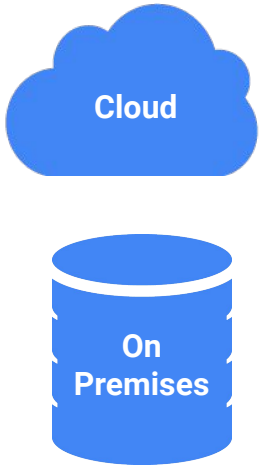
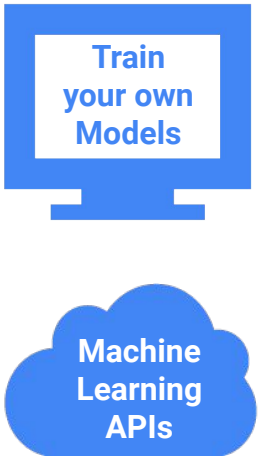
Data Warehouse is the foundation of something bigger

Data Warehouses/Lakes



Machine Intelligence



Predictive

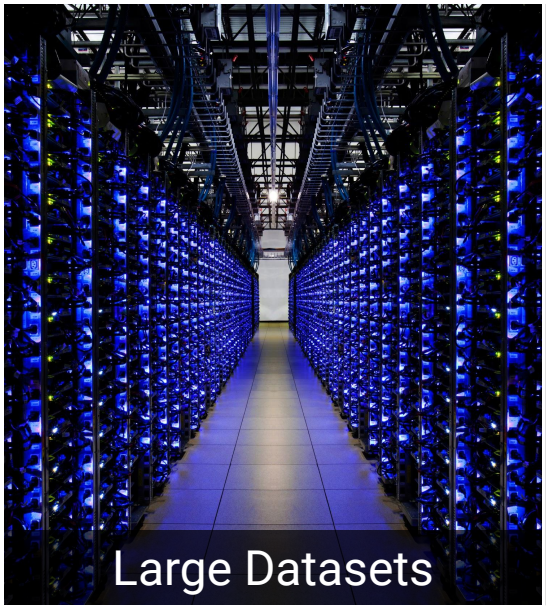
\+

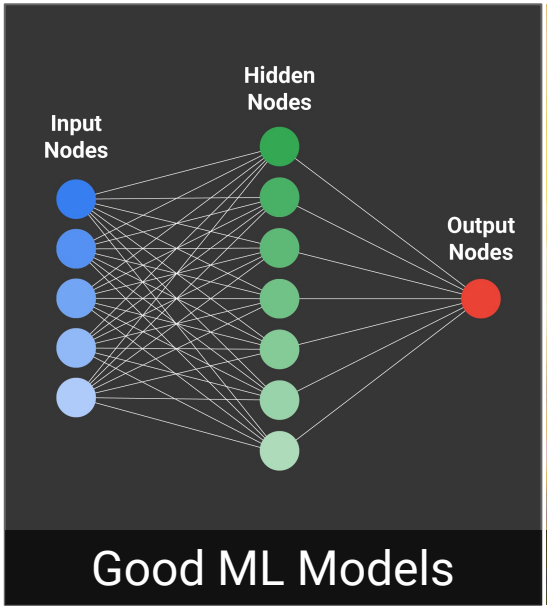
Prescriptive analytics

=

Advanced analytics

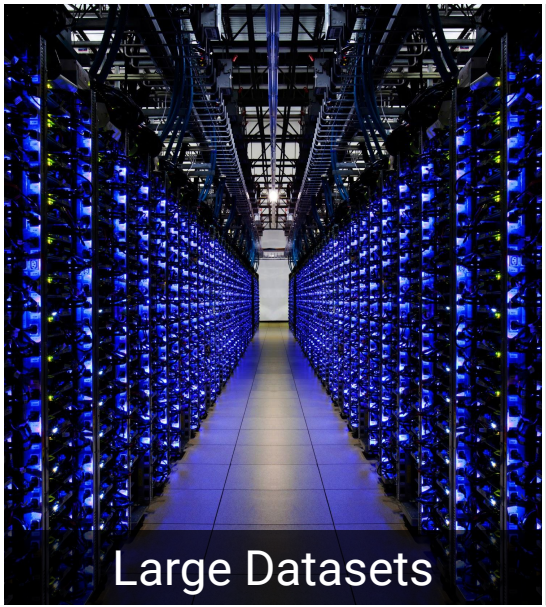
Keys to successful ML

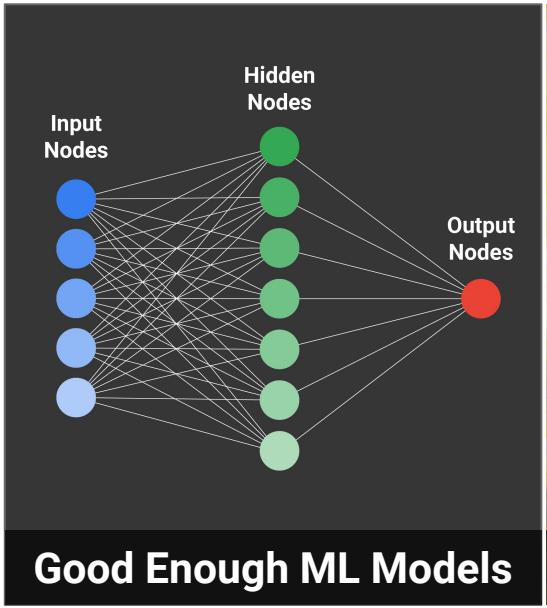




Google Cloud

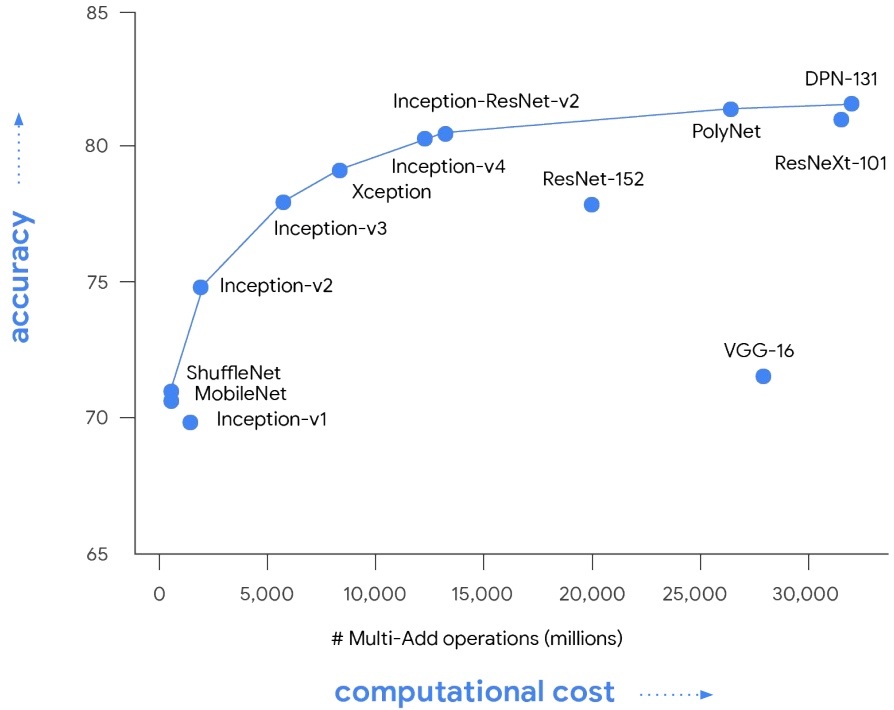
Keys to successful ML





Google Cloud

Increasing complexity and compute needs



Google Cloud

Current solutions don’t provide enterprises with the computing power they need

80%

of recent AI advances can be attributed to more available computer power[1]

Attendees at recent TensorFlow events admitted their computer power is limited[2]

1. Dileep George, Cofounder of the machine learning startup, Vicarious
2. TensorFlow Fall Symposium 2016, TensorFlow Dev Summit 2017

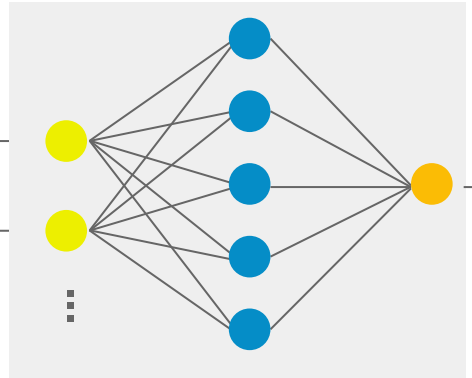
Google Cloud


TensorFlow

The point of ML is to make predictions

	mintemp	maxtemp	rain	day_1	day_2	day_3	day_4	day_5	day_6	day_7
104	28.9	37.9	0.01	0	0	0	1	0	0	0
9	32.0	43.0	0.00	0	1	0	0	0	0	0
114	35.1	48.0	0.00	1	0	0	0	0	0	0
11	30.0	37.9	0.00	0	0	0	0	0	0	0
316	19.0	39.9	0.05	0	0	1	0	0	0	0

Input Feature

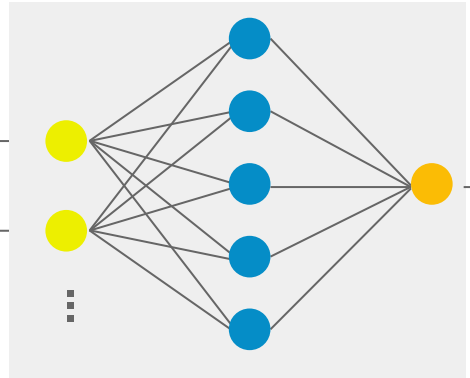


Predicted Value

We write model which makes predictions

	mintemp	maxtemp	rain	day_1	day_2	day_3	day_4	day_5	day_6	day_7
104	28.9	37.9	0.01	0	0	0	1	0	0	0
9	32.0	43.0	0.00	0	1	0	0	0	0	0
114	35.1	48.0	0.00	1	0	0	0	0	0	0
11	30.0	37.9	0.00	0	0	0	0	0	0	0
316	19.0	39.9	0.05	0	0	1	0	0	0	0

Input Feature

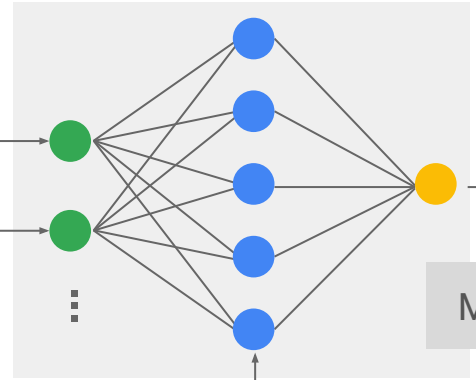


Predicted Value

Tensorflow helps you write the models

	mintemp	maxtemp	rain	day_1	day_2	day_3	day_4	day_5	day_6	day_7
104	28.9	37.9	0.01	0	0	0	1	0	0	0
9	32.0	43.0	0.00	0	1	0	0	0	0	0
114	35.1	48.0	0.00	1	0	0	0	0	0	0
11	30.0	37.9	0.00	0	0	0	0	0	0	0
316	19.0	39.9	0.05	0	0	1	0	0	0	0

Input Feature



Model

Predicted Value

Cost

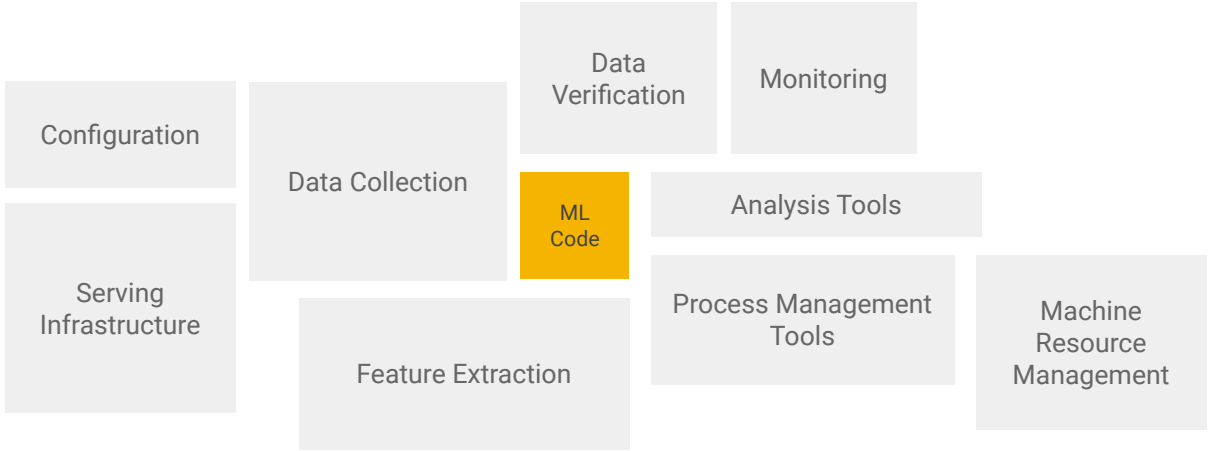
True Value

Update model based on Cost

ML
Code

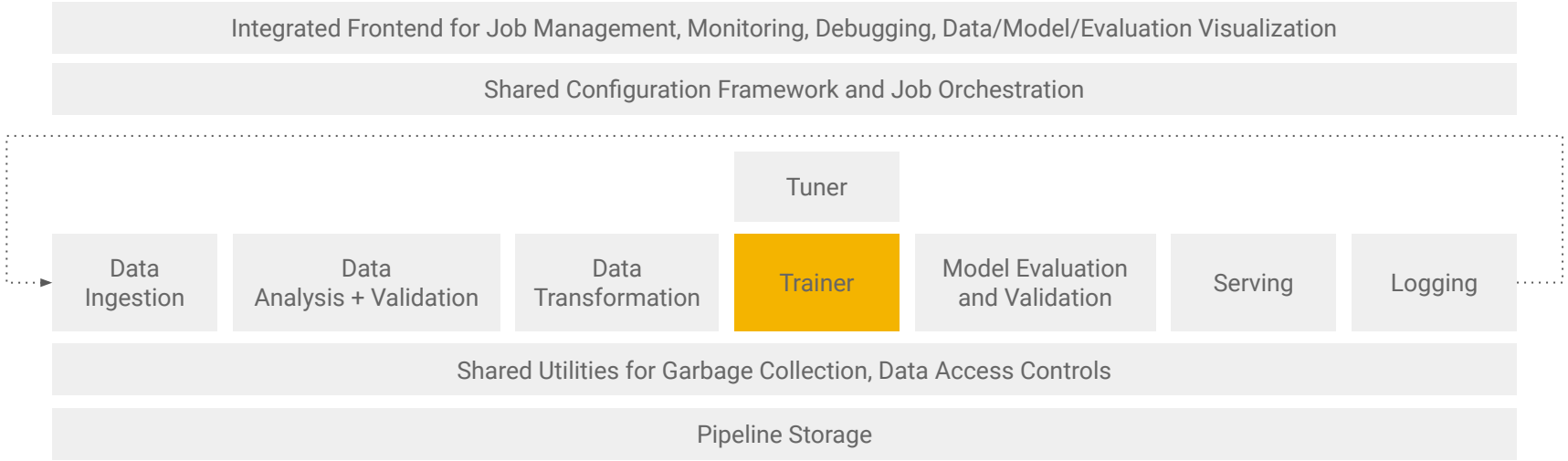
In addition to the actual ML...

...you have to worry about so much more.

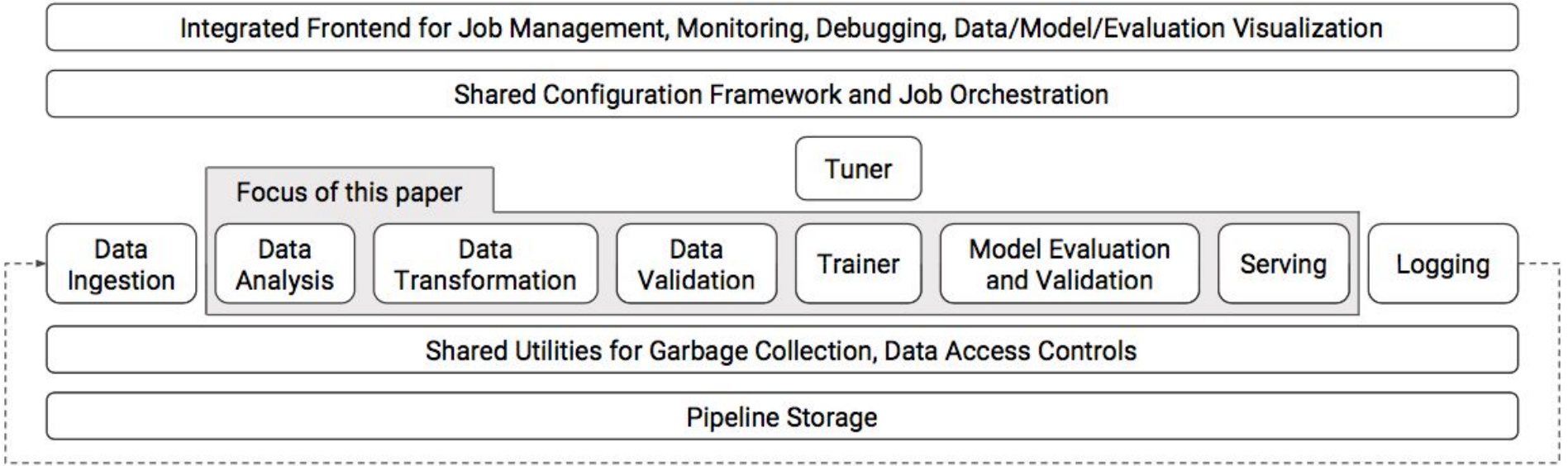


Source: Sculley et al.: Hidden Technical Debt in Machine Learning Systems

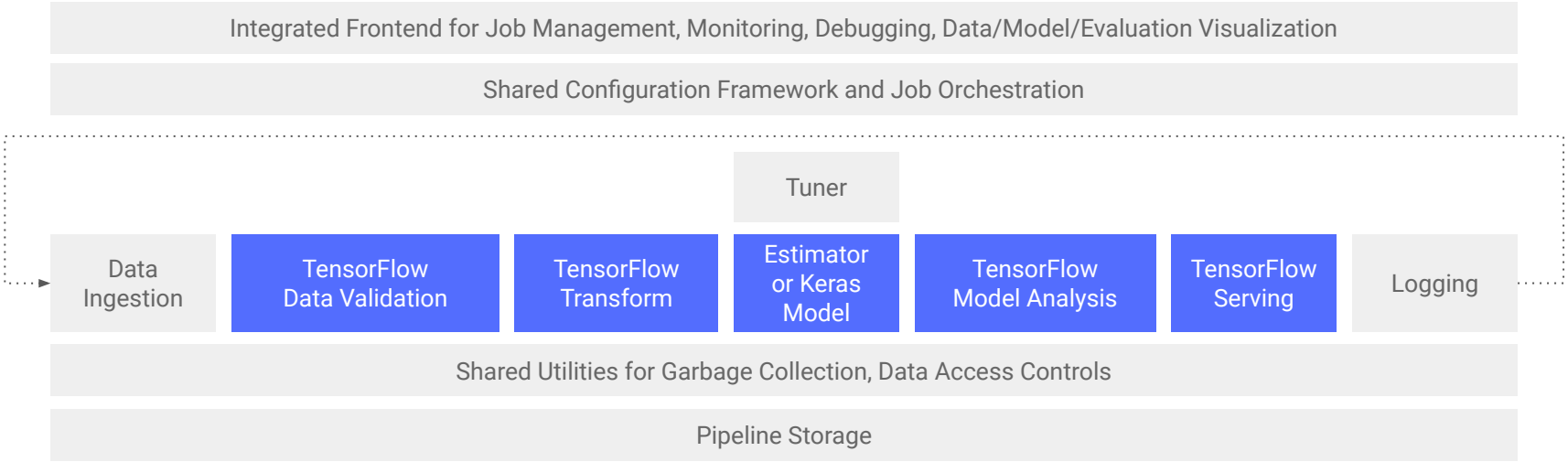
TFX is the solution to this problem...



TFX is the solution to this problem...



…and we are making it available to you.

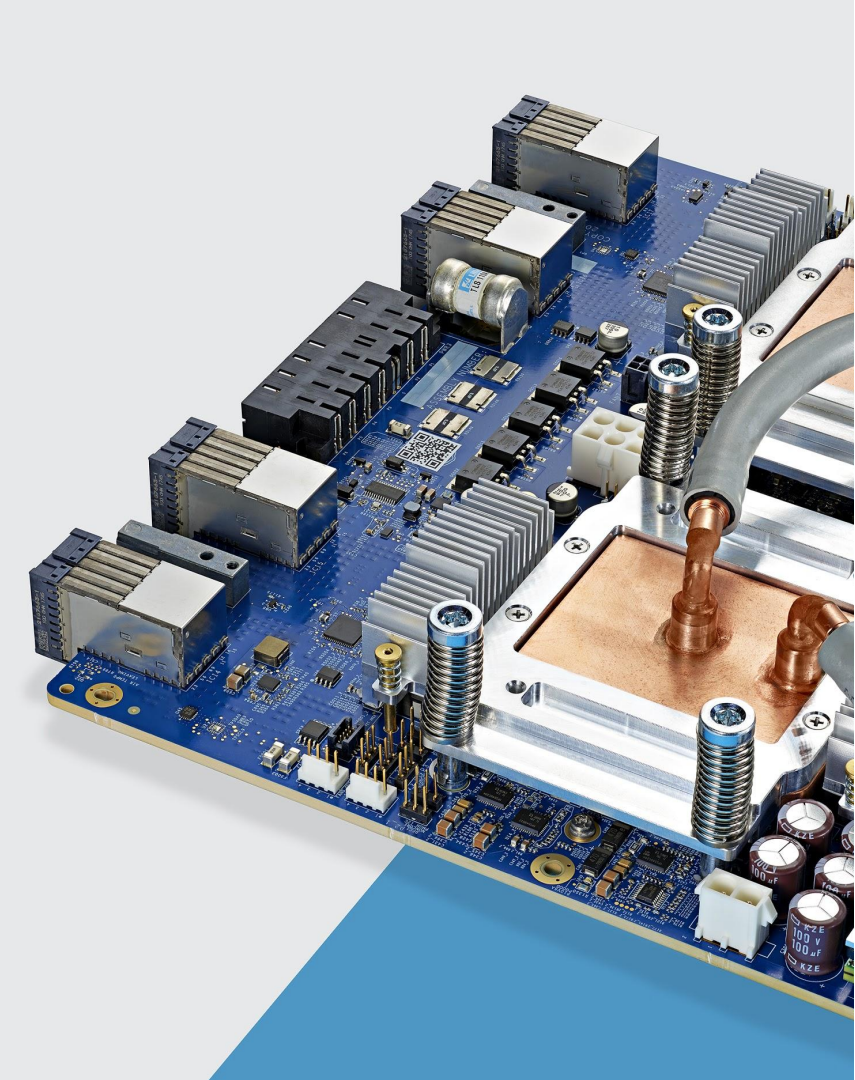


Cloud TPU

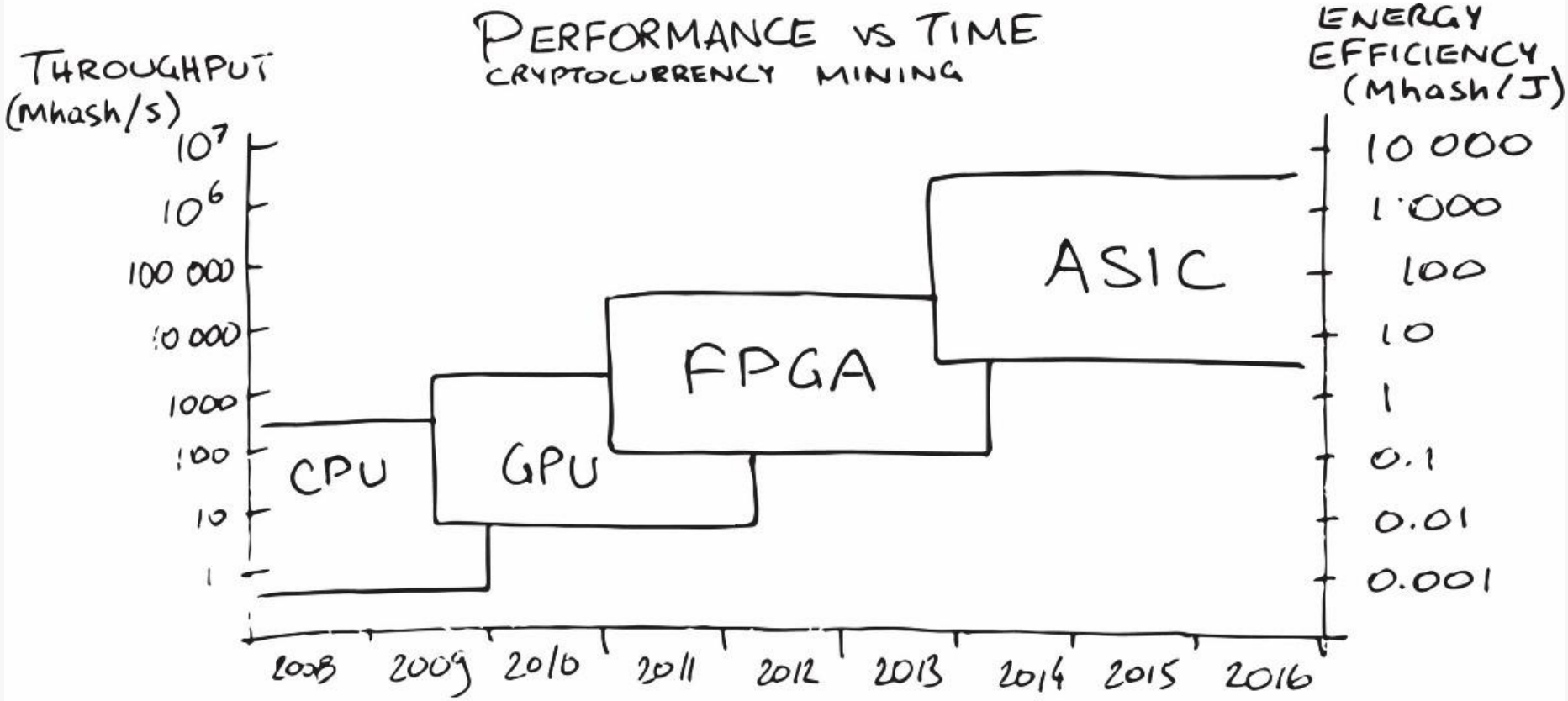
Custom ASIC by Google to train and execute deep neural networks

- Built for AI on Google Cloud
- Fast, iterative development
- Offers proven, Google-qualified reference models, optimized for performance, accuracy, and quality

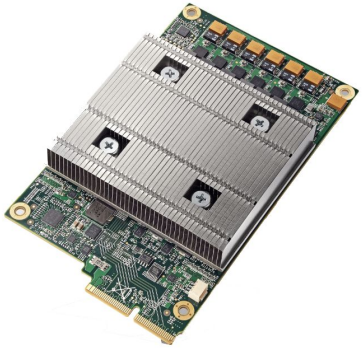
Google Cloud



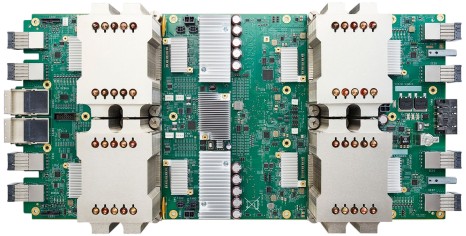
Cloud TPU is ASIC



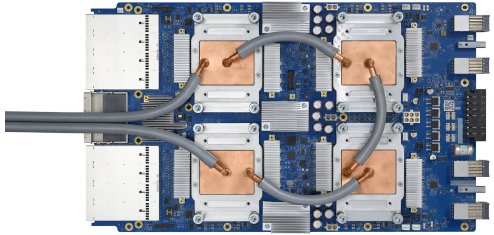
Expanding the AI frontier of performance



