Empowering Students and Researchers to Build What Is Next

Brian A Gibson
bagibson@google.com | linkedin.com/in/bagibson | twitter.com/bagibson

@bagibson

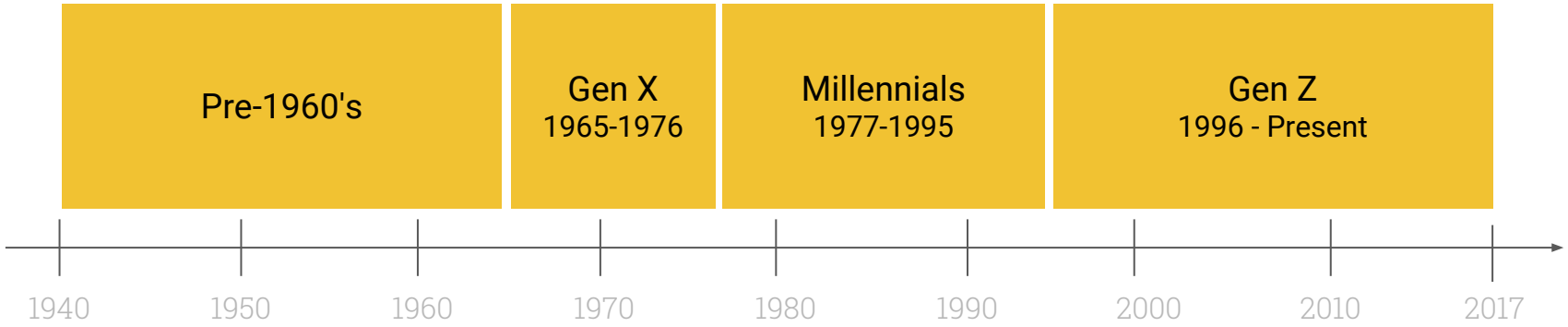
Photo © 2010 J. Ronald Lee.

@bagibson

1999

Generation Z is entering higher education

Are we ready to prepare them to build what is next?



@bagibson

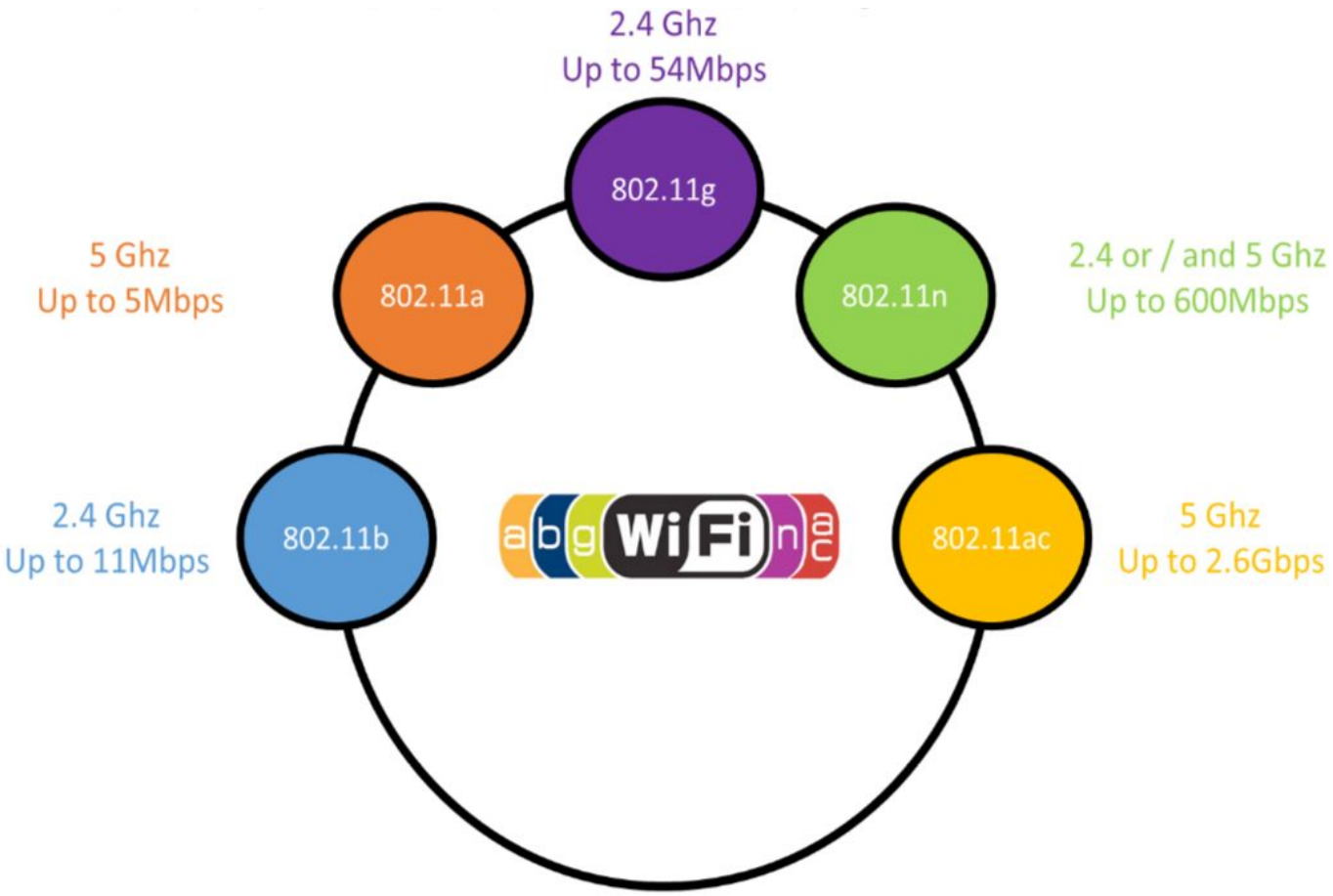
Nokia 6110

No colour
No internet
No wireless networking
No browser
No bluetooth
No apps....

