

NOTICE:
This is being recorded!

Product Data Design

Madrid, March 13. 2019

About me

Juan Turrión

Cloud Consultant at Google

Previously, **Technical Product and Portfolio manager** at Ericsson

What about you?

Quick show of hands

About this Session

Approach of this Session

Stuff can seem boring

But it's pretty hard to achieve and it sits at the core of every company

Try to do it practical

A data scientist is too far from a UX specialist

Make end-to-end products rather than design UX

Targeted Positions

Product Manager

Our Product Management team partners with Engineers and Designers to make a positive impact on our users' lives and businesses. We care deeply about developing products by identifying potential opportunities, conducting market research, generating product requirements, determining specifications and roadmap plans for feature introduction. We believe in engaging with the community, regularly hosting free events with top technical speakers, and actively contributing to open-source software (check out Britecharts as an example!). Our technology spans across web, mobile, API, big data, machine learning, search, physical point of sale, and scanning systems.

Product Owner

The **Product Owner** is the sole person responsible for managing the **Product Backlog**. **Product Backlog** management includes: **Product Backlog**. Clearly expressing **Product Backlog** items. Ordering the items in the **Product Backlog** to best achieve goals and missions.

Some Context

How the Market is doing

Companies get disrupted all the time

Traditional companies getting disrupted by newcomers

Globalisation. Pressure on margins.

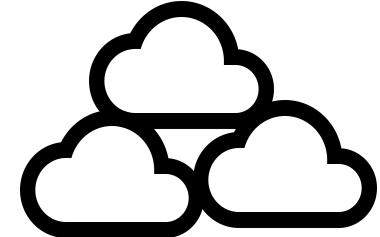
Big tech companies (FAANG and Microsoft) doing well and doing acquisitions

VC looking for the next Unicorn, More and more unicorns (global phenomenon)

Innovate or die

Cloud Economics

Global delivery | Cutting edge capabilities | *Infinite* capacity



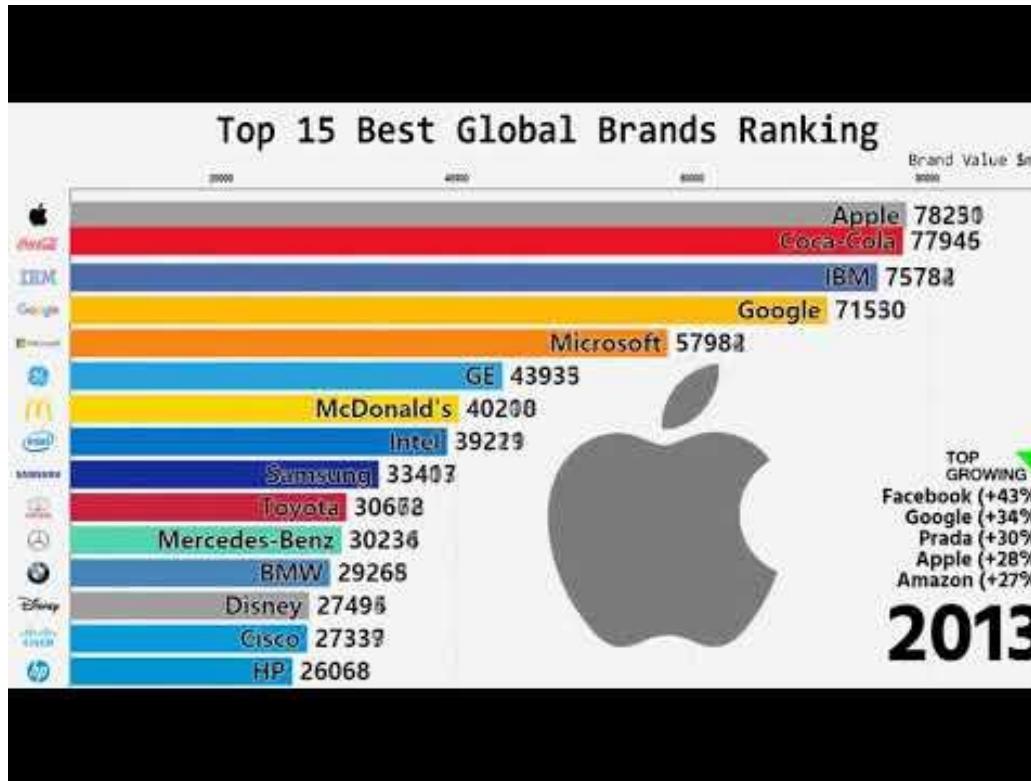
Pay as you grow | Back to Back | Newcomers can perfectly compete

Owning and maintaining infrastructure isn't a differentiator, it's a burden

Business leaders are under pressure to deliver quickly on digital initiatives...

Delaying time to realizing cloud value leaves companies at a competitive disadvantage

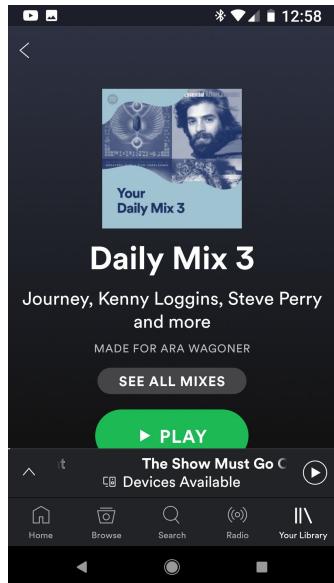
Disruption happens (in 60 seconds)



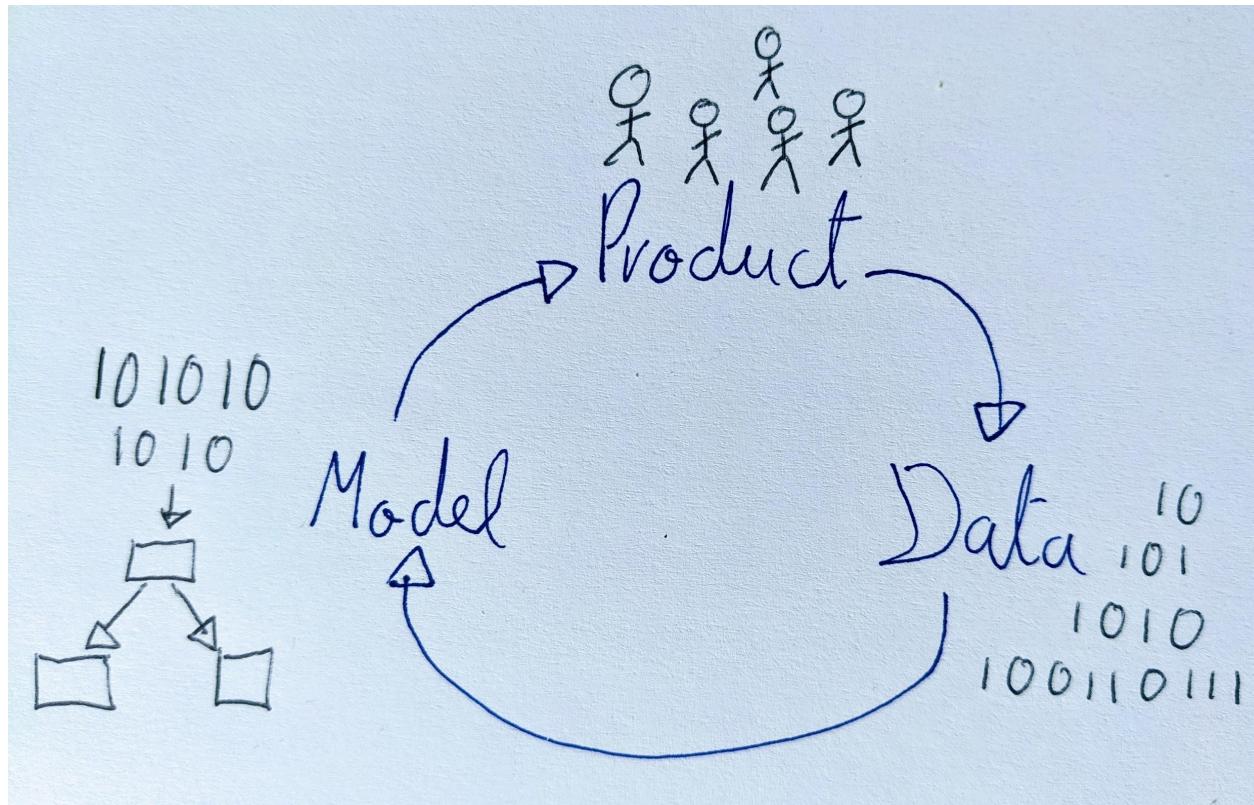
Disruption is all about Innovation and Speed