TPU v1
(2015)
92 teraops
First Generation



TPU v2
(2017)
180 teraflops
Available via Google Cloud



TPU v3
(2018)
420 teraflops
Available via Google Cloud

Google Cloud

Power of Machine Learning on GCP

1. Tensorflow Library
    a. Scale for Data & Compute
    b. Distributed data & training
    c. Monitor training visually with TensorBoard
2. Custom hardware for Tensorflow
    a. Speed & Cost
3. Cloud Engine - Built for experimentation
4. Rest of Cloud

Three ways for AI on Google Cloud

MY data + MY model


Cloud TPUs


Compute Engine


Cloud Dataproc

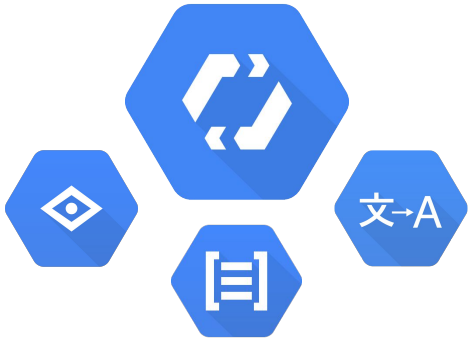

Kubernetes Engine


Cloud ML Engine


BigQuery ML

MY data + Google's models

AutoML



Google's data + Google's models


Cloud Translation API


Cloud Vision API


Cloud Speech API


Cloud Video Intelligence API


Data Loss Prevention API


Cloud Speech Synthesis API


Cloud Natural Language API


Dialogflow

Customisation

Ease of Use

Build your own models

Train our state-of-the-art models

Call our perception APIs

Three ways for AI on Google Cloud

MY HUGE **data** + MY **model**


Cloud TPUs


Compute Engine


Cloud Dataproc

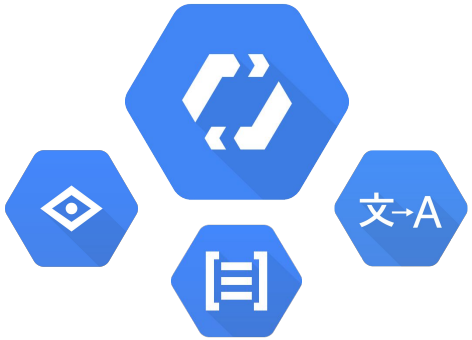

Kubernetes Engine


Cloud ML Engine


BigQuery ML

MY subset **data** + Google's **models**

AutoML



Google's **data** + Google's **models**


Cloud Translation API


Cloud Vision API


Cloud Speech API


Cloud Video Intelligence API


Data Loss Prevention API


Cloud Speech Synthesis API

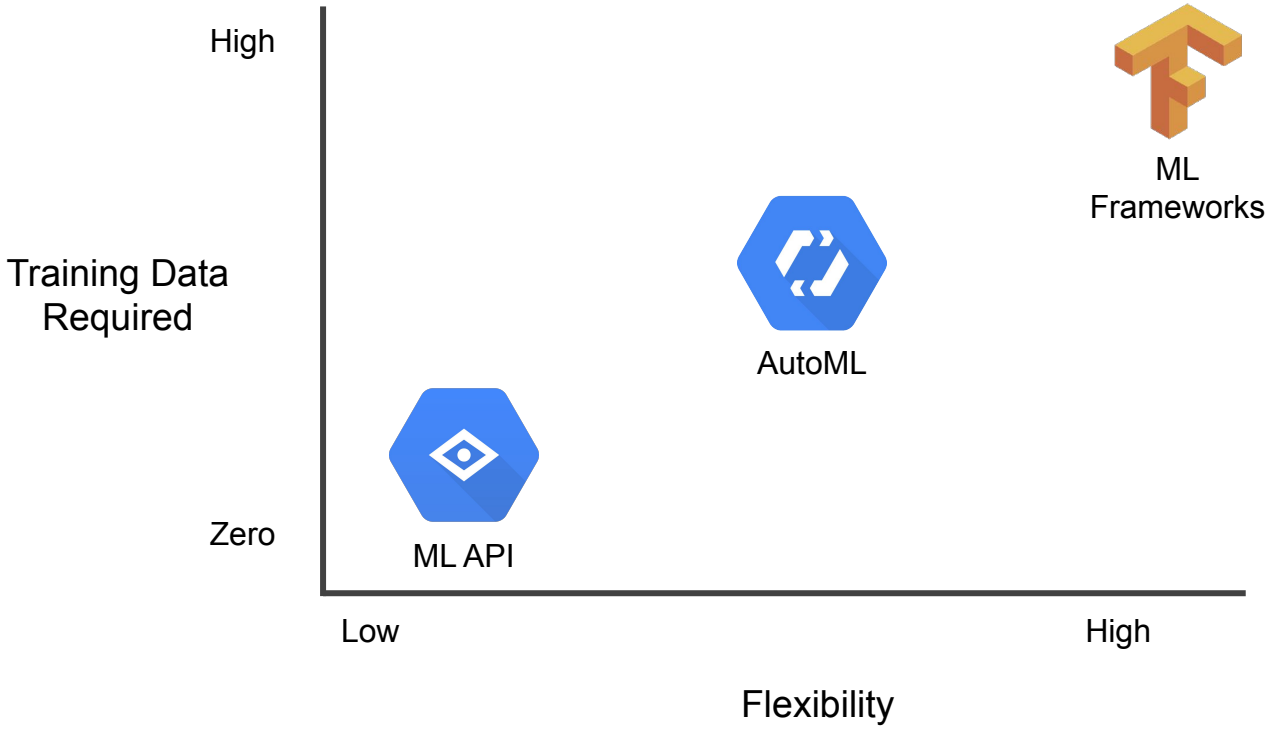

Cloud Natural Language API


Dialogflow

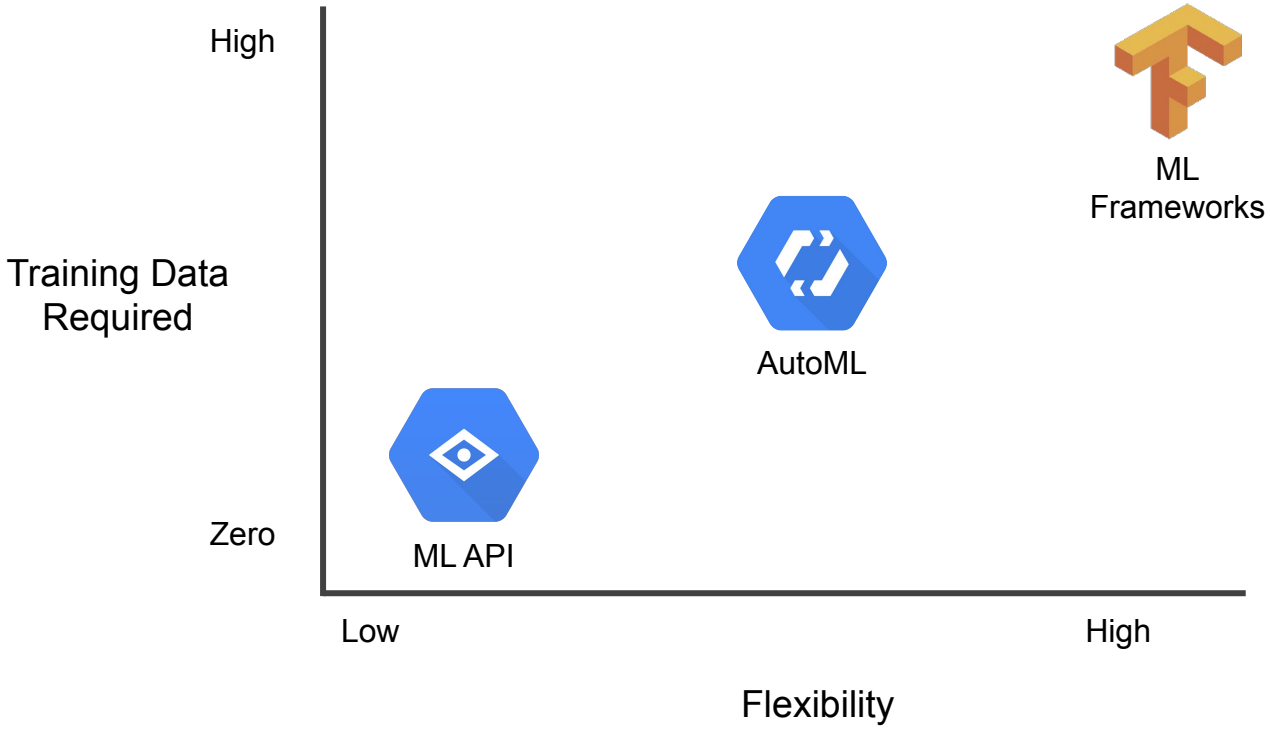
Customisation — Ease of Use

Build your own models — Train our state-of-the-art models — Call our perception APIs

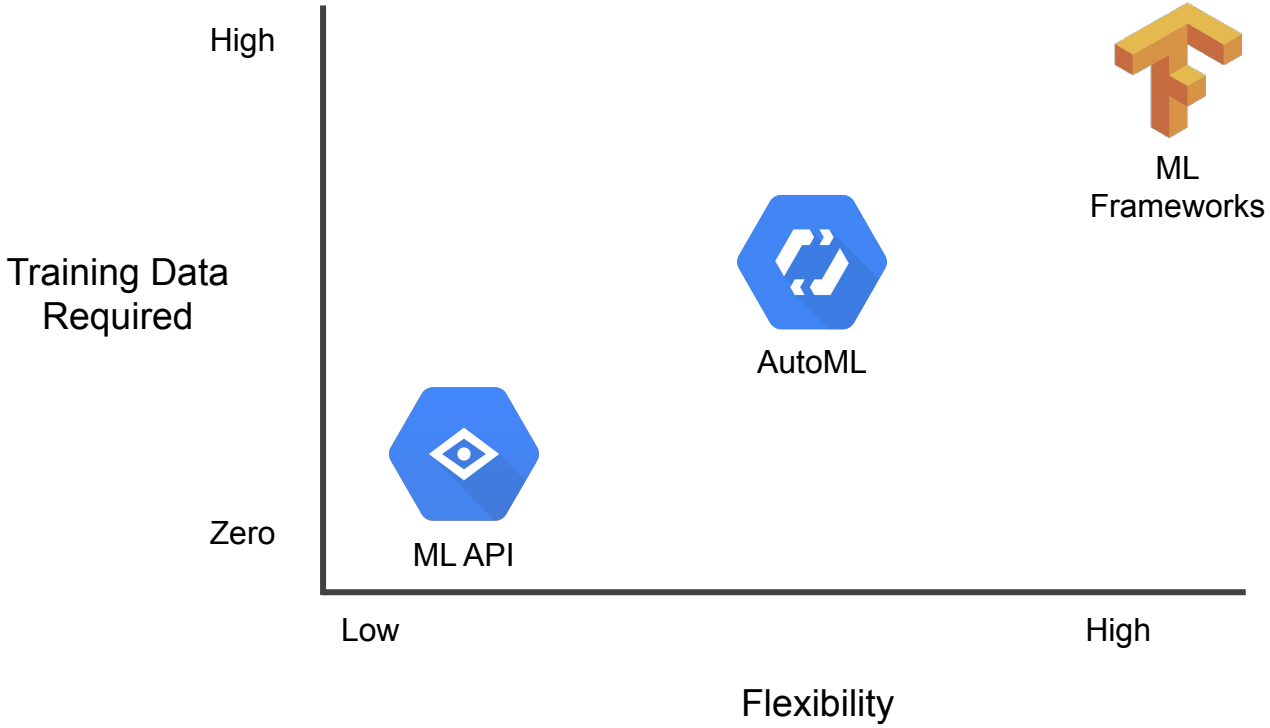
Training Data Required vs Flexibility



Training Time Required vs Flexibility



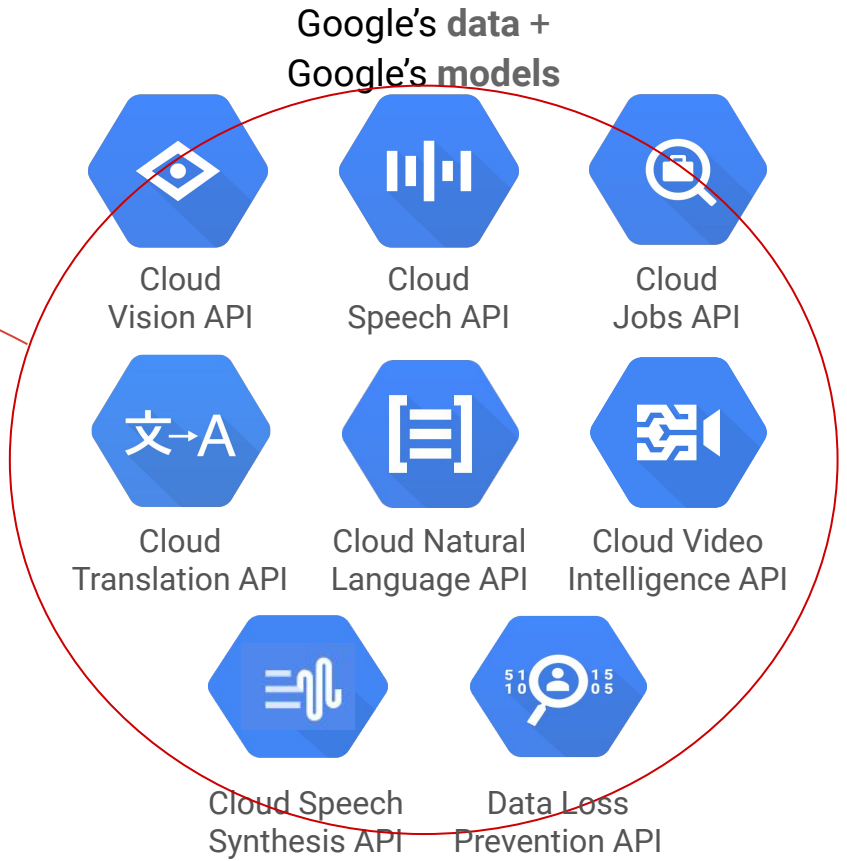
Time for Model Dev vs Flexibility




The Rule #1 of Machine Learning Club is:
You don’t build a model if you don’t have to

Use our pre-trained perception APIs

Use straight away in your apps, or integrate in ETL




Cloud Vision API
Insight from images with our powerful
Cloud Vision API

Cloud Vision API

Cloud Vision API **quickly classifies images into thousands of categories** (e.g. "sailboat", "lion", "Eiffel Tower"), detects individual objects and faces within images, and finds then reads printed words contained within images in many languages and scripts.

With Cloud Vision API, you can:

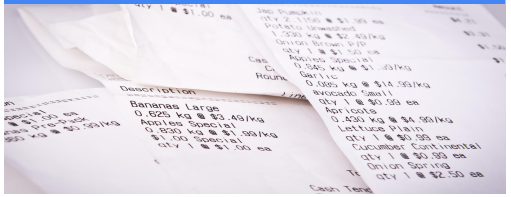
- **Understand the content of an image**
- **Get value from images**
- **Easily build apps using those data**

Vision API is available through an easy to use REST API to analyze images stored anywhere, or integrate with your image storage on Google Cloud Storage.

Faces
Faces, facial landmarks, emotions

Label
Detect entities from furniture to transportation

OCR
Read and extract text, with support for > 10 languages



Logos
Identify product logos

Safe Search
Detect explicit content - adult, violent, medical and spoof

Landmarks & Image Properties
Detect landmarks & dominant color of image

Cloud Vision API

Call API from anywhere, with support for embeddable images, and Google Cloud storage

Cloud Speech API

Speech to text conversion

Features

Automatic Speech Recognition

Automatic Speech Recognition (ASR) powered by deep learning neural networking to power your applications like voice search or speech transcription.

Global Vocabulary

Recognizes over 80 languages and variants with an extensive vocabulary.

Streaming Recognition

Returns partial recognition results immediately, as they become available.

Inappropriate Content Filtering

Filter inappropriate content in text results.

Real-time or Buffered Audio Support

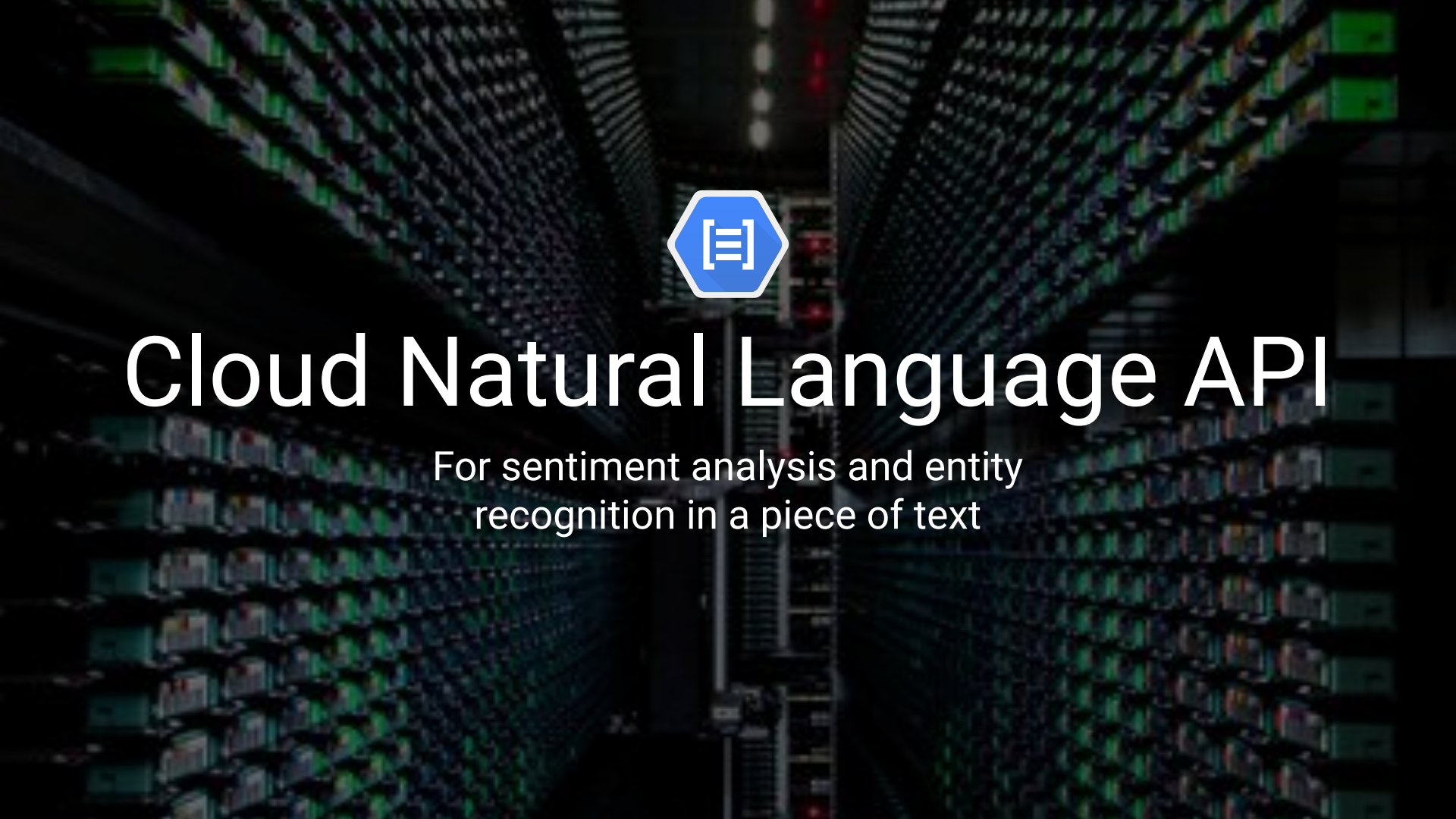
Audio input can be captured by an application's microphone or sent from a pre-recorded audio file. Multiple audio file formats are supported, including FLAC, AMR, PCMU and linear-16.

Noisy Audio Handling

Handles noisy audio from many environments without requiring additional noise cancellation.

Integrated API

Audio files can be uploaded in the request and, in future releases, integrated with Google Cloud Storage.



Cloud Natural Language API

For sentiment analysis and entity recognition in a piece of text

About Cloud Natural Language API

- Cloud Natural Language API lets businesses and developers **understand the structure and meaning of text across languages**
- With Natural Language API, you can **identify the most important entities** from people, location, events, along with the sentiment of the text
- **Ease of use:** Cross platform and easy to use via REST APIs, enabling access from any platform or device
- **Proven:** Cloud Natural Language API uses the same technology as Search, Google Now and Inbox-Smart Reply

Features

Syntax Analysis

Extract sentence, identify parts of speech and create dependency parse trees for each sentence.

Entity Recognition

Identify entities and label by types such as person, organization, location, events, products and media.

Sentiment Analysis

Understand the overall sentiment of a block of text.

Integrated REST API

Access via REST API. Text can be uploaded in the request or integrated with Google Cloud Storage.

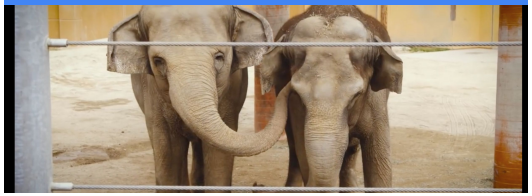

Cloud Video Intelligence
Analyze your video with Machine Learning

Cloud Video Intelligence API

- **Quickly understand video content** by encapsulating powerful machine learning models in an easy to use REST API.
- **Accurately annotate videos stored in Google Cloud Storage** with video and frame-level (1 fps) contextual information.
- Cloud Video Intelligence API enables businesses to **make sense of large amount of video files in a very short amount of time**.
- Available through an easy to use REST API to analyze videos stored anywhere, or integrate with your image storage on Google Cloud Storage.

Label Detection

Detect objects, such as dog, flower, human, in the video.



Face Detection

Detect faces throughout the length of the video.

Shot Change Detection

Detect scene changes within the video.

Video Segmentation

Segment long-running videos to provide annotations for specified time segments.

Integrated REST API

Request one or more annotation types per image.

Regionalization

Specify a region where processing will take place (for regulatory compliance).

Cloud Video Intelligence API

Call API from anywhere.


文→A
Cloud Translation API
Dynamically translate between
thousands of available language pairs

Cloud Translation API

Fast, Dynamic Translation

- Google Translation API provides a simple programmatic interface for translating an arbitrary string into any supported language.
- Translation API is highly responsive, so websites and applications can integrate with Translation API for fast, dynamic translation of source text from the source language to a target language (e.g., French to English).
- Language detection is also available in cases where the source language is unknown.

Features

Programmatic Access

Accessible via a standard Google REST API. See sample code and libraries for ten different programming languages including Python, Objective C and Ruby.

Text Translation

Supports more than 90 languages and thousands of language pairs.

Language Detection

Detect a document's language and translate it using a RESTful API.

Continuous Updates

Behind the scenes, Translation API is learning from logs analysis and human translation examples.

Adjustable Quota

Easily increase your quota from 2M characters per day to 50M per day or request a higher quota

Pre-Trained Machine Learning Models

Fully trained ML models from Google Cloud that allow a general developer to take advantage of rich machine learning capabilities with simple REST based services.

Cloud Translate

Cloud Vision

Cloud Natural Language

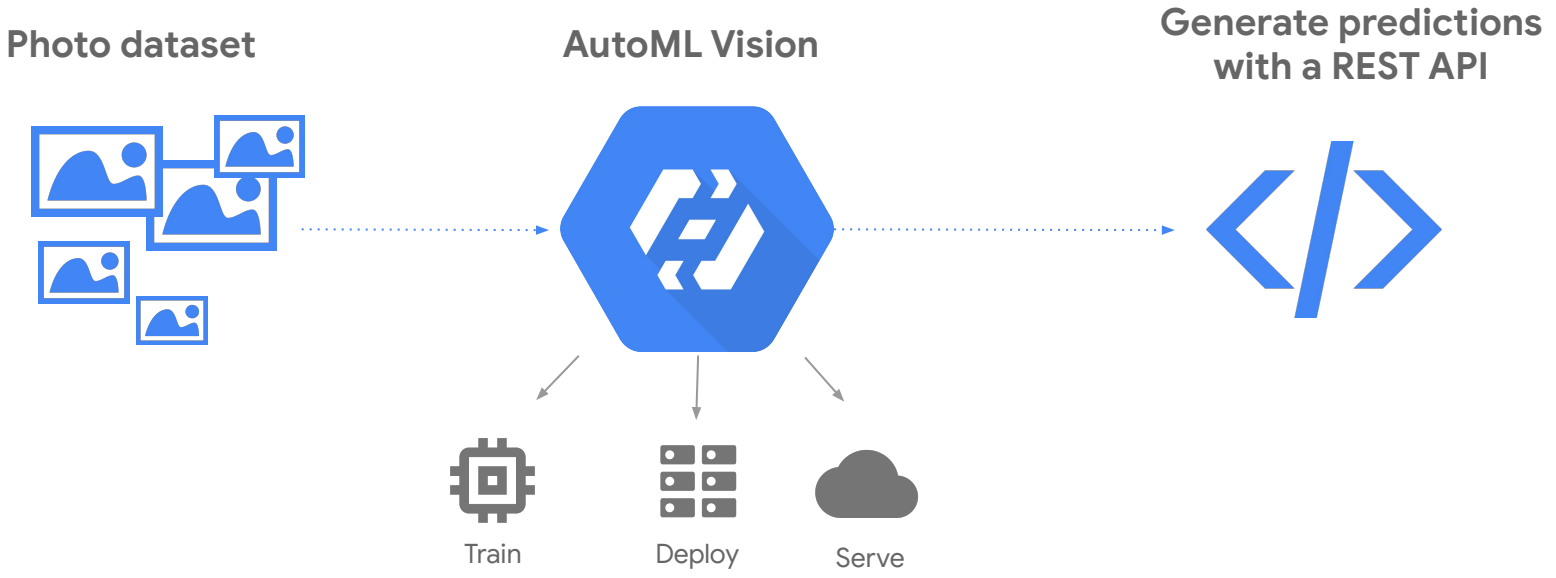
Cloud Speech

Cloud Video Intelligence

Cloud Jobs API

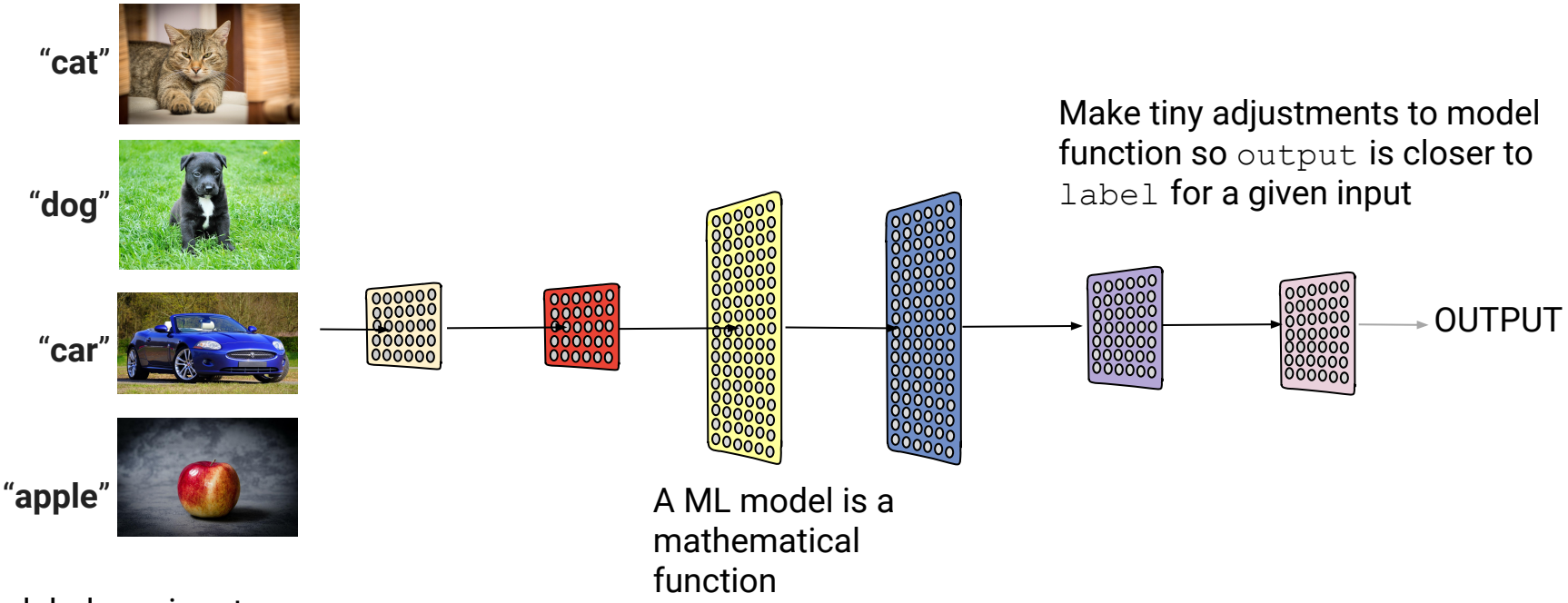
Stay tuned...

