# Superpower User Manual



Marc Cohen
mco@google.com
@marcacohen
mcohen.io





# Machine Learning Background





The brown quick fox jumps over the lazy dog.



### Syntax <u>rules</u> for adjective order in English:

- Quantity or number
- Quality or opinion
- Size
- Age
- Shape
- Colour
- Proper adjective (often nationality, other place of origin, or material)
- Purpose or qualifier

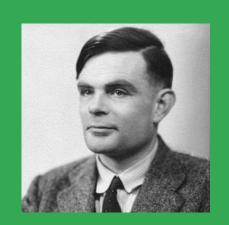


# Machine learning is learning from examples and experience.



"Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?"

- Alan Turing, 1950



## Cats vs. Dogs

Chris Bishop: Even the most

2009: sophisticated computers can't tell a dog

from a cat

2012: How Many Computers to Identify a Cat? 16,000

2015: Microsoft, Google Beat Humans at Image Recognition

2016: Machine Learning Algorithms Outperform Inexperienced Radiologists

Just this month...

\_\_\_

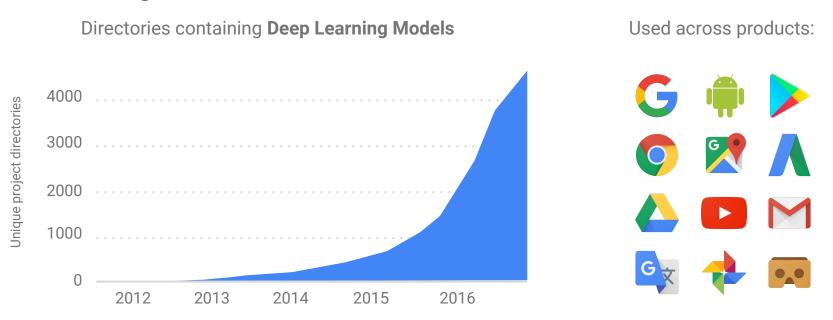
# Google AI can spot advanced breast cancer more effectively than humans

It's up to 99 percent accurate in the right conditions.

## Revolution

- Moore's Law + GPUs
- Cloud computing
- More/richer data
- Github + Open Source
- New/refined techniques
- "Perfect Storm"

