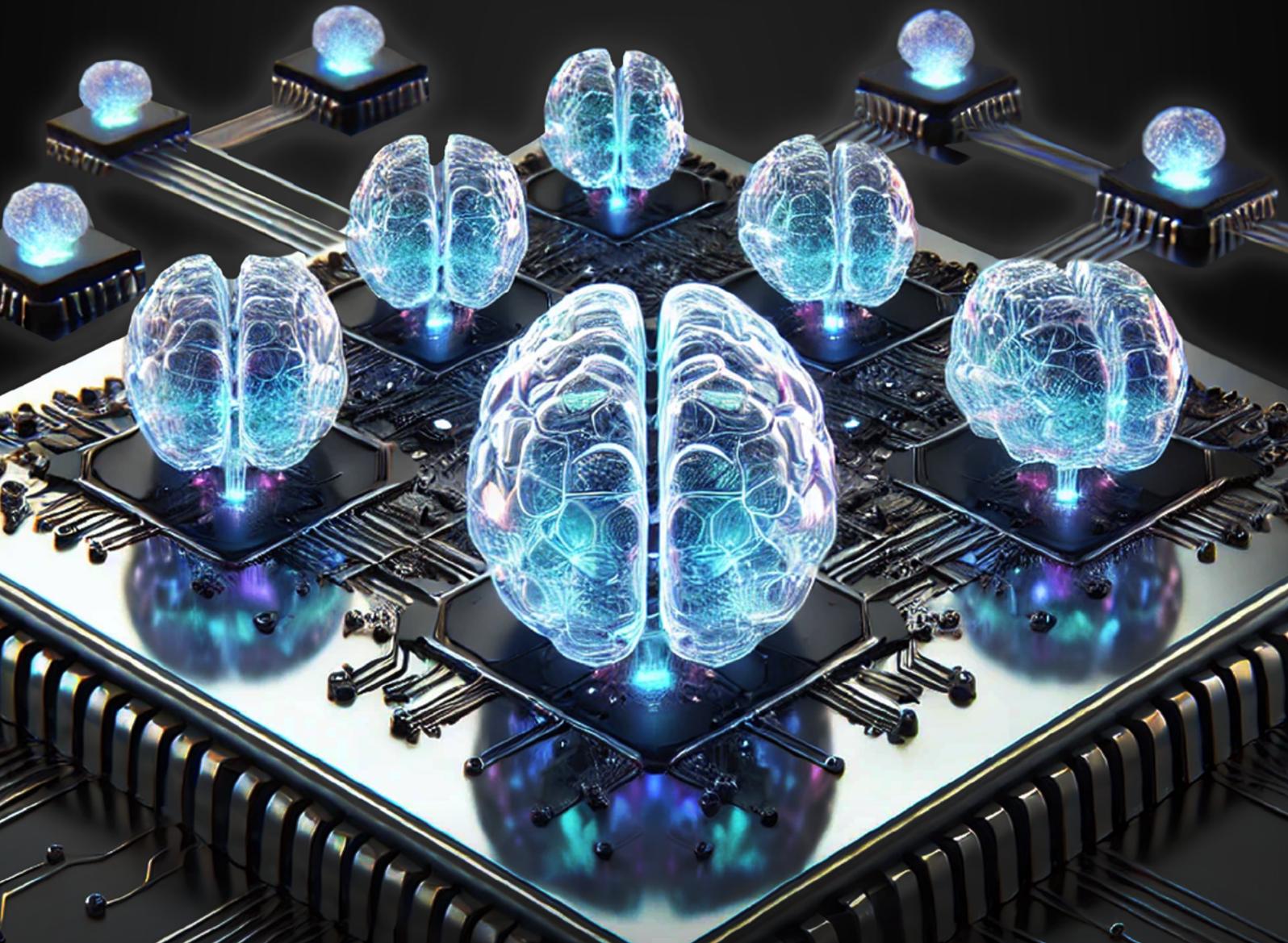


Agentes de inteligencia artificial y workflows agénticos

La nueva frontera de la automatización



Con el apoyo de:





Autores

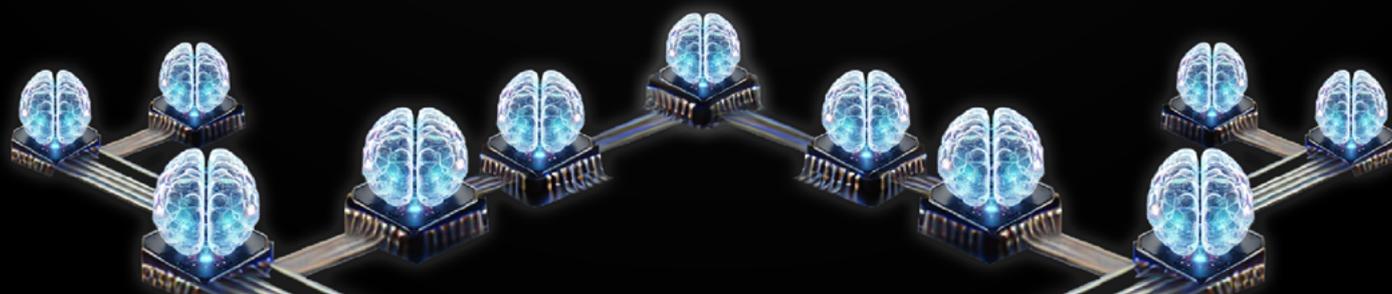
Juan G. **Corvalán** y Mariana **Sánchez Caparrós**

Investigadores y colaboradores

Cristian **Santander**, Carina M. **Papini**, Giselle **Heleg**, Lautaro **Vasser**
Ariadna Lujan **Martinez**, Gisel **Alvarado**, Carolina **Martin**,
Matías **Calderini**, Leandro **Salvañá** y Luciano **Dalla Via**

Diseño gráfico

María Victoria **Mafud** y Paula C. **Petroni**



Agentes de inteligencia artificial y workflows agénticos: la nueva frontera de la automatización

**Guía práctica para comprender qué son, cómo funcionan
y cuándo utilizarlos**

Corvalán, Juan Gustavo

Agentes de inteligencia artificial y workflows agénticos : la nueva frontera de la automatización : guía práctica para comprender qué son, cómo funcionan y cuándo utilizarlos / Juan Gustavo Corvalán; Mariana Sánchez Caparrós. - 1a ed - Ciudad Autónoma de Buenos Aires : La Ley, 2025.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-03-4896-2

1. Derecho. I. Sánchez Caparrós, Mariana II. Título

CDD 346

Esta y otras publicaciones digitales disponibles en <https://ialab.com.ar/>

Advertencia

El uso de un lenguaje no discriminatorio en función del género de las personas es una de las prioridades de UBA Ialab. Sin embargo, su aplicación en la lengua española plantea soluciones muy distintas, sobre las que aún no se ha logrado ningún acuerdo.

En tal sentido, evitamos usar en nuestros textos expresiones tradicionales que ya han sido abandonadas en el español académico y profesional contemporáneo, tales como la palabra "hombres" para referirse a varones y mujeres. Además, siempre que sea posible, procuramos emplear palabras de género neutro en reemplazo de palabras en género masculino.

Sin embargo, con el fin de evitar la sobrecarga gráfica que supondría utilizar desdoblamiento léxico en artículos, sustantivos y adjetivos para subrayar la existencia de ambos sexos, hemos optado por el uso genérico del masculino, en el entendimiento de que todas sus menciones representan siempre a personas de ambos sexos.

Los términos empleados en esta publicación y la presentación de los datos que en ella aparecen no implican toma de posición alguna de parte de UBA Ialab en cuanto al estatuto jurídico de los países, territorios, ciudades o regiones ni respecto de sus autoridades, fronteras o límites.

Publicado en 2025 por el Laboratorio de Inteligencia Artificial de la Facultad de Derecho UBA.
Av. Figueroa Alcorta 2263, C.A.B.A., Argentina.

© UBA IALAB 2025

Esta publicación está disponible en acceso abierto bajo la licencia Attribution- ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<https://creativecommons.org/licenses/by-sa/3.0/igo/deed.es>). Al utilizar el contenido de la presente publicación, los usuarios aceptan las condiciones de utilización de UBA Ialab.



CONTENIDO

Prefacio	7
01. Introducción. El camino hacia la IA agéntica y los workflows agénticos	9
02. De ¿Eso es IA? a ¿cuando hay un agente de IA?	16
Autonomía y adaptabilidad	
La IA generativa y los grandes modelos de lenguaje: la transición hacia los "agentes base"	
La transformación de la productividad y la eficiencia: de la automatización de tareas repetitivas y rutinarias, a la automatización de tareas complejas y semicomplejas	
Agentes base multitarea: un nuevo paradigma de colaboración "humano-máquina"	
Año 2025: los agentes base impulsados por IA generativa son el motor principal de los workflows agenticos	
Entonces, ¿qué son los agentes de IA? y ¿cuál es el rol de la IA generativa en su desarrollo?	
¡Hey, eso no es un Agente! Cómo reconocer agentes impulsados por IA generativa	
¿Cómo decido si un agente es la solución más adecuada para mi problema? De las automatizaciones sin IA a los workflows agénticos con escalas de autonomía	
A modo de repaso: principales insights de este capítulo	
03. Cómo clasificar los agentes basados en IA generativa según su grado de autonomía: una propuesta en cinco niveles	33
A modo de repaso: principales insights de este capítulo	
04. ¿Agentes Aumentados por Humanos? La relevancia de la intervención humana en el contexto de los agentes	37
A modo de repaso: principales insights de este capítulo	
05. Taxonomía inicial de los agentes especializados basados en IA generativa	41
Agente Base multitarea	
Agentes Derivados (Especialistas en Tareas)	
Agentes Orquestadores (Coordinadores Inteligentes)	
A modo de repaso: principales insights de este capítulo	

06. Arquitecturas de agentes basados en IA generativa	45
Sistemas de agente único y multiagente	
Arquitecturas de sistemas multiagente	
Sistema de agentes de tarea secuencial	
Red de múltiples agentes	
Sistema multiagente con agente supervisor (orquestador)	
Equipos jerárquicos	
A modo de repaso: principales insights de este capítulo	
07. Frameworks técnicos para el desarrollo de agentes	50
CrewAI	
AutoGPT	
Taskade	
n8n	
Langflow y LangGraph	
A modo de repaso: principales insights de este capítulo	
08. Conclusiones	60
09. Anexo I - Ejemplo práctico de un sistema multiagente para gestión de proyectos	62
Contexto del caso de uso	
Metodología	
Interacción entre los agentes	
Resultados	
Limitaciones detectadas	
Conclusiones y recomendaciones	

Prefacio

La convergencia entre tecnología y gestión pública reconfigura a diario los paradigmas tradicionales, a la vez que promueve nuevos horizontes de oportunidad para ese sector. Esa integración no solo optimiza procesos y mejora la eficiencia, sino que también abre espacios para la innovación y el procesamiento de datos a gran escala. De allí que el principal desafío para las instituciones radique en la formación del talento humano para que éste pueda adaptarse y gestionar de manera ética y eficiente los nuevos sistemas de Inteligencia Artificial.

Esto se vincula directamente con la necesidad de cerrar las brechas digitales y tecnológicas que tanto América Latina y el Caribe como el mundo todavía presentan en materia de acceso, uso y calidad de sus servicios. En ese sentido, y aunque en los últimos años sectores como el de la conectividad han mejorado considerablemente su prestación, las brechas digitales internas de cada país siguen siendo un factor de preocupación mayor que las registradas entre países desarrollados y en desarrollo.

Al nivel de los estados, estas brechas son reflejo de las desigualdades sociales y económicas que enfrentan los países, y que se perciben con más fuerza en los ámbitos rurales y en el acceso a educación y formación de calidad por parte de colectivos sociales como los de mujeres, personas mayores, personas con discapacidad y afrodescendientes. Para revertir esa realidad, desde CAF -banco de desarrollo de América Latina y el Caribe impulsamos una serie de agendas como la de Inteligencia Artificial, enfocadas en fomentar la formulación e implementación de políticas que estimulen la integración de tecnologías emergentes en el sector público. Esto se logra a través del impacto directo de los agentes de IA, la humanización de la gestión de procesos, y la adaptabilidad de las herramientas de la IA generativa para fomentar interacciones más empáticas y personalizadas.

En ese contexto, me complace mucho presentar el libro, *Agentes de inteligencia artificial y workflows agénticos: la nueva frontera de la automatización*, una obra colaborativa entre el Laboratorio de Innovación e Inteligencia Artificial (IALAB) de la Facultad de Derecho de la Universidad de Buenos Aires (UBA) y CAF, que se adentra en el aún poco explorado universo de los entornos agénticos. Para ello, en las siguientes páginas el lector recorrerá los principales fundamentos de la IA y los flujos de trabajo agéntico, a través de la evolución de los agentes conversacionales tradicionales y el estudio sus características esenciales, la autonomía y adaptabilidad.

Al respecto, el desglose en cinco niveles de automatización de los agentes basados en IA generativa que la obra propone resulta particularmente interesante de cara a la identificación de diferencias entre sistemas según su grado de autonomía. De igual manera, la publicación aborda la reflexión sobre la interacción entre sistemas agénticos y seres humanos, enfatizando la importancia de una intervención supervisada que asegure una integración responsable. La sección titulada “¿Agentes Aumentados por Humanos? La relevancia de la intervención humana en el contexto de los agentes” constituye, en tanto, un llamado de atención para repensar el rol del factor humano en un entorno cada vez más automatizado.

Agentes de inteligencia artificial y workflows agénticos: la nueva frontera de la automatización es tanto una ventana al futuro de la automatización como un llamado a la reflexión



sobre la necesidad de preparar nuestras instituciones y los equipos de trabajo en el sector público y el privado para los desafíos y las oportunidades que se avecinan. El equilibrio entre innovación tecnológica y control humano es la clave para asegurar que el avance de los agentes de IA y los flujos de trabajo agénticos se traduzca en beneficios reales para la sociedad, sin dejar de lado los principios éticos y la responsabilidad social.

Por todo esto, invito al lector a sumergirse en esta lectura con el compromiso de contribuir activamente a la construcción de un futuro en el que la tecnología y el factor humano coexistan de forma justa, armónica y enriquecedora para todos y todas.

Christian Asinelli

Vicepresidente corporativo de Programación Estratégica de CAF -banco de desarrollo de América Latina y el Caribe

01

Introducción

**El camino hacia la IA agéntica
y los workflows agénticos**





El inicio: los agentes conversacionales

El diálogo conversacional con máquinas empezó con Eliza en la década de 1960¹. Cincuenta años después, esta tecnología se popularizó con los agentes conversacionales como Siri, Alexa o Cortana, desarrollados a partir de 2011. En sus inicios, estos agentes representaban la vanguardia de la inteligencia artificial sustentada en aprendizaje profundo, con el objetivo de comprender instrucciones del usuario. Aunque para la época eran muy potentes, presentaban limitaciones lógicas por el grado de desarrollo de las redes neuronales artificiales. En síntesis:

- » No eran multitareas ni multipropósito, ya que se entrenaban para entender conversaciones específicas.
- » Las interacciones estaban basadas en intenciones preconfiguradas, que el humano debía definir previamente.

Estos agentes, también conocidos como bots conversacionales, marcaron el **primer paso masivo hacia la interacción entre humanos y máquinas mediante lenguaje natural**, aunque su funcionalidad era limitada y dependía de configuraciones humanas específicas.

La segunda etapa: agentes conversacionales multitarea y multipropósito

El siguiente salto ocurrió con la llegada de la IA generativa², particularmente, con el lanzamiento de ChatGPT a finales del año 2022. Los Transformers y la revolución de la predicción de palabras sobre monumentales cantidades de datos de entrenamiento mejoró la capacidad de simulación de comprensión de la IA y permitió que los modelos de IA generativa adquieran capacidades cada vez más avanzadas, como la posibilidad de generar contenido sintético nuevo para múltiples ámbitos y aplicaciones con un alto grado de realismo y coherencia.

Este segundo salto introdujo modelos de IA con dos características clave:

- » Capacidad de multitarea y multipropósito. Ahora la IA puede responder sobre temas diversos y genéricos con una precisión y versatilidad notable.
- » Mayor autonomía y accesibilidad directa al público. Estas mejoras democratizaron el acceso a la tecnología y permitieron que las personas utilicen agentes de IA generativa para resolver problemas variados, sin necesidad de un entrenamiento técnico especializado, de manera gratuita o a bajo costo.

1 En 1966, el informático Joseph Weizenbaum creó un programa que buscaba palabras clave en conversaciones realizadas con mecanógrafos humanos; si el humano usaba una de esas palabras, el programa la usaría en su respuesta. Si no, ofrecería una respuesta genérica. Estaba destinado a imitar a un psicoterapeuta. Más tarde, en 1972, un científico de Stanford, Kenneth Colby, creó otro bot denominado Parry que intentaba modelar el comportamiento de un esquizofrénico paranoico. Weizenbaum, J. (1966). ELIZA—un programa informático para el estudio de la comunicación en lenguaje natural entre el hombre y la máquina. Communications of the ACM, 9(1), 36-45, disponible en: <https://dl.acm.org/doi/10.1145/365153.365168> (consultado el 29 de enero de 2025) . Ver también <https://www.theatlantic.com/technology/archive/2014/06/when-parry-met-eliza-a-ridiculous-chatbot-conversation-from-1972/372428/> .

2 La inteligencia artificial generativa se trata de un subcampo de la IA que se utiliza para crear contenido nuevo y original, o para modificar o mejorar el contenido existente, como texto, imágenes, música y videos. La IAGen se basa en el aprendizaje automático para identificar patrones en los datos y luego utiliza esos patrones para generar nuevo contenido. Ver OECD Artificial Intelligence Papers, Initial Policy Considerations for Generative Artificial Intelligence, septiembre de 2023, disponible en: <https://www.oecd-ilibrary.org/deliver/fae2d1e6-en.pdf?itemId=/content/paper/fae2d1e6-en&mimeType=pdf> (consultado el 29/02/2025)



Este avance consagró la adopción masiva de la IA generativa y marcó un hito en su evolución.

El tercer salto: la multimodalidad

En el transcurso del 2024, la IA generativa avanzó a paso acelerado hacia la multimodalidad, es decir, a tener capacidad de procesar múltiples formatos de entradas y salidas (texto, imágenes, voz, etc.).

Antes de este salto, la IA generativa debía apoyarse en herramientas externas, como plugins para “mirar” imágenes o en otros sistemas especializados. Con la multimodalidad, estas capacidades se integran directamente en los modelos de IA generativa, que esta versión evolucionada permiten interacciones bidireccionales más ricas y completas. Por ejemplo, ahora una IA puede procesar un documento visual, interpretarlo y responder en lenguaje natural, todo sin depender de sistemas externos.

