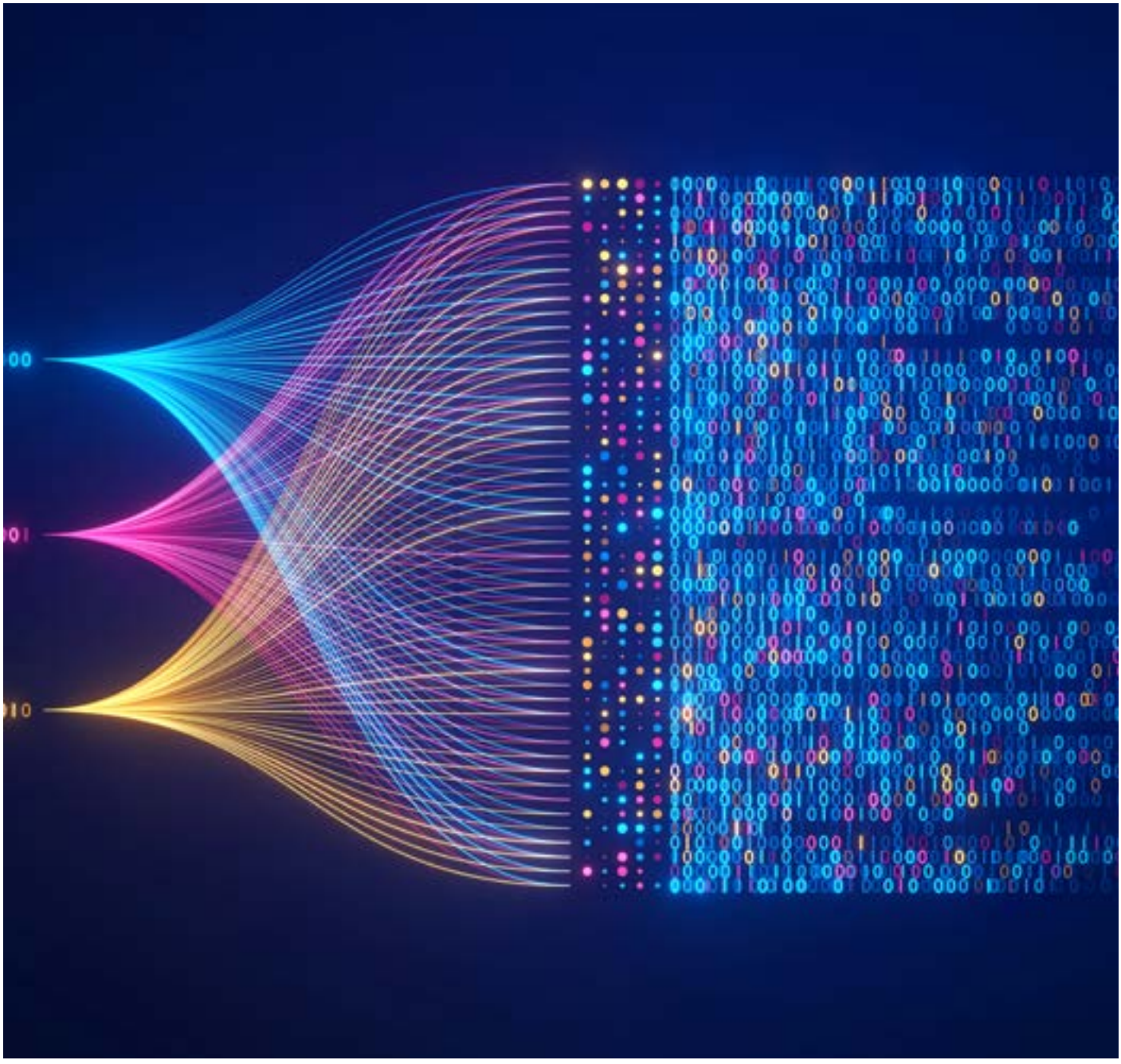
Autonomous enterprises and beyond

The revolution of autonomous agents in business

EVIDEN



CONTENT

Looking into a crystal ball

Imagine you are willing to launch a new product. However, instead of asking your R&D department to conduct a market survey and design the most suitable product, requiring your factory to schedule supply, production, and delivery plans, and urging your marketing team to create a sales website and launch an online ad campaign, you simply enter a single prompt.