# Rapidly accelerating use of deep learning at Google

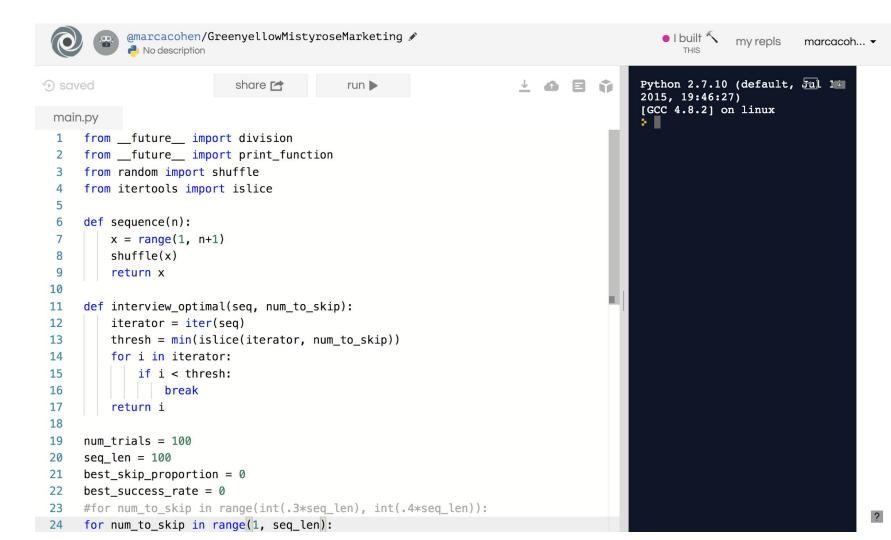




# My First Superpower

```
AN INTRODUCTION TO COMPUTER PROGRAMMING
FLIP-1
                                                                                            IN BASIC LANGUAGE
2 Fak Y=1 Ta 10
                                                                                          second edition
5 LET C=0
10 FOR K=1 TØ 50
                                                                                                JAMES S. COAN
20 LET F=INT(2*RND(1))
30 IF F=1 THEN 60
40 PRINT "T";
50 Ga To 100
56 KEM C COUNTS THE NUMBER OF HEADS
60 LET C=C+1
70 PRINT "H";
100 NEXT X
110 PRINT
120 PRINT "HEADS "; C; "OUT OF 50 FLIPS"
125 NEXT Y
130 END
RUN
FLIP-1
TIPRITARRIER FOR THITTITIA THANKS TITTINIA A THAN TIRE TO
HEADS - 21 OUT OF 50 FLIPS
HTTH THE CTRTTATHATTT HAR TTTTTTTHAHATTRTAHATAHATBELTAHA
HEADS 26 BUT OF 50 FLIPS
THEFT HETT TO THE TOTAL TO THE T
HEADS 17 BUT OF SO FLIPS
THEFT FOR THE THEFT THE THEFT THE TENTH THE THEFT TO THE THAT IT.
HEADS 21 OUT OF 50 FLIPS
TIMETITITE THE LIETHER HITCHE THE TERRITITE THE SECTION TO
HEADS 24 JUT OF 50 FLIPS
HINTHHHANNITATITITATINHHANNITATITATITATITATITATIA
HEADS 26 BUT OF 50 FLIPS
HTTTTTHTTTTHKTTHTHKKKTHTHKKTHTHKKTATTHTHKHTATTHT
HEADS 22 OUT OF 50 FLIPS
ТАТНИННАЯ ТАТИВАЯТТТИНТИТИЧНИЯ ТИНИТЯЯ ТАНИНИЯ ТТТИТЯ
HEADS 34 JUT OF 50 FLIPS
HEADS 24 BUT OF 50 FLIPS
TINETH THE THEFT IS IN A PROPERTY OF THEFT IS A PROPERTY OF THE THEFT IS THE THEFT 
HEADS 26 OUT OF 50 FLIPS
                                                                                                       HAYDEN
```

MANE



### The real world...

Kubernetes Engine Tutorials



### Deploying a containerized web application

This tutorial shows you how to package a web application in a Docker container image, and run that container image on a Kubernetes Engine cluster as a load-balanced set of replicas that can scale to the needs of your users.

#### Objectives

To package and deploy your application on Kubernetes Engine, you must:

- 1. Package your app into a Docker image
- 2. Run the container locally on your machine (optional)
- 3. Upload the image to a registry
- 4. Create a container cluster
- 5. Deploy your app to the cluster
- 6. Expose your app to the Internet
- 7. Scale up your deployment
- 8. Deploy a new version of your app

# Why is it so COMPLICATED?

- 1. Hardware provisioning
- 2. Software installation
- 3. OS upgrades
- 4. Security patches
- 5. System & network admin
- 6. Scaling up & down
- 7. Paying for stuff you don't use
- 8. Dealing with failures
- 9. Managing clusters
- 10. Packages and containers
- 11. CAP Theorem
- 12. Load Balancing
- 13. Service Mesh
- 14. Data Sharding
- 15. Backup and Recovery
- 16. Authentication & Authorization
- 17. Hybrid Operation
- 18. Multiple servers, disks, network

19.



#### **Developer Cheat Sheet**

v2018.9.14

Created by the Google Developer Relations Team

Maintained at https://medium.com/@gregsramblings

✓ @gregsramblings

#### Sites and Blogs

Google Cloud Home Page https://cloud.google.com Google Cloud Blog https://cloud.google.com/blog GCP Medium Publication https://medium.com/google-cloud Apigee Blog https://apigee.com/about/blog Firebase Blog https://firebase.googleblog.com G Suite Developers Blog https://gsuite-developers.googleblog.com GCP .NET Home https://cloud.google.com/dotnet GCP Go Home https://cloud.google.com/go GCP Java Home https://cloud.google.com/java GCP Node Home https://cloud.google.com/node GCP PHP Home https://cloud.google.com/php GCP Python Home https://cloud.google.com/python GCP Ruby Home https://cloud.google.com/ruby Google Cloud Certifications https://cloud.google.com/certification Google Cloud System Status https://status.cloud.google.com Google Cloud Training https://cloud.google.com/training Google Codelabs https://codelabs.developers.google.com Google Developers Blog https://developers.googleblog.com Google Maps Platform Blog https://mapsplatform.googleblog.com Google Open Source Blog https://opensource.googleblog.com Google Security Blog https://security.googleblog.com Kaggle Home Page https://www.kaggle.com Kubernetes Blog https://kubernetes.io/blog Regions and Network Map https://cloud.google.com/about/locations

#### **Podcasts**

Firebase

**GCP Podcast** https://gcppodcast.com **Kubernetes Podcast** https://kubernetespodcast.com

#### YouTube Channels

Google Cloud Platform https://www.youtube.com/googlecloudplatform Google Developers https://www.youtube.com/GoogleDevelopers G Suite https://www.youtube.com/GSuite Apigee https://www.youtube.com/apigee/ https://www.voutube.com/firebase

#### Compute Products

Compute Engine Virtual Machines, Disks, Network App Engine Managed App Platform Kubernetes Engine Managed Kubernetes/Containers Shielded VMs Hardened VMs **Cloud Functions** Event-driven serverless functions

#### **Identity and Security Products**

Access Transparency Cloud Data Loss Prevention API Classify, Redact Sensitive Data Cloud IAM Cloud Identity Cloud Identity-Aware Proxy Cloud Key Management Service Cloud Resource Manager Cloud Security Scanner Security Key Enforcement

