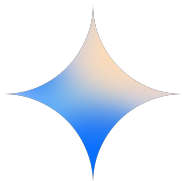CME Group

MARKET TECH FORUM

KEYNOTE

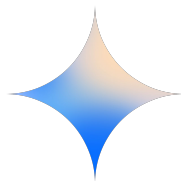# AI in the Cloud and How it Relates to You and Your Business

SPEAKER

Craig Stires, Director, Artificial Intelligence, Google Cloud

What makes Google Cloud AI the right choice?

A deep history of research and innovation

Google's generative AI builds on over 20 years of internal R&D across search, databases, hardware, and AI

A deeply integrated, vertically differentiated portfolio

From super scalable AI infra, to world-class models from a variety of sources, to an intuitive AI development and deployment platform, to AI-powered assistants – you will find the right AI for your skills and needs.
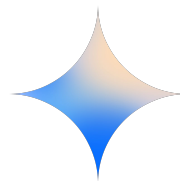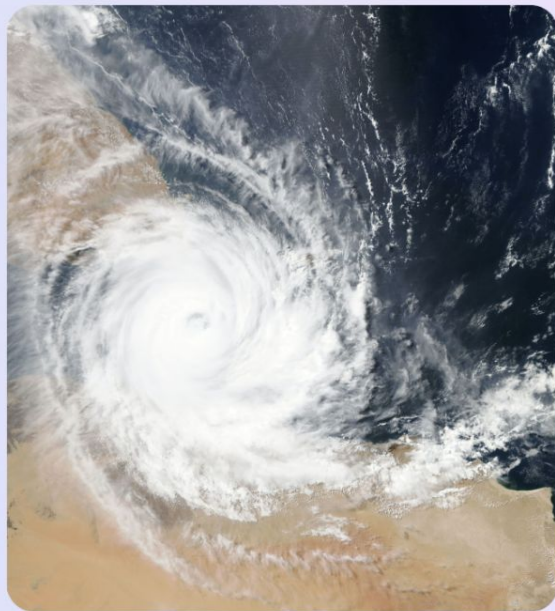
Built on a foundation of enterprise readiness
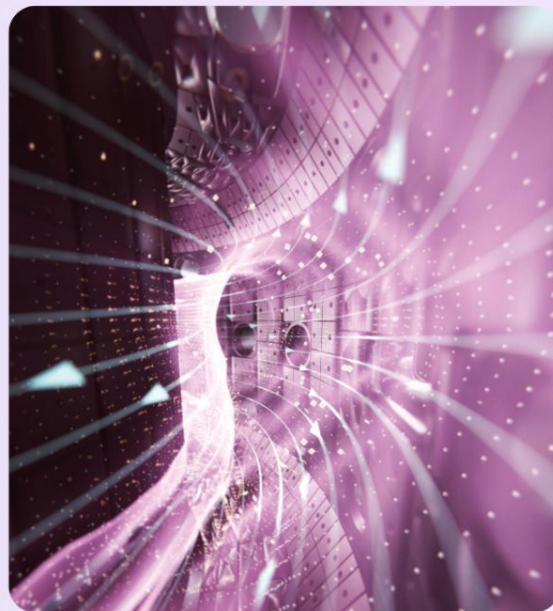
Go to production with the peace of mind provided by data governance and indemnity, security and compliance, and safety and responsible AI.

Google Cloud

Alphacode
competitive-level
programming

GraphCast
weather prediction

Plasma Control
to accelerate
fusion science

This revolution started at Google and we continue to innovate

2017
Transformer

Google invents Transformer kickstarting LLM revolution

2018
BERT

Google's groundbreaking large language model, BERT
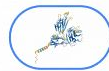
2019
T5

Text-to-Text Transfer Transformer LLM 10B P model open sourced

2020
LaMDA

Google LaMDA model trained to converse

2021
AlphaFold

AlphaFold predicts structures of all known proteins

2022
PaLM

Industry leading large language model

2023
Bard

A conversational AI Service powered by LaMDA.