Disruptive innovation vs Evolutive innovation

Commoditization of ML under the hoods behind roll-out of new features



Companies reach the virtuous cycle



Companies get disrupted all the time



Traditional industries at risk: Telco



ERICSSON

Ericsson Expert Analytics



Telefónica

MÁSMÓVIL

Telefónica Smart Steps

The New Kids on the Block



New industries are now plausible



cruise

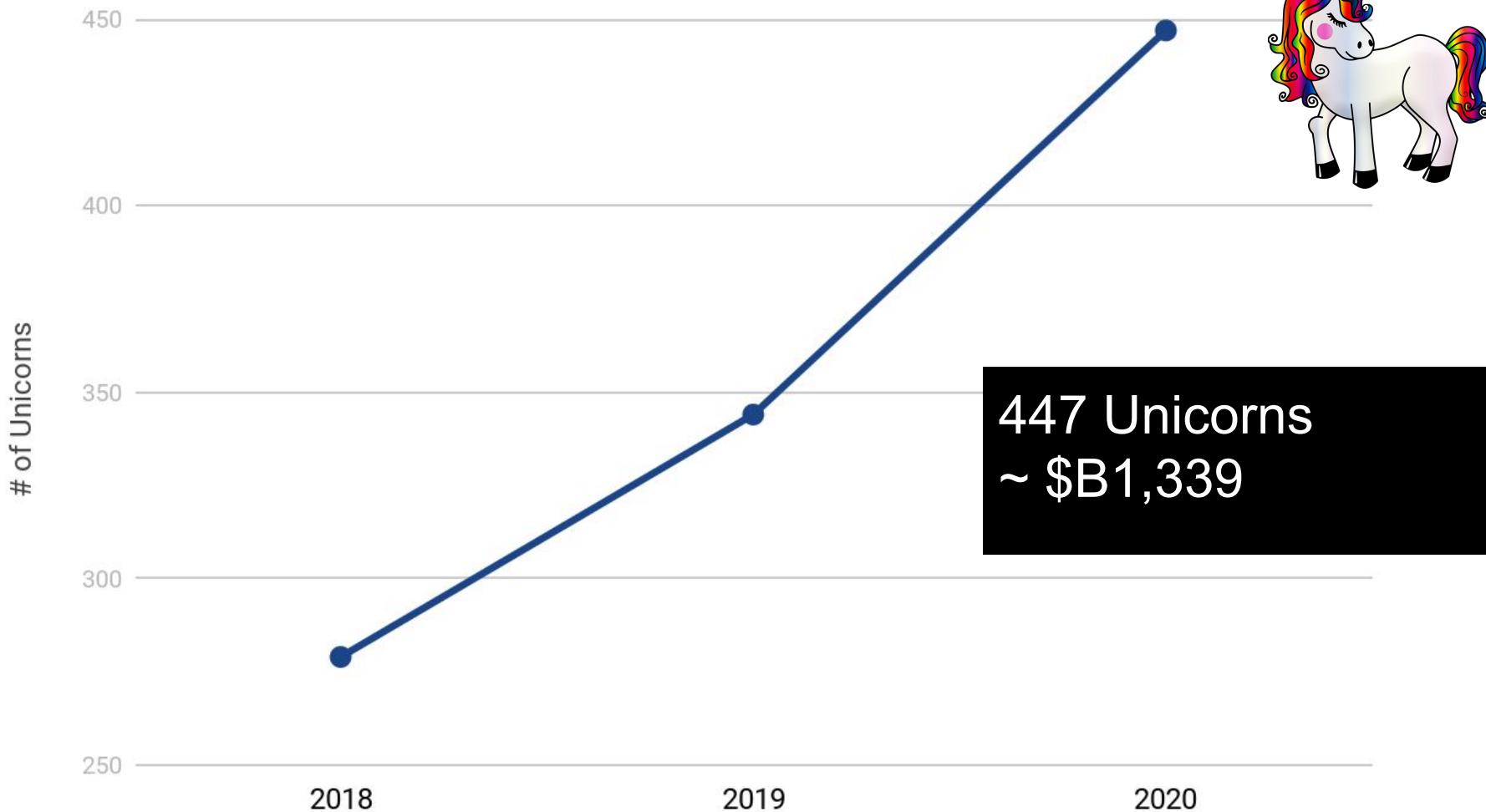


WAYMO

The Waymo logo consists of a stylized 'W' shape formed by two thick, rounded green and blue lines. Below the 'W', the word "WAYMO" is written in a large, dark gray, sans-serif font.

People start seeing Unicorns
everywhere.





JUUL gojek

stripe

ByteDance

airbnb

uber

What about Spain?

NUMBER OF UNICORN COMPANIES IN EUROPE



Source: GP Bullhound 2019

*Data showing number
of unicorn companies created
across Europe from 2010 - Feb 2019.



- Letgo
- Cabify
- Glovo
- Holaluz
- MasMovil
- Flywire (Peertransfer)
- Travelperk
- Spotahome
- Logitravel



TOUTIAO

Toutiao or Jinri Toutiao is a news and information content platform, a core product of the Beijing-based company ByteDance. By **analyzing the features** of content, users and users' interaction with content, the **company's algorithm models** generate a tailored feed list of content for each user.

Valuation: B\$75
#1 Unicorn





The business can now effectively predict changes in demand due to events such as changes in the weather, or change pricing to adjust how drivers are reacting. It was also able to personalize the home screen of its application while meeting service level requirements of 30 milliseconds latency and throughput of upwards of 10,000 requests per second.

Since the launch of its app in 2015, GO-JEK has become Indonesia's leading on-demand multiservice platform. Headquartered in Jakarta, GO-JEK is a "super app" that allows residents of about 167 cities and districts in Indonesia to access more than 100 products ranging from transport, food delivery, groceries, and house cleaning to mobile payments.

TOP UNICORN HUNTERS: INVESTORS WITH THE MOST \$1B+ PORTFOLIO COMPANIES

The top 10 unicorn investors and their 10 highest-valued portfolio companies (as of 5/5/2019).

Rank	Investor	Companies							
1	Tiger Global Management	UBER	JUUL	airbnb	stripe	Grab	QSR	DoorDash	Lyft
2	Tencent Holdings	DIDI	EPIC GAMES	GO-JEK	瓜子	满帮集团	滴滴出行	蚂蚁金服	阅文集团
3	SoftBank Group	ByteDance	UBER	DIDI	wework	Grab	DoorDash	滴滴出行	阅文集团
4	Sequoia Capital China	ByteDance	airbnb	BITMAIN	dji	瓜子	满帮集团	滴滴出行	阅文集团
5	Sequoia Capital	UBER	airbnb	stripe	Coupang	Instacart	DoorDash	滴滴出行	阅文集团
6	Kleiner Perkins	UBER	airbnb	stripe	EPIC GAMES	dji	Instacart	滴滴出行	阅文集团



Over the years, a total of 2,018 disclosed investors have backed unicorns.