AutoML to the rescue

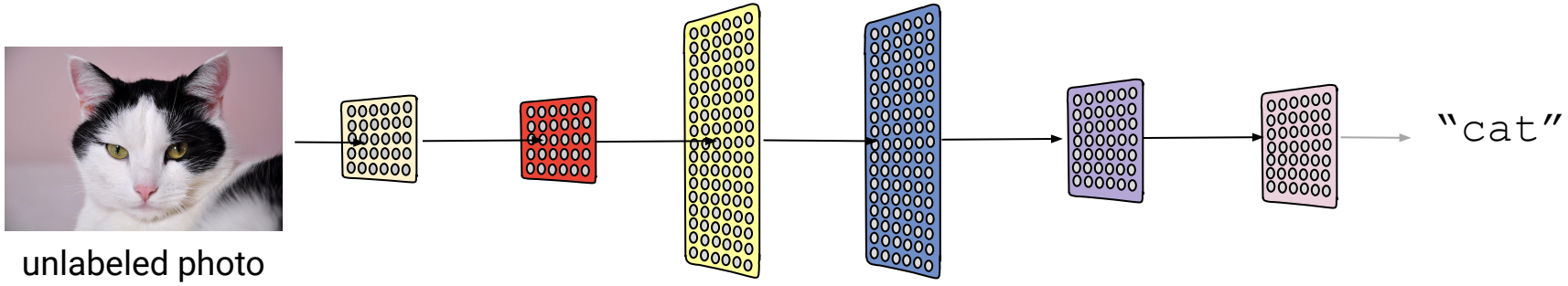


Google Cloud

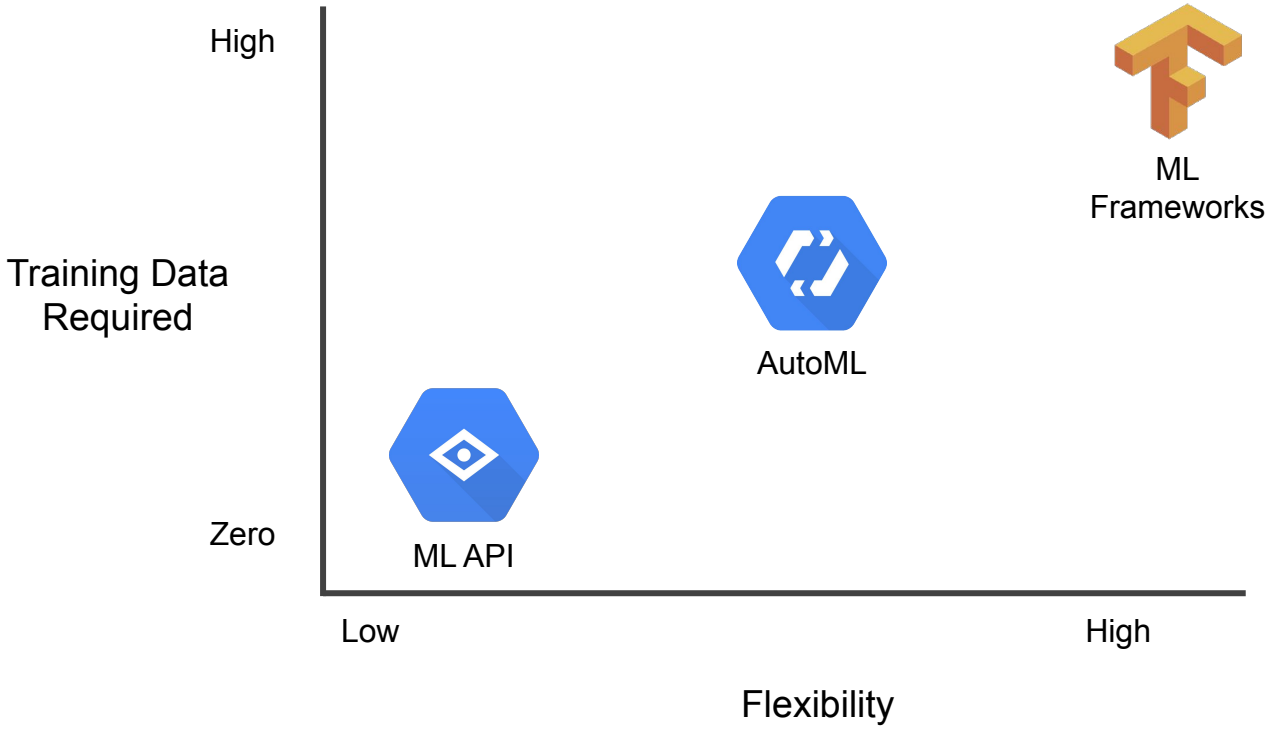
Stage 1: Train an ML model with examples



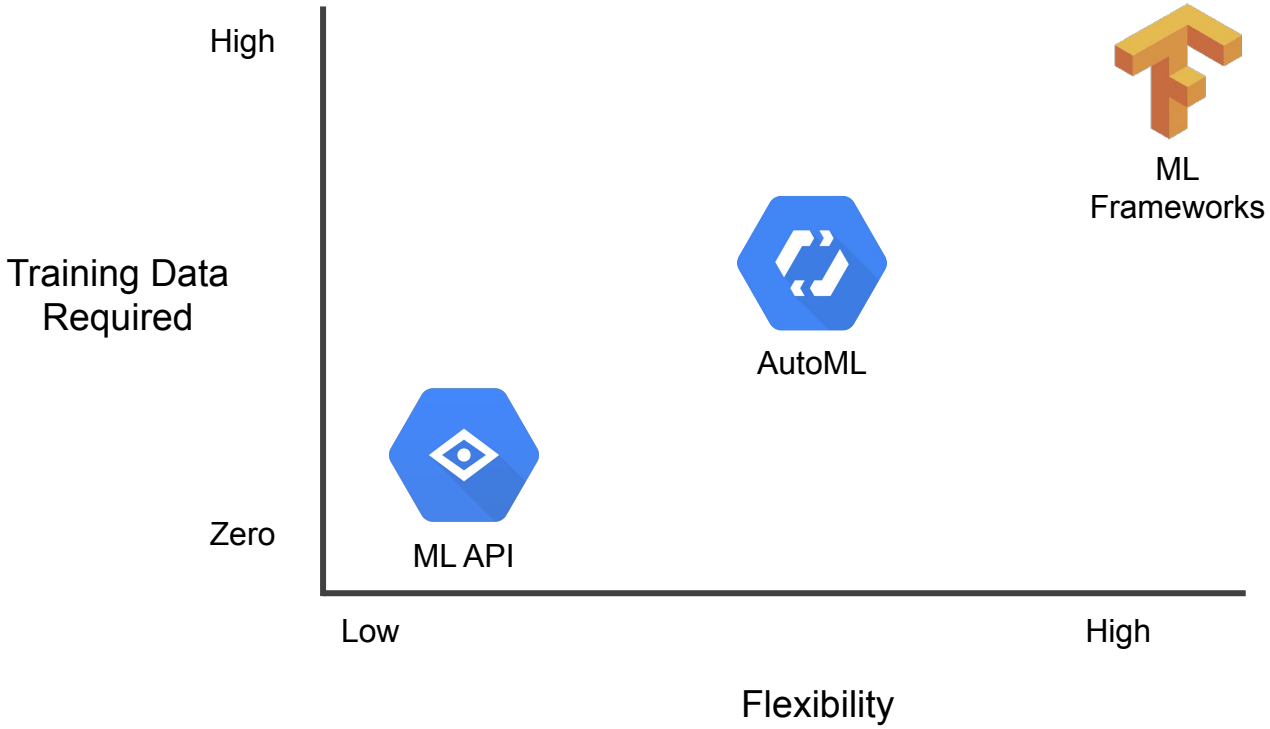
Stage 2: Predict with a trained model



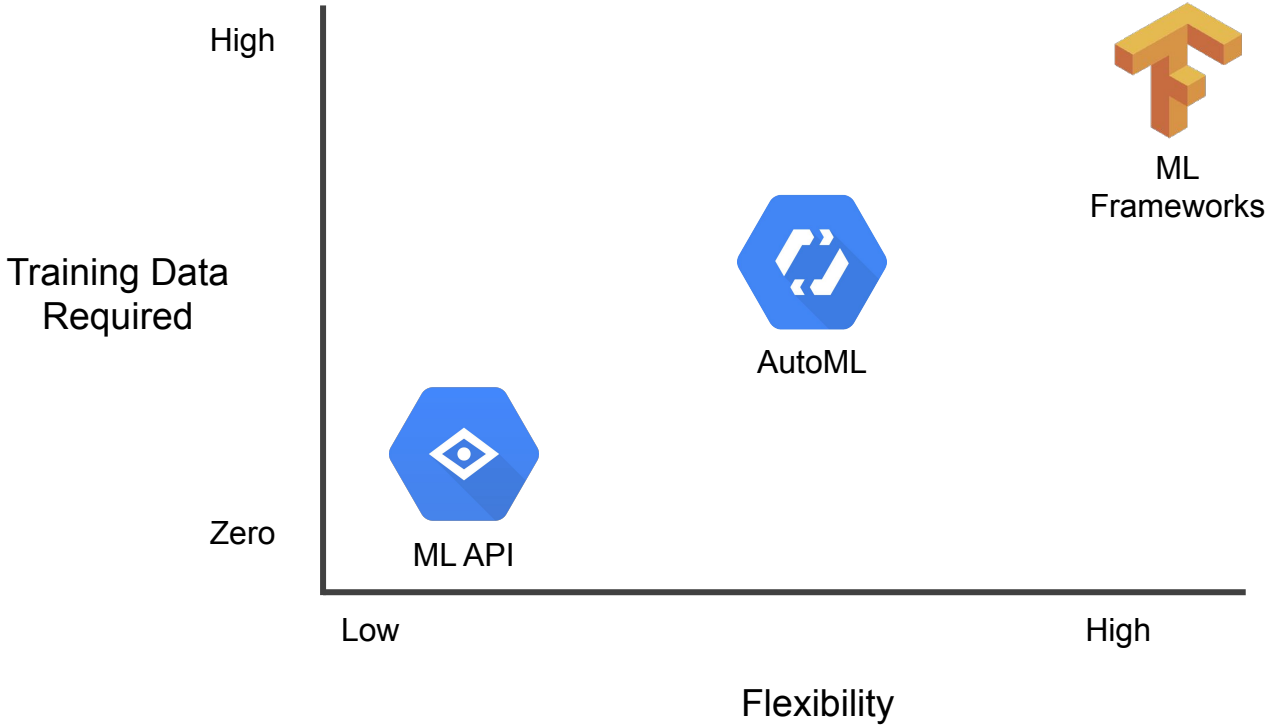
Training Data Required vs Flexibility



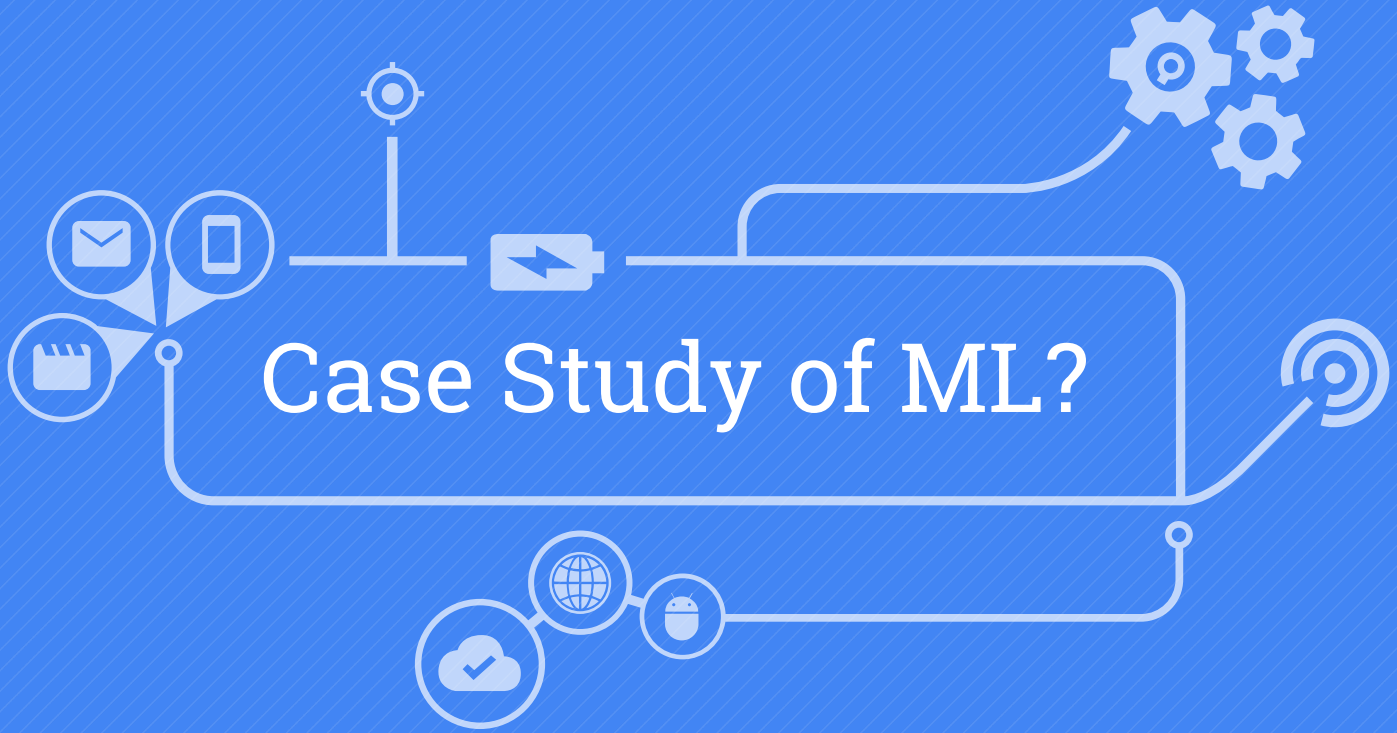
Training Time Required vs Flexibility



Time for Model Dev vs Flexibility



Case Study of ML?



Google

Let's say I'm a meteorologist ...

Google Cloud

I want to predict weather trends and flight plans from images

Google Cloud

Can we use the cloud to analyze clouds?

Google Cloud

There are 10+ different types of clouds

Google Cloud

There are 10+ different types of clouds

Google Cloud

There are 10+ different types of clouds

Google Cloud

There are 10+ different types of clouds

Google Cloud

There are 10+ different types of clouds

Google Cloud

They all indicate different weather patterns

Google Cloud

Let’s try the Vision API

Google Cloud

Let’s try the Vision API

Sky	99%
Cloud	98%
Daytime	97%
Blue	97%
Atmosphere	93%

Google Cloud

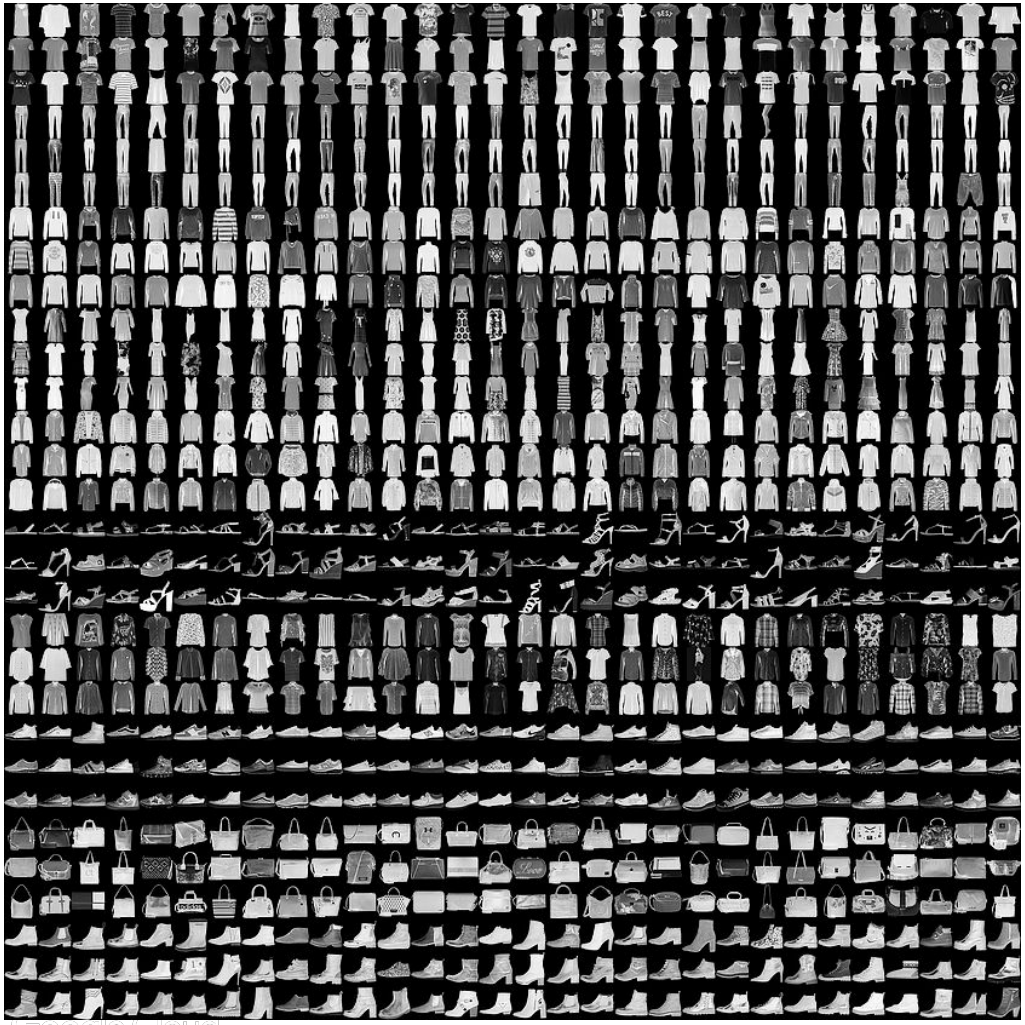
stom Models > Datasets > clouds

Dataset: clouds

IMPORT LABEL TRAIN EVALUATE PREDICT EXPORT

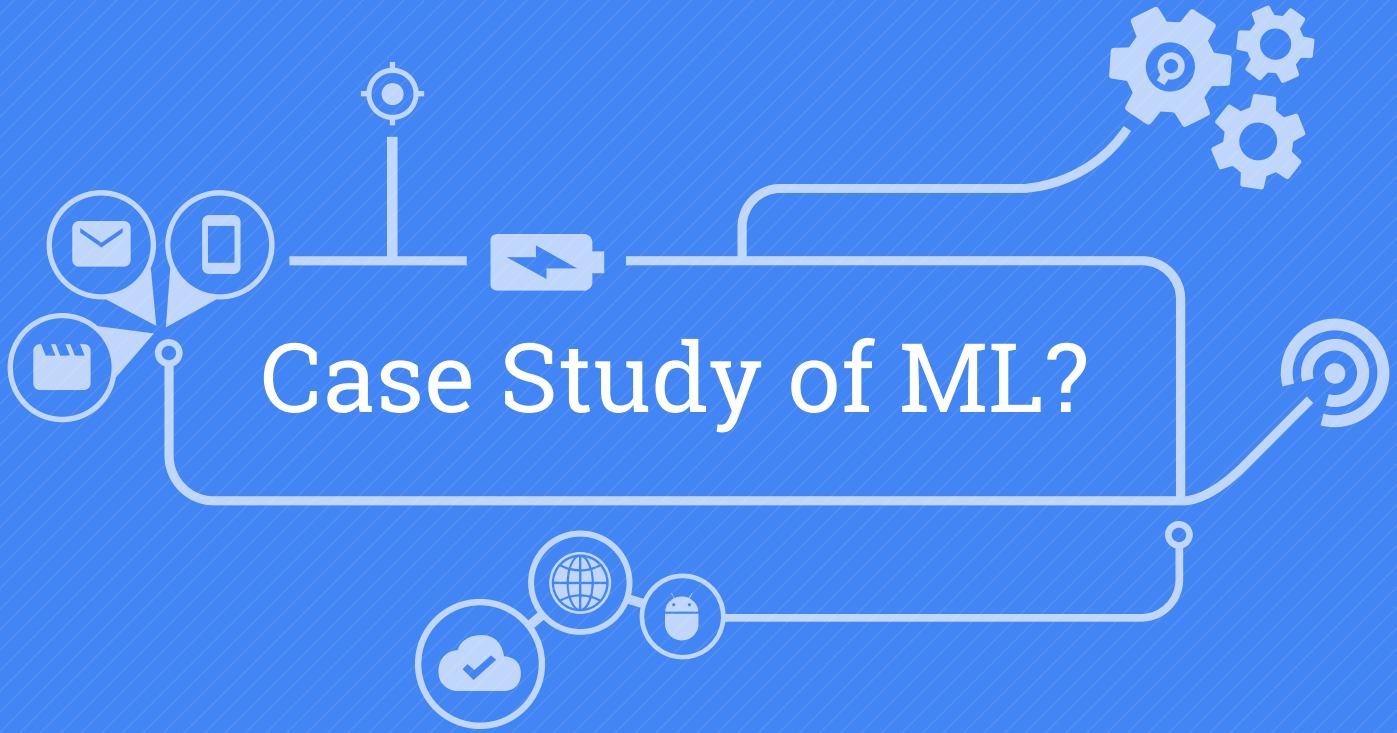
Query online

0.993 cirrus



https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/keras/basic_classification.ipynb

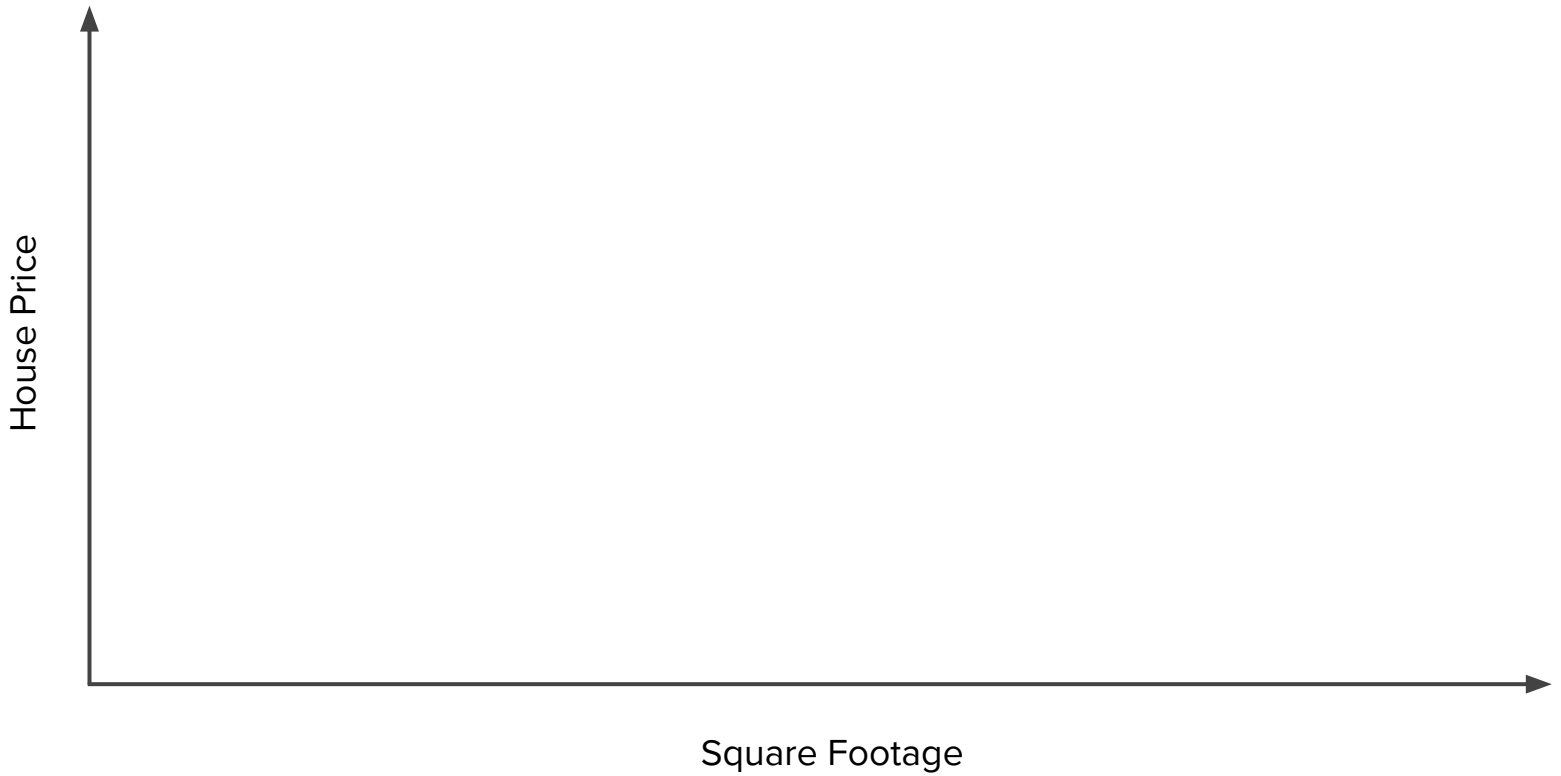
Case Study of ML?