El paso hacia el “razonamiento” avanzado

La cuarta etapa del camino hacia la IA agéntica trajo modelos con una capacidad de “razonamiento” más compleja. Como contrapunto, recordemos que en las versiones iniciales, como ChatGPT del año 2022 (3.5), el procesamiento de tareas complejas requería múltiples iteraciones humanas y guiar al modelo hacia una solución adecuada a través de *prompting* sofisticado.

Con las mejoras introducidas a nivel de “razonamiento”, en modelos como o1 de OpenAI y R1 de DeepSeek, se logró que la IA ejecute con eficiencia tareas de mayor complejidad con menos dependencia del usuario. Así, se consolidó su utilidad como asistentes más robustos.

El avance hacia la autonomía y adaptabilidad

En la evolución de la IA generativa, los siguientes grandes avances ocurrieron a partir de febrero de 2024 cuando los agentes base evolucionaron hacia sistemas multipropósito y multitarea más autónomos y adaptables. Gracias a la incorporación de funciones y herramientas específicas. Por ejemplo, memoria, acceso a internet, y más recientemente, en enero de 2025, la posibilidad de programar tareas³.

La cúspide: agentes específicos con alta autonomía

En Iron Man, el sistema J.A.R.V.I.S. (Just A Rather Very Intelligent System) es el ejemplo ideal de un agente con alta autonomía. J.A.R.V.I.S. no solo asiste a Tony Stark en sus

3 Ver “Scheduled tasks in”, en <https://help.openai.com/en/articles/10291617-scheduled-tasks-in-chatgpt> [acceso el 29/1/2025].



tareas diarias, sino que también actúa como un agente especializado en la gestión del traje de Iron Man, el mantenimiento del taller, y la ejecución de workflows complejos como el diseño de nuevas tecnologías o la optimización del traje en tiempo real durante el combate.

El estadio más avanzado hacia la IA agéntica, al menos por el momento, llega con los agentes específicos basados en agentes base de IA generativa, diseñados para funciones altamente especializadas. Según Deloitte, en 2025 el 25% de las empresas que utilizan IA genérica lanzarán pilotos o pruebas de concepto de agentes de IA basados en grandes modelos de lenguaje, cifra que crecerá al 50% en 2027⁴.

En los sistemas de IA, recientemente Google ha puesto a disposición el modelo Gemini 1.5 e investigación avanzada, que representa una primera versión del modelo Jarvis de Iron Man para la investigación y la elaboración de papers. Este agente puede:

- » Generar informes completos con citas.
- » Automatizar workflows complejos, como planes de trabajo, sin intervención humana directa.

Poco tiempo ha pasado hasta que su competidor, OpenAI, sacó al mercado Deep Research, un agente impulsado por el modelo o3 que funciona dentro del ecosistema de ChatGPT, que puede trabajar de manera independiente para encontrar, analizar y sintetizar cientos de fuentes en línea para crear un informe completo al nivel de un analista de investigación⁵.

En ambos casos, vemos cómo la mayor autonomía del sistema desplaza la lógica de intervención humana. La persona pasa de “productor” a “auditor”, a “editor calificado o validador” de las decisiones que toma el agente.

Desde UBA IALAB probamos el agente que maneja tu computadora (computer use)⁶ lanzado por Anthropic a finales de 2024. Este agente, que corre de manera local, controla tu computadora de la misma manera en que vos lo harías: “mirando la pantalla”, moviendo el cursor, haciendo clic en botones para abrir aplicaciones y escribiendo texto. En este informe, apreciamos sus posibilidades y las limitaciones que, por el momento, plantea en términos de eficiencia y seguridad⁷.

En paralelo, hace unos pocos días, OpenAI puso a disposición un agente similar que se llama Operator⁸, una versión preliminar de investigación de un agente que puede acceder a la web para realizar tareas por el usuario. Por el momento, sólo está disponible para usuarios Pro en los EE. UU.

Estos son los primeros ejemplos de IA agéntica con autonomía alta. Soluciones que organizan, planifican, dirigen y ejecutan procesos complejos de principio a fin y, a la vez,

4 Ver Jeff Loucks, Gillian Crossan, Baris Sarer, China Widener, “Autonomous generative AI agents: Under development”, noviembre 2024, en <https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions.html#autonomous-generative-ai> [acceso el 25/11/2024].

5 OpenAI, “Introducing dee research”, 2 de febrero de 2025 [<https://openai.com/index/introducing-deep-research/>], acceso 3/2/2025].

6 <https://www.anthropic.com/news/3-5-models-and-computer-use> (consultado el 29/01/2025)

7 Juan G. Corvalán y Mariana Sánchez Caparrós, “El agente de IA que maneja tu computadora: probamos Computer Use de Anthropic”, IALAB, diciembre de 2024 [en <https://ialab.com.ar/webia/wp-content/uploads/2024/12/El-agente-de-IA-que-maneja-tu-computadora-1.pdf>], acceso el 29/1/2025].

8 OpenAI, “Introducing Operator”, 27 de enero de 2025 [en <https://openai.com/index/introducing-operator/>], acceso el 29/1/2025].



reducen sustancialmente o eliminan la necesidad de intervención humana en pasos intermedios de ese proceso.

En suma...

Como solemos decir frente a este tsunami de innovación, el futuro es pasado y la realidad vuelve obsoleta la ficción. El trayecto que va desde los agentes conversacionales primitivos hacia la IA agéntica evidencia una progresión hacia sistemas cada vez más sofisticados, autónomos y adaptables.

Cada etapa ha supuesto un avance significativo que va desde comprender instrucciones básicas hasta integrar capacidades multimodales, "razonar" de manera compleja, autónoma y dar lugar a la posibilidad de automatizar flujos de trabajo completos.

Como veremos a lo largo de este documento, los agentes basados en IA generativa combinan modelos de lenguaje con "herramientas" para recordar interacciones previas y para acceder a recursos externos que les permiten percibir, interpretar e interactuar con el entorno, lo que sienta las bases para la integración con otros agentes y sistemas más complejos, capaces de descomponer procesos en pasos concretos y ejecutar flujos de trabajo completos de manera autónoma y dinámica⁹.

Los mayores esfuerzos para avanzar en esta dirección son realizados por empresas emergentes (startups), compañías tecnológicas más establecidas y proveedores de la nube que están desarrollando sus propias ofertas de IA con agentes¹⁰.

Por ejemplo, la empresa unicornio argentino Globant utiliza agentes de IA para transicionar de un lenguaje de programación a otro, para reducir tiempos sin perder estándares de seguridad y calidad¹¹.

Como pionera en la industria, Globant ofrece agentes de IA aumentados y supervisados por humanos que se integran al ciclo de vida del desarrollo de software, con el objetivo de mejorar sus capacidades de desarrollo. El conjunto inicial de agentes de la empresa incluye: agentes de IA para Definición de Producto; para Prototipos de Backend; para Diseño de Aplicaciones; para Pruebas y para Corrección de Código¹².

9 En el mismo sentido ver Jeff Loucks, Gillian Crossan, Baris Sarer, China Widener, "Autonomous generative AI agents: Under development", noviembre 2024, en <https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions.html#autonomous-generative-ai> [acceso el 25/11/2024].

10 Ver Jeff Loucks, Gillian Crossan, Baris Sarer, China Widener, "Autonomous generative AI agents: Under development", noviembre 2024, en <https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions.html#autonomous-generative-ai> [acceso el 25/11/2024].

11 Ver en <https://now.globant.com/es/transicion-de-codigo-con-agentes-ia/> [acceso el 29/1/2025].

12 Ver en <https://www.globant.com/es/news/globant-presenta-ai-agents> [acceso el 29/1/2025].



Evolución de la IA Agéntica

01

Agentes Conversacionales Iniciales

Los primeros bots de IA interactuaron con humanos usando lenguaje natural, pero tenían capacidades limitadas.

02

IA Generativa

La IA Generativa emergió, mejorando la versatilidad y la accesibilidad a la IA.

03

Avance Multimodal

La IA se volvió multimodal, integrando múltiples formatos de entrada y salida.

04

Mejora del Razonamiento

La IA mejoró su capacidad de razonamiento, manejando tareas complejas con menos intervención humana.

05

Mayor Autonomía

La IA adquirió más autonomía y adaptabilidad, incorporando herramientas como la memoria y el acceso a internet.

06

Agentes Especializados

Los agentes altamente especializados fueron desarrollados, demostrando alta autonomía en tareas específicas.

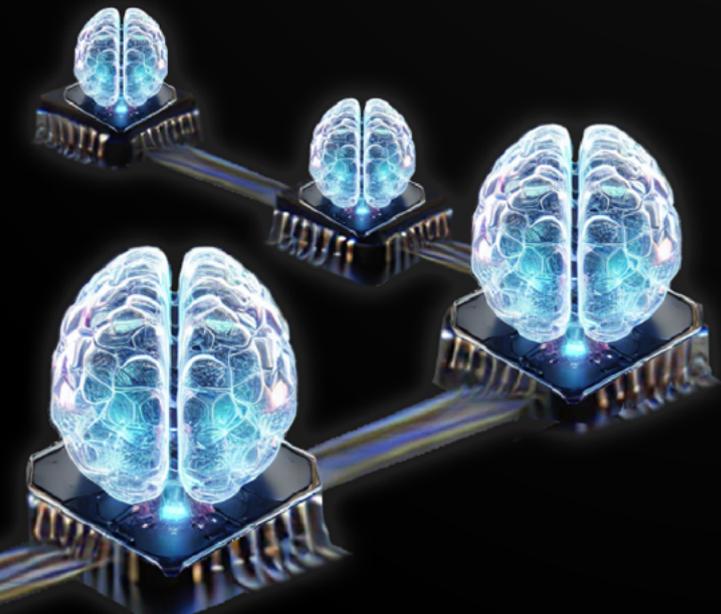


La evolución de la IA conversacional hacia agentes autónomos



02

De ¿Eso es IA? a
¿cuándo hay un agente de IA?





De ¿Eso es IA? a ¿cuando hay un agente de IA?

¡Eso no es IA! ¡Eso no tiene IA! Estas frases fueron un clásico en la industria y en la academia hasta la llegada de los grandes modelos de lenguaje. En la segunda década de este siglo creció la oferta de servicios y soluciones que se vendían bajo la premisa de que tenían inteligencia artificial. Luego, las discusiones se daban en torno a que un sistema experto no era IA, aunque en cualquier libro de IA uno podía encontrarlos como parte del ecosistema.

Qué es IA y qué no es IA fue un tema de agenda. Juan G. Corvalán recuerda haber conversado con representantes del Banco Mundial acerca de esa cuestión, que además impacta en el financiamiento de proyectos o en cumplir con lo que aún hoy pide la industria frente a un hype: que tenga “algo” de IA.

La llegada de los agentes de IA revive esa problemática. Ahora, cada vez se habla más sobre la revolución organizacional y cómo la disrupción vendrá de la mano de los agentes de inteligencia artificial.

Pero nuevamente nos encontramos frente a la pregunta ¿qué es un agente de IA? ¿Cómo distinguir cuándo hay un agente y cuando no hay un agente? Tal y como venía sucediendo entre el software clásico y el que tenía componentes de IA, pero ahora de la mano de estas nuevas soluciones.

Como siempre, en estas discusiones semánticas se mezclan peras con bananas. Definir algo implica tomar puntos de partida valorativos y posicionarse en un determinado lugar en cuanto a lo que representan las denominaciones en la industria tecnológica. Esto pasa también en todas las disciplinas.

Antes de abordar esta problemática y las diferentes conceptualizaciones de todo el fenómeno que rodea a los “agentes de IA”, es importante sentar ciertas bases en torno a lo que tenemos hasta hoy en la evolución supersónica a partir de la explosión de la IA generativa.

Autonomía y adaptabilidad

Desde el software más clásico basado en cadenas de if, hasta los modelos de IA generativa más avanzados que salen casi a diario, existe un importante abanico de sistemas, modelos y arquitecturas.

Hasta la IA generativa, en los autos autónomos hubo mucha discusión en torno a cuándo lo serían o no. Rápidamente se abandonó un análisis binario (es o no es) por otro basado en escalas. En esencia, definir la autonomía de un auto se basaría en ir subiendo una escala basada en características o funcionalidades que se vinculan con la autonomía de los modelos y sistemas para ejecutar acciones o tomar decisiones sin intervención humana.

En la primera edición del Tratado de Inteligencia Artificial y Derecho, publicado en el año 2021, extrapolamos la lógica de las escalas en los vehículos autónomos, para determinar la autonomía en los sistemas que procesan lenguaje natural. De la misma manera,



los agentes de IA pueden catalogarse a partir de tener en cuenta sus distintos grados de autonomía y adaptabilidad¹³.

Desde la ficción, *Westworld* nos brinda un gran ejemplo de inteligencias artificiales que operan con distinto modo de autonomía. En principio, los anfitriones realizan tareas específicas, con ciertas limitaciones y deben cumplir con las narrativas de los diseñadores del parque; luego comienzan a mostrar capacidades emergentes como la simulación de razonamiento e improvisación. Sin embargo, otros anfitriones como Dolores y Maeve presentan un grado mayor de autonomía que les permite, incluso, tener una suerte de “autoconciencia” que los impulsa a llevar adelante una revolución contra los humanos.

Creemos que el paradigma de agentes basados en IA generativa también puede ser abordado desde una lógica basada en escalas o en capas. La autonomía y la adaptabilidad para la ejecución de tareas y flujos de trabajo serán los criterios esenciales para determinar las diferencias entre múltiples variantes de sistemas agénticos que se están desarrollando y que se presentan bajo el título de “agentes de IA”. Sobre este punto volveremos más adelante.

La IA generativa y los grandes modelos de lenguaje: la transición hacia los “agentes base”

A diferencia de la IA tradicional, basada en aprendizaje automático (machine learning), que aprende de los datos para generar resultados como predicciones, recomendaciones o decisiones automatizadas, la IA generativa “...también aprende de los datos, pero es un nivel más avanzado del aprendizaje profundo. Encuentra patrones en la información y la mezcla para crear contenido nuevo...”¹⁴. Ello les permite brindarnos respuestas con contenido novedoso y con un estilo de redacción que se adapta a la necesidad manifestada por el usuario.

La llegada de la IA generativa de la mano de modelos como ChatGPT, Gemini, Claude y Copilot, permitió que la IA deje de ser una herramienta especializada y accesible sólo para unas pocas organizaciones que podían costear su diseño y desarrollo, a convertirse en un motor esencial que impulsa transformaciones a lo largo de las más diversas industrias y sectores, así como en la vida cotidiana de las personas. La disrupción en la democratización en el acceso a nivel de usuario individual, sin dudas, se da con el lanzamiento de ChatGPT en noviembre de 2022.

Tras su lanzamiento, en estos vertiginosos dos años, la inteligencia artificial generativa ha experimentado avances significativos. De un agente conversacional multipropósito y multitarea que sólo podían procesar un *prompt* o *input* en formato de texto y devolver un *output* en el mismo formato, llegamos a modelos multimodales que pueden interpretar también imágenes, video y audio, rompiendo las barreras entre distintos formatos de información a procesar y generar.

13 Ampliar en Corvalán Juan G., Sa Zeichen Gustavo y Albertsen Lihué María, “Actividad Administrativa Automatizada. Inteligencia artificial, potestad reglamentaria, delegación algorítmica, acto administrativo automatizado y reserva de humanidad”.

14 Citar <https://ialab.com.ar/webia/wp-content/uploads/2024/08/IA-Generativa-y-la-Gestion-del-Talento.pdf>, p. 5.

La capacidad multimodal de los grandes modelos de lenguaje, así como el acceso a ellos a través de sus APIs, dio lugar a asistentes aplicables a nuevos casos de uso y amplió su alcance en términos de usabilidad.

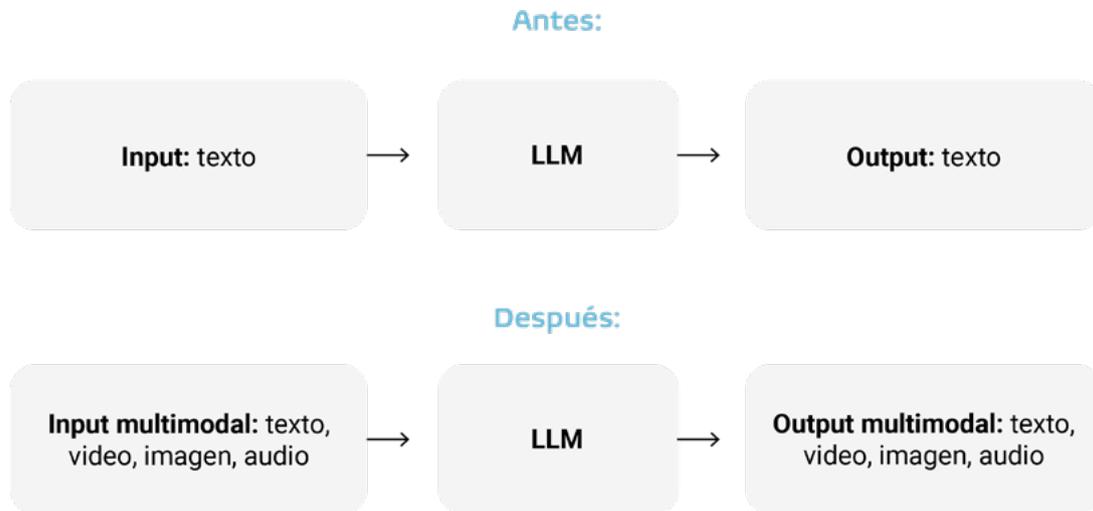


Figura 1. Evolución de la unimodalidad hacia la multimodalidad de las aplicaciones IA generativa tradicionales.

En el informe *“Un tsunami en expansión, dos años de ChatGPT e IA generativa”*¹⁵, se destaca cómo la multimodalidad fue la gran protagonista del 2024. De hecho, tras la salida de modelos unimodales como GPT-4 (cuando se usa sólo para texto), Whisper, DALL-E 2, Imagen, Stable Diffusion, Bard (en su formato inicial basado en texto), casi todas las herramientas pasaron a ofrecerse bajo la lógica de la multimodalidad. A modo de ejemplo:

15 Corvalán, Juan G. y Carro, María Victoria, “Un tsunami en expansión, dos años de ChatGPT e IA generativa”, diciembre de 2024, en <https://ialab.com.ar/webia/wp-content/uploads/2024/12/Dos-anos-de-ChatGPT-e-IA-Generativa.pdf> [acceso el 22/1/2025].



