

Machine Intelligence at Google Scale

Vision, Video, NLP, Speech, Dialogflow
TensorFlow, Cloud ML Engine, AutoML

Guillaume Laforge

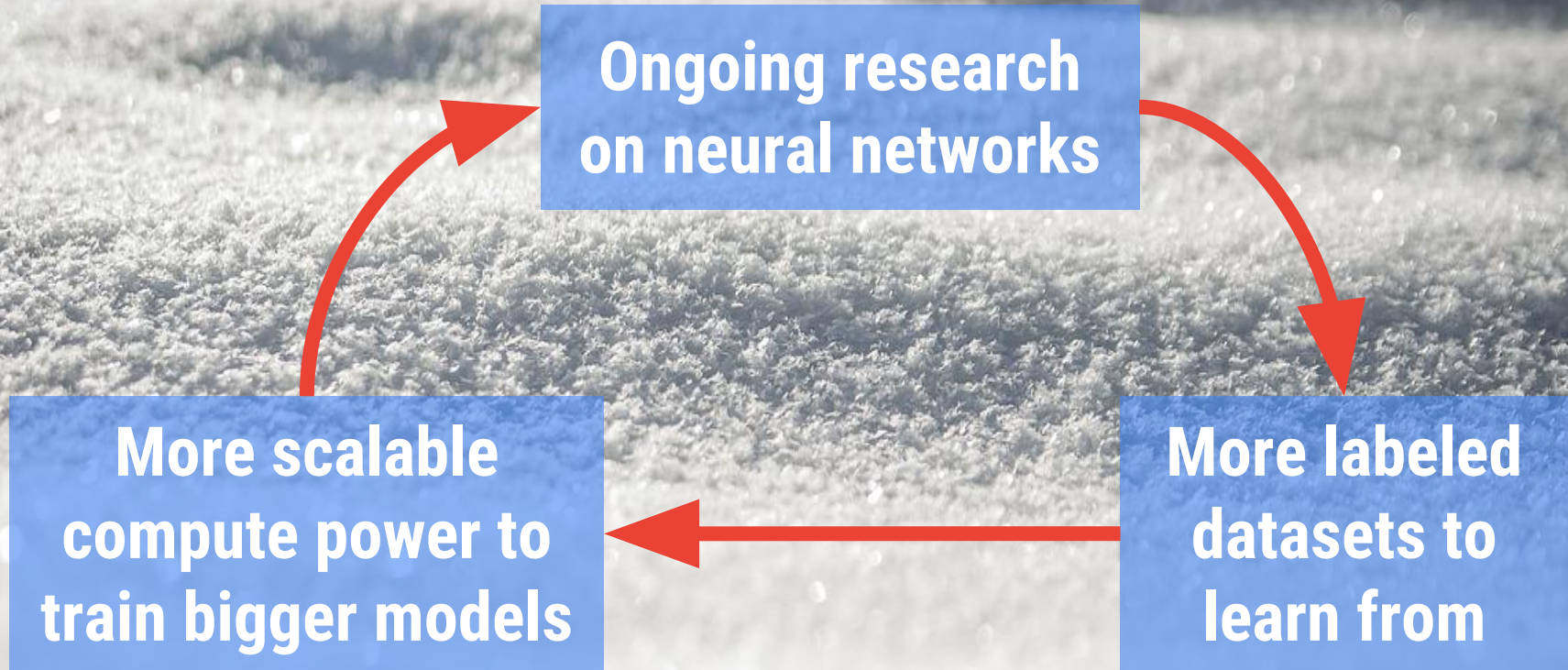
Developer Advocate
Google Cloud

@glaforge



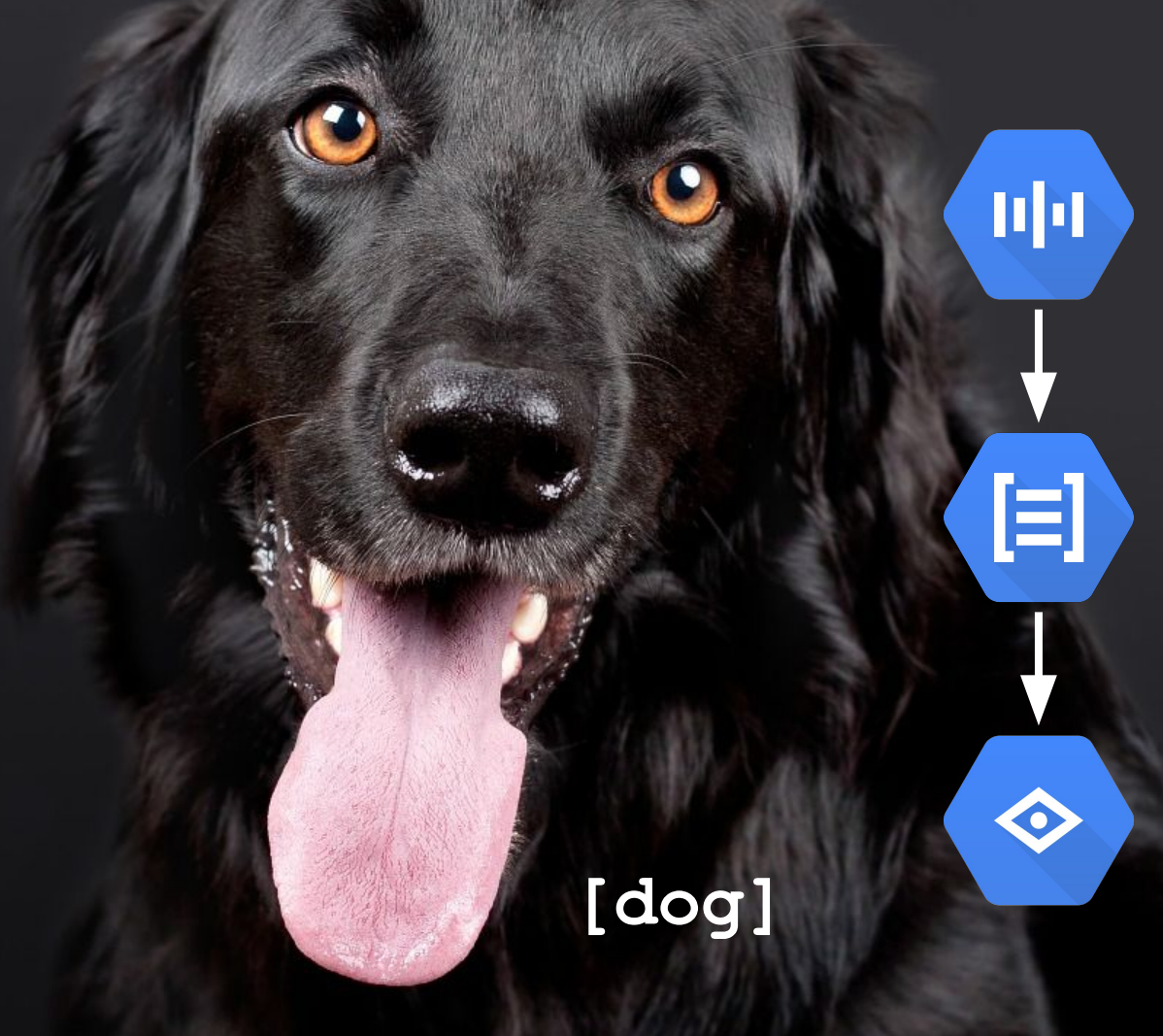
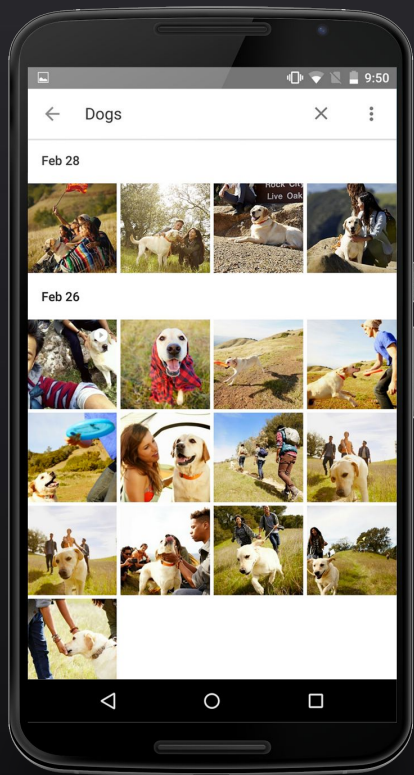
AlphaGo

