

# NETAPP GEN AI TOOLKIT WITH SUPPORT FOR GOOGLE CLOUD NETAPP VOLUMES



## Unlocking enterprise data for generative AI in Google Cloud

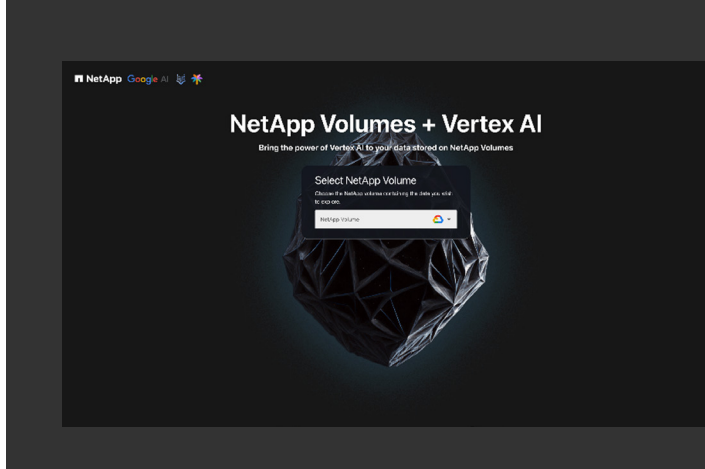
### THE CHALLENGE

To gain a competitive advantage with generative AI (GenAI), you need a way to achieve ultrahigh relevance and produce high-quality results. It's not just about having "more data," but rather about combining publicly available data with your proprietary data, and the unique insights exclusive to your organization. This combination enhances the relevance of learning and results, establishing an ever-expanding intelligent foundation.

As with other open-source software, organizations can build using DIY tools or partner with system integrators. In either scenario, a well-compiled toolkit can make it easier to integrate these complex technologies, allowing organizations to derive greater value from their data.

### THE OPPORTUNITY

NetApp provides AI-ready storage that can power retrieval-augmented generation (RAG) operations and offers the agility to support both general and domain-specific large language models (LLMs). In its integration with Google Cloud, NetApp® storage enables you to innovate faster and smarter by securely augmenting public data with your proprietary data—for increased relevance and pinpoint accuracy from your LLM workloads—and rapidly create intelligent, containerized modern applications.



### THE SOLUTION

The NetApp GenAI Toolkit, which is in preview, empowers you to take advantage of your private data stored on Google Cloud NetApp Volumes by using the foundation models provided by Google Cloud's Vertex AI machine-learning platform. The GenAI Toolkit comes with a chatbot app that's hosted in a NetApp artifact registry and a Terraform module in a GitHub repository that automates deployment of the chatbot app in your environment.

The Toolkit, along with the accompanying reference architecture, allows you to implement RAG operations more quickly while enabling secure, consistent, and automated workflows that connect data stored in Google Cloud NetApp Volumes with Vertex AI. The result is an enhanced ability to generate unique, high-quality, and ultra relevant competitive insights.

## Industry-leading capabilities

The NetApp GenAI Toolkit helps optimize RAG processes with industry-leading capabilities, including:



1

**Common data footprint everywhere.** ONTAP everywhere means you can easily include data from any environment to power your RAG efforts. NetApp ONTAP® data management software lets you use common operational processes while reducing risk, cost, and time to results.



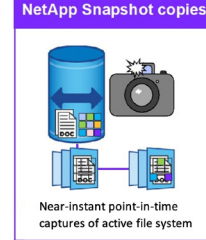
2

**Automated classification.** The NetApp BlueXP™ classification service streamlines data categorization, classification, and cleansing for both the ingest and inferencing phases of the data pipeline. With this approach, the right data is used for queries, and sensitive data is protected according to your organization's policy.



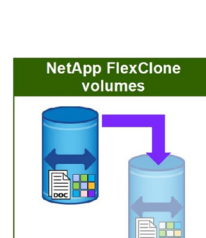
3

**Fast, scalable snapshot copies.** NetApp Snapshot™ technology creates near-instant space-efficient, in-place copies of vector stores and databases for interval-based A/B testing and recovery. You can perform point-in-time analysis or, if data is inconsistent, immediately roll back to a previous version.



4

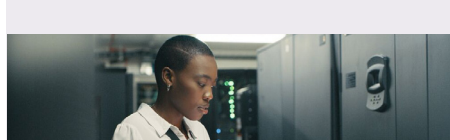
**Real-time cloning at scale.** NetApp FlexClone® technology can create instant clones of vector index stores for parallel processing of A/B prompt testing and result validation. With cloning, you can safely make uniquely relevant data instantly available for queries from different users, without affecting the core production data.



## Advantages of integrating the GenAI Toolkit with Vertex AI

### Enhanced language understanding and generation

Vertex AI offers powerful, sophisticated language models that are trained on vast amounts of diverse data. By integrating these models into the GenAI Toolkit, you can benefit from superior language understanding, contextual awareness, and generation capabilities.

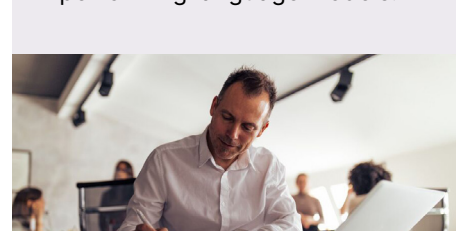


### Improved accuracy and relevance

The LLMs in Vertex AI have been extensively trained on a wide range of tasks and domains, enabling them to capture intricate semantic relationships and generate highly accurate and relevant output. With the GenAI Toolkit's integration, you can use these models to produce more precise and contextually appropriate results, whether you're generating product descriptions, customer support responses, or content recommendations.