Responsible AI

Google DeepMind Research

Google Cloud

AI for Capital Markets will apply through all participants

Client Experience

Personalized Interactions

Curated content

Knowledge & Productivity

Improved speed & scope of research

Scaled development & analytics

Alpha Generation

Create, test, and deploy strategies

Reasoning & decisioning

Multi-modal with self service → High precision knowledge discovery

Client Experience

Personalized client segmentation engagement via external marketing, blogs, and graphics generation on demand.

Automate low-touch routine tasks and gain visibility into trading opportunities

Translate financial markets reports and news in real time

Engage clients directly with relevance

Inbox ×

Cymbal Investments Co
to J. Smith

07:31 (3 hours ago)

Mr. Smith,

We have reviewed your existing existing portfolio allocations and see that many of your 5 yr treasury positions are maturing soon. Rolling these into 2 year treasuries could earn an additional $5,000 (+10%) in interest income annually.

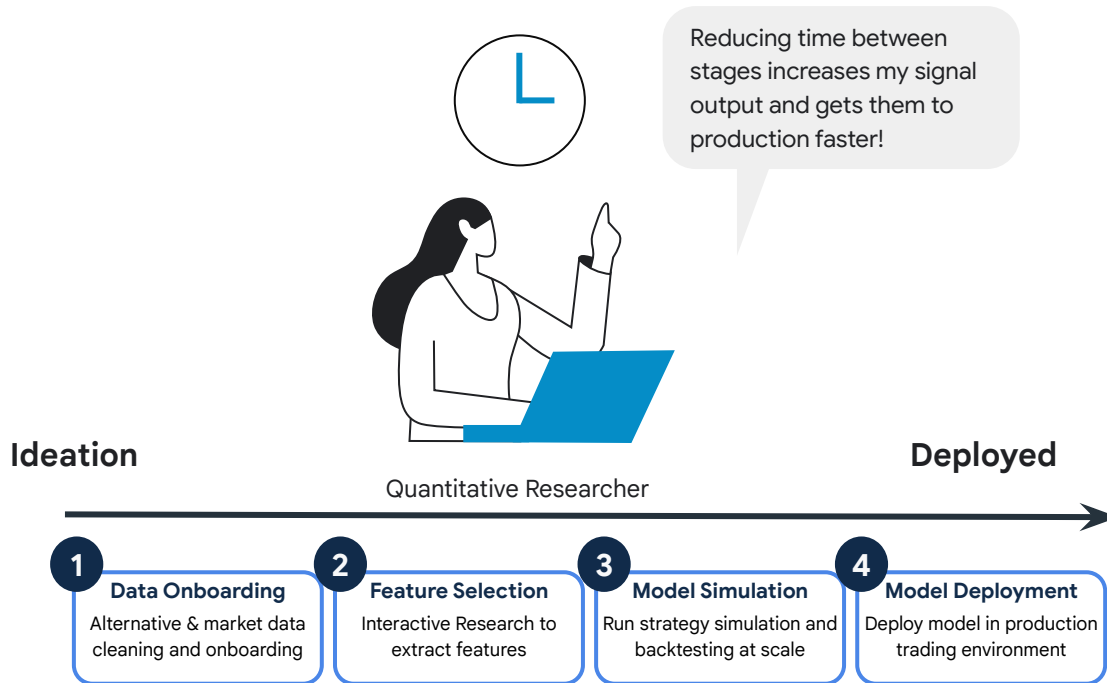**Earn an additional $5,000 in investment income with your next treasury roll.**

Google Cloud

Increasing QR Productivity to Generate Alpha

Two of the most important areas for quantitative firms when partnering with us are:

1) **Researcher productivity**
2) **Price Performance ratio**

We get increase productivity by streamlining data onboarding, eliminating the job queue and providing immediate resources.

We improve price performance ratio through tailored financial engineering and spot / CUD strategies.

Reducing time between stages increases my signal output and gets them to production faster!

