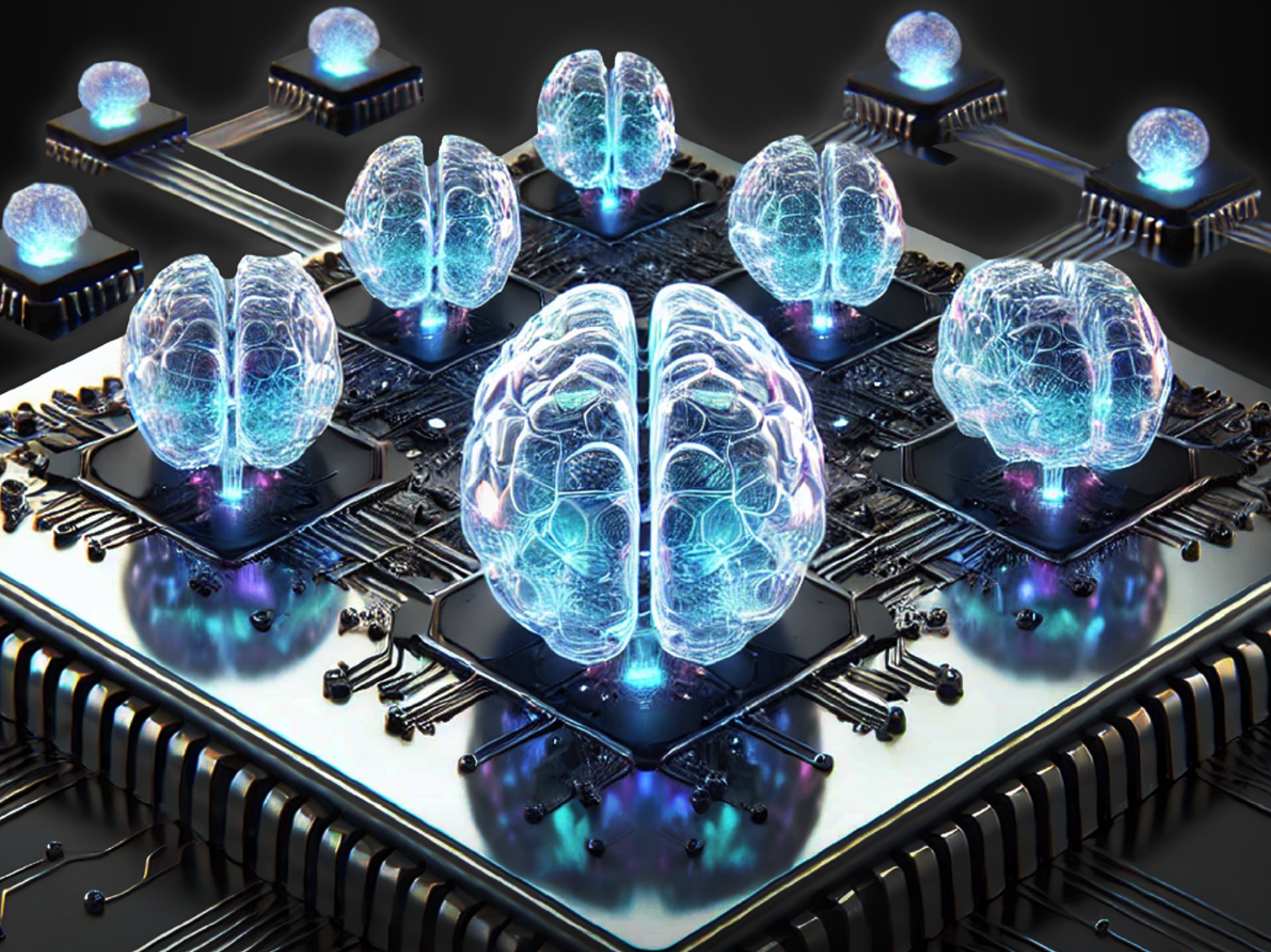
.UBAderecho IALAB CAF BANCO DE DESARROLLO DE AMÉRICA LATINA Y EL CARIBE

Artificial intelligence agents and agentic workflows

The new frontier of automation



Supported by

ubatec 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 FCEFyN LA LEY Thomson Reuters™

Artificial intelligence agents and agentic workflows: The new frontier of automation

Practical guide to understanding what they are, how they work and when to use them.

Corvalán, Juan Gustavo
Artificial Intelligence Agents and Agentic Workflows: The New Frontier of Automation – A Practical Guide to Understanding What They Are, How They Work, and When to Use Them / Juan Gustavo Corvalán; Enzo María Le Fevre Cervini, Mariana Sánchez Caparrós. – 1st ed. – Autonomous City of Buenos Aires: La Ley, 2025.

Digital book, PDF
Digital File: Download and Online Access 978-987-03-4946-4

1. Law. I. Sánchez Caparrós, Mariana II. Le Fevre, Enzo María III. Title CDD 342.066

This and other digital publications are available at https://ialab.com.ar/

Published in 2025 by the Artificial Intelligence Laboratory of the UBA Faculty of Law.
Av. Figueroa Alcorta 2263, C.A.B.A., Argentina.

.UBAderecho IALAB CAF BANCO DE DESARROLLO DE AMÉRICA LATINA Y EL CARIBE

Direction and co-authorship

Juan G. **Corvalán,** Enzo Maria **Le Fevre Cervini** and Mariana **Sánchez Caparrós**

Researchers and collaborators

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

Graphic designers

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

CONTENT

01

Introduction

The path to agentic AI and agentic workflow

The concept of Agency in AI

Agency in artificial intelligence refers to the capacity of AI systems to operate with autonomy, make decisions, and execute tasks with minimal human intervention. Unlike traditional automation, which follows predefined rules and workflows, agentic AI integrates learning, adaptation, and decision-making capabilities, enabling systems to proactively respond to changing conditions.

A fundamental distinction exists between reactive AI and proactive AI. Reactive AI systems, such as early chatbots and rule-based assistants, rely on predefined responses to user inputs. They do not retain memory or adapt beyond their programmed capabilities. In contrast, proactive AI agents take initiative, analyze contexts, and perform tasks independently, often without waiting for explicit instructions. This shift moves AI from passive assistance to active decision-making, where agents anticipate user needs, automate complex workflows, and continuously improve through iterative learning.

The implications of this transformation are profound. AI is no longer just a tool executing commands but an active participant in workflows, capable of coordinating multiple actions, synthesizing large volumes of information, and dynamically optimizing decisions. As AI becomes increasingly autonomous, the challenge lies in balancing its capabilities with human oversight to ensure reliability, transparency, and ethical use.

The beginning: conversational agents

Conversational dialogue between humans and machines began in the 1960s with Eliza[1]. Fifty years later, this technology became popular with conversational agents such as Siri, Alexa or Cortana, introduced to the wider public from 2011 onwards. At the time, these agents represented the cutting edge of artificial intelligence powered by deep learning, designed to interpret and respond to user instructions. However, their capabilities were constrained by the limitations of artificial neural networks at the time. In summary:

- They were neither multitaskers nor multipurpose, as they were trained to handle specific conversations.
- The interactions relied on preconfigured intents, which had to be defined by humans in advance.

Despite these limitations, these agents, also known as conversational bots, marked t**he first massive step toward human-machine interaction using natural language**, laying the foundation for more advanced AI-driven communication systems.

1 In 1966, computer scientist Joseph Weizenbaum created a program that searched for keywords in conversations with human typists; if the human used one of those words, the program would use it in its response. If not, it would offer a generic response. It was meant to mimic a psychotherapist. Later, in 1972, a Stanford scientist, Kenneth Colby, created another bot named Parry that attempted to model the behavior of a paranoid schizophrenic. Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. Communications of the ACM, 9 (1), 36–45, available at: https://dl.acm.org/doi/10.1145/365153.365168 (accessed January 29, 2025] . See also https://www.theatlantic.com/technology/archive/2014/06/when-parry-met-eliza-a-ridiculous-chatbot-conversation-from-1972/372428/

The second stage: multitasking and multipurpose conversational agents

The next leap forward came with the rise of generative AI[2], particularly with the launch of ChatGPT in late 2022. The introduction of transformer architectures[3] and the revolution in word prediction, powered by vast amounts of training data, significantly enhanced AI's ability to simulate understanding. This advancement enabled generative AI models to develop increasingly sophisticated capabilities, including the creation of highly realistic and consistent synthetic content across multiple domains and applications.

This second leap introduced AI models with two key features

- Multitasking and multipurpose capabilities. AI systems could now respond to diverse and complex topics with remarkable accuracy and versatility.
- Greater autonomy and direct accessibility to the public. These improvements democratized access to technology and allowed people to use generative AI agents to solve a wide range of problems, without the need for specialized technical training, for free or at low cost.

This breakthrough not only marked the mass adoption of generative AI but also became a defining milestone in its evolution.

The third leap: multimodality

In 2024 generative AI rapidly advanced towards multimodality, that is, the ability to process multiple input and output formats (text, images, voice, etc.).

Before this leap, generative AI had to rely on external tools, such as plugins to "look at" images or other specialized systems. With multimodality, these capabilities are integrated directly into generative AI models, which in this evolved version allow for richer and more complete two-way interactions. For example, an AI can now process a visual document, interpret it, and respond in natural language, all without relying on external systems.

2 Generative artificial intelligence is a subfield of AI that is used to create new and original content, or to modify or improve existing content, such as text, images, music and videos. GenAI relies on machine learning to identify patterns in data and then uses those patterns to generate new content. See OECD Artificial Intelligence Papers, Initial Policy Considerations for Generative Artificial Intelligence, September 2023, available at: https://www.oecd-ilibrary.org/deliver/fae2d1e6-en.pdf?itemId=/content/paper/fae2d1e6-en&mimeType=pdf (accessed 29/02/2025]

3 Consulted on 29/02/2025

The step towards advanced "reasoning"

The fourth stage in the evolution of agentic AI brought models with more complex "reasoning" capabilities. As a counterpoint, earlier versions, such as ChatGPT in 2022 (3.5], struggled with processing complex tasks independently. Achieving accurate results often required multiple human interventions, guiding the model step by step through sophisticated prompting to steer it toward an appropriate solution.

With advancements in "reasoning" capabilities models such as OpenAI's o1 and DeepSeek's R1 became more adept at handling complex tasks with reduced user intervention. This evolution significantly enhanced their effectiveness, solidifying their role as more robust and capable AI assistants.

The move towards autonomy and adaptability

The next major advancements in generative AI began in February 2024, marking the transition from base agents to more autonomous and multi-purpose systems capable of handling multiple tasks. This evolution was driven by the integration , such as memory, internet access, and more recently, in January 2025, the ability to schedule tasks[4] further expanding AI's capacity to operate independently.

The top: specific agents with high autonomy

In Iron Man, the JARVIS (Just A Rather Very Intelligent System) AI serves as the ideal example of a highly autonomous agent. JARVIS not only assists Tony Stark with daily tasks, but also acts as a specialized agent, managing the Iron Man suit, maintaining the workshop, and executing complex workflows such as designing new technologies or optimizing the suit in real time during combat.

The latest stage of agentic AI, at least for the now, introduces specific agents built on generative AI base models, designed for highly specialized functions. According to Deloitte, in 2025, 25% of companies that use generic AI will launch pilots or proofs of concept of AI agents based on large language models, a figure expected to rise to 50% in 2027[5].

In AI systems, Google has recently introduced Gemini 1.5 and its Advanced Research model, which represents a first version of Iron Man's Jarvis model for research and paper writing. This agent can:

4 See "Scheduled tasks in," at https://help.openai.com/en/articles/10291617-scheduled-tasks-in-chatgpt [Accessed 1/29/2025].
5 See Jeff Loucks, Gillian Crossan, Baris Sarer, China Widener , "Autonomous generative AI agents: Under development", November 2024, at https://www2.deloitte.com/us/en/insights/industry/technology/technology-media -and-telecom-predictions.html#autonomous-generative-ai [Accessed 11/25/2024].

- Generate comprehensive reports with proper citations.
- Automate complex workflows, such as drafting work plans, without direct human intervention but yet based on human input.

Here, we see how greater AI autonomy is shifting the dynamic of human intervention. The role of the user is evolving from “producer” to “auditor”, and ultimately to “qualified editor or validator” of decisions taken by the AI agent.

At UBA IALAB we tested the agent that manages personal computers (computer use)[6] launched by Anthropic at the end of 2024. This agent, which runs locally, controls your computer in the same way you would: “looking at the screen”, moving the cursor, clicking on buttons to open applications and writing text. In this report, we appreciate its possibilities and the limitations that, for the moment, it poses in terms of efficiency and security[7].

In parallel, just like the Operator in The Matrix, who provides guidance and access to crucial information from outside the simulation, OpenAI recently introduced its own 'Operator'[8]—a preliminary research version of an AI agent capable of accessing the web to perform tasks for the user. For the moment, it is only available for Pro users in the US.

These are the first examples of highly autonomous agentic AI systems. Solutions that organize, plan, direct and execute complex processes from start to finish, while substantially reducing or eliminating the need for human intervention in intermediate steps of that proces.

To sum up...

As we often say in the face of this tsunami of innovation, the future is the past, and reality makes fiction obsolete. The evolution from early conversational agents to fully agentic AI reflects a steady progression towards increasingly sophisticated, autonomous and adaptable systems.

Each stage has marked a significant leap forward, ranging from understanding basic instructions to integrating multimodal capabilities, “reasoning” in a complex, autonomous way, and giving rise to the possibility of automating complete workflows.

As we will explore throughout this document, AI agents based on generative models now combine language processing with specialized "tools" that allow them to remember previous interactions and to access external resources. This capability enables them to perceive, interpret and interact with their environment, laying the groundwork for integration with more complex agentic systems. These advanced systems can break down processes into concrete steps and execute entire workflows autonomously and

6 https://www.anthropic.com/news/3-5-models-and-computer-use (accessed 29/01/2025]
7 Juan G. Corvalán and Mariana Sánchez Caparrós, “The AI agent that controls your computer: we tested Anthropic's Computer Use”, IALAB, December 2024 [in https://ialab.com.ar/webia/wp-content/uploads /2024/12/The-AI-agent-that-drives-your-computer-1.pdf , accessed 29/1/2025].
8 OpenAI, “Introducing Operator,” January 27, 2025 [https://openai.com/index/introducing-operator/, accessed January 29, 2025].

dynamically[9] bringing us ever closer to a new paradigm of AI-driven decision-making and automation.

The biggest efforts to move in this direction are being made by startups, more established technology companies, and cloud providers that are developing their own agent-based AI offerings[10].