How did we escape the AI winter?

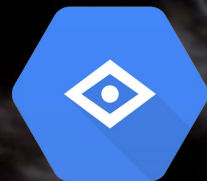
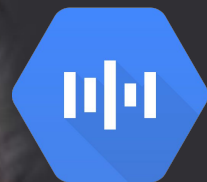




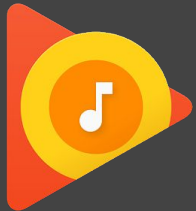
Google Photos



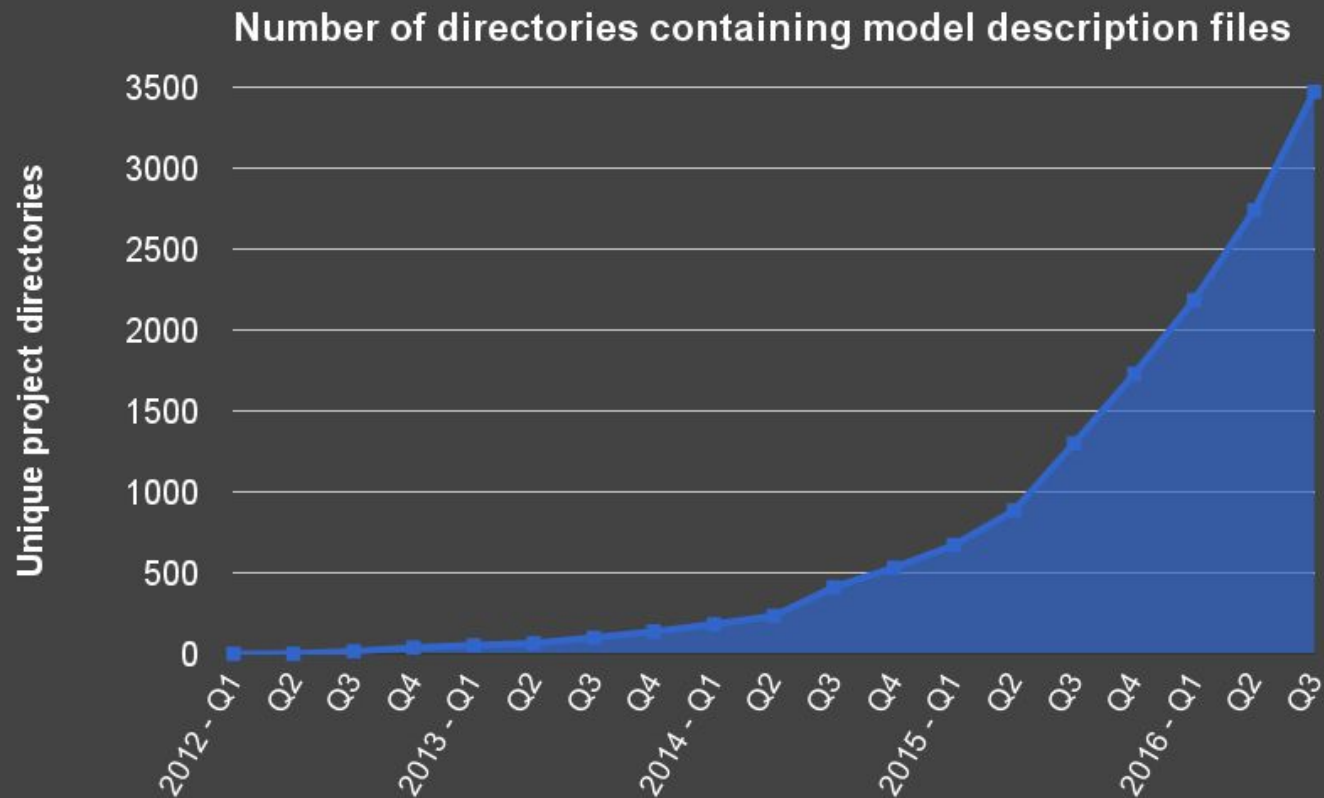
[dog]



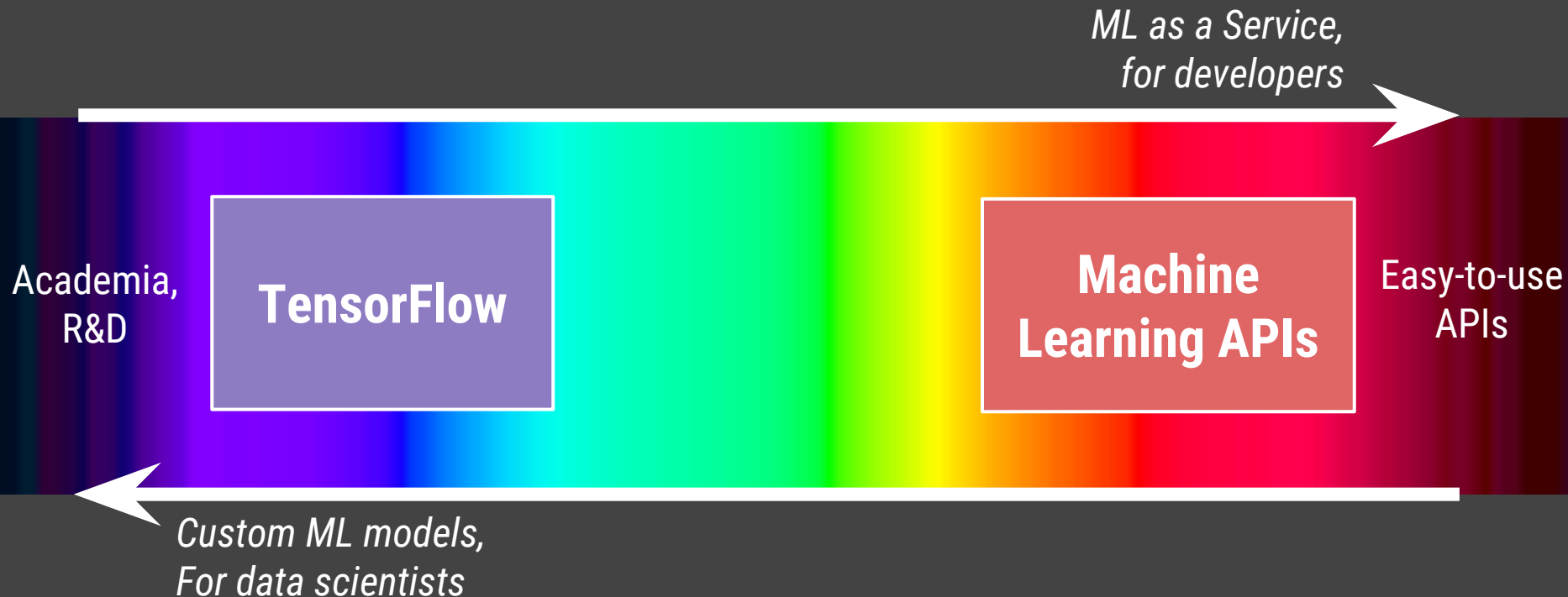
Machine Learning is everywhere at Google



