

Servizi Google Cloud per il Machine Learning, Use cases

Put the best of Google's AI technologies to work

Vertex AI

Put the best of Google's AI technologies to work

Unified development and deployment platform
for data science and machine learning

Q2 2022

Agenda

1

Shaping your AI strategy

2

Vertex AI overview

3

Transforming with Vertex AI

Shaping your AI strategy

01

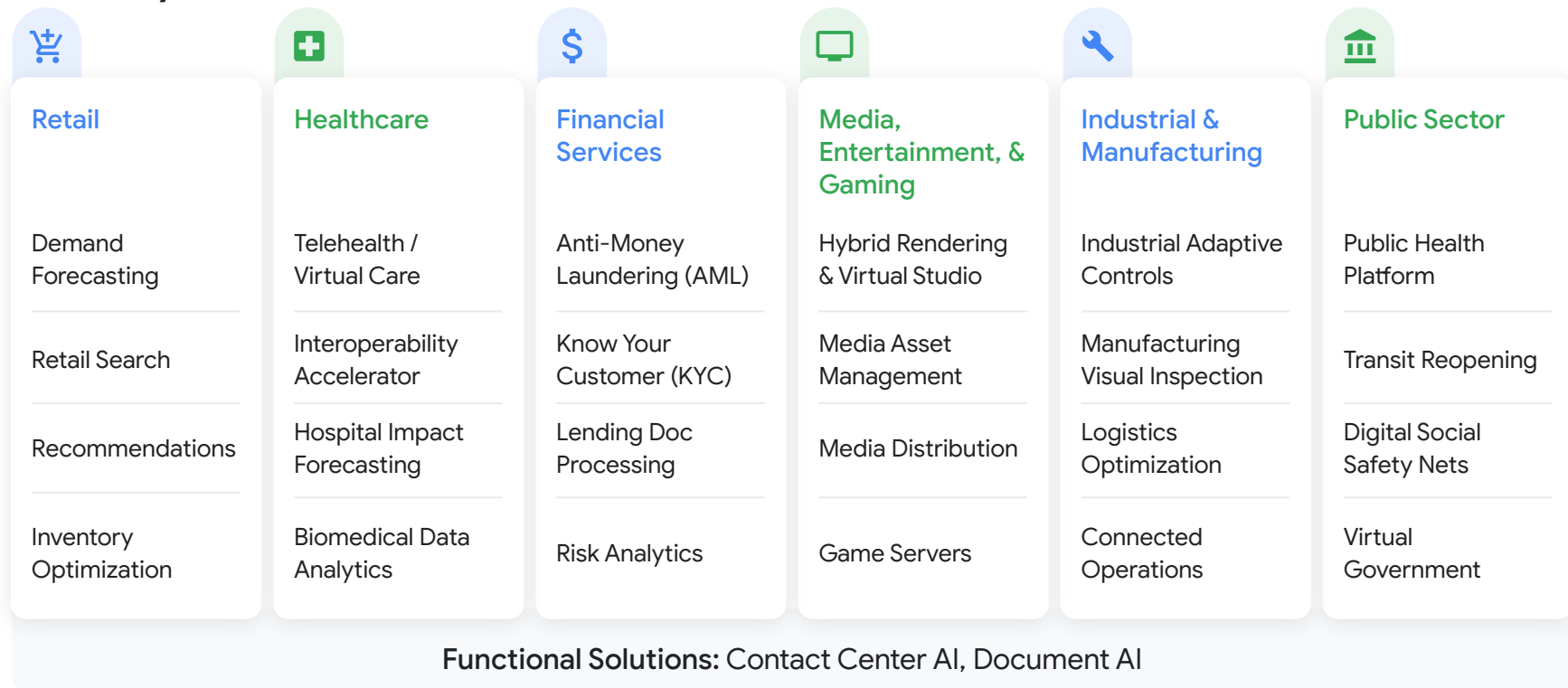
“

March 2020

Thanks to Google Cloud AI Platform and the work done with it over the last 12 months, our team has been able to deploy a new ML model in production to respond to the COVID-19 crisis in less than 2 days.”

Head of Data Science, EMEA-based e-commerce company

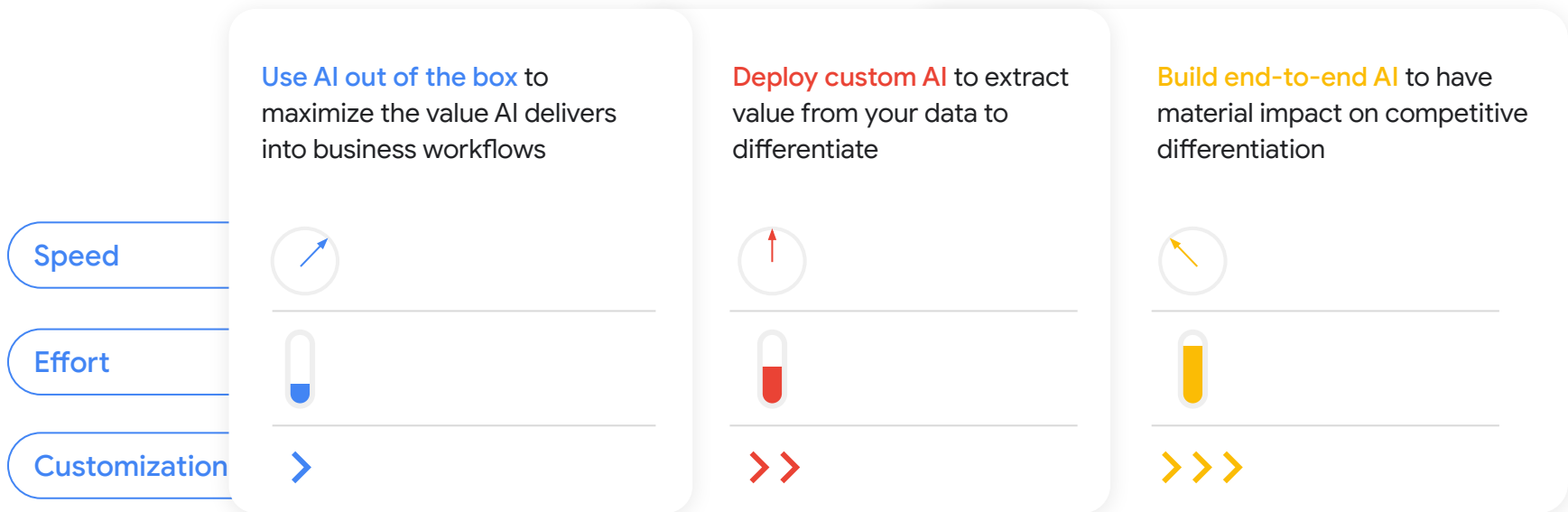
There are AI opportunities in every industry