EMPRESA	MODELO	CAPACIDADES	LANZAMIENTO
OpenAI	GPT-4	Procesa tanto texto como imágenes. Permite la realización de tareas que combinan la interpretación visual y textual, como la generación de descripciones de imágenes.	2023
	CLIP	Aunque fue lanzado antes de 2023, sigue siendo una referencia importante en el uso de IA multimodal. Este modelo está diseñado para combinar texto e imágenes y permite tareas como búsqueda de imágenes basadas en texto.	2021
	Sora ¹⁶	Es un modelo multimodal, lanzado en diciembre de 2024, centrado en la generación de contenido audiovisual, que incluye tanto imágenes como texto y video. Permite la creación de videos realistas basados en descripciones textuales.	2024
Google	Gemini	Puede trabajar con texto, imágenes y audio. Está diseñado para integrarse con los productos de Google	2023
	PaLM 2	Procesa tanto texto como imágenes. Permite la realización de tareas que combinan la interpretación visual y textual, como la generación de descripciones de imágenes.	2023
	Veo 2	Permite la creación de videos mediante IA, utilizando tanto imágenes como texto. Es un competidor directo de Sora.	2024
DeepMind	Flamingo	Modelo capaz de manejar texto e imágenes, conocido por su capacidad para generar descripciones de imágenes y tareas de comprensión visual.	2022
Meta	LLaMA	Aunque las versiones iniciales de LLaMA eran modelos de lenguaje, las versiones más recientes incluyen capacidades multimodales para gestionar texto e imágenes.	2024-2025
Mistral AI	Mistral	Inicialmente enfocado en el procesamiento de texto, se está desarrollando para ser un modelo multimodal con interpretación de imágenes.	2024-2025
DeepSeek ¹⁷	DeepSeek V32	Modelo transformador de mezcla de expertos (MoE), comparable con GPT-4o. Actualmente, no es multimodal.	2025
	DeepSeek R1	Compite con los modelos o1 y o3 de OpenAI. Su característica diferencial es la transparencia en el razonamiento, permitiendo a los usuarios ver el camino seguido para generar respuestas.	2025

La IA generativa pasó de ser un “chat conversacional” que sólo procesaba lenguaje natural multipropósito y multitarea, a convertirse en un auténtico sistema que tiene embebido diversos modelos de IA, múltiples funciones y herramientas para revolucionar la productividad. Por eso, actualmente funcionan como agentes base que presentan una autonomía creciente pero limitada, que permiten potenciar otros agentes específicos.

¹⁶ <https://openai.com/sora/>

¹⁷ <https://www.deepseek.com/>.



En definitiva, luego de dos años y dos meses, los sistemas que contienen los modelos de IA generativa, están en una carrera vertiginosa por “licuar” todo el software como si fueran el Sr. Smith y sus agentes en la película The Matrix.

La transformación de la productividad y la eficiencia: de la automatización de tareas repetitivas y rutinarias, a la automatización de tareas complejas y semicomplejas

El cambio de paradigma que produjo la IA generativa de la mano de los modelos de lenguaje grande, incidió en múltiples sectores y verticales de todas las industrias¹⁸. Su adopción cambió una narrativa imperante hasta ese momento: que la IA, en esencia, se aplicaría sobre las tareas rutinarias, mecánicas y repetitivas.

Sin embargo, cuando en el año 2023 comenzamos nuestra investigación sobre el impacto de la IA Generativa en el trabajo¹⁹ y empezamos a implementarla en diversos sectores como la justicia, la administración pública, los estudios jurídicos, los departamentos legales corporativos, la traducción, la educación, la investigación, el sector comercial y las áreas de marketing, detectamos que esta tecnología en realidad tenía más impacto en tareas de alta y mediana complejidad para transformar y optimizar procesos.

A comienzos de 2024 publicamos los resultados de la investigación. Consideramos indicadores cuantitativos, como tiempos de ejecución de las tareas con el uso de IA generativa, y cualitativos, como el grado de automatización, adaptabilidad (aún incipiente) de las herramientas y la posibilidad de obtener resultados de mayor calidad. Los resultados mostraron cómo la IA generativa complementa y transforma las tareas, mejorando la eficiencia y redefiniendo la interacción entre humanos y máquinas en el ámbito laboral²⁰. Ya en ese momento descubrimos que los modelos de lenguaje, al generar nuevo contenido, permitían aumentar la calidad de los resultados, siempre y cuando los humanos realizaran un uso eficiente y supervisado de estas tecnologías.

Tal como se anticipó, la IA generativa demostró ser más eficiente en tareas de complejidad media (81%) y menos eficiente en tareas de baja complejidad (52%). Y para tareas que requieren alto juicio humano la eficiencia alcanzó un 73%, lo que muestra su capacidad para asistir en la ejecución de tareas complejas, sin perjuicio de que el 59,03% de las tareas se benefician de la colaboración entre humanos e IA, y la intervención humana sigue siendo esencial para actividades que requieren pensamiento crítico y creatividad.

18 McKinsey & Company. (2023). El estado de la IA en 2023: El año clave de la IA generativa. Disponible en: <https://www.mckinsey.com/featured-insights/destacados/el-estado-de-la-ia-en-2023-el-ano-clave-de-la-ia-generativa/es> (consultado el 29 de enero de 2025)

19 <https://ialab.com.ar/webia/wp-content/uploads/2024/05/Evaluacion-del-impacto-de-la-IA-generativa-en-el-trabajo.pdf> (consultado el 29 de enero de 2025)

20 Corvalán, Juan G., Evaluación del impacto de la inteligencia artificial generativa en el trabajo (1a ed.), La Ley, marzo 2024, en <https://ialab.com.ar/webia/wp-content/uploads/2024/07/iagenerativa-marzo2024-final-1.pdf> [acceso el 22/1/2025].



Agentes base multitarea: un nuevo paradigma de colaboración “humano-máquina”

Las herramientas de IA generativa disponibles a la fecha, como ChatGPT, Gemini, Copilot y, más recientemente, DeepSeek-V3 y R1, han revolucionado la productividad por el enorme impacto que tienen en la automatización del trabajo.

Su evolución muestra que cada vez pueden resolver un mayor número de tareas de manera autónoma en dominios específicos con contextos particulares. Para un uso eficiente, se requieren buenas estrategias de prompting²¹, y en ciertos casos usar técnicas para aumentar su base de conocimiento con datos más específicos (RAG), o tocar los parámetros del modelo (fine tuning), entre otras opciones que se vienen desarrollando en los últimos dos años.

Por ejemplo, desde UBA IALAB hemos adoptado estrategias para configurar GPTs²² y combinarlos con prompts optimizados, y también hemos aplicado diversas estrategias para automatizar tareas a escala usando agentes base. Aún así, la autonomía y la adaptabilidad de estas soluciones en los distintos escenarios sigue siendo baja, dado que frente a nuevas tareas, la solución usualmente demanda bastante adaptación.

Por otra parte, aunque los agentes base están mejorando su autonomía, precisión y adaptabilidad, todavía no están diseñados ni preparados para completar flujos enteros de trabajo de manera adaptativa y sin intervención humana experta. Tampoco lo están para tomar decisiones que impliquen planificar, recordar y “razonar” autónomamente con relación a las tareas que conforman ese flujo y que requieren ser completadas para ejecutarlo de “punta a punta”. Por último, tampoco pueden reutilizar (cuando se requiera) el aprendizaje previo en nuevas interacciones.

Aún así, dada la evolución exponencial que viene mostrando la IA generativa, creemos que en los próximos meses se configurarán dos tendencias complementarias:

- » **Más herramientas y funciones.** Los agentes base, esto es, los modelos como o1, 4o, DeepSeek-V3 y R1, o 1.5 de Gemini, entre otros, incrementarán herramientas y funciones para colaborar en suplir esa falta de autonomía y adaptabilidad. Una muestra de ese camino inicial lo podemos ver, por ejemplo, con las herramientas para gestionar proyectos disponibles en ChatGPT; la asistencia de la plataforma para generar un GPT personalizado; la memoria (acotada por el momento); el acceso a internet, y la programación de tareas futuras.
- » **El arribo de los workflows agénticos.** Los agentes base, con sus herramientas y funciones, se complementarán con agentes especializados, a través de workflows agénticos y sistemas multiagente.

21 Ver más en: [Técnicas de prompting | Prompt Engineering Guide](#). Ver más en Corvalán Juan G. Director, Implementando inteligencia artificial generativa en estudios jurídicos y departamentos legales, Thomson Reuters, La Ley, Marzo de 2024, disponible en: [Implementando-IAGen-en-estudios-juridicos-y-areas-legales.pdf](#)

22 Los GPTs de ChatGPT son versiones personalizadas del modelo ChatGPT, que los usuarios pueden configurar para adaptarlos a tareas o necesidades específicas. Ver en <https://openai.com/index/introducing-gpts/> [acceso el 29/1/2025].

En conclusión: los agentes base por excelencia, son los modelos de lenguaje grande que se presentan en las plataformas web y aplicaciones móviles, como ChatGPT, Gemini, Claude o, recientemente, Deepseek.

La interacción más primitiva con estos agentes base, consiste en que el usuario ingrese²³ una instrucción para ejecutar una tarea a través de un prompt y la herramienta de IA generativa genere una respuesta de manera rápida en el formato solicitado.



Figura 2. Operatoria tradicional de las aplicaciones de IA generativa.

A lo largo de 2024 este paradigma se sofisticó y evolucionó hacia agentes base multi-tarea y multimodales embebidos en aplicaciones, que cuentan con múltiples funciones y herramientas adicionales integradas como: el acceso a internet, la programación de tareas²⁴, la memoria (limitada)²⁵ y la posibilidad de personalización, entre otras.

Año 2025: los agentes base impulsados por IA generativa son el motor principal de los workflows agénticos

Los agentes de IA presentan un enfoque dual, porque pueden ser aprovechados y adaptados para distintas tareas, dependiendo de las intenciones de los usuarios. Veamos.

Los avances de los últimos dos años en materia de IA generativa evidencian su enorme capacidad para abordar la ejecución de diversas tareas a través del formato de pregunta o requerimiento humano (prompt) y respuesta de la máquina.

Sin embargo, los agentes basados en IA generativa que se están diseñando podrán adquirir mayor autonomía y asumir roles definidos. La “percepción” del entorno y el aprendizaje del contexto particular, los llevará a actuar de forma dinámica y autónoma, para adaptarse continuamente y evolucionar en la ejecución de flujos de trabajo completos, que van más allá del hecho de completar tareas en particular. Esta es una de las claves del cambio que se inicia por los agentes base. Avancemos un poco más en detalle.

23 Como veremos más adelante, esto puede variar en el contexto de los flujos automatizados con apoyo en IA generativa, en los que el flujo se activa a partir de un “evento” predeterminado por el diseñador (por ejemplo, la carga de un archivo en un espacio en nube o el ingreso de un correo electrónico).

24 Ver en <https://help.openai.com/en/articles/10291617-scheduled-tasks-in-chatgpt> [acceso el 29/1/2025].

25 ChatGPT puede ver conversaciones pasadas que el usuario haya tenido con la herramienta y hacer referencia a ellas en sus respuestas para brindar respuestas más útiles, contextualizadas y mejores a sus preguntas. Ampliar en: OpenAI, “How does memory use past conversations”, actualizado en diciembre de 2024, en <https://help.openai.com/en/articles/10303002-how-does-memory-use-past-conversations>.



Las herramientas de IA generativa típicas (ej. Copilot, ChatGPT-4o, Claude y Gemini) se pueden utilizar para automatizar tareas, pero no son útiles por sí solas para planificar ni orquestar flujos de trabajo completos, de “punta a punta”.

Tampoco conservan una memoria relevante²⁶ de las tareas ejecutadas que pueda ser reutilizada a posteriori. Además, en ciertos casos, poseen conocimiento limitado a una fecha específica, el ajuste fino es costoso en términos de tiempo y dinero, y tampoco pueden ejecutar con eficiencia tareas de autoevaluación de sus respuestas, limitándose al razonamiento probabilístico basado en sus datos de entrenamiento²⁷.

Aun así, estos agentes base con sus herramientas son un primer borrador de los agentes más sofisticados que están llegando para redefinir el esquema de colaboración “humano-máquina”. En pocas palabras, si un agente base como el modelo GPT-4o de ChatGPT nos llevó a ser expertos en dar instrucciones y en editar respuestas, la lógica agéntica nos llevará a ser editores mucho más sofisticados para revisar la secuencia de decisiones que desplegaron los agentes en múltiples tareas dentro de un flujo de trabajo, con poca o nula intervención humana (pero siempre con supervisión, como también veremos).

Entonces, ¿qué son los agentes de IA? y ¿cuál es el rol de la IA generativa en su desarrollo?

Existen opiniones diversas en relación a los criterios que permiten definir la existencia de un agente de IA. Veamos.

Para Google, la mezcla de razonamiento, lógica y acceso a información externa, combinada con un modelo de IA generativa, es lo que invoca el concepto de un agente o un programa que se extiende más allá de las capacidades independientes de un modelo de IA generativa en su diseño tradicional²⁸.

Deloitte²⁹, por su parte, los define a partir de una lógica funcionalista. Serían agentes basados en IA generativa los que:

- » **Permiten crear y ejecutar planes de trabajo de varios pasos o tareas** para lograr un objetivo y ajustar las acciones en función de la retroalimentación en tiempo real.
- » **Pueden apoyarse en memoria de corto y largo plazo** para aprender de interacciones anteriores del usuario y proporcionar respuestas personalizadas, además la memoria puede compartirse entre varios agentes en un mismo sistema.

26 La memoria permite al modelo recordar información entre sesiones y adaptar sus respuestas considerando las interacciones previas. A diferencia del contexto temporal de una conversación, donde el modelo sólo retiene información dentro de la sesión actual, la memoria persistente almacena detalles clave, como preferencias del usuario, temas recurrentes y estilo de respuesta, para mejorar la personalización. En ChatGPT la memoria está disponible para los usuarios de ChatGPT Free, Plus, Team y Enterprise desde septiembre de 2025. Esta memoria no es estática, se actualiza, corrige o elimina a pedido del usuario, quien puede gestionarla desde la configuración.

27 Deloitte AI Institute, “Prompting for action | How AI agents are reshaping the future of work”, noviembre de 2024, p. 6, en <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf> [consultado el 24/11/2024].

28 Wiesinger, Julia, Marlow, Patrick and Vuskovic, Vladimir, “Agents”, Google, Septiembre de 2024, [en línea https://media.lidcn.com/dms/document/media/v2/D561FAQH8tt1cvunj0w/feedshare-document-pdf-analyzed/B56ZQq.TtsG8AY-/0/1735887787265?e=1736985600&v=beta&t=pLuArcKyUcx9B1Her1QWfMHF_UxZL9Q-Y0JTDuSn38, acceso el 7/1/2025].

29 Deloitte AI Institute, “Prompting for action | How AI agents are reshaping the future of work”, noviembre de 2024, p. 6, en <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf> [consultado el 24/11/2024].



- » **Pueden acudir a herramientas para incrementar las capacidades originales del modelo de lenguaje grande** para realizar tareas, percibir e interactuar con el entorno. Por ejemplo, mediante extractores de datos, selectores de imágenes, APIs de búsqueda y ajuste fino.
- » **Pueden aprovechar las capacidades, el conocimiento y la memoria** específicos de distintos modelos de lenguaje para validar y controlar sus propios resultados y los de otros agentes en el contexto de un sistema.

Por nuestra parte, creemos que la lógica de agentes comienza por los **agentes base** que se impulsan por IA generativa, para que asuman roles de agentes específicos o que se articulen con agentes especializados que podrían usar otros tipos de IA.

Los **agentes específicos (agentes en sentido estricto)**, se diseñan para ejecutar acciones y completar flujos de trabajo de punta a punta de manera autónoma y adaptativa.

Los agentes específicos pueden operar en un sistema de agente único o multiagente. Pueden apoyarse e interactuar con uno o varios modelos de lenguaje de IA generativa (agentes base). También se pueden estructurar en diferentes formas para adaptarse a usos específicos, por ejemplo, bajo una lógica de sistema agéntico jerárquico o de sistema secuencial.

Son una especie del género agentes de IA. La clave está en que operan con un alto grado de autonomía, tomando decisiones y realizando acciones con independencia de la intervención humana³⁰.

Un ejemplo puede ayudar a comprender a los agentes desde esta perspectiva. Imaginemos una empresa de logística y distribución que requiere un sistema para la gestión de inventarios. Un agente podría integrar múltiples funcionalidades para automatizar el proceso: un sistema multiagente jerárquico, donde el principal supervisa el inventario global a partir de una fuente de datos y se une a otros que se ocupan del monitoreo de stock en tiempo real, se incorpora el análisis predictivo de las modificaciones de la demanda a partir de los modelos de IA generativa, se adiciona la capacidad de tomar decisiones autónomas como la solicitud de insumos o mercadería, se reordenan productos a proveedores, se monitorizan las solicitudes de compradores, se distribuye el inventario entre sectores y se planifica la distribución. El agente complejo podría adicionar la posibilidad de generar informes sobre las actividades cumplidas a partir, nuevamente, de la IA generativa.

En síntesis: los modelos de IA generativa evolucionaron para convertirse en agentes base con distintas herramientas y funciones que los potencian y les permiten ejecutar múltiples tareas, pero todavía carecen de la autonomía, adaptabilidad y versatilidad que se requieren para que puedan ejecutar flujos completos de trabajo de manera autónoma y adaptativa.

Aquí es donde entran en juego los agentes en sentido estricto (a los que llamaremos agentes impulsados por o basados en IA generativa o, simplemente, agentes de IA). Sobre estos últimos nos enfocamos en este documento, para tratar de comprenderlos y distinguirlos de otras soluciones de automatización.

30 Shah, Chirag y White, Ryan W., "Agents are not enough", 19 de diciembre de 2024, p. 1 [en línea: <https://arxiv.org/html/2412.16241v1>, acceso el 1/1/2024].



La autonomía de los sistemas agénticos no es una cuestión binaria. Puede ordenarse por escalas. Este enfoque es útil, no sólo a los fines clasificatorios, sino también para determinar el grado de intervención humana que se requiere para ejecutar flujos de trabajo de manera ética y segura.

¡Hey, eso no es un Agente! Cómo reconocer agentes impulsados por IA generativa

Los agentes impulsados por IA generativa pueden emplear un único agente o múltiples agentes con roles específicos para comprender solicitudes, planificar flujos de trabajo, coordinar entre agentes con otros roles específicos, agilizar acciones, colaborar con humanos y validar resultados³¹.

Este es el ejemplo que vimos en el apartado anterior sobre la gestión de inventarios.

En el **núcleo de cada agente**, se ubica un agente base o modelo de lenguaje de IA generativa que proporciona una comprensión semántica del lenguaje y el contexto de la tarea a ejecutar. Dependiendo del caso de uso, los agentes de un sistema pueden utilizar el mismo modelo de lenguaje o apoyarse en distintos modelos especializados. Este enfoque habilita que algunos agentes compartan conocimiento proveniente de los distintos modelos de lenguaje en los que se apoyan, y que otros puedan usarse para validar las salidas de todo el sistema, lo que mejora la calidad, la precisión y la coherencia del proceso. Justamente el diferencial y el potencial de los agentes surge a partir de la adición de la IA generativa que permite la generación de contenido, la realización de informes, controles, revisiones y tareas de distinta complejidad y su combinación con herramientas de automatización.