....but it did have snake!

@bagibson

@bagibson


2.4 Ghz
Up to 54Mbps
802.11g
5 Ghz
Up to 5Mbps
802.11a
802.11n
2.4 or / and 5 Ghz
Up to 600Mbps
2.4 Ghz
Up to 11Mbps
802.11b
abg WiFi n ac
802.11ac
5 Ghz
Up to 2.6Gbps

@bagibson

AltaVista HOME - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: http://www.altavista.com/

AltaVista® Connections — The most powerful and useful guide to the Net

October 23, 1999 PDT

My AltaVista Shopping.com Zip2.com

Ask AltaVista® a question. Or enter a few words in any language

Help
Advanced Text Search

Search For: Web Pages Images Video Audio

Search

Search tip: use image search

Example: **When precisely will the new millennium begin?**

ALTAVISTA CHANNELS - My AltaVista - Finance - Travel - Shopping - Careers - Health - News - Entertainment

FREE INTERNET ACCESS - Download Now New - Support **USEFUL TOOLS** -Family Filter - Translation - Yellow Pages - People Finder - Maps - Usenet - Check Email

DIRECTORY

Automotive

Business & Finance

Computers & Internet

ALTAVISTA HIGHLIGHTS

POWER SEARCH

▸ BIG changes coming to AltaVista 10/25 -Info inside!

TRY THESE SEARCHES...

Search for **Halloween** in images

Document: Done

Search the web using Google

Google Search | I'm feeling lucky

More Google!

Copyright ©1999 Google Inc.

@bagibson

Generation Z only knows ubiquitous technology

2G

www.

amazon

WIKIPEDIA
The Free Encyclopedia

Google

Gmail
by Google

YouTube

4G

Instagram

WhatsApp

Pre-2000 | 2000 | 2002 | 2004 | 2006 | 2008 | 2010 | 2012 | 2014 | 2016

Generation Z

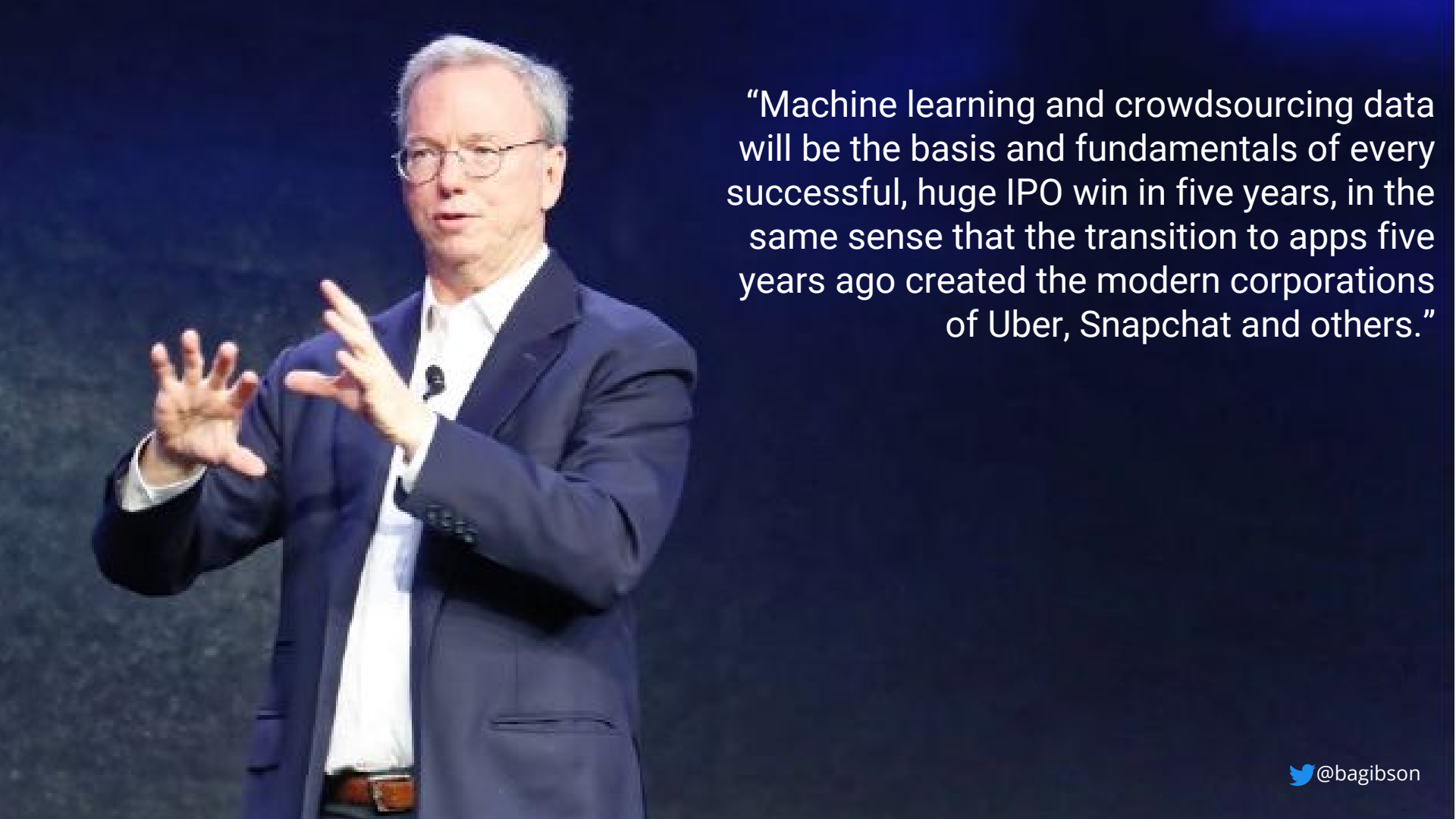
What kind of a world should university prepare them for?