Audit Cloud Provider Access Resource Access Control Manage Users, Devices & Apps Identity-based App Signin Hosted Key Management Service Cloud Project Metadata Management App Engine Security Scanner Two-step Key Verification Two-factor Authentication (2FA) Device

#### **Management Tools Products**

Titan Security Key

Cloud APIs **APIs for Cloud Services Cloud Billing** Billing and Cost Management Tools Cloud Billing AP Programmatically Manage GCP Billing Cloud Console Web-based Management Console Cloud Deployment Manager Templated Infrastructure Deployment Cloud Mobile App iOS/Android GCP Manager App Cloud Shell Browser-based Terminal/CLI Stackdriver Debugger Live Production Debugging Stackdriver Error Reporting App Error Reporting Stackdriver Logging Centralized Logging Stackdriver Monitoring Infrastructure and Application Monitoring Stackdriver Profiler CPU and heap profiling Stackdriver Transparent SLIs Monitor GCP Services Stackdriver Trace App Performance Insights

#### **Developer Tools**

CLI for GCP Cloud SDK Cloud Build Continuous integration/delivery platform **Cloud Source Repositories** Hosted Private Git Repos Cloud Tools for IntelliJ IntelliJ GCP Tools Cloud Tools for PowerShell PowerShell GCP Tools Visual Studio GCP Tools Cloud Tools for Visual Studio Cloud Tools for Eclipse **Eclipse GCP Tools** Container Registry Private Container Registry/Storage Gradle App Engine Plugin Gradle App Engine Plugin Mayen App Engine Plugin Maven App Engine Plugin

#### Migration (to GCP)

Google Transfer Appliance Dentable Data Transport Boy

#### Al and Machine Learning

Cloud AutoML Natural Language 
Custom text classification models Cloud AutoML Translate Custom domain-specific translation Cloud AutoML Vision Custom image classification models Cloud Job Discovery Job Search with ML Cloud Machine Learning Engine Managed Platform for ML Text Parsing and Analysis Cloud Natural Language Cloud Speech-To-Text Convert Audio to Text Cloud Text-To-Speech Convert Text to Audio Cloud TPU Specialized Hardware for ML Cloud Translation API Language Detection and Translation Cloud Video Intelligence Scene-level Video Annotation Cloud Vision API Image Recognition and Classification **Dialogflow Enterprise Edition** Create Conversational Interfaces

#### **Data and Analytics Products**

Cloud Composer Managed Workflow Orchestration Service Cloud Dataflow Stream/batch data processing Cloud Datalab Managed Jupyter Notebook **Cloud Dataprep** Visual data wrangling Cloud Dataproc Managed Spark and Hadoop Cloud Pub/Sub Global Real-time Messaging Google BigQuery Data Warehouse/Analytics Google Data Studio Collaborative Data Exploration/Dashboarding Google Genomics Managed Genomics Platform

#### **Databases Products**

Cloud Bigtable Petabyte-scale, low-latency nonrelational Cloud Datastore Horizontally Scalable Document DB Cloud Firestore Strongly-consistent Serverless Document DB Cloud Memorystore Managed Redis Cloud Spanner Horizontally Scalable Relational DB Cloud SQL Managed MySQL and PostgreSQL

#### Storage Products

Cloud Storage Object Storage and Serving Nearline Archival Occasional Access Storage Coldline Archival Rare Access Storage Persistent Disk VM-attached Disks Cloud Filestore Managed NFS Server

#### **Networking Products**

Cloud DNC

Carrier Peering Peer with a carrier Direct Peering Peer with GCP **Dedicated Interconnect** Dedicated private network connection Partner Interconnect Connect on-premises network to VPC Cloud Armor **DDoS Protection** Cloud CDN Content Delivery Network

Drogrammable DNC Carring

#### Google Maps Platform

**Directions API Get Directions Between Locations** Distance Matrix API Calculate Travel Times Geocoding API Convert Address to/from Coordinates **Derive Location Without GPS** Geolocation API Maps Embed API Web Embedded Maps Dynamic Web Maps Maps JavaScript API Maps SDK for Android Maps SDK for Android Maps SDK for iOS Maps SDK for iOS Maps Static API Web Static Maps Maps Unity SDK Unity SDK for Games Maps URLs **URL Scheme for Maps** Metadata About Places (REST) Places API Places Library, Maps JavaScript API Metadata About Places (JavaScript) Places SDK for Android Places SDK for Android

Places SDK for iOS Places SDK for iOS Roads API Metadata About Roads Street View API Street View API