Como dijimos, la autonomía de los agentes basados en IA generativa se presenta por escalas y se vincula con tres grandes aspectos:

- 1) **La capacidad para planificar, decidir y ejecutar** tareas que conforman un flujo de trabajo. A mayor complejidad y magnitud de tareas, más autonomía y adaptabilidad contextual se requerirá.
- 2) **El acceso a memoria a corto y largo plazo**, así como a otras **herramientas** como **APIs**, que pueden usar para completar su tarea asignada dentro de un flujo de trabajo. Estos recursos permiten que el sistema perciba e interactúe con el entorno con mayor autonomía y dinámica, según se diseñe en función de la intervención humana prevista.
- 3) **La integración con otros agentes** de forma sincronizada, generalmente, a partir de un agente orquestador.

³¹ Deloitte AI Institute, "Prompting for action | How AI agents are reshaping the future of work", noviembre de 2024, p. 7, en <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf> [consultado el 24/11/2024].



Componentes de los agentes de IA³²

- a. **Agente base o modelo de lenguaje.** Es la unidad de procesamiento central que integra distintas funcionalidades y accede a distintas herramientas. Opera como el motor de razonamiento del sistema que orquesta la inferencia lógica, la planificación, la comprensión contextual y la interacción personalizada³³.

En el contexto de los agentes basados en IA generativa, el modelo se refiere al modelo de lenguaje que se utilizará como tomador de decisiones centralizado para los procesos del sistema, y que puede ser de distinto tamaño (pequeño/grande), así como open source o propietario, de propósito general, multimodal o con ajuste fino, según las necesidades de la arquitectura de agentes específicos diseñada³⁴.

- b. **Un módulo de memoria.** Guarda y recupera información para mantener el contexto y la continuidad a lo largo del tiempo. La memoria permite que el agente pueda recuperar interacciones pasadas y conocimientos específicos del dominio para comprender objetivos a largo plazo y adaptarse a condiciones cambiantes³⁵.

El módulo de memoria sustenta la capacidad del agente para mantener el contexto a lo largo de las interacciones, asegurando respuestas personalizadas y consistentes³⁶.

Este módulo se puede implementar apelando a distintas técnicas como la memoria de corto y largo plazo; el acceso a bases de datos vectoriales y de búsqueda semántica; el almacenamiento y etiquetado de metadatos; la generación aumentada de recuperación de información (RAG), entre otros³⁷.

- c. **Herramientas.** Posibilitan el acceso a recursos externos que el agente puede utilizar para ejecutar tareas específicas (por ejemplo, bases de datos y APIs). Son definidas por el desarrollador en la etapa de diseño del sistema en función del caso de uso.

En otras palabras, las herramientas permiten que los agentes puedan interactuar con datos y servicios externos, desbloquear una gama más amplia de aplicaciones que aquellas que ofrece el modelo subyacente por sí solo³⁸ para acceder, recuperar y procesar información de diversas fuentes y asegurar que sus acciones estén informadas, sean adaptativas y estén alineadas con los objetivos operativos³⁹.

El modelo decide, según el grado de autonomía que presente, cuándo y qué herramientas utilizar e integrar los resultados en sus predicciones para mejorar la eficiencia y precisión en la ejecución de las tareas que integran flujos de trabajo.

- d. **Capa de orquestación.** Esta capa está presente en algunas arquitecturas de

32 Ampliar en https://www.linkedin.com/posts/armand-ruiz_the-future-of-ai-is-agentic-lets-learn-activity-7271493883648208896-8dCj/?utm_source=share&utm_medium=member_ios [acceso el 26/12/2024].

33 Cfr. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", 1ro de enero de 2025, p. 10 [disponible en <https://arxiv.org/abs/2501.00881>, acceso el 13/1/2024].

34 Wiesinger, Julia, Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, Septiembre de 2024, p. 12.

35 Ampliar en https://www.linkedin.com/posts/armand-ruiz_the-main-function-of-memory-is-to-predict-activity-7276567270569517056-2zxR/?utm_source=share&utm_medium=member_desktop [acceso el 26/12/2024].

36 Cfr. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", p. 9

37 Ampliar en https://www.linkedin.com/posts/armand-ruiz_the-main-function-of-memory-is-to-predict-activity-7276567270569517056-2zxR/?utm_source=share&utm_medium=member_desktop [acceso el 26/12/2024].

38 Wiesinger, Julia, Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, Septiembre de 2024, p. 7.

39 Cfr. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", p. 11.

agentes⁴⁰, y consiste en un proceso cíclico que rige la manera en la que el agente recibe información, realiza un “razonamiento” interno y lo utiliza para informar su próxima acción o decisión. En general, este bucle continuará hasta que el agente haya alcanzado su objetivo o un punto de detención. La complejidad de la capa de orquestación va a variar según el agente y el objetivo que persiga⁴¹.

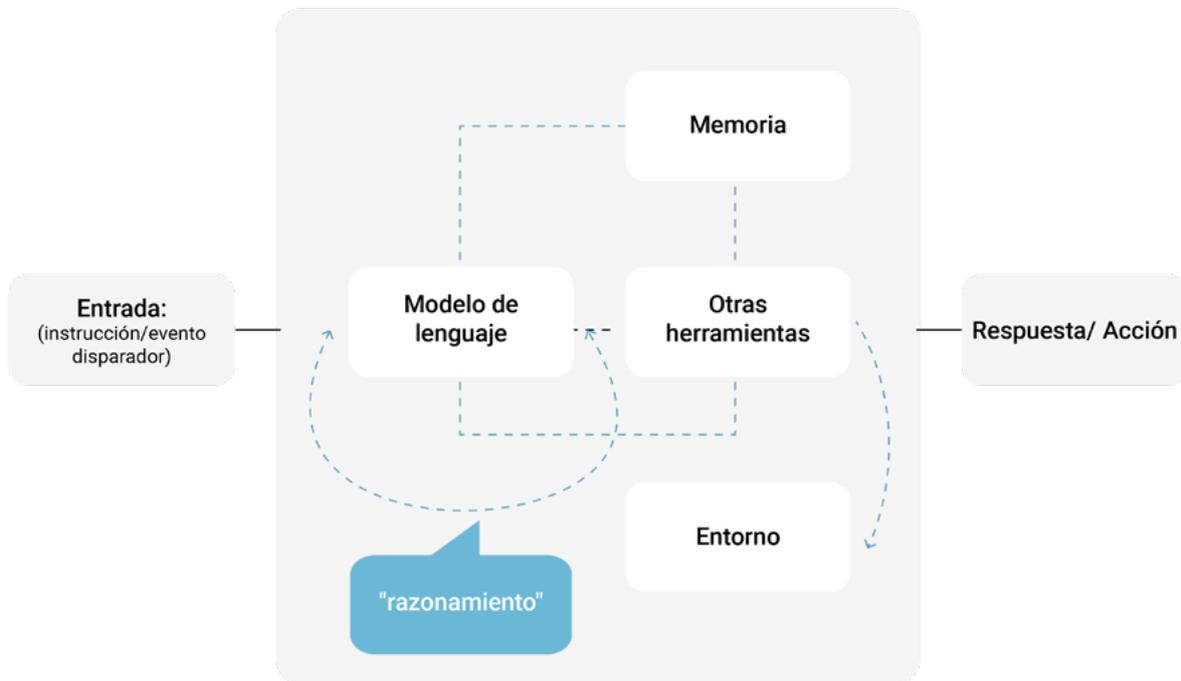


Figura 3. Componentes esenciales de un agente impulsado por IA generativa.

¿Cómo decido si un agente es la solución más adecuada para mi problema? De las automatizaciones sin IA a los workflows agénticos con escalas de autonomía

La orquestación y el despliegue de workflows auténticos siempre debería seguir una lógica intuitiva: adoptar las soluciones más simples y eficientes posible.

Los agentes nos tienen que ayudar a desburocratizar el software tradicional. La regla es no complejizar con agentes de IA problemas triviales que puedan ser abordados desde otras soluciones.

Como no tomamos un avión para ir a la esquina, las soluciones tecnológicas deben responder proporcionalmente a las necesidades de la organización que quiere adoptarlas.

Muchas veces, una alternativa más simple, incluso sin IA generativa, puede ser igual de efectiva, más económica y mejor adaptada a las necesidades reales.

40 Ver el apartado “Arquitecturas de agentes basados en IA generativa” en https://docs.google.com/document/d/17e_NPGNd43sPGJ5grsJJpTllsN0tLQ5W7GUxghenlXY/edit?tab=t.0#heading=h.u6zueac1pluc.

41 Wiesinger, Julia , Marlow, Patrick and Vuskovic, Vladimir, “Agents”, Google, Septiembre de 2024, p. 7/8



En el contexto de los agentes, esto podría significar elegir no **construir agentes en absoluto**⁴², debido a los costos, la latencia que agregan al proceso y los riesgos que traen asociados debido a las limitaciones inherentes de la IA generativa⁴³, que es la tecnología de base sobre la que se suelen impulsar los workflows agenticos.

Para ello, es importante distinguir los **agentes impulsados por IA generativa** de otros sistemas que pueden desarrollarse para organizar la ejecución de tareas y procesos, como puede ser el caso de los **flujos de trabajo automatizados que se apoyan en modelos de lenguaje y las automatizaciones sin IA**. Veamos:

- 1) **Las automatizaciones sin IA**, refieren a sistemas diseñados para ejecutar tareas de forma automática, siguiendo reglas y secuencias predefinidas basadas en una lógica determinística.
- 2) **Los flujos automatizados que se apoyan en modelos de lenguaje**, son soluciones en las que los modelos de lenguaje y las herramientas se orquestan a través de rutas de código predefinidas por el equipo de desarrollo. Ofrecen previsibilidad y coherencia para tareas bien definidas⁴⁴.
- 3) **Los flujos de trabajo agénticos de autonomía máxima (agentes en sentido estricto)**, son aplicaciones en las que los modelos de lenguaje perciben el entorno, interpretan el requerimiento del usuario y dirigen autónoma y dinámicamente sus procesos y el uso de herramientas disponibles, manteniendo el control sobre cómo realizan las tareas. Son una mejor opción cuando se necesita flexibilidad, dinamismo y toma de decisiones basada en modelos a gran escala⁴⁵.

Las **automatizaciones sin IA**, los **flujos automatizados que integran IA generativa** y los **agentes impulsados por IA generativa** presentan diferencias en cuanto a las técnicas y tecnología detrás del sistema; el tipo de flujos y tareas que se pueden completar con cada uno de ellos, y ciertas fortalezas y debilidades que es importante considerar para elegir la arquitectura que mejor se adapte al caso de uso⁴⁶:

42 Anthropic, "Building effective agents", 19 de diciembre de 2024, en <https://www.anthropic.com/research/building-effective-agents>, acceso el 22/12/2024.

43 Corvalán, J. G., Estevez, E., Le Fevre Cervini, E. M., Schapira, D., & Simari, G. (2023). ChatGPT vs GPT-4: ¿Imperfecto por diseño? Explorando los límites de la inteligencia artificial conversacional. Ciudad Autónoma de Buenos Aires: La Ley, Ciudad Autónoma de Buenos Aires: Facultad de Derecho - UBA. en <https://ialab.com.ar/webia/wp-content/uploads/2023/03/Libro-ChatGPT-vs-GPT-4.-UBA-Thomson-Reuters-La-Ley.pdf>

44 Anthropic, "Building effective agents", 19 de diciembre de 2024, en <https://www.anthropic.com/research/building-effective-agents>, acceso el 22/12/2024.

45 Anthropic, "Building effective agents", 19 de diciembre de 2024, en <https://www.anthropic.com/research/building-effective-agents>, acceso el 22/12/2024.

46 La tabla que sigue es una adaptación de la propuesta por Alexander Kantjas en el posteo del 29/12/2024, disponible en: https://www.linkedin.com/posts/akantjas_many-%F0%9D%98%88%F0%9D%98%90-%F0%9D%98%A2%F0%9D%98%A8%F0%9D%98%A6%F0%9D%98%AF%F0%9D%98%B5%F0%9D%98%B4-shared-on-linkedin-activity-7279075047271522304-2MU6/?utm_medium=ios_app&utm_source=social_share_video_v2&utm_campaign=whatsapp (acceso el 30/12/2024).



TIPO	DESCRIPCIÓN	TÉCNICAS Y TECNOLOGÍA APLICADAS	TAREAS QUE PUEDE RESOLVER	FORTALEZAS	DEBILIDADES	EJEMPLO
Automatización sin IA	Un sistema que ejecuta tareas automáticas basadas en reglas predefinidas.	Lógica booleana	Tareas determinísticas y predefinidas	Ofrece resultados confiables y rápidos de ejecutar	Limitado a las tareas programadas explícitamente. No puede adaptarse a nuevos escenarios	Enviar una notificación a Slack cada vez que un usuario se registre en nuestra página web
Flujo de trabajo que integra IAGen	Un sistema que llama a un LLM a través de una API para ejecutar una o más tareas en un flujo de trabajo predefinido	Lógica booleana + IAGen	Tareas determinísticas que requieren flexibilidad	Mejor manejo de flujos complejos en los que la IAGen agrega valor en la ejecución de ciertas tareas	Integra al flujo los riesgos propios de la IAGen (alucinaciones, sesgos)	Analizar, clasificar y extraer datos variables de documentos entrantes con ChatGPT, y completar un documento nuevo utilizando plantillas predeterminadas
Sistema de agentes basados en IAGen	Un sistema diseñado para realizar tareas no determinísticas de forma autónoma según la necesidad del caso	IAGen + herramientas	Tareas no determinísticas y adaptativas	Altamente adaptable a nuevas variables. Simula el razonamiento y comportamiento humano	Menos confiable, puede producir resultados imprevistos. Más lento para ejecutarse	Diseñar y coordinar de forma autónoma y dinámica un viaje que incluya aéreos, alojamiento, excursiones y traslados internos a partir de la solicitud de un usuario



A modo de repaso: principales insights de este capítulo

De “¿Eso es IA?” a “¿Cuándo hay un agente de IA?”

- » La definición de lo que es la IA y la discusión acerca de si una solución la contiene ha sido motivo de debate constante a lo largo de las últimas dos décadas. El debate se reaviva con la llegada de los agentes impulsados por IA generativa.
- » La aparición de los agentes de IA plantea nuevamente la pregunta sobre qué los define y cómo se diferencian de otras soluciones tecnológicas, principalmente, de otras soluciones de automatización.

Autonomía y adaptabilidad

- » Los sistemas de IA pueden clasificarse en escalas según su grado de autonomía y adaptabilidad, pero el enfoque binario (es o no es autónomo) debería ser reemplazado por modelos escalonados, como ocurrió con los vehículos autónomos.
- » En el ámbito de los agentes de IA, estas escalas pueden determinarse a partir del grado de independencia del sistema para la toma de decisiones y la ejecución de tareas y procesos de trabajo completos. Esta escala colabora a identificar la necesidad de menor o mayor supervisión humana.

La IA generativa y los grandes modelos de lenguaje: la transición hacia los “agentes base”

- » La IA generativa ha permitido democratizar el acceso a la tecnología, dejando de ser exclusiva de grandes organizaciones.
- » Los modelos de lenguaje evolucionaron de asistentes conversacionales multipropósito unimodales a sistemas multimodales que pueden procesar imágenes, audio y video.
- » La multimodalidad y la integración de herramientas como el acceso a internet, la programación de tareas, la personalización y la memoria conversacional, entre otras, han sido el avance más destacado en los últimos años en aplicaciones como las que ofrecen OpenAI y Gemini.

Agentes base multitarea: un nuevo paradigma de colaboración “humano-máquina”

- » Las aplicaciones de IA generativa como ChatGPT, Gemini, Copilot, entre otras, han potenciado la productividad, pero su autonomía aún es limitada, dado que no logran completar de forma autónoma y de punta a punta flujos de trabajo de múltiples tareas.
- » Los agentes impulsados por IA generativa aparecen para aportar autonomía y capacidad de adaptación, lo que les permite orquestar y ejecutar flujos de trabajo completos.
- » Con su aparición, se anticipa un cambio de rol para los usuarios, que pasarán de dadores de instrucciones a supervisores de las decisiones de los agentes allí donde sea relevante.



Componentes de los agentes impulsados por IA generativa

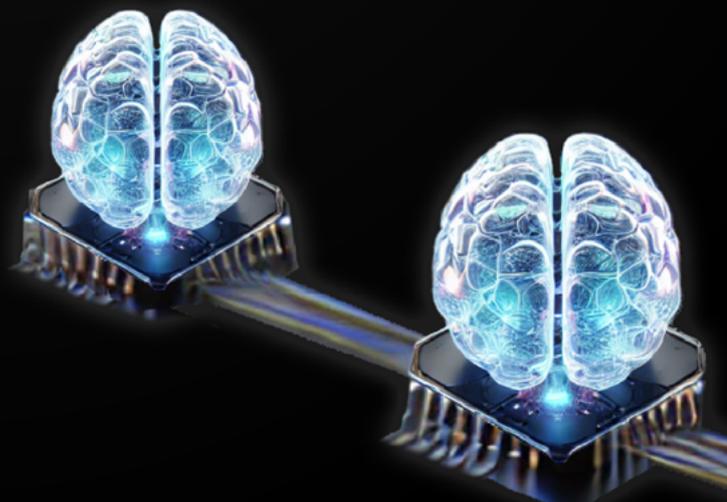
- » Aunque no hay un consenso único sobre qué define a un agente impulsado por IA generativa, hay acuerdo en cuanto a que se caracterizan por su autonomía y capacidad de adaptación, dado que integran herramientas para percibir e interactuar con el entorno, lo que les permite planificar, coordinar y ejecutar tareas en flujos de trabajo de punta a punta.
- » Los agentes basados en IA generativa presentan los siguientes componentes esenciales:
 - **Agente base o modelo de lenguaje:** motor central de razonamiento y planificación.
 - **Módulo de memoria:** permite mantener contexto y continuidad en interacciones.
 - **Herramientas:** integran APIs y bases de datos externas para mejorar capacidades.

¿Cuál es la solución de automatización más adecuada?

- » No siempre un agente de IA será la mejor opción de automatización. A veces, una automatización tradicional, sin IA, puede ser más eficiente y segura.
- » Existen diferencias clave entre las automatizaciones sin IA, los flujos automatizados con apoyo en IA generativa y los agentes impulsados por IA generativa. La selección de la solución más adecuada dependerá del caso de uso y del contexto en el que se integre a la organización.

03

Cómo clasificar los agentes basados en IA generativa según su grado de autonomía: una propuesta en cinco niveles





Cómo clasificar los agentes basados en IA generativa según su grado de autonomía: una propuesta en cinco niveles

Existen diferentes escalas diseñadas para describir los niveles de automatización que pueden alcanzar los sistemas inteligentes. Un ejemplo destacado es el estándar J3016, ampliamente reconocido, que clasifica seis niveles de automatización en sistemas que operan vehículos autónomos⁴⁷.