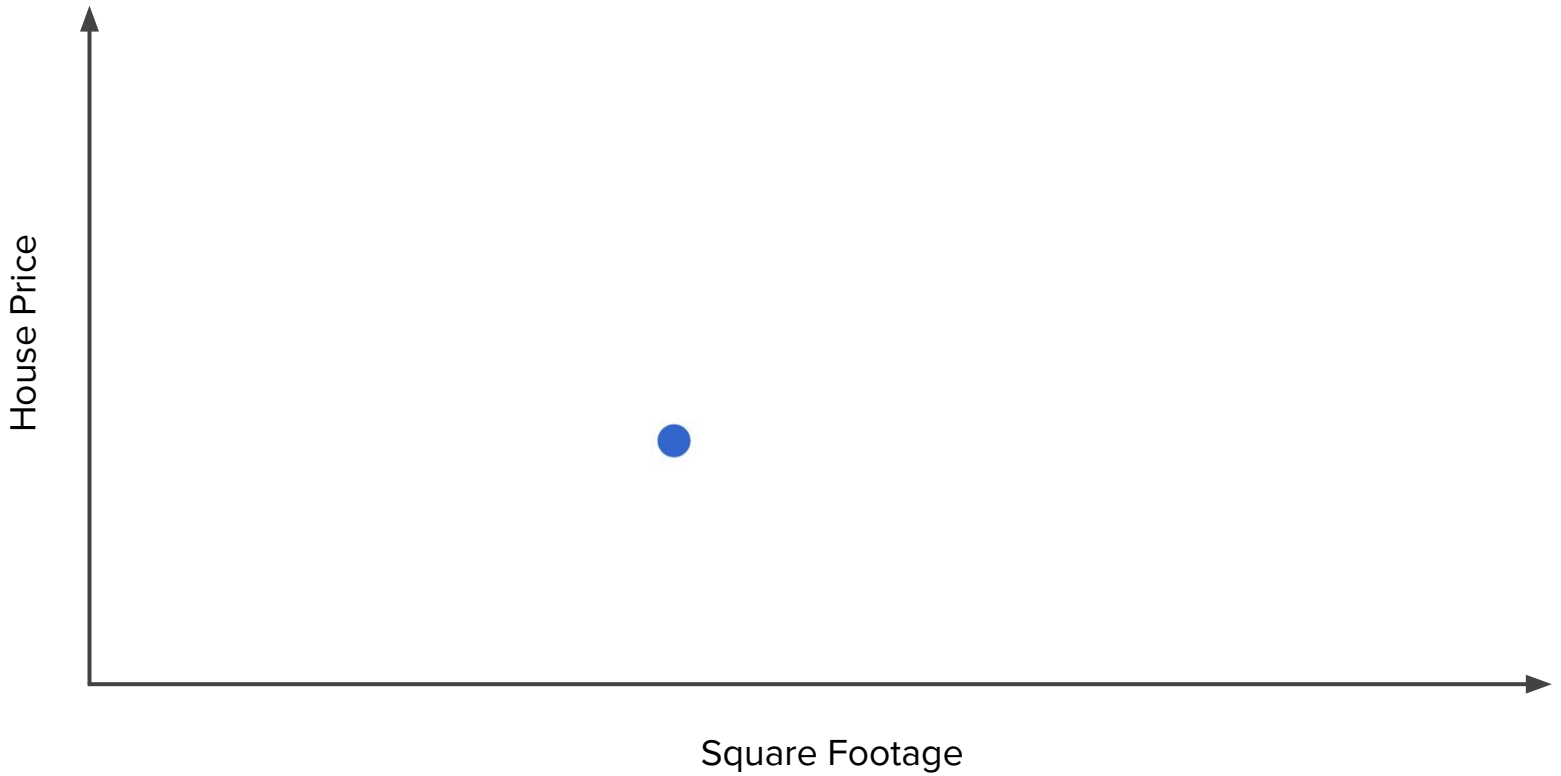
Google

How Much Is My House Worth?

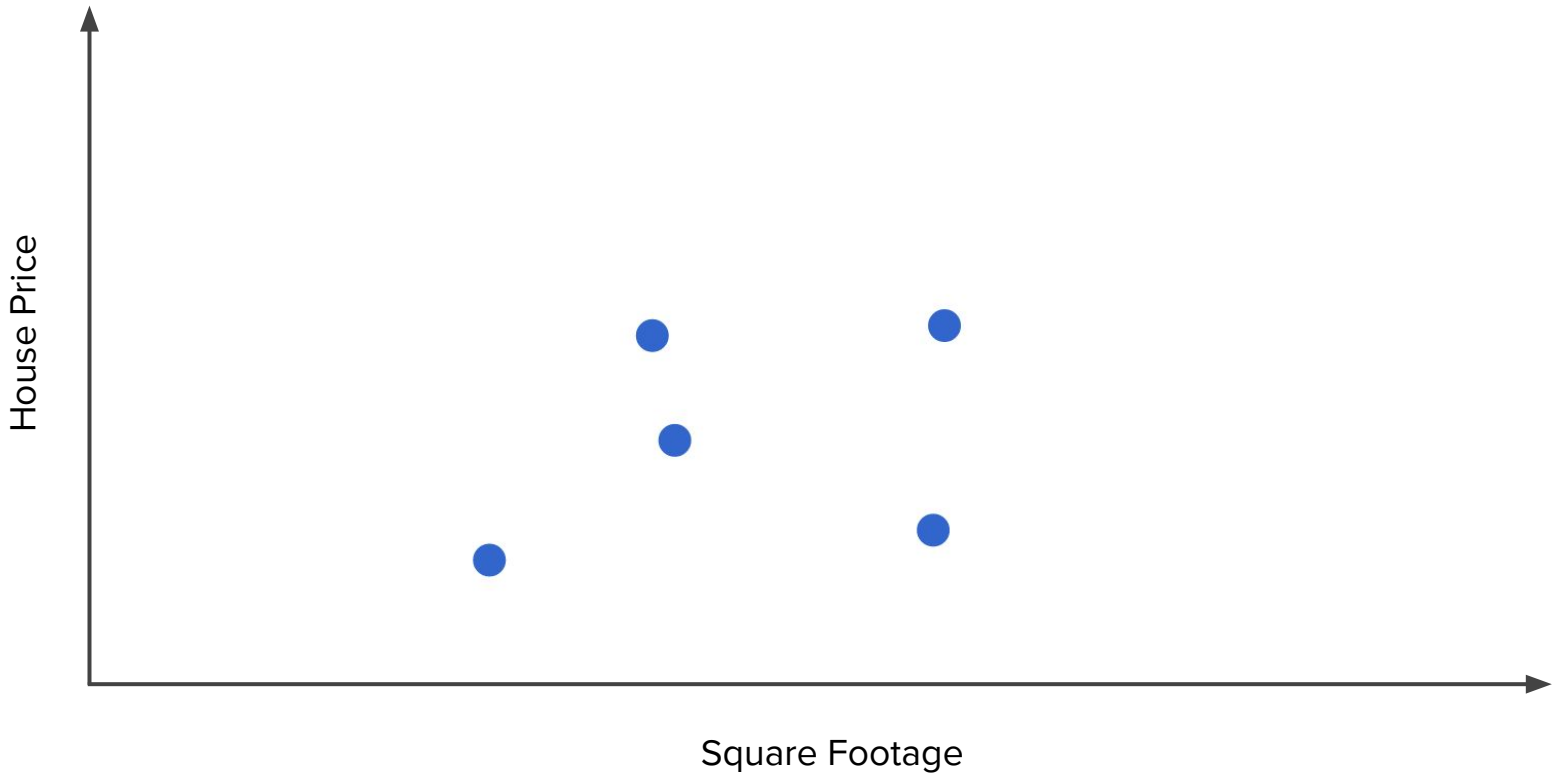
How Much Is My House Worth?



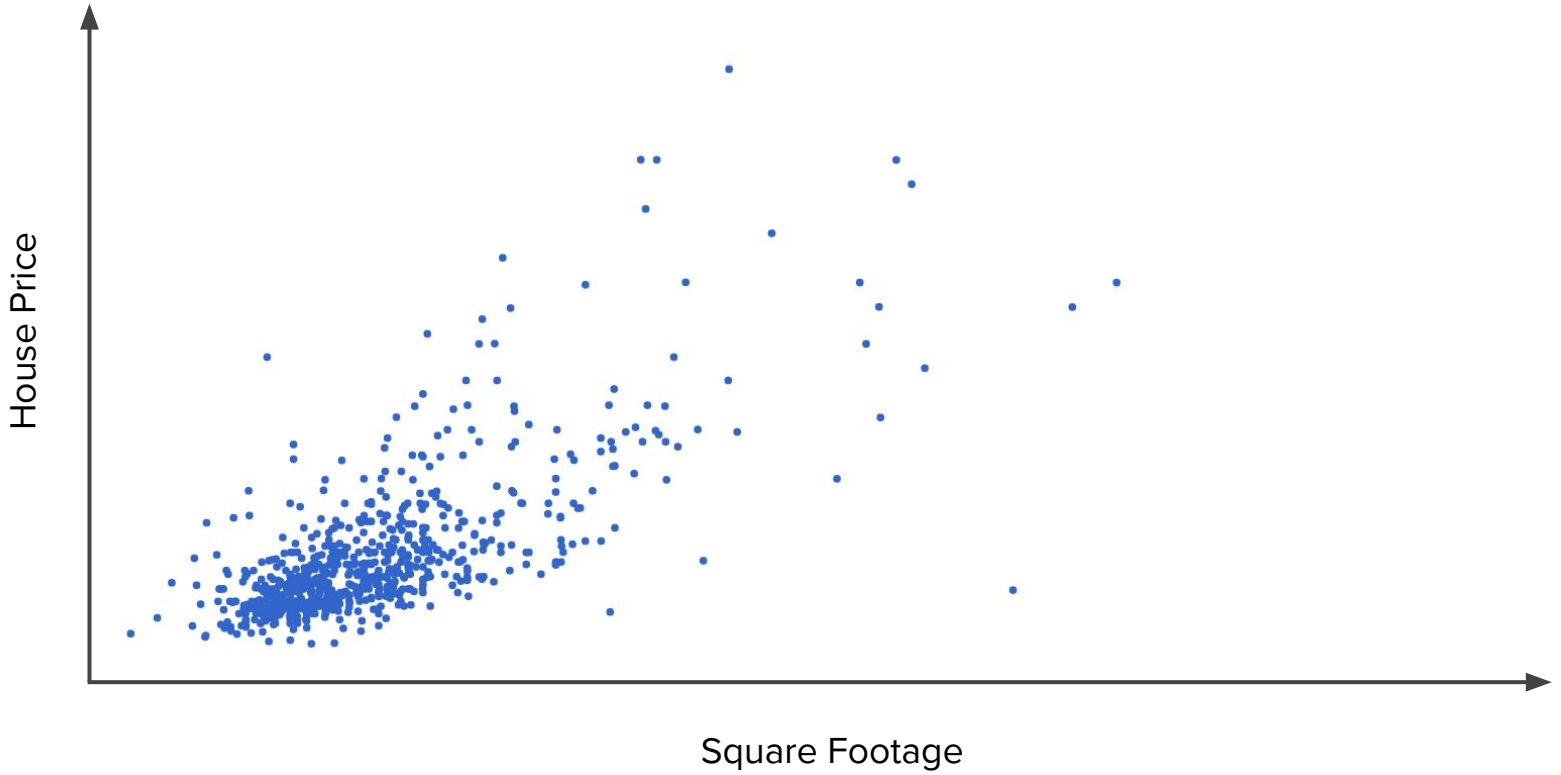
How Much Is My House Worth?



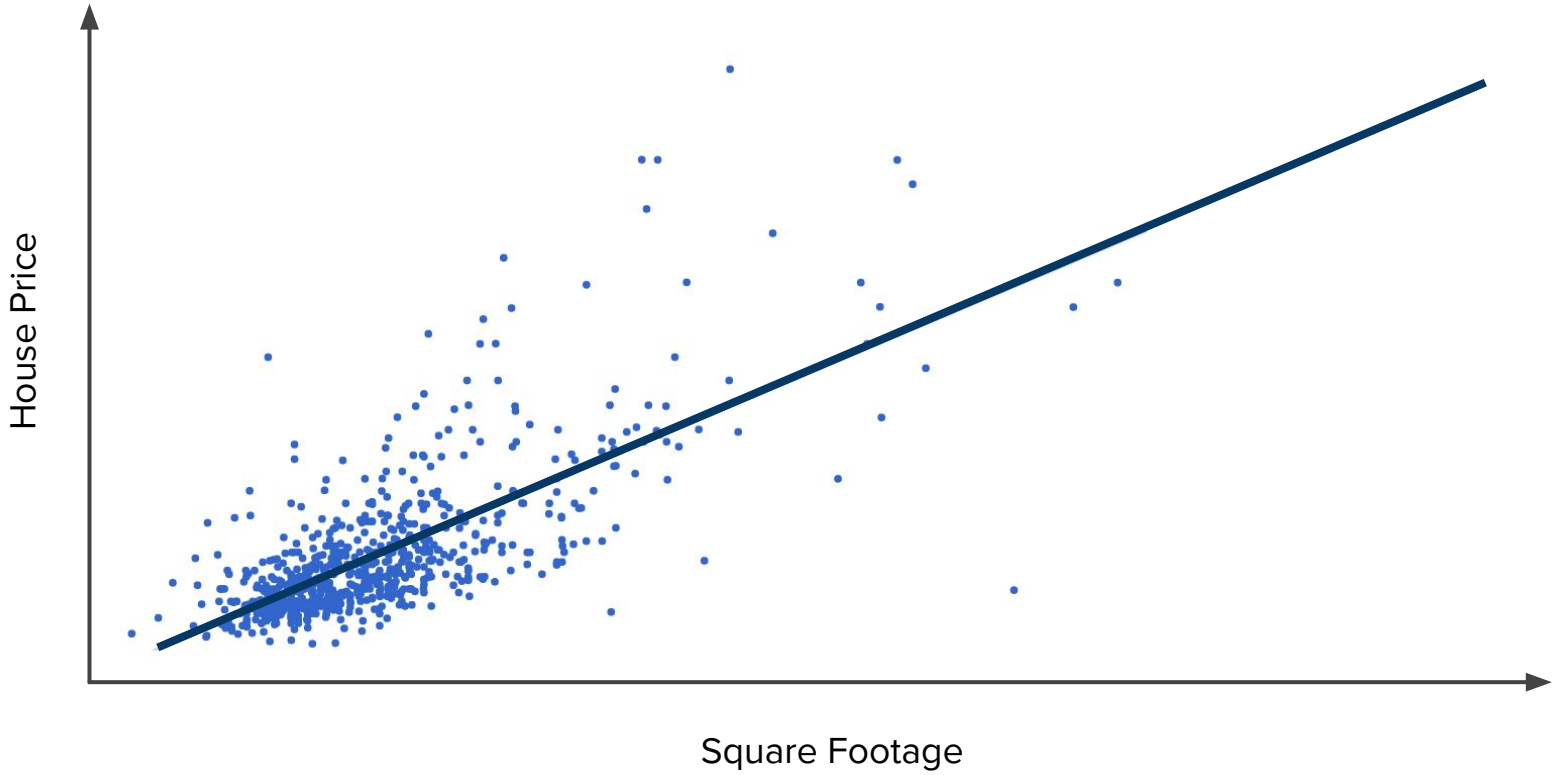
How Much Is My House Worth?



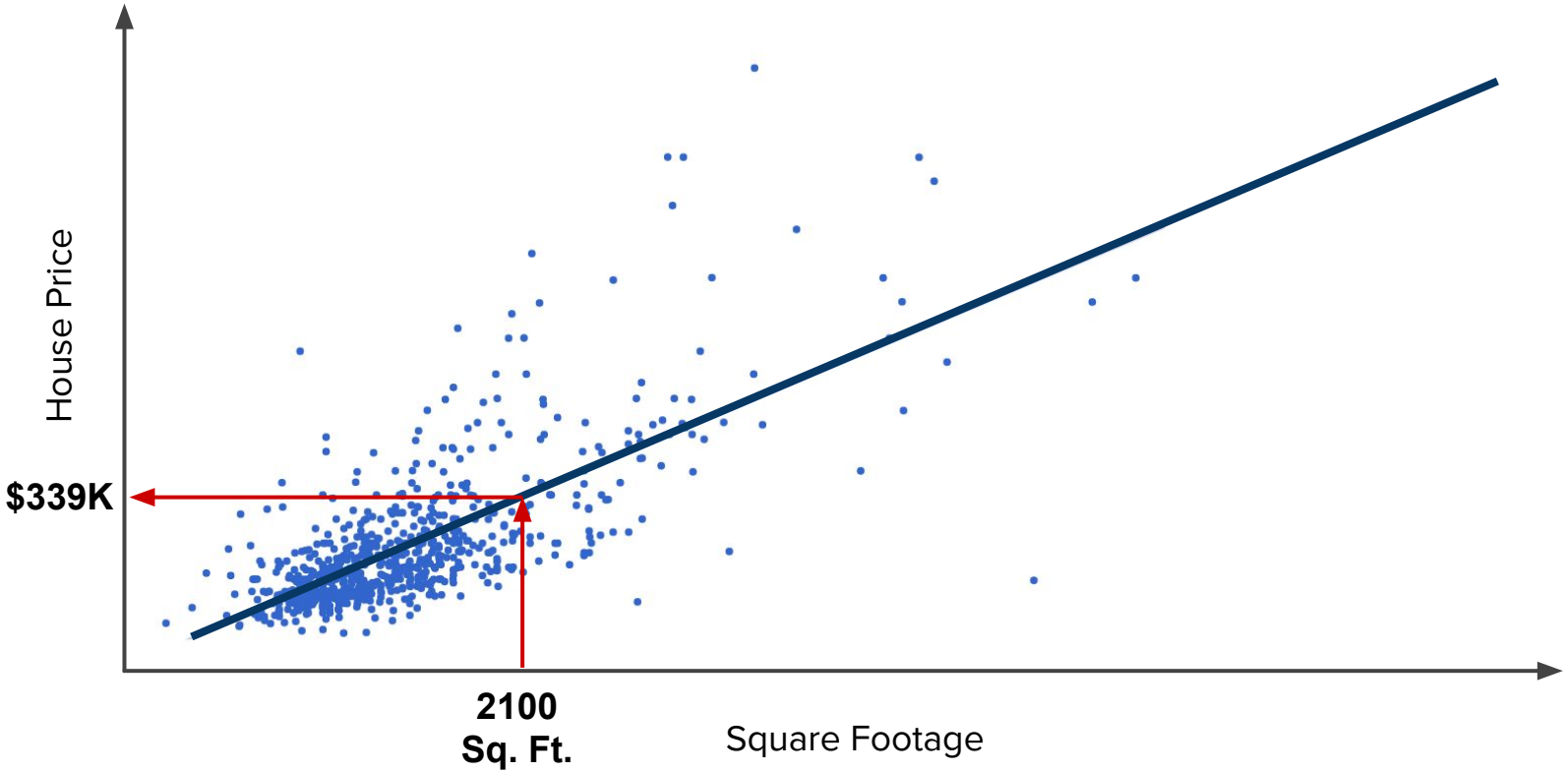
How Much Is My House Worth?



How Much Is My House Worth?



How Much Is My House Worth?



How does a machine learn?

PREDICTING HOUSE SALE PRICES

Sq. Footage	Price
1,000	$100,000
3,000	$300,000

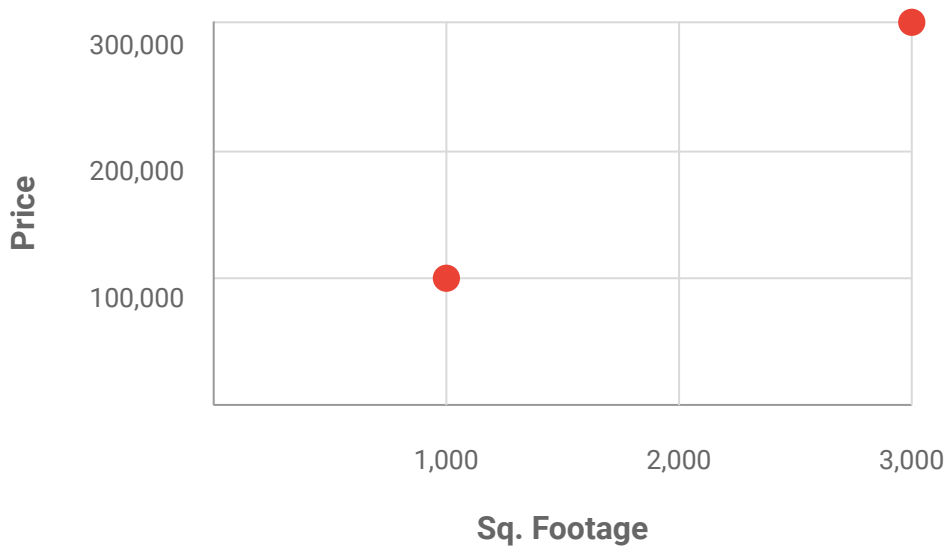
How much would a 2,000 sq ft. house sell for?

How does a machine learn?

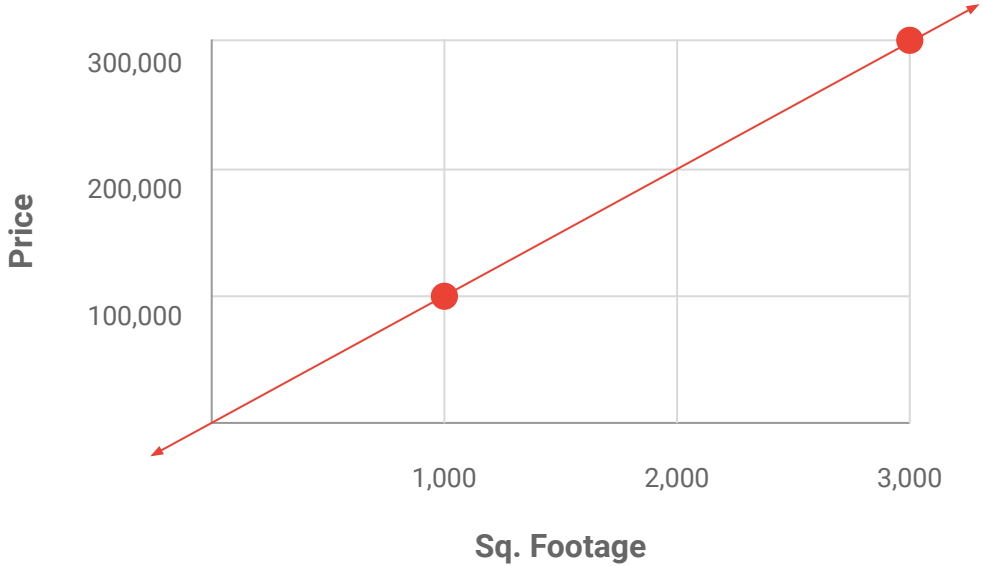
PREDICTING HOUSE SALE PRICES

Sq. Footage	Price
1,000	$100,000
3,000	$300,000
2,000	**$200,000**

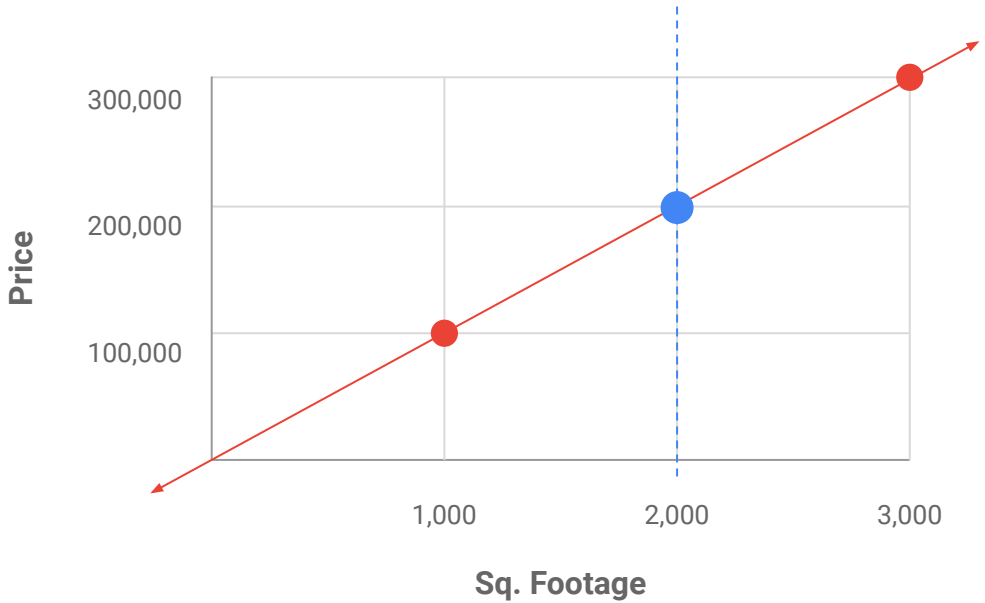
1. Look at the Data



2. Find a pattern



3. Predict



Revisiting our data

PREDICTING HOUSE SALE PRICES

Sq. Footage	Price
1,000	$100,000
3,000	$300,000

Three ways for AI on Google Cloud

MY HUGE **data** + MY **model**


Cloud TPUs


Compute Engine


Cloud Dataproc

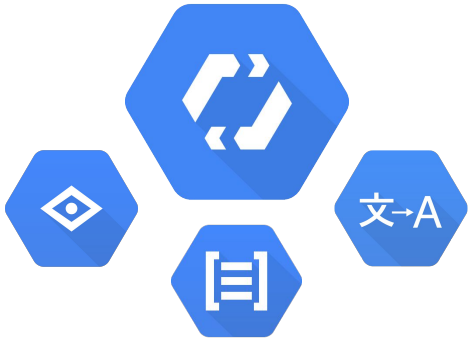

Kubernetes Engine


Cloud ML Engine


BigQuery ML

MY subset **data** + Google's **models**

AutoML



Google's **data** + Google's **models**


Cloud Translation API


Cloud Vision API


Cloud Speech API


Cloud Video Intelligence API


Data Loss Prevention API


Cloud Speech Synthesis API

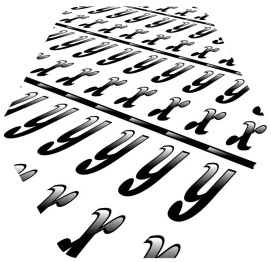

Cloud Natural Language API

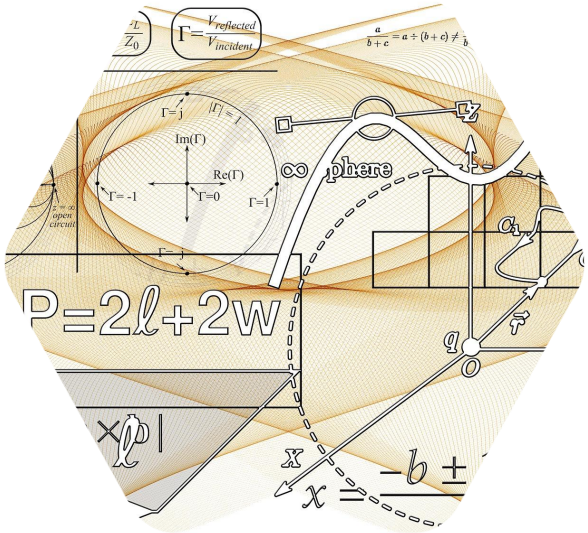

Dialogflow

Customisation — Ease of Use

Build your own models — Train our state-of-the-art models — Call our perception APIs

