

Contextual AI and Google Cloud Partner to Bring Generative AI to the Enterprise

SUNNYVALE, Calif., Aug. 21, 2023 /PRNewswire/ -- Contextual AI, the company building AI that works for work, today announced a strategic partnership with Google Cloud as its preferred cloud provider to build, run, and scale its growing business and to train its large language models (LLMs) for the enterprise. Contextual AI [came out of stealth mode in June 2023](#) to build the next generation of foundation models that provide fully customizable, trustworthy, privacy-aware AI that lets companies focus on the work that matters. The company selected Google Cloud for its leadership and open approach to generative AI, as well as the comprehensiveness of its compute infrastructure, purpose-built for AI/ML.

AI workloads require large amounts of time-consuming computation, both to train the underlying machine learning models and to serve those models once they are trained. As part of the partnership, Contextual AI will build and train its LLMs with the choice and flexibility offered through Google Cloud's extensive portfolio of GPU VMs, specifically [A3 VMs](#) and A2 VMs, which are based on the NVIDIA H100 and A100 Tensor Core GPUs, respectively. Contextual AI will also leverage Google Cloud's custom AI accelerators, Tensor Processor Units (TPUs), to build its next generation of LLMs.

Contextual AI enables enterprises to unlock the true potential of AI by grounding language models in their internal knowledge bases and data sources. Built on Google Cloud, Contextual Language Models (CLMs) will craft responses that are tailored to an enterprise's data and institutional knowledge, which results in higher accuracy, better compliance, less hallucination, and the ability to trace answers back to source documents. For example, a customer service agent can leverage CLMs to answer a user's questions with greater precision by relying only on approved data sources such as the user's account history, company policies, and prior tickets that are similar or a financial advisors can automate reporting workflows to provide personalized recommendations based on a client's unique portfolio and history, proprietary market insights, and other private data assets.

"Building a large language model to solve some of the most challenging enterprise use cases requires advanced performance and global infrastructure," said Douwe Kiela, chief executive officer, Contextual AI. "As an AI-first company, Google has unparalleled experience operating AI-optimized infrastructure at high performance and at global scale which they are able to pass along to us as a Cloud customer."

Contextual AI is helping its customers—many of which are Fortune 500 companies—solve shared pain points when it comes to AI, including hallucinations, attribution, compliance, latency, and data privacy. Contextual AI's LLMs take into consideration data privacy while providing customization and efficiency. Co-founder Douwe Kiela helped pioneer the retrieval augmented generation (RAG) technique that underpins Contextual AI's text-generating AI technology. RAG allows enterprise customers to build custom LLMs on top of their data, ensuring that data remains secure, using external sources to generate responses that take context into consideration.

"At Google Cloud, we believe that enabling the next generation of generative AI services requires a purpose-built, AI-optimized infrastructure stack, spanning hardware, software, and services," added Mark Lohmeyer, VP/GM, Compute and ML Infrastructure, Google Cloud. "We're proud to offer customers unparalleled flexibility and performance, and excited to support Contextual AI's world-class team of AI innovators as they build next generation LLMs for the enterprise on Google Cloud."

Additional Resources

- Learn more about Google Cloud's AI infrastructure offerings [here](#).
- Keep up with the latest Google Cloud news on our [newsroom](#) and [blog](#).

About Google Cloud

Google Cloud accelerates every organization's ability to digitally transform its business and industry. We deliver enterprise-grade solutions that leverage Google's cutting-edge technology, and tools that help developers build more sustainably. Customers in more than 200 countries and territories turn to Google Cloud as their trusted partner to enable growth and solve their most critical business problems.

SOURCE Google Cloud

For further information: press@google.com

<https://www.googlecloudpresscorner.com/2023-08-21-Contextual-AI-and-Google-Cloud-Partner-to-Bring-Generative-AI-to-the-Enterprise>