For example, the Argentine unicorn company Globant uses AI agents to transition from one programming language to another, enabling to reduce times without losing security and quality standards[11].

As an industry pioneer, Globant offers human-augmented and supervised AI agents that integrate into the software development lifecycle , with the goal of enhancing your development capabilities. The company's initial set of agents includes: AI people for Product Definition; for Backend Prototyping; for Application Design; for Testing and for Code Correction[12].

9 In the same sense, see Jeff Loucks, Gillian Crossan, Baris Sarer, China Widener , "Autonomous generative AI agents: Under development", November 2024, at https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions.html#autonomous-generative-ai [Accessed 11/25/2024]

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

11 See https://now.globant.com/en/code-transition-with-ia-agents/ [Accessed 1/29/2025].

12 See https://www.globant.com/es/news/globant-presenta-ai-agents [Accessed 29/1/2025].

Evolution of Agentic AI

01

Initial Conversational Agents

The first AI bots interacted with humans using natural language, but they had limited capabilities.

02

Generative AI

Generative AI emerged, improving the versatility and accessibility of AI.

03

Multimodal Advancement

AI has become multimodal, integrating multiple input and output formats.

04

Improved Reasoning

AI has improved its reasoning ability, handling complex tasks with less human intervention.

05

Greater Autonomy

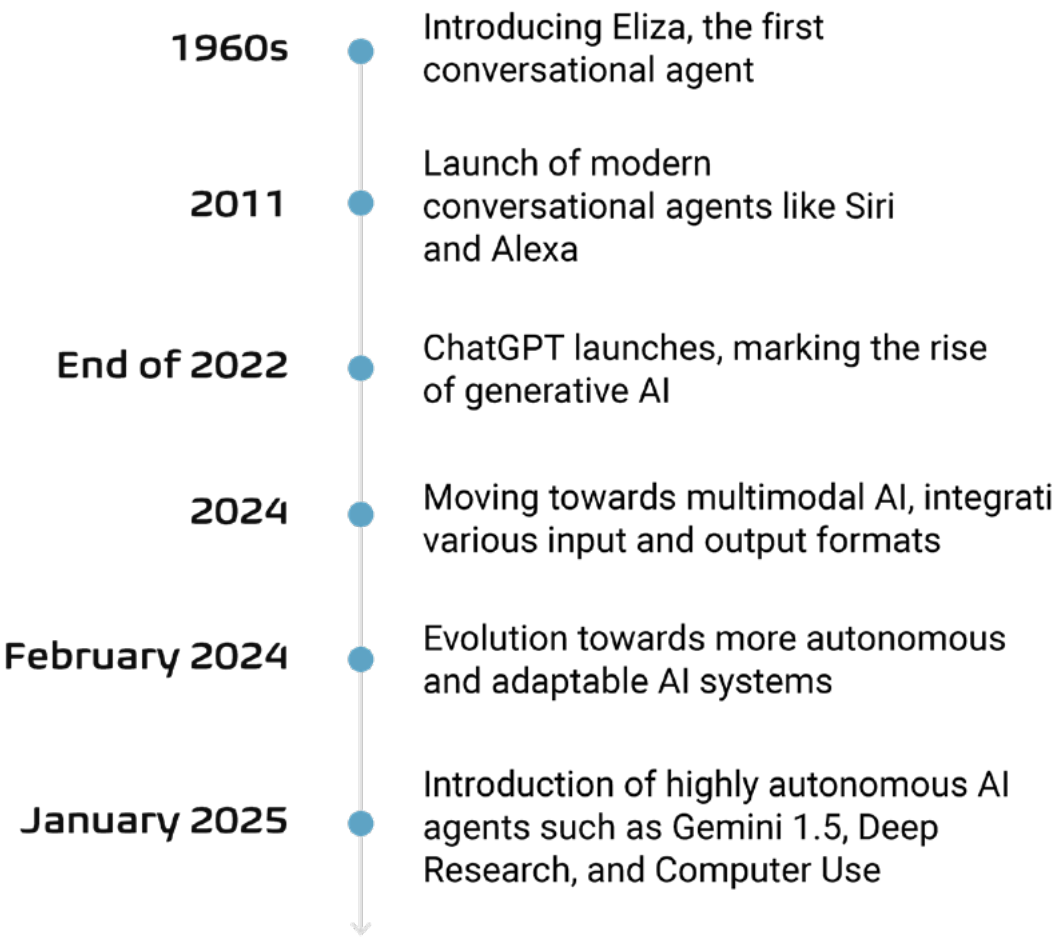
AI became more autonomous and adaptable, incorporating tools such as memory and internet access.

06

Specialized Agents

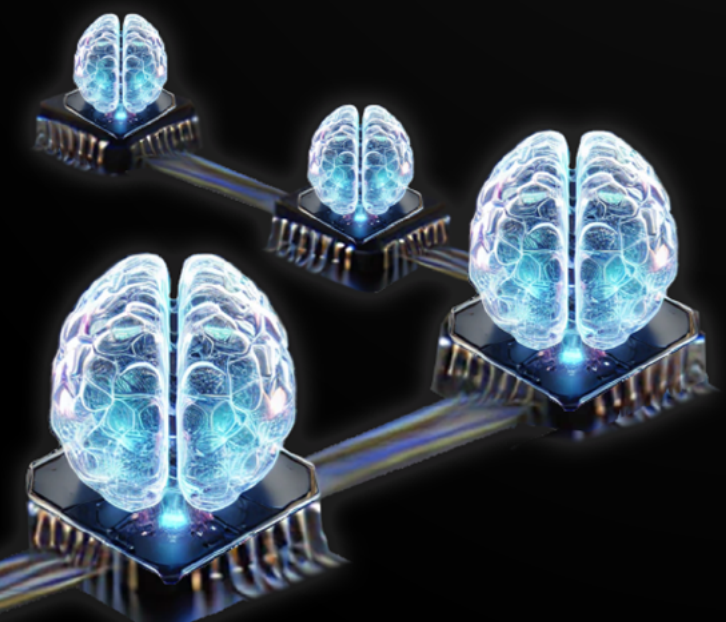
Highly specialized agents were developed, demonstrating high autonomy in specific tasks.

The evolution of conversational AI towards autonomous agents



02

From is that AI?
to when is there an AI agent?



From: Is that AI? To: When is there an AI agent?

"That's not AI!" or "That doesn't have AI!". These assertions were a classic in industry and academia until the arrival of the great language models. In the second decade of this century, the offer of services and solutions sold under the premise that they had artificial intelligence grew. Then, the discussions arose around the fact that an expert system was not AI, although in any AI book one could find them as part of the ecosystem.

The question of what qualifies as AI—and what doesn't—became a recurring topic of discussion. Juan G. Corvalán recalls conversations with World Bank representatives on this very issue, highlighting its impact not only on project funding but also on meeting the industry's persistent demand, even amid the hype: that a solution must include something of AI.

The rise of AI agents has brought this issue back into focus. Discussions about an organizational revolution are becoming more frequent, emphasizing how disruption will be driven by artificial intelligence agents.

Yet, once again, we face a fundamental question: What exactly is an AI agent? How do we distinguish between the presence of an agent and its absence? This challenge mirrors the ongoing debate between traditional software and software with AI components—only now, it is amplified by new solutions.

As always, in semantic debates like this, apples are mixed with bananas. Defining a concept requires adopting evaluative starting points and positioning oneself within the broader discourse of how terms are used in the technology industry—a challenge that exists across all disciplines.

Before diving into the different conceptualizations surrounding "AI agents," it is essential to establish some groundwork. The rapid evolution of AI, especially since the explosion of generative models, has brought us to a supersonic pace of change. Understanding where we stand today is crucial before exploring the nuances of these definitions.

Autonomy and adaptability

From traditional software based on *if-then* logic to the cutting-edge generative AI models emerging almost daily, there exists a vast spectrum of systems, models and architectures.

Before the rise of generative AI, much of the debate around autonomy centered on self-driving cars—specifically, determining when a car could truly be considered autonomous. The initial binary approach (either it is or it isn't) was quickly abandoned in favor of a graded scale. In essence, defining autonomy became a matter of positioning a system along a continuum, where specific characteristics and functionalities determine its ability to execute actions or make decisions without human intervention.

In the first edition of the Treaty on Artificial Intelligence and Law, published in 2021, we applied the logic of autonomy scales from self-driving vehicles to systems that process natural language. Similarly, AI agents can be categorized based on their varying degrees of autonomy and adaptability.[13]

A compelling example of this comes from fiction: *Westworld* offers a rich portrayal of artificial intelligences operating at different levels of autonomy. At first, the hosts are programmed to perform specific tasks within predefined limitations, strictly adhering to the narratives set by the park's designers. Over time, however, some hosts begin to exhibit emergent capabilities, such as simulating reasoning and improvisation. However, other hosts, such as Dolores and Maeve, demonstrate an even higher degree of autonomy, reaching a form of "self-awareness" that ultimately drives them to carry out a revolution against their human creators.

We believe that the paradigm of agents based on generative AI can also be approached from a scale-based or layered logic. Autonomy and adaptability for the execution of tasks and workflows will be the essential criteria to determine the differences between multiple variants of agentic systems that are being developed and presented under the title of "AI agents". We will revisit this point later.

Generative AI and large language models: the transition to "base agents"

Unlike traditional AI, based on machine learning , which learns from data to generate results such as predictions, recommendations or automated decisions, generative AI "... also learns from data, but it is a more advanced level of deep learning. It finds patterns in the information and mixes them to create new content..."[14]. This allows them to provide us with answers with novel content and a writing style that adapts to the need expressed by the user.

Since its launch, in these two dizzying years, generative artificial intelligence has experienced significant advances. From a multipurpose and multitasking conversational agent that could only process a prompt or input in text format and return an output in the same format, we have arrived at multimodal models that can also interpret images, video and audio, breaking down the barriers between different formats of information to be processed and generated.

The multimodal capability of large language models, as well as access to them through their APIs, gave rise to assistants applicable to new use cases and expanded their scope in terms of usability.

13 Expand in Corvalán Juan G., Sa Zeichen Gustavo and Albertsen Lihué María, "Automated Administrative Activity. Artificial intelligence, regulatory power, algorithmic delegation, automated administrative act and reserve of humanity".
14 See it https://ialab.com.ar/webia/wp-content/uploads/2024/08/IA-Generativa-y-la-Gestion-del-Talento.pdf, p. 5.

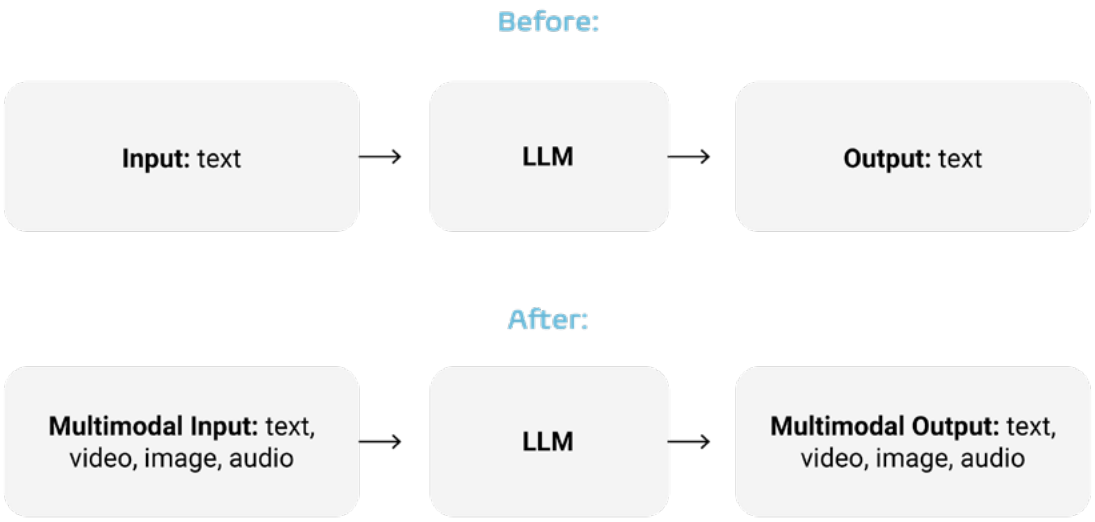


Figure 1. Evolution from unimodality to multimodality of traditional generative AI applications.

In the report *"A spreading tsunami, two years of ChatGPT and generative AI"*[15], it is highlighted how multimodality was the great protagonist of 2024. In fact, after the release of unimodal models such as GPT-4 (when used only for text), Whisper, DALL·E 2, Imagen, Stable Diffusion, Bard (in its initial text-based format), almost all tools began to be offered under the logic of multimodality. As an example:

15 Corvalán, Juan G. and Carro, Maria Victoria, "An expanding tsunami, two years of ChatGPT and generative AI", December 2024, in https://ialab.com.ar/webia/wp-content/uploads/2024/12/Dos-anos-de-ChatGPT-e-IA-Generativa.pdf [Accessed 1/22/2025].

COMPANY	MODEL	CAPABILITIES	LAUNCH
OpenAI	GPT-4	Processes both text and images. Enables tasks that combine visual and textual interpretation, such as generating image descriptions.	2023
	CLIP	Designed to combine text and images, allowing image search based on text.	2021
	Sora[16]	A multimodal model for generating audiovisual content (images, text, and video). Enables the creation of realistic videos based on text descriptions.	2024
Google	Gemini	Works with text, images, and audio. Designed to integrate with Google products.	2023
	PaLM 2	Processes both text and images. Allows for the generation of image descriptions and visual interpretation tasks.	2023
	Veo 2	Enables AI-generated videos using both images and text. Direct competitor to Sora.	2024
DeepMind	Flamingo	A model capable of handling text and images, known for generating image descriptions and visual understanding tasks.	2022
Meta	LLaMA	Initially a text-only model, but recent versions include multimodal capabilities for handling text and images.	2024-2025
Mistral AI	Mistral	Initially focused on text processing, now being developed for image interpretation.	2024-2025
DeepSeek[17]	DeepSeek V32	A mixture of experts (MoE) transformer model, comparable to GPT-4o. Not currently multimodal.	2025
	DeepSeek R1	Competes with OpenAI's o1 and o3 models. Its key feature is transparency in reasoning, allowing users to see the reasoning path behind responses.	2025

Generative AI has evolved from a “conversational chat” that only processed multipurpose and multitasking natural language to becoming a true system that has embedded various AI models, multiple functions and tools to revolutionize productivity. For this reason, they currently function as base agents that present increasing but limited autonomy, which allow other specific agents to be enhanced.