@bagibson

“We are on the cusp of a 4th industrial revolution”

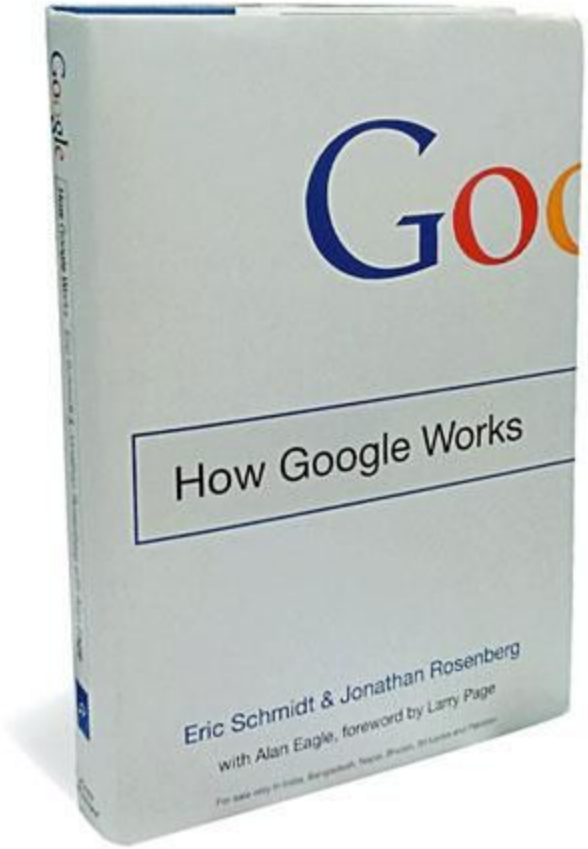
- World Economic Forum
The Future of Jobs

@bagibson


“Machine learning and crowdsourcing data will be the basis and fundamentals of every successful, huge IPO win in five years, in the same sense that the transition to apps five years ago created the modern corporations of Uber, Snapchat and others.”
@bagibson

Helping students build what’s next

@bagibson



"Three powerful technology trends have converged to fundamentally shift the playing field in most industries"

1. The internet has made information free, copious, ubiquitous. Almost everything is online.

2. Mobile devices & internet have made global reach and continuous connectivity widely available.

3. Cloud computing has put practically infinite compute power and storage at everyone's disposal, on an inexpensive, pay as you go model.

@bagibson

Hitting the limits early on

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page

Computer Science Department,
Stanford University, Stanford, CA 94305, USA
sergey@cs.stanford.edu and page@cs.stanford.edu

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/ To improve a search engine is a challenging task. Search engines index tens to hundreds of millions of pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and web proliferation, creating a web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale web search engine – the first such detailed public description we know of to date. Apart from the problems of scaling traditional search techniques to web-scale data, there are new technical challenges involved with using the additional information present in hypertext to produce better search results. This paper addresses this question of how to build a practical large-scale system which can exploit the additional information present in hypertext. Also we look at the problem of how there is effectively deal with uncontrolled hypertext collections where anyone can publish anything they want.

Keywords

World Wide Web, Search Engines, Information Retrieval, PageRank, Google.

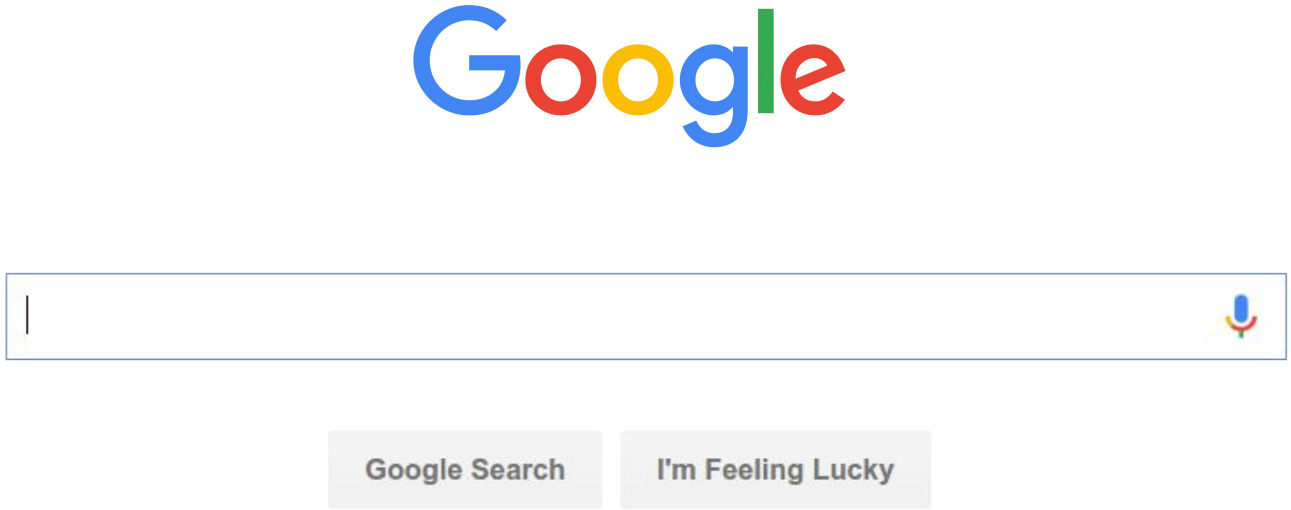
1. Introduction

(Note: There are two versions of this paper – a longer full version and a shorter printed version. The full version is available on the web and the conference CD-ROM.)
The web creates new challenges for information retrieval. The amount of information on the web is growing rapidly, as more many users get web servers and write HTML documents to describe their works. Web users are likely to surf the web using its link graph, often starting with high quality human maintained indices such as Yahoo! or with search engines. Human maintained lists cover popular topics effectively but are subjective, expensive to build and keep up to date, and cover relatively few of all topics. Automated search engines that rely on keyword matching usually return too many low quality matches. To make matters worse, some advertisers attempt to gain people's attention by taking measures meant to mislead automated search engines. We assume that the average web user will not want to wade through the problems of existing systems. It makes especially heavy use of the additional structure present in hypertext to provide much higher quality search results. We chose our system name, Google, because it is a common spelling of googol, or 10^{100} and fits well with our goal of building very large-scale search