AI maturity phases



Mapping the business objective to an AI absorption strategy

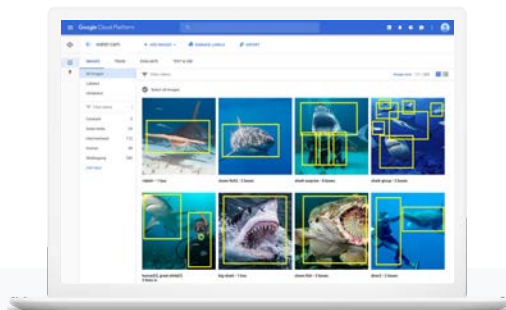


Google Cloud AI for every level of expertise



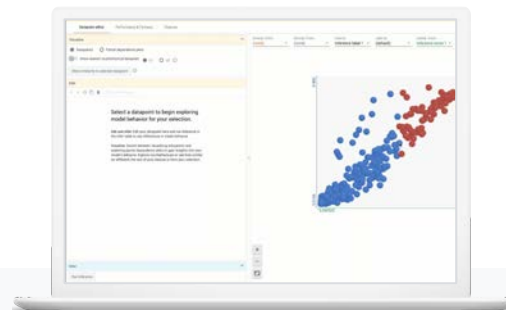
Pre-trained APIs & AI Solutions

No training data needed, get started right away



Custom AI with AutoML

Easily create custom models
(A no-code approach)