This prompt activates your AI system, which automatically decomposes, or simply breaks down all these processes into subtasks. It then autonomously submits these subtasks to the best AI applications within your ecosystem and returns your offering, ready to deploy and launch.

Sci-Fi? Not so fast.

Early experiments are already underway that come close to turning this dream into a reality. Solutions such as BabyGPT, AutoGen, and CrewAI are emerging, providing basic functionalities for this purpose. Our specialists estimate that such technologies could fully mature within one to three years, paving the way for what analysts may describe as autonomous business.

What is the path forward? How can you foresee, adapt, and thrive in this revolution? This paper offers a glimpse into the upcoming trends and, most importantly, the best practices to benefit from them, rather than being disrupted.

Autonomous agents: The next big thing in business

For two years, large language models (LLMs) have demonstrated incredible power in analyzing data to generate text, sounds, images, videos, designs and even strategies. This has not only enabled revolutionary conversational user assistance with tools like ChatGPT, but also spurred advances in content creation, software development, and product ideation.

It quickly became apparent that these models could be leveraged to create powerful assistant applications to aid our daily work, as seen with Microsoft CoPilot, Google Gemini, Amazon Q, and similar tools. Additionally, it soon emerged that they could also directly generate prompts (instructions given to the AI), call functions, and (with some AI reasoning capabilities) could address complex queries.

All this has given rise to a new concept: Autonomous Agents enhanced by LLMs, capable of planning, orchestrating, and executing complex actions, making decisions, and acting without human intervention.

Language: The next universal lingua franca?

Now, this latest development has introduced a revolutionary perspective:

With the ability to decompose a problem (presented in natural language) into questions whose answers can be found in a database or documentary corpus, doesn't this situation make text the lingua franca for interacting with anything that can be AI-powered, including computers, and for requesting anything? (This is assuming request results can be generated digitally first, and perhaps physically tomorrow with the likely commoditization of 3D printing.)

Already, most enterprise software publishers, such as Salesforce, SAP, ServiceNow, and Microsoft, have begun to offer interfaces that allow for querying their systems in natural language. Concurrently, it is becoming increasingly easy to develop agents to access other enterprise data and to facilitate their cooperation.

Autonomous agents are therefore on the verge of profoundly transforming the business landscape, connecting to business applications like ERP, CRM and HRM, and to the underlying IT infrastructure made up of databases, CMS, HRM, business code, and external resources too, like the APIs and online data sources. All of them can continuously exchange information among themselves and with users to enable a new level of automation within the enterprise, within the extended enterprise through APIs, and across the entire networked world.

Autonomous agents through a technical lens

The most widespread framework for autonomous agents, notably implemented by AWS (Agents for BedRock), Microsoft (Semantic Kernel Planners), Salesforce (Einstein Copilot), or LangChain, is inspired by the scientific paper, ReAct: Synergizing Reasoning and Acting in Language Models.

In brief, a ReAct autonomous agent has access to tools, services or applications to gather information or perform tasks in response to a request. **For example, if a user asks an autonomous agent to prepare and coordinate maintenance operations for industrial equipment, the autonomous agent could plan the use of the following tools:**

Scheduling	Content Management System	ERP	Databases	Code Interpreter
A scheduling system to provide maintenance tasks assigned to technicians	A content management system to obtain information on procedures for tasks such as the required tooling, spare parts and other prerequisites	An ERP to get details on the availability and location of spare parts	Databases to extract the history of events, maintenance activities, sensor data, and more	Possibly a code interpreter to run code generated by the agent, for example, to create diagrams

Depending on the use case, the tools may also include information extracted from specialized data sources, simulators and other scientific codes, data generators, programs, scripts for performing specific actions or calculations, machine learning models or algorithms, and more.