The popular imagination of what ML is





Lots of data

Complex mathematics in multidimensional spaces

Magical results


MACHINE LEARNING POC
imgflip.com

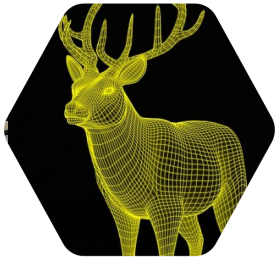

MACHINE LEARNING PRODUCTION
imgflip.com

In reality, ML is

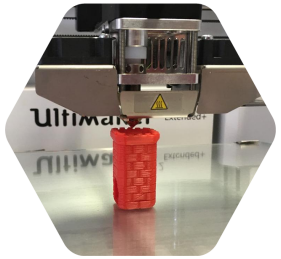
Define objectives

Collect data

Understand and prepare the data



Create the model



Refine the model

Serve the model

In reality, ML is

Define objectives

Collect data

Understand and prepare the data

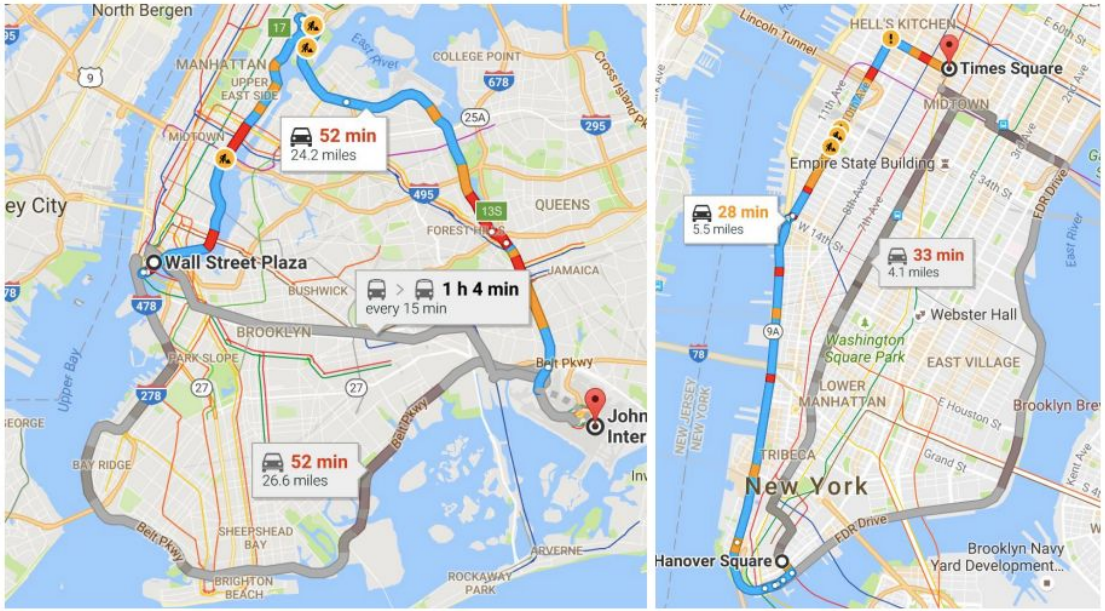
Create the model

Refine the model

Serve the model

Estimate taxi fares

NYC Taxi & Limousine Commission



Goal is to **estimate the taxi fare** based on pickup and drop locations, as well as other trip information...

http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

In reality, ML is

Define objectives

Collect data

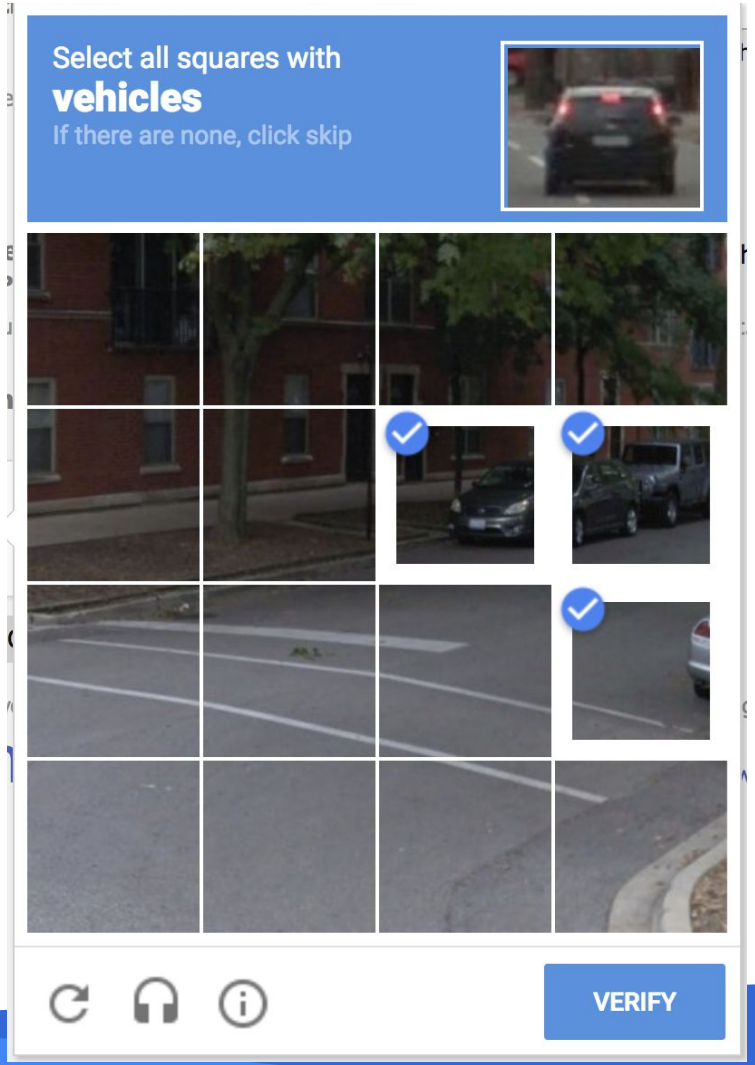
Understand and prepare the data

Create the model

Refine the model

Serve the model

In reality, ML is...



In reality, ML is

Define objectives

Collect data

Understand and prepare the data

Create the model

Refine the model

Serve the model

Serverless platform for all stages of the analytics data lifecycle

Ingestion

Cloud Pub/Sub

Cloud Storage

Processing

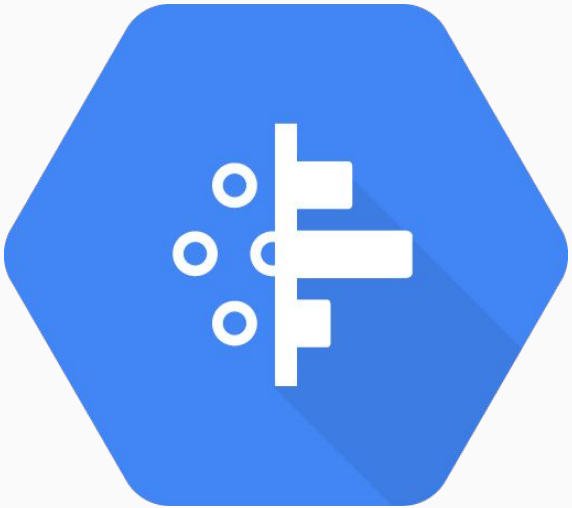
Cloud Dataflow

Analysis

BigQuery

No manual sharding
No capacity guessing
No idle resources
No maintenance window
No manual scaling
No file management

Introducing: self-service data preparation



Cloud Dataprep

- Visually explore and interact with data
- Instantly understand data distribution and patterns
- Quickly identify data quality issues
- Get automatic data transformation suggestions
- Standardize, structure and join datasets easily with a guided approach

Dataprep Example

Two datasets: Retail transactions & Ad impression logs
Goal: Find purchases that followed an ad for that product to that user

Retail purchase transactions

0-57	[{"salesOrderLineKey":"1","productKey":"1110067200000-182","dueDate":"2005-10-07T00:00:00Z","shipDate":"2005
0-588	[{"salesOrderLineKey":"1","productKey":"1145836800000-296","dueDate":"2006-09-13T00:00:00Z","shipDate":"2006
0-503	[{"salesOrderLineKey":"1","productKey":"1289606400000-221","dueDate":"2011-10-30T00:00:00Z","shipDate":"2011
0-623	[{"salesOrderLineKey":"1","productKey":"1213660800000-460","dueDate":"2008-09-11T00:00:00Z","shipDate":"2008

Ad impressions

4:00	https://theglobeandmail.com/sed/accumsan/felis/ut.json?pid=1157241600000-568
0:00	http://cam.ac.uk/neque/vestibulum/eget/vulputate.aspx?pid=1109721600000-237
8:00	https://odnoklassniki.ru/maecenas.jsp?pid=1281312000000-61
9:00	https://ft.com/hac.png?pid=1140393600000-91

Cloud Dataprep demo

BigQuery: 100% serverless data warehouse

Google BigQuery

- Google Cloud's Enterprise Data Warehouse for Analytics
- Petabyte-Scale and Fast Convenience of SQL
- Encrypted, Durable and Highly Available
- Fully Managed and Serverless

Powerful Data Exploration

Cloud Datalab

- Notebook interface
- Leverage existing Jupyter modules and knowledge
- Suitable to interactive data science and machine learning
- Closely integrated with BigQuery and Cloud ML

Understanding & Preparing data

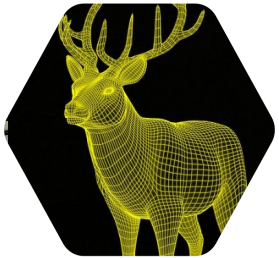
- Describe dataset
- Explore and visualise the data
- Create training, validation, and test datasets
- Set a baseline

In reality, ML is

Define objectives

Collect data

Understand and prepare the data



Create the model

Refine the model

Serve the model

First experiments

- Basic features
- Small training set
- Simple Linear Regression model

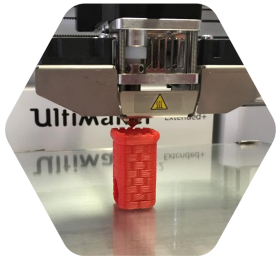
In reality, ML is

Define objectives

Collect data

Understand and prepare the data

Create the model



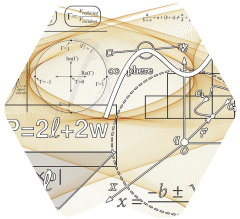
Refine the model

Serve the model

What do you think we can do to improve the model?

Refine the model

Feature engineering



Better algorithms

More examples, more data

Hyperparameter tuning

Refine the model

Feature engineering

Better algorithms

More examples, more data

Hyperparameter tuning

Feature engineering

Cleanse

Rows
Invalid data
Several missing features

Columns
(Near) zero variance
Many missing values
Many distinct values
(consider grouping or encoding)

Tune

Scaling
Handling missing values
Handling outliers
Handling skewed distribution

Transform

Discretization (binning)
Equal width
Equal Size
Clustering
Supervised

Encoding
Indicators (one-hot)
Learning with Counts
Embedding

Extract

(kernel-based) PCA
PLS (supervised)
Embedding
(Auto-encoders)
Hash Bucketization

Select

Filter methods
Wrapper methods

Construct

Polynomial expansion

Interactions (Crossing)
Multiplication, ratios
AND, OR, NOT

Feature comparison flags

Data-specific
Text
Image
Audio

Recency, Frequency, Intensity (RFI) metric
Datetime elements
Distances with locations

Feature engineering

Cleanse

Rows
Invalid data
Several missing features

Columns
(Near) zero variance
Many missing values
Many distinct values
(consider grouping or encoding)

Tune

Scaling
Handling missing values
Handling outliers
Handling skewed distribution

Transform

Discretization (binning)
Equal width
Equal Size
Clustering
Supervised

Encoding
Indicators (one-hot)
Learning with Counts
Embedding

Extract

(kernel-based) PCA
PLS (supervised)
Embedding
(Auto-encoders)
Hash Bucketization

Select

Filter methods
Wrapper methods

Construct

Polynomial expansion

Interactions (Crossing)
Multiplication, ratios
AND, OR, NOT

Feature comparison flags

Data-specific
Text
Image
Audio

Recency, Frequency, Intensity (RFI) metric
Datetime elements
Distances with locations

Feature engineering

Cleanse

Rows
Invalid data
Several missing features

Columns
(Near) zero variance
Many missing values
Many distinct values
(consider grouping or encoding)

Tune

Scaling
Handling missing values
Handling outliers
Handling skewed distribution

Transform

Discretization (binning)
Equal width
Equal Size
Clustering
Supervised

Encoding
Indicators (one-hot)
Learning with Counts
Embedding

Extract

(kernel-based) PCA
PLS (supervised)
Embedding
(Auto-encoders)
Hash Bucketization

Select

Filter methods
Wrapper methods

Construct

Polynomial expansion

Interactions (Crossing)
Multiplication, ratios
AND, OR, NOT

Feature comparison flags

Data-specific
Text
Image
Audio

Recency, Frequency, Intensity (RFI) metric
Datetime elements
Distances with locations

Feature engineering

Rows
Invalid data
Several missing features

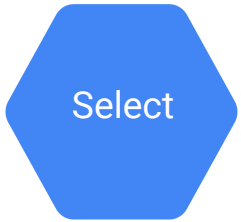
Columns
(Near) zero variance
Many missing values
Many distinct values
(consider grouping or encoding)

Scaling
Handling missing values
Handling outliers
Handling skewed distribution

Discretization (binning)
Equal width
Equal Size
Clustering
Supervised

Encoding
Indicators (one-hot)
Learning with Counts
Embedding

(kernel-based) PCA
PLS (supervised)
Embedding
(Auto-encoders)
Hash Bucketization



Filter methods
Wrapper methods

Polynomial expansion

Interactions (Crossing)
Multiplication, ratios
AND, OR, NOT

Feature comparison flags

Data-specific
Text
Image
Audio

Recency, Frequency, Intensity (RFI) metric
Datetime elements
Distances with locations

Feature engineering

Cleanse

Rows
Invalid data
Several missing features

Columns
(Near) zero variance
Many missing values
Many distinct values
(consider grouping or encoding)

Tune

Scaling
Handling missing values
Handling outliers
Handling skewed distribution

Transform

Discretization (binning)
Equal width
Equal Size
Clustering
Supervised

Encoding
Indicators (one-hot)
Learning with Counts
Embedding

Extract

(kernel-based) PCA
PLS (supervised)
Embedding
(Auto-encoders)
Hash Bucketization

Select

Filter methods
Wrapper methods

Construct

Polynomial expansion

Interactions (Crossing)
Multiplication, ratios
AND, OR, NOT

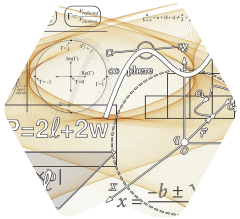
Feature comparison flags

Data-specific
Text
Image
Audio

Recency, Frequency, Intensity (RFI) metric
Datetime elements
Distances with locations

Refine the model

Feature engineering



Better algorithms

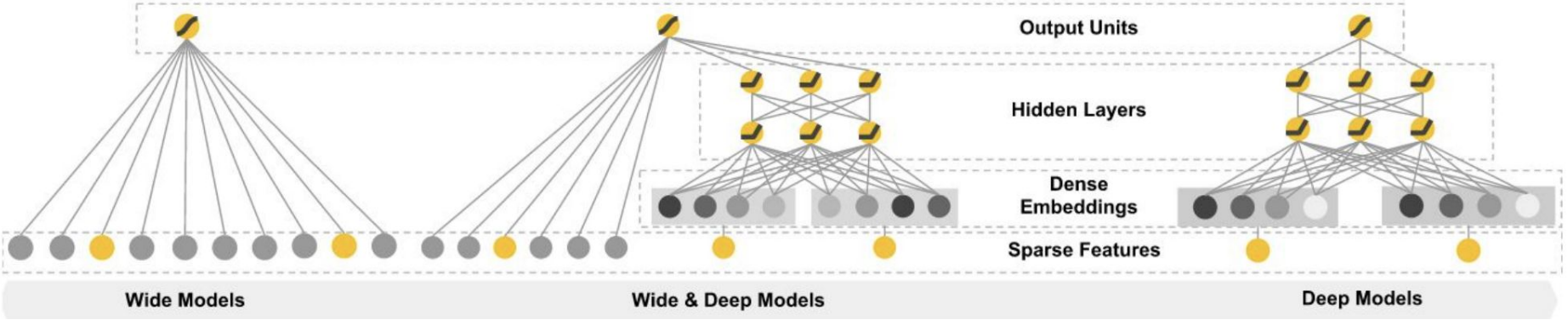
More examples, more data

Hyperparameter tuning

Better algorithms: DNN

- **Consider different optimisation algorithms**: Adam, Adagrade, derivative-free methods, etc.
- **Consider different activation functions**: relu, radial basis functions (RDF), etc.
- **Consider different model architectures**: number of layers and nodes per layer, connectivity, etc.
- **Consider different loss metric:** quadratic loss, cross entropy, bayesian information reward, etc.
- **Consider handling overfitting**: regularisation, early-stopping, dropouts, etc.
- **Consider Linear Combined DNN**: for deep (dense) and wide (sparse) features

Linear combined DNN

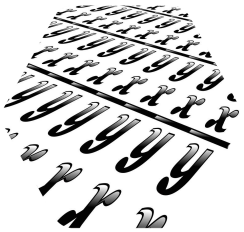


https://research.googleblog.com/2016/06/wide-deep-learning-better-together-with.html

Refine the model

Feature engineering

Better algorithms



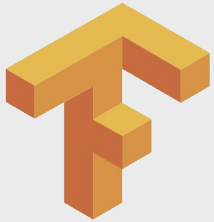
More examples, more data

Hyperparameter tuning

Fully-managed data processing service

Cloud Dataflow

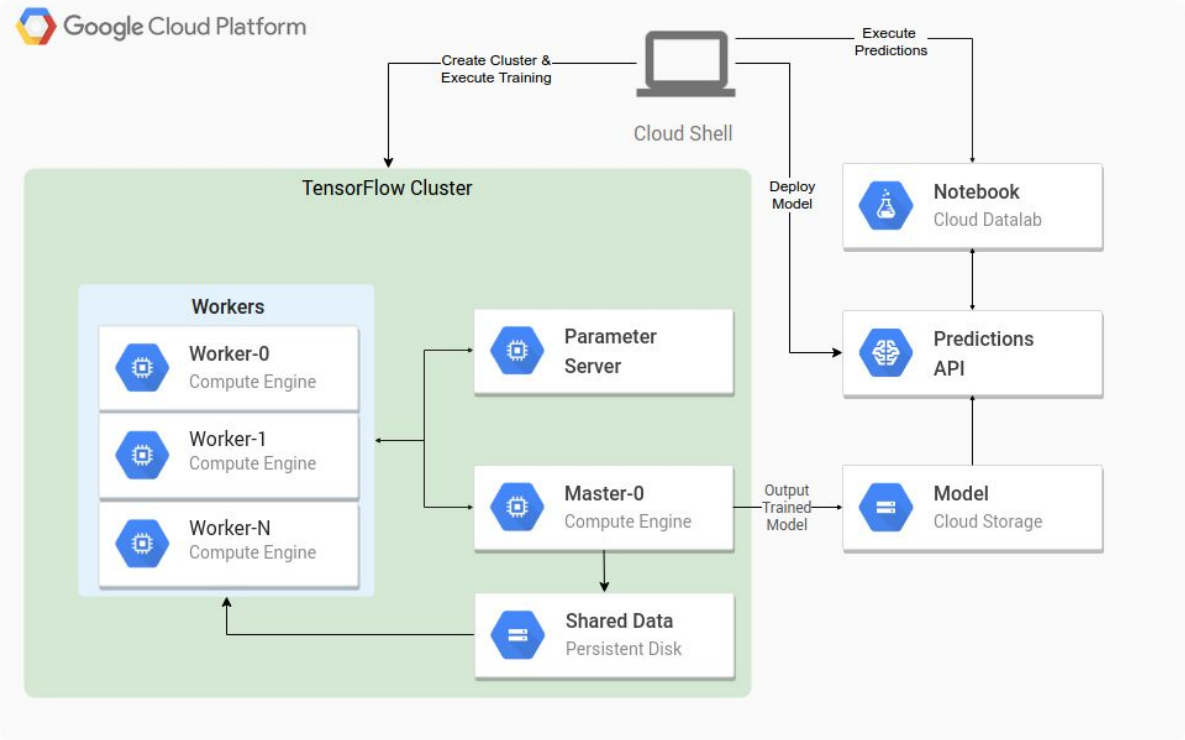
- Combines batch and streaming models
- Autoscaling and dynamic task sharding
- Connectors to Bigtable, BigQuery, Cloud Storage
- Portable via open-source Apache Beam SDK



Preparing big data for training

- Dataflow to export data from BigQuery to GCS

Distributed TensorFlow on Compute Engine



https://cloud.google.com/solutions/running-distributed-tensorflow-on-compute-engine

Machine Learning on any data, of any size

Cloud ML Engine

- Services are designed to work together
- Managed distributed training infrastructure that supports CPUs and GPUs
- Automatic hyperparameter tuning
- Portable models with TensorFlow

Running locally

train locally

output directory

```
gcloud ml-engine local train \
   --module-name trainer.task --package-path trainer/ \
   -- \
   --train-files $TRAIN_DATA --eval-files $EVAL_DATA --train-steps 1000 --job-dir $MODEL_DIR
```

training data

evaluation data

Single trainer running in the cloud

train in the cloud

region

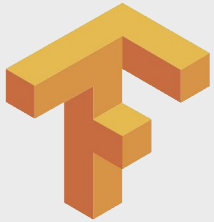
```
gcloud ml-engine jobs submit training $JOB_NAME --job-dir $OUTPUT_PATH \
   --runtime-version 1.0 --module-name trainer.task --package-path trainer/ --region $REGION \
   -- \
   --train-files $TRAIN_DATA --eval-files $EVAL_DATA --train-steps 1000 --verbosity DEBUG
```

Google cloud storage location

Distributed training in the cloud

```
gcloud ml-engine jobs submit training $JOB_NAME --job-dir $OUTPUT_PATH \
   --runtime-version 1.0 --module-name trainer.task --package-path trainer/ --region $REGION \
   --scale-tier STANDARD_1
   -- \
   --train-files $TRAIN_DATA --eval-files $EVAL_DATA --train-steps 1000 --verbosity DEBUG
```

distributed



Training with Cloud ML Engine

Refine the model

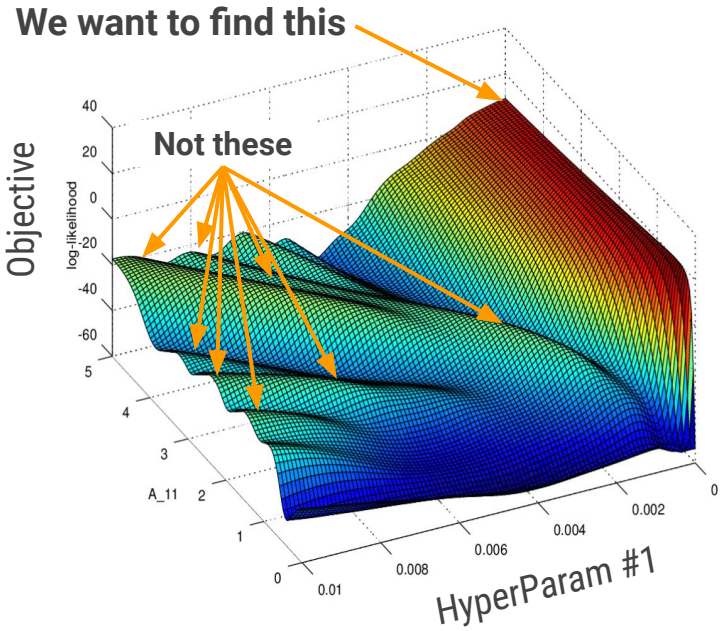
Feature engineering

Better algorithms

More examples, more data

Hyperparameter tuning

Hyperparameter tuning



- Automatic hyperparameter tuning service
- Build better performing models faster and save many hours of manual tuning
- Google-developed search (Bayesian Optimisation) algorithm efficiently finds better hyperparameters for your model/dataset

https://cloud.google.com/blog/big-data/2017/08/hyperparameter-tuning-in-cloud-machine-learning-engine-using-bayesian-optimization

Hyperparameter tuning

```
gcloud ml-engine jobs submit training $JOB_NAME --job-dir $OUTPUT_PATH \
   --runtime-version 1.0 --module-name trainer.task --package-path trainer/ --region $REGION \
   --scale-tier STANDARD_1 --config $HPTUNING_CONFIG
   -- \
   --train-files $TRAIN_DATA --eval-files $EVAL_DATA --train-steps 1000 --verbosity DEBUG
```

hypertuning

Hyperparameter tuning

hptuning_config.yaml

```yaml
trainingInput:
 hyperparameters:
   goal: MAXIMIZE
   hyperparameterMetricTag: accuracy
   maxTrials: 4
   maxParallelTrials: 2
   params:
     - parameterName: first-layer-size
       type: INTEGER
       minValue: 50
       maxValue: 500
       scaleType: UNIT_LINEAR_SCALE

...
```

task.py

```
...

      # Construct layers sizes with exponetial decay
      hidden_units=[
          max(2, int(hparams.first_layer_size *
                     hparams.scale_factor**i))
          for i in range(hparams.num_layers)
      ],

...

parser.add_argument(
      '--first-layer-size',
      help='Number of nodes in the 1st layer of the DNN',
      default=100,
      type=int
 )

...
```

In reality, ML is

Define objectives

Collect data

Understand and prepare the data

Create the model

Refine the model

Serve the model

Deploying the model

Creating model

```
gcloud ml-engine models create $MODEL_NAME --regions=$REGION
```

Creating versions

```
gcloud ml-engine versions create v1 --model $MODEL_NAME --origin $MODEL_BINARIES \
  --runtime-version 1.0

gcloud ml-engine models list
```

Predicting

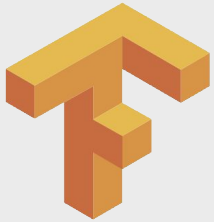
```
gcloud ml-engine predict --model $MODEL_NAME --version v1 --json-instances ../test.json
```

Using REST:

POST https://ml.googleapis.com/v1/{name=projects/**}:predict

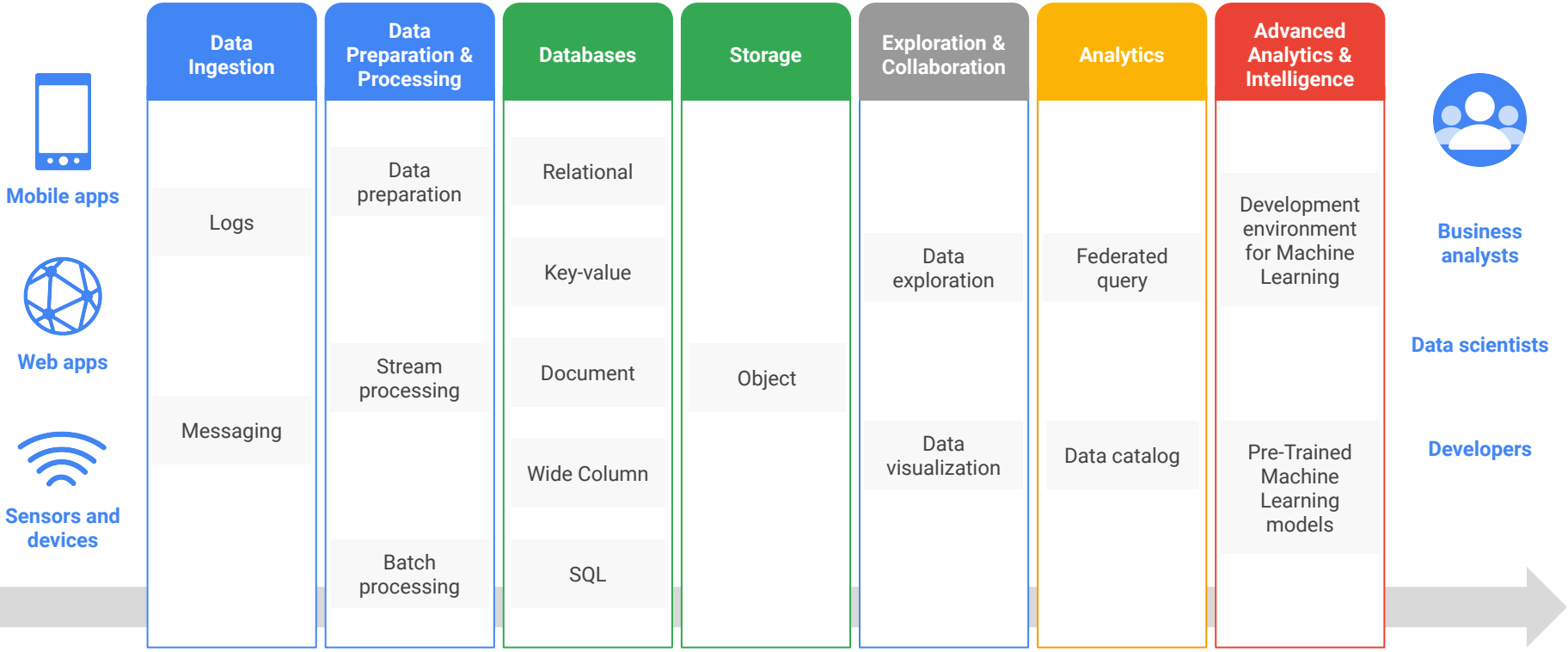
JSON format (in this case):

```
{"age": 25, "workclass": "private", "education": "11th", "education_num": 7, "marital_status":
"Never-married", "occupation": "machine-op-inspector", "relationship": "own-child", "gender": "
male", "capital_gain": 0, "capital_loss": 0, "hours_per_week": 40, "native_country": "
United-States"}
```