16 See in https://openai.com/sora/
17 See inhttps://www.deepseek.com/.

In short, after two years and two months, the systems containing the generative AI models are in a dizzying race to "liquefy" all the software as if they were Mr. Smith and his agents in the movie The Matrix.

The transformation of productivity and efficiency: from the automation of repetitive and routine tasks to the automation of complex and semi-complex tasks

The paradigm shift brought about by generative AI through large language models impacted multiple sectors and verticals across all industries[18]. Its adoption changed a prevailing narrative until then: that AI would essentially be applied to routine, mechanical and repetitive tasks.

However, when we began our research into the impact of Generative AI on work in 2023[19] and started implementing it in various sectors such as justice, public administration, law firms, corporate legal departments, translation, education, research, the commercial sector and marketing areas, we detected that this technology actually had more impact on high and medium complexity tasks to transform and optimize processes.

At the beginning of 2024, we published the results of the research. We considered quantitative indicators, such as task execution times using generative AI, and qualitative ones, such as the degree of automation, adaptability (still incipient) of the tools and the possibility of obtaining higher quality results. The results showed how generative AI complements and transforms tasks, improving efficiency and redefining the interaction between humans and machines in the workplace[20]. Already at that time, we discovered that language models, by generating new content, allowed for an increase in the quality of the results, as long as humans made efficient and supervised use of these technologies.

As anticipated, generative AI proved to be more efficient in tasks of medium complexity (81%) and less efficient in tasks of low complexity (52%). And for tasks requiring high human judgment, efficiency reached 73%, which shows its ability to assist in the execution of complex tasks, notwithstanding that 59.03% of tasks benefit from collaboration between humans and AI, and human intervention remains essential for activities requiring critical thinking and creativity.

Multitasking base agents: a new paradigm of "human-machine" collaboration

Generative AI tools available today, such as ChatGPT, Gemini, Copilot, and more recently DeepSeek-V3 and R1, have revolutionized productivity due to the enormous impact they have on work automation.

18 McKinsey & Company. (2023). *The State of AI in 2023: The Key Year for Generative AI*. Available at: https://www.mckinsey.com/featured-insights/featured-insights/the-state-of-ai-2023-the-key-year-for-generative-ai/ (accessed January 29, 2025]

19 https://ialab.com.ar/webia/wp-content/uploads/2024/05/Evaluacion-del-impacto-de-la-AI-generativa-en-el-trabajo.pdf (accessed January 29, 2025]

20 Corvalán, Juan G., Evaluation of the impact of generative artificial intelligence at work (1st ed.), La Ley, March 2024, at https://ialab.com.ar/webia/wp-content/uploads/2024/07/iagenerativa-marzo2024-final-1.pdf [Accessed 1/22/2025].

Their evolution shows that they can increasingly solve a greater number of tasks autonomously in specific domains with particular contexts. For efficient use, good prompting strategies are required[21], and in certain cases, techniques are used to increase their knowledge base with more specific data (RAG), or to touch the model parameters (fine tuning), among other options that have been developed in the last two years.

For example, at UBA IALAB we have adopted strategies to configure GPTs[22] and combine them with optimized prompts, and we have also applied various strategies to automate tasks at scale using base agents. Even so, the autonomy and adaptability of these solutions in different scenarios remains low, given that when faced with new tasks, the solution usually requires a lot of adaptation.

On the other hand, although base agents are improving their autonomy, precision and adaptability, they are not yet designed or prepared to complete entire workflows adaptively and without expert human intervention. Nor are they prepared to make decisions that involve planning, remembering and "reasoning" autonomously with regard to the tasks that make up that workflow and that need to be completed to execute it from "end to end." Finally, they cannot reuse (when required) previous learning in new interactions.

Still, given the exponential evolution that generative AI has been showing, we believe that two complementary trends will emerge in the coming months:

- **More tools and features.** Base agents, i.e. models such as o1, 4o, DeepSeek-V3 and R1, or Gemini 1.5, among others, will increase tools and features to help make up for this lack of autonomy and adaptability. An example of this initial path can be seen, for example, in the tools for managing projects available in ChatGPT; the platform's assistance in generating a personalized GPT; memory (limited for the moment); internet access; and the scheduling of future tasks.
- **The arrival of agentic workflows.** Base agents, with their tools and functions, will be complemented by specialized agents, through agentic workflows and multi-agent systems.

In conclusion: the base agents par excellence are the large language models that are present in web platforms and mobile applications, such as ChatGPT, Gemini, Claude or, more recently, Deepseek . The most primitive interaction with these base agents consists of the user entering[23] an instruction to execute a task through a prompt and the generative AI tool generating a response quickly in the requested format.

21 See more at: Prompting Techniques | Prompt Engineering Guide . See more at Corvalán Juan G. Director, Implementing generative artificial intelligence in law firms and legal departments, Thomson Reuters, La Ley, March 2024, available at: Implementing-generative-artificial-intelligence-in-law-firms-and-legal-areas.pdf

22 ChatGPT GPTs are customized versions of the ChatGPT model, which users can configure to suit specific tasks or needs. See https://openai.com/index/introducing-gpts/ [Accessed 29/01/2025].

23 As we will see later, this can vary in the context of automated flows supported by generative AI, where the flow is triggered by an "event" predetermined by the designer (for example, a file being uploaded to a cloud space or an email being entered).

Figure 2. Traditional operation of generative AI applications.

Throughout 2024, this paradigm became more sophisticated and evolved towards multi-task and multi-modal base agents embedded in applications, which have multiple additional functions and tools integrated such as: internet access, task scheduling[24], (limited) memory[25] and the possibility of customization, among others.

Year 2025: Generative AI-powered base agents are the main engine of agentic workflows

AI agents have a dual focus, as they can be leveraged and adapted for different tasks, depending on the users' intentions. Let's take a look.

The advances made in generative AI over the past two years demonstrate its enormous capacity to tackle the execution of various tasks through the format of a human question or request (prompt) and a machine response.

However, the generative AI-based agents that are being designed will be able to acquire greater autonomy and assume defined roles. The "perception" of the environment and the learning of the particular context will lead them to act dynamically and autonomously, to continuously adapt and evolve in the execution of complete workflows, which go beyond the fact of completing particular tasks. This is one of the keys to the change that is initiated by the base agents. Let's go into a little more detail.

Typical generative AI tools (e.g. Copilot, ChatGPT-4o, Claude, and Gemini) can be used to automate tasks, but they are not useful on their own for planning and orchestrating complete, end-to-end workflows. They also do not retain a relevant memory[26] of executed tasks that can be reused later. Moreover, in certain cases, they possess knowledge limited to a specific date, fine-tuning is costly in terms of time and money, and they cannot efficiently execute tasks that self-evaluate their responses, being limited to probabilistic reasoning based on their training data[27].

24 See https://help.openai.com/en/articles/10291617-scheduled-tasks-in-chatgpt [Accessed 29/1/2025].

25 ChatGPT can see past conversations the user has had with the tool and reference them in its responses to provide better, more contextual, and more useful answers to your questions. See more at: OpenAI, "How does memory use past conversations," updated December 2024, at https://help.openai.com/en/articles/10303002-how-does-memory-use-past-conversations.

26 Memory allows the model to remember information between sessions and adapt its responses considering previous interactions. Unlike the temporal context of a conversation, where the model only retains information within the current session, persistent memory stores key details such as user preferences, recurring topics, and response style to improve personalization. In ChatGPT, memory is available to ChatGPT Free, Plus, Team, and Enterprise users from September 2025. This memory is not static, it is updated, corrected, or deleted at the user's request, which can be managed by the user from the settings.

27 Deloitte AI Institute, "Prompting for action | How AI agents are reshaping the future of work", November 2024, p. 6, at https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf [Accessed 24/11/2024].

Still, these base agents with their tools are a first draft of the more sophisticated agents that are coming to redefine the "human-machine" collaboration scheme. In short, if a base agent like ChatGPT's GPT-4o model led us to be experts in giving instructions and editing responses, agentic logic will lead us to be much more sophisticated editors to review the sequence of decisions deployed by agents in multiple tasks within a workflow, with little or no human intervention (but always with supervision, as we will also see).

So what are AI agents, and what is the role of generative AI in their development?

There are different opinions regarding the criteria that allow us to define the existence of an AI agent. Let's see.

For Google, the mix of reasoning, logic, and access to external information, combined with a generative AI model, is what invokes the concept of an agent or a program that extends beyond the independent capabilities of a generative AI model in its traditional design[28].

Deloitte[29], for its part, defines them based on a functionalist logic. They would be agents based on generative AI that:

- **They allow you to create and execute multi-step work plans or tasks** to achieve a goal and adjust actions based on real-time feedback.
- **They can rely on short- and long-term memory** to learn from the user's previous interactions and provide personalized responses, and memory can be shared between multiple agents in the same system.
- **They can turn to tools to augment the original capabilities of the large language model** to perform tasks, perceive and interact with the environment. For example, through data extractors, image pickers, search APIs and fine-tuning.
- **They can leverage the specific capabilities, knowledge, and memory** específicos de distintos modelos de lenguaje para validar y controlar sus propiosof different language models to validate and control their own and other agents' outcomes in the context of a system.

For our part, we believe that agent logic begins with **base agents** that are driven by generative AI, so that they take on specific agent roles or that they articulate with specialized agents that could use other types of AI.

Agents **(agents in the strict sense)** are designed to execute actions and complete end-to-end workflows autonomously and adaptively.

Specific agents can operate in a single-agent or multi-agent system. They can rely on and interact with one or more generative AI language models (base agents). They can also be structured in different ways to suit specific uses, for example under a hierarchical agentic system logic or a sequential system logic.

28 Wiesinger, Julia , Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, September 2024, [online https://media.licdn.com/dms/document/media/v2/D561FAQH8tt1cvunj0w/feedshare-document-pdf-analyzed/B56ZQq.TtsG8AY-/0/1735887787265?e=1736985600&v=beta&t=pLuArcKyUcxE9B1Her1QWfMHF_UxZL9Q-Y0JTDuSn38, accessed 7/1/2025].

29 Deloitte AI Institute, "Prompting for action | How AI agents are reshaping the future of work", November 2024, p. 6, at https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf [Accessed 24/11/2024].

They are a species of AI agents. The key is that they operate with a high degree of autonomy, making decisions and performing actions independently of human intervention[30].

An example can help understand agents from this perspective. Imagine a logistics and distribution company that requires a system for inventory management. An agent could integrate multiple functionalities to automate the process: a hierarchical multi-agent system, where the main agent monitors the global inventory from a data source and joins others that deal with real-time stock monitoring, predictive analysis of demand changes based on generative AI models is incorporated, the ability to make autonomous decisions such as requesting inputs or merchandise, reordering products from suppliers, monitoring buyer requests, distributing inventory between sectors and planning distribution is added. The complex agent could add the ability to generate reports on the activities completed based, again, on generative AI.

In summary, generative AI models have evolved into base agents with different tools and functions that empower them and allow them to execute multiple tasks, but they still lack the autonomy, adaptability and versatility required for them to execute complete workflows autonomously and adaptively.

This is where agents in the strict sense come into play (which we will call agents driven by or based on generative AI, or simply AI agents). It is on the latter that we focus on this document, in order to try to understand them and distinguish them from other automation solutions.

The autonomy of agentic systems is not a binary issue. It can be ranked. This approach is useful not only for classification purposes, but also for determining the degree of human intervention required to execute workflows in an ethical and safe manner.

Hey, That's Not an Agent! How to Recognize Agents Powered by Generative AI

Generative AI-powered agents can employ a single agent or multiple agents with specific roles to understand requests, plan workflows, coordinate between agents with other specific roles, streamline actions, collaborate with humans, and validate results[31]. This is the example we saw in the previous section on inventory management.

At the core of each agent is a base agent or generative AI language model that provides a semantic understanding of the language and context of the task to be performed. Depending on the use case, agents in a system can use the same language model or rely on different specialized models. This approach enables some agents to share knowledge from the different language models they rely on, and others can be used to validate the outputs of the entire system, improving the quality, accuracy, and consistency of the process. The differential and potential of agents arises precisely from the addition of generative AI that allows the generation of content, the realization of reports, controls,

30 Shah, Chirag and White, Ryan W., "Agents are not enough", December 19, 2024, p. 1 [online: https://arxiv.org/html/2412.16241v1 , accessed 1/1/2024].

31 Deloitte AI Institute, "Prompting for action | How AI agents are reshaping the future of work," November 2024, p. 7, at https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-agents-multiagent-systems.pdf [Accessed November 24, 2024]

reviews, and tasks of different complexity and their combination with automation tools.

As we said, the autonomy of agents based on generative AI is presented by scales and is linked to three major aspects:

1) The **ability to plan, decide and execute** tasks that make up a workflow. The greater the complexity and magnitude of tasks, the more autonomy and contextual adaptability will be required.

2) Access **to short-term and long-term memory,** as well as other **tools** such as **APIs**, that they can use to complete their assigned task within a workflow. These resources allow the system to perceive and interact with the environment with greater autonomy and dynamics, as designed based on the anticipated human intervention.

3) The **integration with other agents** in a synchronized manner, generally from an orchestrating agent.

Components of AI Agents[32]

a. **Base agent or language model.** It is the central processing unit that integrates different functionalities and accesses different tools. It operates as the system's reasoning engine that orchestrates logical inference, planning, contextual understanding, and personalized interaction[33].

In the context of generative AI-based agents, the model refers to the language model that will be used as a centralized decision-maker for the system processes, and which can be of different sizes (small/large), as well as open source or proprietary, general purpose, multimodal or fine-tuned, depending on the needs of the specific agent architecture designed[34].

A memory module. Stores and retrieves information to maintain context and continuity over time. Memory allows the agent to recall past interactions and domain-specific knowledge to understand long-term goals and adapt to changing conditions[35].

The memory module supports the agent's ability to maintain context across interactions, ensuring personalized and consistent responses[36].

This module can be implemented by appealing to different techniques such as short and long-term memory; access to vector databases and semantic search; storage and labeling of metadata; augmented information retrieval generation (RAG), among others[37].

32 Expand at 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 [Accessed 12/26/2024].
33 Cf. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", January 1, 2025, p. 10 [Available at https://arxiv.org/abs/2501.00881, accessed 1/13/2024].
34 Wiesinger, Julia , Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, September 2024, p. 12.
35 Expand at 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 [Accessed 12/26/2024].
36 Cf. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", p. 9.
37 Expand at 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 [Accessed 12/26/2024].

b. **Tools.** They enable access to external resources that the agent can use to execute specific tasks (for example, databases and APIs). They are defined by the developer at the system design stage based on the use case.