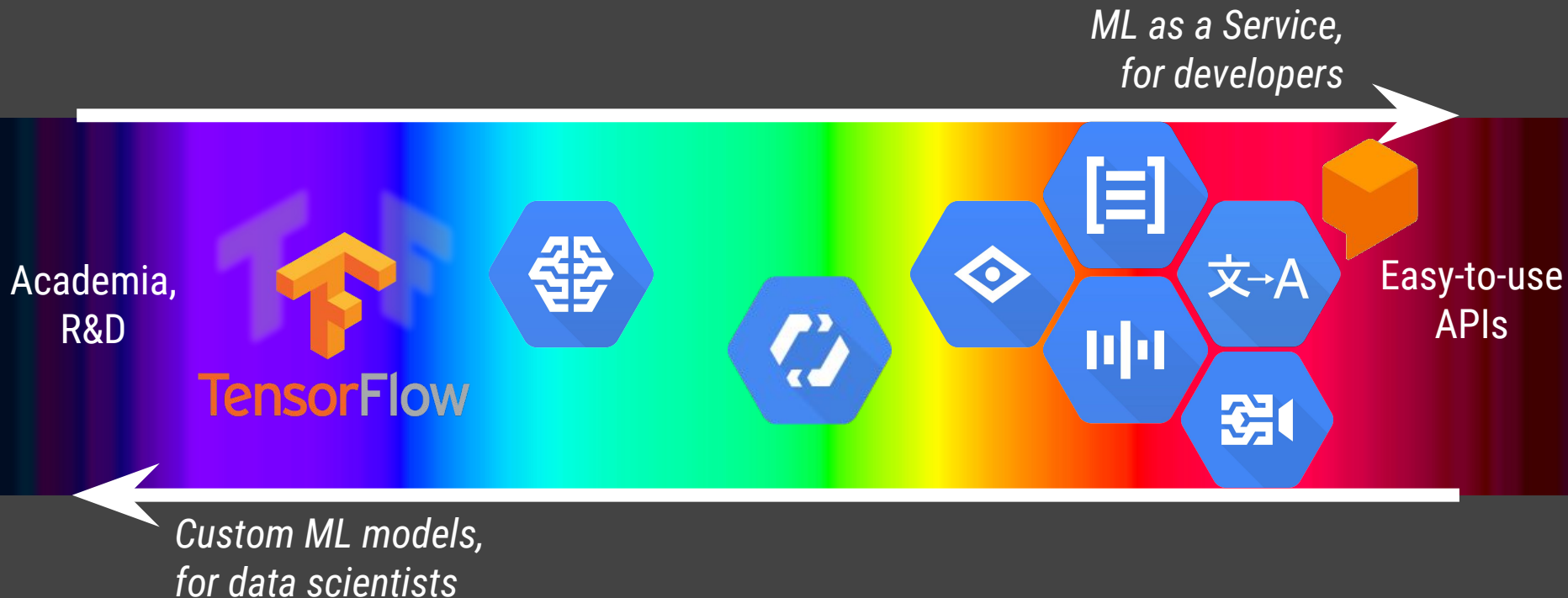
Machine Learning is everywhere at Google



The Machine Learning Spectrum



The Machine Learning Spectrum

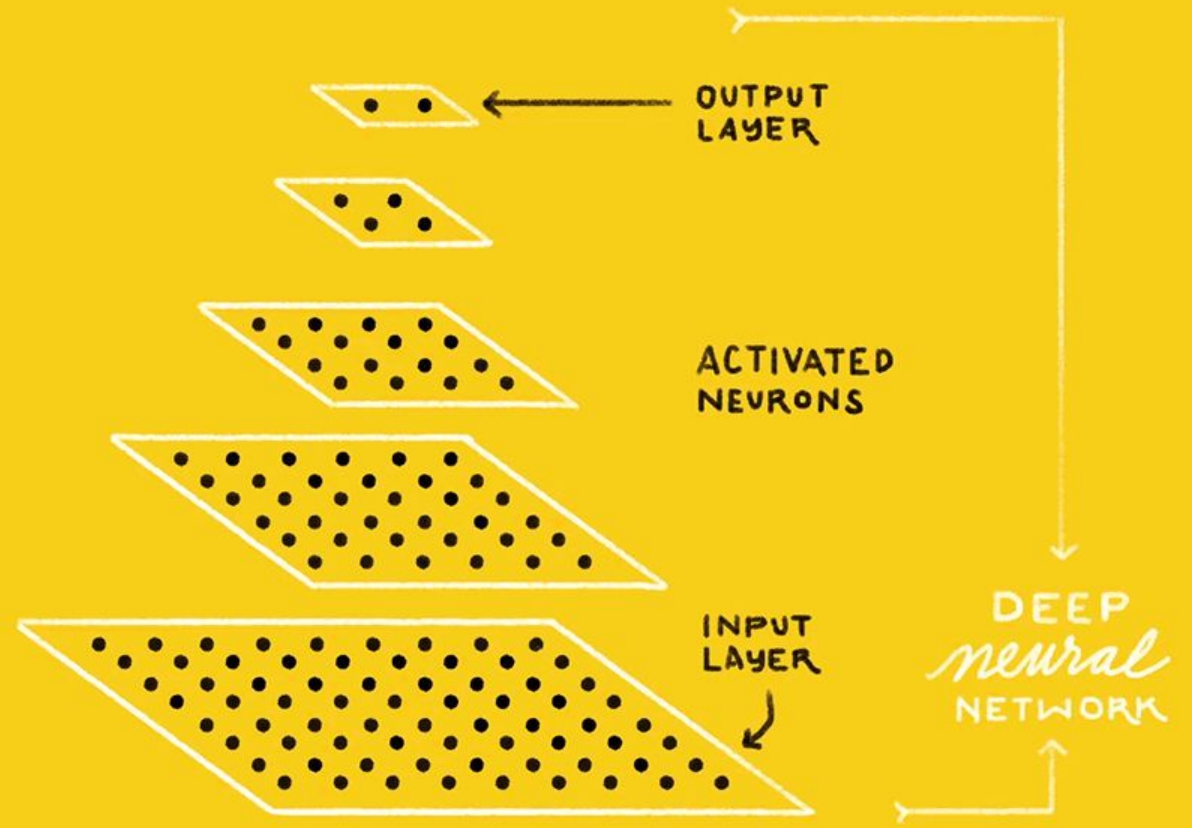


Machine learning is learning from
examples and *experience*

IS THIS A
CAT or DOG?