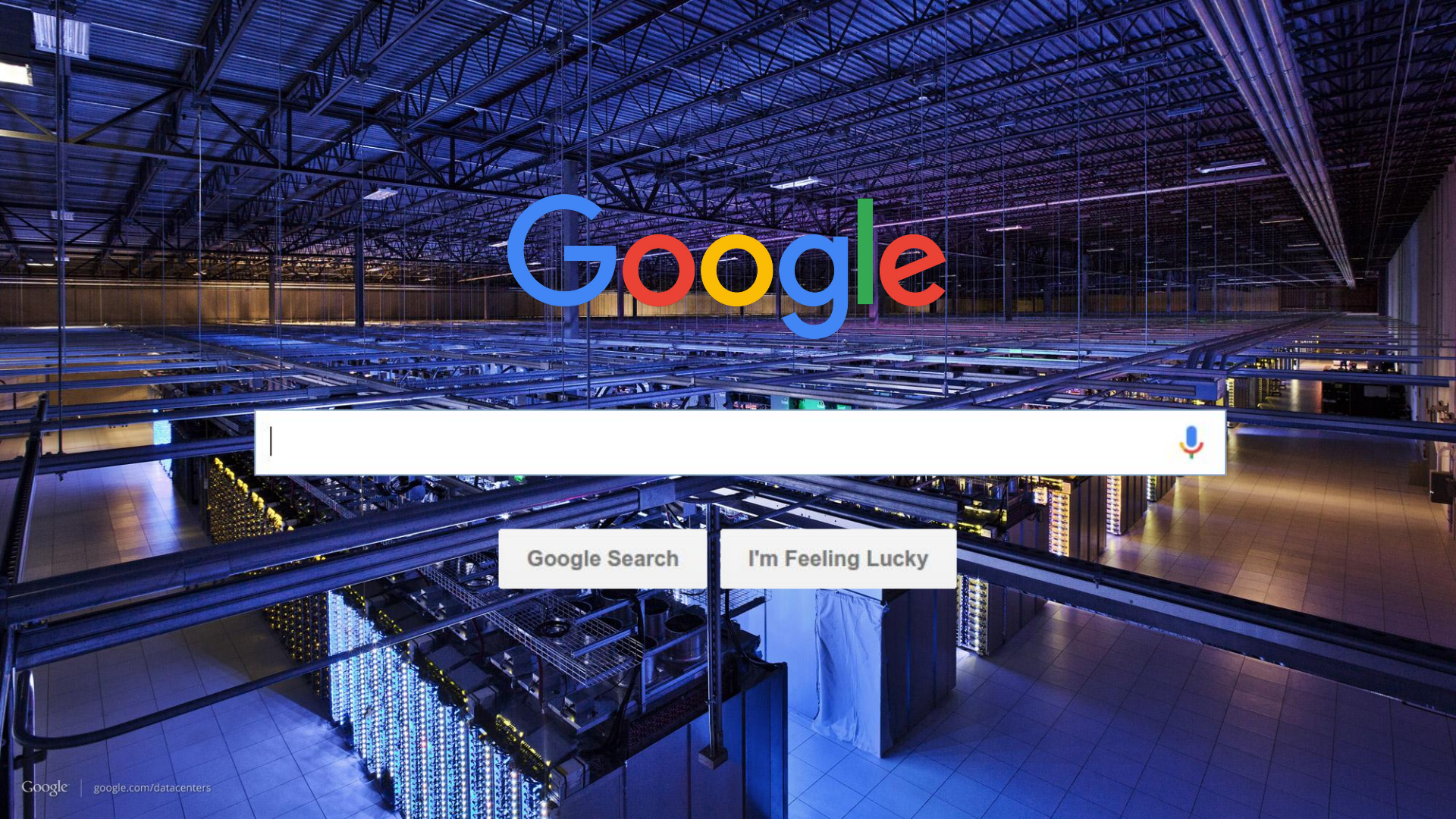
The Anatomy of a Large-Scale Hypertextual Web Search Engine

1996, Sergey Brin and Lawrence Page
Computer Science Department, Stanford University, Stanford, CA 94305