Ideation

Deployed

Quantitative Researcher

1 **Data Onboarding**
Alternative & market data cleaning and onboarding

2 **Feature Selection**
Interactive Research to extract features

3 **Model Simulation**
Run strategy simulation and backtesting at scale

4 **Model Deployment**
Deploy model in production trading environment

Google Cloud

Knowledge and Productivity

Explore topics with natural language, assist in sentiment analysis, and aggregate financial data

Accelerate developers in creating complex and scalable cap market applications.
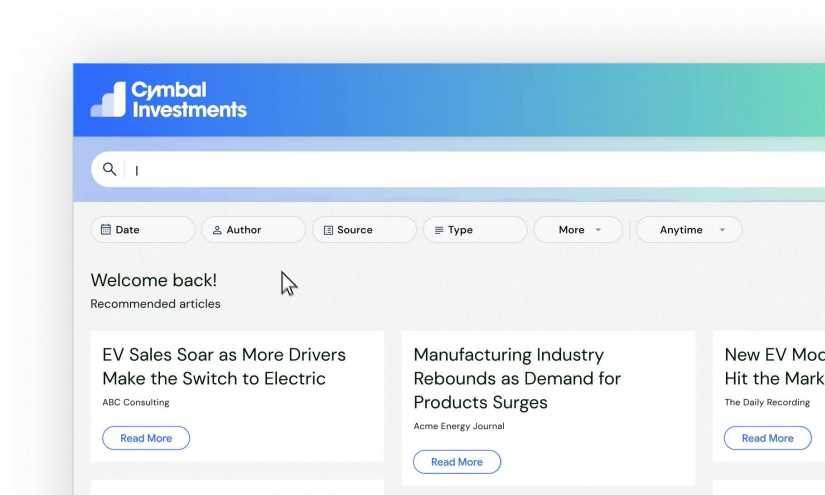
On demand call report summarization

Automate pitch deck creation based upon successful prior pitches and internal client knowledge

I need to assess the semiconductor industry

Portfolio manager

Summarized view of research from internal and external data sources

Google Cloud

You control your data, not Google

We ensure responsible use of data for Cloud AI/ML with these commitments to customers:

By default, we do not use customer data to train our models, in accordance with *GCP Terms* and *Cloud Data Processing Addendum*. *Your customer data is only processed according to your instructions.*
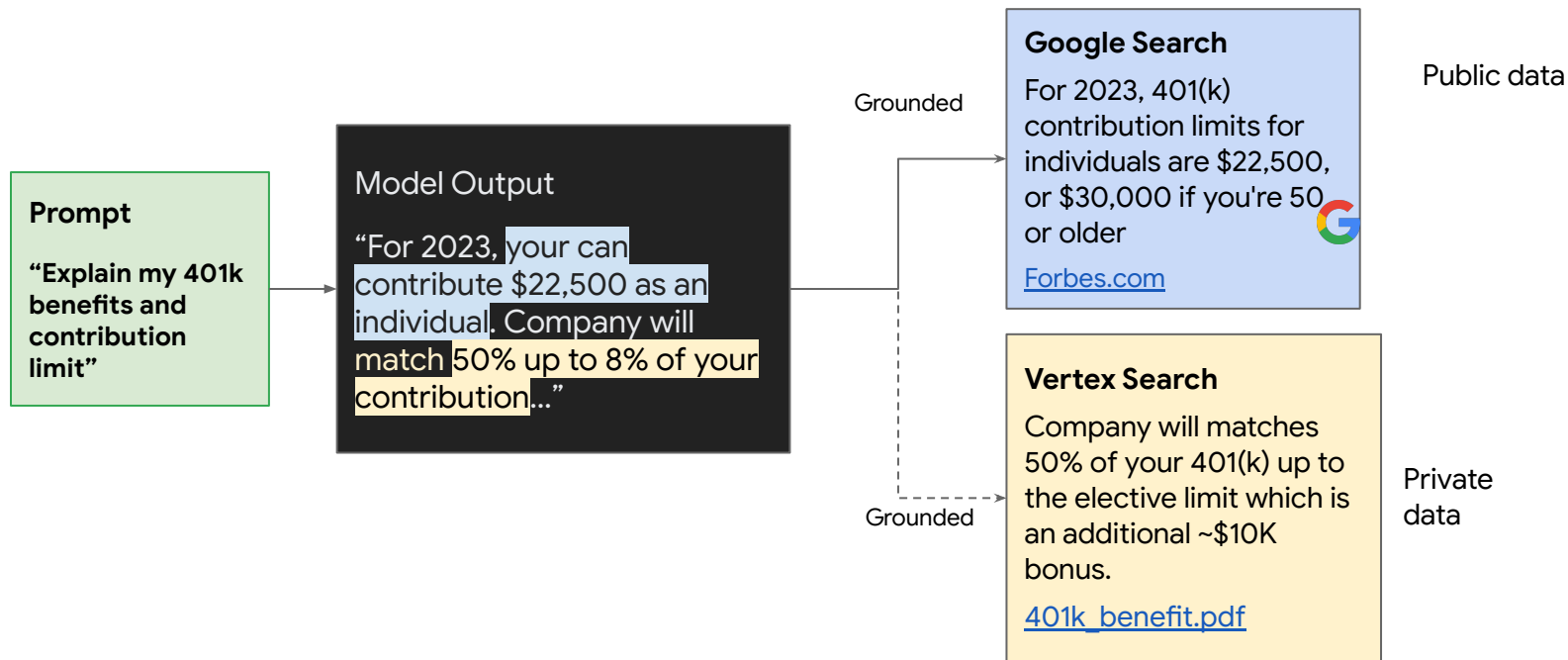
You benefit from our privacy experience and commitments to transparency, compliance with regulations such as the GDPR, and privacy best-practices.

Our extensive data governance & security controls provide you with confidence in the protection of your data from other customers, users, attackers, and unauthorized access by Google employees.

Grounding is a core building block for Vertex AI Search & Conversation

Prompt

"Explain my 401k benefits and contribution limit"

Model Output

"For 2023, your can contribute $22,500 as an individual. Company will match 50% up to 8% of your contribution..."

Grounded

Grounded

Google Search

For 2023, 401(k) contribution limits for individuals are $22,500, or $30,000 if you're 50 or older

Forbes.com

Public data

Vertex Search

Company will matches 50% of your 401(k) up to the elective limit which is an additional ~$10K bonus.

401k_benefit.pdf

Private data

Customization and **expanded grounding** and **compliance capabilities** for customers to develop even more powerful and secure search, chat, and personalized recommendation applications (*blog post*)