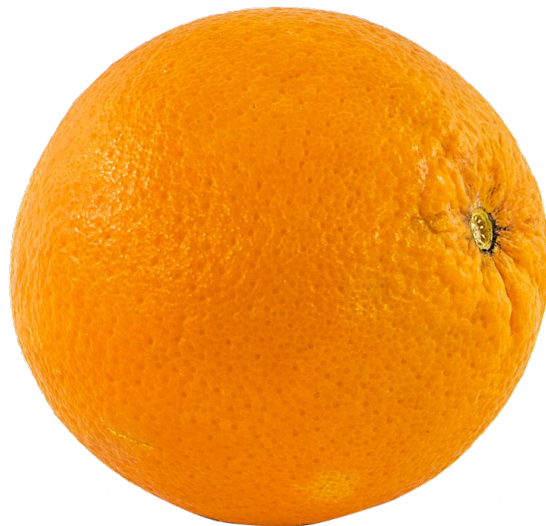
CAT DOG



Let's try some
human-powered
image detection



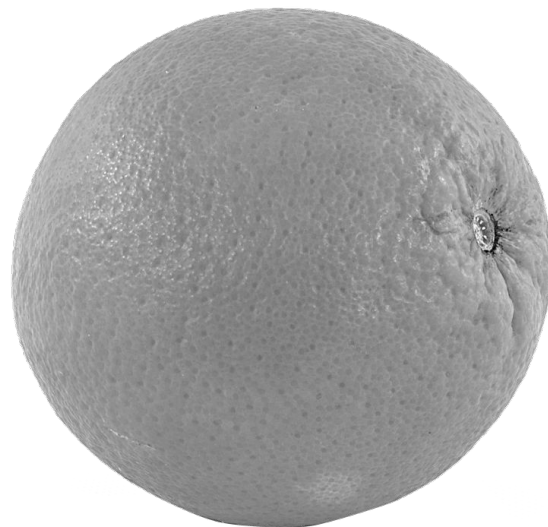
How would we do this without ML?



CC-BY-SA 2.0 Wikimedia Commons

https://commons.wikimedia.org/wiki/File:Apple_in_lightbox.png

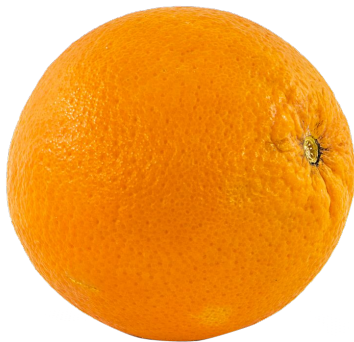
How would we do this without ML?



CC-BY-SA 2.0 Wikimedia Commons

https://commons.wikimedia.org/wiki/File:Apple_in_lightbox.png

How would we do this without ML?



CC-BY-SA 2.0 Wikimedia Commons

https://commons.wikimedia.org/wiki/File:Apple_in_lightbox.png

What about a dog and a mop? Easy, right?



CC-BY 4.0 Wikimedia Commons https://commons.wikimedia.org/wiki/File:Mop_and_bucket.jpg

Not so fast...



CC-BY-SA-2.5 Wikimedia Commons https://commons.wikimedia.org/wiki/File:Komondor_Westminster_Dog_Show_crop.jpg
CC-BY-2.0 Wikimedia Commons [https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_\(12487315865\).jpg](https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_(12487315865).jpg)
CC-BY-2.0 Petful <https://www.flickr.com/photos/petsadviser-pix/16395099127>
CC-BY-SA-2.0 Jeffrey Beall <https://www.flickr.com/photos/denverjeffrey/6903790333>

Machine Learning tools by Google at your disposal

Use your own data to train models



TensorFlow



Cloud Machine Learning Engine



Cloud AutoML

Machine Learning as an API



Cloud Vision API



Cloud Speech API



Dialogflow

Conversational Interfaces



Cloud Natural Language API



Cloud Translation API



Cloud Video Intelligence

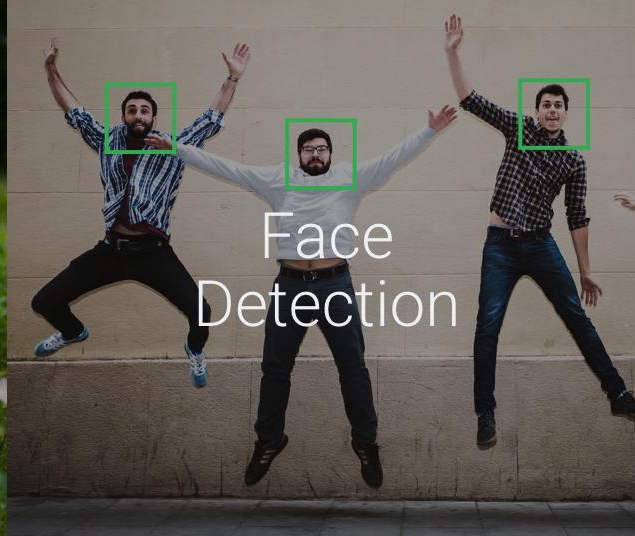


Vision API

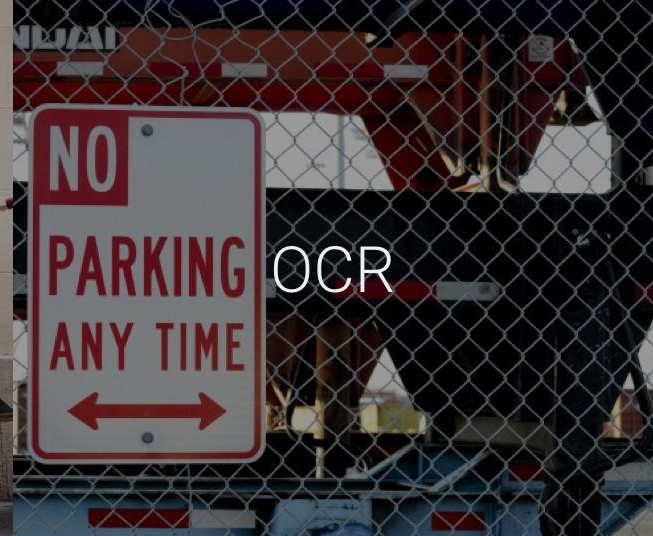
Complex image detection with
a simple REST request