@bagibson



@bagibson


Google
Google Search
I'm Feeling Lucky
Google
google.com/datacenters

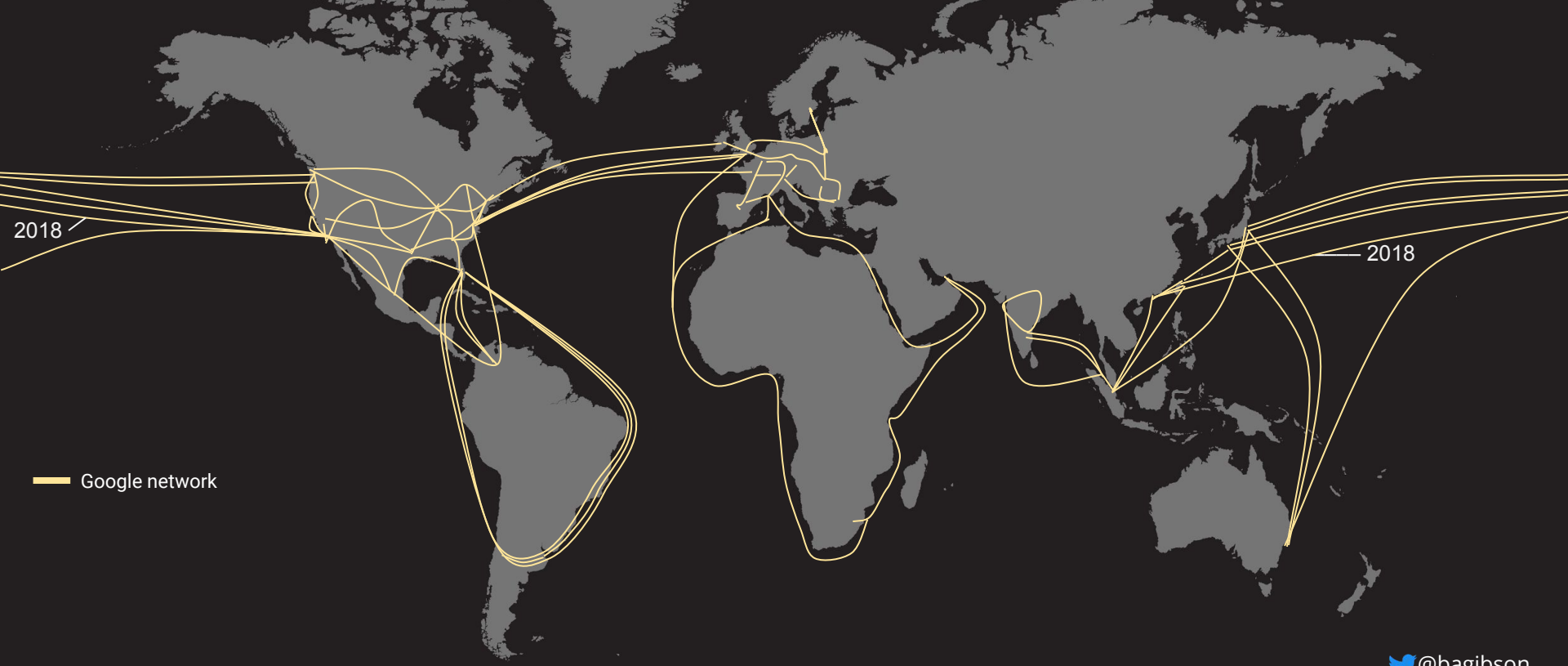
Google

7 Cloud products with 1 billion users

@bagibson

Google Cloud Platform Backbone

Google Cloud Platform customers enjoy the incredible performance of Google's global network, including hundreds of thousands of miles of fiber, and four owned undersea cables, 3x more than any other cloud.



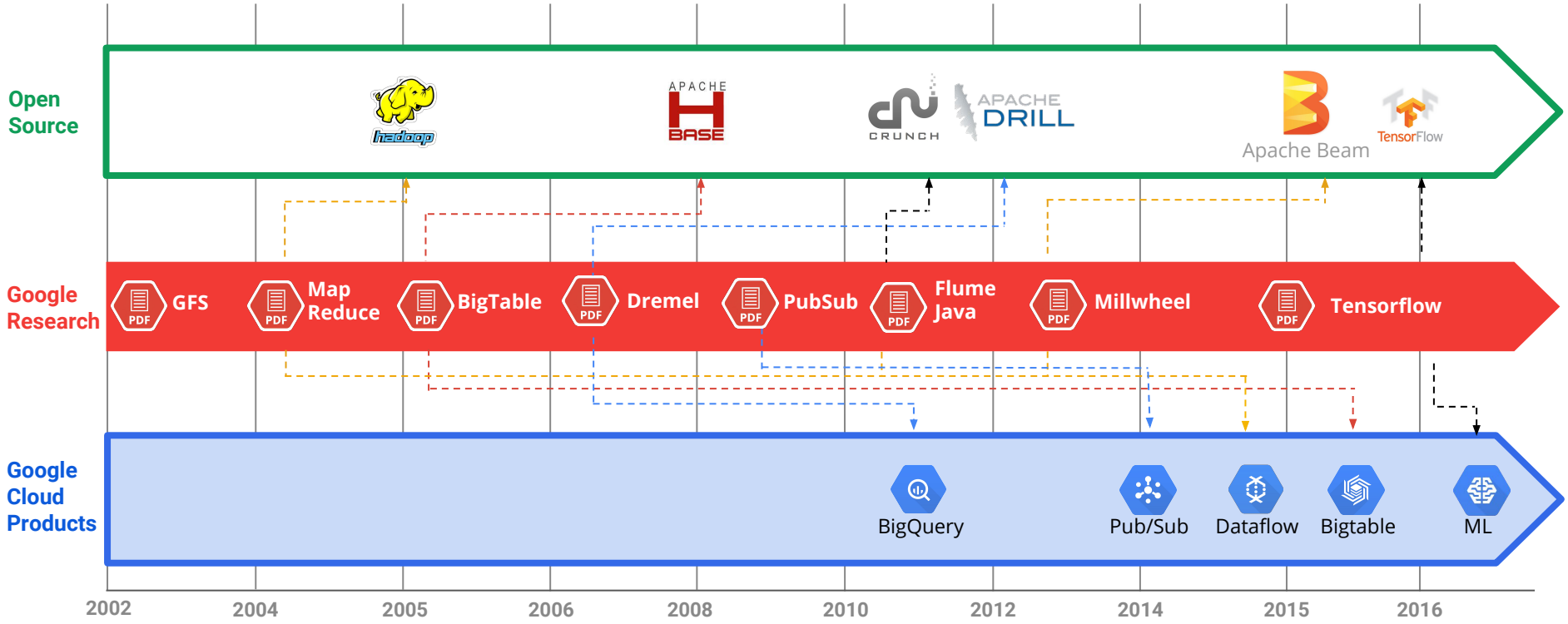
@bagibson

$29.4 Billion

3 Year Trailing CAPEX Investment

@bagibson

15+ years of solving Data Problems



@bagibson

More than 60 Google Cloud Platform services

Compute

- Compute Engine
- App Engine
- Container Engine
- Container Registry
- Cloud Functions