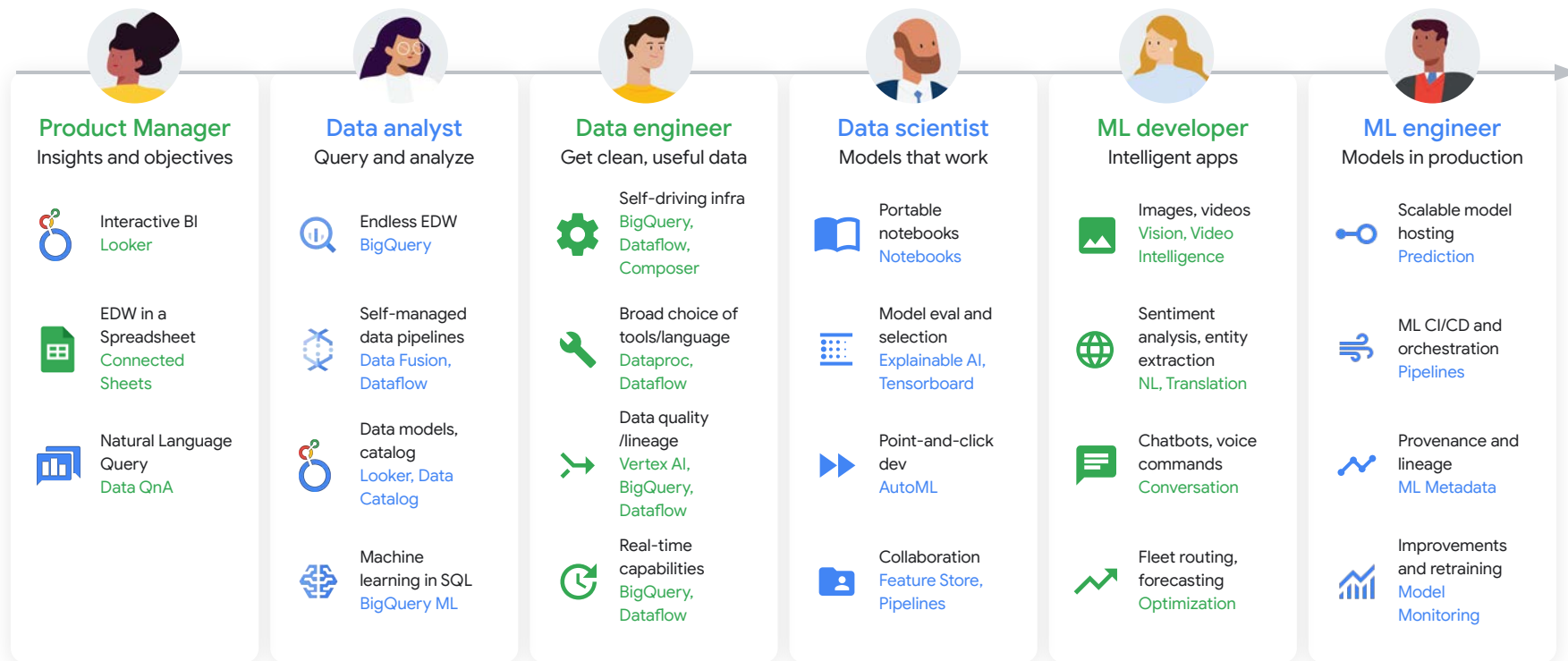
End-to-end AI with core tools

Help data scientists and ML engineers build and deploy AI

Vertex AI overview

02

A platform for all users and intents throughout the ML lifecycle





Vertex AI

Applications

Vision and Video

Conversation

Language

Structured Data

Custom machine learning

Workbench

AutoML

NAS

Prediction

ML Metadata

Data Labeling

Training

Explainable AI

Feature Store

Model Monitoring

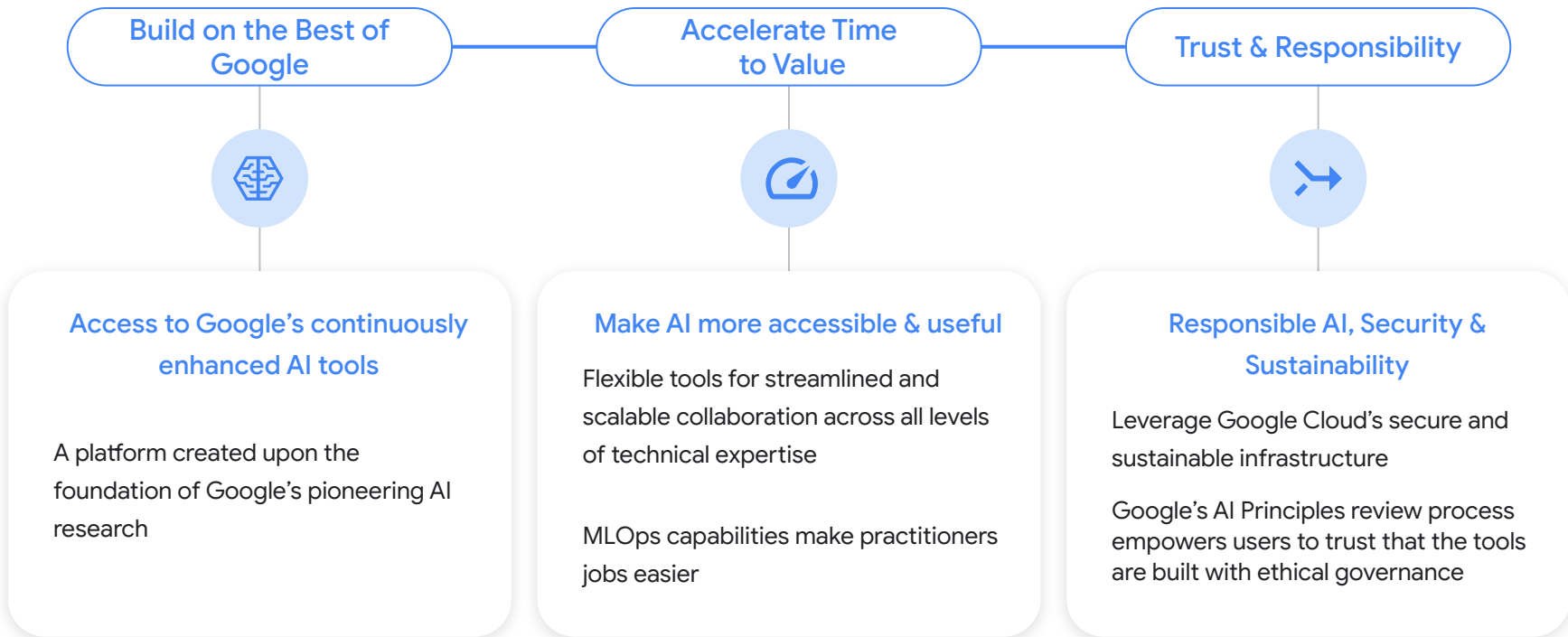
Experiments

Vizier (Optimization)