Label
Detection



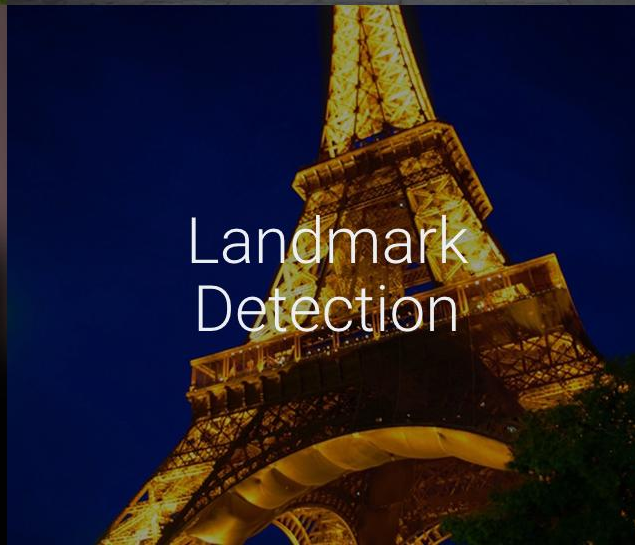
Face
Detection



OCR



Explicit Content
Detection



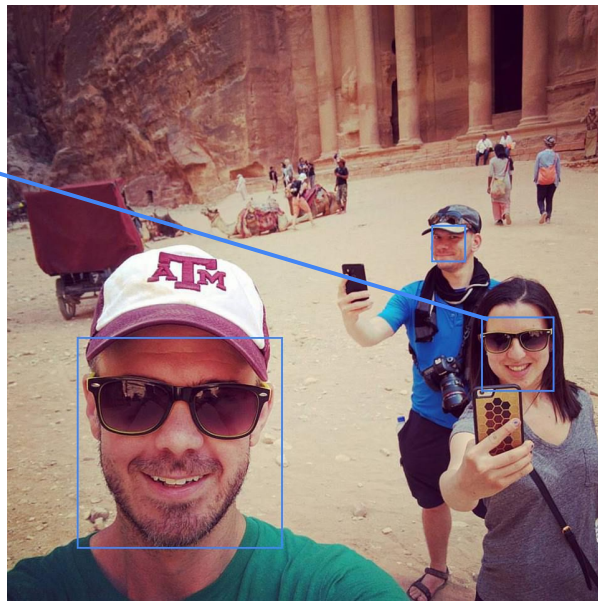
Landmark
Detection



Logo Detection

Face detection

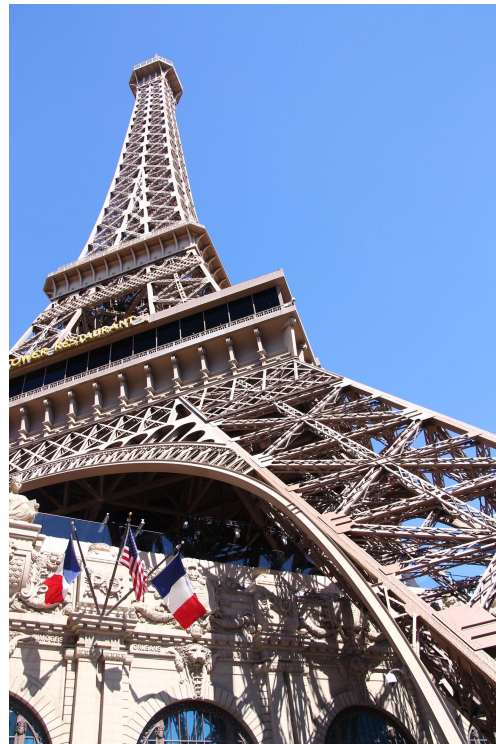
```
"faceAnnotations" : [  
  {  
    "headwearLikelihood" : "VERY_UNLIKELY",  
    "surpriseLikelihood" : "VERY_UNLIKELY",  
    "rollAngle" : -4.6490049,  
    "angerLikelihood" : "VERY_UNLIKELY",  
    "landmarks" : [  
      {  
        "type" : "LEFT_EYE",  
        "position" : {  
          "x" : 691.97974,  
          "y" : 373.11096,  
          "z" : 0.000037421443  
        }  
      }  
    ],  
    ...  
  },  
  ...  
],  
"boundingPoly" : {  
  "vertices" : [  
    {  
      "x" : 743,  
      "y" : 449  
    }  
  ],  
  ...  
}
```



```
"detectionConfidence" : 0.93568963,  
"joyLikelihood" : "VERY_LIKELY",  
"panAngle" : 4.150538,  
"sorrowLikelihood" : "VERY_UNLIKELY",  
"tiltAngle" : -19.377356,  
"underExposedLikelihood" : "VERY_UNLIKELY",  
"blurredLikelihood" : "VERY_UNLIKELY"
```

Landmark detection

```
"landmarkAnnotations": [  
  {  
    "mid": "/m/0348s6",  
    "description": "Paris Hotel and Casino",  
    "score": 80,  
    "boundingPoly": {  
      "vertices": [  
        {  
          "x": 117,  
          "y": 479  
        },  
        ...  
      ]  
    },  
    "locations": [  
      {  
        "latLng": {  
          "latitude": 36.11221,  
          "longitude": -115.172596  
        }  
      }  
    ]  
  }  
]
```



CC-BY-SA-3.0 Wikimedia Commons <https://commons.wikimedia.org/wiki/File:Las-Vegas-Paris-Hotel-Eiffel-Tower-8307.jpg>

Web annotations

```
{  
  "entityId": "/m/0gff2yr",  
  "score": 5.92256,  
  "description": "ArtScience Museum"  
}
```

```
{  
  "entityId": "/m/016ms7",  
  "score": 1.44038,  
  "description": "Ford Anglia"  
}
```



CC-BY 2.0 Rev Stan: <https://www.flickr.com/photos/revstan/6865880240>

```
{  
  "entityId": "/m/0h898pd",  
  "score": 7.4162,  
  "description": "Harry Potter (Literary Series)"  
}
```

Web annotations



CC-BY 2.0 Rev Stan: <https://www.flickr.com/photos/revstan/6865880240>

```
"fullMatchingImages": [{  
  "url":  
    "https://upload.wikimedia.org/wikipedia/commons/6/6d/Flying_Ford_Angl  
    ia_from_Harry_Potter_and_the_Chamber_of_Secrets_at_the_ArtScience_Mus  
    eum,_Singapore_-_20120608.jpg",  
  "score": 0.34952533  
},  
  ...  
]
```

```
"partialMatchingImages": [{  
  "url":  
    "https://muckysock.files.wordpress.com/2012/06/img_2730.jpg",  
  "score": 0.887808  
},  
  ...  
]
```

```
"pagesWithMatchingImages": [{  
  "url":  
    "https://www.haikudeck.com/harry-potter-and-chamber-of-secrets--educa  
    tion-presentation-SKZRnA02UH",  
  "score": 53.212971  
},  
  ...  
]
```

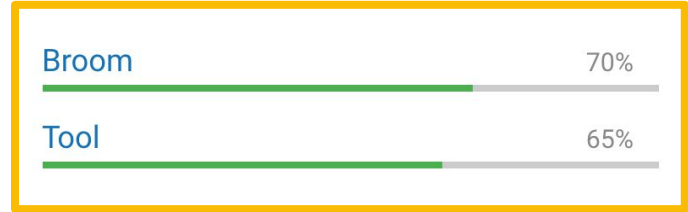
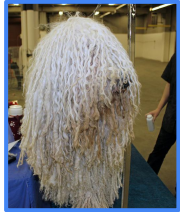
In case you were wondering...