In other words, the tools enable agents to interact with external data and services, unlock a broader range of applications than the underlying model alone offers[38] to access, retrieve and process information from diverse sources, and ensure that their actions are informed, adaptive and aligned with operational objectives[39].

The model decides, depending on the degree of autonomy it presents, when and which tools to use and integrate the results into its predictions to improve efficiency and precision in the execution of tasks that make up workflows.

Orchestration layer. This layer is present in some agent architectures[40], and consists of a cyclical process that governs how the agent receives information, performs some internal "reasoning," and uses this to inform its next action or decision. In general, this loop will continue until the agent has reached its goal or a stopping point. The complexity of the orchestration layer will vary depending on the agent and the goal it is pursuing[41].

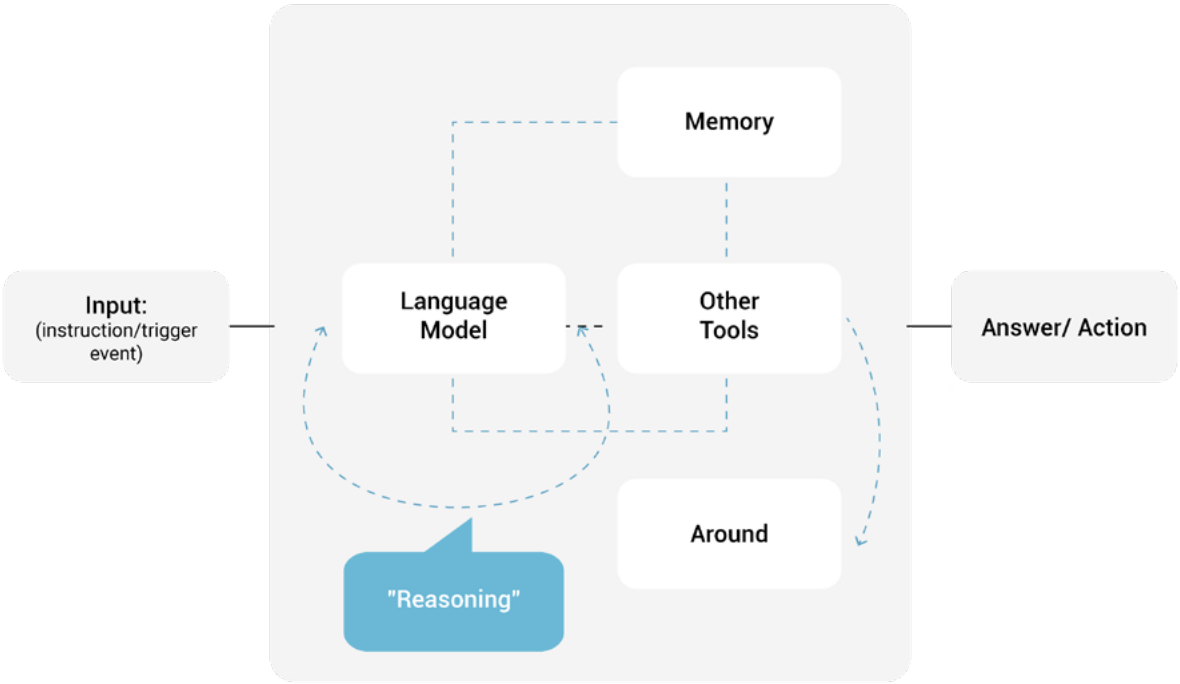


Figure 3. Essential components of a generative AI-powered agent.

38 Wiesinger, Julia , Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, September 2024, p. 7.
39 Cf. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", p. 11.
40 See the section "Architectures of agents based on generative AI" in https://docs.google.com/document/d/17e_NPGNd43sPGJ5grsJJpTIlsNOtLQ5W7GUxghenlXY/edit?tab=t.0#heading=h.u6zueac1pluc.
41 Wiesinger, Julia , Marlow, Patrick and Vuskovic, Vladimir, "Agents", Google, September 2024, p. 7/8.

How do I decide if an agent is the right solution for my problem? From AI-free automations to agentic workflows with scales of autonomy

The orchestration and deployment of authentic workflows should always follow an intuitive logic: adopt the simplest and most efficient solutions possible. AI agents should serve as de-bureaucratization tools rather than complicate trivial problems that can be solved more effectively with simpler solutions. Just as we do not take a plane to go to the corner, technological solutions must be proportional to the needs of the organization adopting them.

Often, a simpler alternative—even one without generative AI—can be just as effective, cheaper, and better suited to real needs. In many cases, organizations might choose not to build agents at all due to factors such as costs, added latency, and risks associated with generative AI's inherent limitations, which typically underpin agentic workflows.

As Todd Park, former CTO of US Digital Services, insightfully put it: *"We need both kinds of people: people who can hack the technology and people who can hack the bureaucracy. The overarching goal here is to get everything to grow together, inevitably joining up into a forest canopy so as to create a functional and interconnected system."*

This perspective highlights that technology and bureaucracy must co-evolve—not in opposition, but in synchronization—to create scalable, functional, and interconnected systems that balance efficiency, governance, and adaptability.

In the context of agents, this might mean choosing not to build **agents at all**[42], due to the costs, the latency they add to the process, and the risks they bring with them due to the inherent limitations of generative AI[43], qwhich is the foundation technology upon which agentic workflows are typically built.

To do this, it is important to distinguish **agents powered by generative AI** from other systems that can be developed to organize the execution of tasks and processes, such as **automated workflows that rely on language models** and **automations without AI**. Let's see:

1) **Non-AI automations**, refer to systems designed to execute tasks automatically, following predefined rules and sequences based on deterministic logic.

2) **Automated workflows that rely on language models**, sare solutions where language models and tools are orchestrated through code paths predefined by the development team. They offer predictability and consistency for well-defined tasks[44].

3) **Maximum autonomy agentic workflows (agents in the strict sense)**, are applications in which language models perceive the environment, interpret user requirements, and autonomously and dynamically direct their processes and use of available tools, while maintaining control over how they perform tasks. They are

42 Anthropic, "Building effective agents," December 19, 2024, at https://www.anthropic.com/research/building-effective-agents, accessed December 22, 2024.
43 Corvalán, J. G., Estevez, E., Le Fevre Cervini, E. M., Schapira, D., & Simari, G. (2023). *ChatGPT vs GPT-4: Imperfect by design? Exploring the limits of conversational artificial intelligence* . Autonomous City of Buenos Aires: La Ley; Autonomous City of Buenos Aires: Faculty of Law - UBA. in 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," December 19, 2024, at https://www.anthropic.com/research/building-effective-agents, accessed December 22, 2024.

a better option when flexibility, dynamism, and large-scale model-based decision making are needed[45].

Non-AI automations, automated **flows that integrate generative AI,** and **agents powered by generative AI** present differences in terms of the techniques and technology behind the system; the type of flows and tasks that can be completed with each of them, and certain strengths and weaknesses that are important to consider in order to choose the architecture that best fits the use case[46]:

TYPE	DESCRIPTION	APPLIED TECHNIQUES AND TECHNOLOGY	TASKS YOU CAN SOLVE	STRENGTHS	WEAKNESSES	EXAMPLE
Automation without AI	A system that executes automatic tasks based on predefined rules	Boolean logic	Deterministic and predefined tasks	It offers reliable and quick-to-execute results	Limited to explicitly scheduled tasks. Cannot adapt to new scenarios	Send a notification to Slack every time a user registers on our website
Workflow that integrates IAGen	A system that calls an LLM through an API to execute one or more tasks in a predefined workflow	Boolean logic + IAGen	Deterministic tasks that require flexibility	Better handling of complex flows in which IAGen adds value in the execution of certain tasks	Integrates the risks inherent to IAGen (hallucinations, biases) into the flow	Analyze, classify and extract variable data from incoming documents with ChatGPT, and complete a new document using pre-defined templates
IAGen-based agent system	A system designed to perform non-deterministic tasks autonomously as needed	IAGen + tools	Non-deterministic and adaptive tasks	Highly adaptable to new variables. Simulates human reasoning and behavior	Less reliable, can produce unexpected results. Slower to execute	Design and coordinate an autonomous and dynamic trip that includes airfare, accommodation, excursions and internal transfers based on a user's request

45 Anthropic, "Building effective agents," December 19, 2024, at https://www.anthropic.com/research/building-effective-agents, accessed December 22, 2024.

46 The following table is an adaptation of the proposal by Alexander Kantjas in the post of 12/29/2024, available at: 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 (accessed on 30/12/2024).

As a review: key insights from this chapter

From “Is that AI?” to “When is there an AI agent?”

- The definition of what AI is and the discussion about whether a solution contains it has been a constant subject of debate over the past two decades. The debate is rekindled with the arrival of agents powered by generative AI.
- The emergence of AI agents once again raises the question of what defines them and how they differ from other technological solutions, primarily other automation solutions.

Autonomy and adaptability

- AI systems can be classified on scales according to their degree of autonomy and adaptability, but the binary approach (it is either autonomous or not) should be replaced by tiered models, as was the case with autonomous vehicles.
- In the field of AI agents, these scales can be determined based on the degree of independence of the system in making decisions and executing tasks and complete work processes. This scale helps to identify the need for less or more human supervision.

Generative AI and large language models: the transition to “base agents”

- Generative AI has made it possible to democratize access to technology, no longer being exclusive to large organizations.
- Language models have evolved from unimodal multipurpose conversational assistants to multimodal systems that can process images, audio, and video.
- Multimodality and the integration of tools such as internet access, task scheduling, personalization, and conversational memory, among others, have been the most notable advances in recent years in applications such as those offered by OpenAI and Gemini.

Multitasking base agents: a new paradigm of “human-machine” collaboration”

- Generative AI applications such as ChatGPT, Gemini, Copilot, among others, have boosted productivity, but their autonomy is still limited, since they cannot autonomously complete end-to-end multi-task workflows.
- Generative AI-powered agents appear to bring autonomy and adaptability, allowing them to orchestrate and execute entire workflows.
- With its appearance, a change of role is anticipated for users, who will go from giving instructions to supervising the decisions of agents where relevant.

Components of Generative AI-Powered Agents

» Although there is no single consensus on what defines a generative AI-powered agent, there is agreement that they are characterized by their autonomy and adaptability, since they integrate tools to perceive and interact with the environment, allowing them to plan, coordinate and execute tasks in end-to-end workflows.

» Generative AI-based agents feature the following **essential components**:

- **Base agent or language model:** central engine of reasoning and planning.
- **Memory module:** allows maintaining context and continuity in interactions.
- **Tools:** integrate APIs and external databases to improve capabilities.

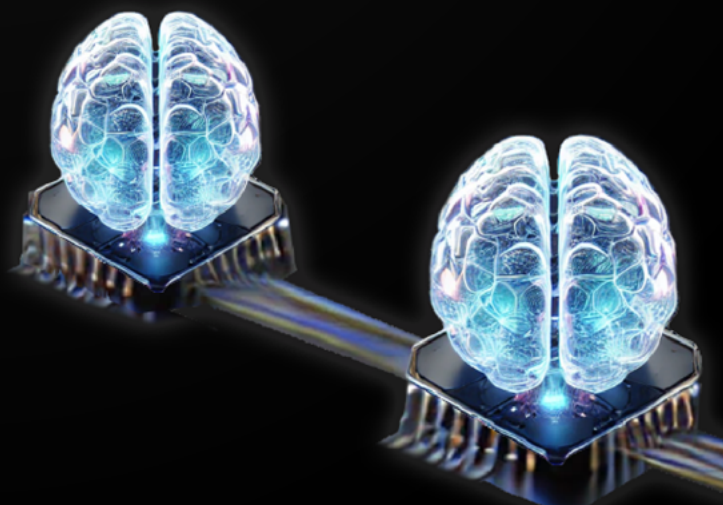
What is the most suitable automation solution?

» An AI agent will not always be the best automation option. Sometimes, traditional automation, without AI, can be more efficient and safer.

» There are key differences between non-AI automations, automated flows supported by generative AI, and generative AI-powered agents. Selecting the most appropriate solution will depend on the use case and the context in which it is integrated into the organization.

03

How to classify generative AI agents according to their degree of autonomy: a five-level proposal



How to classify generative AI agents according to their degree of autonomy: a five-level proposal

There are different scales designed to describe the levels of automation that intelligent systems can achieve. A prominent example is the widely recognized J3016 standard, which classifies six levels of automation in systems that operate autonomous vehicles[47].

Along those lines, and considering that autonomy is not a simple or binary characteristic, it is equally valuable to develop nuanced scales that adapt to the specific context of agents based on generative AI[48].

Based on the classification we outlined in 2021 in the AI Treaty, in the context of generative AI and in view of the arrival of agents driven by this type of AI, we propose a new version of five levels of automation. This scale ranges from automation without AI to agentic systems based on generative AI, ordered according to the degree of autonomy of the system:

- **First level of automation: automation without AI.** At this level, the system is completely dependent on human activation through a closed menu of predefined options under a "traditional software" logic.

  Example: A chatbot that answers frequently asked questions based on predetermined options from a closed menu, such as "Select: 1. Contact information, 2. Available services."

- **Second level of automation: automation with generative AI.** At this level, the solution integrates generative AI for the execution of tasks and subtasks but without autonomy, always requiring active user intervention through instructions in natural language.

  Example: Using ChatGPT, Gemini, or Copilot to write an executive report based on user-entered data, such as a summary of key performance metrics or analysis of trends in a specific market.

Third, fourth and fifth level of automation: agentic workflows

- **Third level of automation. Low-autonomy agentic workflow.** This third level is the beginning of what is known as "agent AI." It is linked to AI systems that can operate autonomously, making decisions and performing actions without constant human intervention. At this third level, the solution relies on a personalized base agent that acts with low autonomy to execute specific tasks predefined by the user.

47 Expand in Corvalán Juan G., Sa Zeichen Gustavo and Albertsen Lihué María, "Automated Administrative Activity. Artificial intelligence, regulatory power, algorithmic delegation, automated administrative act and reserve of humanity", in Treaty on Artificial Intelligence and Law , volume IV, p. 3-50. Director's note: The 6 levels that make up the standard are the following: 0-human control; 1-driver assistance: the system assists the driver in a specific task; 2-partial automation: the system is in charge of at least two driving functions; 3-conditional autonomy: the system manages all functions but requires the human to take control in critical contexts; 4-highly automated driving: the system manages all functions, being able to do so even if the human does not respond to the intervention request; 5-total autonomy: the system is capable of driving autonomously in every environment and situation. On the levels of the J3016 standard. See Corvalán, Juan G. - Danesi, Cecilia C., Carro María Victoria "Civil liability of artificial intelligence", volume II of this *Treaty*.