Time Zone API Convert Coordinates to Timezone

#### **G Suite Platform**

App Maker Assistive App Building **Apps Script** Extend and Automate Everything Editor Add-ons Extend Docs, Sheets, Slides Gmail Add-ons Contextual Apps in Gmail Conversational Bots in Chat **Hangouts Chat Bots** Calendar API Create and Manage Calendars Classroom API Provision and Manage Classrooms Drive API Read and Write Files Gmail API Enhance Gmail and Inbox Sheets API Read and Write Spreadsheets Slides API Create and Edit Presentations **Drive Picker Drive File Selection Widget** Admin SDK Manage G Suite Settings Email Markup Interactive Email using Schema.org G Suite Marketplace Storefront for Integrated Applications

#### Other G Suite APIs/SDKs Mobile Products (Firebase)

Cloud Firestore Document Store and Sync Cloud Functions for Firebase **Event-driven Serverless Applications** Cloud Storage for Firebase Object Storage and Serving Crashlytics Crash Reporting and Analytics Firebase A/B Testing Create A/B Test Experiments Firebase App Indexing App / Google Search Integration Firebase Authentication Drop-in Authentication Send Device Notifications Firebase Cloud Messaging Firebase Dynamic Links Link to App Content Eirobaca Hasting Web Hosting with CDN/SSI

Contacts, Google+, Tasks, Vault...

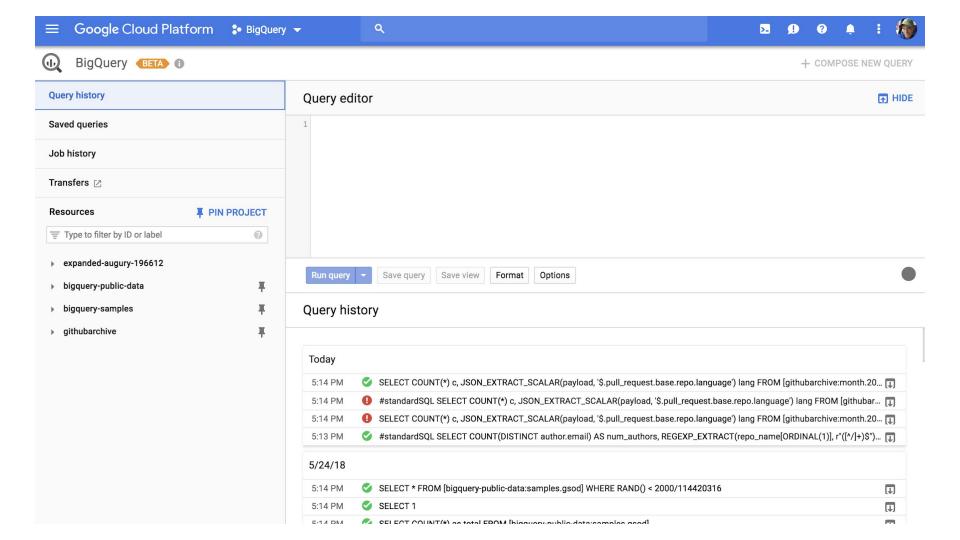
# What I want:

- Services, not servers
- Unlimited automatic scaling
- Pay only for what I use
- Endlessly extensible
- Relegate infrastructure

# In other words...

- 1. I don't want to think about stuff I don't care about, and...
- 2. I want to specify the "what", not the "how".

# Superpowered Services







# The **HTTP Archive** Tracks How the **Web is Built**.

We periodically crawl the top sites on the web and record detailed information about fetched resources, used web platform APIs and features, and execution traces of each page. We then crunch and analyze this data to identify trends — learn more about our methodology.

**View Reports** 

#### **HTTPS** Requests



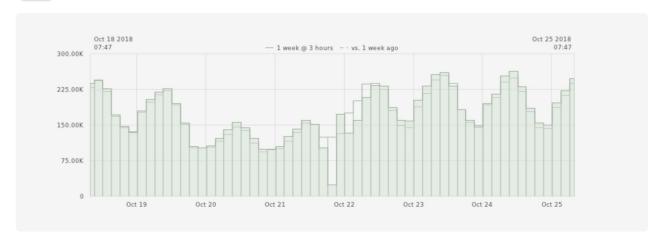
#### Timeseries of HTTPS Requests





Show Table





Open-source developers all over the world are working on millions of projects: writing code & documentation, fixing & submitting bugs, and so forth. GH Archive is a project to **record** the public GitHub timeline, **archive it**, and **make it easily accessible** for further analysis.