Google Cloud

Your data, your terms!
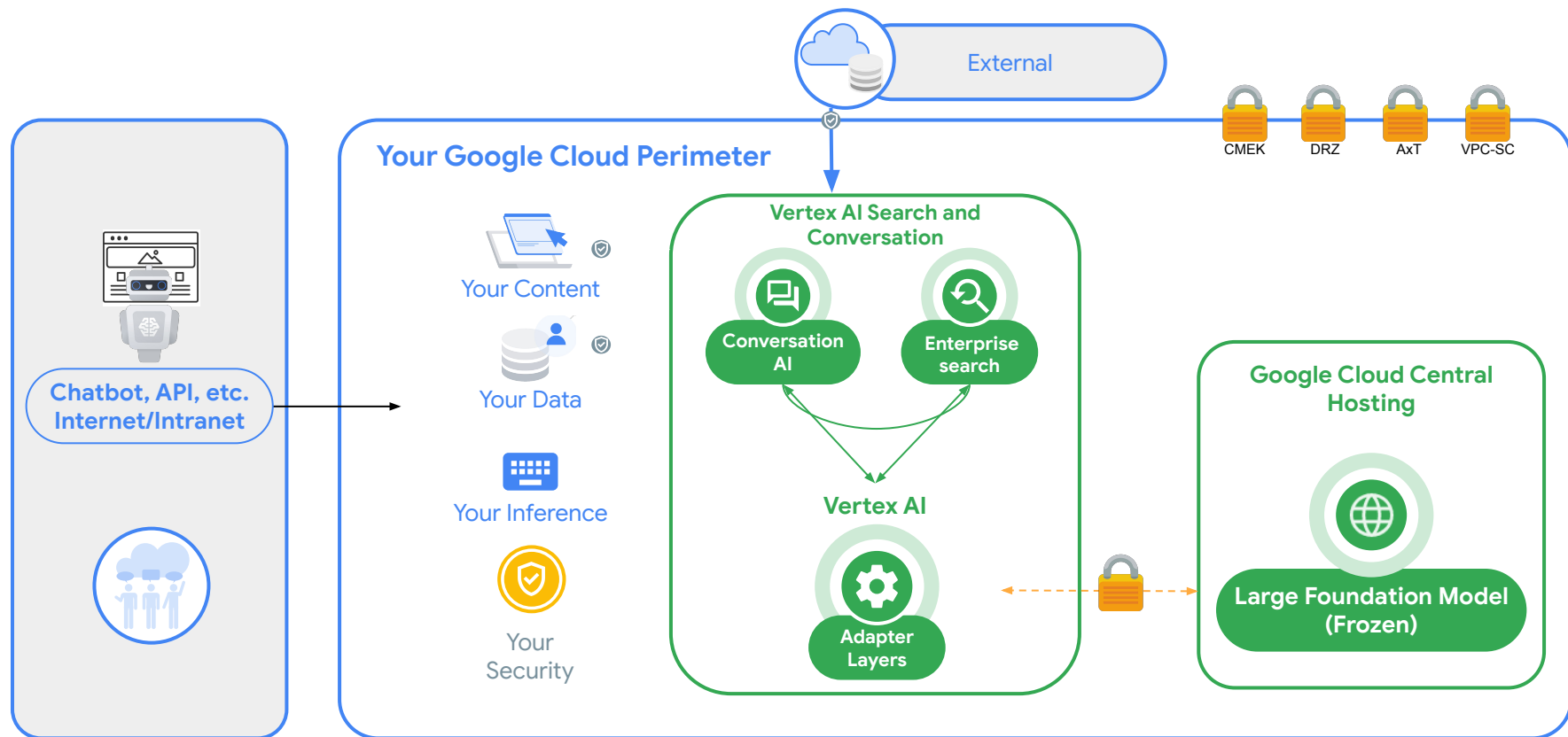
Gemini 1.5 Pro

Mid-size multimodal model with breakthrough long-context understanding

Gemini 1.5 Pro delivers dramatically enhanced performance and represents a step change in our foundation model approach, including:

- A **new Mixture–of-Experts (MoE) architecture** that provides more efficient training and serving, while increasing model performance
- An **expanded context window** (up to 1 million tokens) for complex reasoning across vast amounts of information
- **Better understanding and reasoning across modalities** including text, code, image, audio and video
- **Extensive ethics and safety testing** that builds on novel research on safety risks and leverages red-teaming techniques to test for a range of potential harms

130+ enterprise-ready foundation models in Vertex AI Model Garden

Vertex AI Model Garden

Gemini Foundation Models	Gemini 1.0 Pro	Gemini 1.5 Pro	Gemini 1.0 Ultra	
Google Foundation Models	PaLM 2	Imagen 2	Chirp	Codey	Embeddings
Google Task Specific Models	Speech-to-Text Text-to-Speech Natural Language Translation Doc AI OCR Occupancy analytics Vision Video Intelligence				
Google Domain Specific Models	MedLM — Life Science and Healthcare Sec-PaLM — Cybersecurity				
Partner & Open Ecosystem	Llama 2 Code Llama Technology Innovation Institute — Falcon Claude 2 Pre-announce MISTRAL AI_ Gemma				

- **Choice and flexibility** with Google, open source, and third-party foundation models

- **Multiple modalities** to match every use case

- **Multiple model sizes** to match cost and efficacy needs

- **Domain-specific models** for specialized industries

- Enterprise ready with **safety, security, and responsibility**

- Decrease time to value with **fully integrated platform**

Google Cloud

We closely track and engage in AI policy, regulations, and standards.

