

AI & MACHINE LEARNING

Expanding Google Cloud AI to make it easier for developers to build and deploy AI

Rajen Sheth

VP, Product Management, Cloud AI and Industry Solutions

April 10, 2019



Every year, more and more businesses look to AI to help them solve complex business challenges. Whether they're using AI to anticipate demand, predict when equipment will need routine maintenance, or deliver better customer experiences, they all have one thing in common: they need a workforce that can help them do it.

Our goal has always been to make AI simpler, faster, and more useful for businesses. This means easy-to-use AI solutions that make it simple for enterprises to adopt them.



Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)



Blog

Menu ▾

- [AutoML Natural Language](#)
 - [Custom entity extraction \(beta\)](#)
 - [Custom sentiment analysis \(beta\)](#)

Introducing AI Platform: build AI applications, then run them in the cloud or on premises

When approaching AI projects, businesses grapple with a variety of problems—from unstructured data to siloed teams to complex deployments. They need a place that brings all these things together in a way that makes ML easier and more collaborative.

Today, we're announcing [AI Platform](#) in beta, a comprehensive, end-to-end development platform that helps teams prepare, build, run, and manage ML projects via the same shared interface. Whether you're a developer, data scientist, or data engineer, you can collaborate on model sharing, training, and scaling workloads from the same dashboard within Cloud Console.

With AI Platform, you can ingest streaming or batch data, and use a built-in [labeling service](#) to easily label training data—like images, videos, audio, and text—by applying classification, object detection, entity extraction, and other processes. You can import your data directly into AutoML, or use Cloud Machine Learning Engine, now part of AI Platform, to train and serve your own custom-built ML models on GCP. AI Platform



Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

enterprises to use AI

[Tables](#), now available in beta, lets you build and deploy state-of-the-art machine learning models on structured tabular datasets with zero code. With just a few clicks, you can ingest data from BigQuery and other GCP storage services into AutoML Tables and build and deploy ML models in just days versus weeks. The codeless interface guides you through the full end-to-end machine learning lifecycle, making it easy for anyone on your team—whether data scientist, analyst, or developer—to build models and reliably incorporate them into broader applications.

For an ever deeper look at AutoML Tables, read our [data analytics blog post](#).

Extending AutoML Vision to the edge

Optimizing machine learning models to run on edge devices, like connected sensors or cameras, can be challenging because these devices often grapple with latency and unreliable connectivity. Last year, we announced AutoML Vision to make it easier for developers to create custom ML models for image recognition. Today we're announcing AutoML Vision Edge to simplify training and deployment of high-accuracy, low-latency custom ML models for (on premises or remote) edge devices. AutoML Vision Edge supports a variety of devices and can take advantage of Edge TPUs for faster inference. For example, LG CNS is using AutoML Vision Edge to create manufacturing intelligence solutions that detect defects in everything from LCD screens to optical films to automotive fabrics on the assembly line.

Enabling powerful content discovery and engaging experiences with AutoML Video

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

understand the overall opinion, feeling or attitude expressed in a block of text, tuned to your organization's own domain-specific sentiment scores.

Continuing to make machine learning faster with the latest accelerators

We continue to invest in the infrastructure that makes machine learning possible for you. Our [Cloud TPUs](#), custom-built to quickly train ML models, lets you iterate at scale to achieve higher classification accuracy, at a lower cost. Our third generation liquid-cooled TPUs are now generally available, and all Cloud TPUs are also generally available in [Google Kubernetes Engine](#) (GKE), which is a new and flexible way to run your containerized ML workloads, giving you the flexibility to switch between on-prem and cloud-based training. GCP is also the first cloud provider to offer the new [NVIDIA Tesla T4](#), now generally available across eight regions.

A fully-featured, user-centric ecosystem for machine learning

As part of today's announcements, we're also working with numerous partners—including Accenture, Atos, Cisco, Gigster, Intel, NVIDIA, Pluto 7, SpringML, and UiPath—to build Kubeflow pipelines to grow and extend AI Hub. It takes a robust partner

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)



Blog

Menu 

Follow Us




Google

[Privacy](#)

[Terms](#)

[About Google](#)

[Google Cloud Products](#)

Language 



Help