GitHub provides 20+ event types, which range from new commits and fork events, to opening new tickets, commenting, and adding members to a project. These events are aggregated into hourly archives, which you can access with any HTTP client:

Query	Command
Activity for 1/1/2015 @ 3PM UTC	wget http://data.gharchive.org/2015-01-01-15.json.gz
Activity for 1/1/2015	wget http://data.gharchive.org/2015-01-01-{023}.json.gz
Activity for all of January 2015	want http://data.aharchive.org/2015_01_{01} 31}_{01} 31}_{01} 32} ison oz













#### Machine Learning as an API



Cloud Vision API



Cloud Speech API



Cloud Natural Language API



Cloud Translation API



Cloud Video Intelligence API

### Use your own data to train models



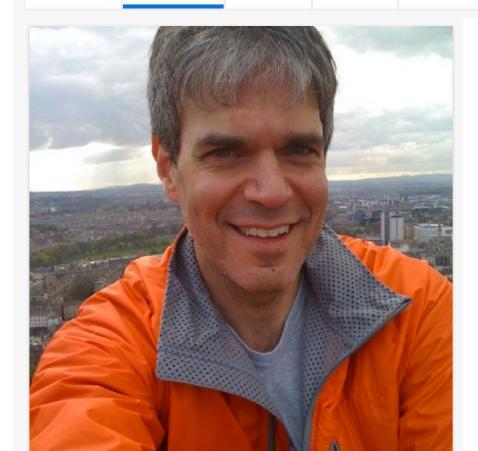
TensorFlow



Cloud Machine Learning Engine



Faces Labels Web Text Document Properties Safe Search JSON



Why Google Products Solutions Pricing Security

#### Al & Machine Learning Products

Contact sales

Cloud AutoML Product Overview AutoML Documentation

Audit Logging

AutoML Vision

Cloud AutoML Products AutoML Natural Language AutoML Translation

AutoML

#### Cloud AutoML



Cloud AutoML Products

Cloud AutoML makes the power of machine learning available to you even if you have limited knowledge of machine learning. You can use AutoML to build on Google's machine learning capabilities to create your own custom machine learning models that are tailored to your business needs, and then integrate those models into your applications and web sites.

You can use the following AutoML products to create custom machine learning models:

#### Cloud AutoML Products

#### AutoML Natural Language

AutoML Natural Language enables you to train your own, custom machine learning models to classify documents according to labels that vou define.

#### **AutoML Translation**

AutoML Translation enables you to create your own, custom translation models so that translation queries return results specific to your domain

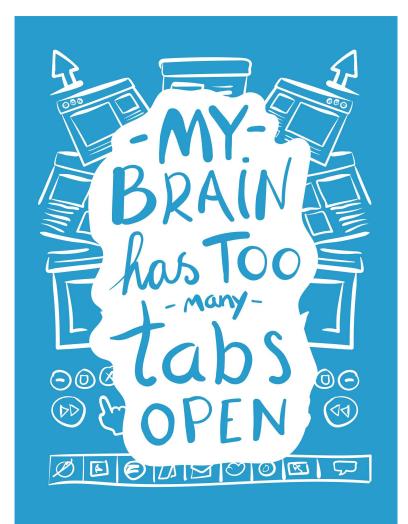
#### AutoML Vision

AutoML Vision enables you to train your own, custom machine learning models to classify your images according to labels that you define.

## The Meiko Detector



# Superpowered App



# curate<sup>2</sup>

### /kjʊ(ə) reit/ •

#### verb

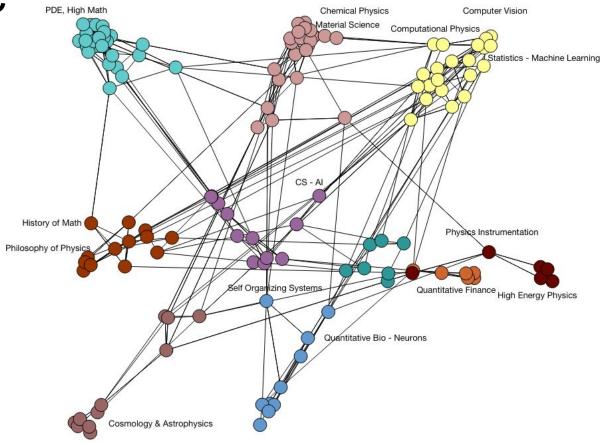
verb: **curate**; 3rd person present: **curates**; past tense: **curated**; past participle: **curated**; gerund or present participle: **curating** 

select, organize, and look after the items in (a collection or exhibition). "both exhibitions are curated by the Centre's director"