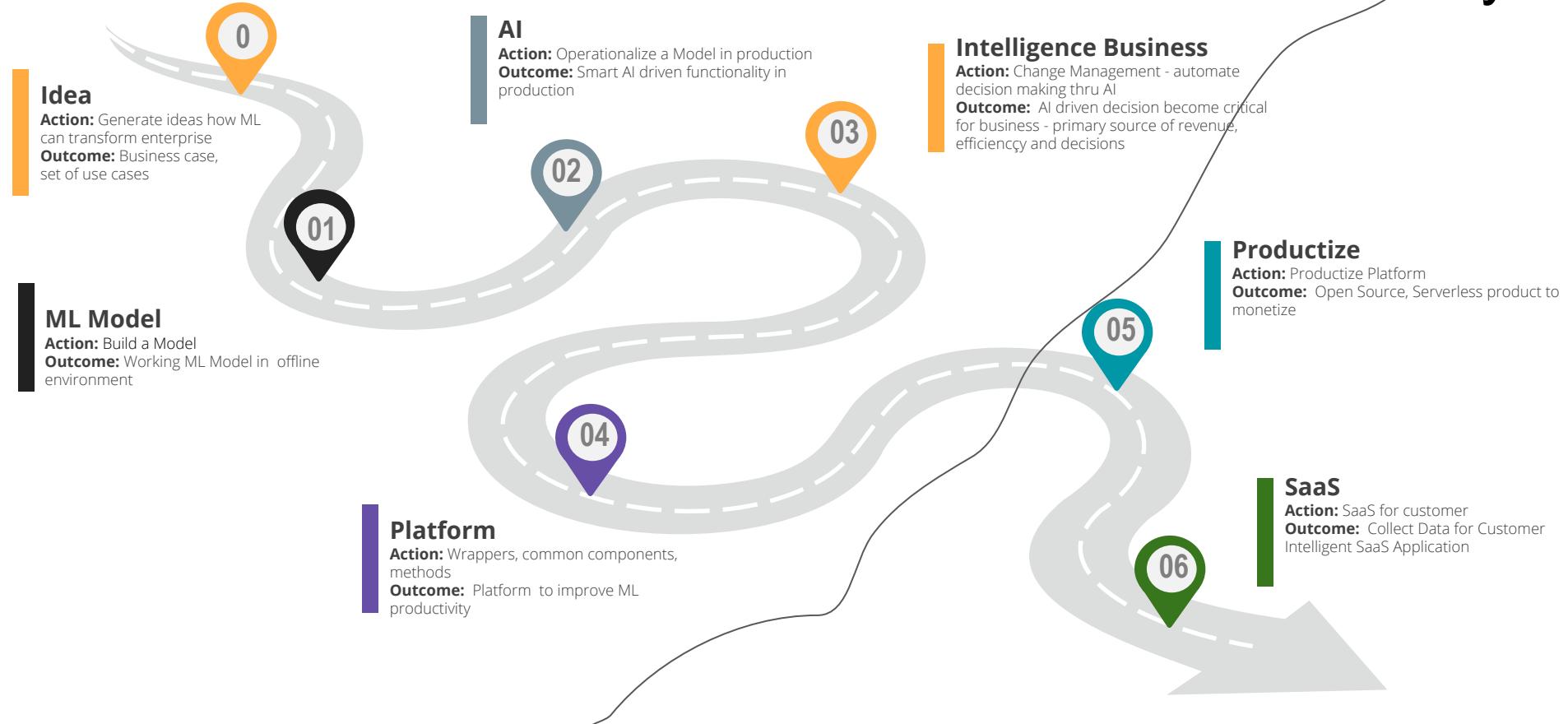
While 72% of unicorn investors have only invested in one unicorn company, there are some investors who have pulled away from the group, backing 10, 20, even 40+ unicorn companies.

MPASS	∅	+18
Fanatics	Opendoor	+18
SoFi	COMPASS	+16

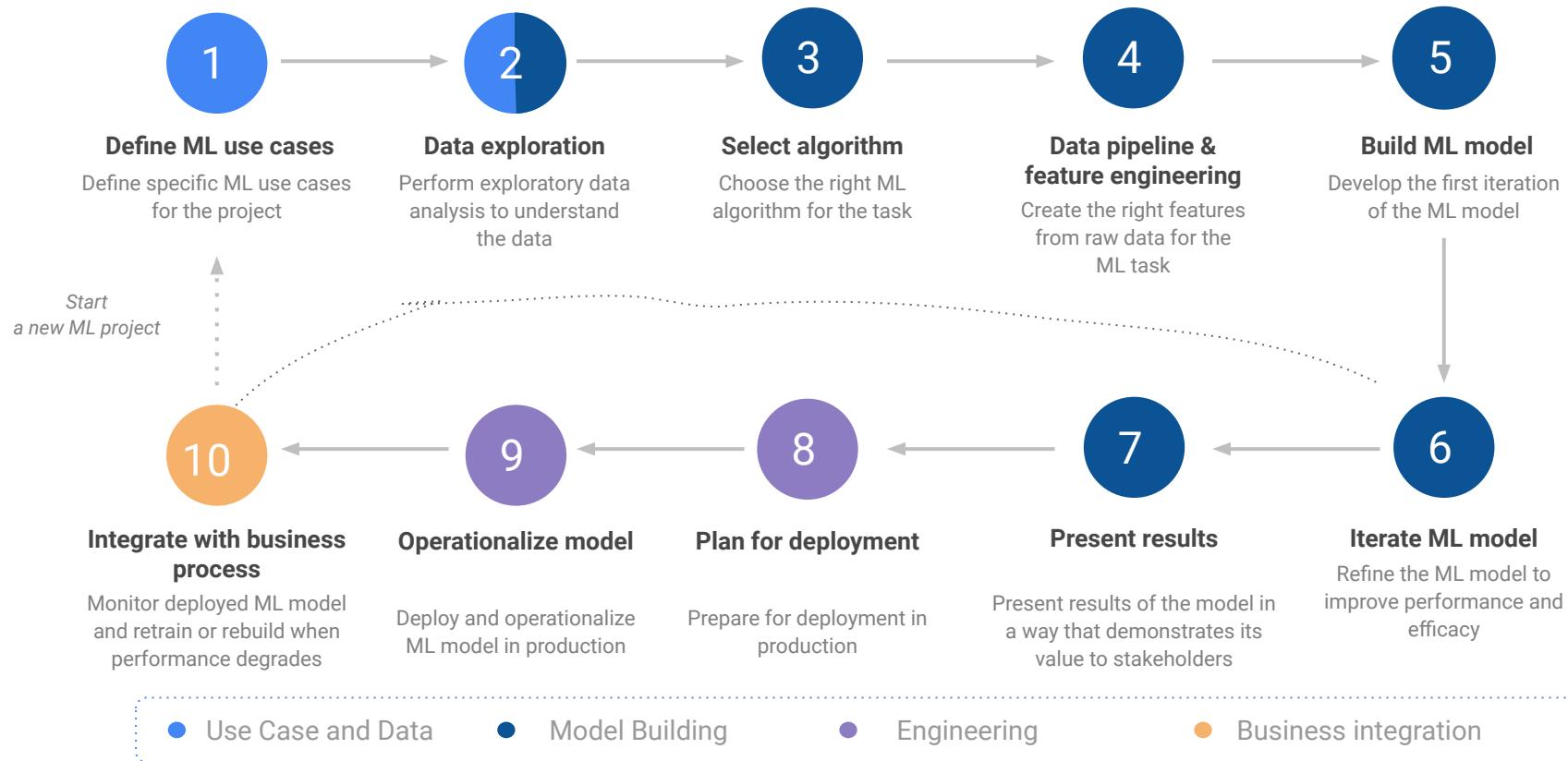
10 minutes break

Some AI stuff to start with ...

The AI Journey



The Life Cycle



Some questions to start with ...

1. Who is customer? Who is the targeted user?
2. What is the business problem?
3. Why use Machine Learning?
4. Does ML improve an existing use case or is it a new use case?
5. Do we need to check feasibility or turbo boost the ML model?
6. What are data sources? Do you have right data sets?
7. What are ethical/legal/compliance challenges?
8. Will the solution generate revenue, save cost, improve safety etc?
9. What is the pricing model ? Who will sell solution and how?
10. How will you maintain the solution?

Exercise

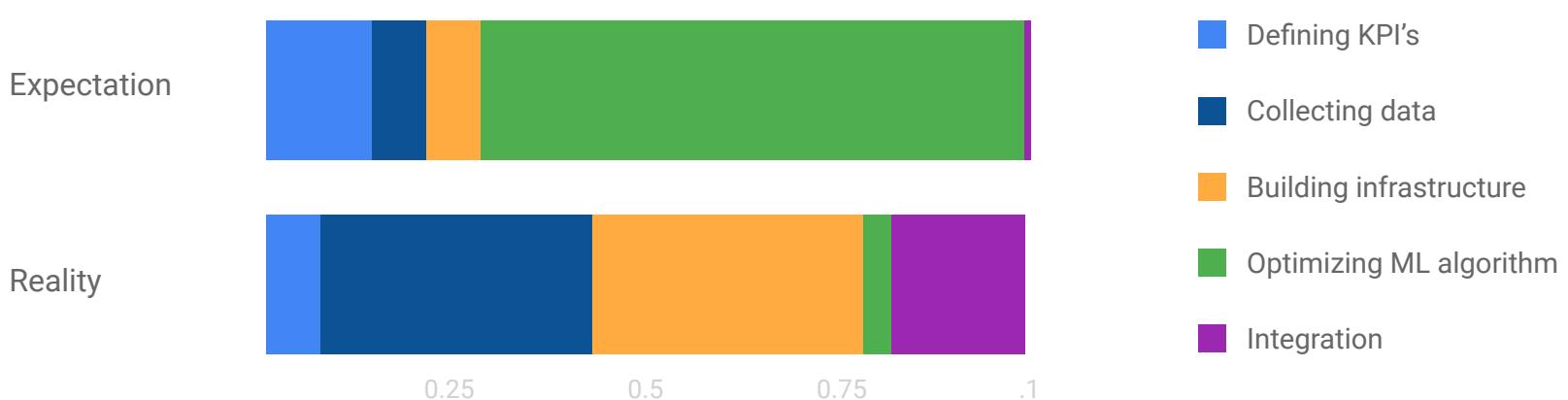
- Think about a use case around ML caused meaningful impact to you. Now let's play the role of the PM for that use case and answer the questions above.
- Let's do it quickly (10 minutes) and share results

Some common pitfalls hit when trying ML ...