More complex autonomous agents can act as tools for other agents, for example, to automate a process using a Robotic Process Automation (RPA) solution, or by analyzing screenshots and acting like a human. Many applications powered by LLMs can, in fact, be used as tools orchestrated by autonomous agents.

Agents can collaborate directly with each other and even request for (and get) direct assistance from human experts, if needed.

For example, consider a multi-agent system whose goal is to create software products:

One agent takes a line requirement as input and generates user stories, competitive analyses, requirements, data structures, APIs and documents.	Agents have different roles — some are architects, some are developers, and others are coders.	Each agent executes an operating procedure designed by humans, coded as instructions in the prompts provided to the LLM.	The entire process of a software company is provided with carefully orchestrated operating procedures.

More generally, autonomous agents can rethink and automate entire workflows with digital tools and LLMs to detect and act on their environment. Compared to standalone LLMs or traditional RPA, autonomous agents can directly control other enterprise systems and are not limited by predefined rules. This can fundamentally change how a business operates, enabling it to deploy automation on a larger scale.

The future is just a few prompts away

This statement holds multiple promises and opportunities.

After introducing robotics in factories and RPA in back offices, are we on the verge of reinventing all enterprise processes and moving closer to an almost people-less enterprise, where the workforce constitutes robots and digital agents?

Indeed, autonomous agents can provide versatile, intelligent assistance to users, enabling them to perform complex tasks or tasks that may require the combination of information from disparate, multiple systems and applications that are not interconnected.

At the enterprise level, they will initially direct user queries to specialized GenAI applications, gradually enabling an increased level of automation of various processes. But we are only at the beginning of this revolution. The path to autonomous business may be closer than we think.

Of course, human supervisors will still be necessary. However, if they are to remain competitive, enterprises and organizations will need to leverage these technologies before their competitors do. Organizations that fail to adapt may risk falling behind their competitors who leverage powerful ecosystems of autonomous agents that can operate continuously, at high speeds, and with near-zero marginal costs.

As these technologies proliferate, we must rapidly begin to address their societal implications. Given the rapid pace of innovation, policymakers, industry leaders, and the public should engage in open discussions immediately to develop strategies that will minimize potential social drawbacks.

Six steps closer to success

How can an organization exploit the full potential of this revolution?

Autonomous agents offer multiple promises to enterprises, but also present some challenges, namely ensuring that these agents are part of a long-term strategy for integrating AI within the company and that the risks and costs associated with this evolution are managed effectively.

Indeed, the true value of autonomous agents will lie in their ability to address complex requests involving multiple agents. This requires considering several technical and operational factors. Let's explore some of these key factors:

Technical Interoperability: Although autonomous agents primarily exchange text (prompts and responses), they need to be able to communicate with a variety of tools and services using standardized protocols and interfaces, ensuring strict access controls. They may also need to efficiently exchange datasets, images, and other non-textual content. Agent design, therefore, requires developers with strong technical proficiency.

Semantic Interoperability: Agents must respond to prompts from the user or other agents in a precise, accurate, and contextually appropriate manner, relying on ontologies or taxonomies to structure and interconnect the view of concepts, entities, and relationships within a specific domain. This holistic approach must be considered from the start.

Security: Autonomous agents could be used for unethical purposes too. Moreover, it's crucial to secure autonomous agents against attacks such as prompt injection. If the regulations are not in tandem with technology, self-imposed safeguards are necessary to ensure appropriate and safe use.

Code Integrity: The ability of autonomous agents to generate and execute code opens up an almost infinite realm of possibilities. However, this also raises potential risks related to security and compliance, as they could introduce vulnerabilities or non-compliant behaviors that might negatively impact the system or user. Agents have to be secured by design.

Continuous Learning: The integration of autonomous agents with various tools and services is an ongoing and evolving process. Continuous learning and improvement through user feedback, data analysis, and experimentation are necessary to optimize and refine their integration strategies and approaches.