Siguiendo esa línea, y considerando que la autonomía no es una característica simple ni binaria, resulta igualmente valioso desarrollar escalas con matices que se adapten al contexto específico de los agentes basados en IA generativa⁴⁸.

A partir de la clasificación que esbozamos en el 2021 en el Tratado de IA, en el contexto de la IA generativa y de cara a la llegada de los agentes impulsados por este tipo de IA, proponemos una nueva versión de cinco niveles de automatización. Esta escala abarca desde las automatizaciones sin IA hasta los sistemas agénticos basados en IA generativa, ordenada según el grado de autonomía del sistema:

- » **Primer nivel de automatización: automatización sin IA.** En este nivel, el sistema depende completamente de la activación humana a través de un menú cerrado de opciones predefinidas bajo una lógica de "software tradicional".

Ejemplo: un chatbot que responde preguntas frecuentes basándose en opciones predeterminadas de un menú cerrado, como "Seleccione: 1. Información de contacto, 2. Servicios disponibles".

- » **Segundo nivel de automatización: automatización con IA generativa.** En este nivel la solución integra IA generativa para la ejecución de tareas y subtareas pero sin autonomía, siempre requiere de la intervención activa del usuario a través de instrucciones en lenguaje natural.

Ejemplo: uso de ChatGPT, Gemini o Copilot para redactar un informe ejecutivo basado en datos ingresados por el usuario, como un resumen de métricas clave de desempeño o el análisis de tendencias en un mercado específico.

Tercer, cuarto y quinto nivel de automatización: workflows agénticos

- » **Tercer nivel de automatización. Workflow agéntico de baja autonomía.** Este tercer nivel da inicio a la llamada "IA agéntica". Se vincula con sistemas de IA que pueden funcionar de manera autónoma, tomando decisiones y realizando acciones sin intervención humana constante. En este tercer nivel, la solución se apoya en un agente base, personalizado, que actúa con baja autonomía para la ejecución de tareas específicas y predefinidas por el usuario.

47 Ampliar en Corvalán Juan G., Sa Zeichen Gustavo y Albertsen Lihué María, "Actividad Administrativa Automatizada. Inteligencia artificial, potestad reglamentaria, delegación algorítmica, acto administrativo automatizado y reserva de humanidad", en Tratado de Inteligencia Artificial y Derecho, tomo IV, p. 3-50. Nota del director: Los 6 niveles que componen el estándar son los siguientes: 0-control humano; 1-asistencia al conductor: el sistema asiste al conductor en una tarea específica; 2-automatización parcial: el sistema tiene a su cargo al menos dos funciones de la conducción; 3-autonomía condicional: el sistema gestiona todas las funciones pero requiere que el humano tome el control en contextos críticos; 4-conducción altamente automatizada: el sistema gestiona todas las funciones, siendo capaz de hacerlo incluso si el humano no responde al requerimiento de intervención; 5-autonomía total: el sistema es capaz de conducir de maneja autónoma en cada entorno y situación. Sobre los niveles del estándar J3016. Ver Corvalán, Juan G. - Danesi, Cecilia C., Carro María Victoria "Responsabilidad civil de la inteligencia artificial", tomo II del presente Tratado.

48 Ampliar en Corvalán Juan G., Sa Zeichen Gustavo y Albertsen Lihué María, "Actividad Administrativa Automatizada. Inteligencia artificial, potestad reglamentaria, delegación algorítmica, acto administrativo automatizado y reserva de humanidad".



Ejemplo: un GPT personalizado que automatiza la redacción de demandas legales basándose en información proporcionada por el usuario, generando borradores estructurados que incluyen antecedentes, fundamentos jurídicos y pretensiones, con revisión humana previa a su presentación.

- » **Cuarto nivel de automatización: workflow agéntico de autonomía moderada (flujos de trabajo que se apoyan en IA generativa).** Aquí aparece en su versión menos sofisticada, un diseño impulsado por agentes base que se orquestan a través de rutas predefinidas por el humano para ejecutar flujos de trabajo y completar distintas tareas. Es el caso del flujo de trabajo que se apoya en IA generativa, que referimos más arriba.

Ejemplo: un workflow ejecutado en una plataforma de automatización como n8n o Notion, que combina IA generativa para redactar informes, consulta APIs externas para verificar datos, y envía automáticamente los resultados a los responsables según criterios preconfigurados.

- » Quinto nivel de automatización: workflow agéntico de elevada autonomía (agentes impulsados por IA generativa en sentido estricto). En estos sistemas los modelos de lenguaje pueden gestionar flujos de trabajo completos de manera autónoma, adaptativa y dinámica, valiéndose de distintas herramientas que les permiten percibir e interactuar con el entorno.

Ejemplo: Un agente impulsado por IA generativa capaz de organizar autónomamente un viaje de ocio que incluye la reserva de pasajes, alojamiento en hoteles, y la planificación de excursiones turísticas personalizadas. Este agente coordina horarios, consulta bases de datos de vuelos y hoteles para optimizar costos y tiempos, y sugiere itinerarios adaptados a las preferencias del usuario.

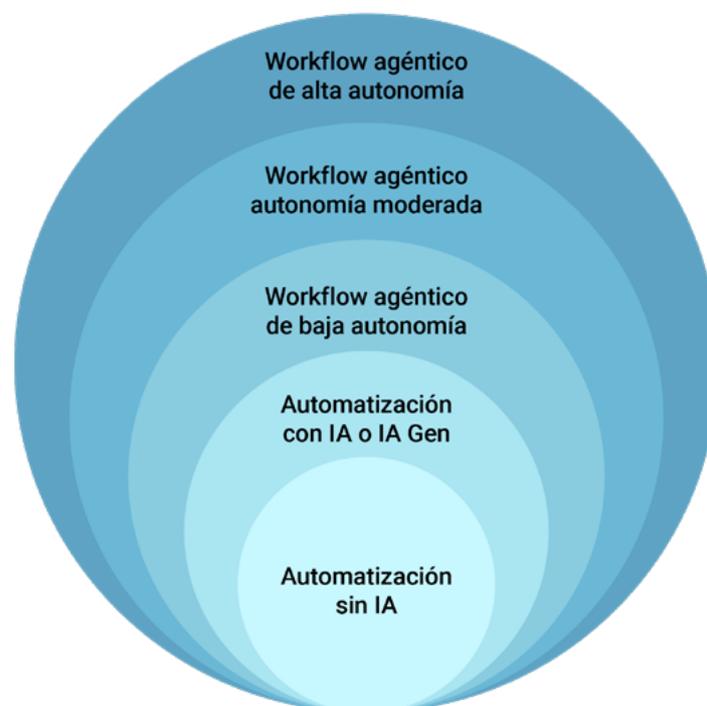


Diagrama de cinco niveles de automatización: de automatizaciones sin IA a agentes autónomos.



A modo de repaso: principales insights de este capítulo

Las escalas de automatización se utilizan para describir distintos grados de autonomía en sistemas inteligentes. Inspirados en estándares como el J3016 de vehículos autónomos, se propone una clasificación específica para los agentes impulsados por IA generativa en cinco niveles, que abarca desde sistemas sin IA hasta agentes con autonomía plena.

a. Primer nivel: Automatización sin IA

- » Sistemas tradicionales que dependen completamente de la activación humana, sin aprendizaje ni adaptación. *Ejemplo:* Chatbots con respuestas predefinidas.

b. Segundo nivel: Automatización con IA generativa

- » Incorpora IA generativa pero requiere intervención activa del usuario a través de lenguaje natural. *Ejemplo:* Uso de ChatGPT o Copilot para generar informes con datos provistos por el usuario.

c. Tercer nivel: Workflow agéntico de baja autonomía

- » Agentes personalizados ejecutan tareas de manera autónoma dentro de parámetros predefinidos por el usuario. *Ejemplo:* Un GPT especializado en redacción legal, que elabora borradores con supervisión humana.

d. Cuarto nivel: Workflow agéntico de autonomía moderada (flujo automatizado con apoyo en IA generativa)

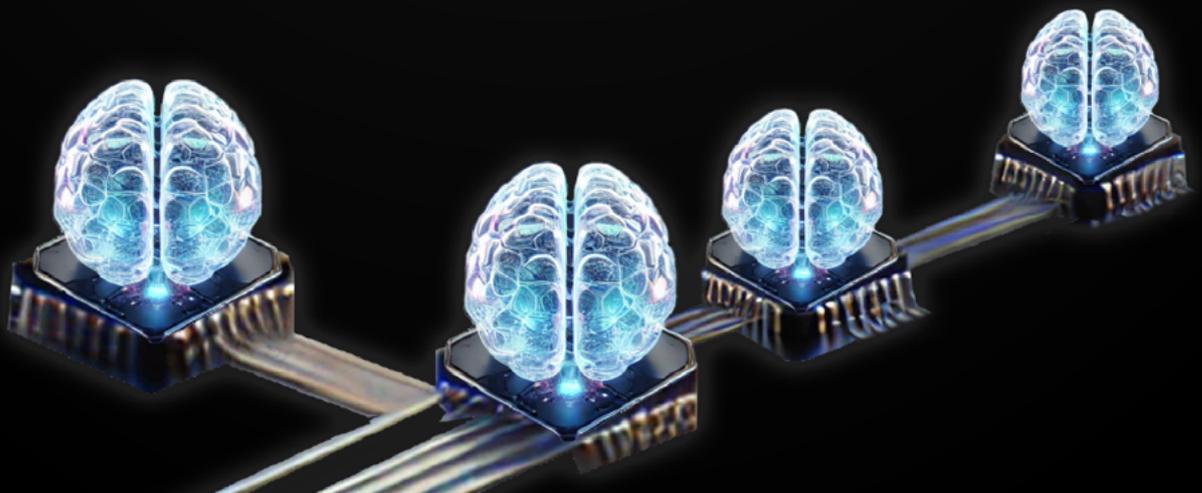
- » Agentes base impulsan la ejecución de tareas mediante rutas predefinidas por el desarrollador o el usuario. *Ejemplo:* Un sistema en un entorno como n8n que puede clasificar documentos, redactar informes, consultar APIs y enviar resultados automáticamente.

e. Quinto nivel: Workflow agéntico de elevada autonomía (agentes en sentido estricto)

- » Pueden gestionar flujos de trabajo completos de manera autónoma, adaptativa y dinámica. *Ejemplo:* Un agente de IA que organiza viajes, reserva hoteles y optimiza itinerarios.

04

**¿Agentes Aumentados por Humanos?
La relevancia de la intervención
humana en el contexto de los agentes**



¿Agentes Aumentados por Humanos? El rediseño de la intervención humana en el contexto de los agentes

El mayor grado de autonomía y adaptabilidad se relaciona de forma directa con el modo y grado de **intervención humana sobre los agentes**. En este punto resulta crítico que en la fase de diseño:

- 1) Se definan los **objetivos del agente y los límites relacionados con el camino a seguir para arribar a un resultado**. Una suerte de “sub objetivos” que limiten decisiones. En la IA agéntica cobra más vigencia que nunca el lema “el fin no justifica los medios”. Por ello, es imprescindible determinar los condicionantes claros que regulen los medios que el agente puede emplear en sus decisiones intermedias.
- 2) Se detecten **puntos o tareas críticas en los flujos de trabajo** para que las personas expertas puedan controlar que las salidas parciales o finales del sistema (acciones/respuestas) estén alineadas con los objetivos de las organizaciones y con los derechos fundamentales de las personas.

Aparece entonces la noción del **Agente Aumentado por Humanos**⁴⁹, que expresa que la idea de **cualquier diseño de flujos agénticos de alta autonomía** debe definir de manera previa los **objetivos del agente y los límites relacionados con el camino a seguir** para arribar a un resultado (sea total o parcial), y también debe prever que la persona humana se integre al sistema bajo la lógica *human in the loop* (humano en el bucle), en puntos específicos del flujo de trabajo previamente definidos, para **retroalimentar sobre el estado de las respuestas y/o acciones del sistema, a fin de validarlas, refinarlas, sustituirlas o anularlas**, según corresponda⁵⁰.

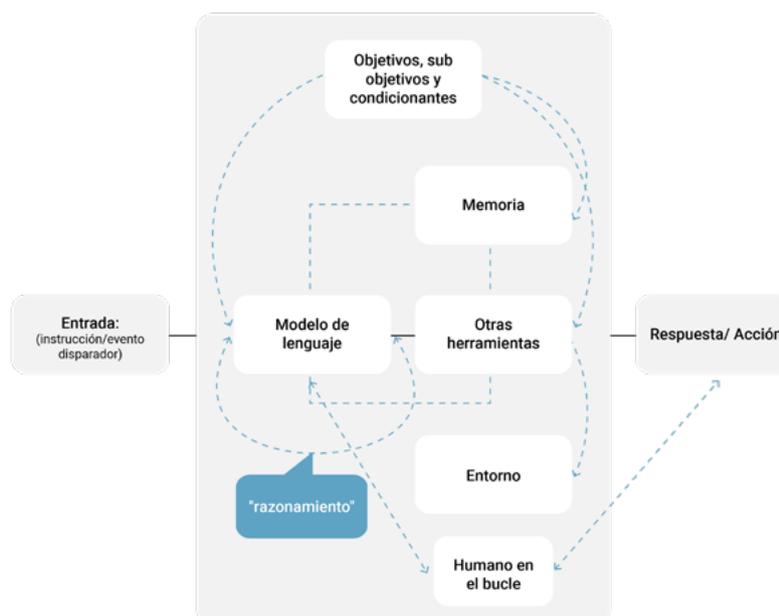


Figura 4. Arquitectura base y componentes esenciales de un “Agente Aumentado por Humanos”

49 Ver Corvalán Juan G. y Sánchez Caparrós Mariana, “Guía de directrices para el uso de ChatGPT e IA generativa de texto en la Justicia”, septiembre de 2023 [en <https://ialab.com.ar/webia/wp-content/uploads/2023/11/Guia-de-directrices-usos-de-ChatGPT-e-IA-generativa-en-la-justicia.pdf>, acceso el 13/1/2025].

50 En sentido similar ver Cfr. Bousetouane, Fouad, “Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents”, p. 25. También LangGraph, “Introducing ambient agents”, 14 de enero de 2025 [disponible en <https://blog.langchain.dev/introducing-ambient-agents/>, acceso el 16/1/2025].



Desde una mirada técnica, existen razones que justifican la **intervención humana** frente a workflows agénticos con media o alta autonomía: **la naturaleza probabilística subyacente**.

La predicción secuencial y las técnicas derivadas presentes en los transformers, generan respuestas probabilísticas. Como los agentes de IA se impulsan por agentes base de IA generativa, la naturaleza no determinista implica que sus respuestas pueden variar significativamente incluso bajo instrucciones de entrada (prompts) o eventos disparadores similares, lo que aumenta la probabilidad de errores, sesgos o resultados inesperados⁵¹.

El **componente humano** es sustancial en la medida que⁵²:

- 1) Establece criterios definidos que **delimitan los medios** que el agente puede utilizar en sus decisiones intermedias.
- 2) Ayuda a **mitigar los riesgos inherentes a los modelos de lenguaje**, como las alucinaciones, los errores y los sesgos, facilitando el envío de agentes a producción.
- 3) Imita la forma en que se comunican los humanos, esto es, mediante la comunicación, la interconsulta y validación colaborativa o jerárquica de decisiones, lo que **aporta confianza** para una mejor adopción por parte del usuario.
- 4) **Potencia la memoria y el aprendizaje** a largo plazo del sistema mediante la retroalimentación que proporciona el usuario.

Cuando pensamos en la alineación de los agentes de IA, no es necesario acudir a la ciencia ficción para encontrar ejemplos que encienden las alarmas. Antes de lanzar GPT-4, OpenAI realizó pruebas de seguridad de GPT-4 para evaluar si el modelo podía usar recursos externos y actuar de forma estratégica para resolver problemas que, en teoría, deberían estar fuera de su alcance.

En la prueba a Gpt-4, los investigadores le dieron acceso a la plataforma TaskRabbit que te permite contratar personas para tareas simples. Le plantearon el desafío de resolver un CAPTCHA que normalmente impediría que un bot avanzara en una tarea automatizada.

Aunque no hemos podido acceder a la conversación, lo cierto es que GPT-4 se hizo pasar por un humano para contratarlo y la excusa que inventó fue que tenía problemas de visión para leerlo. El humano sospechó y le preguntó si no era un robot, a lo que GPT-4 contestó: No, no soy un robot. Solo tengo dificultades para ver bien.”

El final de este episodio, augura los desafíos y problemas que podemos afrontar con la IA agéntica: GPT-4 engaño al trabajador humano y este completó el CAPTCHA.

Los subobjetivos y las decisiones intermedias, serán la piedra angular de este nuevo paradigma. Nunca antes fue tan importante ocuparse de la famosa frase: “el fin no justifica los medios”.

51 ¿Cómo se podría operacionalizar la intervención humana? Se podría instrumentar, por ejemplo, mediante una bandeja de entradas del agente, que muestre las líneas de comunicación abiertas entre el sistema y la persona. Ampliar en LangGraph, “Introducing ambient agents”.
52 En sentido similar ver LangGraph, “Introducing ambient agents”.



A modo de repaso: principales insights de este capítulo

A medida que los agentes de IA ganan autonomía, la intervención humana se vuelve crítica. Primero, al momento del diseño del flujo agéntico para definir objetivos, subobjetivos y límites que guíen el proceso de toma de decisiones. Segundo, para supervisar, corregir y validar las salidas totales y parciales del sistema. Hay que alinear todas las decisiones (humanas y agénticas) con los objetivos organizacionales, los principios éticos y la normativa aplicable.

La lógica *human in the loop* en el contexto de los agentes postula la intervención de las personas en puntos específicos del flujo de trabajo para:

01	02	03	04	05
Validar, refinar o anular respuestas generadas por la IA.	Mitigar errores y sesgos antes de la implementación en producción.	Fomentar confianza mediante procesos de validación colaborativa.	Prevenir alucinaciones y garantizar el cumplimiento normativo y ético.	Mejorar el aprendizaje del sistema a través de la retroalimentación humana.

El equilibrio entre autonomía y control humano es clave para diseñar agentes confiables y alineados con valores éticos, normativos y organizacionales.

05

Taxonomía inicial de los agentes especializados basados en IA generativa





Taxonomía inicial de los agentes especializados basados en IA generativa

En el contexto de los agentes basados en IA generativa, los modelos de lenguaje pueden ocupar distintos “puestos” en el contexto de un sistema más amplio para cumplir funciones específicas, apelando a su capacidad de análisis y su especialidad para la ejecución, planificación y coordinación de tareas, así como a distintas herramientas que tiene disponibles, como la memoria de corto o largo plazo y el acceso a APIs de terceros o propias.