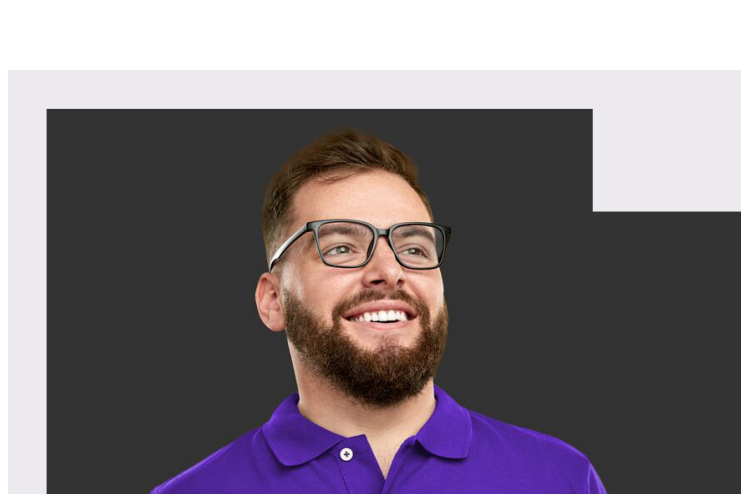
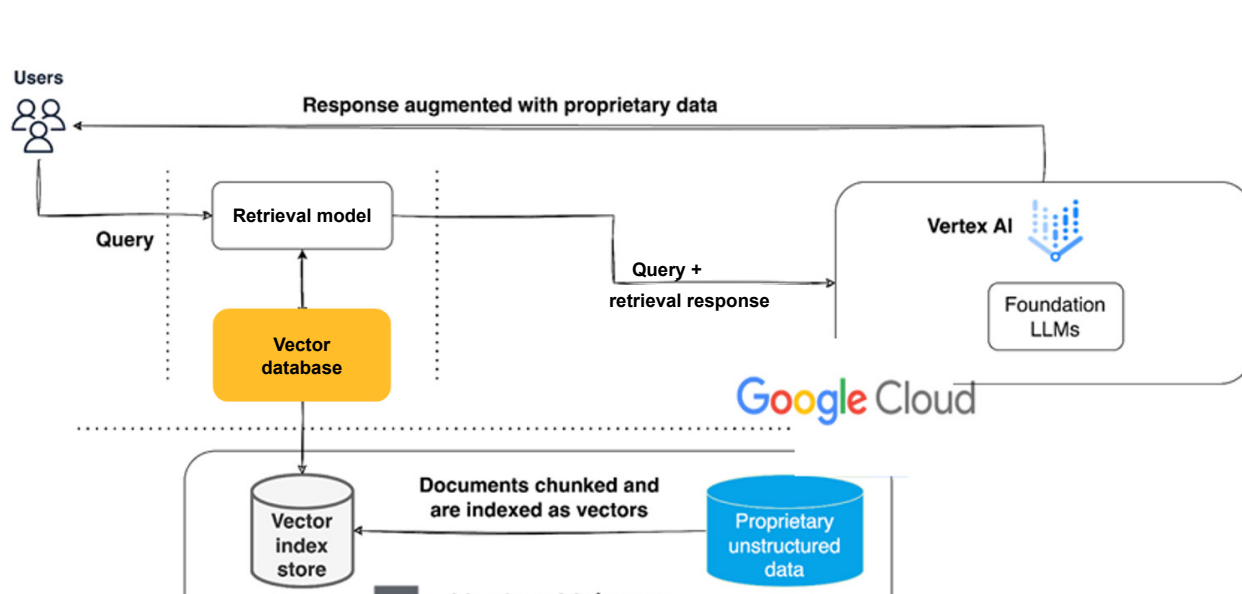
### Continual learning and improvement

Vertex AI LLMs are continually updated and fine-tuned based on the latest advancements in natural language processing. By integrating with Vertex AI, the GenAI Toolkit provides you with access to the most up-to-date and high-performing language models.



## Use case: Securely extending private data to Google Cloud Vertex AI

- Innovate faster and smarter by augmenting vast sums of public data with your valuable, proprietary data—for increased relevance from your LLM workloads.
- Create a temporary instance of selected data with NetApp cloning technology to harness the power of generative AI for enterprise data—without the need for coding or engineering skills.
- When bad data is detected, revert the volume and vector indices back to an earlier Snapshot copy.
- Continuously refresh the data pipeline so that RAG-enriched foundation LLMs include the latest private data.
- Build a chatbot application on your own data without waiting for skilled AI professionals.



The integration of the GenAI Toolkit with Vertex AI represents a powerful synergy that empowers you to harness advanced language generation capabilities while maintaining data privacy and security. To learn more and request demo, [contact a specialist](#).

## Additional resources

[Private RAG – Unlocking Generative AI for the Enterprise](#)

[Google Cloud NetApp Volumes with Google Cloud Vertex AI](#)

[Advancing Secure Gen AI and RAG](#)



[Contact Us](#)



**About NetApp**  
NetApp is the intelligent data infrastructure company, combining unified data storage, integrated data services, and CloudOps solutions to turn a world of disruption into opportunity for every customer. NetApp creates silo-free infrastructure, harnessing observability and AI to enable the industry's best data management. As the only enterprise-grade storage service natively embedded in the world's biggest clouds, our data storage delivers seamless flexibility. In addition, our data services create a data advantage through superior cyber resilience, governance, and application agility. Our CloudOps solutions provide continuous optimization of performance and efficiency through observability and AI. No matter the data type, workload, or environment, with NetApp you can transform your data infrastructure to realize your business possibilities. [www.netapp.com](http://www.netapp.com)

© 2024 NetApp, Inc. All rights reserved. NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners. NA-1108-0524