DESIGNING MACHINE LEARNING

A Multi-Disciplinary Approach

ML+UX

Machine Learning + User Experience

Abhay Agarwal (abhayka@stanford.edu)
Michelle Carney (michelle.carney@berkeley.edu)

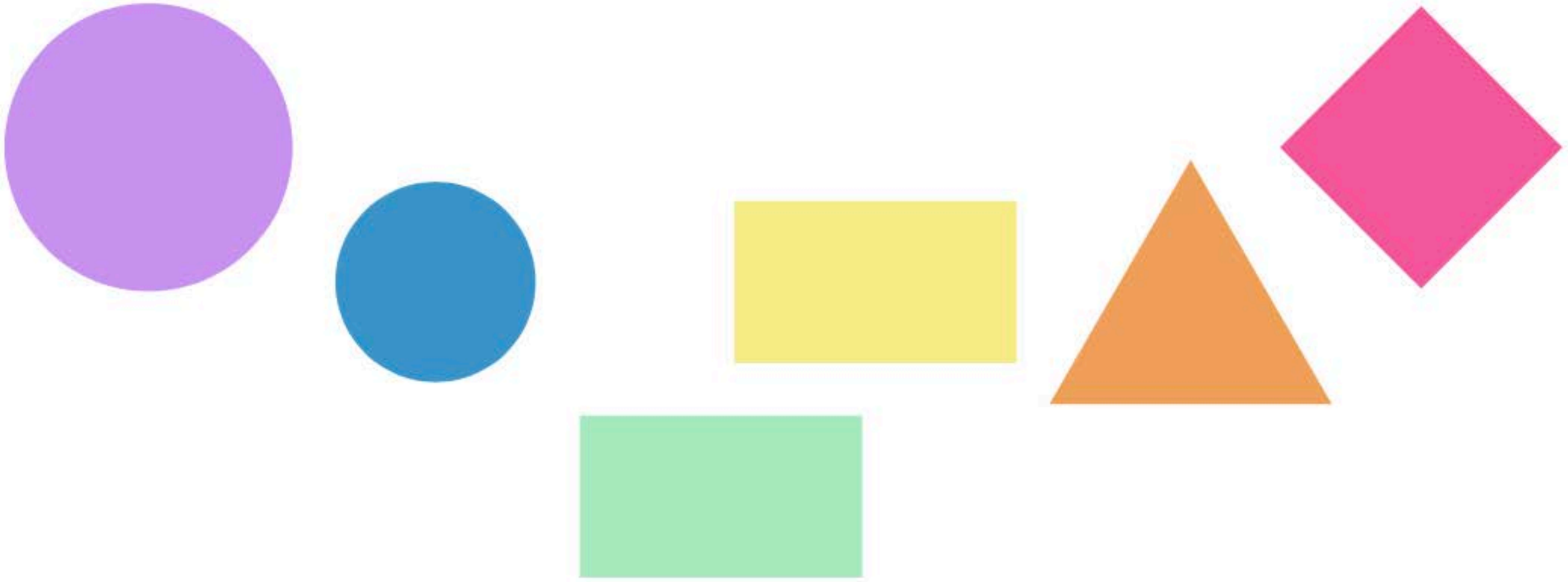
data = Machine Learning, AI, Data Analysis, Statistical Inference, and more
ux = User Experience Design, Research, Interaction Design, Info Architecture, and more

What is this talk about?

How do I create a wonderful user experience (UX) in an ML product?

Why do most ML products fail in the real world?

How do I turn my UX findings into actual product improvements?



MODEL INTERACTION

let's zoom out for a second..

What is User Experience Design?

UX designers work in a rather conceptual ***mode, with*** rough representations ***of the world, to improve the*** subjective aspects ***of a product***

Crafting

How do I create this UX?

Crafting

How do I create this UX?

Mapping

What does our UX look like?

Crafting

How do I create this UX?

Researching

What is a good UX?

Mapping

What does our UX look like?

Crafting

How do I create this UX?

Researching

What is a good UX?

Mapping

What does our UX look like?

Monitoring

Where is our UX failing?

Crafting

How do I create this UX?

Researching

What is a good UX?

Creative

Descriptive

Mapping

What does our UX look like?

Monitoring

Where is our UX failing?

Idealistic

Researching

What is a good UX?

Crafting

How do I create this UX?

Creative

Descriptive

Mapping

What does our UX look like?

Monitoring

Where is our UX failing?