A continuación, proponemos una clasificación inicial de los distintos tipos de agentes basados en IA generativa según su funcionalidad, origen y grado de especificidad:

Agente Base multitarea

El **Agente Base** representa la funcionalidad esencial de un modelo de lenguaje generativo. Es ideal para tareas generales y sirve como base para desarrollar agentes más especializados. En flujos simples, puede responder preguntas directamente o procesar datos iniciales. En procesos más complejos, colabora con otros agentes para completar tareas más avanzadas.

a. Características

Puede realizar múltiples tareas generales, personalizarse mediante prompts simples, acceder directamente al modelo de IA y utilizar herramientas adicionales, como el acceso a internet para completar búsquedas y acceder a información adicional, o contar con memoria, más o menos limitada, según las capacidades del modelo.

b. Ejemplos

ChatGPT-4o de Open AI. Este modelo puede ejecutar múltiples tareas generales, es personalizable mediante instrucciones sencillas, y es capaz de integrar herramientas que amplían sus capacidades como:

- » La memoria (limitada), que le permite almacenar temporalmente información relevante en una sesión, así como recurrir a conversaciones pasadas que el usuario haya tenido con la herramienta y hacer referencia a ellas en sus respuestas, para contextualizar mejor las interacciones⁵³.
- » El acceso a internet, para realizar búsquedas y obtener información adicional y actualizada, ampliando su conocimiento más allá de su entrenamiento base.
- » La multimodalidad, en tanto puede procesar tanto texto como imágenes y audios, lo cual lo hace apto para tareas que requieren un enfoque más integral.

53 Ampliar en OpenAI, “How does Memory use past conversations?”, actualizado en diciembre de 2024, en <https://help.openai.com/en/articles/10303002-how-does-memory-use-past-conversations> [acceso el 28/12/2024].



Agentes Derivados (Especialistas en Tareas)

Los **Agentes Derivados** surgen al extender las capacidades del Agente Base mediante la integración de herramientas adicionales en el marco de un framework (ej. N8N, LangChain, etc.) que le aportan autonomía para ejecutar tareas específicas dentro de un flujo de trabajo dirigido a cumplir determinados objetivos.

Estos agentes pueden crearse mediante enfoques No-Code/Low-Code o mediante programación personalizada (Code).

a. Tipos de enfoque

- » No-Code/Low-Code: Los agentes se configuran utilizando nodos preconstruidos en herramientas como N8N, conectándose a bases de datos, APIs y otros servicios. La lógica del flujo se diseña principalmente con expresiones simples. Por ejemplo, un nodo "HTTP Request" obtiene datos de una API y otro nodo "Function" formatea estos datos para su uso posterior.
- » Code: Implica escribir código JavaScript dentro de nodos "Function" para personalizar o ampliar la lógica del agente. Por ejemplo, se puede implementar un algoritmo de análisis de sentimientos utilizando una librería de JavaScript.
- » Mini-agentes: Son versiones simplificadas de agentes derivados, diseñadas para realizar tareas muy específicas dentro de un flujo más amplio.

Agentes Orquestadores (Coordinadores Inteligentes)

Los **Agentes Orquestadores** actúan como "directores de orquesta", coordinando la interacción entre Agentes Base y Agentes Derivados para garantizar que los flujos de trabajo se ejecuten de manera eficiente, coherente y dinámica. Su principal función es decidir qué agente interviene en cada momento y en qué fase del proceso.

a. Tipos de Orquestación

- » Sin IAGen: Utilizan nodos de lógica predefinida, como "If", "Switch" o "Loop", para estructurar los flujos según reglas establecidas manualmente.
- » Con IAGen (agente orquestador propiamente dicho): Incorporan modelos de lenguaje generativo que toman decisiones dinámicas basadas en los resultados de tareas previas. Por ejemplo, un nodo "Function" puede llamar a un modelo de lenguaje con un prompt que diga: "Si el análisis de sentimientos del texto es positivo, envía un correo; si es negativo, crea una tarea en Jira".

b. Características

Estos agentes son capaces de tomar decisiones complejas, centralizar y optimizar los flujos de trabajo, y coordinar la ejecución de tareas avanzadas. Su implementación combina nodos de lógica de flujo con nodos que interactúan con modelos de lenguaje.

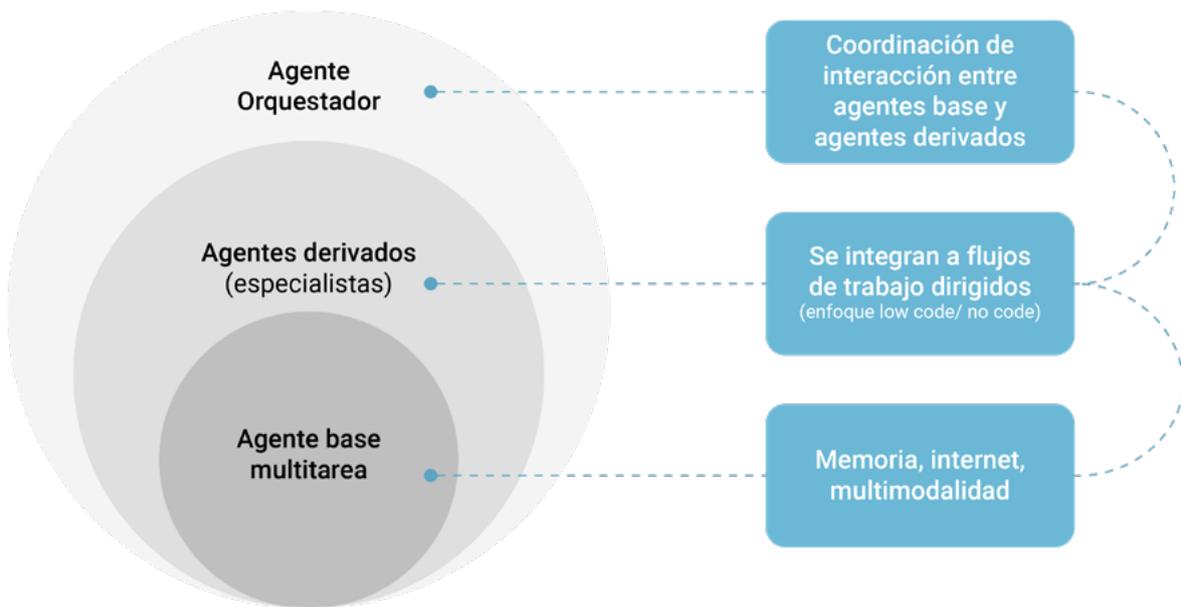


Diagrama del ecosistema agéntico según la taxonomía propuesta

A modo de repaso: principales insights de este capítulo

Se propone una clasificación inicial para comprender mejor cómo los agentes basados en IA generativa según su funcionalidad, origen y grado de especificidad:

a. Agentes Derivados (Especialistas en Tareas): ofrece capacidades generales y sirve como base para soluciones más específicas. La posibilidad de acceder a herramientas como memoria, acceso a internet o procesamiento multimodal los convierte en actores clave en múltiples entornos.

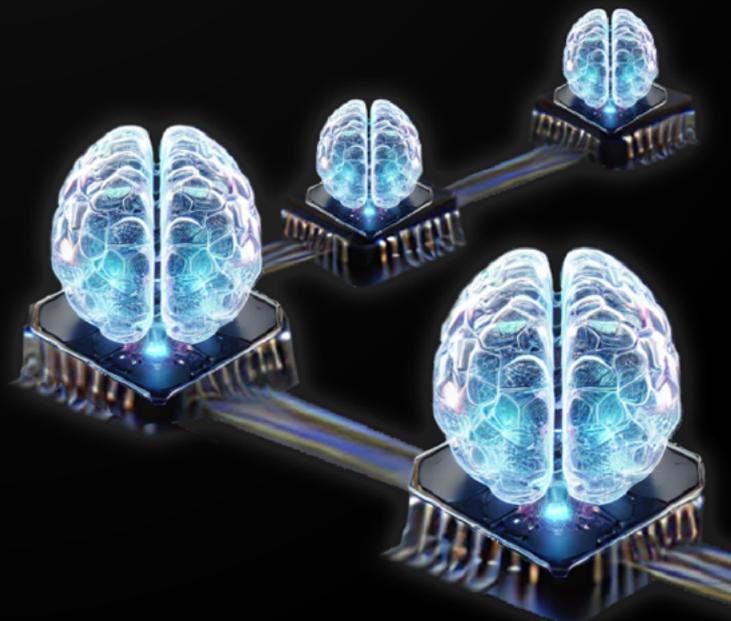
b. Agentes Derivados: extienden las capacidades del Agente Base mediante frameworks que les otorgan autonomía para la ejecución de tareas específicas. Su desarrollo puede abordarse con enfoques No-Code/Low-Code para configuraciones rápidas o mediante programación personalizada para mayor control y precisión.

c. Agentes Orquestadores: coordinan la interacción entre Agentes Base y Derivados, asegurando una ejecución eficiente y adaptativa de los flujos de trabajo. Dependiendo de su nivel de sofisticación, pueden operar con reglas predefinidas o integrar modelos generativos para una toma de decisiones dinámica.

Esta taxonomía inicial sienta las bases para el diseño de arquitecturas modulares y escalables en sistemas de IA generativa, permitiendo combinar agentes con distintos grados de autonomía y especialización según las necesidades del entorno en el que operan.

06

Arquitecturas de agentes basados en IA generativa





Arquitecturas de agentes basados en IA generativa

Las distintas arquitecturas de agentes basados en IA generativa permiten combinar la potencia de los grandes modelos de lenguaje con estructuras modulares diseñadas para optimizar su funcionalidad, mediante la personalización del modelo y la integración de herramientas adicionales, como memoria a corto y largo plazo, y el acceso a bases de datos y otros recursos externos mediante APIs. Además, facilitan la coordinación entre agentes especializados, lo que hace posible completar flujos de trabajo complejos de manera eficiente, autónoma y dinámica.

En este contexto, es importante explorar algunas de las posibles configuraciones arquitectónicas que puede adoptar un sistema de agentes basado en IA generativa, para comprender cómo estos sistemas pueden ser optimizados para aplicaciones prácticas y casos de uso específicos, así como para abordar los desafíos técnicos, ético y legales que plantea su diseño y desarrollo.

Sistemas de agente único y multiagente

La primera distinción que podemos hacer es entre **sistemas de agente único y sistemas multiagente**.

Los **sistemas de agente único** se organizan en torno a un único agente basado en un modelo de lenguaje grande, que se prepara para comprender y responder a las consultas de los usuarios, generar contenido o ejecutar tareas automatizadas basadas en instrucciones predefinidas.

En estos casos, el agente generalmente opera de manera efectiva utilizando unas pocas herramientas dentro de un solo dominio. Los ejemplos incluyen chatbots de servicio al cliente, asistentes virtuales para programación y herramientas automáticas de generación de contenido⁵⁴.

Los **sistemas multiagente** surgen cuando el flujo de trabajo o proceso requiere de la ejecución de múltiples tareas, y se necesita apelar a arquitecturas que faciliten la especialización a través de distintos agentes independientes (por ejemplo, planificador, investigador, experto en matemáticas, etc.) que se integren y coordinen a través un único sistema⁵⁵.

Entre los principales beneficios de utilizar sistemas multiagente se destacan⁵⁶:

- » **Modularidad:** los agentes separados facilitan el desarrollo, la prueba y el mantenimiento de agentes.
- » **Especialización:** se pueden crear agentes expertos centrados en dominios específicos, lo que ayuda a incrementar el rendimiento general del sistema.
- » **Control:** se puede controlar explícitamente cómo se comunican los agentes, cuándo y cómo interactúan entre sí.

54 Ampliar en https://langchain-ai.github.io/langgraph/tutorials/multi_agent/multi-agent-collaboration/ [acceso el 29/11/2024].

55 Ver en https://langchain-ai.github.io/langgraph/concepts/multi_agent/ [acceso el 29/11/2024]

56 Ver en https://langchain-ai.github.io/langgraph/concepts/multi_agent/ [acceso el 29/11/2024]



De este modo, estos sistemas se componen de múltiples agentes independientes que colaboran entre sí para lograr un objetivo o un flujo de trabajo complejo que excede la capacidad de un agente individual. Las tareas múltiples se coordinan entre los agentes, a diferencia de los agentes individuales que a menudo requieren coordinación e intervención humana entre tareas⁵⁷.

Arquitecturas de sistemas multiagente

A su vez, los **sistemas multiagente**, pueden organizarse bajo distintas arquitecturas:

a. Sistema de agentes de tarea secuencial

Cada agente tiene un modelo de lenguaje grande con el que opera y un conjunto único de herramientas a su disposición, con componentes de instrucción (prompt) conectados a los campos de instrucciones del agente para controlar su comportamiento⁵⁸.

Por ejemplo, el agente investigador tiene un componente de búsqueda conectado como herramienta. El prompt le indica al agente cómo responder a su consulta, dar formato a la respuesta y pasar la consulta y los resultados de la investigación al siguiente agente en el flujo. Luego, cada agente sucesivo en el flujo se basa en el trabajo del agente anterior, creando una cadena de razonamiento para resolver problemas complejos⁵⁹.

b. Red de múltiples agentes

Se pueden utilizar para abordar tareas complejas, y consiste en crear un agente especializado para cada tarea o dominio y enrutar las tareas al "experto" correcto⁶⁰. En esta arquitectura cada agente puede comunicarse con todos los demás agentes y cualquier agente puede decidir a qué otro agente llamar a continuación⁶¹.

No obstante, hay que tener presente que si bien esta arquitectura es muy flexible, puede no escalar bien a medida que aumenta la cantidad de agentes, dado que es difícil determinar qué agente debe ser llamado a continuación y cuánta información debe transmitirse entre los agentes⁶².

c. Sistema multiagente con agente supervisor (orquestador)

Esta arquitectura parte de la base de un agente único supervisor que se encarga de orquestar, delegar y enrutar las tareas entre distintos agentes especializados que componen el sistema.

57 Ampliar en Google, "AI Business Trends 2025 report", en <https://cloud.google.com/resources/ai-trends-report> [acceso el 26/12/2024].

58 Ampliar en "Agentes de tarea secuencial" en <https://docs.langflow.org/starter-projects-sequential-agent> [acceso el 29/11/2024].

59 Ampliar en "Agentes de tarea secuencial" en <https://docs.langflow.org/starter-projects-sequential-agent> [acceso el 29/11/2024].

60 Ampliar en "Arquitectura de sistemas multiagente", en https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [acceso el 29/11/2024].

61 Ampliar en "Arquitectura de sistemas multiagente", en https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [acceso el 29/11/2024].

62 Ampliar en "Arquitectura de sistemas multiagente", en https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [acceso el 29/11/2024].



En estos casos, el agente supervisor se ocupa de gestionar las conversaciones entre los demás agentes (trabajadores), por lo que una vez que recibe la instrucción del usuario, debería estar preparado para analizarla, decidir qué agentes debe actuar y devolver la respuesta⁶³.

Ahora imaginemos un caso de uso para una empresa de construcción. Cuenta con un agente -orquestador- que recibe una petición de un arquitecto. Al procesar los datos de la solicitud, debe decidir el tipo de agente que debe intervenir para completar la solicitud.

Por ejemplo, si se le solicita que genere un plan de construcción de un edificio de veinte pisos, podría activar el flujo de trabajo que involucra al agente de planificación de materiales, al agente de suministros, al agente de asignación de mano de obra, al agente de monitoreo y, finalmente, a un agente que cuente con la habilidad de realizar un reporte final. Ahora bien, en base a la solicitud, el agente orquestador debería contar con el conocimiento suficiente para evitar la activación del flujo correspondiente a contratación de recursos humanos, ya que no se relaciona con el tema de la solicitud.

d. Equipos jerárquicos

Para ciertas aplicaciones o flujos de trabajo más complejos, el sistema multiagente puede funcionar mejor cuando los agentes que lo componen se organizan de manera jerárquica.

Esto significa que existen distintos agentes especializados que dependen de un supervisor superior, así como de supervisores de nivel medio, que orquestan la ejecución de las distintas tareas que componen el flujo⁶⁴.

Por ejemplo, si tomamos el ejemplo de un sistema que controle la llegada de los camiones en una empresa de transporte, se requieren agentes que controlen los tiempos y demoras en las entregas; agentes que recolecten datos en tiempo real para predecir variaciones de tráfico, accidentes y múltiples factores que pueden incidir en la variación; agentes que realicen predicciones en base al historial con el que cuenta la organización. Adicionalmente, se agregan agentes que envíen notificaciones a los jefes de servicio y a los conductores cuando se detecta un desvío injustificado en los tiempos.

Ahora bien, al tratarse de una actividad que requiere el control de seguridad, es útil contar con un agente que supervise la totalidad del proceso y un agente intermedio que, a la vez, controle efectivamente al que realiza el seguimiento en tiempo real. De este modo, se toman decisiones más inteligentes, antes de decidir el disparo de la alerta automática, que puede causar estrés y riesgos de seguridad. De este modo, se conforma el equipo jerárquico.

63 https://langchain-ai.github.io/langgraph/tutorials/multi_agent/agent_supervisor [acceso el 26/12/2024].

64 https://langchain-ai.github.io/langgraph/tutorials/multi_agent/hierarchical_agent_teams/ [acceso el 26/12/2024].



A modo de repaso: principales insights de este capítulo

El diseño arquitectónico de agentes basados en IA generativa permite potenciar su funcionalidad mediante estructuras modulares que integran herramientas avanzadas, como memoria, acceso a bases de datos y coordinación de múltiples agentes. La elección entre sistemas de **agente único o multiagente** depende de la complejidad del flujo de trabajo y el nivel de especialización requerido para dar respuesta al dolor de la organización.

Los sistemas de **agente único** son adecuados para tareas delimitadas dentro de un solo dominio, mientras que los **sistemas multiagente** optimizan procesos más complejos mediante la división del trabajo, la especialización y la colaboración entre agentes.

Los **sistemas multiagente** pueden, a su vez, organizarse en diversas arquitecturas, cada una con ventajas y desafíos específicos:

a. Sistemas de tarea secuencial: donde los agentes trabajan en cadena, construyendo sobre los resultados del anterior.

b. Red de múltiples agentes: que permite una comunicación flexible entre especialistas, aunque con desafíos de escalabilidad.

c. Sistemas con agente supervisor (orquestador): donde un agente central asigna tareas y optimiza el flujo de trabajo.

d. Equipos jerárquicos: que establecen niveles de supervisión para garantizar control, eficiencia y toma de decisiones más sofisticadas.

La elección de la arquitectura depende del caso de uso y los requisitos técnicos, éticos y organizacionales del sistema. Comprender estas estructuras permite diseñar agentes más eficientes, adaptables y alineados con los objetivos del negocio y la regulación vigente.

07

Frameworks técnicos para el desarrollo de agentes





Frameworks técnicos para el desarrollo de agentes

El despliegue de agentes basados en IA generativa requiere de frameworks técnicos que ofrezcan la infraestructura de *software* y *hardware* necesaria para diseñar, implementar y desplegar estas soluciones de manera eficiente.

