

# Codeium Enterprise on VMware Private AI Foundation with NVIDIA

## Improve developer productivity using AI without compromising privacy and security

Why invest in an AI-powered code assistant?



**Increased productivity**



**Quicker onboarding**



**Freed-up time for senior engineers**



**Happier workforce**

Generative AI has fantastic potential today. One of the top benefits enterprises are excited about is coding assistants. However, privacy and security of their data remains one of their top concerns. This is why deploying Codeium Enterprise on VMware Private AI Foundation with NVIDIA drives the next-generation software to accelerate software development. Enterprises will deploy private and secure AI models, while still getting access to the entire variety of Integrated Development Environments and programming languages.

This joint solution enables personalization of the system on company's private data or proprietary source code and provides results that utilize internal best practices, utility libraries or other relevant syntax and semantics. This increases value over SaaS solutions whose promises of zero-data retention lead to minimal personalization potential.

### Codeium Overview

Codeium is a powerful AI code assistant that helps accelerate time to delivery of products & services by helping with code generation, debugging, testing, modernization, and more. Codeium can immediately accelerate time to delivery for OKRs and end business impact. With code suggestions, IDE-integrated chat, and inline refactoring capabilities, Codeium has a wide set of functionalities, available on the largest set of IDEs (including Visual Studio Code, JetBrains, Visual Studio, Eclipse, XCode, Jupyter, Vim/Neovim, and more) and supporting the most programming languages (over 70). Codeium provides detailed analytics dashboards for each organization to track the immediate impact of Codeium to their software development. Customers consistently report a 15-25% of decrease in pull request cycle time, a meaningful representation of developer and team productivity.

Codeium is able to set your technical organization up for scale by establishing standards, reducing tech debt, and accelerating modernization through its context awareness system that can reason over entire (and even multiple) repositories as well as custom organizational guidelines to encode best practices of your particular software development. This context aware engine increases the value of the coding assistant by over 35% towards code

Used by 700k+  
active developers

45%

of new committed code is  
written by Codeium

25 hours

saved per month for top 10%  
of users

1000+

enterprise customers

## Customer Stories:

17%

decrease in PR cycle time

31%

improvement in debugging  
time

40%

decrease in time spent writing  
unit tests

\* Data sources <sup>2</sup>

generation, refactoring, code explanation, and test generation.<sup>1</sup> By providing organizations with context-aware AI, Codeium customers have observed 4x faster onboarding times and happier workforces.

Codeium does not want its enterprise customers to compromise on their enterprise standards and helps mitigate operational risk in AI strategy by ensuring security via its on-premises deployment, no training on user code or non-permissively licensed code, providing attribution & audit logging, and more. Having scaled private deployments to tens of thousands of developers within individual organizations, Codeium provides white glove service in change management and AI adoption, all while never forcing vendor lock-in to other parts of the customer's tech stack.

Codeium provides meaningful value to individual developers and software development organizations. It is often the first generative AI tool adopted by organizations due to the clear, quantifiable ROI.

## VMware Private AI Foundation with NVIDIA

VMware Private AI Foundation with NVIDIA is a joint GenAI platform that is built and run on the industry-leading private cloud platform, [VMware Cloud Foundation™ \(VCF\)](#). This platform enables enterprises to run RAG workflows, fine-tune and customize LLM models, and run inference workloads in their data centers, addressing privacy, choice, cost, performance, and compliance concerns. It simplifies GenAI deployments for enterprises by offering an intuitive automation tool, deep learning VM images, vector database, and GPU monitoring capabilities. Its key benefits include:

- Enable privacy, security, and compliance of AI models
- Simplify GenAI deployment and optimize costs
- Get accelerated performance for GenAI models regardless of the chosen LLMs

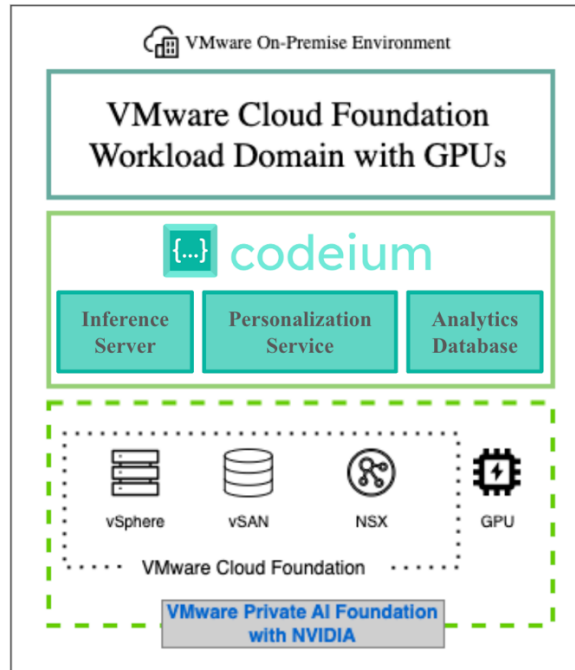
## Architecture Design

With self-hosted deployment, you can run Codeium in a fully air-gapped manner, with no code or data leaving your on-premises servers or virtual private cloud. You get the most advanced AI-powered code acceleration toolkit without compromising the security.

Codeium Enterprise can be deployed as a container-based workload on a Tanzu Kubernetes Grid (TKG) cluster using the [Codeium Helm Chart](#). The TKG cluster is provisioned and lifecycle-managed through the [vSphere IaaS control plane](#) that transforms vSphere, a core element of VCF, to a platform for running Kubernetes workloads natively on the hypervisor layer. The high-level architecture of deploying Codeium Enterprise on VMware Private AI Foundation with NVIDIA is illustrated in the figure below.

<sup>1 & 2</sup> Internal studies conducted by Codeium.

Figure 1: High-Level Architectural Diagram



## How to get started

Listed below is a series of resourceful references that help you understand the technology components that constitute the foundation of this AI-powered solution.

1. [Codeium Enterprise](#)
2. [VMware Private AI Foundation with NVIDIA \(solution brief\)](#)
3. [VCF](#)
4. [NVIDIA AI Enterprise](#)

## Summary

Codeium can help organizations move faster to ship more and achieve more business outcomes, stay competitive with peers and set the engineering organization up for scale, all without the risks that are present with other generative AI solutions. Running Codeium Enterprise on VMWare Private AI Foundation with NVIDIA securely within your tenant and personalized to your company, you can get the best of all worlds of quality, value, and security.

For more information on Codeium, please contact the Codeium team at [codeium.com/contact/enterprise](https://codeium.com/contact/enterprise). Alternatively, you can visit [vmware.com/products/cloud-infrastructure/private-ai-foundation-nvidia](https://vmware.com/products/cloud-infrastructure/private-ai-foundation-nvidia) to learn more about [VMware Private AI Foundation with NVIDIA](#) to unlock the power of generative AI.