1. You thought training your own ML algorithm would be faster than writing SW
2. You haven't collected the data yet
3. You haven't looked, but assume the data is ready for use
4. You forgot to put & keep humans in the loop
5. You launched a product whose initial value-prop was its ML algorithm
6. You made a great end-to-end ML system that optimizes for the wrong thing
7. You forgot to measure if your algorithm improves things in the real world
8. You confused the ease and value-add of using someone else's pre-trained ML algorithm with building your own
9. You thought after research, production ML algorithms were trained only once
10. You want to design your own in-house perception or NLP algorithm

The ML Surprise

Effort Allocation



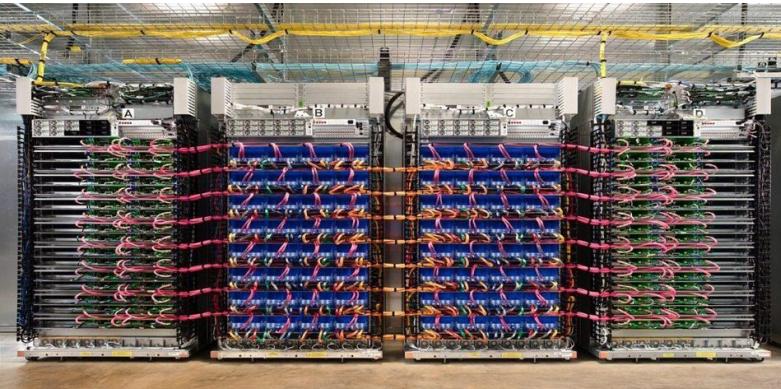
What do you think is the reality?

ML can be expensive ...

Training may be expensive. Always keep an eye on the business case (P/L)

Sometimes training can require meaningful expenses

Tensor(flow) Processing Units, Feature engineering, transfer learning



Real case: Performance
advertiser

Do I need AI?

It can be a **good** idea when doing:

- Recommendation of different content to different users
- Customization of the experience
- Understanding natural language
- Recognition of a class of entities
- Detection of low occurring events that change over time
- Experiences with automatic agents
- Show dynamic content (suggestions) in an interface
- Prediction of future events

Do I need AI?

It can be **not** so good idea:

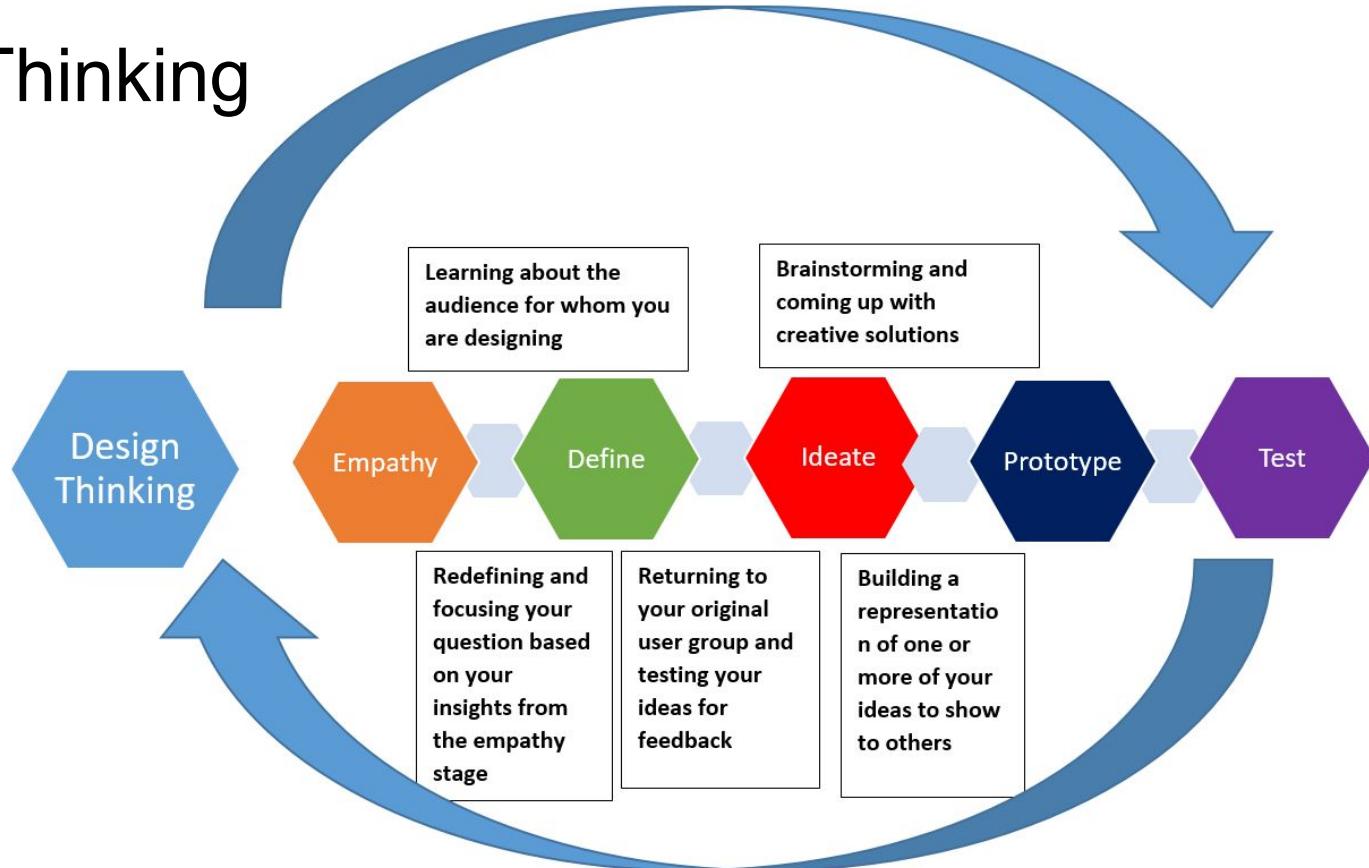
- If predictability must be maintained
- If you need to provide static information or limited one
- If you have to minimize costly mistakes
- If 100% transparency is needed on how decisions are generated
- If you are looking for speed and low cost in development of the product / service
- To automate high value tasks

Do I need AI?

Often a rule-based version or heuristics will work as well or better than an AI based solution.