Dog	99%
Mammal	93%
Dog Breed	91%
Vertebrate	91%
Komondor	77%
Dog Like Mammal	76%
Glen Of Imaal Terrier	51%
Dog Crossbreeds	51%

CC-BY-SA-2.5 Wikimedia Commons https://commons.wikimedia.org/wiki/File:Komondor_Westminster_Dog_Show_crop.jpg
CC-BY-2.0 Wikimedia Commons [https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_\(12487315865\).jpg](https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_(12487315865).jpg)
CC-BY-2.0 Petful <https://www.flickr.com/photos/petsadviser-pix/16395099127>
CC-BY-SA-2.0 Jeffrey Beall <https://www.flickr.com/photos/denverjeffrey/6903790333>

In case you were wondering...



CC-BY-SA-2.5 Wikimedia Commons https://commons.wikimedia.org/wiki/File:Komondor_Westminster_Dog_Show_crop.jpg
CC-BY-2.0 Wikimedia Commons [https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_\(12487315865\).jpg](https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_(12487315865).jpg)
CC-BY-2.0 Petful <https://www.flickr.com/photos/petsadviser-pix/16395099127>
CC-BY-SA-2.0 Jeffrey Beall <https://www.flickr.com/photos/denverjeffrey/6903790333>

In case you were wondering...



CC-BY-SA-2.5 Wikimedia Commons https://commons.wikimedia.org/wiki/File:Komondor_Westminster_Dog_Show_crop.jpg
CC-BY-2.0 Wikimedia Commons [https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_\(12487315865\).jpg](https://commons.wikimedia.org/wiki/File:2014_Westminster_Kennel_Club_Dog_Show_(12487315865).jpg)
CC-BY-2.0 Petful <https://www.flickr.com/photos/petsadviser-pix/16395099127>
CC-BY-SA-2.0 Jeffrey Beall <https://www.flickr.com/photos/denverjeffrey/6903790333>



Natural Language API

Extract entities, sentiment,
and syntax from text

Extract entities

Joanne "Jo" Rowling, pen names J. K. Rowling and Robert Galbraith, is a British novelist, screenwriter and film producer best known as the author of the Harry Potter fantasy series

Extract entities

*Joanne "Jo" Rowling, pen names **J. K. Rowling** and **Robert Galbraith**, is a **British** novelist, screenwriter and film producer best known as the author of the **Harry Potter** fantasy series*

Extract entities

```
{
  "name": "Joanne 'Jo' Rowling",
  "type": "PERSON",
  "metadata": {
    "mid": "/m/042xh",
    "wikipedia_url": "http://en.wikipedia.org/wiki/J._K._Rowling"
  }
}
```

Joanne "Jo" Rowling, pen names *J. K. Rowling* and *Robert Galbraith*, is a **British** novelist, screenwriter and film producer best known as the author of the **Harry Potter** fantasy series

```
{
  "name": "British",
  "type": "LOCATION",
  "metadata": {
    "mid": "/m/07ssc",
    "wikipedia_url": "http://en.wikipedia.org/wiki/United_Kingdom"
  }
}
```

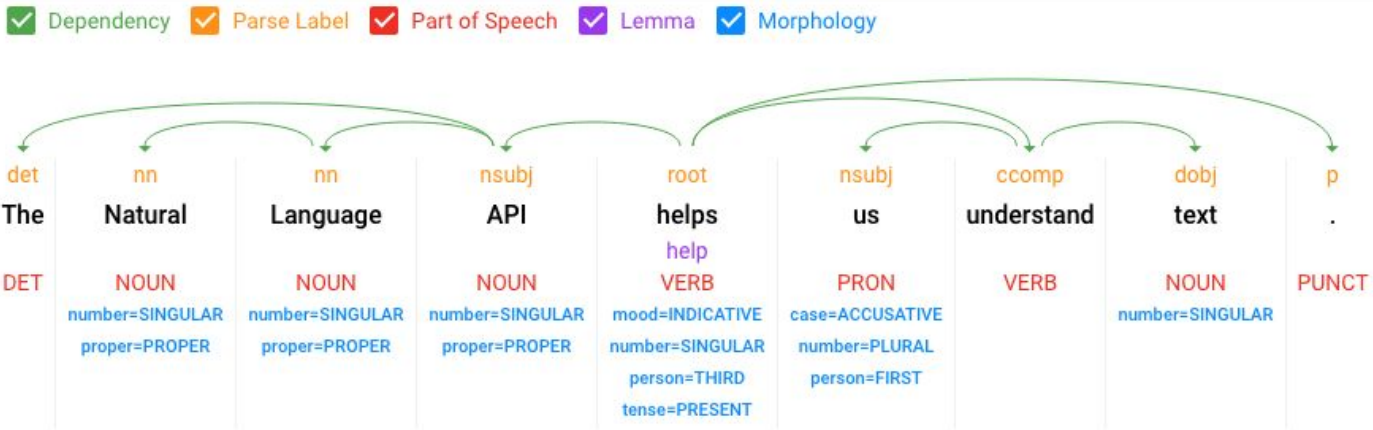
```
{
  "name": "Harry Potter",
  "type": "PERSON",
  "metadata": {
    "mid": "/m/078ffw",
    "wikipedia_url": "http://en.wikipedia.org/wiki/Harry_Potter"
  }
}
```

Analyze sentiment

"The food was excellent, I would definitely go back!"

```
{  
  "documentSentiment": {  
    "score": 0.8,  
    "magnitude": 0.8  
  }  
}
```


Analyze syntax





Speech API

Speech to text transcription in
over 80 languages

Speech API features

Speech Recognition

Recognizes over **110 languages** & variants.

Powered by deep learning neural networking to power your applications.

Real-time results

Can **stream text results**, returning partial recognition results as they become available.

Can also be run on buffered or archived audio files.

Noise Robustness

No need for signal processing or noise cancellation before calling API.

Can **handle noisy audio** from a variety of environments.

Context-Aware

Can provide **context hints** for improved accuracy.

Especially useful for device and app use cases.