Pipelines

Matching Engine

Why organizations choose Vertex AI



Transforming with Vertex AI

03

What customers get with Vertex AI

Comprehensive, end-to-end platform for everything AI

01

Velocity of
models in
production

02

Google's
best-in-class
algorithms
as a service

03

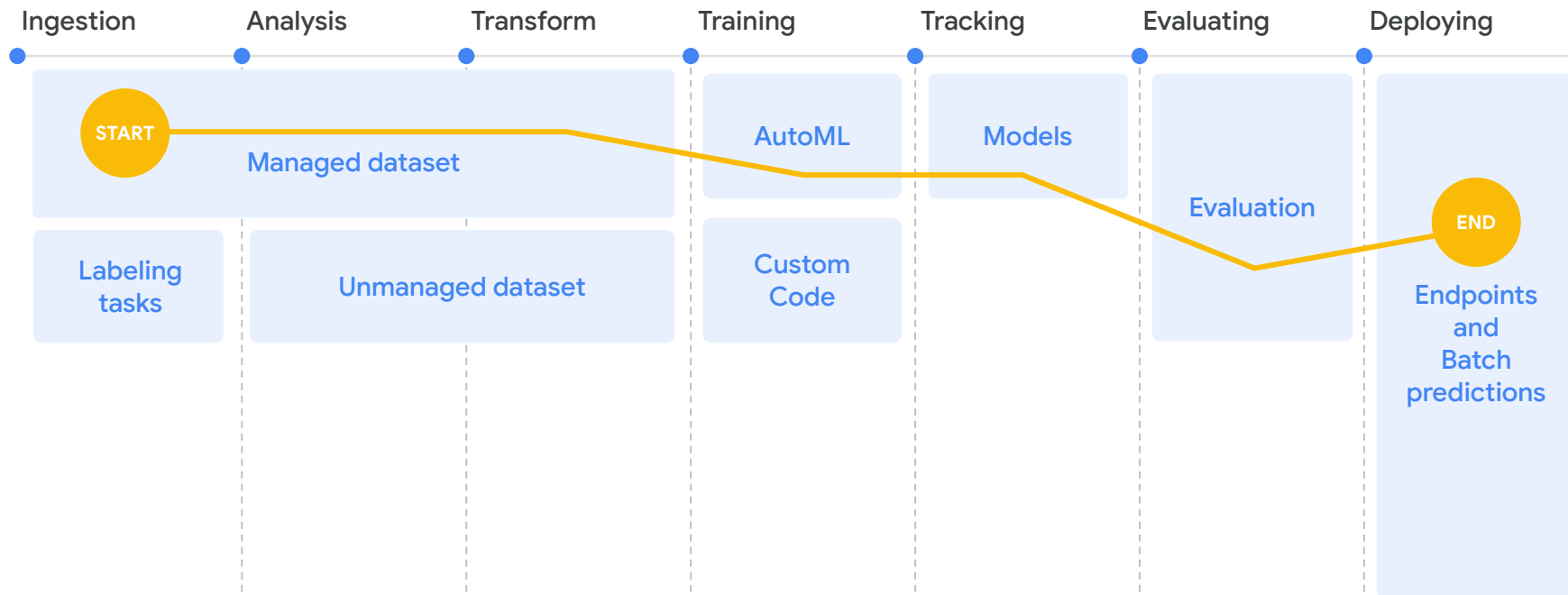
Scaling ML
workflows

04

Model
management
and governance

1

Get models into production faster with the new, unified workflow

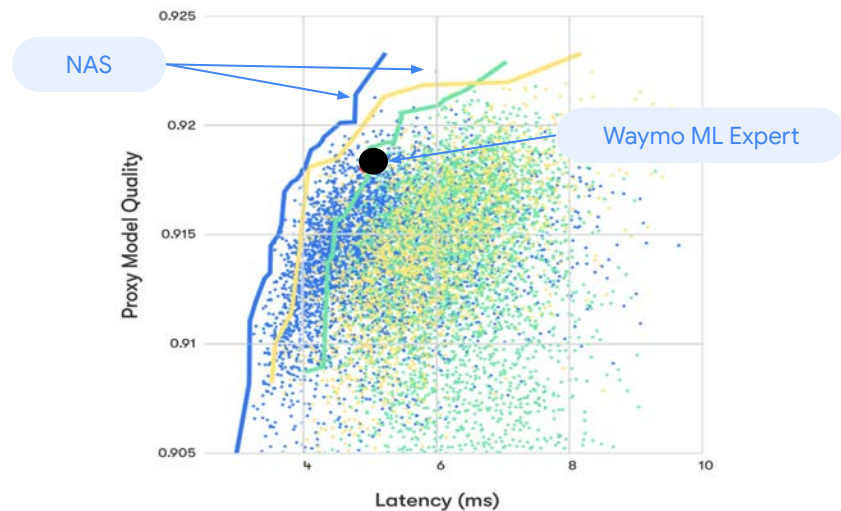


2 Access to Google's best-in-class algorithms like NAS

Use of Neural Architecture Search (NAS) at Waymo

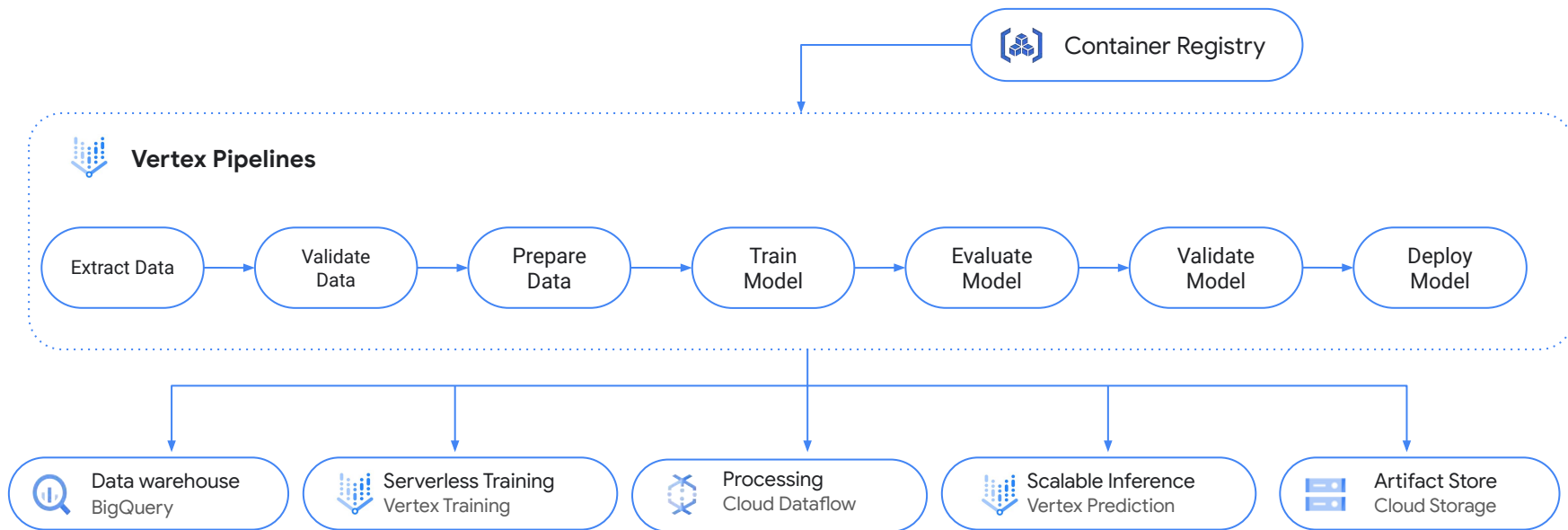
"Going from months of engineering time to generate and fine tune a architecture manually to "automatically generating" neural nets with NAS"

- ⌚ 20–30% lower latency/same quality
- 📉 8–10% lower error rate/same latency
- 📅 NAS model in 2 weeks vs months (1 year of GPU time) searching over 10k architectures



3

Scaling ML workflows with **Vertex Pipelines** and the **MLOps suite** of services



4

Vertex AI is well positioned to help organizations with model governance and compliance

Example Area	What organizations need	How Vertex AI can help
Model transparency	Documentation	ML Metadata
	General explanation of system logic	Explainable AI