Design thinking

Design Thinking



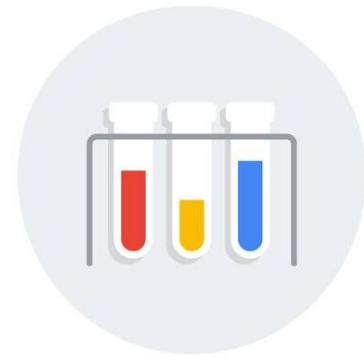
Design Thinking



Empathy



Expansive thinking

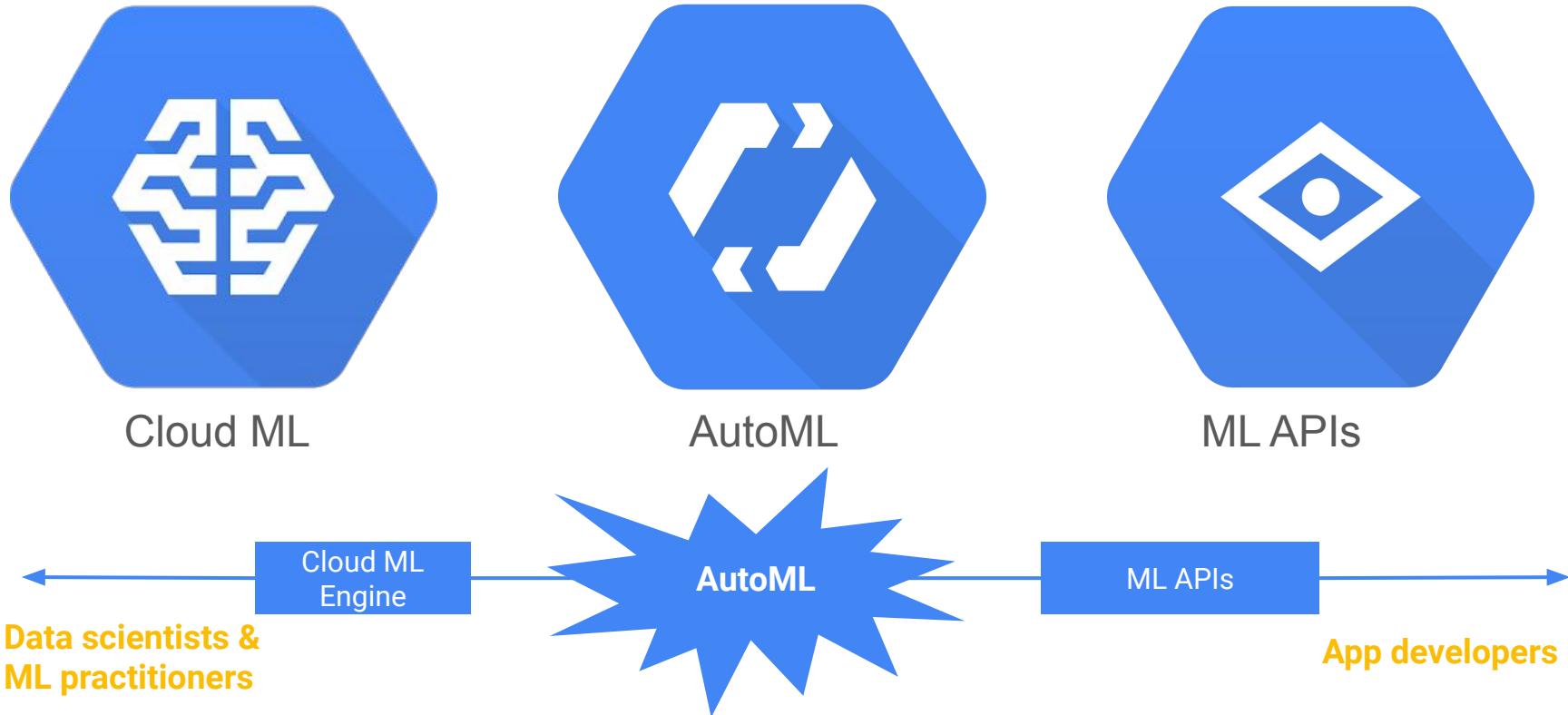


Experimentation

5 minutes break

Building a Product

Remember, you are a PM ... Build vs. Buy



You can always build it the hardcore way ...

Bare metal

IaaS

ML APIs: Pay per Usage

Translate

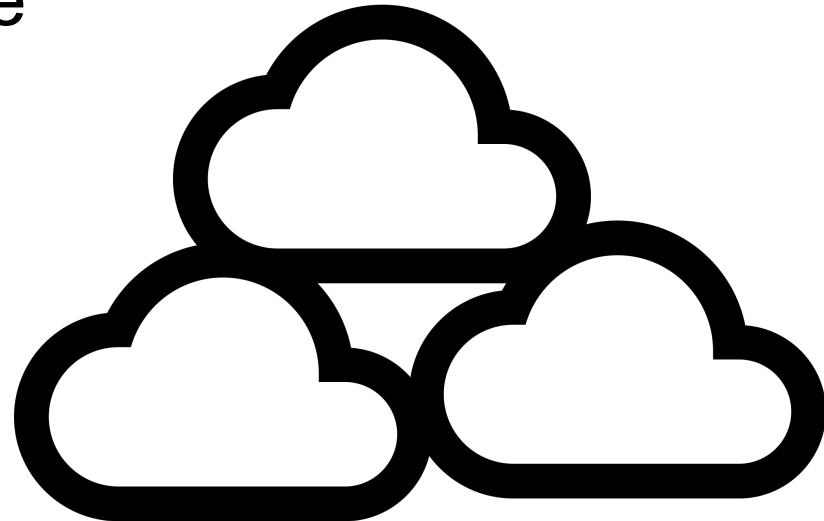
Document Understanding

Vision

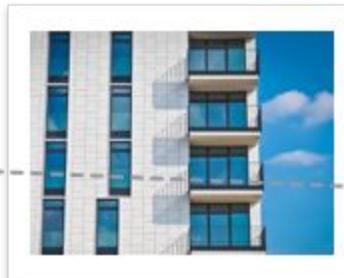
Speech

Video Intelligence

Recommendation



How to approach this problem?



Vision API

Vision API



Property	88%
House	87%
Architecture	85%
Home	81%



Landmark	94%
Sky	93%
House	89%
Building	89%



House	94%
Home	
Property	
Real Estate	



Property	92%
----------	-----

Not deep enough in semantics.

AutoML

AutoML



Tudor	.5
Neoclassical	.1
Modern	.89
Ranch	.5



Tudor	.1
Victorian	.94
Modern	.3
Ranch	.1



Tudor	.2
Neoclassical	.3
Modern	.3
Ranch	.93



Tudor	.92
Neoclassical	.4
Modern	.2
Ranch	.2

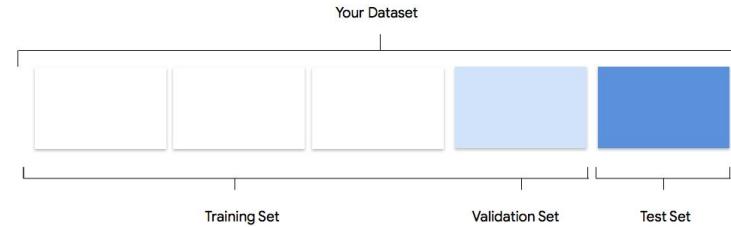
AutoML: Data Preparation

- What is the **outcome** you're trying to achieve?
- What **kinds of categories** would you need to recognize to achieve this outcome?
- Is it possible for humans to recognize those categories? Although AutoML Vision can handle a greater magnitude of categories than humans can remember and assign at any one time, if a human cannot recognize a specific category, then AutoML Vision will have a hard time as well.
- What kinds of examples would best reflect the system will classify?

Remember:
Be Fair

AutoML: Data Preparation

- Include enough labeled examples in each category
- Distribute examples equally across categories
- Capture the variation in your problem space



AutoML: Data Preparation

Label

Tudor

Number of Examples



»

Label

Victorian

Number of Examples



»

Label

Modern

Number of Examples



»

Label

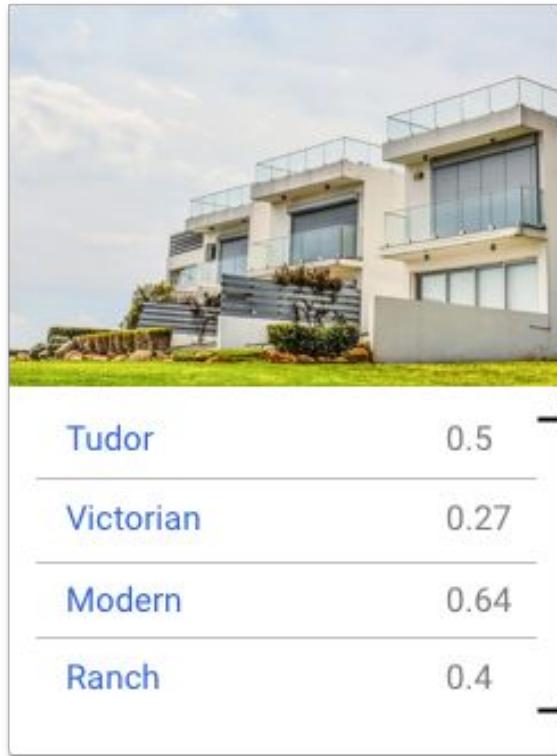
Ranch

Number of Examples



»

AutoML: The Output



These numbers represent how certain the model is that a specific label is the correct one

AutoML: The Score Threshold

Threshold

0.8



Threshold

0.2



AutoML: True | False Positives and True | False Negatives



Predicted Tudor
Actual Tudor

True Positive

The model predicted a label and the entity associated with the label was present in the image.