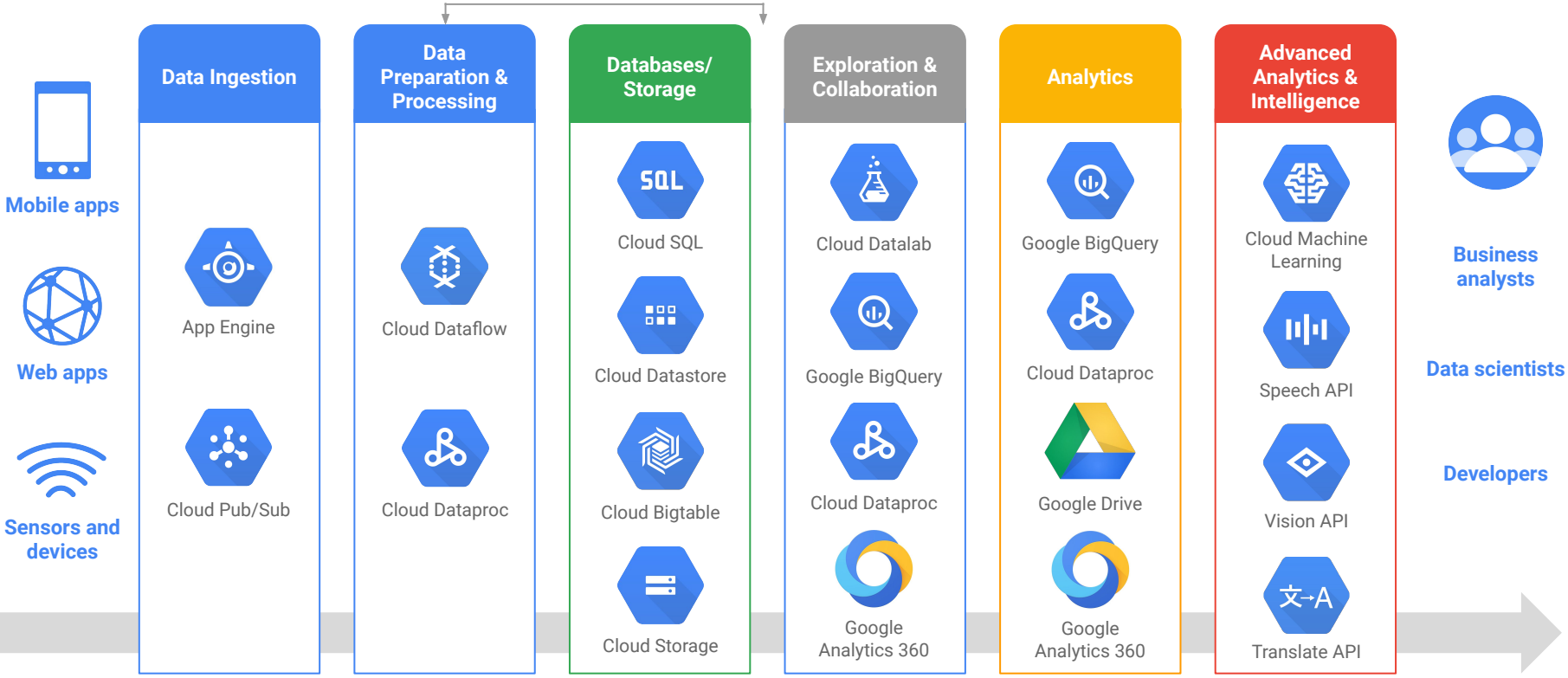


Predicting taxi fares

Google Cloud Data Platform



Google Cloud Data Platform



Applied Data Science Challenges

Applied data science challenges

The need for an integrated data science environment

Training and tuning ML at scale

Deploying, managing and consuming ML models as APIs

Choosing the right, flexible, effective ML framework

Making sense of Unstructured data to enrich my information

Transitioning from PoC to production + Automation

Integrating with other data systems and services

Time, cost and complexity to build ML pipelines

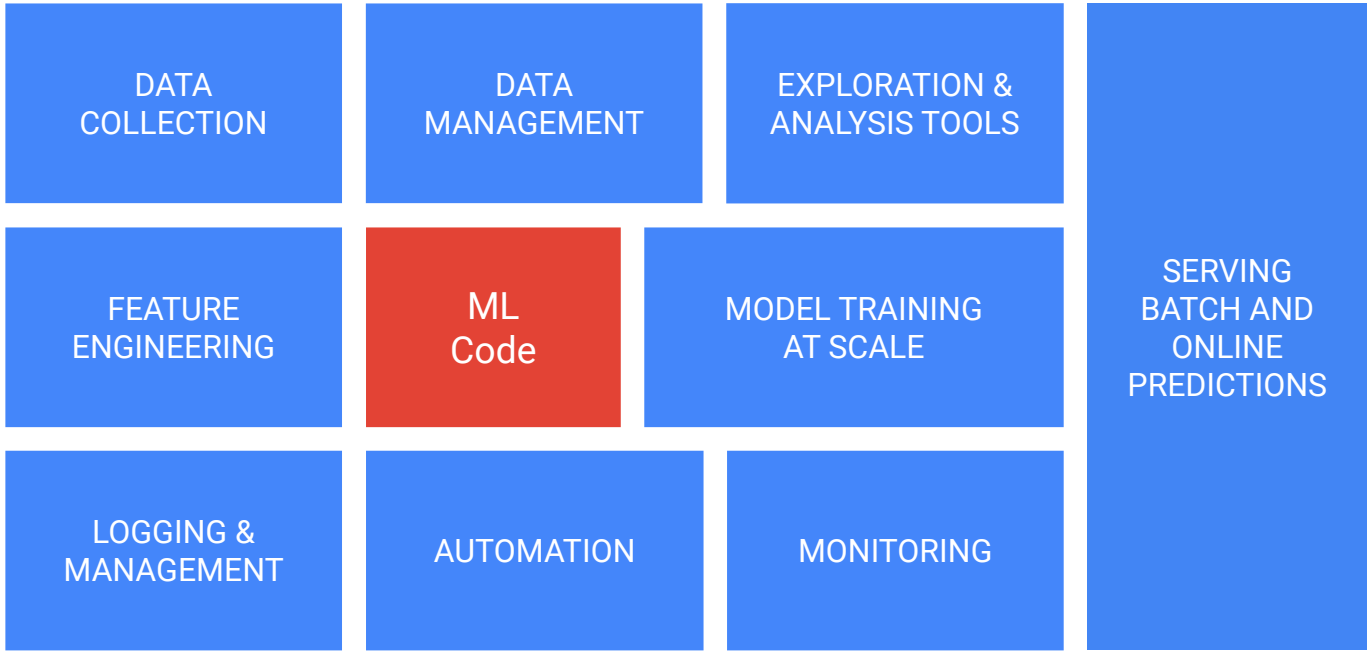
Google Cloud

Experimental data science

Google Cloud

Adapted from: Hidden Technical Debt in Machine Learning Systems

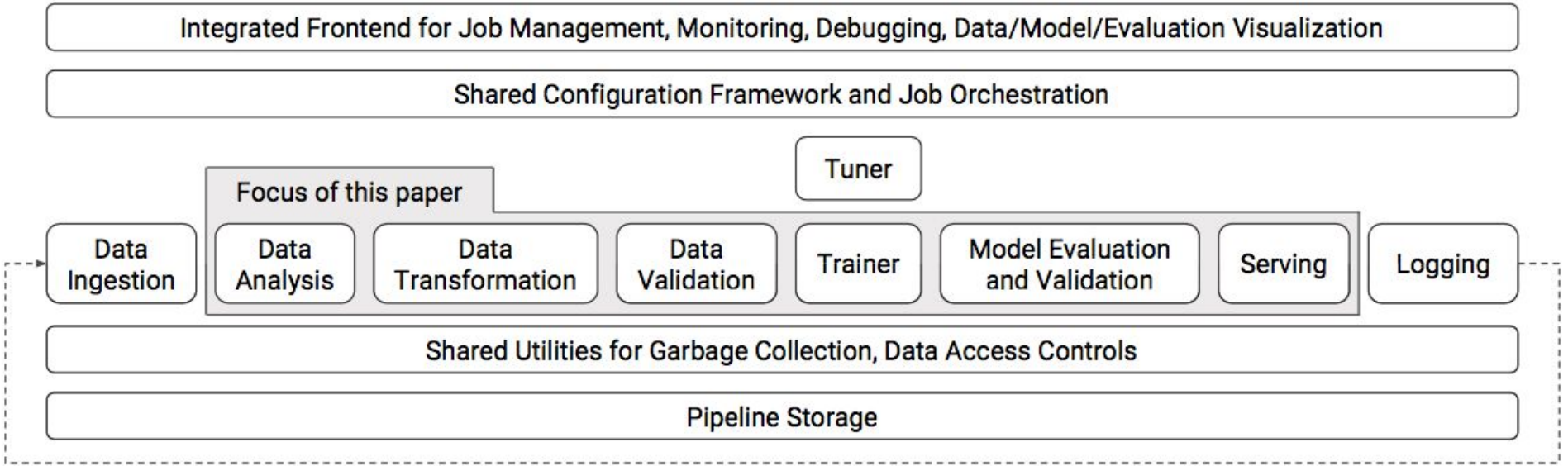
Operational ML systems



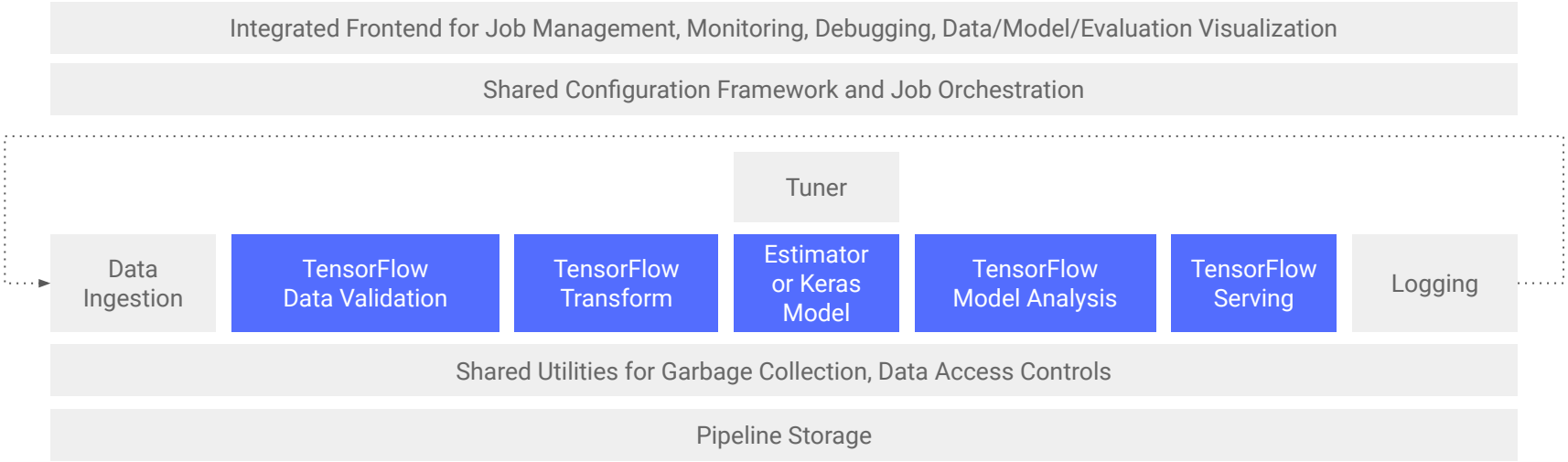
Google Cloud

Adapted from: Hidden Technical Debt in Machine Learning Systems

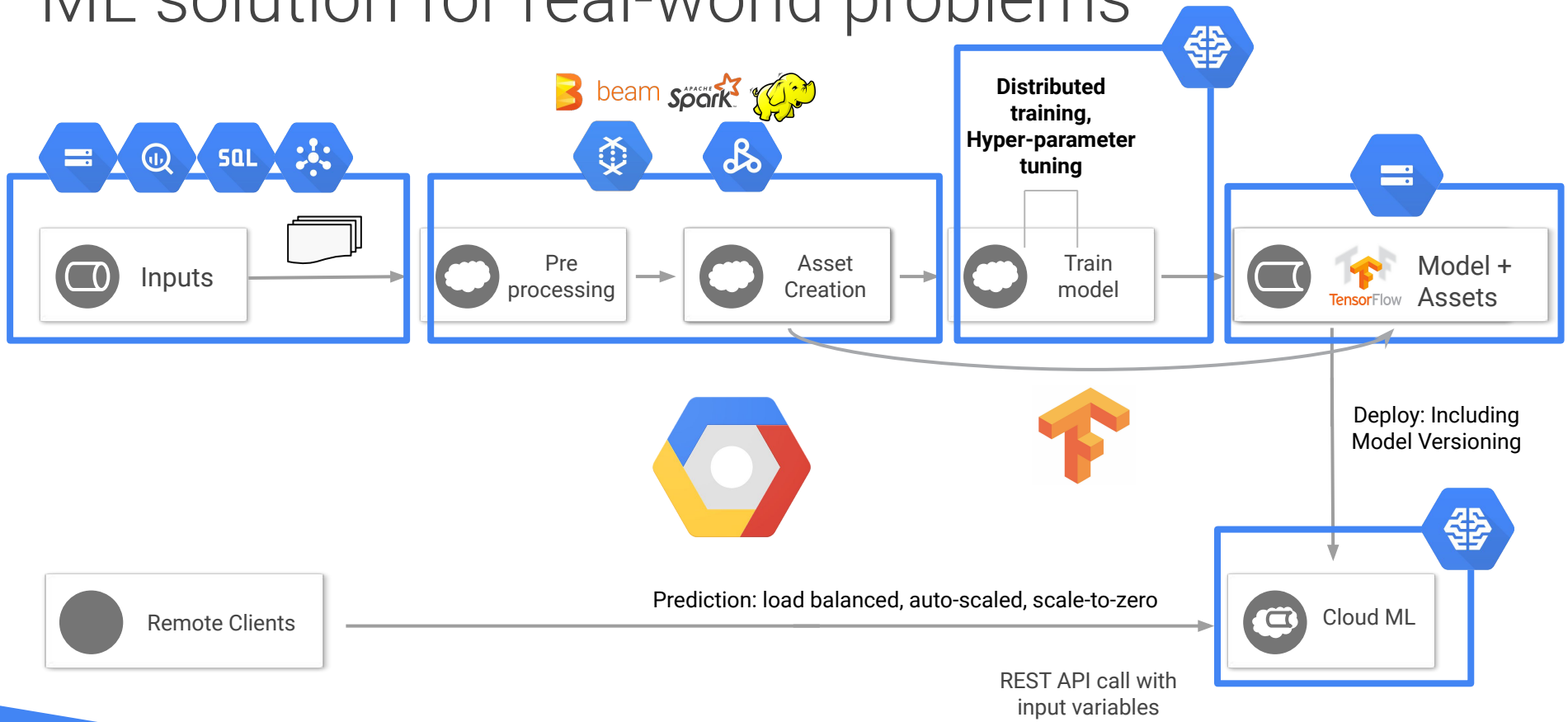
TFX is the solution to this problem...



…and we are making it available to you.