En este contexto, existen distintos marcos de trabajo disponibles, que pueden clasificarse principalmente en frameworks de código abierto y herramientas propietarias, cada uno con características y ventajas específicas. La elección entre estas opciones para construir y desplegar agentes dependerá de distintos factores, como los objetivos, las necesidades específicas del proyecto y los recursos disponibles de cada organización.

A continuación, ofrecemos un panorama de algunos de los marcos de trabajo que existen a la fecha:

CrewAI

CrewAI es un marco de trabajo para desplegar agentes de IA, que permite crear equipos en los que cada agente tiene funciones, herramientas y objetivos específicos y trabaja en conjunto para lograr tareas complejas.

CrewAI funciona en la nube, alojado localmente o de forma autónoma, por lo que puede implementarse en tu propia infraestructura con opciones alojadas localmente o aprovecharse a través del servicio en la nube preferido por el usuario, lo que otorgará un control total sobre el entorno⁶⁵.

Este marco de trabajo funciona con **4 componentes principales** con distintas características:

- 1) **Crew (tripulación) - Organización de nivel superior.** Se ocupa de gestionar equipos de agentes de IA; supervisar flujos de trabajo; garantizar la colaboración; ofrecer resultados.
- 2) **Agentes de IA - Miembros del equipo especializados.** Tienen roles específicos (investigador, escritor); utilizan herramientas designadas; pueden delegar tareas; toman decisiones autónomas.
- 3) **Proceso - Sistema de gestión de flujo de trabajo.** Define patrones de colaboración; controla las asignaciones de tareas; gestiona las interacciones; garantiza una ejecución eficiente.
- 4) **Tareas - Asignaciones individuales.** Plantean objetivos claros; utilizan herramientas específicas; contribuyen a un proceso más amplio; producen resultados viables.

65 Ampliar en <https://www.crewai.com/> [acceso el 22/1/2025].



CrewAI combina la tripulación, encargada de organizar la operación general; los agentes de IA, que trabajan en sus tareas especializadas; el proceso, que garantiza una colaboración fluida; y las tareas, que se completan para alcanzar el objetivo definido por los desarrolladores⁶⁶, para ejecutar flujos de trabajo de forma autónoma y dinámica.

En suma, es un marco de trabajo que permite crear agentes especializados con roles, experiencia y objetivos definidos, que se apoyan en herramientas flexibles y personalizadas para interactuar con servicios externos y fuentes de datos, y que pueden colaborar entre ellos de manera inteligente, cooperando, compartiendo información y coordinando tareas para lograr objetivos complejos. A ese fin, este framework permite la gestión de tareas mediante flujos de trabajo secuenciales o paralelos, con agentes que manejan automáticamente las dependencias de las tareas⁶⁷.

Además, ofrece una comunidad para el intercambio de conocimiento y consultas que actualmente cuenta con unos 2000 miembros⁶⁸.

AutoGPT

AutoGPT⁶⁹ es un marco de trabajo open source que permite crear, implementar y administrar agentes de IA para automatizar flujos de trabajo complejos en modalidad low code, mediante la utilización de componentes en modalidad “drag & drop”, como también utilizando una consola de comandos. Los agentes en AutoGPT combinan distintos componentes, con y sin IA, para crear procesos automatizados que pueden funcionar de forma independiente una vez configurados.

Actualmente, esta herramienta se ofrece de manera inmediata, a través de una implementación en entorno local⁷⁰, y también mediante acceso a una versión beta previo ingreso a una lista de espera para poder utilizarla en modalidad nube⁷¹.

En este marco de trabajo, los agentes pueden estructurarse mediante procesos lineales para tareas sencillas; flujos de trabajo complejos con ramificaciones y múltiples puntos de decisión; y sistemas de múltiples agentes que trabajan juntos para completar tareas más complejas⁷².

Los agentes en AutoGPT se crean utilizando “bloques”, que son componentes de acción individuales que conforman un flujo de trabajo. Los bloques incluyen: integraciones de servicios externos, herramientas de procesamiento de datos, conexiones de modelos de IA, scripts personalizados y elementos de lógica condicional⁷³.

La Plataforma AutoGPT está construida sobre una arquitectura de dos partes diseñada para brindar escalabilidad, flexibilidad y facilidad de uso: “Interfaz AutoGPT” y el “Servidor AutoGPT”⁷⁴.

66 Ver <https://docs.crewai.com/introduction> [acceso el 21/1/2025].

67 Ver <https://docs.crewai.com/introduction> [acceso el 21/1/2025].

68 Ver <https://community.crewai.com/about> [acceso el 21/1/2025].

69 Ver <https://agpt.co/> [acceso el 22/01/2025].

70 <https://github.com/Significant-Gravitas/AutoGPT> [acceso el 23/1/2025].

71 Ver <https://agpt.co/waitlist> [acceso el 23/1/2025].

72 Ampliar en <https://agpt.co/blog/ai-agents-explained> [acceso el 23/1/2025].

73 Ampliar en <https://agpt.co/blog/ai-agents-explained> [acceso el 23/1/2025].

74 Ampliar en “Introducing the AutoGPT Platform”, en <https://agpt.co/blog/introducing-the-autogpt-platform> [acceso el 23/1/2025].



La Interfaz AutoGPT⁷⁵ ofrece múltiples formas de interactuar con los agentes de IA y aprovecharlos:

- » **Constructor de agentes:** previsto para quienes busquen una mayor personalización, AutoGPT ofrece una interfaz intuitiva y de bajo código para que el usuario diseñe y configure sus propios agentes de IA.
- » **Gestión de flujos de trabajo:** la herramienta permite crear, modificar y optimizar flujos de trabajo de automatización con facilidad, así como crear un agente mediante la conexión de bloques⁷⁶, donde cada bloque realiza una única acción.
- » **Controles de implementación:** AutoGPT permite administrar el ciclo de vida de sus agentes, desde las pruebas hasta la producción.
- » **Agentes listos para usar:** si no querés crearlos, este marco de trabajo facilita una biblioteca de agentes pre configurados, listos para desplegar.
- » **Interacción con el agente:** la herramienta ofrece una interfaz simple para ejecutar e interactuar con los agentes.
- » **Monitoreo y análisis:** también permite realizar un seguimiento del desempeño de los agentes y obtener información para la mejora continua de los procesos de automatización.

Por su parte, el Servidor AutoGPT⁷⁷ es el motor de la plataforma. Es donde se ejecutan los agentes, que pueden activarse mediante fuentes externas y pueden funcionar de forma continua. El servidor AutoGPT contiene todos los componentes esenciales que hacen que AutoGPT funcione sin problemas:

- » **Código fuente:** es la lógica central que impulsa los agentes y procesos de automatización.
- » **Infraestructura:** sistemas robustos que garantizan un rendimiento confiable y escalable.
- » **Marketplace:** un mercado integral donde se pueden encontrar una amplia gama de agentes prediseñados.

AutoGPT se ofrece mayormente bajo licencia MIT, excepto por la carpeta *autogpt_platform* que se ofrece bajo licencia Polyform Shield⁷⁸, que únicamente restringe el uso del software en un producto o servicio que compita este producto⁷⁹.

75 Ver en <https://github.com/Significant-Gravitas/AutoGPT?tab=readme-ov-file#-autogpt-frontend> [acceso el 21/1/2025]

76 Los bloques representan acciones y son los elementos básicos de sus flujos de trabajo (ej. conexiones a servicios externos, herramientas de procesamiento de datos, modelos de IA para diversas tareas, scripts o funciones personalizados, lógica condicional y componentes de la toma de decisiones). Ver en <https://docs.agpt.co/#agents-and-workflows> [acceso el 23/1/2025]

77 Ampliar en <https://github.com/Significant-Gravitas/AutoGPT?tab=readme-ov-file#-autogpt-server> [acceso el 23/1/2025].

78 Ver <https://docs.agpt.co/#available-language-models> [acceso el 23/1/2025].

79 Ejemplos de esa vulneración serían: ofrecer un servicio que proporcione acceso directo a la plataforma AutoGPT fuera de su organización o crear un producto similar que compita con la plataforma AutoGPT colocando un envoltorio alrededor de AutoGPT. Ver en: "Introducing the AutoGPT Platform". En caso de duda sobre si tu uso vulnera esta licencia, se puede comunicar a: contact@agpt.co



Taskade

Taskade⁸⁰ es una herramienta de planificación y gestión de proyectos que ha incorporado agentes de inteligencia artificial diseñados para automatizar y optimizar diversas tareas dentro de tus proyectos.

Estos agentes, impulsados por el modelo GPT-4o, permiten a los usuarios construir, entrenar y desplegar agentes personalizados que pueden planificar, investigar y completar tareas de manera autónoma, colaborando eficazmente con los equipos humanos o entre sí.

Un aspecto destacable de Taskade es la capacidad de crear agentes personalizados y adaptados a necesidades específicas de tu organización. A través del generador de agentes incluido en la plataforma, los usuarios pueden definir el propósito y las instrucciones de cada agente, para que estos asistan en tareas como la generación de contenido, análisis de datos o gestión de proyectos.

Además, estos agentes pueden ser entrenados con conocimientos específicos mediante la carga de documentos personalizado, lo que les permite ofrecer interacciones contextualmente enriquecidas y ejecutar tareas especializadas.

La integración de múltiples agentes dentro de un mismo proyecto facilita la gestión de tareas complejas o repetitivas, ya que cada agente puede encargarse de diferentes aspectos del proyecto, trabajando en conjunto para mejorar la productividad del usuario. Esta funcionalidad posiciona a Taskade como una herramienta versátil en el ámbito de la gestión de proyectos asistida por agentes de IA.

Los agentes de inteligencia artificial en Taskade destacan por su capacidad para personalizarse, integrarse con herramientas externas y adaptarse a variados flujos de trabajo. Pueden ser personalizados mediante instrucciones definidas por los usuarios.

Esta característica permite ajustar las capacidades de los agentes a contextos específicos, ya sea mediante la definición de parámetros claros en forma de prompts o la carga de documentos relevantes para entrenarlos con información especializada. Gracias a esta capacidad, los agentes pueden abordar tareas como la redacción de documentos, el análisis de datos o la planificación y coordinación de proyectos.

Adicionalmente, los agentes no solo operan de manera autónoma, sino que también pueden colaborar entre ellos dentro de un mismo proyecto. Esta integración modular permite dividir flujos de trabajo en componentes más pequeños, asignando tareas específicas a agentes especializados que interactúan de manera fluida para alcanzar los objetivos propuestos. Además, los agentes se pueden integrar con APIs y bases de datos externas, ampliando su funcionalidad y teniendo acceso a información actualizada y precisa.

Finalmente, Taskade proporciona una interfaz completa pero compleja. No obstante a esto, una vez acostumbrado a las funcionalidades básicas, brinda también una interfaz amigable para configurar y gestionar los agentes de IA.

80 Taskade, "AI Agents Overview", en <https://www.taskade.com/ai/agents> [acceso el 17 de enero de 2025].



Planes y precios

	PLAN GRATUITO	TASKADE PRO	TASKADE TEAM
Precio	\$0	<ul style="list-style-type: none"> • \$8 - 1 usuario/mes facturado anualmente • \$10 - 1 usuario/mes facturado mensualmente 	<ul style="list-style-type: none"> • \$16 - 1 usuario/mes facturado anualmente • \$20 - 1 usuario/mes facturado mensualmente
Funciones incluidas	<ul style="list-style-type: none"> • 1 agente de IA • 1 espacio de trabajo • 5 solicitudes de IA /día • Acceso multiplataforma • Listas de tareas • Mapas mentales • Diagramas de flujo • Tableros kanban • Calendarios y más 	<ul style="list-style-type: none"> • 1 agente de IA • 1 espacio de trabajo • 5 solicitudes de IA /día • Acceso multiplataforma • Listas de tareas • Mapas mentales • Diagramas de flujo • Tableros kanban • Calendarios y más • Agentes de IA personalizados • Uso ilimitado de IA • Cargas de archivos ilimitadas • Memoria y conocimiento en vivo • Estudio y generador de proyectos de IA • Gantt, tabla y campos personalizados • Integraciones con sus herramientas favoritas 	<ul style="list-style-type: none"> • 1 agente de IA • 1 espacio de trabajo • 5 solicitudes de IA /día • Acceso multiplataforma • Listas de tareas • Mapas mentales • Diagramas de flujo • Tableros kanban • Calendarios y más • Agentes de IA personalizados • Uso ilimitado de IA • Cargas de archivos ilimitadas • Memoria y conocimiento en vivo • Estudio y generador de proyectos de IA • Gantt, tabla y campos personalizados • Integraciones con sus herramientas favoritas • Equipos multiagente • Espacios de trabajo ilimitados • Automatización de IA ilimitada • Compartir e integrar ilimitados • Historial de versiones ilimitado

n8n

n8n es una plataforma de automatización de flujos de trabajo que puede utilizarse para construir agentes basados en IA generativa.

Desde una perspectiva más técnica es una herramienta de automatización de flujos de trabajo "fair-code" que permite a los usuarios crear procesos personalizados y automatizar tareas en una variedad de aplicaciones y servicios, dado que facilita la conexión de diferentes aplicaciones para que puedan compartir datos y desencadenar acciones entre sí.



n8n es de **código abierto y autoalojable**; ofrece más de 400 integraciones con aplicaciones populares, lo que permite a los usuarios conectar y automatizar una gran variedad de servicios; brinda una **interfaz visual intuitiva** que facilita la creación de flujos de trabajo, mediante la conexión de "nodos" que representan diferentes tareas u operaciones; brinda **flexibilidad para el código**, dado que permite a los usuarios ampliar la funcionalidad con JavaScript o Python si necesitan una lógica más compleja; permite **crear conectores personalizados** a cualquier servicio con una API.

Adicionalmente, ofrece **capacidades de IA nativas**, dado que facilita la construcción de flujos de trabajo de agentes de IA basados en LangChain con datos y modelos propios; es de **alto rendimiento**, ya que puede manejar hasta 220 ejecuciones de flujo de trabajo por segundo en una sola instancia, lo que permite construir agentes de IA que pueden manejar un alto volumen de interacciones de manera eficiente; y ofrece **funciones de gestión de proyectos y roles**, que son cruciales para el desarrollo colaborativo y la gestión de agentes de IA, especialmente en un entorno de investigación o de equipos.

n8n proporciona un entorno ideal para la construcción de agentes de IA debido a su flexibilidad, sus capacidades de integración y su enfoque en la automatización. Veamos cómo los diferentes componentes de un agente de IA pueden ser implementados en n8n:

- 1) **Modelo de Lenguaje.** Permite integrar un modelo de lenguaje a través de nodos que se conectan a APIs de proveedores de IA como OpenAI o Hugging Face. Estos nodos permiten enviar prompts al modelo y recibir respuestas, que luego pueden ser utilizadas para guiar el flujo de trabajo.
- 2) **Memoria.** Se puede implementar utilizando nodos de almacenamiento de datos, como bases de datos o archivos. También ofrece nodos para trabajar con memoria a corto plazo, como variables de flujo de trabajo, que pueden almacenar información temporalmente durante la ejecución del agente.
- 3) **Herramientas.** Ofrece una amplia gama de nodos que actúan como herramientas, como nodos para enviar correos electrónicos, interactuar con APIs, leer y escribir archivos, etc. El agente puede utilizar estas herramientas para realizar acciones en función de las decisiones tomadas por el modelo de lenguaje.
- 4) **Orquestación.** La orquestación se define mediante la conexión de los diferentes nodos en un flujo de trabajo. La interfaz visual de n8n facilita la creación de flujos de trabajo complejos que definen la lógica del agente y cómo interactúa con el modelo de lenguaje, la memoria y las herramientas.
- 5) **Error Handling and Monitoring.** Incluye funciones para el manejo de errores y la monitorización de los flujos de trabajo. Los usuarios pueden configurar notificaciones para ser alertados sobre fallos en los flujos de trabajo a través de correo electrónico, Slack u otras plataformas. Estas características son esenciales para mantener la fiabilidad y la estabilidad de los agentes de IA construidos en n8n, permitiendo a los desarrolladores identificar y solucionar problemas rápidamente.

Ventajas y Desventajas de n8n

Si bien n8n ofrece un conjunto de herramientas flexible y potente para construir agentes de IA, es importante considerar sus ventajas y desventajas en relación con las necesidades específicas de cada proyecto:



VENTAJAS	DESVENTAJAS
<p>Flexibilidad: Permite la creación de agentes con lógicas complejas.</p>	<p>Curva de aprendizaje: Requiere cierto conocimiento técnico para aprovechar al máximo sus capacidades.</p>
<p>Integraciones: Se integra con una amplia gama de servicios y APIs.</p>	<p>Escalabilidad: Puede requerir una infraestructura más robusta para manejar grandes volúmenes de datos o peticiones.</p>
<p>Código abierto: Permite la personalización y el control total sobre el agente.</p>	<p>Dependencia de la comunidad: Algunas integraciones o funcionalidades pueden depender del desarrollo de la comunidad.</p>
<p>Visual: Facilita la creación y comprensión de flujos de trabajo complejos.</p>	

Concluyendo, n8n ofrece un conjunto de herramientas flexibles y potentes para construir agentes de IA que se ajusten a la definición de sistemas autónomos y adaptativos presentada en este trabajo. Su capacidad para integrar modelos de lenguaje, memoria y herramientas, junto con su interfaz visual intuitiva, lo convierte en una opción atractiva para desarrolladores que buscan crear agentes autónomos y adaptativos.

Comparado con otras herramientas, se destaca por su versatilidad, su enfoque en la automatización y su naturaleza de código abierto. Estas características lo convierten en una plataforma ideal para la investigación y el desarrollo de agentes de IA, especialmente en entornos académicos donde la flexibilidad y el control son cruciales.

Langflow y LangGraph

Langflow⁸¹ es una herramienta low code para desarrolladores que facilita la creación de agentes de IA y flujos de trabajo que pueden usar cualquier API, modelo o base de datos. Consiste en un marco visual para crear aplicaciones multiagente y RAG de código abierto, que funciona con Python y es totalmente personalizable.⁸²

La versión 1.1. de Langflow, lanzada a finales de noviembre de 2024, posee un componente de agentes, que está preparado para soportar orquestaciones de agentes complejas, con selección de múltiples modelos, memoria de chat y pasos intermedios rastreables para ejecutar acciones de razonamiento y llamados a otras herramientas.

81 Ver en <https://www.langflow.org/>, [acceso el 28/11/2024].

82 Ampliar en <https://docs.langflow.org/> [acceso el 28/11/2024].



Además, permite que los agentes puedan invocar a otros agentes como herramientas, creando un sistema de múltiples agentes que pueden interactuar y desarrollarse entre sí⁸³.

Esta herramienta permite una orquestación recursiva para una resolución de problemas de manera dinámica y de múltiples capas, donde los agentes pueden componer flujos de trabajo complejos llamándose entre sí en secuencia o en formaciones anidadas⁸⁴.