	Ability to explain specific outputs	Explainable AI
User trust	Reassurance that systems are not unfairly biased	Explainable AI, Pipelines
	Facilities for third party auditing	ML Metadata and Pipelines
	Mechanisms to reassure that AI systems are robust	Pipelines, Model Monitoring
Clear accountability	Facilities for human oversight/monitoring	Data labeling service, human in the loop
	Clarity over legal responsibilities, including liability	Compliance with industry standards (FedRamp, PCI DSS, ISO)
	Clearer and more accessible frameworks for redress should a problem arise	Bias optimizations using What-If Tool, Fair-aware in AutoML

How organizations are transforming their teams with Vertex AI



Unified machine
learning and data
science



Streamline and scale
with MLOps

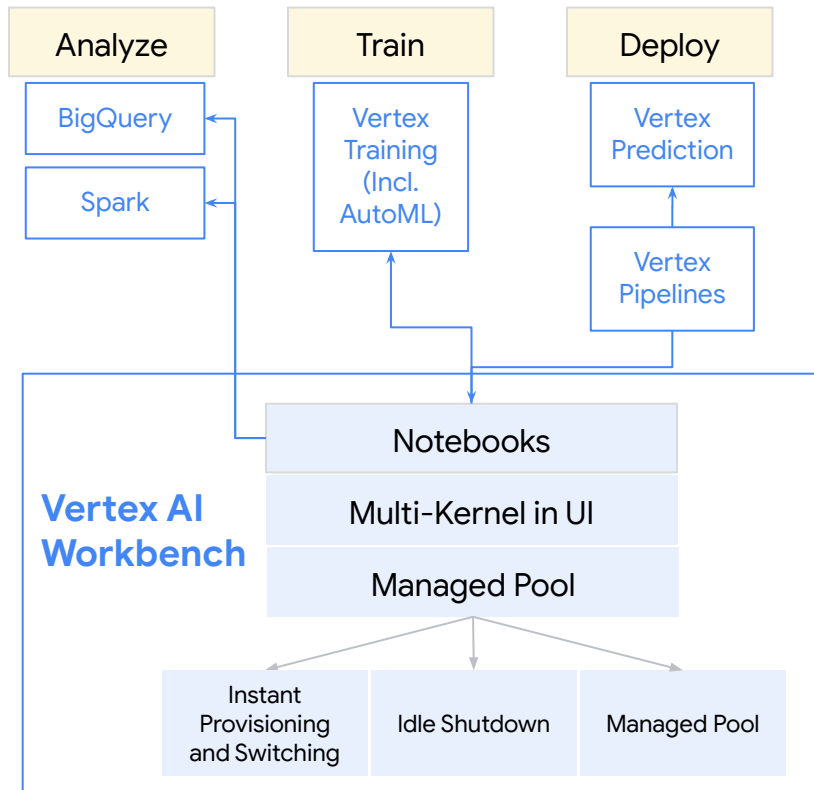


Easy-to-deploy AI
applications

Code

Vertex AI Workbench

A one-stop surface for Data Science



Fully Managed Compute

A Jupyter-based fully managed, scalable, enterprise-ready compute infrastructure with easily enforceable policies and user management



MLOps

Notebooks will be integral to continuous training, and deployment workflows with ML pipelines



Unified Workbench

Seamless visual and code-based integrations with analytics and Vertex AI services