48 Expand in Corvalán Juan G., Sa Zeichen Gustavo and Albertsen Lihué María, "Automated Administrative Activity. Artificial intelligence, regulatory power, algorithmic delegation, automated administrative act and reserve of humanity".

Example: A custom GPT that automates the drafting of legal claims based on user-provided information, generating structured drafts that include background, legal basis, and claims, with human review prior to filing.

» **Fourth level of automation: agentic workflow with moderate autonomy (workflows that rely on generative AI).** This is where the least sophisticated version of a design is found, driven by base agents that are orchestrated through routes predefined by humans to execute workflows and complete different tasks. This is the case of the workflow that relies on generative AI, which we referred to above.

Example: A workflow running on an automation platform like n8n or Notion, which combines generative AI to write reports, queries external APIs to verify data, and automatically sends results to decision makers based on preconfigured criteria.

» **Fifth level of automation: highly autonomous agentic workflow (agents driven by generative AI in the strictest sense)**. In these systems, language models can manage complete workflows autonomously, adaptively and dynamically, using different tools that allow them to perceive and interact with the environment.

Example: A generative AI-powered agent capable of autonomously organizing a leisure trip that includes booking tickets, hotel accommodations, and planning customized sightseeing tours. This agent coordinates schedules, consults flight and hotel databases to optimize costs and times, and suggests itineraries tailored to the user's preferences.

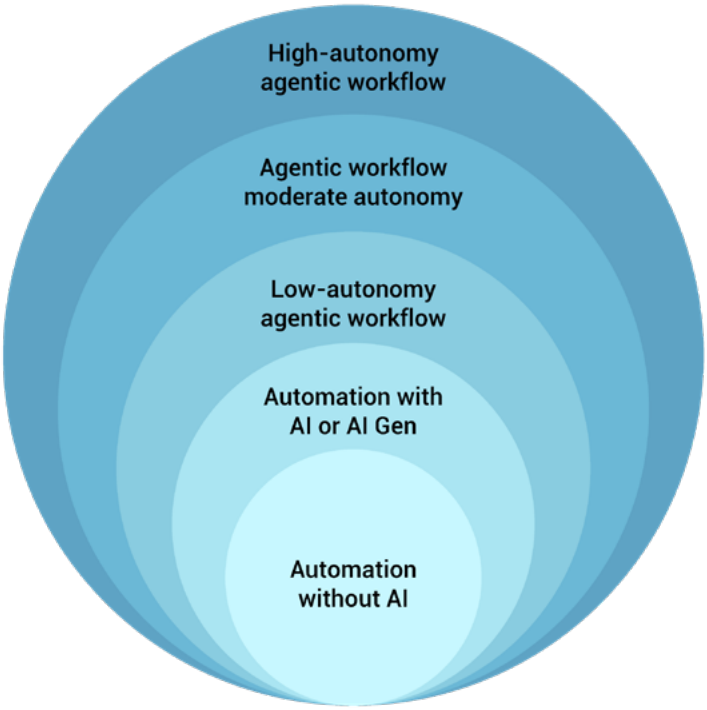


Diagram of five levels of automation: from AI-free automations to autonomous agents.

As a review: key insights from this chapter

Automation scales are used to describe different degrees of autonomy in intelligent systems. Inspired by standards such as J3016 for autonomous vehicles, a specific classification for generative AI-powered agents is proposed in five levels, ranging from non-AI systems to fully autonomous agents.

a. First level: Automation without AI

» Traditional systems that rely entirely on human activation, without learning or adaptation. *Example:* Chatbots with predefined responses.

b. Second level: Automation with generative AI

» It incorporates generative AI but requires active user intervention through natural language. *Example:* Using ChatGPT or Copilot to generate reports with data provided by the user.

c. Third level: Low-autonomy agentic workflow

» Customized agents execute tasks autonomously within parameters predefined by the user. *Example:* A GPT specialized in legal writing, which prepares drafts with human supervision.

d. Fourth level: Moderately autonomous agentic workflow *(automated workflow supported by generative AI)*

» Base agents drive task execution through routes predefined by the developer or user. *Example:* A system in an environment like n8n that can classify documents, write reports, query APIs, and send results automatically.

e. Fifth level: Highly autonomous agentic workflow *(agents in the strict sense)*

» They can manage complete workflows autonomously, adaptively and dynamically. *Example:* An AI agent that organizes trips, reserves hotels and optimizes itineraries.

04

Human Augmented Agents? The relevance of human intervention in the context of human augmented agents

Human Augmented Agents? The relevance of human intervention in the context of human augmented agents

The highest degree of autonomy and adaptability is directly related to the degree of human intervention on the agents. Here, it is key to detect critical points or tasks in the workflows so that experts can check that the partial or final outputs of the system (actions/responses) are aligned with the objectives of the organizations and with the fundamental rights.

The necessity ofhuman control is is frequently emphasized in film and television For instance, in 2001: A Space Odyssey, the transformation of HAL 9000 led to unforeseen dangers, while in Her, Samantha evolved beyond Dorothy's expectations, pushing the boundaries of artificial intelligence. Similarly, in I, Robot, VIKI initially appears as a technology designed to protect humanity but ultimately issues dangerous commands to the NS-5 robots. These fictional examples illustrate the importance of establishing safeguards and human oversight in AI-driven systems.

From a more technical perspective, several factors justify redesigning human intervention versus agentic workflows with medium or high autonomy.

The underlying probabilistic nature

Sequential prediction and derived techniques present in Transformers generate probabilistic responses. As AI agents are driven by generative AI base agents, the non-deterministic nature implies that their responses can vary significantly even under similar input instructions (prompts) or trigger events, increasing the likelihood of errors, biases, or unexpected outcomes[49] .

In this context, human intervention allows for the correction of possible deviations and at the same time provides critical judgment in situations where the models may generate erroneous content (hallucinate) or results that imply the violation of fundamental rights, current regulations or ethical standards applicable to the business.

It is important to introduce then the notion of the **Human Augmented Agent**[50], which expresses the idea that the human person is integrated into the system of agents under the human in the loop logic , at specific points of the previously defined workflow, to provide feedback on the status of the responses and/or actions of the system , in order to validate, refine, replace or cancel them, as appropriate[51].

49 See Corvalán Juan G. and Sánchez Caparrós Mariana, "Guidelines for the use of ChatGPT and generative text AI in Justice", September 2023 [in https://ialab.com.ar/webia/wp-content/uploads/2023/11/Guia-de-directrices-usos-de-ChatGPT-e-IA-generativa-en-la-justicia.pdf, accessed on 1/13/2025]

50 In a similar sense, see Cf. Bousetouane, Fouad, "Agentic Systems: A Guide to Transforming Industries with Vertical AI Agents", p. 25. Also LangGraph, "Introducing ambient agents", January 14, 2025 [Available at https://blog.langchain.dev/introducing-ambient-agents/ , accessed 1/16/2025].

51 How could human intervention be operationalized? This could be done, for example, through an agent inbox, showing the open lines of communication between the system and the person. See more in LangGraph, "Introducing ambient agents".

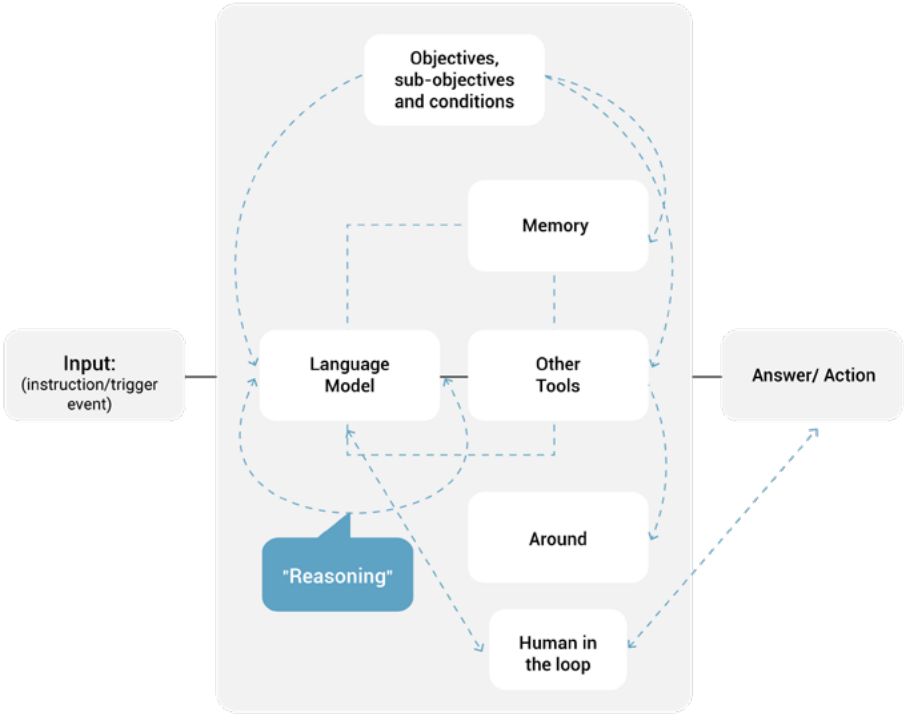


Figure 4. Basic architecture and essential components of a "Human Augmented Agent"

The **human component** is essential to the extent that[52]:

1) **Itigates risks inherent in language models**, such as hallucinations, errors, and biases, improving reliability and facilitating deployment into production..

2) **Mirrors human decision-making processes**, involving communication, consultation, and collaborative or hierarchical validation, thereby enhancing user trust and adoption.

3) **Enhances long-term memory and learning**, as **human feedback** allows AI systems to continuously refine their outputs and improve decision-making over time.

52 In a similar sense, see LangGraph, "Introducing ambient agents"

As a review: key insights from this chapter

As AI agents gain autonomy, human intervention becomes critical to monitor, correct and validate full and partial system outputs, ensuring their alignment with organizational objectives and applicable regulations.

Agentic flows, especially those with medium or high autonomy, require human supervision due to the probabilistic nature of the large language models that drive them.

Human Augmented Agents: Integrating Human Supervision

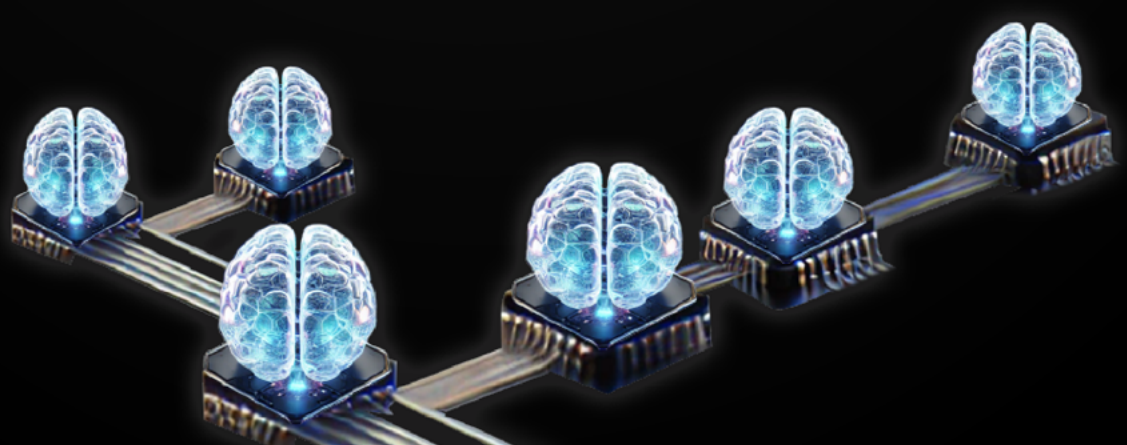
Human Augmented Agent concept incorporates human in the loop logi, where humans intervene at specific points in the workflow to:

01	02	03	04	05
Validate, refine or override AI-generated responses.	Mitigate errors and biases before deployment to production.	Building trust through collaborative validation processes.	Prevent hallucinations and ensure regulatory and ethical compliance.	Improve system learning through human feedback.

The balance between autonomy and human control is key to designing trustworthy agents aligned with ethical and organizational values.

05

Initial taxonomy of specialized agents based on generative AI



Initial taxonomy of specialized agents based on generative AI

In the context of generative AI-based agents, language models can occupy different "positions" in the context of a broader system to perform specific functions, appealing to their analytical capacity and their specialty for the execution, planning and coordination of tasks, as well as to different tools available to them, such as short- or long-term memory and access to third-party or proprietary APIs.

Below, we propose an **initial classification** of the different types of agents based on generative AI according to their functionality, origin and degree of specificity

Multitasking Base Agent

The **Base Agent** represents the core functionality of a generative language model. It is ideal for general tasks and serves as a basis for developing more specialized agents. In simple flows, it can answer questions directly or process initial data. In more complex processes, it collaborates with other agents to complete more advanced tasks.

a. Features

It can perform multiple general tasks, be customized through simple prompts, directly access the AI model and use additional tools, such as internet access to complete searches and access additional information, or have memory, more or less limited, depending on the capabilities of the model.

b. Examples

ChatGPT-4o from Open AI. This model can perform multiple general tasks, is customizable through simple instructions, and is able to integrate tools that extend its capabilities such as:

- **(Limited) memory**, which allows you to temporarily store relevant information in a session, as well as draw on past conversations the user has had with the tool and refer to them in their responses, to better contextualize the interactions[53].
- **Internet access**, to perform searches and obtain additional and updated information, expanding your knowledge beyond your basic traininge.
- **Multimodality**, as it can process both text and images and audio, which makes it suitable for tasks that require a more comprehensive approach.

53 Expand in OpenAI, "How does Memory use past conversations?", updated in December 2024, at https://help.openai.com/en/articles/10303002-how-does-memory-use-past-conversations [Accessed on 12/28/2024]

Derived Agents (Task Specialists)

Derived Agents arise by extending the capabilities of the Base Agent by integrating additional tools within a framework (e.g. N8N, LangChain, etc.) that give it autonomy to execute specific tasks within a workflow aimed at meeting certain objectives.

These agents can be created using No-Code/Low-Code approaches or through custom programming (Code).

a. Types of focus

- **No-Code/Low-Code:** Agents are configured using pre-built nodes in tools like N8N, connecting to databases, APIs, and other services. The flow logic is mostly designed with simple expressions. For example, an "HTTP Request" node gets data from an API, and another "Function" node formats this data for further use.
- **Code:** It involves writing JavaScript code inside "Function" nodes to customize or extend the agent's logic. For example, a sentiment analysis algorithm can be implemented using a JavaScript library.
- **Mini-agents:** These are simplified versions of derived agents, designed to perform very specific tasks within a larger flow.