Google is a **founding sponsor** of the ISO 42000 **family of AI standards** and recognized as a **global leader in AI governance** best practices.
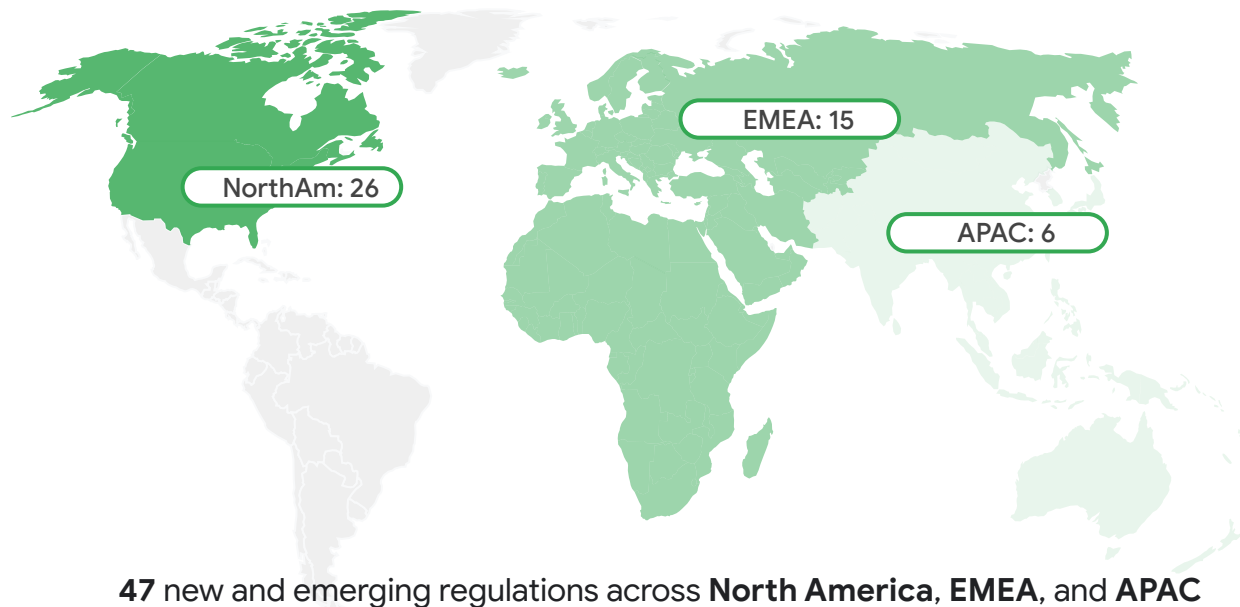
Google Cloud existing certifications

* AI certifications in progress

EMEA: 15

NorthAm: 26

APAC: 6

47 new and emerging regulations across **North America**, **EMEA**, and **APAC**

Why Google?

Govern your data

Google Cloud ensures that you retain control over your proprietary data and how it is used in Google Cloud

Match cost to value

Bigger is not always better when it comes to inference cost. Vertex AI gives you the ability to use the right model for the right job.

Grounded in factuality

It's important to not just provide an answer but to prove where the answer came from. Integrate databases and search results to ground your results.

Easy to get started

Cloud AI gives you a wide range of easy to use tools to get started in days, instead of needing to be a deep learning expert for every use case.

Manageable at scale

Google Cloud provides a platform that integrates into your existing user security, data management, and entitlements regime to manage your myriad data and models.

In summary, Google can help you land production ready and scalable LLM functionality

Google Cloud