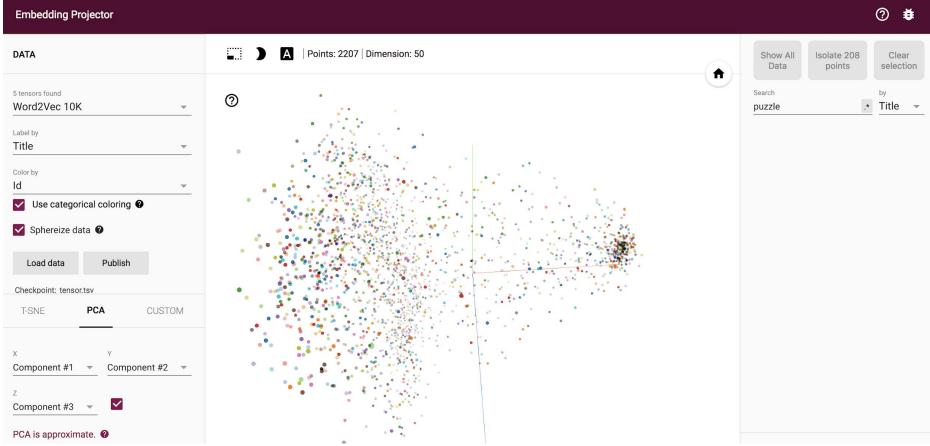
- select the performers or performances that will feature in (an arts event or programme).
   "in past years the festival has been curated by the likes of David Bowie"
- select, organize, and present (online content, merchandise, information, etc.), typically using professional or expert knowledge.

"people not only want to connect when using a network but they also enjoy getting credit for sharing or curating information."

## doc2vec



```
train_corpus = list(read_corpus(train_file))
model = gensim.models.doc2vec.Doc2Vec(vector_size=200, min_count=2, epochs=500)
model.build_vocab(train_corpus)
model.train(train_corpus, total_examples=model.corpus_count, epochs=model.epochs)
model.save_word2vec_format(train_dir + '/vectors.w2v', doctag_vec=True, \
    word_vec=False, fvocab='vocab.txt', binary=False)
model.save(train_dir + '/marc.d2v')
```



### Some results...

how singer won the sewing machine war	
wikipedia beanie babies	
wikipedia monty hall problem	
the weird dangerous isolated life of the saturation diver	0.92
pets who helped solve their owners murders	0.54



# **Three Phases of Digital Curation**

### **Model Training**

- Gather corpus
- Cleanse data
- Train model
- Test model

## Triage

- Regular sampling
- Assess samples
- Save results
- Dashboard

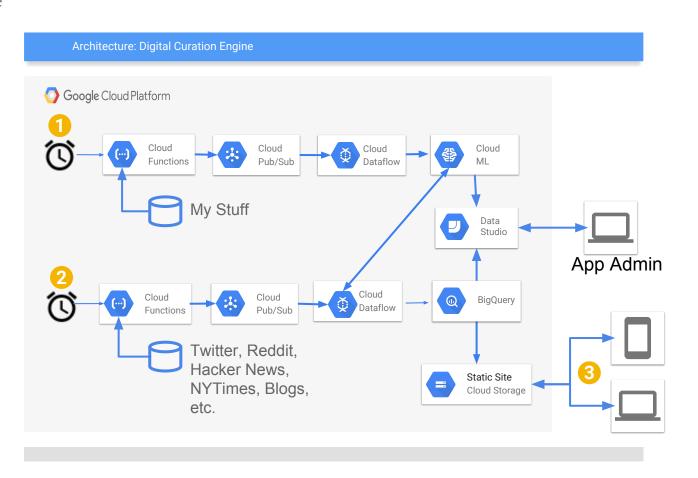
#### UX

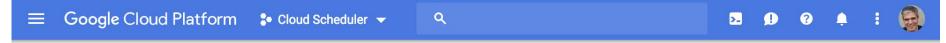
- Chronological list
- Search function
- History+analytics
- Feedback mechanism



#### cur8 - Digital Curation Engine

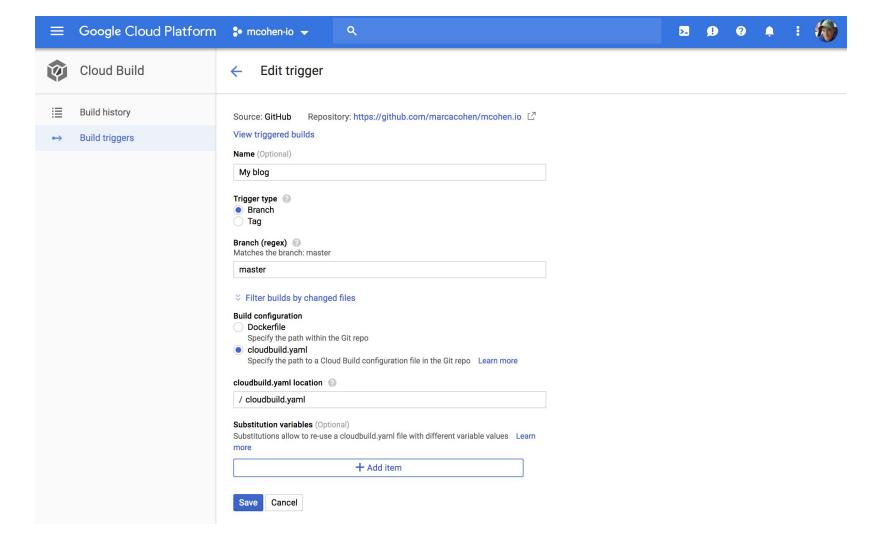
- **1** Model Training
- 2 Triage
- **3** User Experience