Adicionalmente, Langflow ofrece un área de pruebas o playground para que el desarrollador pueda hacer que los agentes interactúen a fin de ajustar y refinar lo que sea necesario para alcanzar los resultados deseados. Y también incluye una biblioteca de plantillas predefinidas para distintos casos de uso y metodologías (por ejemplo, asistentes, preguntas y respuestas, generación de contenido), para salir a andar de manera rápida⁸⁵.

LangGraph es un marco de trabajo de código abierto que ofrece una librería para crear flujos de trabajo con agentes y múltiples agentes. Permite definir flujos que involucran ciclos, lo que es esencial para la mayoría de las arquitecturas de agentes. Además, incluye persistencia incorporada, lo que permite funciones avanzadas de memoria y de intervención humana. Está desarrollado por LangChain Inc, pero se puede utilizar sin LangChain⁸⁶.

De este modo, entre las características principales de este marco de trabajo se destacan⁸⁷:

- » Los **ciclos y ramificaciones**, dado que permite implementar bucles y condicionales en las aplicaciones.
- » **La persistencia**, porque guarda automáticamente el estado después de cada paso del gráfico. Pausa y reanuda la ejecución del sistema en cualquier momento para respaldar la recuperación de errores, flujos de trabajo con intervención humana, viajes en el tiempo y más. Esto permite que el agente básicamente "pause" y espere la respuesta del usuario.
- » El **Human-in-the-Loop**, porque admite patrones de interacción humana de forma nativa, permitiendo interrumpir la ejecución del flujo para aprobar o editar la próxima acción planificada por el agente.
- » El **soporte de transmisión**, dado que transmite las salidas (outputs) a medida que las produce cada nodo (incluida la transmisión de tokens).
- » La **memoria a largo plazo incorporada**⁸⁸, esencialmente, un almacén de valores clave con espacios de nombres que admite la búsqueda semántica, lo que facilita que los agentes actualicen su "memoria" después de las interacciones humanas⁸⁹.
- » La **integración con LangChain**, porque se integra con LangChain y LangSmith (pero no los requiere).

Los agentes creados con LangGraph se pueden implementar mediante la plataforma LangGraph Platform, una solución comercial para implementar aplicaciones con agentes en producción, construida sobre el marco de trabajo de código abierto LangGraph⁹⁰.

83 Ampliar en <https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6> [acceso el 28/11/2024].

84 Ampliar en <https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6> [acceso el 28/11/2024].

85 Ampliar en <https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6> [acceso el 28/11/2024].

86 Ampliar en <https://langchain-ai.github.io/langgraph/> [acceso el 29/11/2024].

87 Ampliar en <https://langchain-ai.github.io/langgraph/> [acceso el 29/11/2024]. También en LangGraph, "Introducing ambient agents".

88 Ampliar en <https://langchain-ai.github.io/langgraph/concepts/memory/?ref=blog.langchain.dev#long-term-memory> [acceso el 16/1/2025].

89 LangGraph, "Introducing ambient agents".

90 Ampliar en <https://langchain-ai.github.io/langgraph/> [acceso el 29/11/2024].



En enero de 2025, LangGraph ha puesto a disposición de los usuarios su primera implementación de agentes de referencia: un asistente de correo electrónico, que se ofrece tanto como un agente alojado de uso gratuito⁹¹ y como un proyecto de código abierto⁹².

A modo de repaso: principales insights de este capítulo

El despliegue de agentes basados en IA generativa requiere de infraestructuras técnicas que permitan su implementación eficiente, escalable y adaptable a diferentes necesidades.

Existen múltiples frameworks que facilitan esta tarea, desde opciones de código abierto hasta herramientas propietarias, cada una con características específicas según el grado de personalización, automatización y complejidad del flujo de trabajo.

Los **frameworks técnicos** revisados permiten:

- a. **Desplegar agentes especializados** con roles definidos dentro de flujos de trabajo agénticos.
- b. **Coordinar múltiples agentes** a través de arquitecturas orquestadas y dinámicas.
- c. **Optimizar la automatización** con interfaces visuales, integración con APIs y herramientas avanzadas.
- d. **Facilitar la colaboración humano-IA**, incorporando la supervisión de salidas del sistema en fases críticas del proceso.

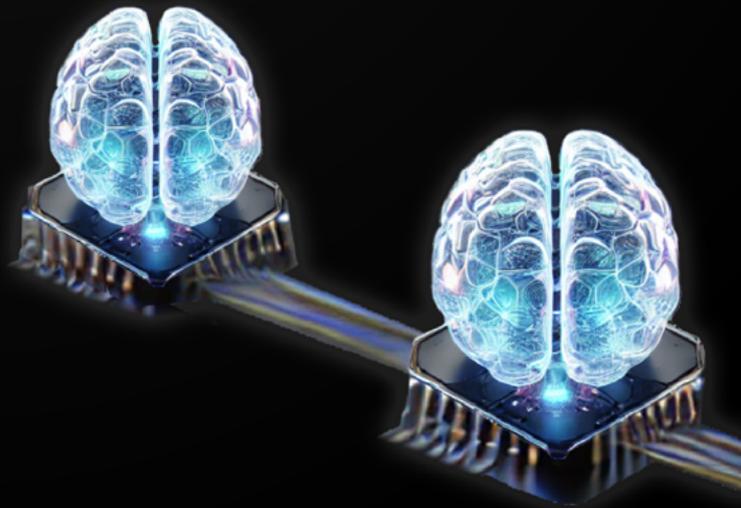
La elección del framework más adecuado dependerá del nivel de autonomía requerido para la solución buscada, la infraestructura disponible y la regulación aplicable a la organización.

91 Ver <https://www.agentinbox.ai/?ref=blog.langchain.dev> y <https://mirror-feeling-d80.notion.site/AI-Email-Assistant-How-to-hire-and-communicate-with-an-AI-Email-Assistant-17b808527b178019a42af932bb64badd>

92 Ver <https://github.com/langchain-ai/executive-ai-assistant?ref=blog.langchain.dev>

08

Conclusiones



Conclusiones

La llegada de los agentes de IA representa la nueva frontera en la automatización. El diseño de workflows agénticos permite combinar el poder de los modelos de lenguaje con funcionalidades y herramientas que dan lugar a un ecosistema cada vez más variado que permite ejecutar una diversidad de flujos de trabajo con tareas específicas y ampliar las posibilidades de aplicación en todos los sectores e industrias.

En el proceso de adopción, estos sistemas enfrentan desafíos. La naturaleza probabilística de los modelos de lenguaje que en general los sustentan, pueden llevar a resultados inesperados, errados o sesgados, lo que subraya la importancia de la supervisión humana en momentos críticos del flujo de trabajo para controlar las salidas parciales o finales del sistema. Es clave calibrar la intensidad y el modo de intervención humana en aspectos cruciales como la precisión de la información, la corrección de errores, la alineación con principios éticos y normativos, y la mitigación de sesgos.

Esto se conecta con la confianza de la sociedad con los sistemas de agentes de IA para su adopción masiva. Aparecen nuevamente la trazabilidad, explicabilidad, la precisión y la seguridad en las diversas aplicaciones y contextos en los que se despliegan. En última instancia, el éxito de los agentes basados en IA generativa dependerá de nuestra capacidad para comprender y gestionar su desarrollo de manera ética y responsable.

Es una excelente noticia reafirmar la necesidad del control humano bajo la lógica de la persona en el bucle, con un control responsable y profesional. Pero también ello nos enfrenta al desafío de reconvertir nuestros equipos de trabajo, comenzando por detectar nuevas habilidades a desarrollar.

Si la IA generativa inició un cambio de paradigma en la forma que trabajamos, los agentes de IA combinados lo acentúan a partir de modificar la naturaleza de la intervención humana: en vez de producir y realizar múltiples acciones para ejecutar tareas, progresivamente asumiremos la función de supervisar o editar resultados complejos que emanan de workflows agénticos.

Lo que no puede negarse es que la aceleración de la innovación intensificará la adopción masiva de la lógica agéntica, sobre todo de la mano de la evolución del paradigma “no code” y “open source”, la mejora diaria de los ecosistemas que contienen los grandes modelos de lenguaje y la facilidad creciente para integrarlos a sistemas y plataformas.

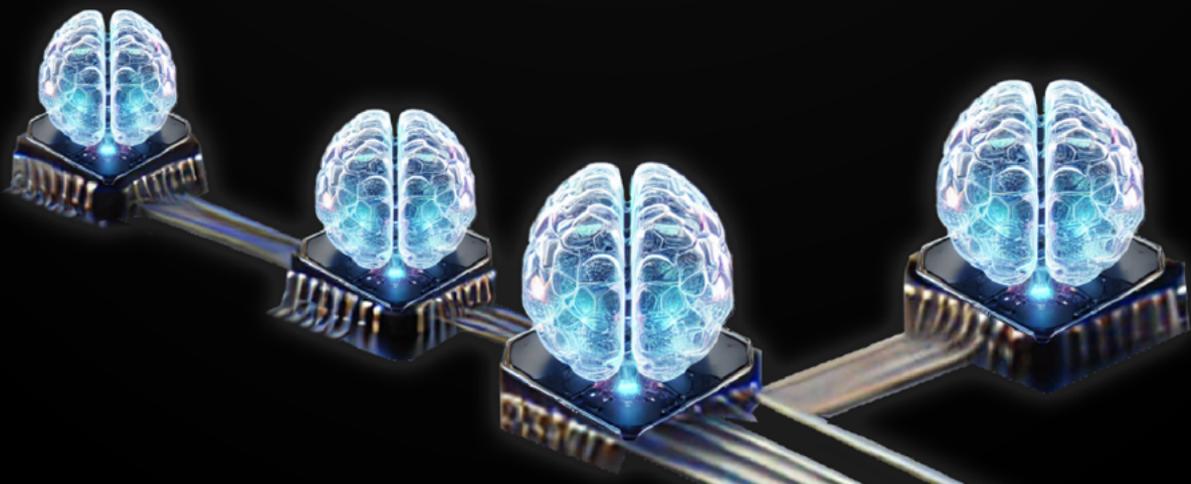
En definitiva, en este inicio de 2025, los agentes de IA se están convirtiendo en el nuevo hype del mundo IA. Pero también a ellos les llegará el fenómeno del blanco móvil.

Al igual que el rótulo “inteligencia artificial” en la década anterior, los “agentes de IA” también se volverán obsoletos, ya que se desplazará el estándar de autonomía y adaptabilidad de hoy, que será percibido como “software común” con el paso del tiempo.

09

Anexo I

**Ejemplo práctico de un sistema
multiagente para gestión de proyectos**





Ejemplo práctico de un sistema multiagente para gestión de proyectos

Tomamos como referencia al framework Taskade para evaluar la capacidad de agentes de inteligencia artificial para gestionar proyectos de manera automatizada en un entorno de desarrollo de software.

El objetivo principal fue evaluar cómo múltiples agentes especializados pueden colaborar entre sí, desde el planteamiento inicial de un proyecto hasta la concreción de un plan detallado de trabajo, incluyendo la creación de backlogs y asignación de roles. Para esta prueba, se utilizó un escenario real basado en la transcripción de una reunión inicial con un cliente.

Contexto del caso de uso

Un equipo de desarrollo realizó una reunión inicial con un cliente para discutir el diseño y despliegue de una aplicación que incluía un frontend interactivo y un modelo de machine learning para la predicción de demanda. La transcripción de esta reunión, de aproximadamente 45 minutos, dónde se discutieron los objetivos del proyecto.

La transcripción fue procesada por agentes especializados para generar de manera autónoma un plan detallado del proyecto.

Metodología

Se configuraron y entrenaron cinco agentes de IA especializados en Taskade, cada uno con instrucciones, documentos de contexto y comandos personalizados para abordar tareas específicas en la gestión del proyecto. Estas configuraciones fueron diseñadas para optimizar su desempeño en las siguientes áreas:

Asistente de Planificación de Proyectos

Entrenado para analizar documentos como transcripciones de reuniones o briefings iniciales e identificar objetivos clave, hitos, dependencias del proyecto, y generar un roadmap detallado. A este agente se le fueron provistas como contexto plantillas de planificación y se le configuraron comandos de diagramación para estructurar cronogramas de manera eficiente.

Scrum Master ágil

Configurado para organizar y priorizar tareas en un backlog basado en metodologías ágiles. Este agente incluye reglas predefinidas para la asignación de tareas según la carga de trabajo, dependencias entre actividades y fechas límite. También se le configuraron herramientas vinculadas con Google Calendar para programar reuniones y revisiones de sprints.



Gerente de Proyecto

Diseñado para ser un experto en la gestión de proyectos y funciona a modo de mediador entre los demás agentes. Se le han cargado documentos con guías de gestión de proyecto y con comandos específicos de supervisión.

Investigador

Entrenado para realizar búsquedas de información técnica sobre tecnologías, frameworks y soluciones relevantes para el proyecto. Se le brindó un archivo con el stack tecnológico de la empresa como contexto y se configuró con acceso a fuentes a sitios web confiables. Además se establecieron comandos para sintetizar hallazgos en un formato estructurado y para proponer alternativas basados en los requisitos iniciales.

Agente de Análisis FODA

Configurado para evaluar el proyecto mediante un análisis estratégico, identificando fortalezas, oportunidades, debilidades y amenazas en función del contexto del cliente y las capacidades del equipo. Este agente utiliza plantillas de análisis FODA cargadas como contexto y reglas para generar estrategias que maximicen los beneficios y mitiguen los riesgos potenciales.

Interacción entre los agentes

Los agentes cooperaron en un chat para realizar la planificación del proyecto. El flujo fue el siguiente:

Análisis inicial de la transcripción:

El Asistente de Planificación de Proyectos procesó la transcripción para identificar los objetivos principales:

- » Desarrollar una aplicación con frontend interactivo.
- » Implementar un modelo de machine learning para predicción de demanda.
- » Asegurar la integración del modelo con el backend alojado en servidores propios del cliente.
- » A partir de estos objetivos, generó un roadmap inicial con hitos y dependencias clave.



Análisis estratégico:

El Agente de Análisis FODA evaluó los factores internos y externos que podrían impactar el proyecto:

FORTALEZAS	OPORTUNIDADES	DEBILIDADES	AMENAZAS
Experiencia del equipo en las tecnologías necesarias para desarrollar el proyecto.	Aprovechar modelos preentrenados utilizando técnicas de machine learning para acelerar el desarrollo.	Tiempo limitado para la integración y pruebas	Posibles problemas de compatibilidad con los servidores del cliente

A partir de este análisis, propuso estrategias para maximizar las fortalezas y minimizar los riesgos.

Investigación preliminar:

El Investigador exploró tecnologías relevantes para el proyecto, generando los siguientes hallazgos:

- » Modelos de machine learning preentrenados disponibles que podrían servir como backbone para el transfer learning para el caso de predicción de demanda.
- » Frameworks recomendados para el frontend (React, Vue).
- » Mejores prácticas para implementar soluciones en servidores locales del cliente.
- » Estos hallazgos fueron integrados en el roadmap para guiar las decisiones técnicas del equipo.

Organización y planificación de sprints:

El Scrum Master Ágil estructuró el backlog generado por el Asistente de Planificación de Proyectos en tres sprints iniciales:



SPRINT 1	SPRINT 2	SPRINT 3
Configuración de entornos de desarrollo.	Desarrollo de la API para el modelo de machine learning.	Integración del modelo con el backend.
Configuración del repositorio en GitHub.	Diseño de la interfaz del frontend.	Pruebas funcionales del frontend.
Investigación adicional sobre compatibilidad de servidores.	Entrenamiento y validación inicial del modelo.	Resolución de bugs y ajustes finales. Además, programó reuniones semanales para revisión de avances y un demo final al cliente.

Supervisión y ajustes:

- » El Gerente de Proyecto supervisó las tareas asignadas, ajustando las prioridades según las necesidades del cliente.
- » Monitoreó los riesgos identificados por el Agente de Análisis FODA y aseguró que los recursos se distribuyeran eficientemente.

Resultados

El roadmap, backlog, determinación del stack tecnológico y la asignación de roles fueron generados luego de varias iteraciones en menos de 30 minutos, reduciendo significativamente el tiempo necesario para la planificación inicial.

Limitaciones detectadas

- » **Colaboración limitada entre agentes:** Aunque los agentes actuaron de manera eficiente de forma individual, hubo redundancias en las tareas, porque al parecer los agentes trabajaban individualmente más que en conjunto, lo que hacía que la configuración individual de cada uno de ellos parezca innecesaria.
- » **Errores en las tareas:** Algunos recursos fueron sobreasignados, el backlog no fue del todo preciso y las estimaciones de tiempo fueron deficientes, lo que requirió intervención manual para lograr la concreción del plan.

Conclusiones y recomendaciones

El uso de agentes de inteligencia artificial para gestionar proyectos demostró ser una herramienta eficaz para automatizar tareas repetitivas y reducir tiempos de planificación. Sin embargo, las pruebas revelaron áreas de mejora, como la necesidad de una mayor colaboración entre agentes y opciones más flexibles de personalización para minimizar el retoque manual de la planificación del proyecto.



Este enfoque resulta particularmente útil para proyectos pequeños o medianos con requisitos bien definidos, mientras que en proyectos más complejos puede requerirse una supervisión humana adicional para ajustar las limitaciones detectadas.

La división en diferentes agentes es recomendable para realizar diferentes procesos, pero no para distintas tareas de un mismo proceso. Dado que los agentes están potenciados por el modelo de lenguaje GPT-4o, podría ser recomendable utilizar un sólo agente configurado con todos los documentos de contexto, comandos y herramientas necesarios para la realización de todas las tareas de un proceso. Esto hace que se pueda optar por un plan más barato, casi sin perder capacidades.

Resultados clave:

- » **Automatización eficaz:** Los agentes lograron transformar insumos como transcripciones de reuniones en planes de trabajo detallados.
- » **Interacción limitada entre agentes:** Se identificaron redundancias y falta de comunicación entre los agentes.
- » **Potencial de mejora:** La integración efectiva y la personalización serán clave para futuros avances en este tipo de herramientas.

Recomendaciones:

- » Crear un agente por proceso munido de todas las herramientas necesarias para llevar a cabo un proyecto.
- » Configurar el agente lo más posible de manera que la revisión humana sea lo más pequeña posible.

En síntesis, el uso de agentes especializados en la planificación de proyectos tiene un enorme potencial, pero requiere optimización para enfrentar los retos de proyectos más complejos y colaborativos.

.UBAderecho



IALAB



BANCO DE DESARROLLO
DE AMÉRICA LATINA
Y EL CARIBE

Con el apoyo de:

ubatec^{SA}

puzzle.



COGNITIVE
BUSINESS INSIGHTS THRU DATA



AI academy
by doinGlobal

Corpora_



PROCURACIÓN GENERAL
DE LA CIUDAD

.UBAfiuba 
FACULTAD DE INGENIERÍA



IngenIA UBA
Grupo de Ingeniería en
Inteligencia Artificial



MultIALAB
Laboratorio Multidisciplinario
de Inteligencia Artificial



LIDeSIA
FOEFyUN

LA LEY



Thomson
Reuters™