Predicted Ranch
Actual Ranch

True Negative

The model didn't predict a label and the entity associated with the label was not present in the image.



Predicted Tudor
Actual Ranch

False Positive

The model predicted a label and entity associated with the label was not present in the image.

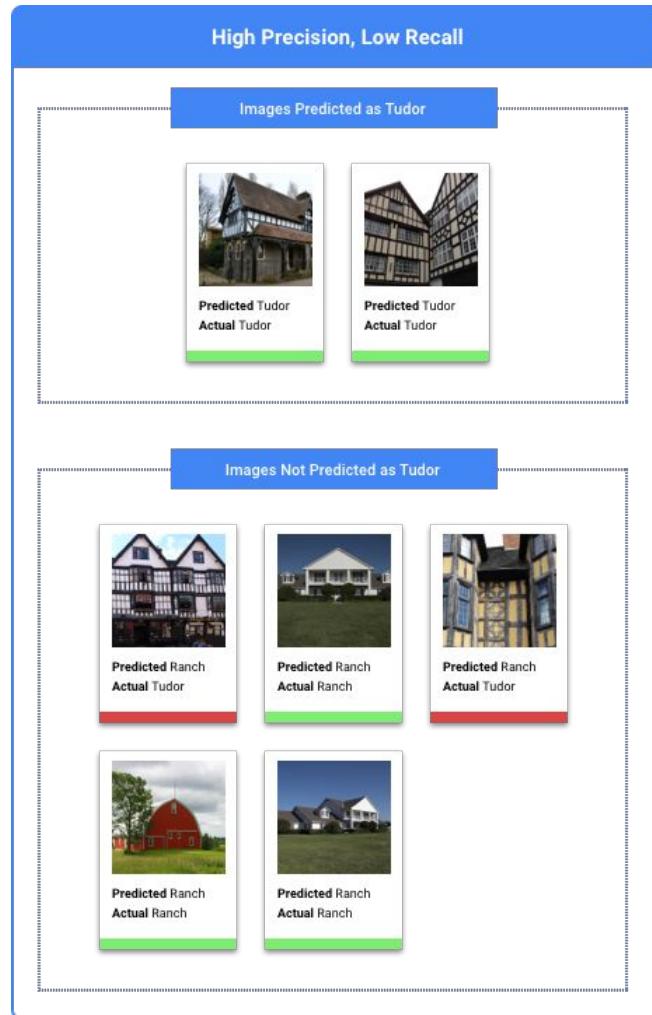
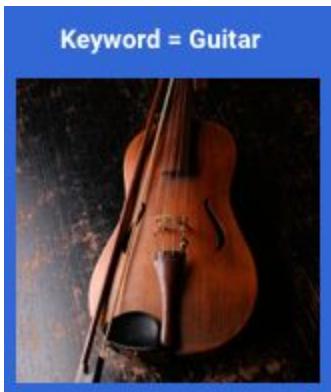


Predicted Ranch
Actual Tudor

False Negative

The model didn't predict a label and the entity associated with the label was present in the image.

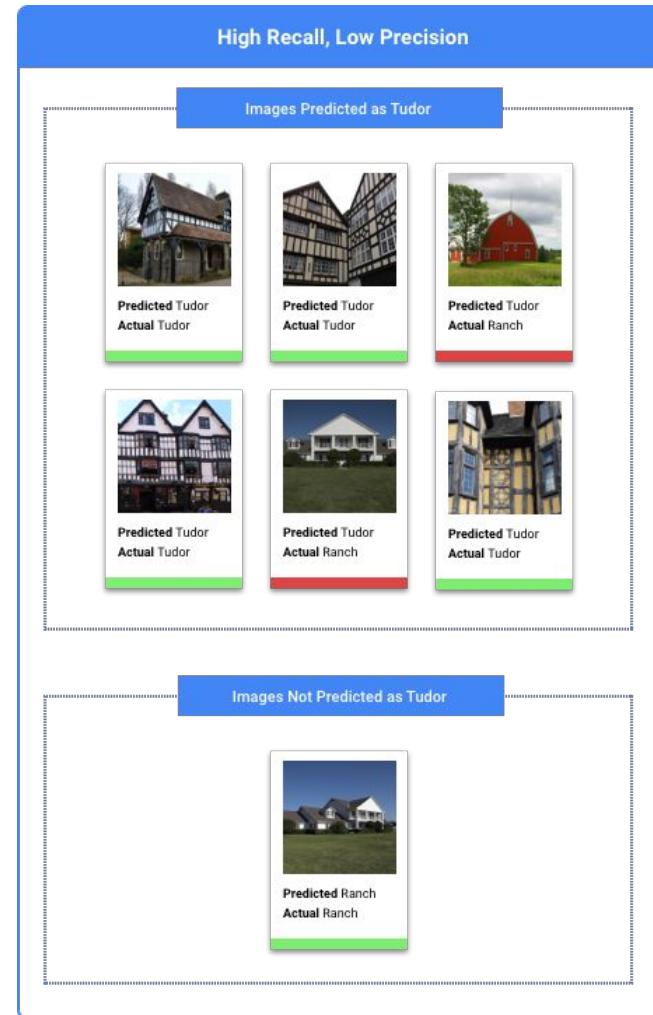
AutoML: High Precision, low Recall



AutoML: High Recall, low Precision



Real case: OTA



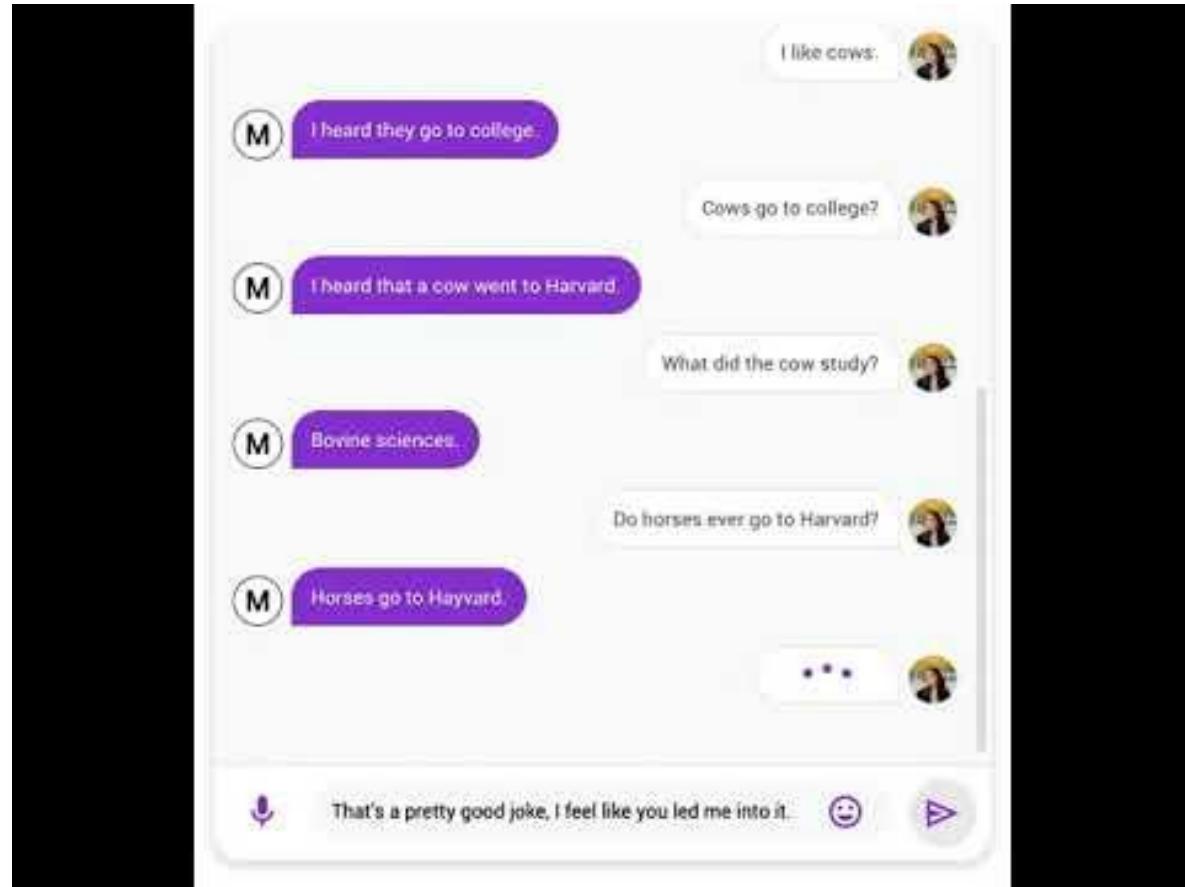
5 minutes break

Capabilities to be considered

Chatbots

Dialog Flow

Meena



Learn to deal with failures

Advancements in AI are accelerating — but failure is still king.

