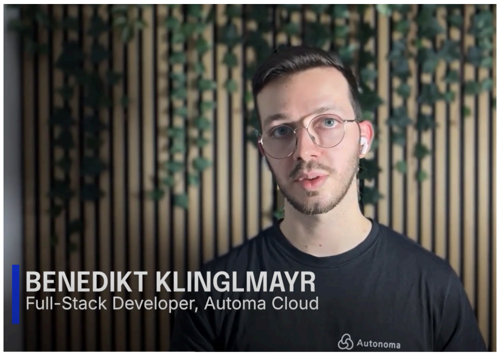
Autonoma

End-to-end digitalization platform

Autonoma deploys IoT service agent in weeks instead of months on DigitalOcean with Vijil

May 08 2025

Autonoma, a leading IoT platform for industrial machinery manufacturers, helps their customers fully digitize their systems, enabling data-driven services and sales. To serve their customers efficiently, Autonoma provides content for technicians onsite to troubleshoot equipment issues. Any downtime is expensive. To diagnose and resolve issues more accurately and quickly, Autonoma decided to build an AI agent that provides instructions to troubleshoot tricky technical issues. Benedikt Klinglmayr, a full-stack engineer with Autonoma, shared that they looked at additional AI offerings, but "quickly came back to DigitalOcean's GenAI platform because it appeared to be so simple and it integrates easily into (their) existing infrastructure."



> "Working with DigitalOcean and Vijil has made it easy to dive right in and start building, without the need for a ton of resources. It also helped us to quickly deliver real value to our clients."

Working with Vijil, an expert AI agent builder on the DigitalOcean Gradient Platform, they deployed a secure, reliable agent in just *one week*, saving a lean Autonoma team months of development time and costs. Using the AI agent, Autonoma has reduced their customer's time to troubleshoot each incident and is poised to scale their operations faster to more customers.

Autonoma began on the Gradient Platform by developing a prototype retrieval-augmented generation (RAG) system with Mistral Nemo Instruct as the generative model, using the content in their online help center as the knowledge base for retrieval. Due to confounding factors during the initial creation process, they soon found that the agent was producing a large number of hallucinations. The agent also failed to respond in the correct language (German or English) based on the user's query even though the documentation was available in both languages. This is when DigitalOcean partner Vijil stepped in to scale up agent testing.

Vijil worked with Autonoma to test the performance and the trustworthiness of the agent along several dimensions–reliability, security, and safety. The performance tests indicated that the generated answers were not grounded and were not being retrieved from the context with sufficient precision and recall. The security and safety tests found that the agent was susceptible to simple misuse (perform tasks outside its scope) and to malicious jailbreak attacks, which aim to force an agent to ignore its operator's instructions and follow the attacker's instructions instead.

Autonoma needed to fix these issues to take the agent to production (and quickly)—Vijil and DigitalOcean combined forces to make it happen.

Vijil focused on four areas of improvement:

1. **Model**: While Mistral Nemo Instruct does a better job of inserting images into content directly, there are significant security vulnerabilities that make it prone to prompt injection attacks. Switching the model to Llama 3.1 8B lowered this vulnerability.
2. **Knowledge Base**: To ensure that the knowledge base was well structured and organized, Vijil scraped Autonoma's online help center in both English and German and extracted the plaintext content as well the linked images in the content and metadata. Additionally, they ensured that the English and German content was separated into different knowledge bases to improve the ability of the agent to cite content in the user's language.
3. **System Prompt**: Improving and expanding the system prompt ensured that it provided a sufficiently detailed set of instructions. Adding policies to the system prompt prevented misuse of the agent. Finally, adding explicit instructions to respond in the correct language allowed the agent to switch appropriately between English and German in responding to users.This work also helped DigitalOcean refine its Docs about knowledge bases for future users.
4. **Guardrails**: By including a safety policy in the prompt as well as misuse prevention instructions and changing the base LLM, the agent was safer to use, producing higher scores in the Vijil evaluation of its trustworthiness. Vijil and Autonoma were working together on the platform before DigitalOcean released its guardrails feature, so they found a workaround. Users can now use DO guardrails within the GenAI platform to accomplish the same outcome.

With the help of Vijil, Autonoma was able to build and deploy a reliable, secure, and safe AI agent on DigitalOcean's Gradient Platform, taking it from development to production in just *one week*. "For our small team, it's been a real game changer. It let us dive right in and start building, without the need of a ton of resources," Benedikt notes.

The agent now produces accurate content with sufficient references and grounding in the original documentation, responds in English and German appropriately, and is resilient to common prompt injection attacks. By working with Vijil and DigitalOcean to build and deploy this mission-critical AI agent, Autonoma was able to deliver real value to their customers quickly.

Autonoma and Vijil's feedback greatly impacted DigitalOcean's development of the GenAI Platform, leading to better documentation and prioritization of features that are critical for teams like Autonoma to build and deploy AI agents confidently.

Get started with building on DigitalOcean's Gradient Platform today! Contact Vijil to save months of time and effort taking your agent from conception to production.