

**New York, USA-based company, provides a single platform to develop, evaluate, and observe AI agents – whether you're just getting started or scaling agents across your enterprise.**

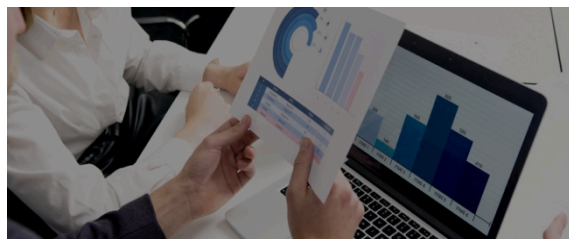
### Modern AI Observability and Evaluation

#### Engagement Overview

Crownstack, a leading technology solutions provider, engaged with Honeyhive in 2025 as its Staff Augmentation Partner to offer quality assurance and DevOps services, ensuring a better user experience for end-users and increasing feature release velocity by completing all quality-related checks on time.

#### About the Customer

Honeyhive, founded in 2022, is headquartered in New York City with operations in San Francisco. HoneyHive is the leading AI observability and evals platform, trusted by next-gen AI startups to Fortune 100 enterprises. They make it easy and repeatable for modern AI teams to debug, evaluate, and monitor AI agents, and deploy them to production with confidence. HoneyHive's founding team brings expertise in AI and infrastructure from Microsoft, OpenAI, Amazon, Amplitude, New Relic, and Sisu.



Industry	Software Development
----------	----------------------

#### Business Challenges

- **Quality Assurance:**
  - The customer aimed to enhance the overall quality of the HoneyHive.ai platform through both manual and automated testing methods. Ensuring complete coverage in test cases and regression test suites was essential for improving platform reliability and minimising recurring issues in AI workflows.
  - Performance optimisation was necessary to manage complex AI workloads effectively, facilitate integrations, and maintain efficient scalability.
  - Additionally, the customer was seeking a QA partner who could not only enhance software quality but also help establish robust testing practices that align with AI-driven development and team workflows.

- **Devops:**
  - The customer needed a scalable and secure infrastructure to support the growing complexity of AI workflows on the HoneyHive.ai platform. Ensuring compliance with industry standards such as HIPAA, ISO 27001, and SOC 2 was critical for operating in regulated environments.
  - Manual infrastructure provisioning processes posed risks of inconsistency, delays, and misconfiguration, impacting development velocity and system reliability.
  - The platform required a robust CI/CD strategy to streamline deployments, reduce time-to-release, and support frequent iteration in agile development cycles.
  - Additionally, the customer sought a DevOps solution that could seamlessly integrate with testing workflows, support AI-specific workloads, and provide observability across distributed environments.

### Business Solution

Our team of QA and DevOps collaborated with the Honeyhive engineering team to establish both a manual and an automated test suite, as well as support their infrastructure.

- **Quality Assurance:**
  - We developed a comprehensive manual test plan and carefully documented all test cases to cover the essential AI workflows, platform features, and user interactions across [HoneyHive.ai](#).
  - Additionally, we implemented a robust automated test suite using the Playwright automation tool, which enables reliable validation of AI pipelines, model integrations, and overall platform functionality, thereby improving efficiency.
  - Our QA team actively participated in the feature planning phase, providing early feedback on testing requirements, edge cases, and AI-specific validation needs. This involvement significantly enhanced the quality of sprint planning and execution.
- **Devops:**
  - Created Terraform modules to provision HIPAA, ISO 27001, and SOC 2-compliant VPCs. Ensured compliance-ready infrastructure with built-in logging, encryption, and access control.
  - Set up Kubernetes-based environments in AWS for deploying and testing the data plane.
  - Designed CI/CD pipelines using GitHub Actions for Infrastructure provisioning and Kubernetes deployments

### Technology Stack

Testing	Manual Testing, Automation Testing using Playwright
Devops	Terraform, GitHub Actions, Kubernetes, AWS
Workflow	Linear, Slack, Github

## Business Outcome

- **Quality Assurance:**

- We significantly improved the end-user experience on HoneyHive.ai by implementing rigorous testing protocols across our AI workflows and integrations.
- Regular regression testing allowed us to detect and resolve potential issues early, resulting in a more stable and reliable platform performance.
- The engineering and QA teams collaborated effectively in remote and agile environments, which fostered strong communication and alignment on objectives related to AI-driven development cycles.
- As a direct result of this integrated approach, we were able to shorten delivery timelines, leading to quicker releases of new AI features and more efficient progress toward project milestones.

- **Devops:**

- We significantly improved platform scalability, deployment speed, and compliance at HoneyHive.ai by implementing automation-first DevOps practices across infrastructure and CI/CD workflows.
- Infrastructure as Code (IaC) using Terraform enabled rapid provisioning of secure, compliance-aligned environments, reducing manual effort and human error.
- GitHub Actions-based CI/CD pipelines streamlined build, test, and deployment cycles, accelerating feature delivery while ensuring consistency across staging and production.
- By standardising Kubernetes environments for AI data plane testing, we empowered teams to validate AI features in production-like conditions.
- As a direct result, we achieved faster deployment cycles, reduced operational overhead, and improved system reliability—enabling quicker and more secure delivery of AI innovations.

## Further Links

[Learn More About Crownstack's Offerings](#)

[Learn More about Honeyhive](#)