Google Cloud Kicks Off Next '23 with a New Way to Cloud

Customer and partner announcements showcase momentum in generative AI

New services and features demonstrate innovation throughout the entire product portfolio

SUNNYVALE, Calif., Aug. 29, 2023 /PRNewswire/ -- Today at Next '23, Google Cloud announced a series of new customers, partners, and product innovations to help every business, government, and user benefit from generative AI and leading cloud technologies.

"We are in an entirely new era of cloud, fueled by generative AI," said Thomas Kurian, CEO, Google Cloud. "Our focus is on putting gen AI tools into the hands of everyone across the organization—from IT, to operations, to security, to the board room. As the industry's most open cloud, our goal is to help companies use AI and other cloud technologies to streamline their operations, increase productivity, and create entirely new lines of business."

Customer momentum around the world
Organizations across industries and around the world are choosing Google Cloud to digitally transform, and this year we've shared stories from **Culture Amp, Deutsche Borse, eDreams ODIGEO, Government of Singapore, HSBC, IHOP, IPG Mediabrands, John Lewis Partnership, The Knot Worldwide, Macquarie Bank, Mayo Clinic, Priceline, Shopify, U.S. Steel,** and **Wendy's**. Today, we are announcing new or expanded relationships with **The Estée Lauder Companies**, **FOX Sports**, **GE Appliances**, **General Motors**, **HCA Healthcare**, and more.

Innovative gen AI startups, like **Replit, Typeface, Jasper,** and more, are increasingly choosing to build on Google Cloud. In fact, today more than half of all funded gen AI startups – including companies like **Contextual**, **CoRover, Elemental Cognition, Fiddler,** and **Quora** are Google Cloud customers. This includes 70% of gen AI unicorns, like **AI21, Anthropic, Cohere**, **Runway,** and others.

New infrastructure and tools to help customers
The advanced capabilities and broad applications that make gen AI so revolutionary demand the most sophisticated and capable infrastructure. Today, Google Cloud announced key infrastructure advancements to help customers, including:

- **Cloud TPUv5e**: This is Google Cloud's most cost-efficient, versatile, and scalable purpose-built AI accelerator to date. Now, customers can use a single Cloud TPU platform to run both large-scale AI training and inferencing.
- **A3 VMs with NVIDIA H100 GPU**: A3 VMs powered by NVIDIA's H100 GPU will be generally available next month, enabling organizations to achieve three times better training performance over prior generation A2.
- **GKE Enterprise**: This enables multi-cluster horizontal scaling required for the most demanding, mission-critical AI/ML workloads.
- **Cross-Cloud Network**: This is a global networking platform that helps customers connect and secure applications across clouds. It is open, workload-optimized, and offers ML-powered security to deliver zero trust.
- **Google Distributed Cloud (GDC)**: GDC is designed to meet the unique demands of organizations that want to run workloads at the edge or in their data centers. The GDC portfolio will bring AI to the edge, with Vertex AI integrations and a new managed offering of AlloyDB Omni on GDC Hosted.

Vertex AI platform gets even better
On top of Google Cloud's world-class infrastructure, the company delivers a comprehensive AI platform—Vertex AI—that enables customers to build, deploy, and scale machine learning models. Customers have access to more than 100 foundation models, including third-party and popular open-source versions, as well as industry-specific models like Sec-PaLM 2 for cybersecurity and Med-PaLM 2 for healthcare and life sciences. New innovations announced today include:

- **Vertex AI Search and Conversation**: Now generally available, the tools enable organizations to create Search and Chat applications using their data in just minutes, with minimal coding and enterprise-grade management and security built in.
- **PaLM 2, Imagen and Codey Upgrades**: This includes updating PaLM 2 to 32k context windows so

enterprises can easily process longer form documents like research papers and books.

- **Tools for tuning**: For PaLM 2 and Codey, Google Cloud is making adapter tuning generally available, and introducing a new method for Imagen, called Style Tuning, so enterprises can create images aligned to brand guidelines with a small amount of reference images.
- **New models**: Google Cloud is announcing availability of Llama 2 and Code Llama from Meta, and Technology Innovative Institute's Falcon LLM, a popular open-source model, as well as pre-announcing Claude 2 from Anthropic. Google Cloud will be the only cloud provider offering both adapter tuning and RLHF for Llama 2.
- **Vertex AI extensions**: Developers can access, build, and manage extensions that deliver real-time information, incorporate company data, and take action on the user's behalf.
- **Digital Watermarking on Vertex AI**: Powered by [Google DeepMind SynthID](#), this is a state-of-the art technology that embeds the watermark directly into the image of pixels, making it invisible to the human eye and difficult to tamper with.
- **[Colab Enterprise](#):** This managed service combines the ease-of-use of Google's Colab notebooks with enterprise-level security and compliance capabilities, helping data scientists accelerate AI workflows.

From the beginning, Google Cloud designed Vertex AI to give users full control and segregation of their data, code and IP. User prompts and data, as well as user inputs at inference time, are not used to improve our models and are not accessible to other customers.

Duet AI in Workspace and Google Cloud
Google Cloud unveiled Duet AI at I/O in May, introducing powerful new features across Workspace and showcasing developer features such as code and chat assistance in Google Cloud. Today, Google Cloud is making [Duet AI in Google Workspace](#) generally available, while expanding the preview capabilities of [Duet AI in Google Cloud](#), with general availability coming later this year.