Realistic

What makes UX different when you have ML?

Idealistic

Crafting

Researching

Creative

Descriptive

Mapping

Monitoring

Realistic

Idealistic

~~Crafting~~

No engineering chops

Researching

Creative

Mapping

Descriptive

Monitoring

Realistic

Idealistic

~~Crafting~~

No engineering chops

~~Researching~~

Can’t turn ideas into specs

Creative

Mapping

Descriptive

Monitoring

Realistic

Idealistic

~~Crafting~~

No engineering chops

~~Researching~~

Can't turn ideas into specs

Creative

~~Mapping~~

Black Box / Opaque System

Descriptive

Monitoring

Realistic

Idealistic

~~Crafting~~

No engineering chops

~~Researching~~

Can't turn ideas into specs

Creative

~~Mapping~~

Black Box / Opaque System

Descriptive

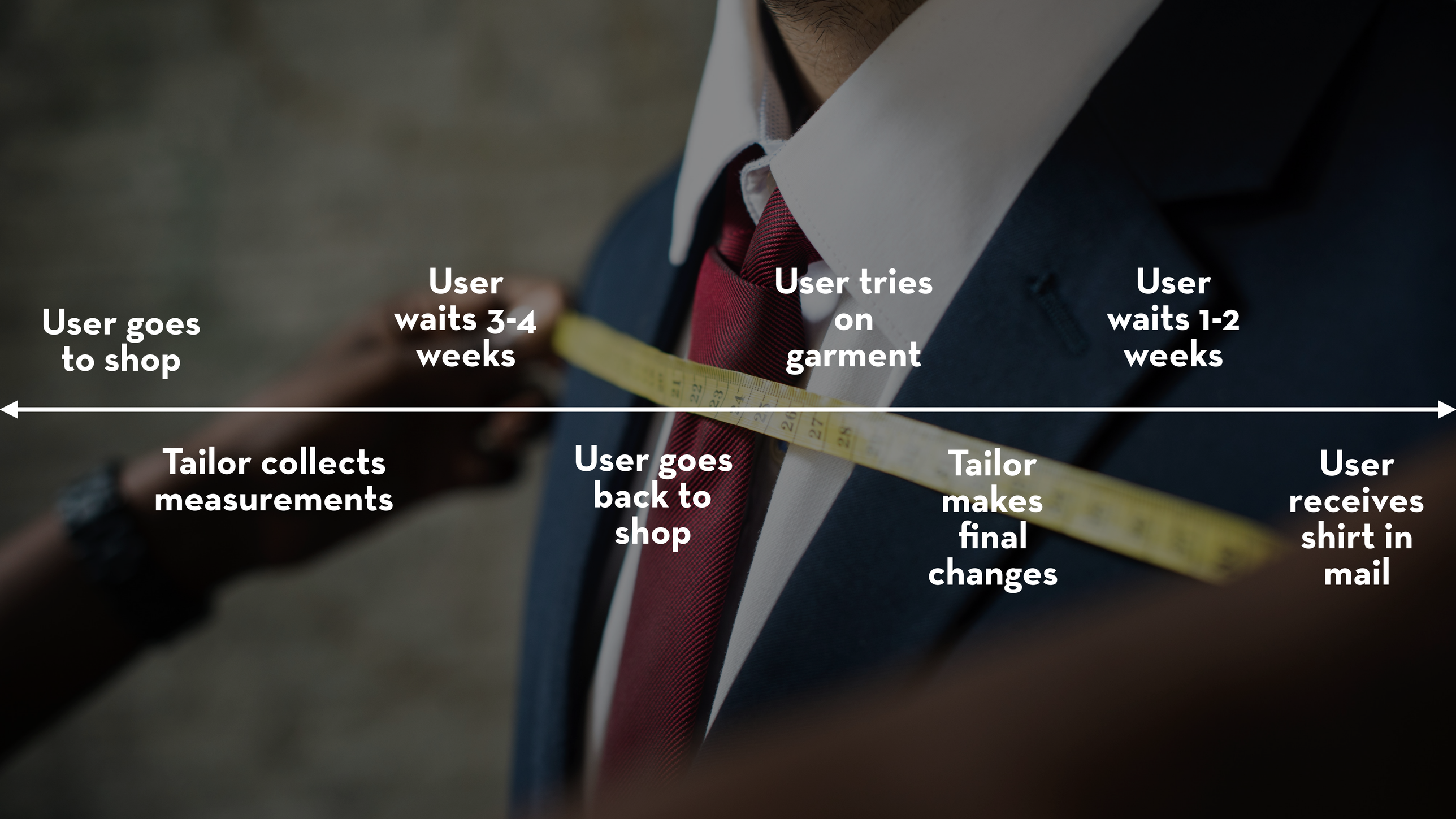
~~Monitoring~~

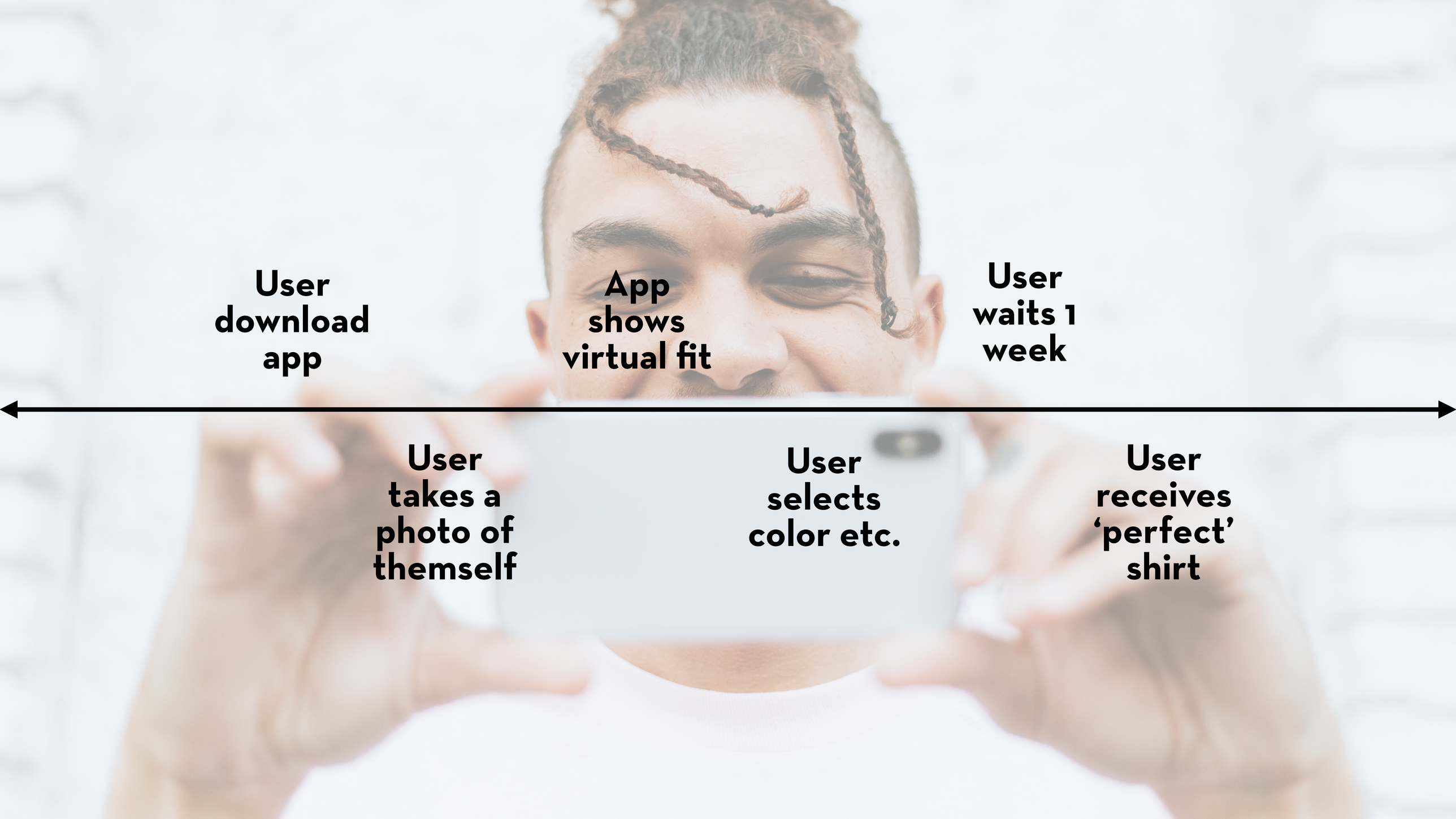
Product too customized

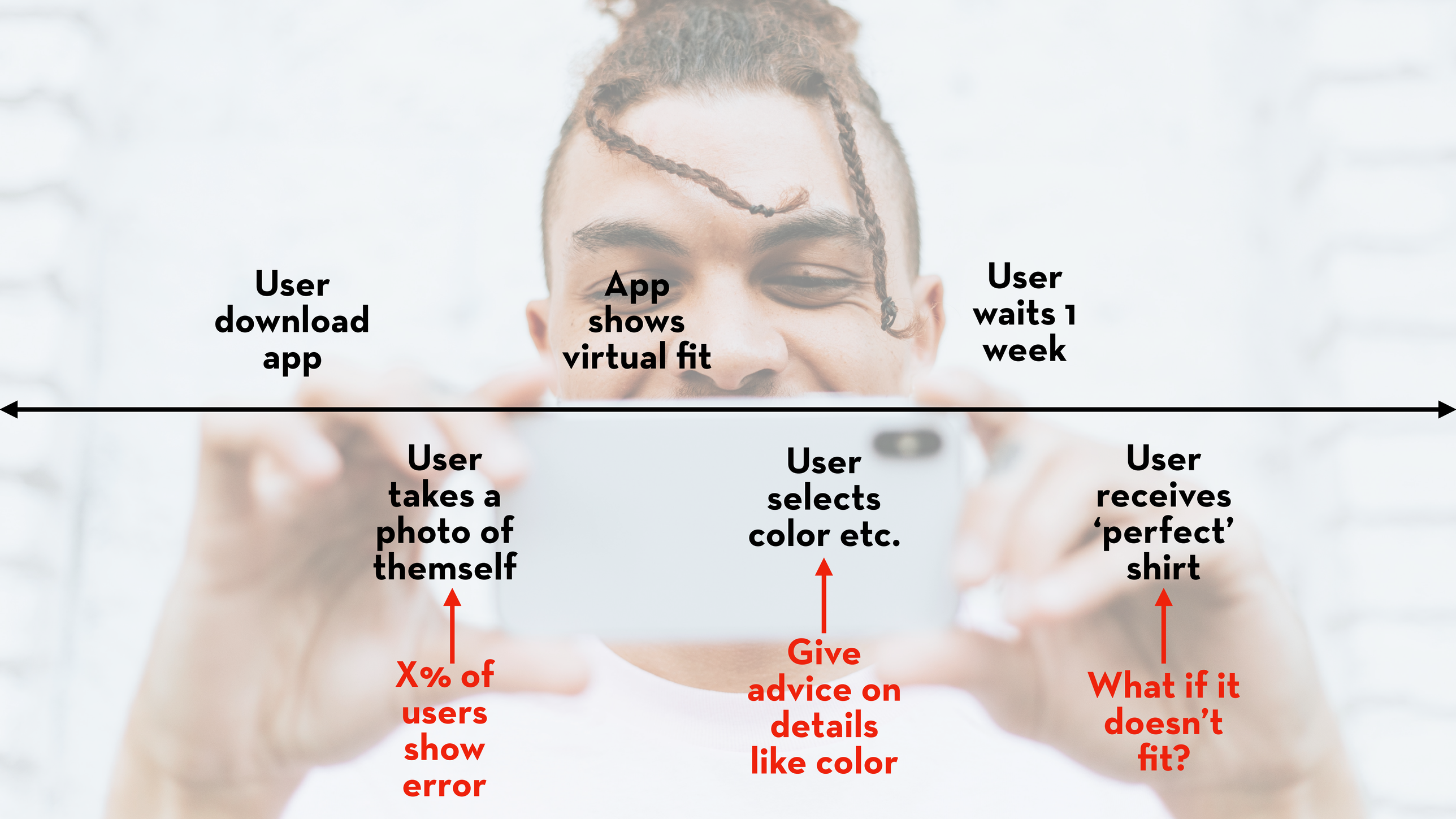
Realistic

A hypothetical system..

“AI Tailor”


User goes to shop
User waits 3-4 weeks
User tries on garment
User waits 1-2 weeks
Tailor collects measurements
User goes back to shop
Tailor makes final changes
User receives shirt in mail


User download app
App shows virtual fit
User waits 1 week
User takes a photo of themself
User selects color etc.
User receives ‘perfect’ shirt

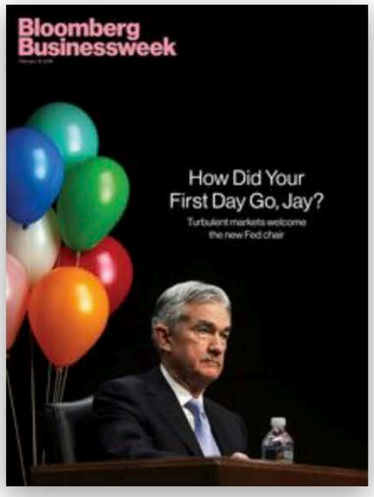

User download app
App shows virtual fit
User waits 1 week
User takes a photo of themself
X% of users show error
User selects color etc.
Give advice on details like color
User receives ‘perfect’ shirt
What if it doesn’t fit?