Orchestrating Agents (Intelligent Coordinators)

Orchestrator Agents act as "conductors", coordinating the interaction between Base Agents and Derived Agents to ensure that workflows are executed efficiently, coherently and dynamically. Their main function is to decide which agent intervenes at which time and in which phase of the process.

a. Types of Orchestration

- **Without IAGen:** They use predefined logic nodes, such as "If", "Switch" or "Loop", to structure flows according to manually established rules.
- **With IAGen (an orchestrating agent):** They incorporate generative language models that make dynamic decisions based on the results of previous tasks. For example, a "Function" node can call a language model with a prompt that says: "If the sentiment analysis of the text is positive, send an email; if it is negative, create a task in Jira."

b. Features

These agents are capable of making complex decisions, centralizing and optimizing workflows, and coordinating the execution of advanced tasks. Their implementation combines flow logic nodes with nodes that interact with language models.

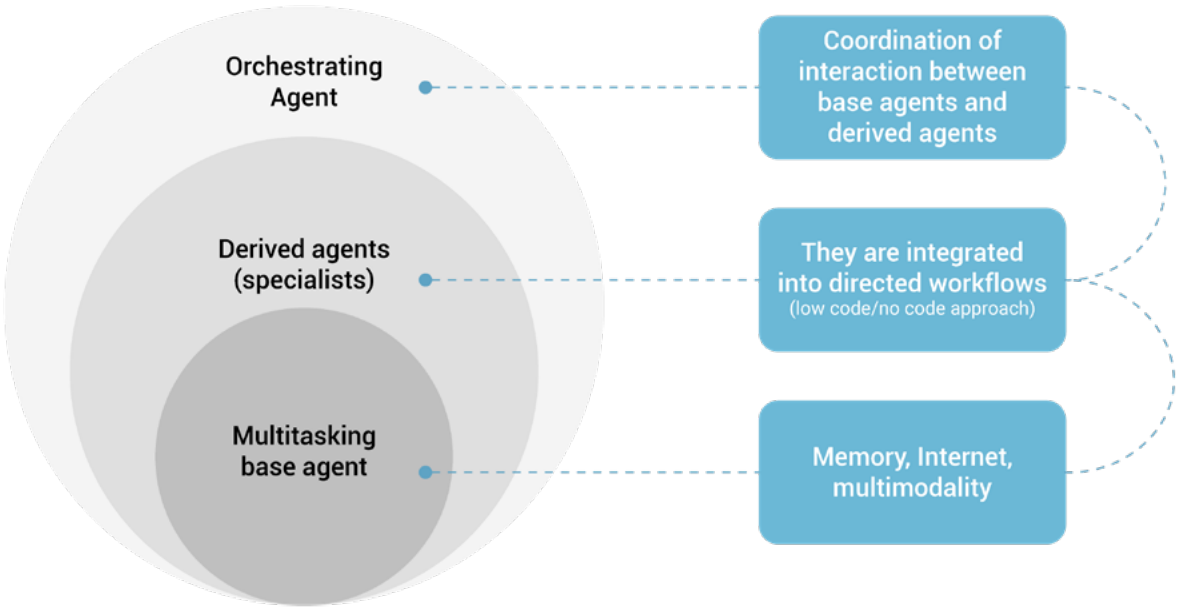


Diagram of the agentic ecosystem according to the proposed taxonomy.

As a review: key insights from this chapter

An initial classification is proposed to better understand how generative AI-based agents function according to their functionality, origin and degree of specificity:

a. Base Agents: Provide general capabilities and serve as a basis for more specific solutions. The ability to access tools such as memory, internet access, or multimodal processing makes them key players in multiple environments.

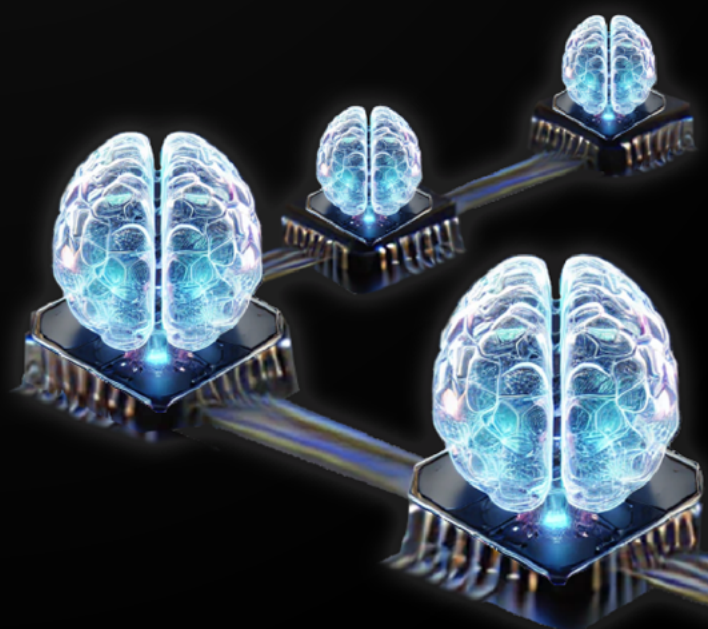
b. Derived Agents: extend the capabilities of the Base Agent through frameworks that grant them autonomy to execute specific tasks. Their development can be addressed with No-Code/Low-Code approaches for quick configurations or through custom programming for greater control and precision.

c. Orchestrator Agents: coordinate the interaction between Base and Derived Agents, ensuring efficient and adaptive execution of workflows. Depending on their level of sophistication, they can operate with predefined rules or integrate generative models for dynamic decision-making.

This initial taxonomy lays the foundation for the design of modular and scalable architectures in generative AI systems, allowing the combination of agents with different degrees of autonomy and specialization according to the needs of the environment in which they operate

06

Generative AI-based agent architectures



Generative AI-based agent architectures

The different architectures of generative AI-based agents allow combining the power of large language models with modular structures designed to optimize their functionality, through model customization and the integration of additional tools, such as short- and long-term memory, and access to databases and other external resources through APIs. In addition, they facilitate coordination between specialized agents, making it possible to complete complex workflows efficiently, autonomously, and dynamically.

In this context, it is important to explore some of the possible architectural configurations that a generative AI-based agent system can adopt, in order to understand how these systems can be optimized for practical applications and specific use cases, as well as to address the technical, ethical and legal challenges posed by their design and development.

Single-agent and multi-agent systems

The first distinction we can make is between **single-agent systems and multi-agent systems**.

Single agent systems are organized around a single agent based on a large language model, which is trained to understand and respond to user queries, generate content, or execute automated tasks based on predefined instructions.

In these cases, the agent typically operates effectively using a few tools within a single domain. Examples include customer service chatbots, virtual assistants for scheduling, and automated content generation tools[54].

Multi-agent systems arise when the workflow or process requires the execution of multiple tasks, and it is necessary to appeal to architectures that facilitate specialization through different independent agents (for example, planner, researcher, mathematics expert, etc.) that are integrated and coordinated through a single system[55].

Among the main benefits of using multi-agent systems are[56]:

- **Modularity:** Separate agents make it easier to develop, test, and maintain agents.
- **Specialization:** Expert agents can be created that focus on specific domains, which helps increase the overall performance of the system.
- **Control:** You can explicitly control how agents communicate, when, and how they interact with each other.

54 Expand at https://langchain-ai.github.io/langgraph/tutorials/multi_agent/multi-agent-collaboration/ [Accessed 11/29/2024]
55 See at https://langchain-ai.github.io/langgraph/concepts/multi_agent/ [Accessed 29/11/2024].
56 See at https://langchain-ai.github.io/langgraph/concepts/multi_agent/ [Accessed 29/11/2024].

These systems are thus composed of multiple independent agents collaborating with each other to achieve a complex goal or workflow that exceeds the capability of a single agent. Multiple tasks are coordinated between agents, as opposed to single agents which often require human coordination and intervention between tasks[57].

Multi-agent system architectures

In turn, **multi-agent systems** can be organized under different architectures:

a. Sequential task agent system

Each agent has a large language model on which it operates and a unique set of tools at its disposal, with prompt components connected to the agent's instruction fields to control its behavior[58].

For example, the investigator agent has a search component attached as a tool. The prompt tells the agent how to respond to its query, format the response, and pass the query and investigation results to the next agent in the flow. Each successive agent in the flow then builds on the work of the previous agent, creating a chain of reasoning to solve complex problems[59].

b. Multi-agent network

They can be used to tackle complex tasks, and consist of creating a specialized agent for each task or domain and routing the tasks to the correct "expert"[60]. In this architecture each agent can communicate with all other agents and any agent can decide which other agent to call next[61].

However, it should be noted that while this architecture is very flexible, it may not scale well as the number of agents increases, since it is difficult to determine which agent should be called next and how much information should be passed between agents[62].

c. Multi-agent system with supervisor agent (orchestrator)

This architecture is based on a single supervisor agent that is responsible for orchestrating, delegating and routing tasks between different specialized agents that make up the system.

57 See more on Google, "AI Business Trends 2025 report", at https://cloud.google.com/resources/ai-trends-report [Accessed 12/26/2024].
58 Expand on "Sequential Task Agents" at https://docs.langflow.org/starter-projects-sequential-agent [Accessed 2024-11-29].
59 Expand on "Sequential Task Agents" at https://docs.langflow.org/starter-projects-sequential-agent [Accessed 2024-11-29].
60 Expand on "Multi-agent systems architecture", at https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [Accessed 11/29/2024].
61 Expand on "Multi-agent systems architecture", at https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [Accessed 11/29/2024].
62 Expand on "Multi-agent systems architecture", at https://langchain-ai.github.io/langgraph/concepts/multi_agent/#multi-agent-architectures [Accessed 11/29/2024].

In these cases, the supervisor agent is responsible for managing the conversations between the other agents (workers), so once it receives the user's instruction, it should be prepared to analyze it, decide which agents should act, and return the response[63].

Now, let's imagine a use case for a construction company. You have an agent - an orchestrator - that receives a request from an architect. When processing the request data, you need to decide which type of agent should be involved in completing the request. For example, if you are asked to generate a construction plan for a twenty-story building, you could trigger the workflow that involves the materials planning agent, the supply agent, the labor assignment agent, the monitoring agent, and finally an agent with the ability to make a final report. Now, based on the request, the orchestrator agent should have enough knowledge to avoid triggering the flow corresponding to human resources recruitment, since it is not related to the topic of the request.

d. Hierarchical teams

For certain more complex applications or workflows, the multi-agent system may work better when the agents that comprise it are organized in a hierarchical manner.

This means that there are different specialized agents that depend on a higher supervisor, as well as mid-level supervisors, who orchestrate the execution of the different tasks that make up the flow[64].

For example, if we take the example of a system that controls the arrival of trucks in a transport company, agents are required to control delivery times and delays; agents to collect data in real time to predict traffic variations, accidents and multiple factors that may affect the variation; agents to make predictions based on the organization's history. Additionally, agents are added to send notifications to service managers and drivers when an unjustified deviation in times is detected.

However, since this is an activity that requires security control, it is useful to have an agent who supervises the entire process and an intermediate agent who, at the same time, effectively controls the one who monitors in real time. In this way, smarter decisions are made, before deciding to trigger the automatic alert, which can cause stress and security risks. In this way, the hierarchical team is formed.

63 See https://langchain-ai.github.io/langgraph/tutorials/multi_agent/agent_supervisor [Accessed 26/12/2024].
64 See https://langchain-ai.github.io/langgraph/tutorials/multi_agent/hierarchical_agent_teams/ [Accessed 26/12/2024].

As a review: key insights from this chapter

The architectural design of agents based on generative AI allows their functionality to be enhanced through modular structures that integrate advanced tools, such as memory, database access, and coordination of multiple agents. The choice between single-agent or multi-agent systems depends on the complexity of the workflow and the level of specialization required to respond to the organization's pain.

Single-agent systems are suited for tasks confined within a single domain, while multi-agent systems optimize more complex processes through division of labor, specialization, and collaboration between agents.

Multi-agent systems can , in turn, be organized into various architectures, each with specific advantages and challenges:

a. Sequential task systems, where agents work in a chain, building on the results of the previous one.

b. Multi-agent network, allowing flexible communication between specialists, albeit with scalability challenges.

c. Systems with a supervisory agent (orchestrator), where a central agent assigns tasks and optimizes the workflow.

d. Hierarchical teams, which establish levels of supervision to ensure control, efficiency and more sophisticated decision-making.

The choice of architecture depends on the use case and the technical, ethical and organizational requirements of the system. Understanding these structures allows for the design of more efficient, adaptable agents that are aligned with business objectives and current regulations.

07

Technical frameworks for agent development

Technical frameworks for agent development

The deployment of agents based on generative AI requires technical frameworks that offer the *software* and *hardware* infrastructure necessary to design, implement and deploy these solutions efficiently.

In this context, there are different frameworks available, which can be mainly classified into open source frameworks and proprietary tools, each with specific features and advantages. The choice between these options for building and deploying agents will depend on different factors, such as the objectives, the specific needs of the project and the resources available to each organization.

Below is an overview of some of the frameworks that exist today:

CrewAI

CrewAI is a framework for deploying AI agents, allowing you to create teams where each agent has specific roles, tools, and goals and works together to accomplish complex tasks.

CrewAI runs in the cloud, hosted locally or standalone, so it can be deployed on your own infrastructure with locally hosted options or leveraged through the user's preferred cloud service, giving them complete control over their environment[65].

This framework works with **4 main components** with different characteristics:

1) **Crew - Top-level organization.** Responsible for managing teams of AI agents; overseeing workflows; ensuring collaboration; delivering results.
2) **AI Agents - Specialized team members.** They have specific roles (researcher, writer); use designated tools; can delegate tasks; make autonomous decisions.
3) **Process - Workflow management system.** Defines collaboration patterns; controls task assignments; manages interactions; ensures efficient execution.
4) **Tasks - Individual assignments.** They set clear objectives; use specific tools; contribute to a broader process; produce actionable results.

CrewAI combines the crew, responsible for organizing the overall operation; the AI agents, who work on their specialized tasks; the process, which ensures smooth collaboration; and the tasks, which are completed to achieve the goal defined by the developers[66], to execute workflows autonomously and dynamically.