Cost Reduction: Autonomous agents can incur significant costs due to the high volume of requests made to LLMs. It is essential for organizations to mitigate these costs through techniques such as advanced caching, LLM optimization, memorization in knowledge graphs, and continuous learning, among others.

Gear up to be the disruptor, not the disrupted

How can you leverage these new opportunities while coping with the new challenges?

To succeed, enterprises need to embrace an emerging discipline: autonomous agents engineering. This involves the secure development and integration of agents into existing business and IT systems, allowing organizations to avail of AI benefits while minimizing risks.

This new discipline aligns with the growing trend of adopting a more industrial approach to deploying applications using generative AI (GenAI), focusing on lifecycle control, security, cost management, and reducing risks, among others.

This trend is exemplified by the growing use of specialized LLMs for tasks like document querying, API calls, code generation, and SQL queries. Since almost every agent shares the same interface (receiving a prompt and returning text), autonomous agents will act as a conductor, directing a user's request to the agent most likely to fulfill it, regardless of where it is executed or how it was developed.

To address the challenges associated with integrating autonomous agents, it is crucial to implement appropriate strategies. In line with its aim to expand possibilities, Eviden's Generative AI acceleration program proposes strategies for optimization, selection, and validation of LLMs, management of technical and semantic interoperability, or risk mitigation. These are specifically designed to tackle the challenge of integrating autonomous agents and extend those we excel in as systems integrators.

Do you want to experience the benefits of autonomous agents and enhance operations and productivity, integrate technologies into your systems effectively and securely, and address technical and operational challenges to maximize your potential today? Contact us to request a discussion with an expert, conduct a workshop, or launch a pilot in your organization!

>> Connect with me to discuss more about the plethora of opportunities presented by Autonomous Agents. **https://www.linkedin.com/in/thierrycaminel/**

>> Read more about Eviden's GenAI solutions: **https://eviden.com/solutions/generative-ai/.**

ABOUT THE AUTHOR

Thierry Caminel is the CTO for AI and Decarbonization at Eviden. He brings 30 years of experience, including early work in symbolic AI and rule engines. After roles in aerospace and IoT during the “AI winter,” he rejoined the AI field at Atos 13 years ago, focusing on Semantic Web, Knowledge Graphs, Big Data, and Machine Learning. He has led an innovation lab and, as a distinguished expert, headed a large community on AI for 6 years. He has supported numerous projects globally and helped build strategic data platform offerings.

Now, as the CTO for AI and Decarbonization at Eviden, he contributes to the organizational roadmap around AI and Automation, and actively utilizes cutting-edge generative AI for customer projects.

Apart from work, Thierry is interested in sustainability, sciences, and technologies.

ABOUT EVIDEN

Eviden is a next-gen technology leader in data-driven, trusted and sustainable digital transformation with a strong portfolio of patented technologies. With worldwide leading positions in advanced computing, security, AI, cloud and digital platforms, it provides deep expertise for all industries in more than 45 countries. Bringing together 47,000 world-class talents, Eviden expands the possibilities of data and technology across the digital continuum, now and for generations to come. Eviden is an Atos Group company with an annual revenue of € 5.1 billion.

Since 2023, Eviden has launched an ambitious Generative AI Acceleration program to help businesses and organizations fully exploit, scale, and leverage the transformative power of Generative AI (Gen AI) with complete trust. Providing End-to-End Generative AI Consulting, fast-to-value solutions, and a modular set of Accelerators, the initiative allows organizations to move beyond the hype and leverage Gen AI for tangible business value, competitive advantage, and innovation across all aspects of their operations.

The program is supported by a robust ecosystem of partnerships with hyperscalers, AI Independent Software Vendors (ISVs), and high-performance processing providers. Notably, Eviden is building the most powerful AI supercomputer in the world and has been ranked as a GenAI service leader by the HSF analyst firm.

More: https://eviden.com/solutions/generative-ai/

Connect with us

eviden.com