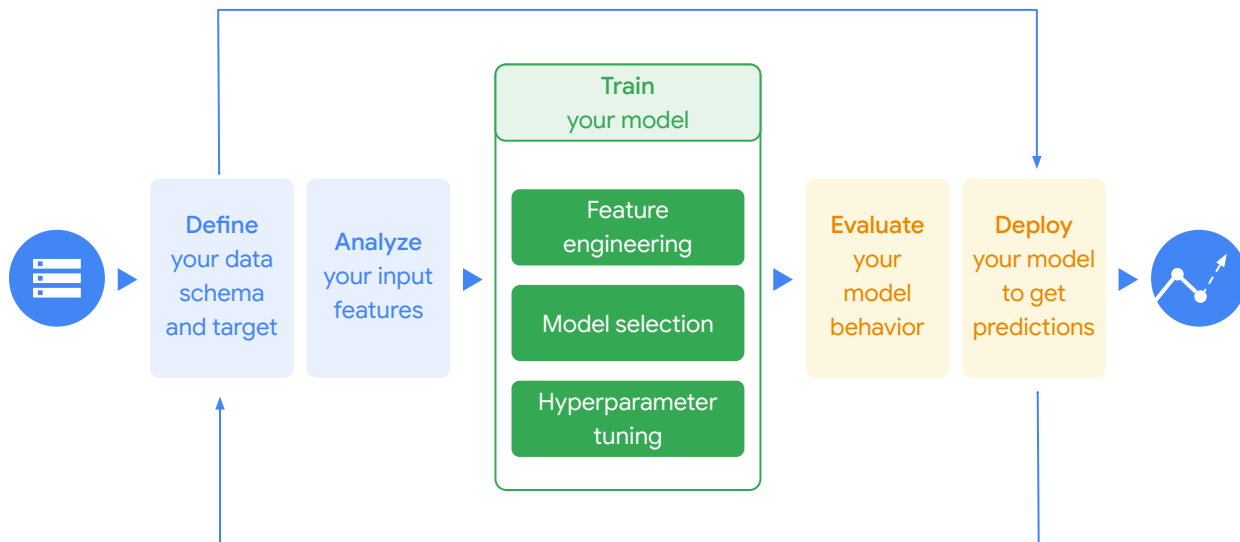
Low/No code

Point and click to build custom,
high-quality models using the

AutoML workflow in **Vertex AI**



AutoML workflow



Automatically search
through Google's whole
model zoo...

Linear, logistic

Feedforward DNN

Wide and Deep NN

Gradient Boosted Decision Tree
(GBDT)

DNN + GBDT Hybrid

Adanet ensemble

Neural + Tree Architecture Search

...and more!

Train

Run and scale your code with high availability using **Vertex AI Training**



Serverless experience,
no provisioning

Rapid cluster orchestration

Built-in logging and monitoring



Ephemeral clusters on-demand

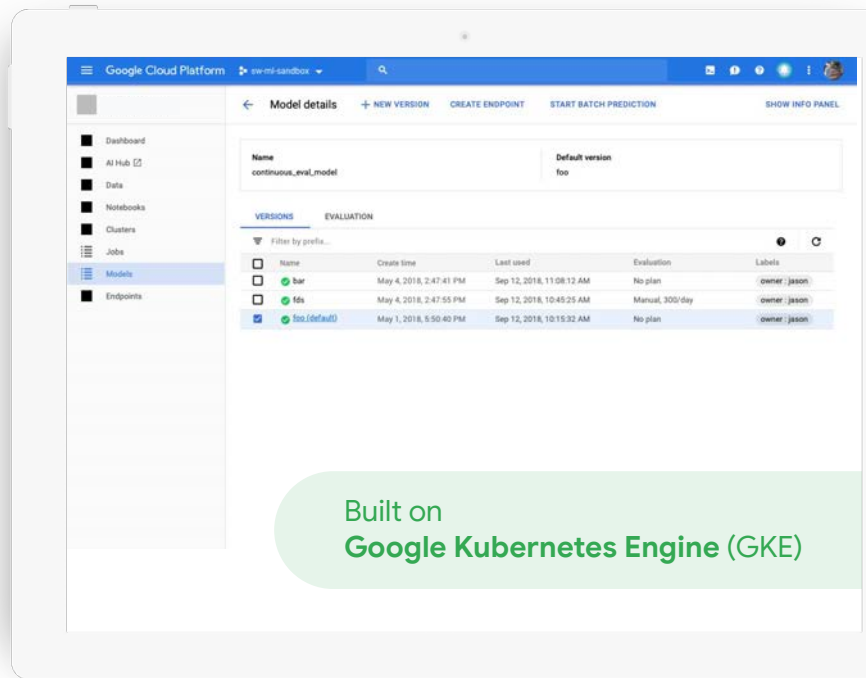
Automatic job queuing

Pay for only what you use

Serve

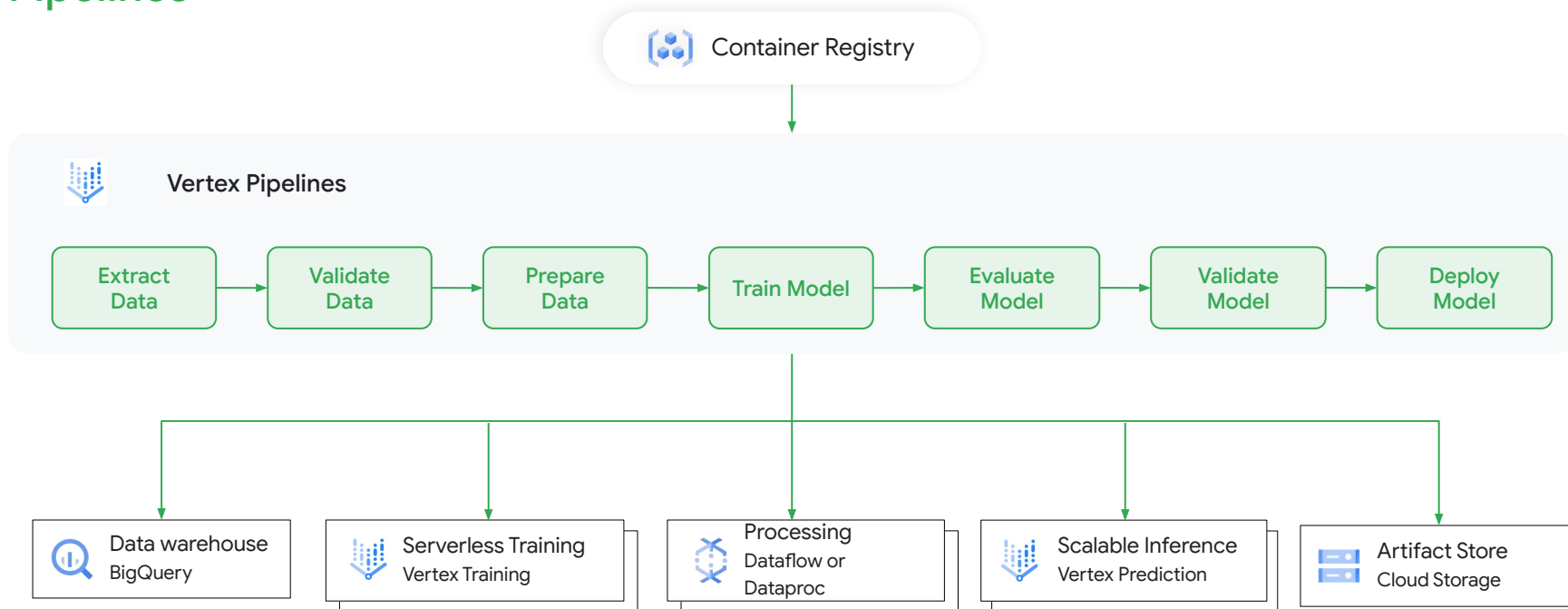
Robust and reliable model hosting with Vertex AI Prediction

