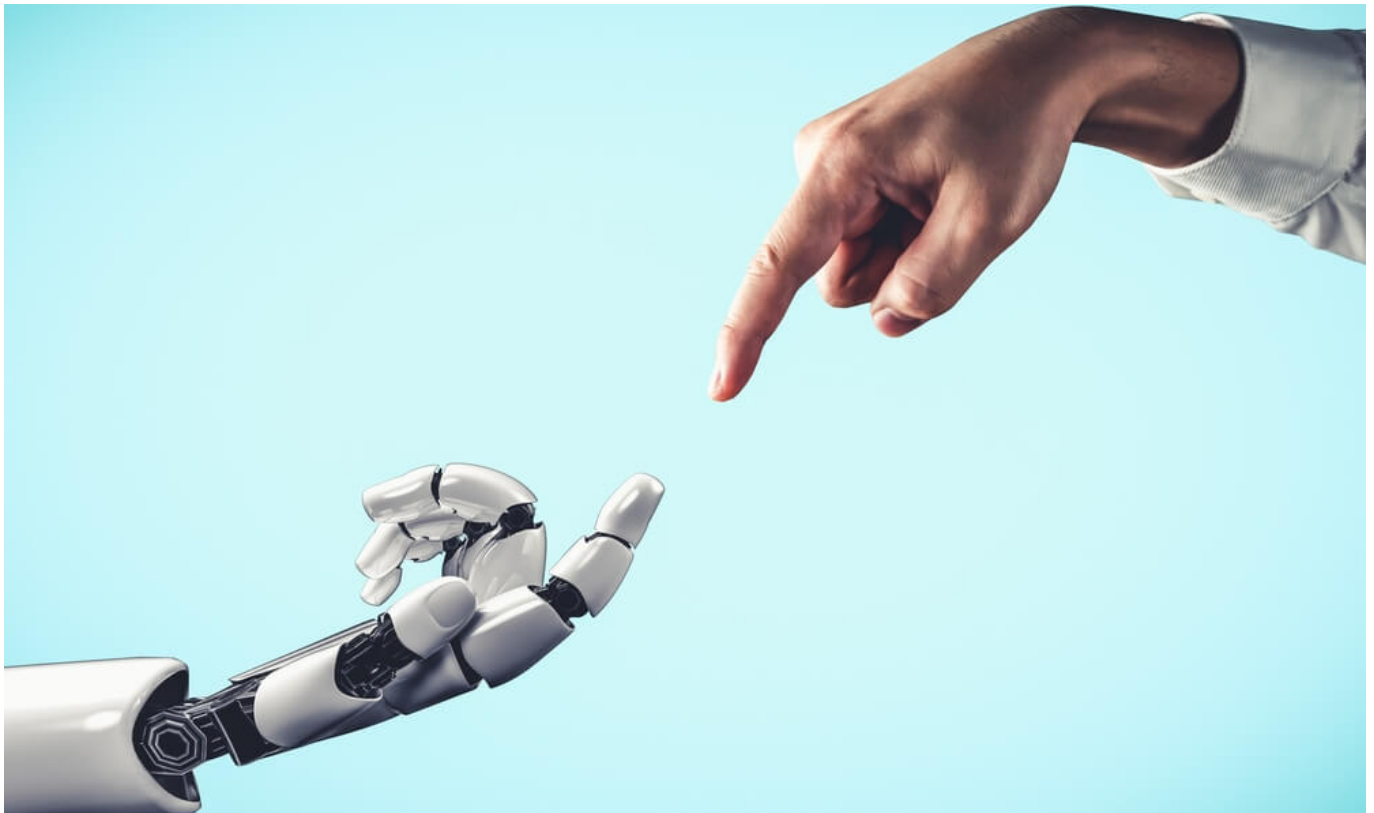




By: *Elise Quevedo*

Are we prepared to entrust machines with unsupervised control?



Two months into 2025, we are having conversations about Artificial Intelligence (AI) moving at lightning speed. Today, I want to discuss the emergence of autonomous AI agents, systems designed to operate without human intervention.

But the main question is, are we prepared to entrust machines with unsupervised control? In this article, I take a deeper dive to get a better understanding of this topic.

The Promise of Autonomous AI Agents

Mirror, mirror on the wall, who is the smartest of them all? Autonomous AI agents are built to independently perform tasks, make decisions, and adapt to new information, but sometimes, it feels like we are reading a Once Upon a Time storybook or waiting for a magic trick to be revealed.

AI agents' potential applications span various sectors, including logistics, customer service, finance, and healthcare. There is a large appeal with the potential to enhance efficiency, reduce operational costs, and tackle complex problems beyond human capability. AI agents aim to handle tasks that traditionally require much human oversight.

For example, consider the strides made by the less-known Gatik, a **company** specialising in autonomous delivery vehicles in North America.

Gatik shows that autonomous logistics can work, but human control is still necessary

Gatik implements self-driving trucks that travel predetermined routes without the need for human drivers in collaboration with big merchants like Walmart, Pitney Bowes, and Loblaw.

Unlike other autonomous vehicle companies that focus on large-scale highway automation,

Gatik targets the "middle-mile" logistics problem, transporting goods from warehouses to retail locations, reducing human labour costs, and increasing efficiency.

But even though these vehicles are autonomous, don't expect to see bumper-to-bumper traffic filled with unattended vehicles. Human intervention is still needed for unexpected scenarios such as road closures or extreme weather conditions. Gatik shows that autonomous logistics can work, but human control is still necessary because the technology is still a work in progress.