Speech timestamps

**Search for text
in audio files**

```
"transcript": "Hello World...",
"confidence": 0.96596134,
"words": [
  {
    "startTime": "1.400s",
    "endTime": "1.800s",
    "word": "Hello"
  },
  {
    "startTime": "1.800s",
    "endTime": "2.300s",
    "word": "World"
  },
  ...
]
```





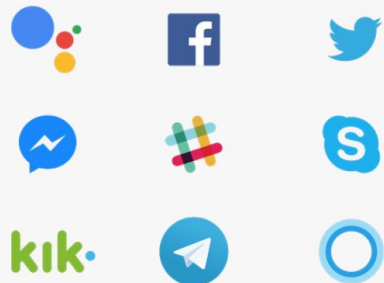
Dialogflow

Build natural and rich
conversational experiences

Build chatbots and conversational interfaces

The screenshot displays the Dialogflow console interface. On the left is a navigation sidebar with the following items: Dialogflow logo, 'deepdive-dialogflow' with a settings gear, 'en' with a plus sign, 'Intents' (highlighted in blue) with a plus sign, 'Entities' with a plus sign, 'Training [beta]', 'Integrations', 'Analytics [new]', 'Fulfillment', and 'Prebuilt Agents'. The main area shows a configuration for a chatbot named '08---entity-list-and-fallback---fastfood'. At the top right of this area is a blue 'SAVE' button and a three-dot menu icon. Below the title, there is a 'Contexts' section with an upward arrow and the text 'Add input context'. Underneath, there is a list of contexts: '1 fastfoodctx' with a close icon and the text 'Add output context'. Below the contexts is a 'User says' section with a search bar labeled 'Search in user says' and an upward arrow. This section contains a list of user expressions: 'Add user expression', 'eat a cheese burger', 'I'd like to eat a cheese burger, a bigmac and a coke', and 'I want to eat a hamburger and a coke'. The words 'cheese burger', 'bigmac', 'coke', and 'hamburger' are highlighted in yellow in the original image.

Build chatbots and conversational interfaces



On any platform

Bring your conversational app to any platform your users are on, such as the Google Assistant, Slack, Cortana, Alexa and Facebook Messenger.



Across devices

Whether your users are on-the-go or at home, engage with them through wearables, phones, cars, speakers and other smart devices.



Around the world

Broaden your reach globally with 14+ supported languages including Spanish, French, and Japanese.



Translation API

Translate text in 100+ languages

airbnb – connecting guests through translation

- 60% of Airbnb bookings connect people who use the app in different languages
- Using the Translation API to translate listings, reviews, and conversations significantly improves a guest's likelihood to book



Calling the translation API

```
import com.google.cloud.translate.*;
import com.google.cloud.translate.Translate.*;

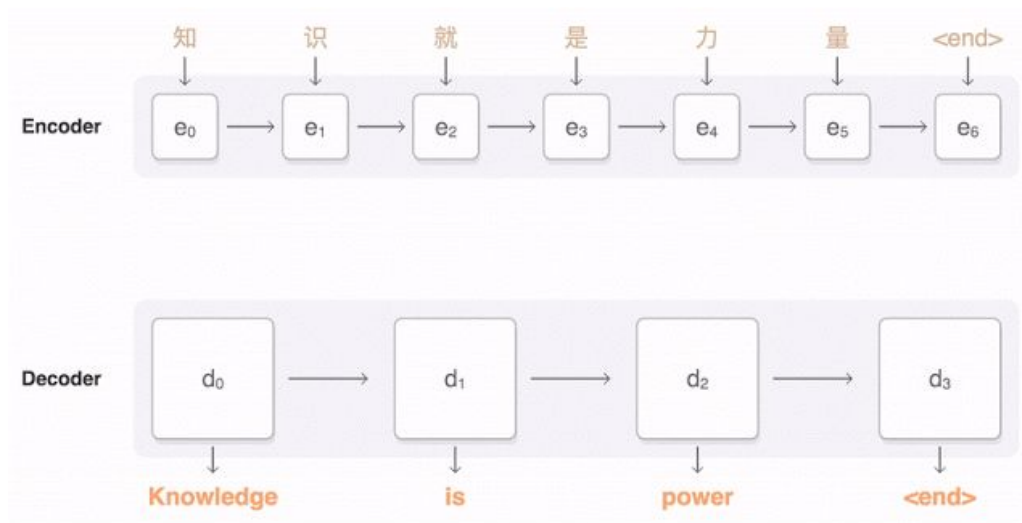
Translate translate =
    TranslateOptions.getDefaultInstance()
        .getService();

String text = "Hello, world!";

Translation translation =
    translate.translate(
        text,
        TranslateOption.sourceLanguage("en"),
        TranslateOption.targetLanguage("de"));

System.out.printf("Translation: %s%n",
    translation.getTranslatedText());
```

Neural machine translation



Learn more: bit.ly/nyt-ai-awakening

Neural machine translation improvements ⚡

Original Spanish Text

El señor Dursley era el director de una empresa llamada Grunnings, que fabricaba taladros. Era un hombre corpulento y rollizo, casi sin cuello, aunque con un bigote inmenso. La señora Dursley era delgada, rubia y tenía un cuello casi el doble de largo de lo habitual, lo que le resultaba muy útil, ya que pasaba la mayor parte del tiempo estirándolo por encima de la valla de los jardines para espiar a sus vecinos

First generation translation

Mr. Dursley was the director of a company called Grunnings, which **made** drills. He was a big beefy man, almost **neckless**, albeit with a huge mustache. Mrs. Dursley was thin and blonde and had a neck **almost twice longer than usual**, so it was very useful, since **he** spent most of the time stretching it over the **fence of the gardens** to spy on **their** neighbors

Neural Machine Translation

Mr. Dursley was the director of a company called Grunnings, which **manufactured** drills. He was a big, plump man, almost **without a neck**, but with a huge mustache. Mrs. Dursley was thin, blond, and had a neck **almost twice as long as usual**, which was very useful, since **she** spent most of the time stretching it over the **garden fence** to spy on **her** neighbors



Video Intelligence API

Understand your video's entities
at shot, frame, or video level

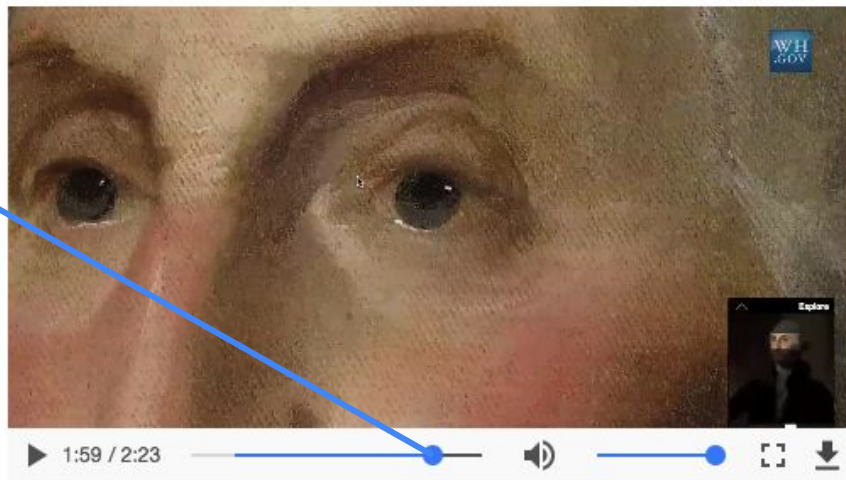
Video API Response: Label detection

```
{  
  "description": "Bird's-eye view",  
  "language_code": "en-us",  
  "locations": {  
    "segment": {  
      "start_time_offset": 71905212,  
      "end_time_offset": 73740392  
    },  
    "confidence": 0.96653205  
  }  
}
```



Video API Response: Label detection

```
{  
  "description": "Portrait",  
  "language_code": "en-us",  
  "locations": {  
    "segment": {  
      "start_time_offset": 116991989  
      "end_time_offset": 118243219  
    },  
    "confidence": 0.8332939  
  }  
}
```