- Serve **online** endpoints for low-latency predictions, or predictions on massive **batches** of data.
- **Built-in security and compliance:** VPC peering and security perimeter. Custom managed encryption keys. Fine-tuned access control.
- **Low TCO:** Scale automatically based on your traffic, and alleviate operational overhead.
- **Intelligent and assistive:** Built-in Model Explainability and proactive model monitoring.
- Log prediction requests and responses to **BigQuery** for monitoring and debugging
- **Fast inference on GPUs:** Support for a broad range of machine types specialized for ML, such as GPUs.



Manage

Simplify MLOps with Vertex AI Pipelines



Easy-to-deploy AI applications



Highest quality models and services

The quality and accuracy of our models compared to others in the market



Easy customization and quick deployment

Making it easy for developers without extensive ML expertise to build custom AI and quickly infuse AI into their applications



Proven business value

Value-driven approach to build productized solutions, and product capabilities that maximize business impact and developer productivity

Easy-to-deploy AI applications



Sight



Vision



AutoML Vision



Video Intelligence



AutoML
Video Intelligence



Language



Translation



AutoML Translation



Natural Language



AutoML
Natural Language



Conversation



Dialogflow



Speech-to-Text



Text-to-Speech



Structured Data



AutoML Tables

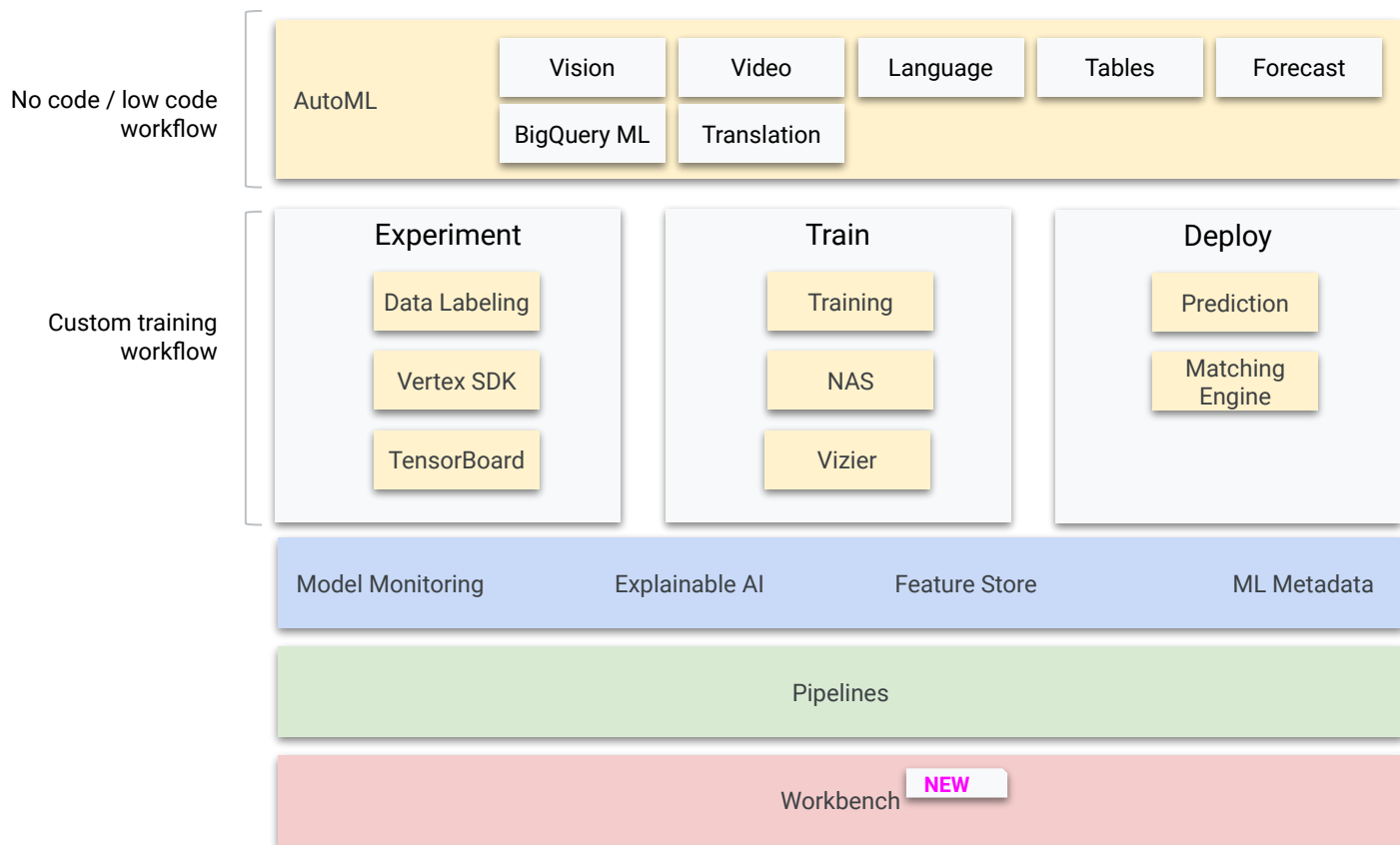


Recommendation AI



Vertex AI

Generally available



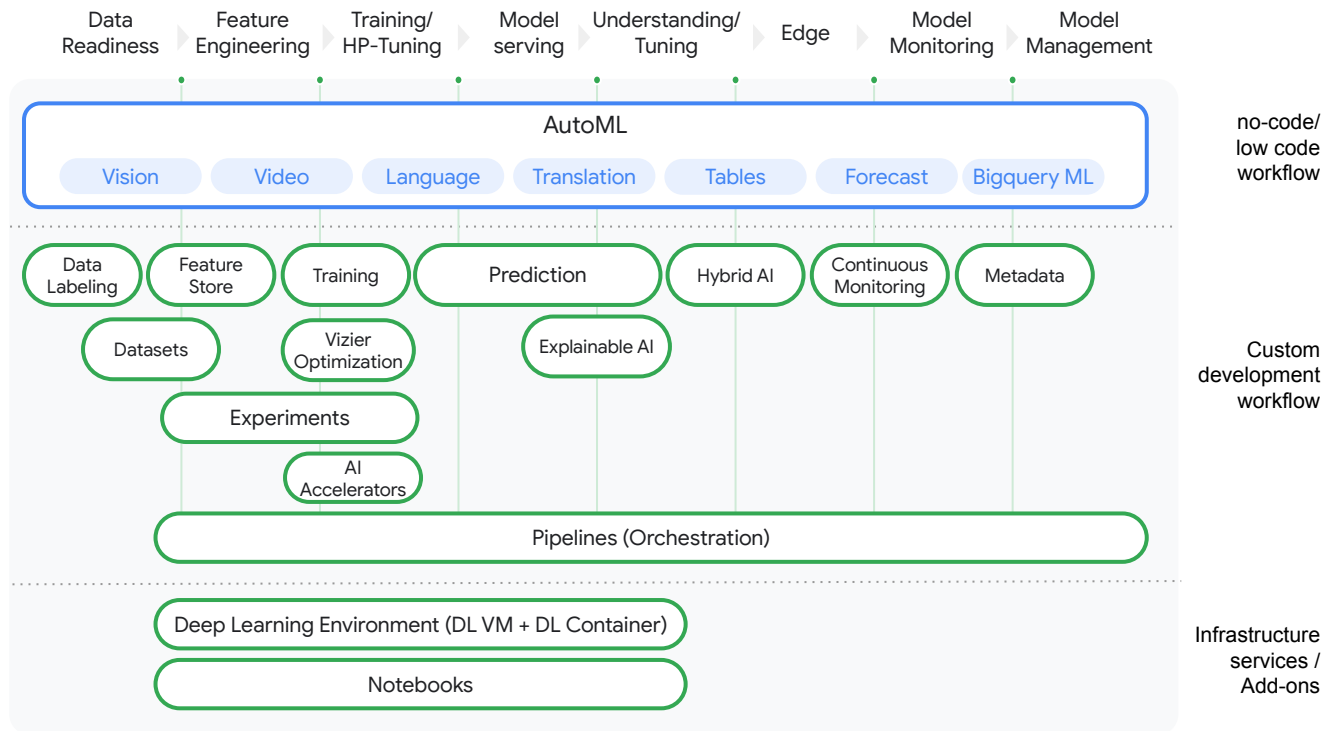
One comprehensive end to end platform for everything AI

One unified experience to create, deploy, and manage models over time, at scale

Tools for all levels of expertise

Accuracy and fairness of predictions and resulting decisions

Flexible and secure





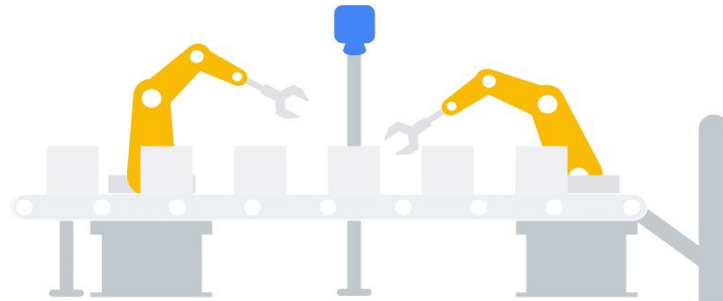
Google Cloud



Visual Inspection

Visual Inspection AI

Transform manufacturing quality



Manufacturers rely on inspection processes for defect detection or assembly conformance throughout the production process



Challenges with current methodologies:



Existing optical inspection stations are **not responsive** to engineering or manufacturing changes



Manual validation inspection is **subjective to operator perception** and experience



Existing optical inspection can only detect a **handful of defects at a time**



No ability to retain manufactured quality data for future reference



High margin of error on manual packaging leads to **customer dissatisfaction**

Artificial Intelligence (AI) and Computer Vision (CV) can address the challenges with existing inspection methodologies

with
AI
and
CV



No AI re-programming necessary, adaptive to changes in products and easily scalable to other production lines



Reduced cognitive load for operators, less defects slipped



Can detect hundreds of areas of interest in a product with a single image capture