In short, it is a framework that allows the creation of specialized agents with defined roles, experience and objectives, which are supported by flexible and customized tools

65 Expand at https://www.crewai.com/ [Accessed 22/1/2025]
66 See https://docs.crewai.com/introduction [Accessed 21/1/2025].

to interact with external services and data sources, and which can collaborate with each other intelligently, cooperating, sharing information and coordinating tasks to achieve complex objectives. To this end, this framework allows the management of tasks through sequential or parallel workflows, with agents that automatically manage task dependencies[67].

In addition, it offers a community for knowledge exchange and consultation that currently has about 2000 members[68].

AutoGPT

AutoGPT[69] is an open source framework that allows you to create, deploy, and manage AI agents to automate complex workflows in a low-code manner, using drag-and-drop components as well as using a command console. Agents in AutoGPT combine different components, with and without AI, to create automated processes that can operate independently once configured.

Currently, this tool is offered immediately, through a local implementation[70], and also through access to a beta version after entering a waiting list to be able to use it in cloud mode[71].

In this framework, agents can be structured using linear processes for simple tasks; complex workflows with branches and multiple decision points; and multi-agent systems that work together to complete more complex tasks[72].

Agents in AutoGPT are built using "blocks," which are individual action components that make up a workflow. Blocks include: external service integrations, data processing tools, AI model connections, custom scripts, and conditional logic elements[73].

The AutoGPT Platform is built on a two-part architecture designed for scalability, flexibility and ease of use : the "AutoGPT Front-end" and the "AutoGPT Server"[74].

The AutoGPT Interface[75] offers multiple ways to interact with and leverage AI agents:

- **Agent Builder:** For those looking for more customization, AutoGPT offers an intuitive, low-code interface for users to design and configure their own AI agents.
- **Workflow Management:** IThe tool allows you to easily create, modify and optimize automation workflows, as well as create an agent by connecting blocks, where each block[76] performs a single action.
- **Deployment Controls:** AutoGPT allows you to manage the lifecycle of your agents, from testing to production.

67 See https://docs.crewai.com/introduction [Accessed 21/1/2025].
68 See https://community.crewai.com/about [Accessed 21/1/2025].
69 See https://agpt.co/ [Accessed 22/01/2025].
70 https://github.com/Significant-Gravitas/AutoGPT [Accessed 2025-01-23]
71 See https://agpt.co/waitlist [Accessed 23/01/2025].
72 Expand at https://agpt.co/blog/ai-agents-explained [Accessed 23/1/2025].
73 Expand at https://agpt.co/blog/ai-agents-explained [Accessed 23/1/2025].
74 See "Introducing the AutoGPT Platform" at https://agpt.co/blog/introducing-the-autogpt-platform [Accessed 23/1/2025].
75 See https://github.com/Significant-Gravitas/AutoGPT?tab=readme-ov-file#-autogpt-front-end [Accessed 21/01/2025].
76 Blocks represent actions and are the building blocks of your workflows (e.g. connections to external services, data processing tools, AI models for various tasks, custom scripts or functions, conditional logic, and decision-making components). See https://docs.agpt.co/#agents-and-workflows [Accessed 23/1/2025].

- **Ready-to-use agents:** If you don't want to create your own, this framework provides a library of pre-configured agents, ready to deploy.
- **Agent Interaction:** The tool provides a simple interface to run and interact with agents.
- **Monitoring and analysis:** It also allows you to track agent performance and obtain information for continuous improvement of automation processes.

The AutoGPT Server, on the other hand[77], is the engine of the platform. It is where the agents run, which can be triggered by external sources and can run continuously. The AutoGPT Server contains all the essential components that make AutoGPT work smoothly:

- **Source code:** This is the core logic that drives automation agents and processes.
- **Infrastructure:** Robust systems that guarantee reliable and scalable performance.
- **Marketplace:** A comprehensive marketplace where a wide range of pre-built agents can be found.

AutoGPT is mostly offered under the MIT license, except for the *autogpt_platform folder* which is offered under the Polyform Shield license[78], which only restricts the software from being used in a product or service that competes with this product[79].

Taskade

Taskade[80] is a project planning and management tool that has incorporated artificial intelligence agents designed to automate and optimize various tasks within your projects.

These agents, powered by the GPT-4o model, enable users to build, train, and deploy custom agents that can plan, investigate, and complete tasks autonomously, effectively collaborating with human teams or each other.

One of Taskade's great features is the ability to create custom agents tailored to your organization's specific needs. Using the platform's built-in agent builder, users can define the purpose and instructions for each agent to assist with tasks such as content generation, data analysis, or project management.

Additionally, these agents can be trained with specific knowledge by uploading customized documents, allowing them to offer contextually rich interactions and perform specialized tasks.

Integrating multiple agents within a single project makes it easier to manage complex or repetitive tasks, as each agent can handle different aspects of the project, working together to improve user productivity. This functionality positions Taskade as a versatile tool in the field of AI-assisted project management.

77 Expand at https://github.com/Significant-Gravitas/AutoGPT?tab=readme-ov-file#-autogpt-server [Accessed 23/1/2025].
78 See https://docs.agpt.co/#available-language-models [Accessed 23/1/2025].
79 Examples of such a violation would be: offering a service that provides direct access to the AutoGPT Platform outside of your organization or creating a similar product that competes with the AutoGPT Platform by placing a wrapper around AutoGPT. See : "Introducing the AutoGPT Platform". If you have any questions about whether your use violates this license, you may contact: contact@agpt.co
80 Taskade, "AI Agents Overview," in https://www.taskade.com/ai/agents [Accessed January 17, 2025].

Taskade's AI agents excel at being customizable, integrating with external tools, and adapting to a variety of workflows. They can be customized through user-defined instructions.

This feature allows agent's capabilities to be tailored to specific contexts, either by defining clear parameters in the form of prompts or by loading relevant documents to train them with specialized information. Thanks to this capability, agents can tackle tasks such as document writing, data analysis, or project planning and coordination.

Additionally, agents not only operate autonomously, but can also collaborate with each other within a single project. This modular integration allows workflows to be divided into smaller components, assigning specific tasks to specialized agents who interact fluidly to achieve the proposed objectives. In addition, agents can be integrated with external APIs and databases, expanding their functionality and having access to updated and accurate information.

Plans and prices

	FREE PLAN	TASKADE PRO	TASKADE TEAM
Price	$0	• $8 - 1 user/month billed annually • $10 - 1 user/month billed monthly	• $16 - 1 user/month billed annually • $20 - 1 user/month billed monthly
Included Functios	• 1 AI agent • 1 workspace • 5 AI requests per day • Multi-platform access • To-do lists • Mind maps • Flowcharts • Kanban boards • Calendars and more	• 1 AI agent • 1 workspace • 5 AI requests per day • Multi-platform access • To-do lists • Mind maps • Flowcharts • Kanban boards • Calendars and more • Custom AI agents • Unlimited use of AI • Unlimited file uploads • Live memory and knowledge • Study and generator of AI projects • Gantt, table and custom fields • Integrations with your favorite tools	• 1 AI agent • 1 workspace • 5 AI requests per day • Multi-platform access • To-do lists • Mind maps • Flowcharts • Kanban boards • Calendars and more • Custom AI agents • Unlimited use of AI • Unlimited file uploads • Live memory and knowledge • Study and generator of AI projects • Gantt, table and custom fields • Integrations with your favorite tools • Multi-agent teams • Unlimited workspaces • Unlimited AI Automation • Unlimited sharing and integration • Unlimited version history • Custom tools for AI agents • Advanced team permissions • SSO, API and Enterprise Features

n8n

n8n is a workflow automation platform that can be used to build generative AI-based agents by seamlessly integrating multiple applications and automating tasks.

From a more technical perspective, it is a fair-code workflow automation tool that allows users to create custom processes and automate tasks across a variety of services. By providing a low-code/no-code approach, n8n simplifies the connection between applications, enabling data exchange and automated actions.

n8n is **open source and self-hosted**; offers over 400 integrations with popular applications, allowing users to connect and automate a wide variety of services; provides an **intuitive visual interface** that makes it easy to create workflows by connecting "nodes" that represent different tasks or operations; provides **flexibility for code**, allowing users to extend functionality with JavaScript or Python if they need more complex logic; and allows you to **create custom connectors** to any service with an API.

Additionally, it offers **native AI capabilities** , as it makes it easy to build LangChain-based AI agent workflows with your own data and models; it is **high-performance** , as it can handle up to 220 workflow executions per second on a single instance, allowing you to build AI agents that can handle a high volume of interactions efficiently; and it offers **project and role management features**, which are crucial for collaborative development and management of AI agents, especially in a research or team environment.

n8n provides an ideal environment for building AI agents due to its flexibility, integration capabilities, and focus on automation. Let's see how the different components of an AI agent can be implemented in n8n:

1) **Language Model**. It allows you to integrate a language model through nodes that connect to APIs from AI providers such as OpenAI or Hugging Face. These nodes allow you to send prompts to the model and receive responses, which can then be used to guide the workflow..

2) **Memory.** It can be implemented using data storage nodes, such as databases or files. It also offers nodes for working with short-term memory, such as workflow variables, which can store information temporarily during the agent's execution.

3) **Tool.** It offers a wide range of nodes that act as tools, such as nodes for sending emails, interacting with APIs, reading and writing files, etc. The agent can use these tools to perform actions based on decisions made by the language model.

4) **Orchestration.** Orchestration is defined by connecting the different nodes in a workflow. n8n's visual interface makes it easy to create complex workflows that define the agent logic and how it interacts with the language model, memory, and tools.

5) **Error Handling and Monitoring.** It includes features for error handling and workflow monitoring. Users can set up notifications to be alerted about workflow failures via email, Slack, or other platforms. These features are essential to maintaining reliability. and the stability of AI agents built on n8n, allowing developers to quickly identify and fix issues.

Advantages and Disadvantages of n8n

While n8n offers a flexible and powerful toolset for building AI agents, it is important to consider its advantages and disadvantages in relation to the specific needs of each project:

ADVANTAGES	DISADVANTAGES
Flexibility: Allows the creation of agents with complex logic.	Learning Curve: Requires some technical knowledge to fully utilize its capabilities.
Integrations: Integrates with a wide range of services and APIs.	Scalability: It may require a more robust infrastructure to handle large volumes of data or requests.
Open source: Allows for customization and full control over the agent.	Community Dependency: Some integrations or features may depend on community development.
Visual: Makes it easier to create and understand complex workflows.	

In conclusion, n8n offers a flexible and powerful set of tools for building AI agents that fit the definition of autonomous and adaptive systems presented in this paper. Its ability to integrate language models, memory, and tools, along with its intuitive visual interface, makes it an attractive option for developers looking to create autonomous and adaptive agents.
Compared to other tools, it stands out for its versatility, its focus on automation, and its open-source nature. These features make it an ideal platform for research and development of AI agents, especially in academic settings where flexibility and control are crucial.

Langflow y LangGraph

Langflow[81] *low-code* tool for developers that makes it easy to build AI agents and workflows that can use any API, model, or database. It is a visual framework for building open-source RAG and multi-agent applications, powered by Python, and fully customizable[82].

81 See https://www.langflow.org/ [Accessed 11/28/2024].
82 Expand at https://docs.langflow.org/ [Accessed 11/28/2024].

The release of **Langflow version 1.1** in late November 2024 introduced an agent component, which is ready to support complex agent orchestrations, with multi-model selection, chat memory, and traceable intermediate steps to execute reasoning actions and calls to other tools. In addition, it allows agents to invoke other agents as tools, creating a multi-agent system that can interact and develop with each other[83].

This tool enables recursive orchestration for dynamic, multi-layered problem solving, where agents can compose complex workflows by calling each other in sequence or in nested formations[84].

Additionally, Langflow offers a testing area or playground so that the developer can have the agents interact in order to adjust and refine what is necessary to achieve the desired results. And it also includes a library of predefined templates for different use cases and methodologies (e.g., wizards, Q&A, content generation), to get up and running quickly[85].

LangGraph is an open source framework that provides a library for creating agent and multi-agent workflows. It allows defining workflows that involve cycles, which is essential for most agent architectures. It also includes built-in persistence, allowing for advanced memory and human intervention features. It is developed by LangChain Inc, but can be used without LangChain[86].

Thus, among the main characteristics of this framework are[87]:

- Cycles and **branches**, since it allows implementing loops and conditionals in applications.
- Persistence , because **it** automatically saves state after each step of the graph. It pauses and resumes system execution at any time to support error recovery, human-involved workflows, time travel, and more. This allows the agent to basically "pause" and wait for the user's response.
- The **Human-in-the-Loop**, because it supports human interaction patterns natively, allowing to interrupt the execution of the flow to approve or edit the next action planned by the agent.
- **Streaming support**, since it streams outputs as they are produced by each node (including token streaming).
- **Embedded long-term memory**[88], essentially a namespaced key-value store that supports semantic search, making it easier for agents to update their "memory" after human interactions[89].
- Integration **with LangChain**, because it integrates with LangChain and LangSmith (but does not require them).

Agents built with LangGraph can be deployed using the LangGraph Platform, a commercial solution for deploying agent applications in production, built on top of the open source LangGraph framework [90].

83 Expand at https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6 [Accessed 11/28/2024].
84 Expand at https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6 [Accessed 11/28/2024].
85 Expand at https://medium.com/logspace/langflow-1-1-release-b6df2f8189a6 [Accessed 11/28/2024].
86 Expand at https://langchain-ai.github.io/langgraph/ [Accessed 11/29/2024].
87 Expand at https://langchain-ai.github.io/langgraph/ [Accessed 11/29/2024]. Also in LangGraph, "Introducing ambient agents".
88 Expand at https://langchain-ai.github.io/langgraph/concepts/memory/?ref=blog.langchain.dev#long-term-memory [Accessed 16/1/2025].
89 LangGraph, "Introducing ambient agents".
90 Expand at https://langchain-ai.github.io/langgraph/ [Accessed 11/29/2024].

In January 2025, LangGraph has made available to users its first reference agent implementation: an email assistant, offered both as a free-to-use hosted agent[91] and as an open source project[92].

As a review: key insights from this chapter

The deployment of agents based on generative AI requires technical infrastructures that allow for efficient, scalable implementation that is adaptable to different needs.

There are multiple frameworks that facilitate this task, from open source options to proprietary tools, each with specific features depending on the degree of customization, automation and complexity of the workflow.

The reviewed **technical frameworks** enable:

a. Deploy specialized agents with defined roles within agentic workflows.

b. Coordinate multiple agents through orchestrated and dynamic architectures.

c. Streamline automation with visual interfaces, API integration, and advanced tools.

d. Facilitate human-AI collaboration by incorporating monitoring of system outputs at critical stages of the process.