Storage and Databases

- Cloud Storage
- Cloud Bigtable
- Cloud Datastore
- Cloud SQL
- Cloud Spanner
- Persistent Disk

Networking

- Virtual Private Cloud
- Cloud Load Balancing
- Cloud CDN
- Cloud Router
- External IP
- Cloud Interconnect
- Cloud DNS
- Networking
- VPN

Big Data

- BigQuery
- Cloud Dataflow
- Cloud Dataproc
- Cloud Datalab
- Cloud Pub/Sub
- Genomics

Machine Learning

- Cloud Machine Learning
- Cloud Vision API
- Cloud Speech API
- Cloud Natural Language API
- Cloud Translation API
- Cloud Jobs API

Identity & Security

- Cloud IAM
- Cloud Resource Manager
- Cloud Security Scanner
- Key Management Service
- Firewall
- BeyondCorp
- Data Loss Prevention
- Identity-Aware Proxy
- Security Key Enforcement

@bagibson

More than 60 Google Cloud Platform services

Management Tools

- Stackdriver
- Monitoring
- Logging
- Error Reporting
- Trace
- Debugger
- Cloud Deployment Manager
- Cloud Endpoints
- Cloud Console
- Cloud Shell
- Cloud Mobile App
- Cloud Billing API
- Cloud APIs

Developer Tools

- Cloud SDK
- Cloud Deployment Manager
- Cloud Source Repositories
- Cloud Tools for Android Studio
- Cloud Tools for IntelliJ
- Cloud Tools for PowerShell
- Cloud Tools for Visual Studio
- Google Plug-in for Eclipse
- Cloud Test Lab

@bagibson

MIT Research

580,000 cores on preemptible VMs

"10^{17} hyperelliptic curves of genus 3 in an effort to find curves whose L-functions can be easily computed, and which have potentially interesting Sato-Tate distributions"

Products used: Google Compute Engine, Cloud Storage, DataStore

MIT Research

580,000 cores on preemptible VMs

That’s roughly 300 years of core-years of compute time for $20,000 □ not bad for a lazy Sunday!

Products used: Google Compute Engine, Cloud Storage, DataStore

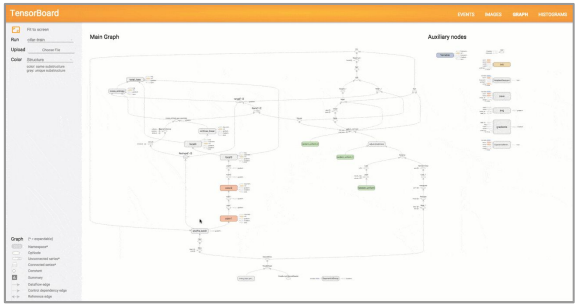
Stanford Genomics

“From our point of view, it’s almost infinite resources. A single user can boot up 5,000 machines. What you can do with Google Genomics — and you can’t do in-house — is run 1,000 genomes in parallel”

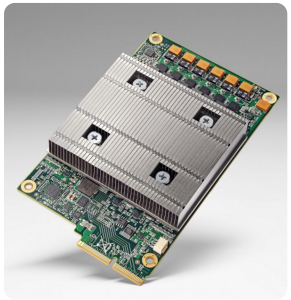
Somalee Datta, bioinformatics director of Stanford University’s Center of Genomics.

The Machine Learning Spectrum

TensorFlow



CloudML



Scale, No-ops Infrastructure

TPU - 7 years ahead of GPU in terms of price/performance

Google Trained Models

Cloud Translate

Cloud Vision

Cloud Natural Language (Beta)

Cloud Speech (Beta)

Use/extend OSS SDK — Build custom models — Use pre-built models

ML researcher — Data Scientist — App Developer

@bagibson

Democratising ML with Pre-Trained 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...

Built on Open Source

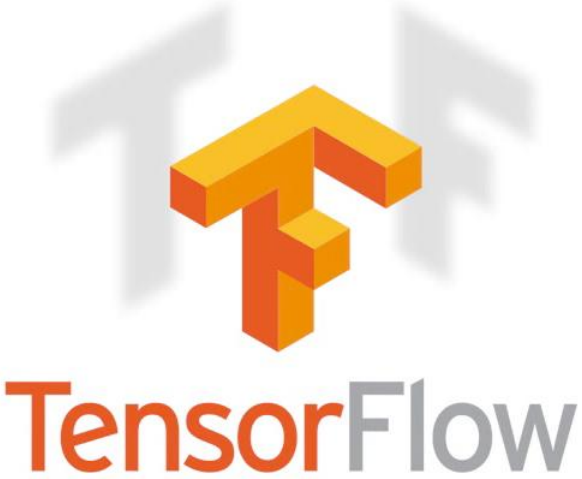
Created by Google Brain team

Most popular ML project on Github

- Over 480 contributors
- 10,000 commits in 12 months

Multiple deployment options:

- Mobile, Desktop, Server, Cloud
- CPU, GPU

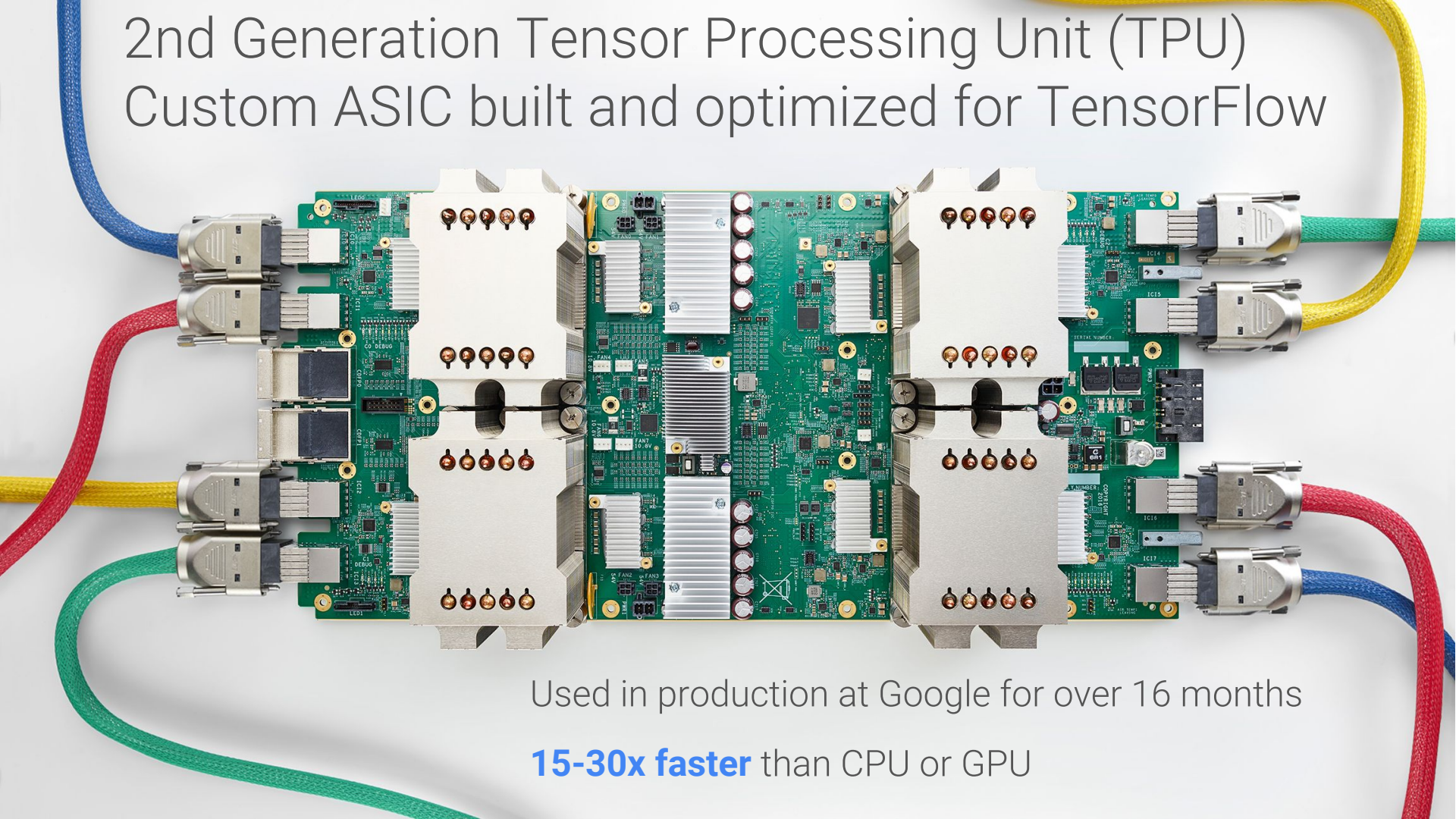


Train Your Own Models

Cloud Machine Learning Engine

Fully Managed ML Infrastructure

2nd Generation Tensor Processing Unit (TPU)
Custom ASIC built and optimized for TensorFlow