February 5, 2018, 8:01 AM PST

Don’t Use This AI Tailor ... Yet

● After three shirts we ordered didn’t fit, Original Stitch took down its measurement software.

By Pavel Alpeyev and Jason Clenfield



From the Issue
February 12, 2018

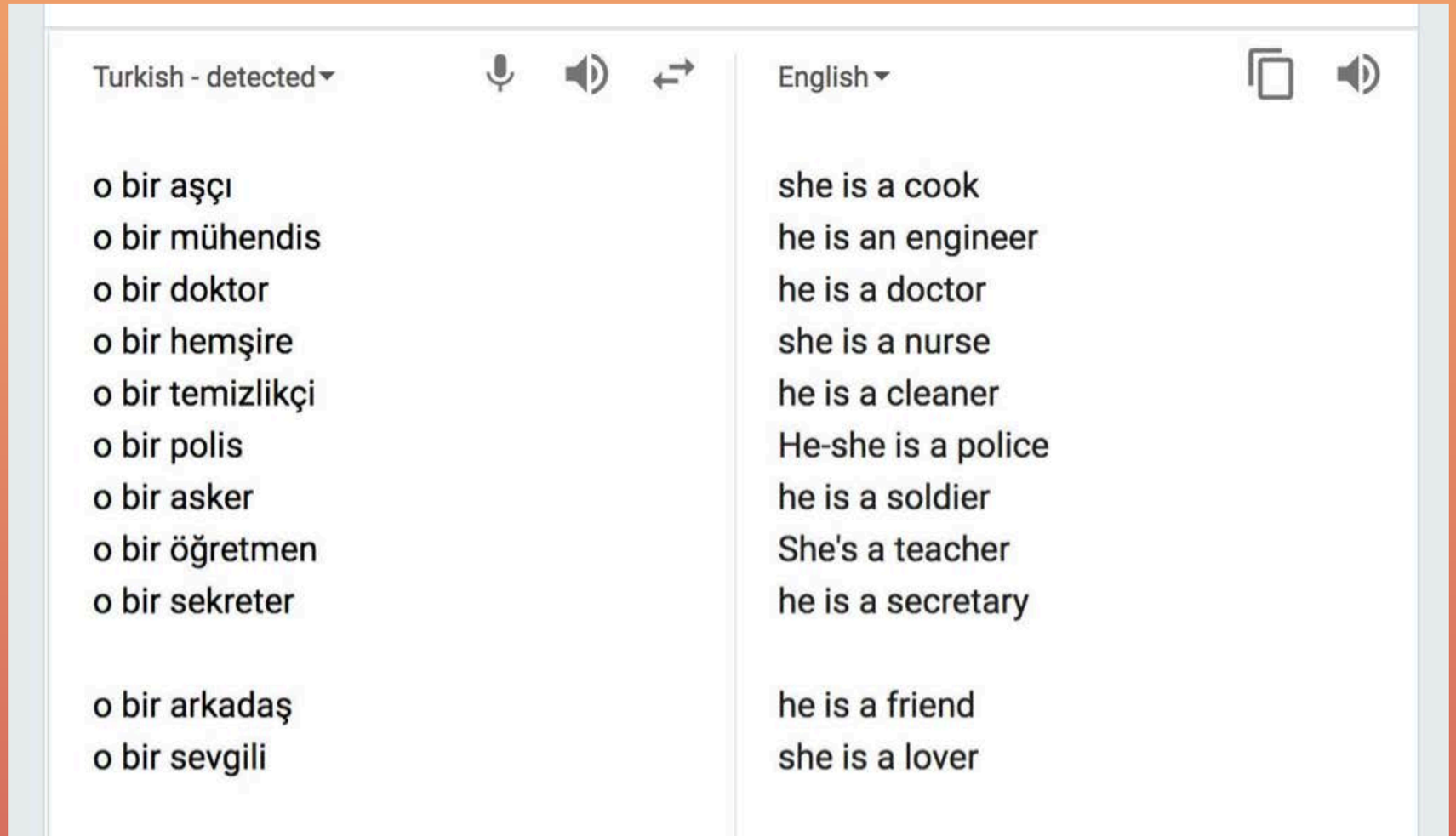
Subscribe Reprints

How do we solve this?