TensorFlow

Google's Open Source framework
for deep neural networks

TensorFlow – Google’s 2nd gen. OSS deep learning library

- Provides APIs in Python and C++ (Java & Go experimental)
 - To describe Machine Learning models
 - To implement Machine Learning algorithms
- Supported:
 - Regression models
 - Neural networks & Deep learning
 - Convolutional Neural Networks
 - Recurrent Neural Networks
 - LSTM Neural Networks

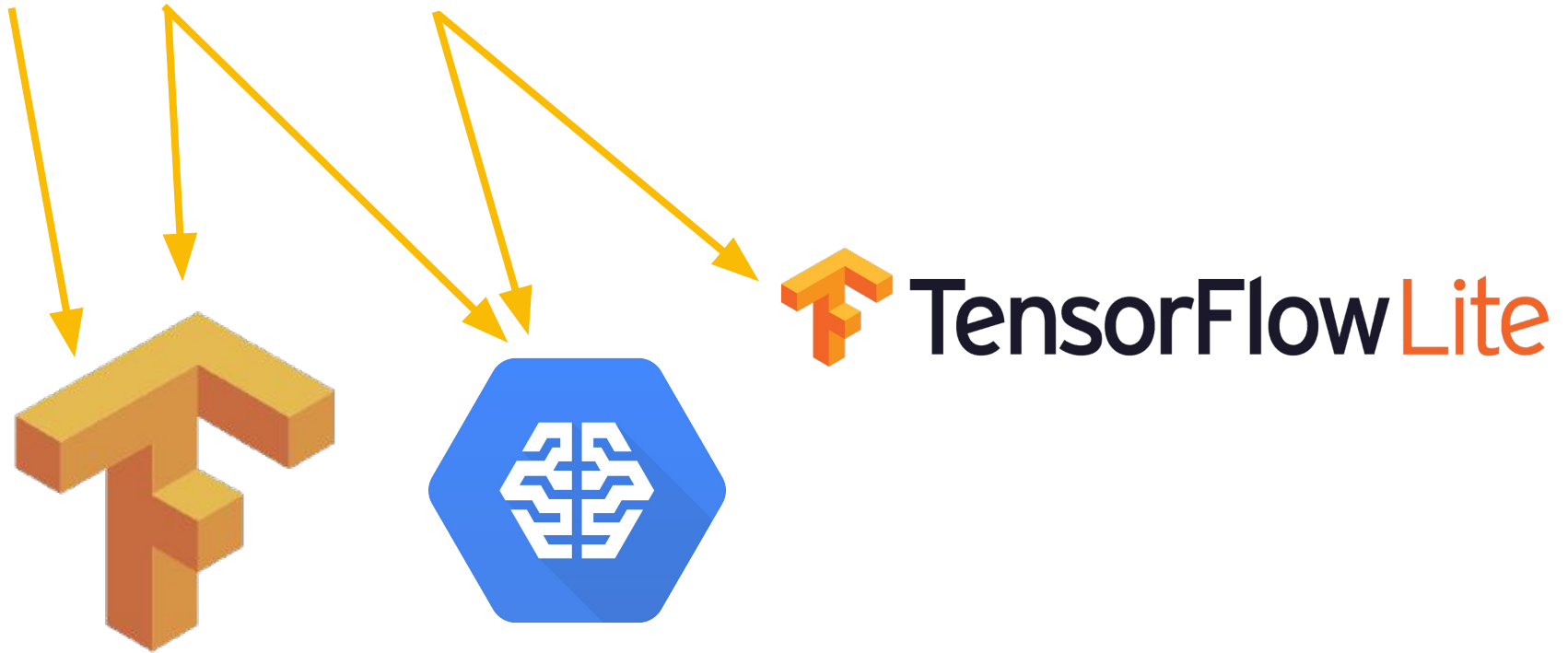




Cloud Machine Learning Engine

Train your models,
run predictions,
directly in the cloud

Build, train and serve your own models



Cloud Machine Learning Engine

Train models and **run predictions**
for your TensorFlow models
in the **cloud**, as a **fully managed service**,
on CPUs, GPUs or **TPUs**



```
gcloud ml jobs submit training job22 --package-path=trainer  
--module-name=trainer.task2 --staging-bucket=gs://ml-demo/jobs  
--config=config.yaml -- --train_dir=gs://ml-demo/jobs/train22
```

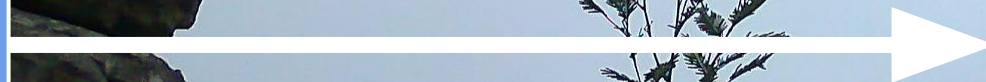

Bridging the Machine Learning gap



Use your own data to customize pre-trained models



API easy to call for a developer



Data scientists with ML background



AutoML [ALPHA]

Fine-tune pre-trained models
with your own datasets



Summary

Guillaume Laforge

Developer Advocate
Google Cloud

@glaforge

Try them all in your browser!

Video – cloud.google.com/video-intelligence

Vision – cloud.google.com/vision

Speech – cloud.google.com/speech

Natural Language – cloud.google.com/natural-language

Dialogflow – dialogflow.com

Translation – cloud.google.com/translation

AutoML – cloud.google.com/automl

TensorFlow – tensorflow.org

ML Engine – cloud.google.com/ml-engine



Machine Learning Crash Course

<https://developers.google.com/machine-learning/crash-course/>



The landing page features a background image of a whiteboard with handwritten mathematical notes and diagrams. The whiteboard includes a scatter plot, the equation $X_3 = X_1 X_2$, a linear regression equation $y = w_0 + w_1 X_1 + w_2 X_2 + w_3$, a definition of a Neural Network, the Sigmoid Activation Function $F(x) = \frac{1}{1 + e^{-x}}$, and the Rectified Linear Unit Activation Function $F(x) = \max(0, x)$. A diagram of a neural network with layers labeled 'Input', 'Hidden Layer 1', 'Hidden Layer 2', and 'Output' is also visible.

Machine Learning Crash Course

with TensorFlow APIs

Google's fast-paced, practical introduction to machine learning

[START CRASH COURSE](#)

[VIEW PREREQUISITES](#)

How Google does Machine Learning

<https://www.coursera.org/learn/google-machine-learning>



How Google does Machine Learning

Overview

Preview

FAQs

Creators

Pricing

Ratings and Reviews

Enroll

Starts Mar 12

[Apply for Financial Aid](#)

About this course: What is machine learning, and what kinds of problems can it solve? Google thinks about machine learning slightly differently -- of being about logic, rather than just data. We talk about such a framing is useful when thinking about building a pipeline of machine learning models. Then, we discuss the five phases of converting a candidate use case to be driven by machine learning, and conclude why it is important the phases not be skipped. We end with a recognition of the biases that machine learning can amplify and how to recognize this.

[^ Show less](#)

Who is this class for: This course is primarily for Data Engineers and programmers interested in learning how to apply machine learning in practice and, more generally, anyone interested in learning how to build and operationalize TensorFlow models.

Created by: Google Cloud

Google Cloud



Level

Intermediate



Commitment

1 week of study, 8-10 hours/week

Thanks
for your
attention