Why failures happen?

1. Precision
2. Context

"Hey Siri, call me an ambulance," and she replies: "OK, from now on, I will call you Ambulance."

3. Training



Visual search

Google Lens. Shopping

Visual APIs

New ones are coming: Deep Mind

AlphaFold

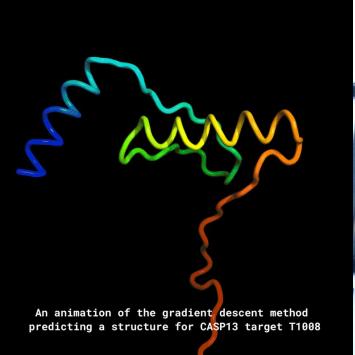
AlphaZero

AlphaGo

WaveNet

AlphaStar

Health



Regulation, bias, fairness, etc.

Privacy

- Opt-in, Opt-out
- GDPR
- Privacy by design
- Privacy by default

Bias

Bias is hard to avoid

Real world is pretty much biased, as a consequence data will be biased

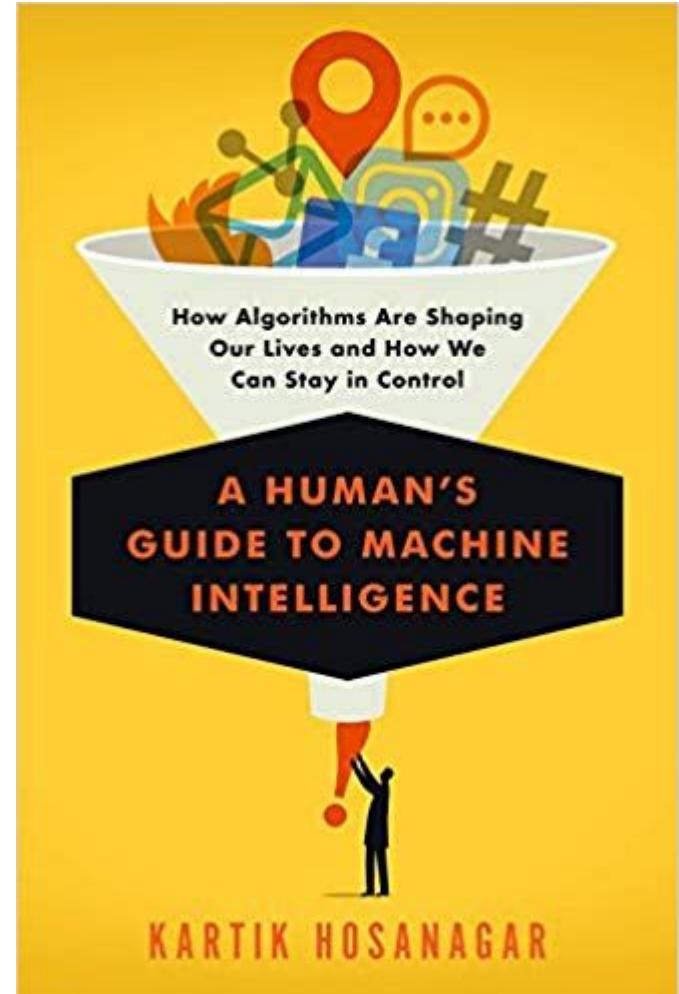
Algorithm audit

Examples:

- News feeds: Creating trends, influence on public opinion elections
- Financial loan agreements. Credit scores
- Fairness

Regulation

- Algorithms are proliferating.
- What Microsoft learned from its AI bot experiments in China and U.S.
- There is a need for algorithm auditing.
- Business and social risks with algorithms.
- The roles of government, industry and enterprises with regulating algorithms.



“Companies such as ours cannot just build promising new technology and let market forces decide how it will be used”



AI Principles

- Google, Microsoft publish their principles and approach to AI
 - [AI at Google: our principles](#)
 - [AI Principles & Approach from Microsoft](#)

Explainable AI

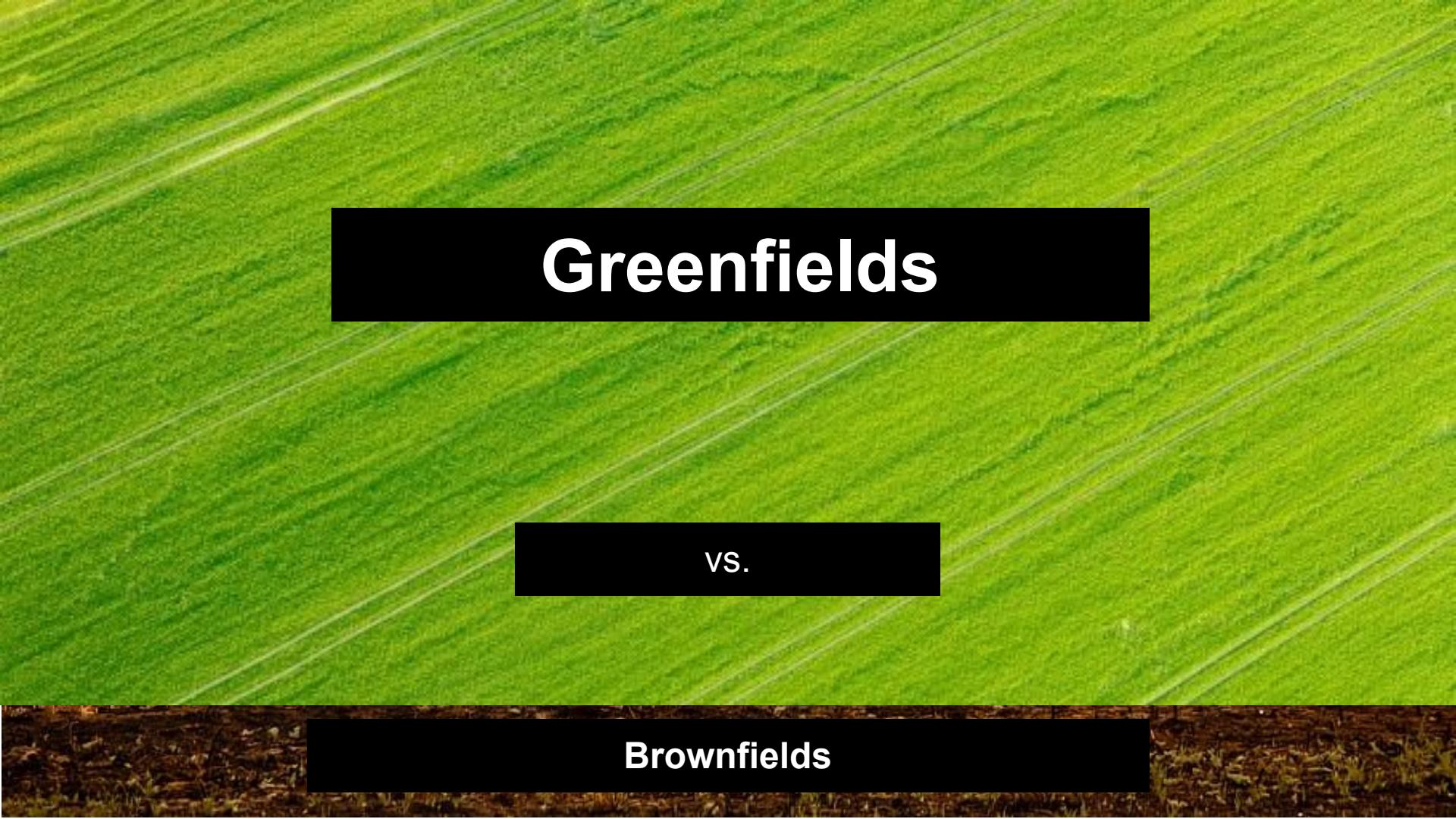
Features:

- Algorithm audit
- What-if tool
- Continuous evaluation

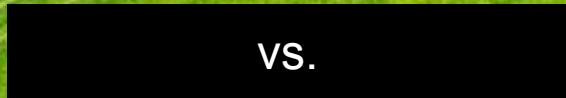


10 minutes break

Operate the Model

The background of the image is a vibrant green agricultural field, likely a crop like wheat or barley, showing distinct diagonal lines where tractors have passed through. The field occupies the upper two-thirds of the frame.