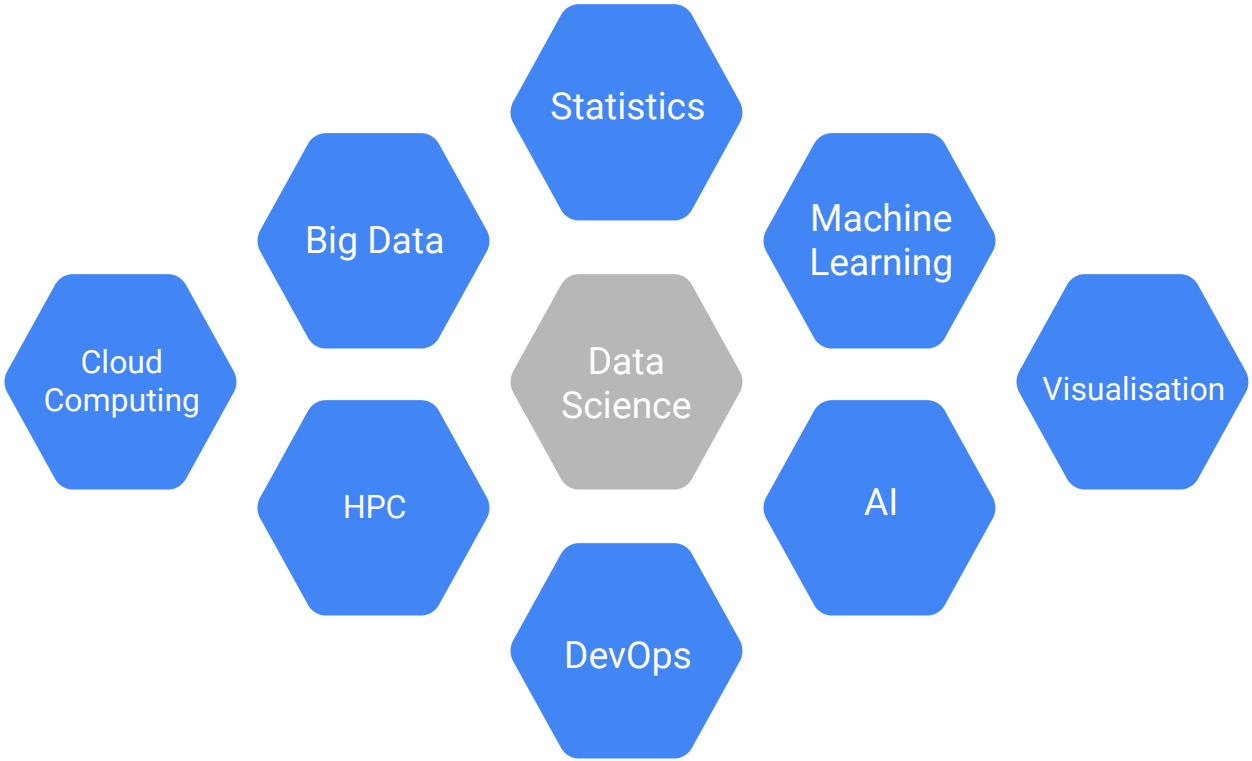


ML solution for real-world problems



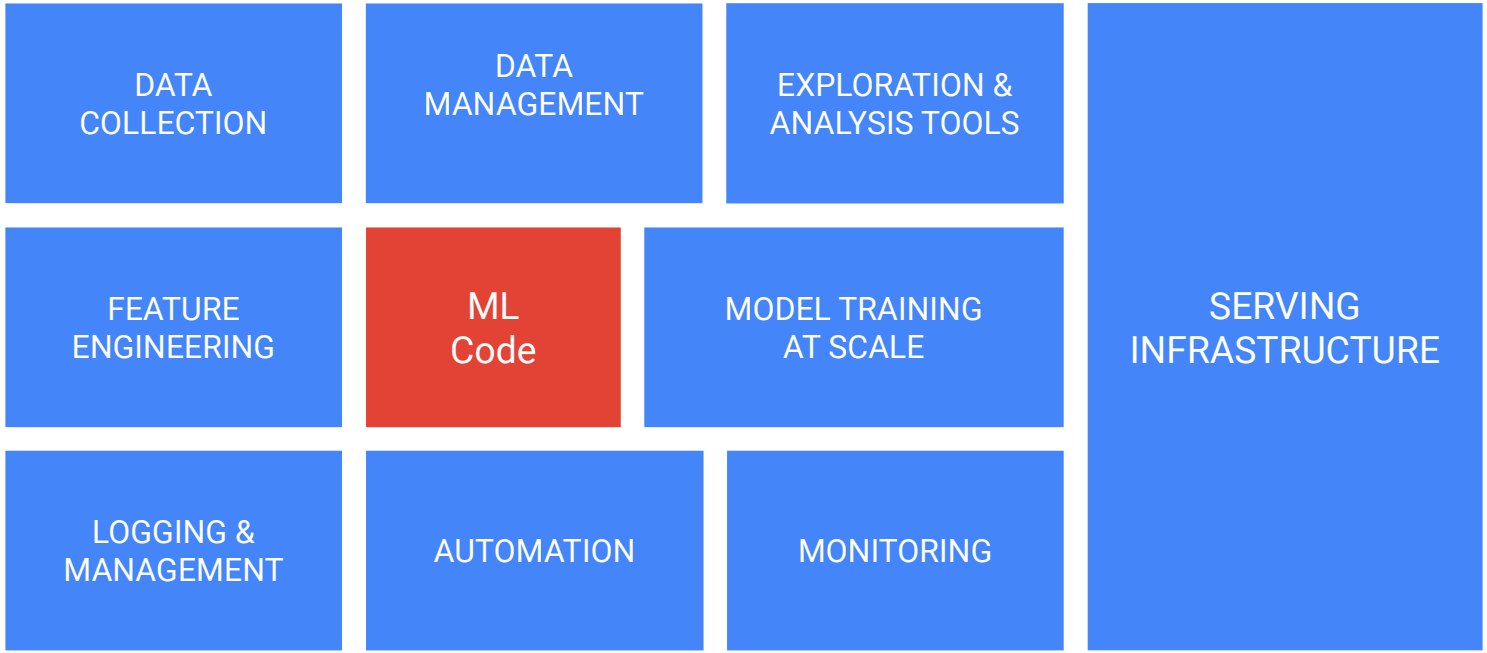
Data science now

From experimentation to operationalisation



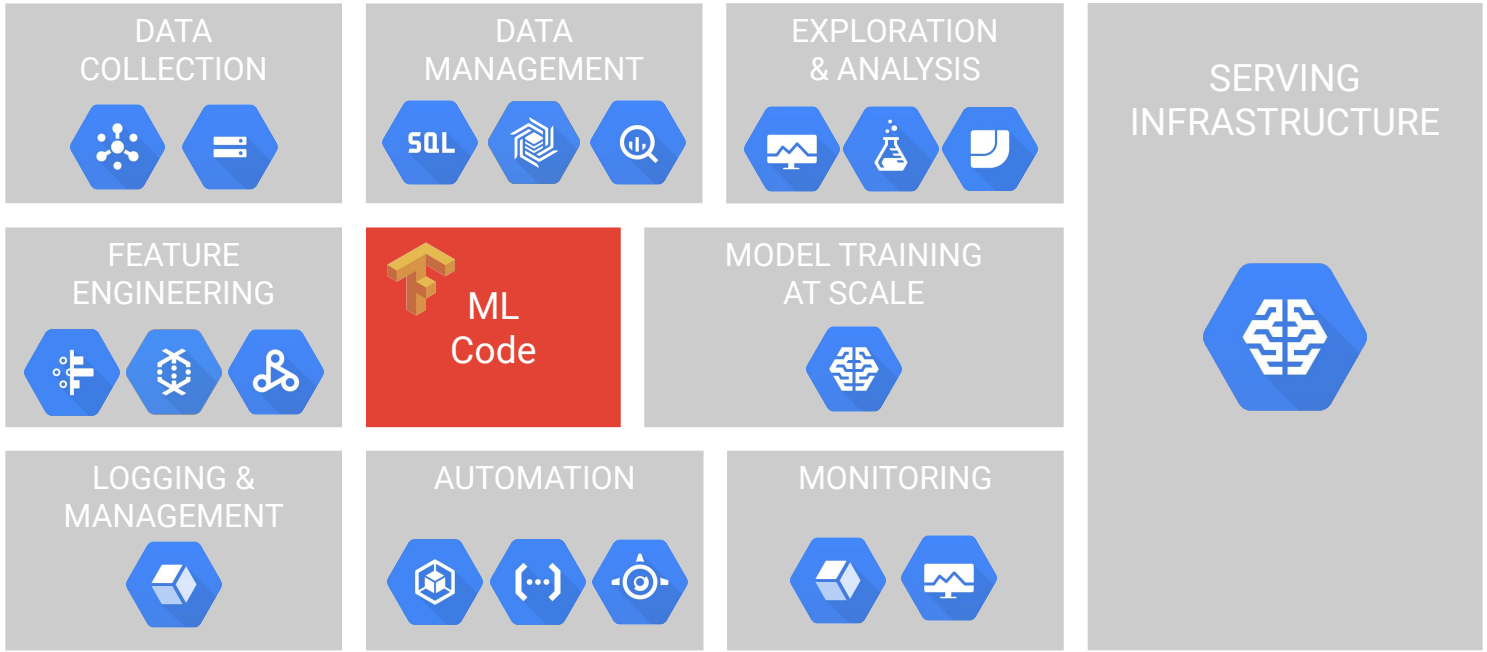
Google Cloud

Operational ML systems



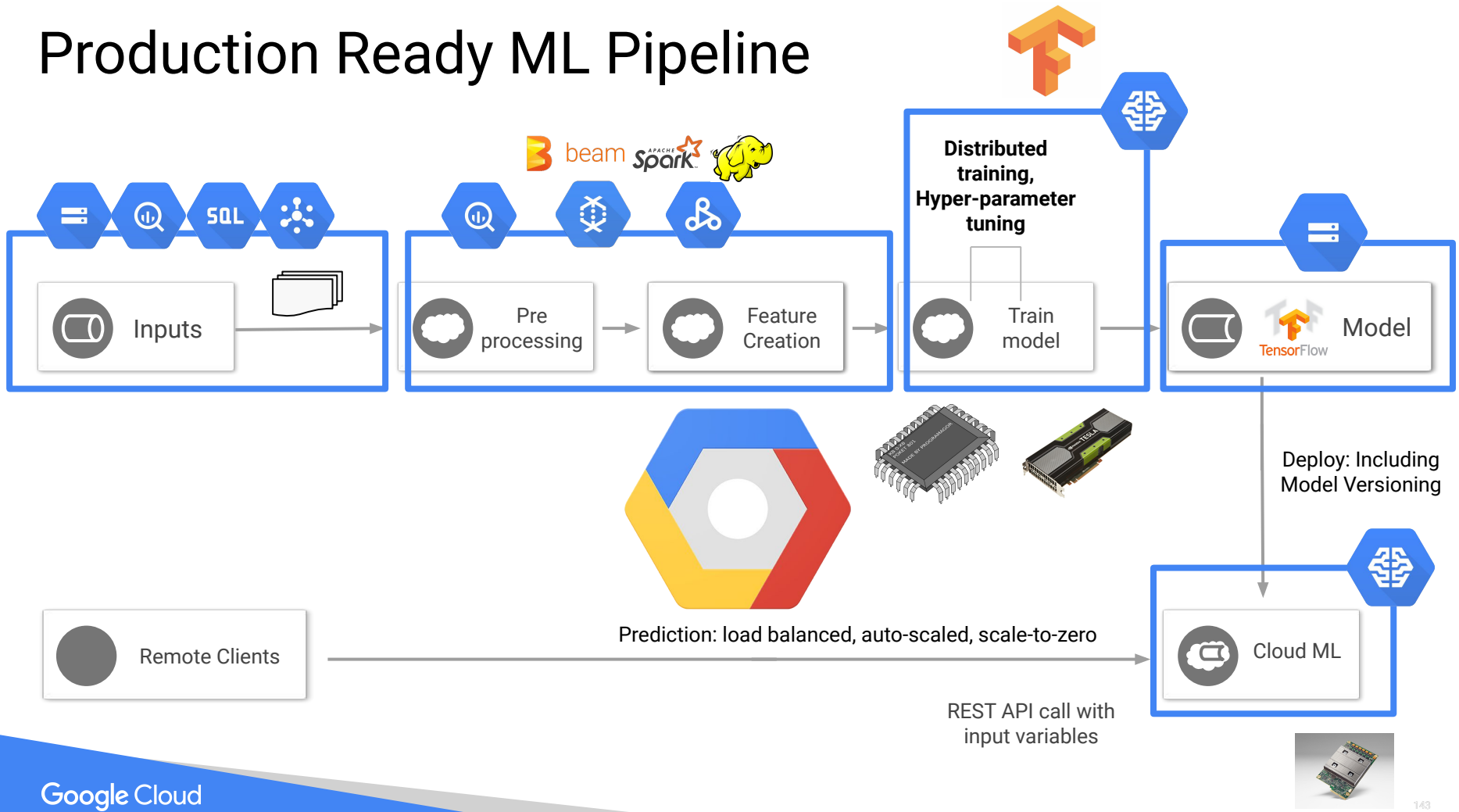
Google Cloud

Operational ML systems

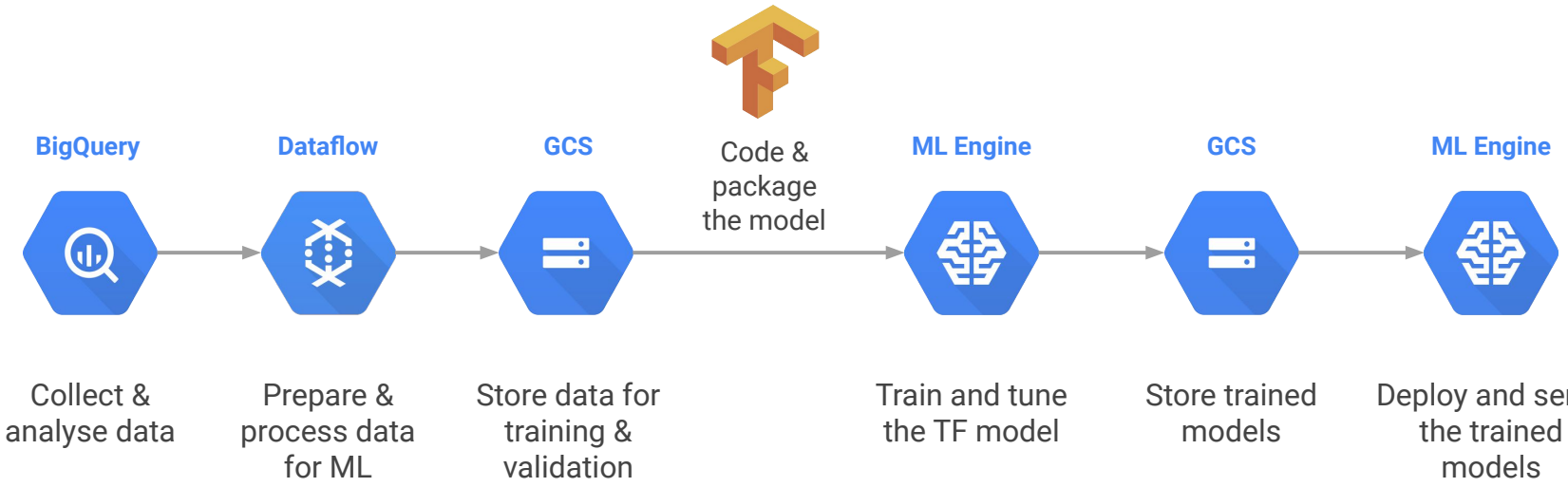


Google Cloud

Production Ready ML Pipeline

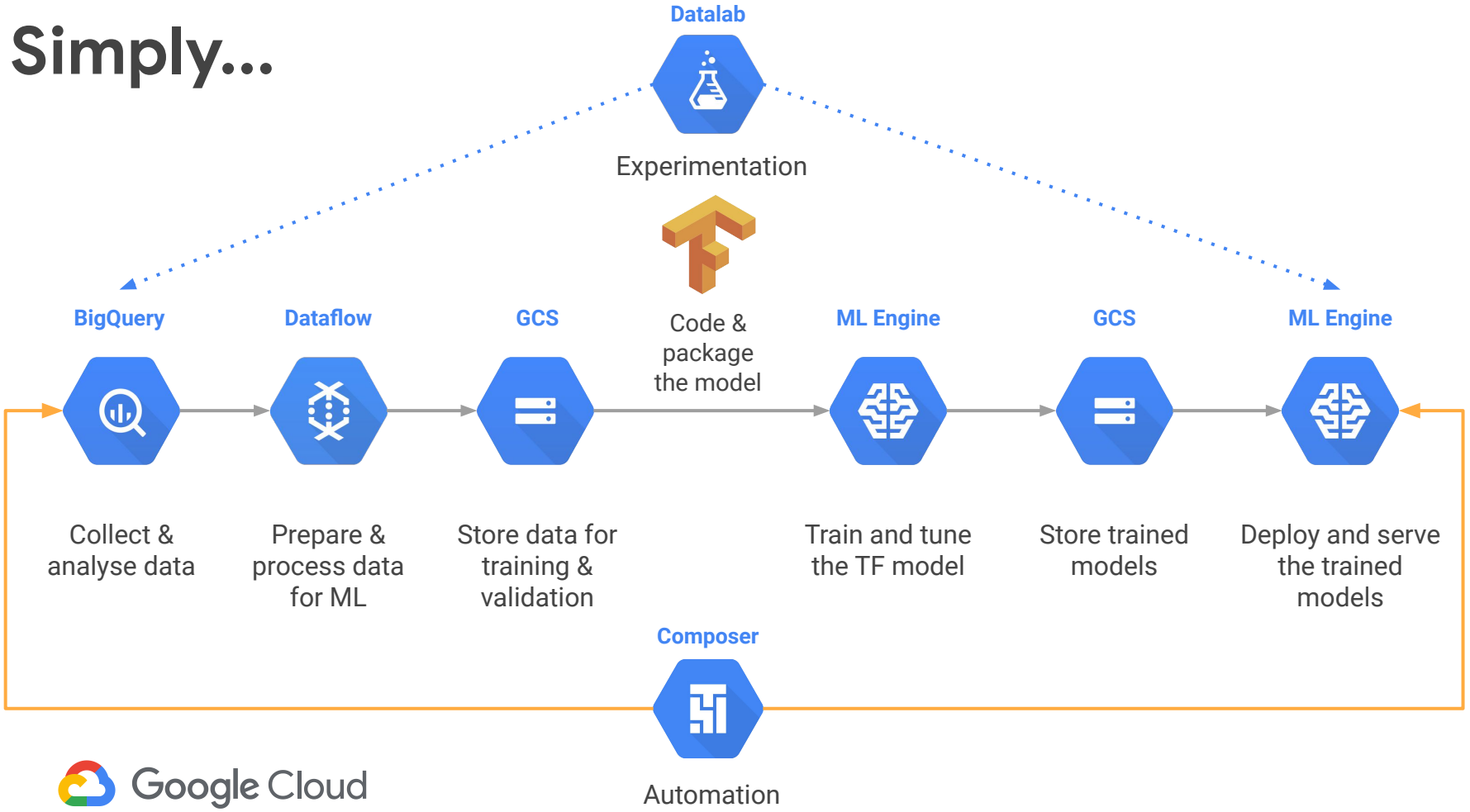


Simply...



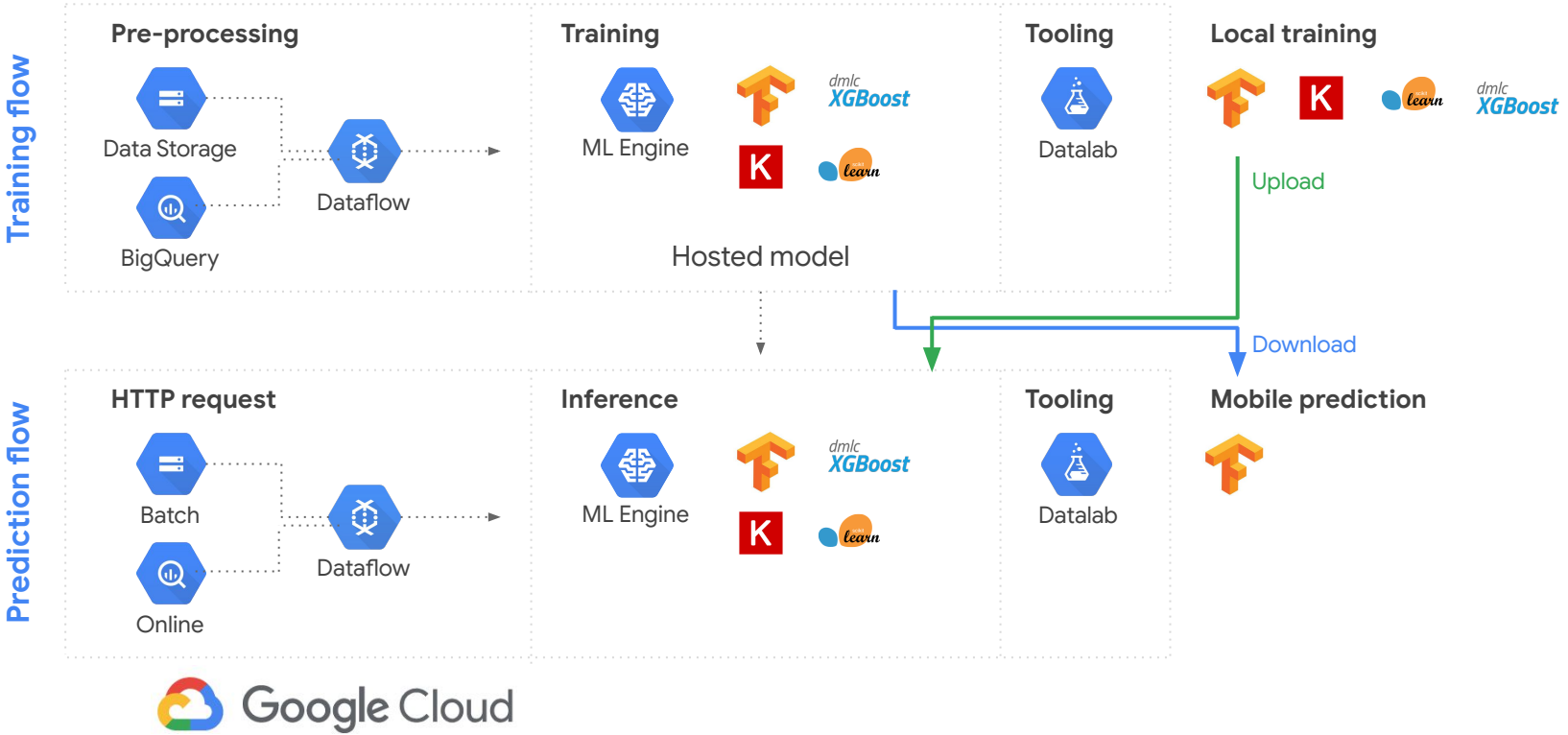
Google Cloud

Simply...



Google Cloud

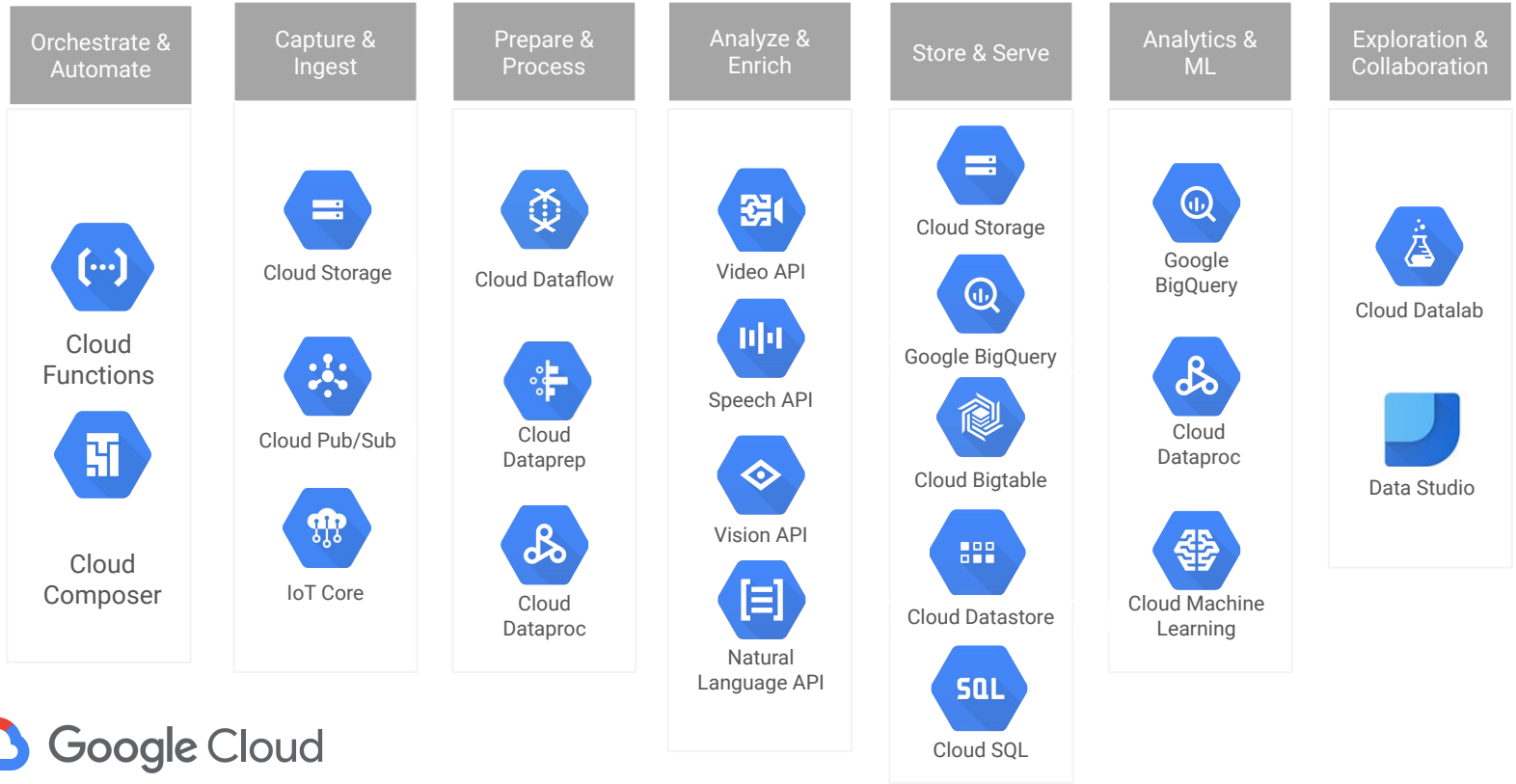
Common training and prediction workflows