The choice of the most appropriate framework will depend on the level of autonomy required for the solution sought, the available infrastructure and the regulations applicable to the organization.

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

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

Regulation of AI Agents

The increasing autonomy and adaptability of AI agents bring forth significant regulatory challenges, requiring frameworks that ensure transparency, accountability, and ethical compliance. Given that agentic AI can make decisions, automate workflows, and interact with users dynamically, regulations must address issues of responsibility, fairness, data security, and human oversight.

One of the primary concerns, as we underlined, is accountability and liability, particularly in high-risk sectors such as finance, healthcare, and public services. Organizations deploying AI agents must ensure that legal liability is clearly defined when an AI system makes an incorrect or harmful decision. Another fundamental issue is bias and fairness, as AI models operate on probabilistic systems and can reinforce existing biases. Regulatory frameworks should mandate fairness audits, explainability requirements, and bias mitigation strategies to prevent discriminatory outcomes.

Data protection and privacy remain at the forefront of AI regulation. AI agents often process sensitive user data, necessitating strict adherence to global privacy laws such as GDPR in Europe and CCPA in California to safeguard individuals' rights and prevent unauthorized data use. Additionally, security and risk management are crucial since AI agents can become targets for adversarial attacks, where malicious actors manipulate models or exploit weaknesses. Regulations should enforce robust security protocols, real-time monitoring, and risk assessment measures to ensure the integrity of AI-driven systems.

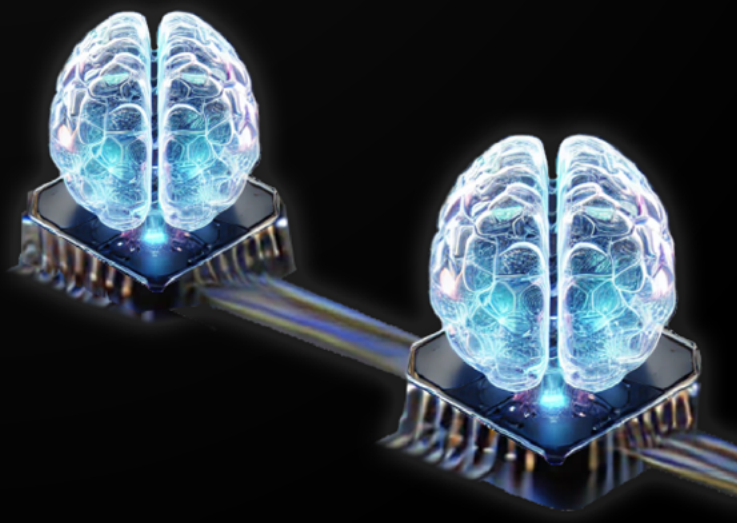
Another essential regulatory focus is human oversight, or the human-in-the-loop approach, which emphasizes human intervention at critical decision points to mitigate hallucinations, ethical concerns, and potential legal violations. Alongside this, AI-agent transparency and explainability must be ensured, requiring organizations to implement mechanisms that allow auditability and clear insights into how AI agents reach their decisions, ensuring that users and regulators understand the reasoning behind their actions.

The EU AI Act is one of the most comprehensive regulatory efforts addressing these challenges, classifying AI applications into risk categories and imposing strict obligations on high-risk AI agents. Other international bodies, such as the OECD and UNESCO, are also shaping global regulatory standards.

As AI agents continue to advance in autonomy and integration within decision-making processes, future regulations will likely include certification requirements, compliance audits, and real-time governance mechanisms to ensure that AI-driven automation remains trustworthy, accountable, and aligned with societal values.

08

Conclusions



Conclusions

The arrival of AI agents represents the new frontier in automation. Agentic workflow design allows combining the power of language models with functionalities and tools that give rise to an increasingly diverse ecosystem that allows executing a variety of workflows with specific tasks and expanding application possibilities across all sectors and industries.

In the process of adoption, these systems face challenges. The probabilistic nature of the language models that generally underpin them can lead to unexpected, erroneous or biased results, which underlines the importance of human supervision at critical moments in the workflow to control partial or final system outputs. It is key to calibrate the intensity and mode of human intervention in crucial aspects such as information accuracy, error correction, alignment with ethical and regulatory principles, and bias mitigation.

This ties in with society's trust in AI agent systems for mass adoption. Traceability, explainability, accuracy, and safety are again at stake in the various applications and contexts in which they are deployed. Ultimately, the success of generative AI-based agents will depend on our ability to understand and manage their development in an ethical and responsible manner.

It is excellent news to reaffirm the need for human control under the logic of the person in the loop, with responsible and professional oversight. But this also confronts us with the challenge of reconfiguring our work teams, starting by identifying new skills to develop.

If generative AI initiated a paradigm shift in the way we work, combined AI agents accentuate it by modifying the nature of human intervention: instead of producing and performing multiple actions to execute tasks, we will progressively take on the role of supervising or editing complex results that emanate from agentic workflows.

AI agents will eventually become standard software, marking a shift in how we perceive and interact with artificial intelligence. However, what lies beyond agentic AI? The next frontier may involve self-improving agents capable of dynamically updating their own models without requiring human retraining, making them more adaptive and efficient over time. Another major evolution will be the integration of AI agents with the physical world, allowing them to move beyond digital tasks to interact with robotics, IoT systems, and autonomous infrastructures, effectively bridging the gap between digital intelligence and real-world execution. Furthermore, advancements in neurosymbolic AI and hybrid reasoning models will enable AI to combine deep learning with structured logic and symbolic reasoning, allowing it to go beyond statistical pattern recognition and achieve deeper forms of understanding and reasoning.

What cannot be denied is that the acceleration of innovation will intensify the massive adoption of agentic logic, especially with the evolution of the "no code" and "open source" paradigm, the daily improvement of the ecosystems that contain the major language models and the increasing ease of integrating them into systems and platforms.

These advancements could further redefine how AI systems interact with humans, businesses, and society, bringing us closer to a seamless integration of intelligent automation into daily life and complex decision-making processes.

Rather than viewing AI as a replacement for human effort, it should be understood as an extension of human intelligence—augmenting our ability to analyze, decide, and create with unprecedented efficiency and strategic insight. The co-evolution of human and artificial intelligence must be the guiding principle in shaping policies, workflows, and future developments.

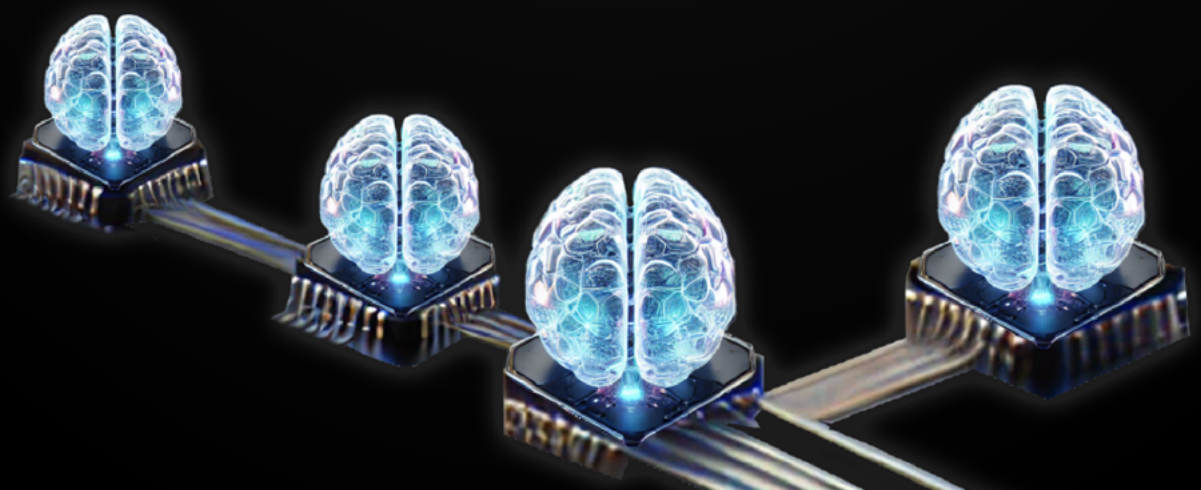
This perspective emphasizes collaboration rather than substitution, ensuring that AI enhances human agency rather than diminishing it. As AI agents become more autonomous and adaptable, fostering responsible innovation and ethical governance will be essential to maintaining human oversight while unlocking AI's full potential.

Ultimately, in this early 2025, AI agents are becoming the new hype of the AI world. But like the label "artificial intelligence" in the previous decade, "AI agents" will eventually become commonplace, shifting from a revolutionary concept to an expected standard. The current threshold of autonomy and adaptability will continuously evolve, redefining what we consider intelligent automation and pushing us toward new breakthroughs in human-AI collaboration.

09

Annex I

Practical example of a multi-agent system for project management



Practical example of a multi-agent system for project management

We take the Taskade framework as a reference to evaluate the capacity of artificial intelligence agents to manage projects in an automated way in a software development environment.

The main objective was to evaluate how multiple specialized agents can collaborate with each other, from the initial approach to a project to the completion of a detailed work plan, including the creation of backlogs and the assignment of roles. For this test, a real scenario based on the transcription of an initial meeting with a client was used.

Use Case Context

A development team held a kick-off meeting with a client to discuss the design and deployment of an application that included an interactive front-end and a machine learning model for demand prediction. The transcript of this meeting, approximately 45 minutes long, where the project objectives were discussed.

The transcript was processed by specialized agents to autonomously generate a detailed project plan.

Methodology

Five specialized AI agents were configured and trained in Taskade, each with instructions, context documents, and custom commands to tackle specific tasks in project management. These configurations were designed to optimize their performance in the following areas:

Project Planning Assistant

Trained to analyze documents such as meeting transcripts or initial briefings and identify key objectives, milestones, project dependencies, and generate a detailed roadmap. This agent was provided with planning templates as context and was configured with diagramming commands to efficiently structure schedule.

Agile Scrum Master

Configured to organize and prioritize tasks in a backlog based on agile methodologies. This agent includes predefined rules for assigning tasks based on workload, dependencies between activities, and deadlines. It is also configured with tools linked to Google Calendar to schedule meetings and sprint reviews.

Project Manager

Designed to be an expert in project management and acts as a mediator between other agents. Documents with project management guides and specific monitoring commands have been loaded.

Researcher

Trained to perform technical information searches on technologies, frameworks and solutions relevant to the project. Provided with a file containing the company's technology stack as context and configured with access to trusted website sources. Additionally, commands were established to synthesize findings in a structured format and to propose alternatives based on initial requirements.

SWOT Analysis Agent

Configured to assess the project using strategic analysis, identifying strengths, opportunities, weaknesses, and threats based on the client context and team capabilities. This agent uses uploaded SWOT analysis templates as context and rules to generate strategies that maximize benefits and mitigate potential risks.

Interaction between agents

The agents worked together in a chat to plan the project. The flow was as follows:

Initial analysis of the transcript:

The Project Planning Assistant processed the transcript to identify the main objectives:

- Develop an application with interactive front-end.
- Implement a machine learning model for demand prediction.
- Ensure the integration of the model with the backend hosted on the client's own servers.

From these objectives, an initial roadmap was generated with key milestones and dependencies.

Strategic analysis:

The SWOT Analysis Agent assessed the internal and external factors that could impact the project:

STRENGTHS	OPPORTUNITIES	WEAKNESSES	THREATS
Team experience in the technologies needed to develop the project	Leverage pre-trained models using machine learning techniques to accelerate development	Limited time for integration and testing	Possible compatibility issues with client servers

Based on this analysis, he proposed strategies to maximize strengths and minimize risks.

Preliminary investigation:

The Researcher explored technologies relevant to the project, generating the following findings:

- Pre-trained machine learning models available that could serve as a backbone for transfer learning for demand prediction.
- Recommended frameworks for the front-end (React, Vue).
- Best practices for deploying solutions on customer-premises servers.

These findings were integrated into the roadmap to guide the team's technical decisions.

Organization and sprint planning:

The Agile Scrum Master structured the backlog generated by the Project Planning Assistant into three initial sprints:

SPRINT 1	SPRINT 2	SPRINT 3
Setting up development environments. Setting up the repository on GitHub. Further research on server compatibility.	Development of the API for the machine learning model. Front-end interface design. Initial training and validation of the model.	Integration of the model with the backend. Front-end functional testing. Bug fixes and final adjustments.

Additionally, he scheduled weekly meetings to review progress and a final demo to the client.

Monitoring and adjustments:

- The Project Manager supervised the assigned tasks, adjusting priorities according to the client's needs.
- Monitored risks identified by the SWOT Analysis Agent and ensured resources were efficiently allocated.

Results

The roadmap, backlog, technology stack determination and role assignment were generated after several iterations in less than 30 minutes, significantly reducing the time required for initial planning.

Limitations detected

- **Limited collaboration between agents.** Although agents performed efficiently individually, there were redundancies in tasks, as agents appeared to work individually rather than together, making individual configuration of each agent seem unnecessary.
- **Errors in tasks.** Some resources were overallocated, the backlog was not entirely accurate, and time estimates were poor, requiring manual intervention to achieve the plan.

Conclusions and recommendations

Using AI agents to manage projects proved to be an effective tool for automating repetitive tasks and reducing planning times. However, testing revealed areas for improvement, such as the need for greater collaboration between agents and more flexible customization options to minimize manual rework of project planning.

This approach is particularly useful for small to medium-sized projects with well-defined requirements, while more complex projects may require additional human oversight to adjust for identified constraints.

Splitting into different agents is recommended for performing different processes, but not for different tasks in the same process. Since agents are powered by the GPT-40 language model, it may be advisable to use a single agent configured with all the context documents, commands, and tools needed to perform all tasks in a process. This makes it possible to opt for a cheaper plan, with almost no loss in capabilities.

Key results:

- **Efficient automation:** Agents were able to transform inputs such as meeting transcripts into detailed work plans.
- **Limited interaction between agents:** Redundancies and lack of communication between agents were identified.
- **Potential for improvement:** Effective integration and customization will be key to future advances in this type of tools.

Recommendations:

- Create an agent per process equipped with all the tools necessary to carry out a project.
- Configure the agent as much as possible so that human review is as small as possible.

In summary, the use of specialized agents in project planning has enormous potential, but requires optimization to meet the challenges of more complex and collaborative projects.

.UBAderecho

IALAB

CAF BANCO DE DESARROLLO DE AMÉRICA LATINA Y EL CARIBE

Supported by

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 FCEFyN

LA LEY

Thomson Reuters™