

LG AI Research Taps Google Cloud to Develop EXAONE 3.0 and ChatEXAONE AI Agent

Uses Google's Cloud TPUs and GPUs to train and serve EXAONE 3.0 family of models with superior performance, cost efficiency and low latency

Entire MLOps process - from training to deployment to operation of the ChatEXAONE - conducted on Google Cloud

SEOUL, South Korea and SUNNYVALE, Calif., Aug. 28, 2024 /PRNewswire/ -- LG AI Research today announced and expanded collaboration with Google Cloud to help develop LG's latest family of EXAONE 3.0 generative AI models, and its enterprise multimodal AI agent, ChatEXAONE, trained and developed on Google Cloud's AI-optimized infrastructure.

The two companies previously collaborated to deliver cutting-edge AI innovation to support real-world industrial applications. In addition, LG AI Research worked closely with Google Cloud in December 2021 to develop EXAONE 1.0, Korea's first bilingual and multimodal generative AI model, and in July 2023 on EXAONE 2.0.

As part of today's announcement, LG's most recent EXAONE 3.0 models and proprietary ChatEXAONE AI agent were trained using a combination of Google's Cloud TPUs and GPUs available via Google Cloud's [AI Hypercomputer](#). As a result, LG is able to deliver breakthrough levels of scalability and performance despite the massive computational demands required for such complex models. Google's best-in-class AI accelerators also enabled LG to effectively shorten the time required for AI model training and build a more efficient model training process. Overall, with the power of Google Cloud's AI Hypercomputer, LG was able to reduce inference processing time by 56%, memory usage by 35%, and operating costs by 72% with EXAONE 3.0 7.8B compared to the previous model. Also, EXAONE 3.0, a bilingual model that can learn and understand both Korean and English, recorded the world-class level performance in Korean.

"EXAONE 3.0 has been trained on 60 million pieces of specialized data sets, 350 million images, patents, software codes, mathematics, and chemistry. Google Cloud's AI Hypercomputer has been instrumental in achieving LG AI Research's goals of AI model lightweight technology and cost efficiency," said Hwayoung (Edward) Lee, Lead of AI Biz Dev and DX Group at the LG AI Research.

Security is a very important factor for LG AI Research, which has LG affiliates dealing with major national technologies such as large batteries, displays, and 5G. LG chose to develop ChatEXAONE with Google Cloud because of its built-in data governance and data privacy controls, and robust AI data principles. Data is also grounded in LG's corporate databases, and all of these processes are operated in Google Cloud's fast and smooth MLOps pipeline.

As a next step in their partnership, LG plans to offer EXAONE 3.0 on the Google Cloud Marketplace, and to explore the development of developer-friendly APIs to meet the growing demand for enterprise AI agents.

"Our collaboration with LG AI Research exemplifies the power of partnership in pushing the boundaries of AI innovation. By providing LG with our cutting-edge AI infrastructure, we're not only accelerating their groundbreaking work but also shaping the future of industries worldwide," said Ki-sung Chi, Country Director of Google Cloud Korea.

About LG AI Research

Launched in December 2020 as the artificial intelligence (AI) research hub of South Korea's LG Group, LG AI Research aims to lead the next epoch of artificial intelligence (AI) to realize a promising future by providing optimal research environments and leveraging state-of-the-art AI technologies. And LG AI Research developed its large-scale AI, EXAONE, a 300 billion parametric multimodal AI model, in 2021. EXAONE, which stands for "Expert AI for Everyone," is a multi-modal large-scale AI model that stands out from its peers due to its ability to process both language and visual data. With one of the world's largest learning data capacities, LG AI Research aims to engineer better business decisions through its state-of-the-art artificial intelligence technologies and its continuous effort on fundamental AI research. For more information, visit <https://www.lgresearch.ai/>.

About Google Cloud

Google Cloud is the new way to the cloud, providing AI, infrastructure, developer, data, security, and collaboration tools built for today and tomorrow. Google Cloud offers a powerful, fully integrated and optimized AI stack with its own planet-scale infrastructure, custom-built chips, generative AI models and development platform, as well as AI-powered applications, to help organizations transform. Customers in more than 200 countries and territories turn to Google Cloud as their trusted technology partner.

SOURCE Google Cloud

For further information: press@google.com

<https://www.googlecloudpresscorner.com/2024-08-28-LG-AI-Research-Taps-Google-Cloud-to-Develop-EXAONE-3-0-and-ChatEXAONE-AI-Agent>