Google Cloud

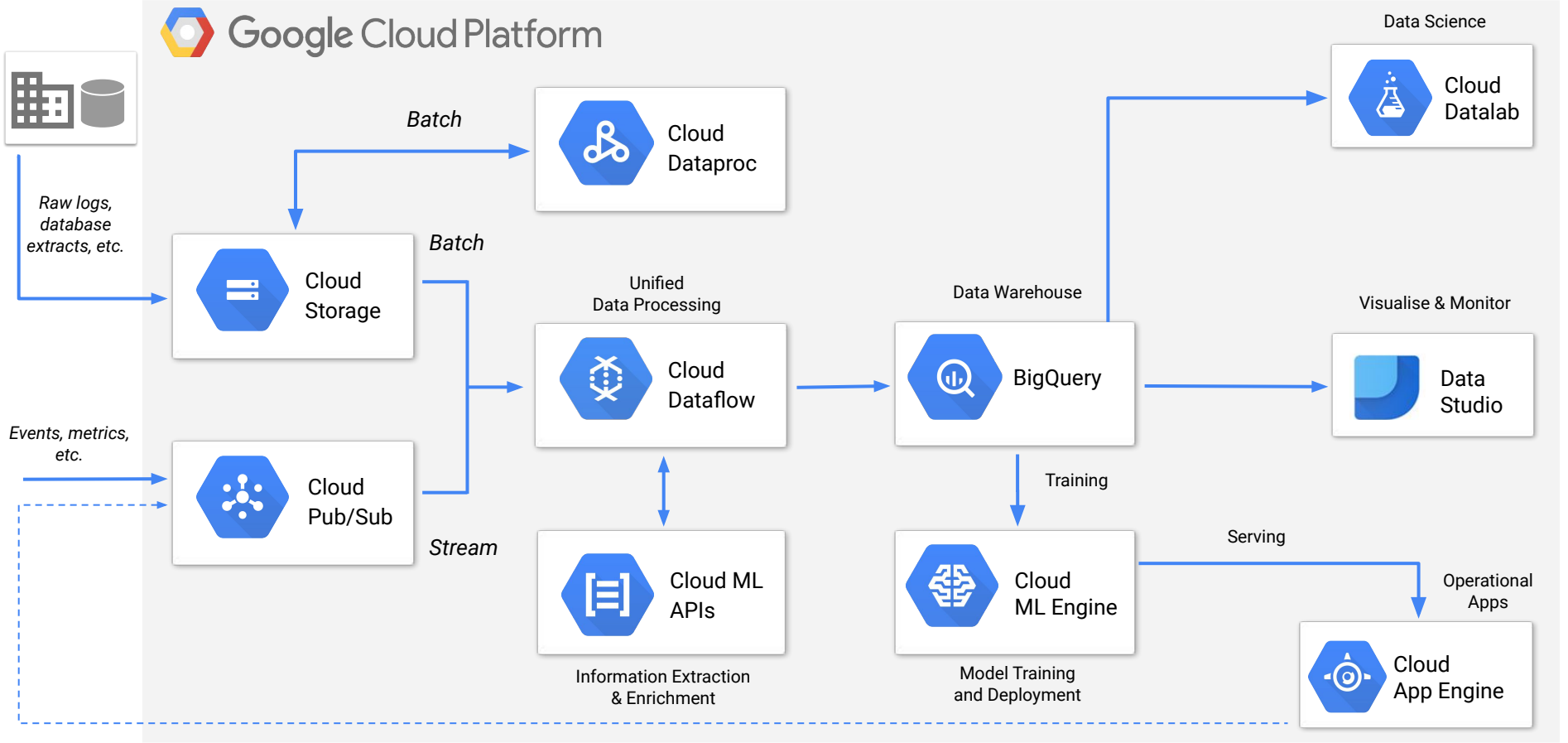
Google Cloud Data & Analytics Ecosystem

Data & analytics ecosystem

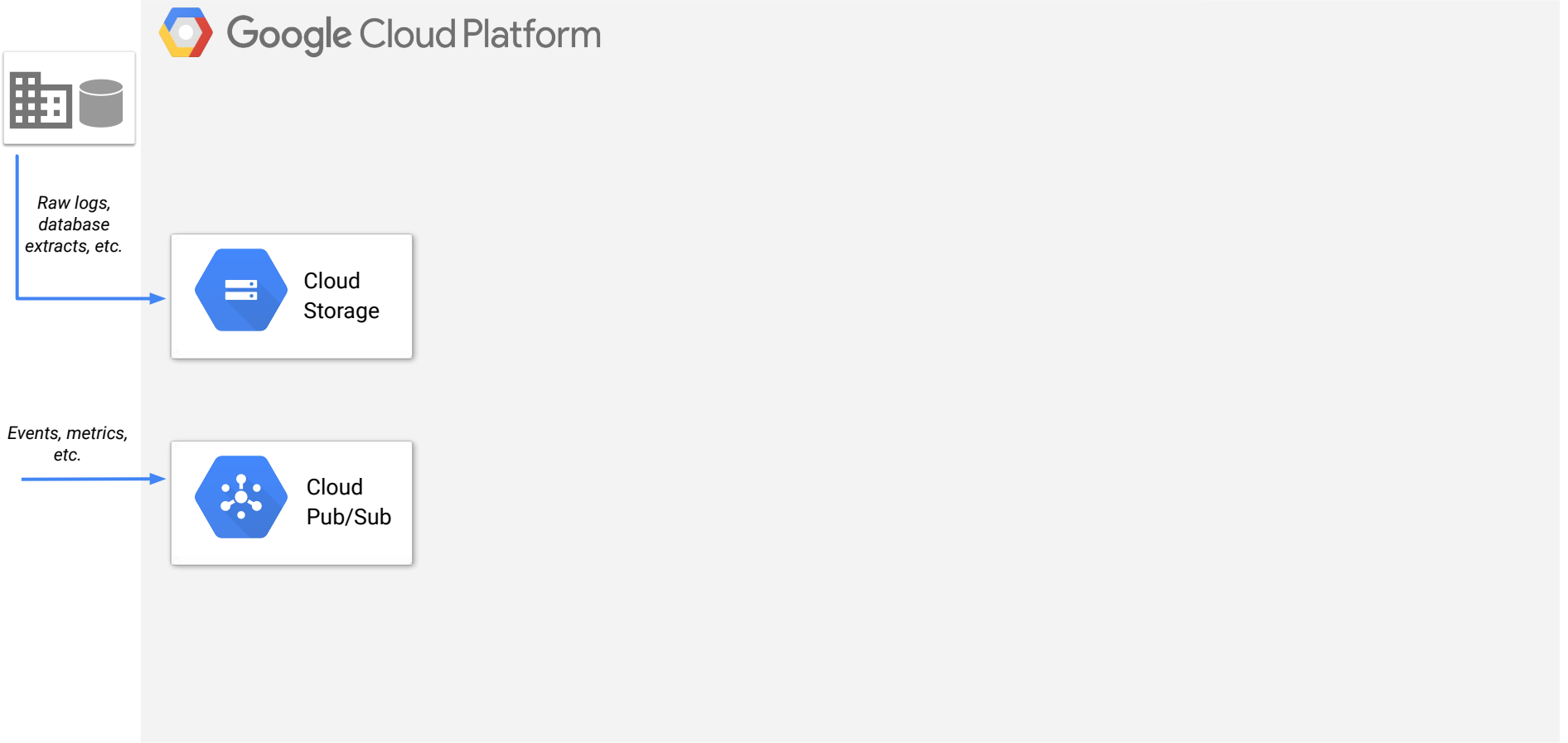


Google Cloud

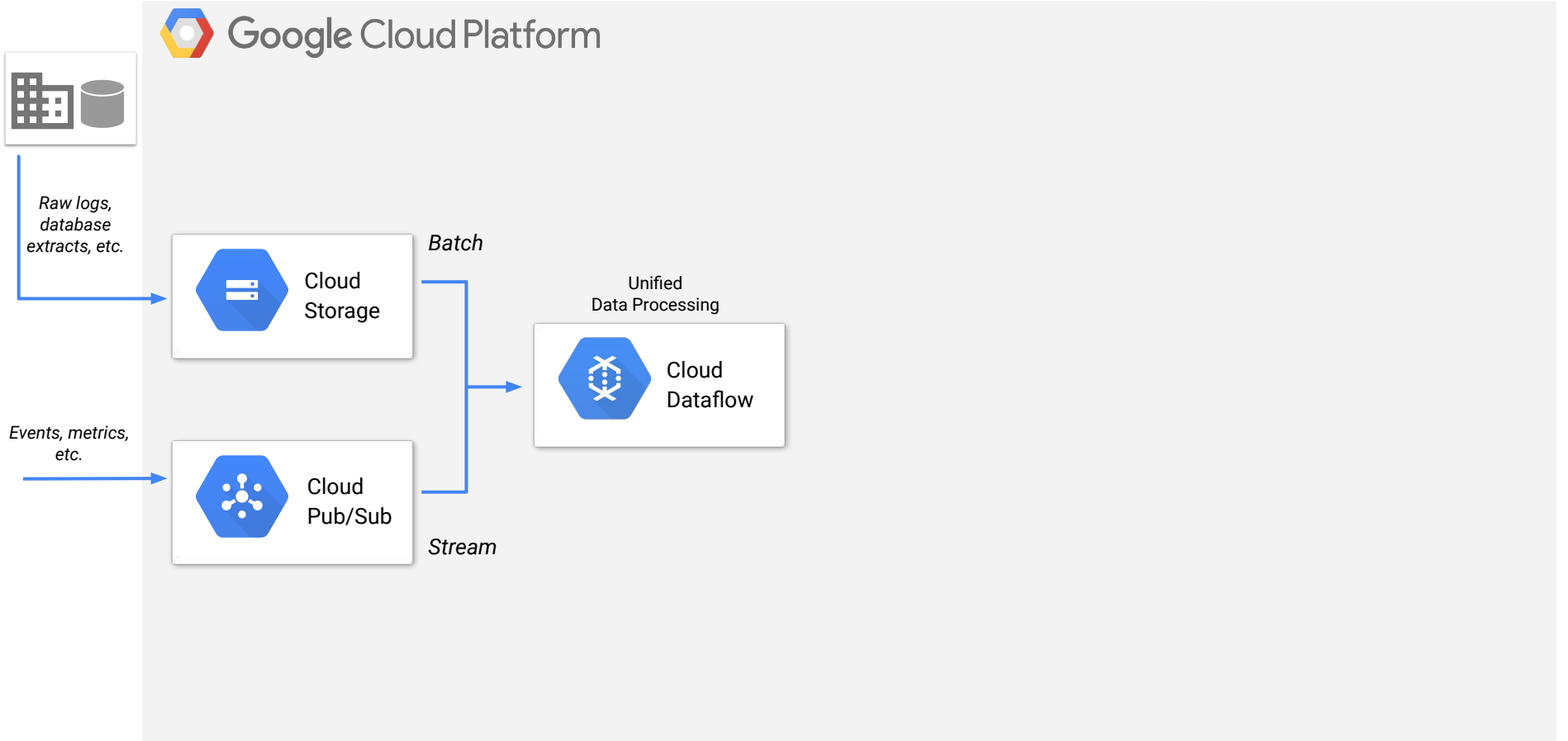
GCP Data Pipelines



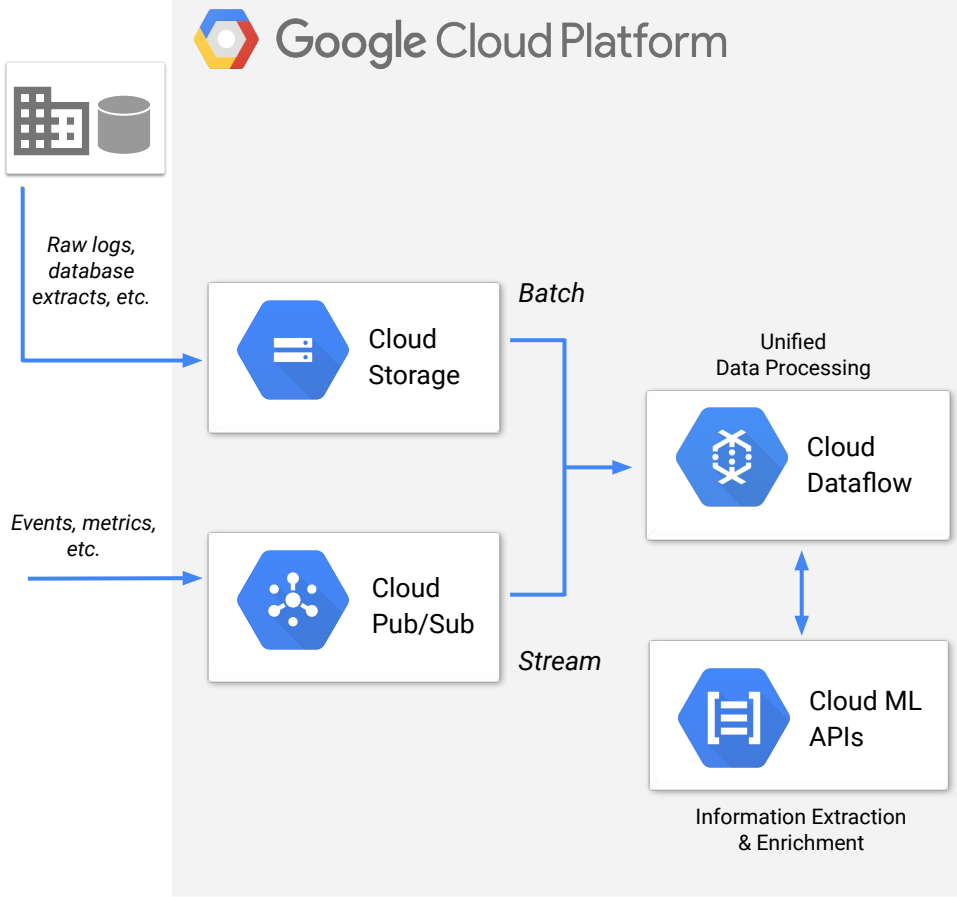
GCP Data Pipelines - Capture & Ingest



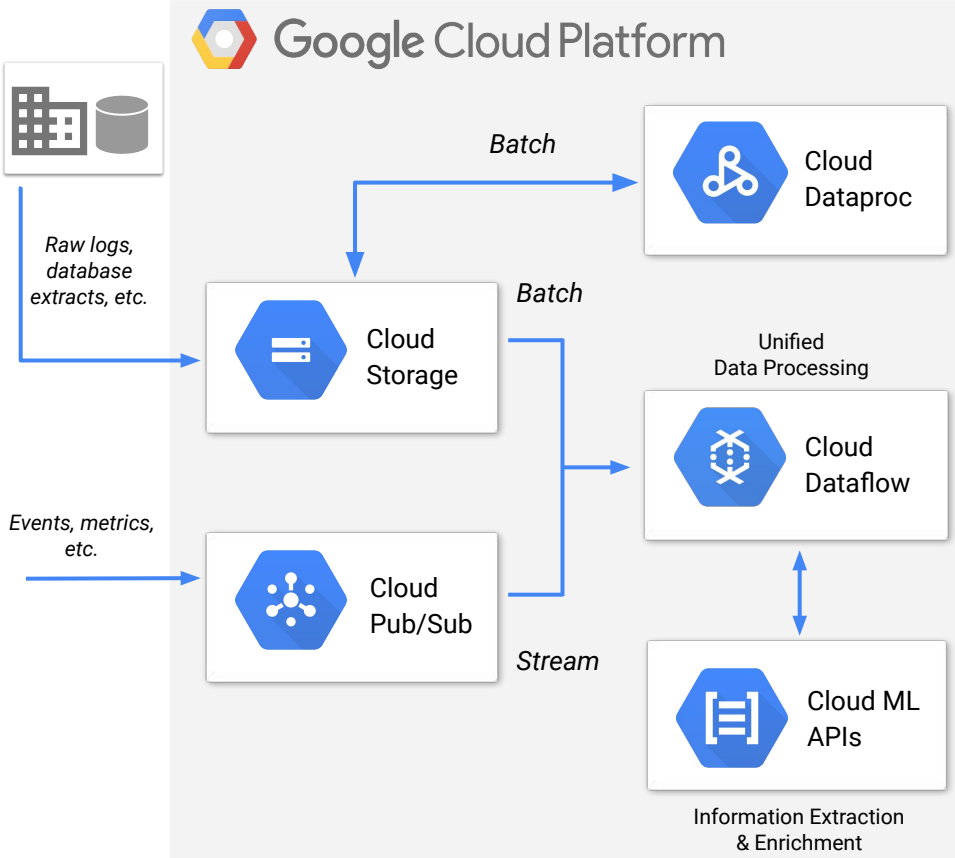
GCP Data Pipelines - Prepare & Process



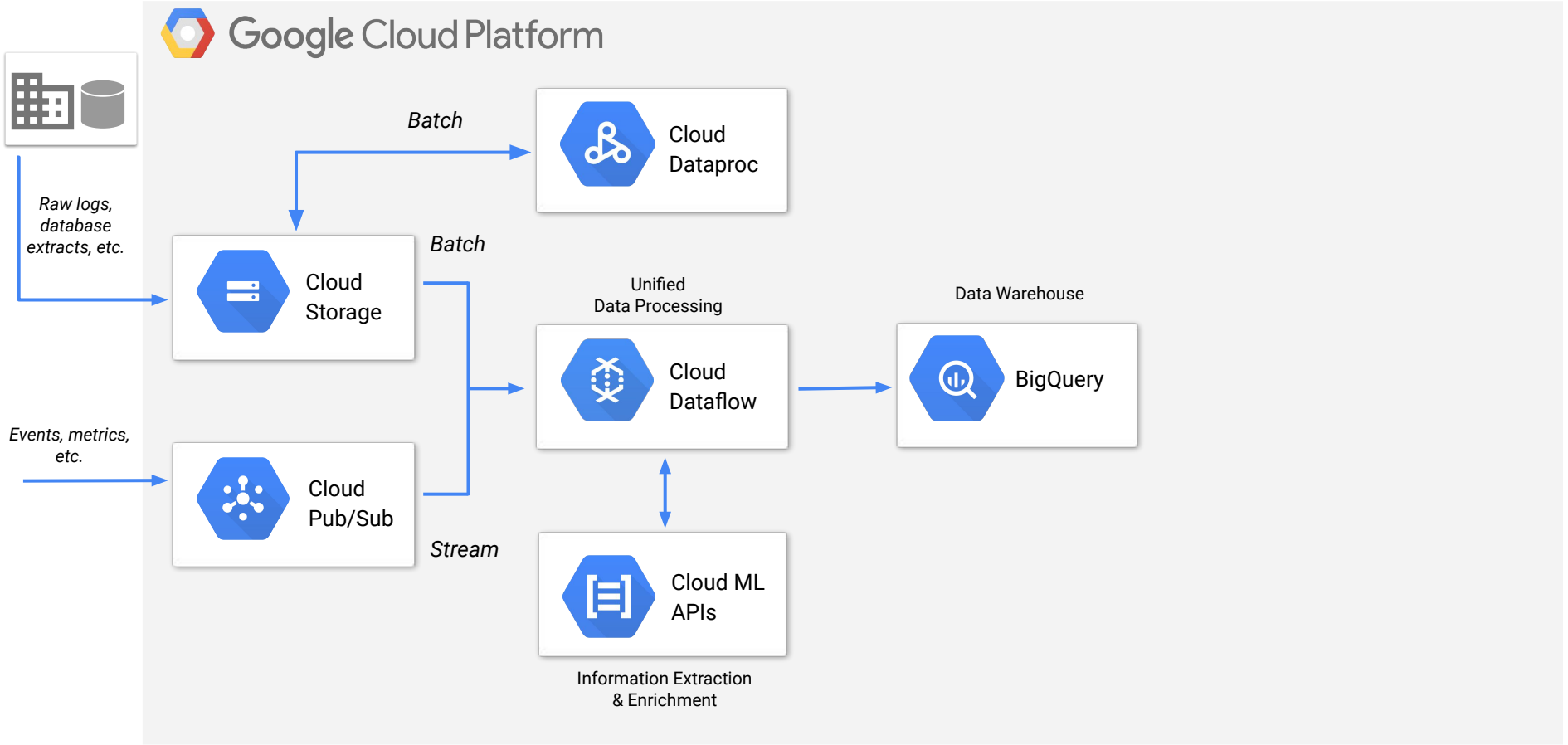
GCP Data Pipelines - Enrich



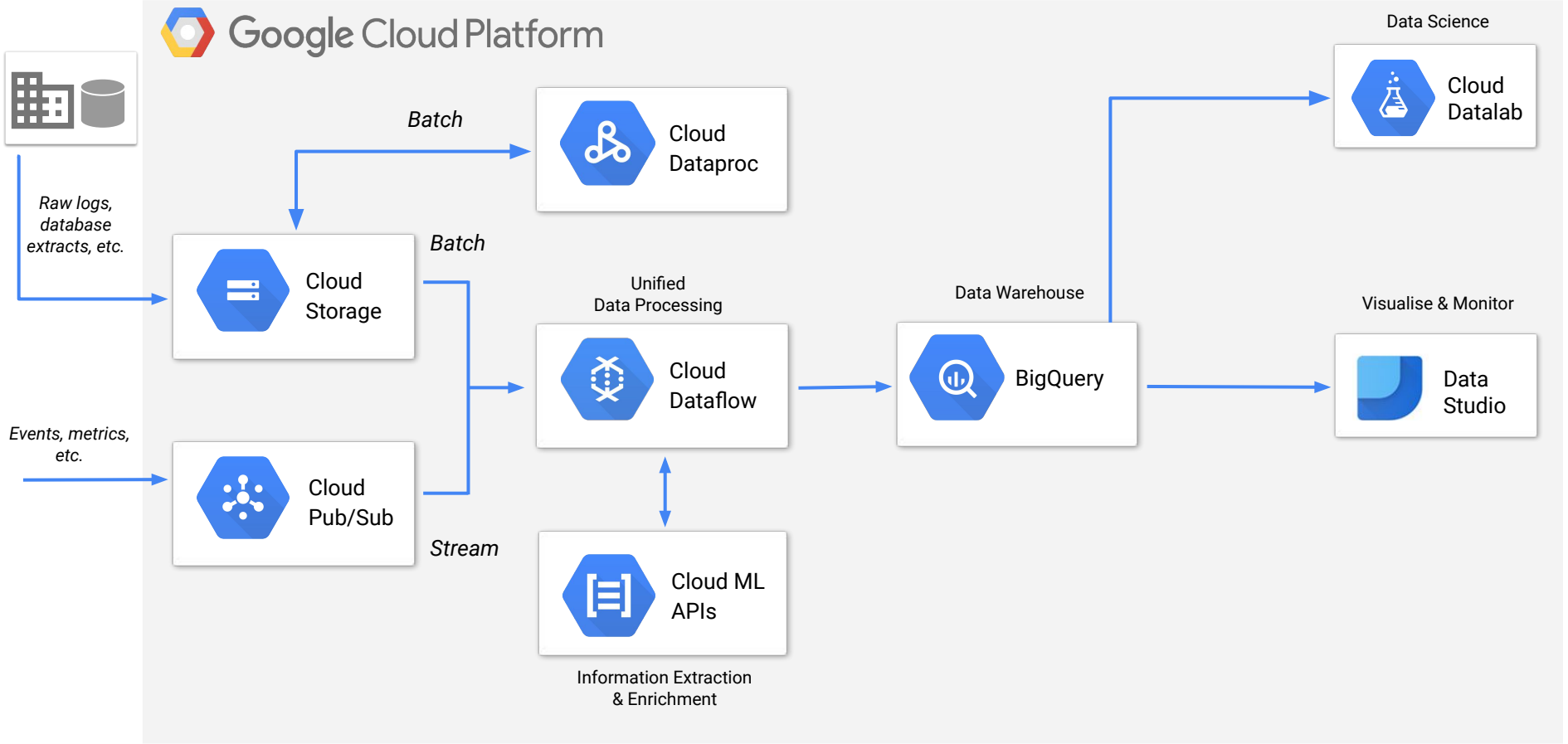
GCP Data Pipelines - Prepare & Process & Enrich



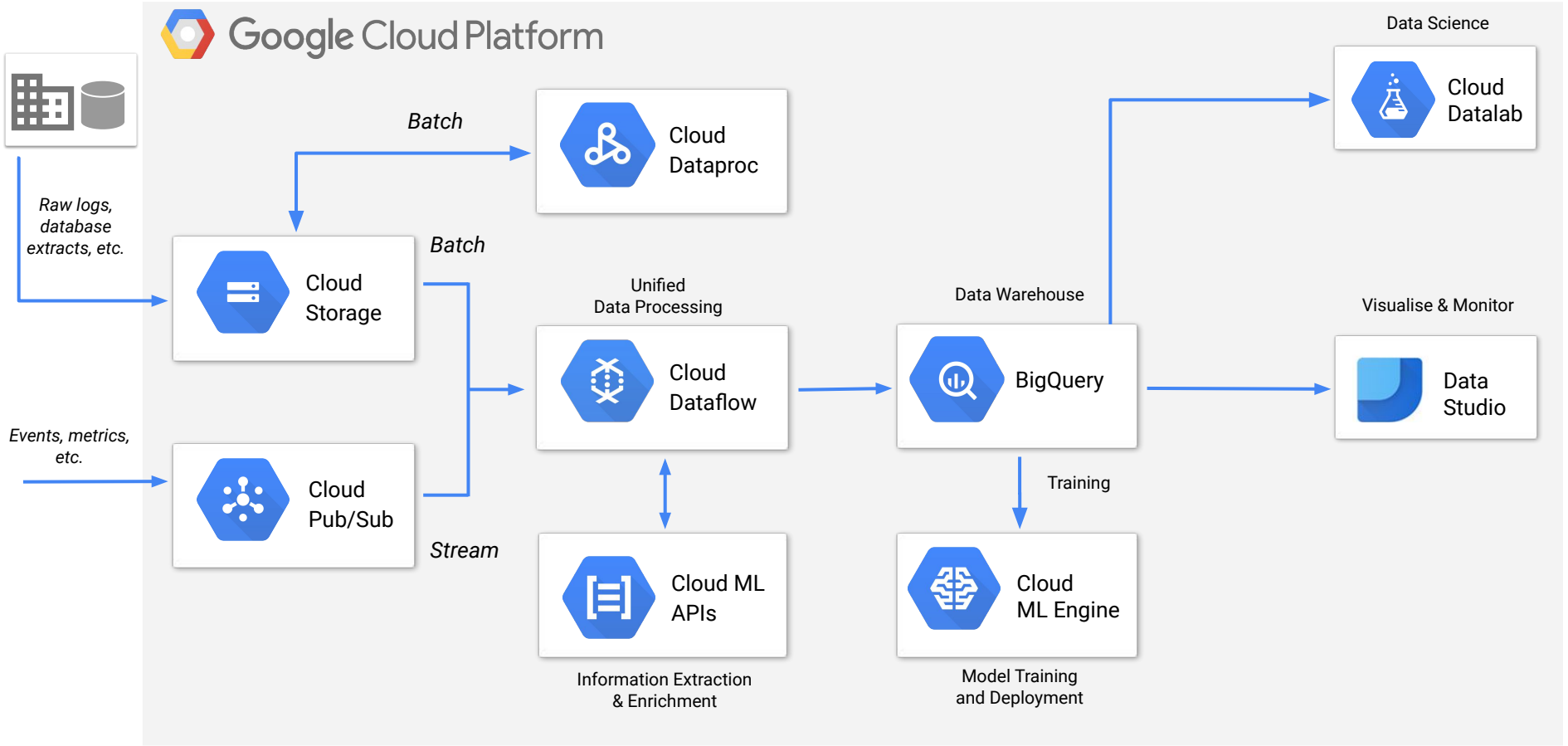
GCP Data Pipelines - Store & Analytics



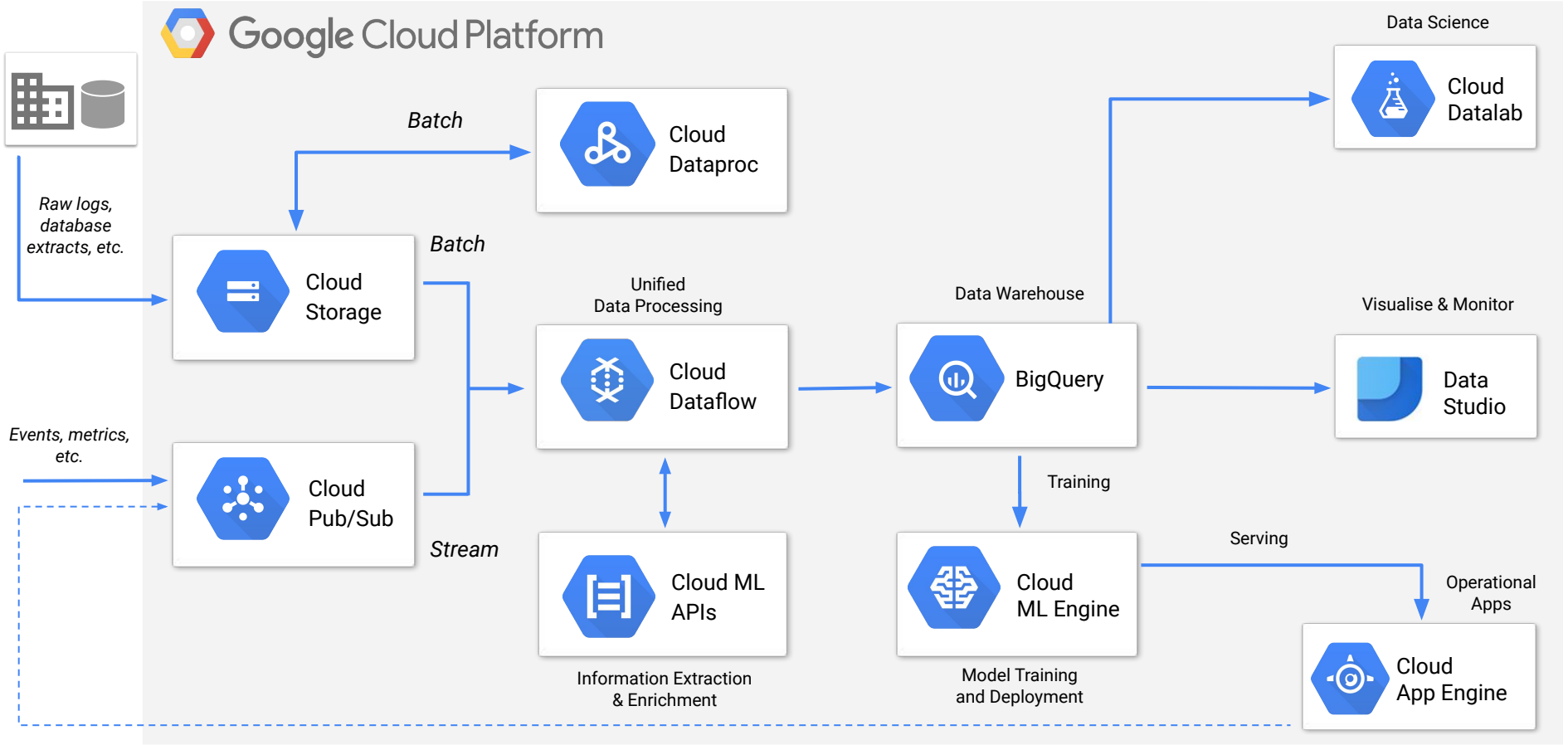
GCP Data Pipelines - Exploration & Collaboration



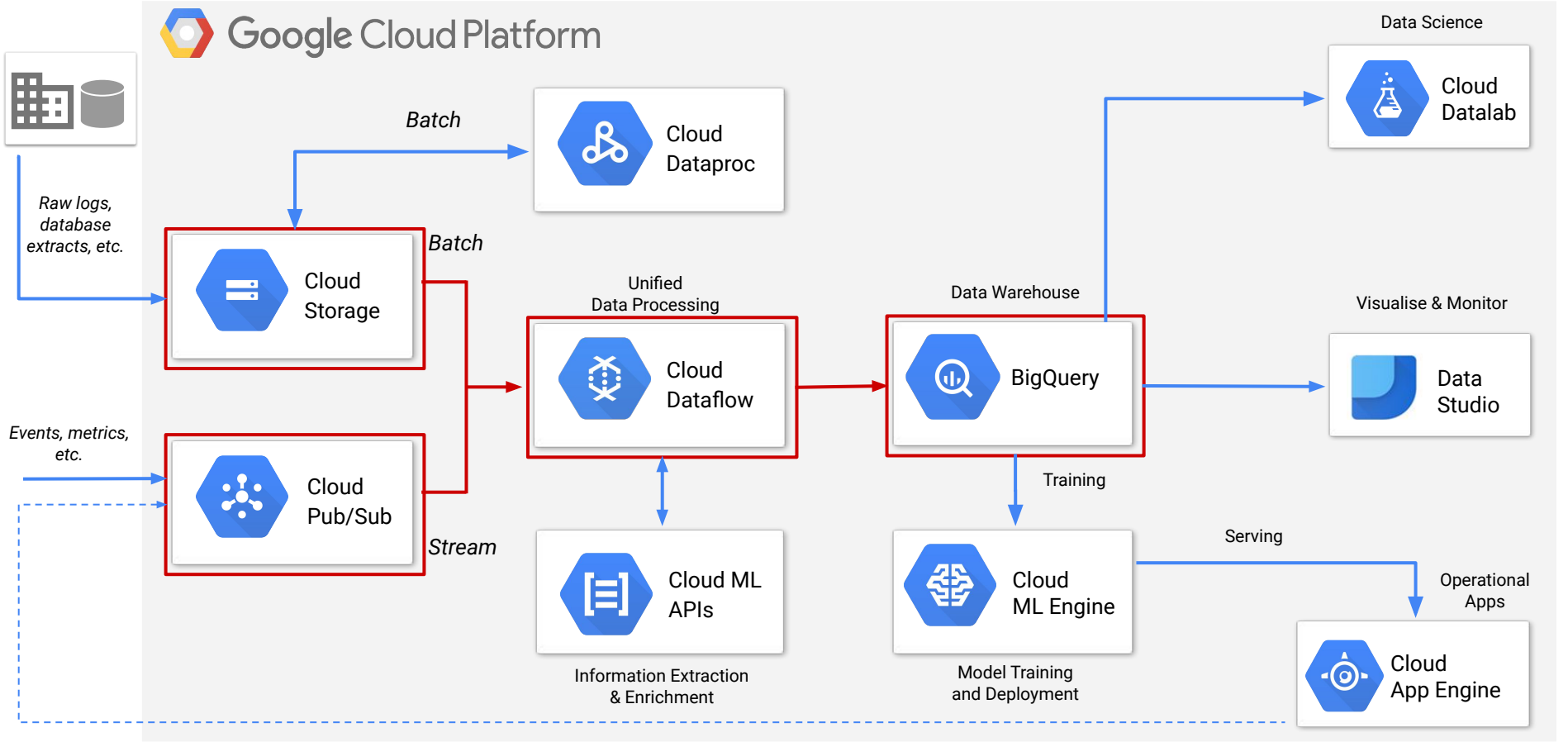
GCP Data Pipelines - Machine Learning



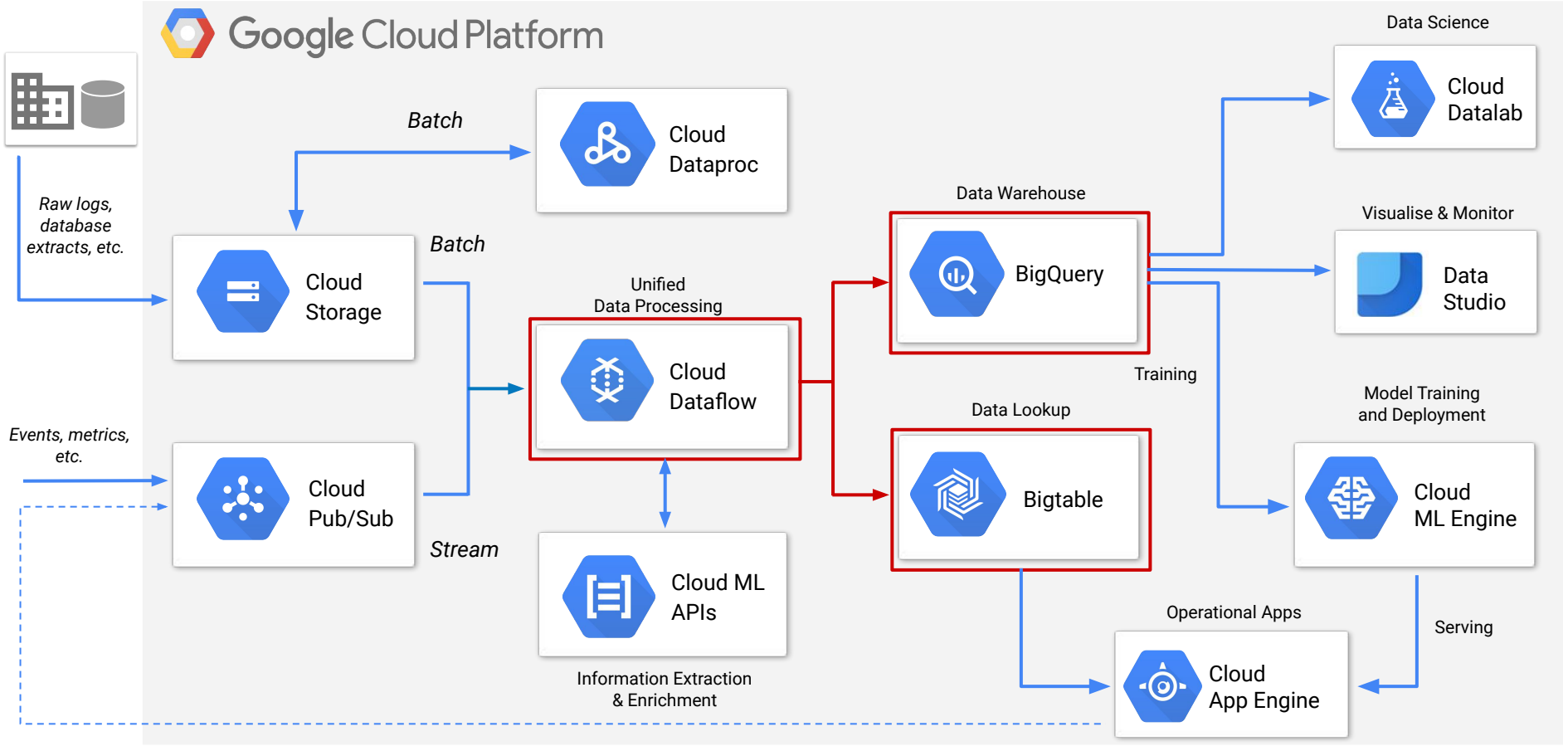
GCP Data Pipelines - Serving ML models



GCP Data Pipelines - Core Services



GCP Data Pipelines - Serving Low Latency data



GCP Data Pipelines - ML ETL integration