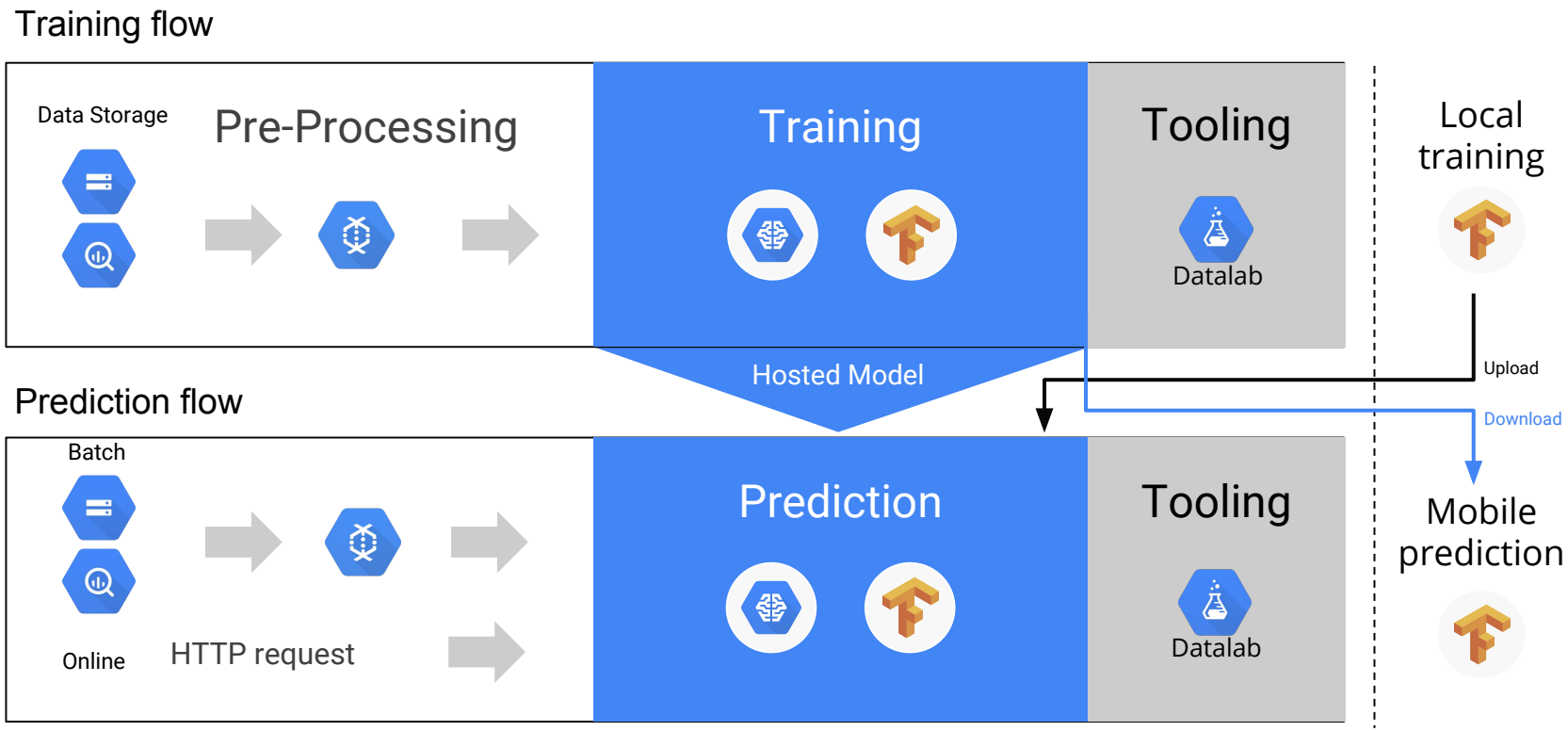


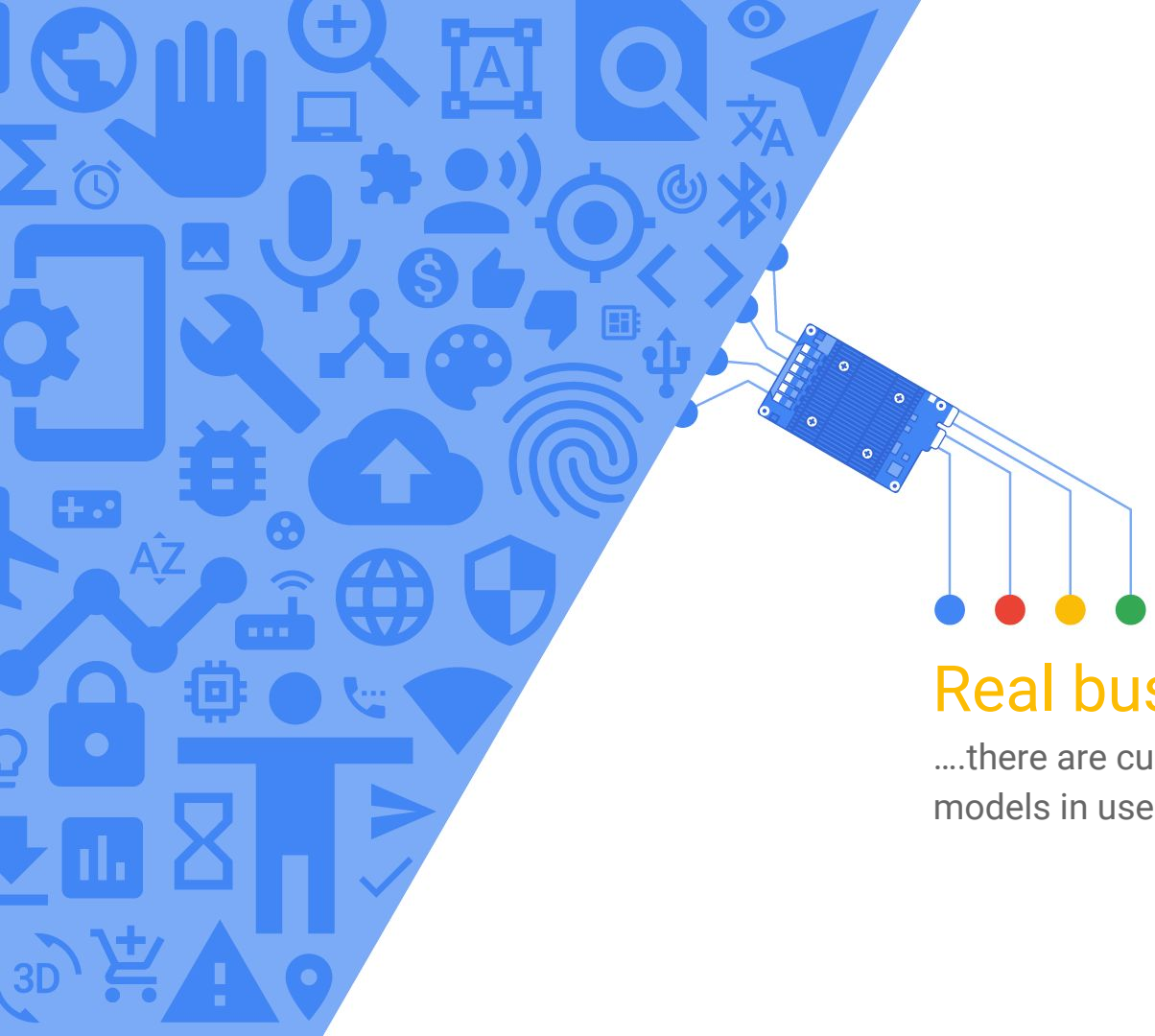
Used in production at Google for over 16 months

15-30x faster than CPU or GPU

A “TPU pod” built with 64 second-generation TPUs delivers up to 11.5 petaflops of machine learning acceleration.

Use your own data to train models





Real business impact...

….there are currently over 4,000 Machine Learning models in use in production in Google today.

Higher education, students & researchers are able to access the power of global computing giants like Google

Google Cloud