#### **Cloud Functions**



#### Create function

Q

https://us-central1-mcohen-io.cloudfunctions.net/function-1

#### Source code

Inline editor

Go 1.11 (Alpha)

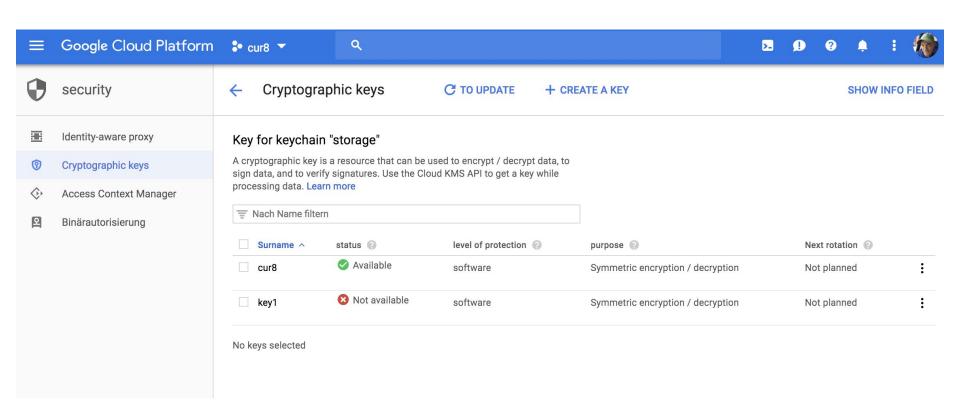
Node.js 6

Node.js 8 (Beta)

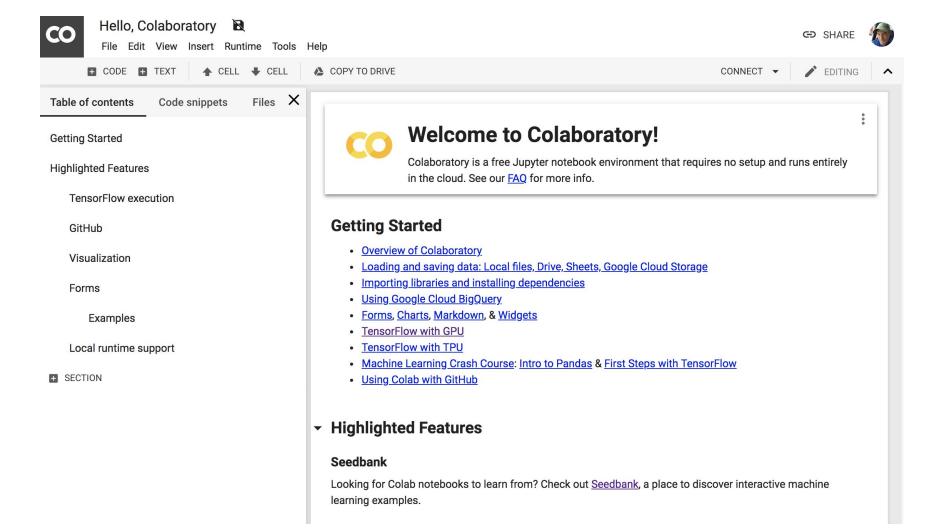
Python 3.7 (Beta)

#### main.py requirements.txt

```
1 def hello world(request):
        """Responds to any HTTP request.
             request (flask.Request): HTTP request object.
             The response text or any set of values that can be
             Response object using
 8
              make response <a href="mailto://flask.pocoo.org/docs/0.12/ar">make response <a href="mailto://flask.pocoo.org/docs/0.12/ar">http://flask.pocoo.org/docs/0.12/ar</a>
 9
10
        request_json = request.get_json()
        if request.args and 'message' in request.args:
11
             return request.args.get('message')
12
        elif request json and 'message' in request json:
13
14
             return request json['message']
15
        else:
16
             return f'Hello World!'
17
```



# Superpowered Learning



CODE TEXT

✓ CONNECTED ▼



♠ CELL ♣ CELL

COPY TO DRIVE

First, you'll need to enable TPUs for the notebook.

Navigate to Edit→Notebook Settings, and select TPU from the Hardware Accelerator drop-down (you can also access Notebook Settings via the command palette: cmd/ctrl-shift-P).