The Reality Check: Challenges and Limitations

Honesty time: despite these advancements, the journey toward fully autonomous AI is filled with challenges. I personally think it is great that we are moving toward a future we only saw in movies in our childhoods.

But we also need to be realistic about the present limitations and the work that is yet to be done.

A big concern here is what is known as "hallucination," where AI models generate outputs that are inaccurate or misaligned with reality. This issue underscores the necessity for human oversight to validate and correct AI-generated content.

So, for those worried that AI and robotics are going to take over the world, we are far from an AI-run utopia.

11x, a London-based **startup** that moved its operations to Silicon Valley, has developed digital workers like "Alice," an AI agent designed to handle outbound sales.

Alice is trained to analyse customer data, craft personalised messages, and engage in lead conversion at a scale far beyond what a human sales team could achieve.

While Alice has shown promise by

outperforming human benchmarks in lead conversion, instances of hallucination have been observed. Alice might, for example, produce excessively tailored communications that misjudge a client's requirements or preferences.

Integrating AI agents into business operations also brings up ethical concerns

Examples where the AI has falsified facts to strengthen its sales argument highlight the limitations of unsupervised AI agents in high-stakes business settings.

So remember, even if businesses like 11x improve these models, these challenges show how important human intervention is and will be in the near future.

Integrating AI agents into business operations also brings up ethical concerns. Bias in AI decision-making is still an issue. AI systems can reinforce societal inequalities by making biased hiring decisions, unfair lending assessments, or discriminatory content recommendations.

So ensuring fairness and transparency in AI-driven decisions will require continuous monitoring and refinements of the algorithms that power these new AI agents.

Tech Giants Leading The Race

A number of tech giants are making significant investments in autonomous AI agents, with differing degrees of success. Their methods and difficulties offer insightful information on the larger field of artificial intelligence.

Google DeepMind

Google's DeepMind division has **pioneered** AI research, pushing the boundaries of what machines can accomplish. By now, everyone with a digital device has heard of Gemini, an AI

system capable of generating images and audio.

Google aims to integrate advanced AI into autonomous agents as it envisions a future where AI systems can autonomously assist with creative tasks, such as designing marketing campaigns or composing music.



Google aims to integrate advanced AI into autonomous agents as it envisions a future where AI systems can autonomously assist with creative tasks

Because multimodal AI systems are complicated, maintaining accuracy and dependability is a constant problem. Questions over whether even the most sophisticated AI models are ready for complete autonomy arise because they occasionally produce offensive or deceptive content. Nonetheless, it is an exciting road we are all watching.

OpenAI

A few weeks ago, AI Agent Operator was **released**; it is intended to manage web tasks like online shopping and reservation services independently. Think of Operator as a virtual assistant; you give it a task, and it will do it.

It demonstrates OpenAI's dedication to useful AI applications. Operators are able to make bookings and finish online purchases, illustrating how AI might simplify routine consumer tasks. There is a but coming, OpenAI has recognized that fully autonomous operation is still a distant goal.

Operator still requires user intervention for

critical actions like logins and payments, indicating that while AI can assist in decision-making, true autonomy has yet to be achieved. This underscores the reality that AI agents remain copilots rather than independent operators.

Microsoft

CEO Satya Nadella envisions AI transforming knowledge work by integrating AI agents with human workflows. I like that Microsoft has taken a different approach from some of its competitors by emphasising AI as a collaborative tool rather than a replacement for human workers.

The company has embedded AI agents within **tools** such as Microsoft 365 Copilot, which help users draft emails, summarise meetings, and organise tasks.

This approach acknowledges that AI agents will require supervision and human oversight to function effectively, reinforcing the idea that AI should augment human work rather than replace it entirely. In my opinion, this is the approach we should use moving forward.

Fetch.AI

This ambitious company's focus is on creating autonomous AI agents for various applications, including the launch of an NFT machine-learning platform for AI-generated art.

Fetch.AI is at the forefront of **exploring** how AI can function in decentralised environments, allowing autonomous agents to interact and trade digital assets without direct human intervention.



Fetch.AI is at the forefront of exploring how AI can function in decentralised environments

Now the but, the practical deployment of fully autonomous agents remains once again a work in progress, with challenges in achieving complete operational independence. The primary concerns are ensuring security and preventing malicious use of AI-generated content.

Moving Forward: Collaboration Between Humans and AI

Even while autonomous AI agents have advanced significantly, they are still not yet prepared to function independently of human oversight. Collaboration is required to address problems, including hallucinations, ethical issues, and the requirement for contextual understanding.

It is more than clear that although they are effective tools, AI agents are not perfect. Adopting an AI-human symbiotic partnership is the best course of action moving forward.

AI agents can handle repetitive and data-intensive tasks, which in turn frees humans to focus on strategic decision-making and creative problem-solving. So instead of viewing AI as a threat to human jobs, we should see it as an enabler that allows us to work smarter, not harder.

As we enter this new era in AI automation, and while the potential benefits are immense, the journey toward fully autonomous AI is a complex and ongoing process

In essence, as we enter this new era in AI automation, and while the potential benefits are immense, the journey toward fully autonomous AI is a complex and ongoing process. Who would have thought we would be talking about the collaboration between humans and machines with such intensity?

As we navigate the challenges and unlock AI's transformative potential in a manner that aligns with our values and societal needs, I will continue to explore, discover, and share my findings so that we can all choose what to embrace moving forward and what we leave behind.

As I always say, together we are stronger.