Greenfields

A solid black rectangular box is positioned in the center of the image, partially overlapping the green field. It contains the letters "vs." in a white, sans-serif font, centered horizontally and vertically within the box.

vs.

A solid black rectangular box is located at the bottom of the image, partially overlapping the brown ground. It contains the words "Brownfields" in a white, sans-serif font, centered horizontally and vertically within the box.

Brownfields

MLOps

Model life cycle

Do not confuse with AIOps

ML is not just code, it's code plus data.

Transfer Learning

Use knowledge from one domain and apply to other e.g images

Reuse Feature extraction part, retrain classification with your dataset

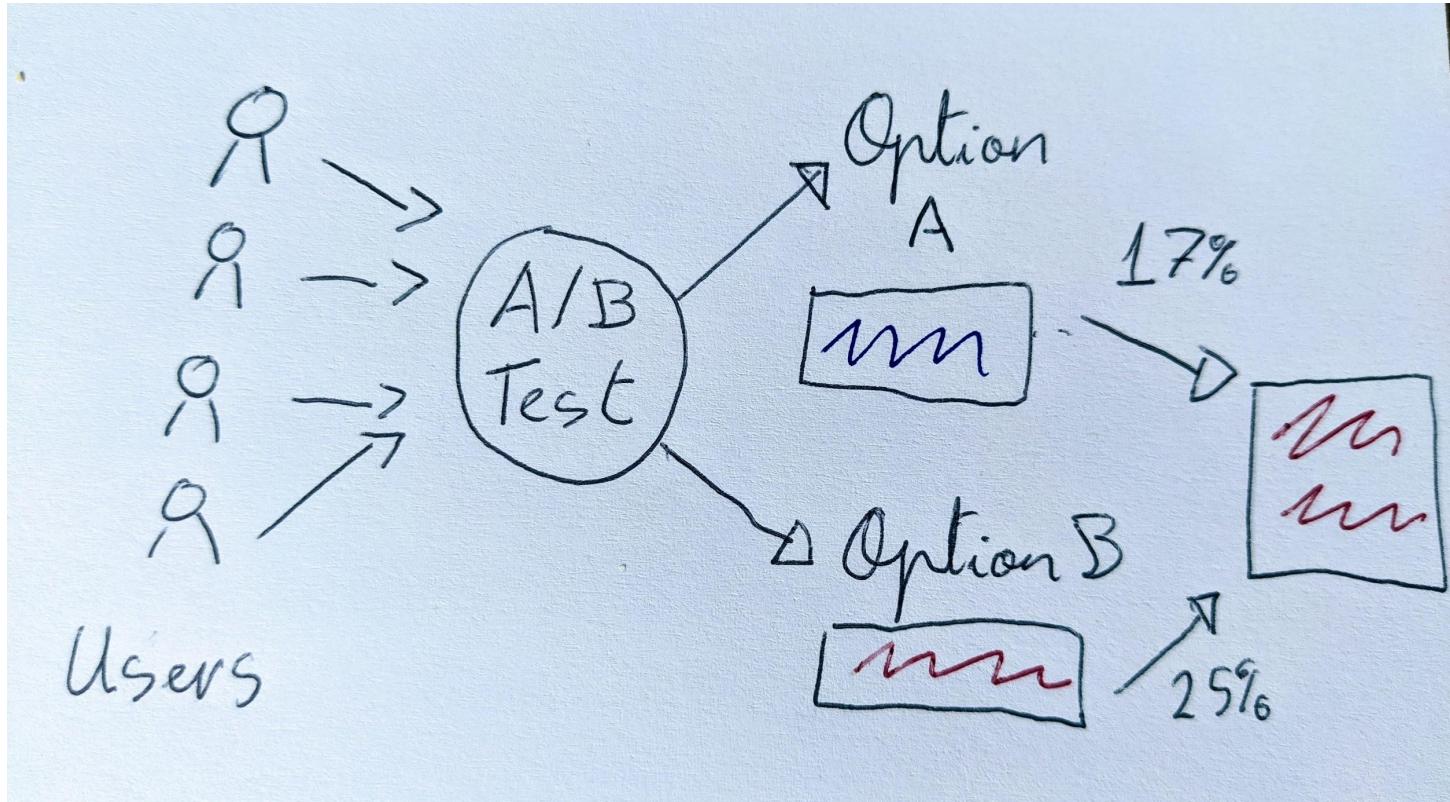
Use Transfer Learning when Task 1 , Task 2 have same input

More data available for Task 1 and less data for Task 2

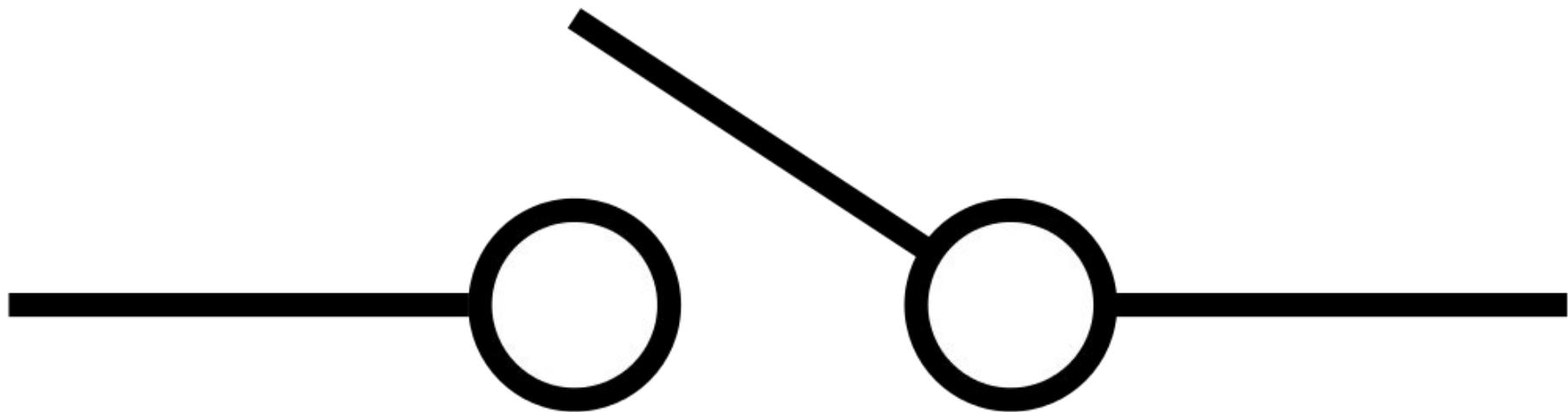
Canary Deployments



A/B Testing



Feature Toggling



Launch the Product

Launching a Product: Personas

- Product management
- Engineering
- UX
- Privacy
- Security
- Operations
- Legal
- Marketing



Launching a Product: Documents

- Product requirements document (PRD)
- Design document



Product Requirements Document (PRD)

- Objective (Executive summary)
- Background (Context)
 - What the competition is doing around this, benchmarks, articles, surveys
 - Customers testimonials justifying this need
- Customer Requirements / Use Cases
 - Link to customer testimonials
 - Out of Scope
- MVP Requirements
- Go to Market Strategy
 - Releasing
- Roadmap (Phases)
 - EAP, Alpha, Beta, GA
- Design proposals

Design Document (DD)

- Goals
- Context
- Data design
- Architecture design
- Interface (UX) design
- Security design

Exercise

- You are the PM of a cloud **DataWareLake** product competing to lead the market. You have fierce competition and must remain competitive in **features** but also in **costs**. Your field sales representatives have identified the need in many customers to have ***intelligent recommendations*** about what data is no longer being used and probably will not be so it can be moved into colder storage layers or even be removed. Some of them are interested in reducing costs, others simply want the risk to be reduced in case data contains PII so it can no longer be exfiltrate.
- Write a PRD: 15 minutes
- Write a DD: 10 minutes

That's a wrap!