Next, we'll check that we can connect to the TPU.

```
[1] import os
   import pprint
   import tensorflow as tf

if 'COLAB_TPU_ADDR' not in os.environ:
        print('ERROR: Not connected to a TPU runtime; please see the first cell in this notebook for instructions!')
   else:
        tpu_address = 'grpc://' + os.environ['COLAB_TPU_ADDR']
        print ('TPU address is', tpu_address)

with tf.Session(tpu_address) as session:
        devices = session.list_devices()

print('TPU devices:')
        pprint.pprint(devices)
```

TPU address is grpc://10.4.220.34:8470

TPU devices:

[\_DeviceAttributes(/job:tpu\_worker/replica:0/task:0/device:CPU:0, CPU, -1, 12857156263210199625),

\_DeviceAttributes(/job:tpu\_worker/replica:0/task:0/device:XLA\_CPU:0, XLA\_CPU, 17179869184, 15227990594443770068),

\_DeviceAttributes(/job:tpu\_worker/replica:0/task:0/device:XLA\_GPU:0, XLA\_GPU, 17179869184, 6960509336819419478),

\_DeviceAttributes(/job:tpu\_worker/replica:0/task:0/device:TPU:0, TPU, 17179869184, 725889995631097262),

\_DeviceAttributes(/job:tpu\_worker/replica:0/task:0/device:TPU:1, TPU, 17179869184, 7406319308873999103),

Q

We call them "seeds". Each seed is a machine learning example you can start playing with. Explore, learn and grow them into whatever you like.

#### O No Setup required

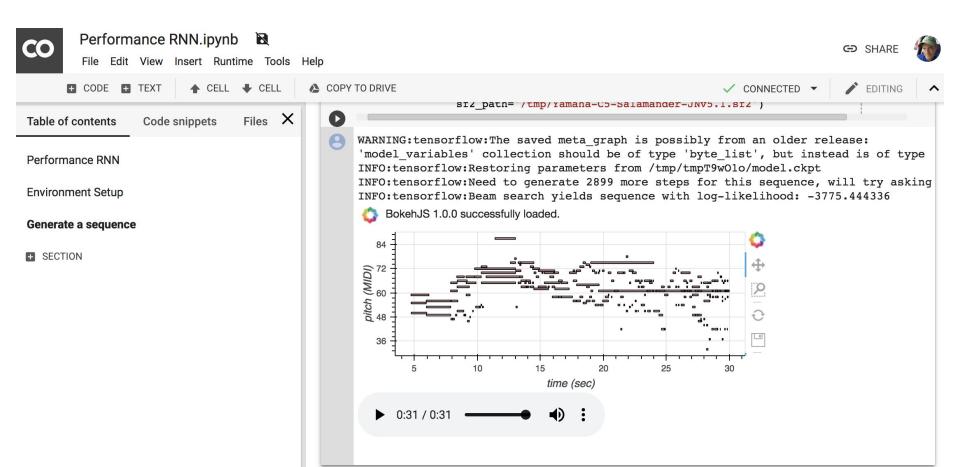
Run and experiment with machine learning code in your browser.

#### O Free GPU Backend

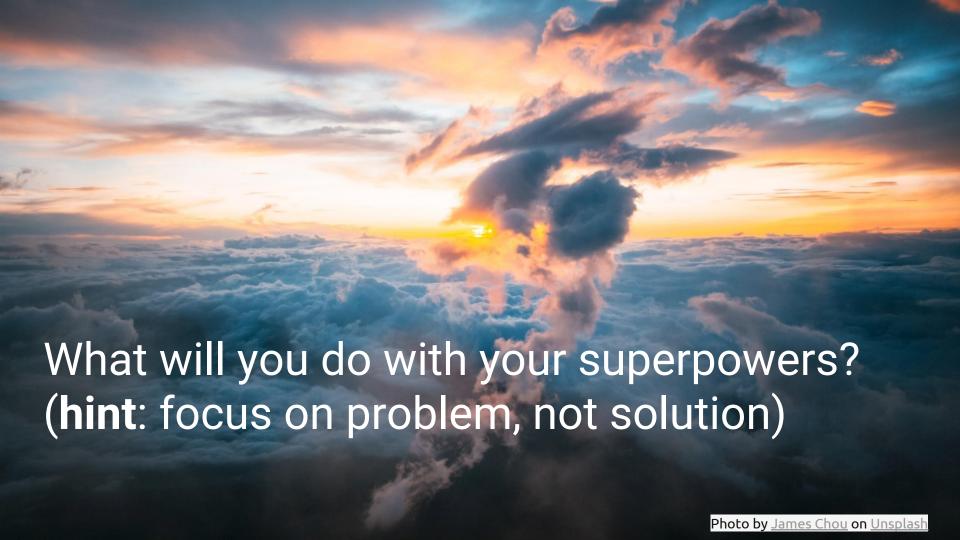
Run and train models, plot results real-time, share with others.

#### O Explore and Learn

Discover new algorithms, extend them to suit your ideas.







# Thanks!



Marc Cohen
mco@google.com
@marcacohen
mcohen.io