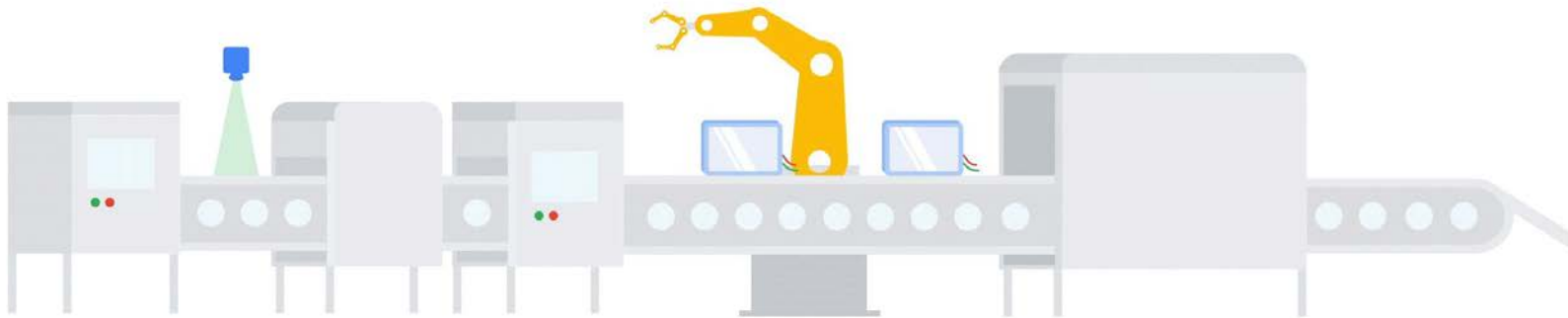


Retain images for root cause analysis or continuous process improvement



Ensure highest quality and conformance on final packaging

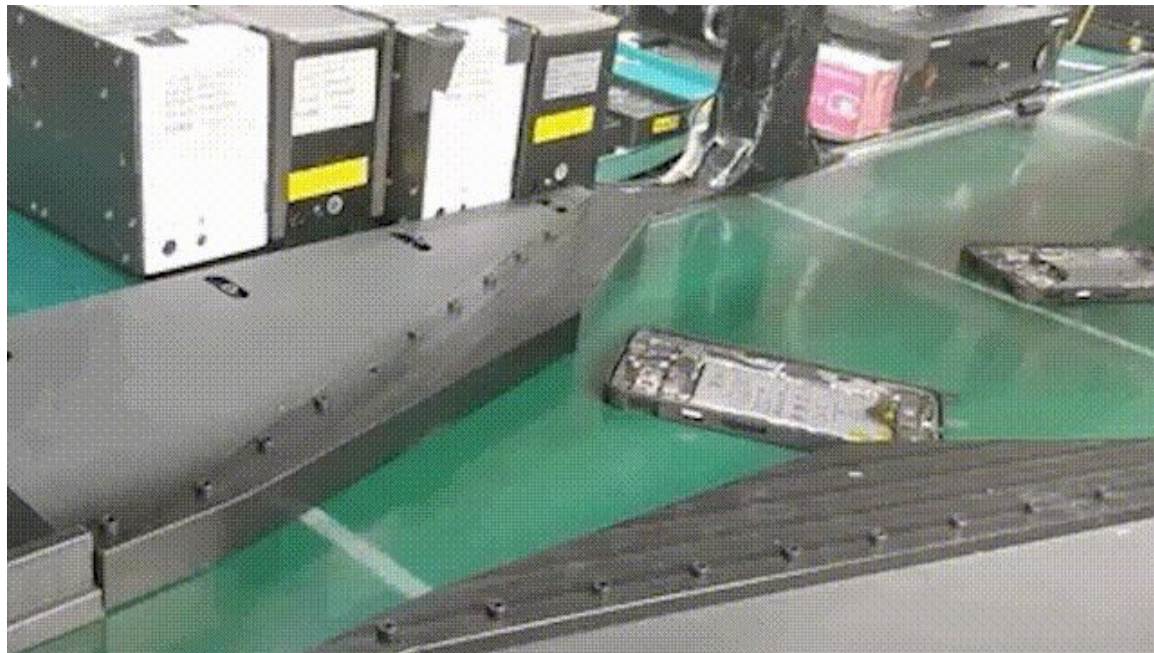
Introducing **Google Cloud Visual Inspection AI**: A solution to transform your manufacturing quality



Top-ranked **Computer Vision** and **Machine Learning** capabilities of Google
to **detect tiniest defects** at an unprecedented **accuracy** and **scale**.

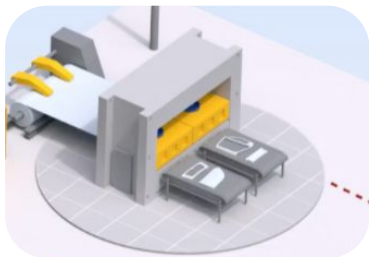
Example: Detecting defects in real-time at a mobile-phone production line

Visual Inspection AI
transform
manufacturing quality



Key benefits of Google Cloud Visual Inspection AI

Purpose-built for production scale



In production trials, **customers improved accuracy by up to 10 times** than general purpose ML platforms

Leading computer vision & AI



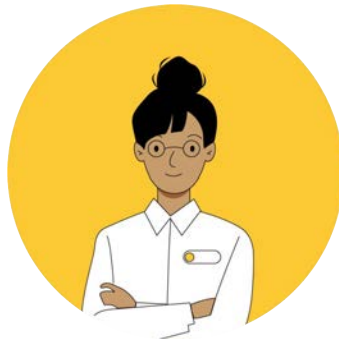
Google Brain



DeepMind

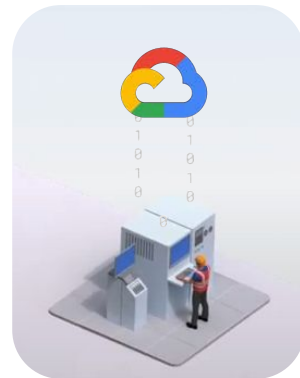
Unlike competing solutions that use simple anomaly detection: **detect, classify, and precisely locate multiple defect types** in a single image

Start quickly



Build accurate models with **up to 300 times fewer human labelled images** than general purpose ML platforms (based on pilots)

Edge-to-Cloud



Integrate your **production quality data with your broader industrial data in Google Cloud** with our edge-to-cloud data solutions

Visual Inspection AI supports three different ways to detect defects

DEEP DIVE

Check for Anomaly

Detect defective pieces
because they appear
“abnormal”



Detect and Locate Defects

Detect and locate multiple
domain specific defects
(dents, cracks, fractures, etc.)

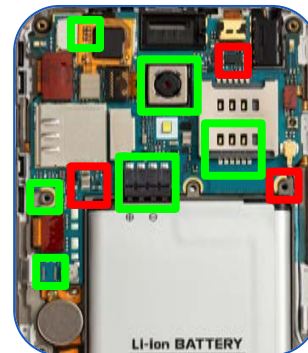


2 particles

3 cracks

Check Assembly

Detect missing and defect
parts within an assembled
product



2 connectors OK

7 solder pins OK

battery clip OK

...

Case screw NOK

Antenna pin NOK

...

Google Visual Inspection AI unlocks new use-cases for manufacturers

Check for Anomaly

Detect defective pieces because they appear “abnormal”

Detect and Locate Defects

Detect and locate multiple domain specific defects (dents, cracks, fractures, etc.)

Check Assembly

Detect missing and defect parts within an assembled product



Visual Inspection

Out of the box

Out of the box

Out of the box

Other cloud inspection offers

Out of the box

DIY

DIY

Use Cases dimostrativi

Public References

AI helps industries make smarter use of data

- Demand forecasting
- Product segmentation
- Fraud detection
- Inventory optimization
- Text summarization
- Risk management
- Defect detection
- Churn prediction
- Complaint categorisation
- Predictive maintenance
- Propensity modeling
- Personalised recommendations
- ...



varian

- Radiation Oncology software provider acquired by Siemens for \$16B
- Use case: Image Segmentation for Cancer Care
- GCP commit for Neural Architecture Search (NAS) service in Vertex AI ([press release](#))

Vertical: Healthcare

Region: AMER

Type: New customer

- ✓ Velocity of models in production
- ✓ Best-in-class algorithms as a service

“We were impressed by the speed and ease of building custom models. We were able to build some of the services, and see the initial results within a day of experimenting with the product. Going forward, we see value in having a unified workflow with all the tools and services needed by our developers and data scientists.”

Mike Kellet, CTO Small Cap Bank, Deutsche Bank



Deutsche Bank

Vertical: Financial Services

Region: EMEA

Type: New customer

- ✓ Velocity of models in production
- ✓ Best-in-class algorithms as a service

“Pipelines let us move faster from ML prototypes to production models, and give us confidence that our ML infrastructure will keep pace with our transaction volume as we scale.”

Hannes Hapke, ML Engineer, Digits Financial, Inc



Vertical: Financial Services

Region: AMER

Type: New customer

- ✓ Velocity of models in production
- ✓ Scaling ML workflows



Thank you

Google Cloud