Agentic AI: 4 reasons why it's the next big thing in AI research

For the last few years generative AI or gen AI, has been the hot new thing among technologists, but a new term has crept into artificial intelligence (AI) development communities recently. "Agentic" is the latest buzzword in AI, and in this case, it's appropriate to believe the buzz. Agentic AI brings together the versatility and flexibility of large language models (LLMs) and the precision of traditional programming.

Agentic AI refers to a system or program that is capable of autonomously performing tasks on behalf of a user or another system by designing its workflow and using available tools. The system has "agency" to make decisions, take actions, solve complex problems and interact with external environments beyond the data upon which the system's machine learning (ML) models were trained.

AI agents draw not only from databases and networks, but can also learn from user behavior, improving over time. Agents' adaptability enables them to handle complex, multistep AI applications that traditional AI would not be able to handle, making them a key part of the modern organization's process automation strategy.

If you were to ask a general-purpose LLM chatbot such as ChatGPT what ice cream maker to buy, that model's natural language processing (NLP) capabilities would enable it to give recommendations based on its training data, which would probably include information scraped from the internet. But you don't want generic advice from the past, you need advice that considers real-time information.

An agentic AI platform consists of an LLM that orchestrates the behavior of multiple agents that can be deployed across various applications. These agents might be more AI models, or they might be simple search tools that can quickly look up information in a knowledge base or online. To continue the preceding example, imagine if an LLM such as GPT also had access to real-time e-commerce data and your payment details.

An agentic AI platform like this could not only tell you that people love a specific ice cream maker that is on sale at a specific retailer, but it could theoretically make the purchase on your behalf. Agentic AI gets us closer to the use cases that we, until recently, thought of as science fiction, where machines can complete complex tasks involving complex workflows, data-driven decision-making and action-taking with minimal human intervention.

There are good reasons to think that the hype around agentic AI is justified. Here are 4:

1. Both flexible and precise

LLMs excel at processing and generating human-like text, making it easier for users to interact with AI using natural language commands. This reduces the need for explicit programming knowledge. LLMs can generate responses or actions based on nuanced, context-dependent understanding, which is useful in scenarios where traditional programming might struggle to cover all edge cases. On top of this, LLMs are creative in tasks such as content generation, code completion summarization and more. This generative capability is hard to replicate with traditional rule-based programming.

Meanwhile, traditional programming is highly structured, deterministic and reliable, making it ideal for tasks that require precision, repeatability and verifiability. Traditional programming languages provide granular control over how tasks are executed, helping to ensure that complex workflows, algorithms or specific system requirements can be explicitly defined and optimized. Traditional programming is often more efficient for tasks that require high performance or unique functions.

Agentic AI systems provide the best of both worlds: using LLMs to handle tasks that benefit from flexibility and dynamic responses, while combining these AI capabilities with traditional programming for strict rules, logic and performance. This hybrid approach enables the AI to be both intuitive and precise. Agents can autonomously perform tasks while adapting to new data or dynamic environments, something that’s challenging for static code. At the same time, critical processes (such as security or calculations) can rely on deterministic, traditional algorithms.

An agentic AI-powered system might contain simple reflex agents that perform one simple task well and consistently. More complex rule-based agents can use current perception and draw on memory, enabling them to receive and store new information, allowing it to perform a wider range of tasks. Learning agents can also ingest new data, but can use this to inform later decisions, improving accuracy over time. A powerful agentic AI platform might involve dozens or even hundreds of agents of varying capabilities working together.

Your subscription will be delivered in English. You will find an unsubscribe link in every newsletter. You can manage your subscriptions or unsubscribe here. Refer to our IBM Privacy Statement for more information.

2. Extended reach

LLMs are typically trained on static data sets that represent a snapshot of information up to a certain point in time. LLMs can’t actively go out and gather new information from the web after their training period. They can only generate responses based on what they already "know." They can't access or update real-time data from external sources on their own.

LLMs cannot directly interact with external tools or data processing systems (such as spreadsheets, cloud platforms or analytics software) or autonomously set up systems to monitor and collect ongoing data (such as IoT sensors, business processes or logs from systems) because they’re not designed to perform continuous tasks.

Agentic AI can be designed to search the web, called application programming interfaces (APIs) or query databases. Agents can fetch real-time information, retrieve updates or pull specific data points that are critical for decision-making. Agents can initiate and manage tasks such as data logging, real-time monitoring and trend analysis. They can proactively track and collect data streams from IoT devices, social media feeds or other systems, providing LLMs with fresh inputs for more informed decision-making and contextual responses.

Agentic AI can use feedback loops where it actively seeks out new data to refine its models or decision-making processes. This might involve periodically querying new sources, collecting user feedback or analyzing real-world outcomes to update and improve its understanding or strategies. This way, the LLM can achieve optimization over time from richer, continually evolving data.

3. Autonomous

With the big brains of LLMs and the targeted capabilities of agents, agentic AI can operate independently and autonomously perform specific tasks without the need for constant human oversight. This enables continuous operation in environments where human supervision is limited or unnecessary. Autonomous systems can maintain long-term goals, manage multistep tasks and track progress over time.

For example, an agentic AI could be tasked with managing a marketing campaign, continuously monitoring performance, adjusting strategies and optimizing results based on feedback without the need for human input at every step.

In healthcare, agents can monitor patient data, adjust treatment recommendations based on new test results and provide real-time feedback to clinicians.

In cybersecurity, agents can continuously monitor network traffic, system logs and user behavior for anomalies that might indicate potential security threats, such as malware, phishing attacks or unauthorized access attempts. In supply chains, AI can autonomously place orders with suppliers or adjust production schedules to maintain optimal inventory levels.

In human resources, agents can analyze the role and background of the new hire to create personalized onboarding training paths. It can adjust content and learning materials based on the individual's prior experience, role requirements and learning pace.

4. Intuitive

One can imagine many business functions currently performed with software as a service (SaaS) products being replaced or augmented by agentic systems, which enable workers to interact with data and perform tasks more efficiently with natural language inputs and simplified user interfaces.

For example, imagine a ticketing system that software developers use to track the progress of projects. Such a system requires many tables, tabs and workflows that aren't always easy to understand at first glance. To find out useful information, users need to hunt for the right data, navigating a complex array of menus to get the information they need. Then, they might need to use that information to create a presentation.

What if, instead of arraying all of that data in tables and tabs, the user only had to ask for the information he or she needs in plain human language?

For example, imagine generating a presentation slide that displays 5 bar graphs representing every completed ticket per employee for the current month, going back 5 years, all without manually sorting through complex data sets.

It might have taken half an hour to fetch that data manually and another half an hour to display it in a tidy format for a slick presentation, but agents could pull this together in seconds.

For organizations struggling to see the benefits of gen AI, agents might be the key to finding tangible business value. Monolithic LLMs are impressive but they have limited use cases in the realm of enterprise AI. It remains to be seen whether the vast sums of money currently being poured into a handful of huge LLMs will be recouped in real-world use cases, but agentic AI represents a promising framework that brings LLMs into the real world, pointing the way to a more AI-powered future.