Workspace is the world's most popular productivity tool, with more than 3 billion users and more than 10 million paying customers who rely on it every day to get things done. With the [introduction of Duet AI](#) just a few months ago, Workspace delivered a number of features to make users more productive, like helping write and refine content in Gmail and Google Docs, create original images in Google Slides, turn ideas into action and data into insights with Google Sheets, foster more meaningful connections in Google Meet, and more. [New enhancements](#) announced today include:

- **Duet AI in Google Meet:** Duet AI will take notes during video calls, send meeting summaries, and even automatically translate captions in 18 languages. Duet AI in Meet also today unveiled studio look, studio lighting, and studio sound.
- **Duet AI in Google Chat:** Users can chat directly with Duet AI to ask questions about their content, get a summary of documents shared in a space, and catch up on missed conversations.

Beyond Workspace, Duet AI can now provide AI assistance across a wide range of Google Cloud products and services—as a coding assistant to help developers code faster, as an expert adviser to help operators quickly troubleshoot application and infrastructure issues, as a data analyst to provide quick and better insights, and as a security adviser to recommend best practices to help prevent cyber threats. [Duet AI in Google Cloud announcements](#) include advancements for:

- **Software development:** Duet AI provides expert assistance across the entire software development lifecycle, enabling developers to stay in flow-state longer by minimizing context switching to help them be more productive. In addition to code completion and code generation, it can assist with code refactoring and building APIs using simple natural language prompts.
- **Application and infrastructure operations:** Operators can chat with Duet AI in natural language across a number of services directly in the Google Cloud Console to quickly retrieve "how to" information about infrastructure configuration, deployment best practices, and expert recommendations on cost and performance optimization.
- **Data Analytics:** Duet AI in BigQuery provides contextual assistance for writing SQL queries as well as Python code, generates full functions and code blocks, auto-suggests code completions and explains SQL statements in natural language, and can generate recommendations based on your schema and metadata.
- **Accelerating and modernizing databases:** Duet AI in [Cloud Spanner](#), [AlloyDB](#) and [Cloud SQL](#), helps generate code to structure, modify, or query data using natural language.
- **Security Operations:** Duet AI is also coming to security products including [Chronicle Security Operations](#), [Mandiant Threat Intelligence](#), and [Security Command Center](#), which can empower security professionals to more efficiently prevent threats, reduce toil in security workflows, and uplevel security talent.

Duet AI was designed using Google's comprehensive approach to help protect customers' security and privacy, as well as Google's [AI principles](#). With Duet AI, *your data is your data*. User code, user inputs to Duet AI, and recommendations generated by Duet AI will not be used to train any shared models nor used to develop any

products.

Simplify analytics at scale with a unified data and AI foundation
Data sits at the center of gen AI, which is why Google Cloud is bringing new capabilities to Google's Data and AI Cloud that will help unlock new insights and boost productivity for data teams. In addition to the launch of Duet AI, which assists data engineers and data analysts across BigQuery, Looker, Spanner, Dataplex, and Google Cloud's database migration tools, other important <u>announcements</u> include:

- **BigQuery Studio**: A single interface for data engineering, analytics, and predictive analysis, BigQuery Studio helps increase efficiency for data teams.
- **AlloyDB AI**: An integral part of AlloyDB, our PostgreSQL-compatible database service, AlloyDB AI offers an integrated set of capabilities for easily building GenAI apps including high-performance and vector queries.
- **New Data Cloud Partners**: Partners, like **Confluent, DataRobot, Dataiku, Datastax, Elastic, MongoDB, Neo4j, Redis, SingleStore,** and **Starburst** are all launching new capabilities to help customers accelerate and enhance gen AI development with data.

Addressing top security challenges
Google Cloud is the only leading security provider that brings together the essential combination of frontline intelligence and expertise, a modern security operations platform, and a trusted cloud foundation, all infused with the power of gen AI, to help drive the security outcomes customers are looking to achieve. Today, Google Cloud announced:

- **Mandiant Hunt for Chronicle:** This service integrates the latest insights into attacker behavior from Mandiant's frontline experts with Chronicle Security Operations' ability to quickly analyze and search security data, helping customers gain elite-level support without the burden of hiring, tooling, and training.
- **Agentless vulnerability scanning:** These posture management capabilities in Security Command Center detect operating system, software, and network vulnerabilities on Compute Engine virtual machines.
- **Network security advancements:** Cloud Firewall Plus adds advanced threat protection and next-generation firewall (NGFW) capabilities to our distributed firewall service, powered by Palo Alto Networks; and Network Service Integration Manager allows network admins to easily integrate trusted third-party NGFW virtual appliances for traffic inspection.

Expanding our ecosystem
Google Cloud's ecosystem is already delivering real-world value for businesses with gen AI, and bringing new capabilities, powered by Google Cloud, to millions of users worldwide. Partners are also using Vertex AI to build their own features for customers – including **Box, Canva, Salesforce, UKG,** and many others. New partner announcements include:

- **DocuSign** is piloting new gen AI features, built with Vertex AI, for its billion-plus users. This includes a new "smart contract assistant" in DocuSign that can summarize, explain, and answer questions about complex contracts and other documents.
- **SAP** is building new solutions utilizing SAP data and Vertex AI that will help enterprises apply gen AI to important business use cases, like streamlining automotive manufacturing or improving sustainability.
- **Workday's** applications for Finance and HR are now live on Google Cloud and they are working with Google Cloud to develop new gen AI capabilities within the flow of Workday, as part of its multicloud strategy. This includes the ability to generate high-quality job descriptions and to bring Google Cloud gen AI to app developers via the skills API in Workday Extend, while helping to ensure the highest levels of data security and governance for customers' most sensitive information.

For more on Google Cloud announcements during Next, please visit the Google Cloud blog or press corner.

Additional Resources

- Google Cloud Next Event Site
- Google Cloud Blog Next updates
- Why Google Cloud?

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/2023-08-29-Google-Cloud-Kicks-Off-Next-23-with-a-New-Way-to-Cloud