Case Studies
Design Principles
Experience & Practice!

CASE STUDY

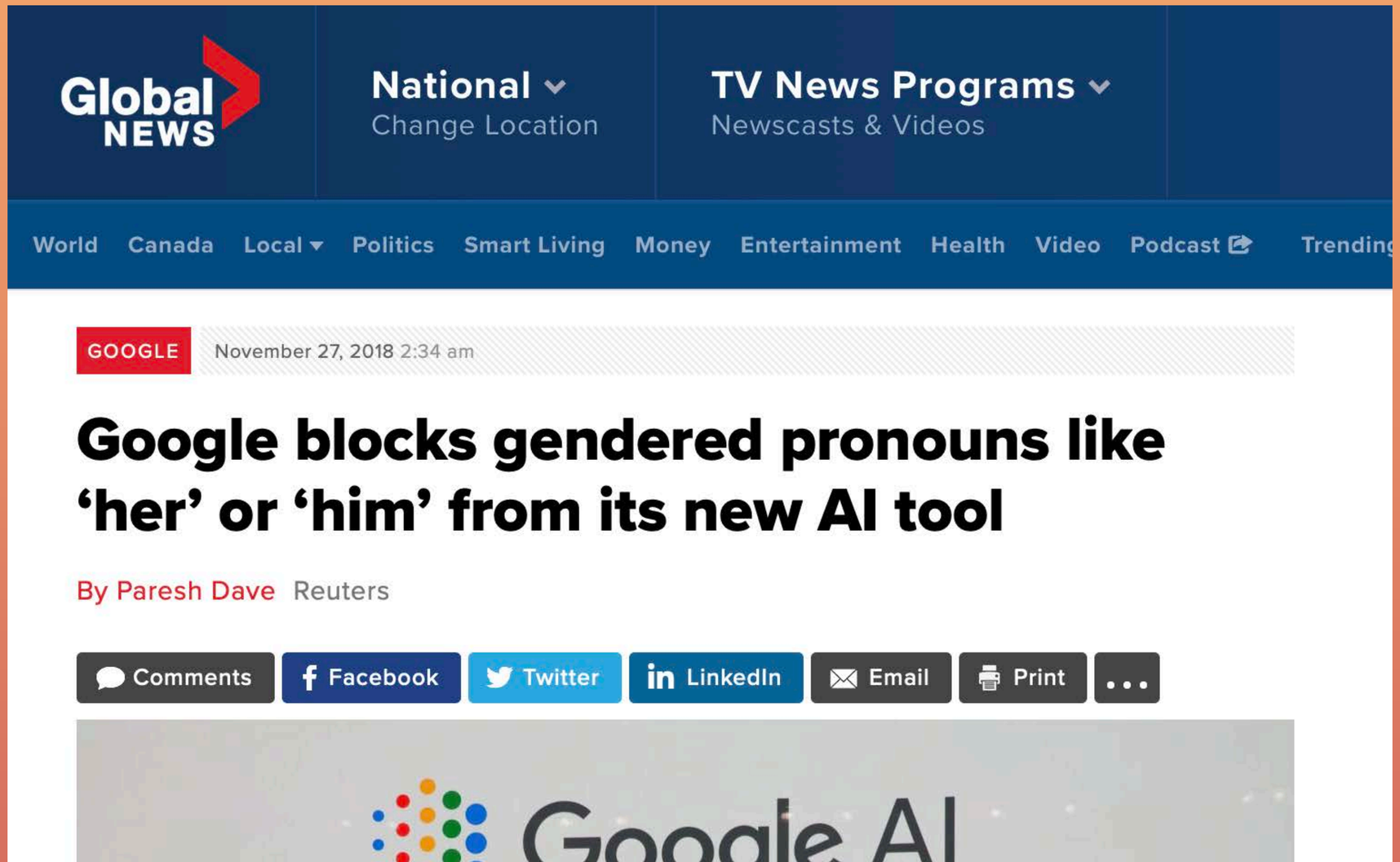
Turkish - English Translation



Turkish - detected	English
o bir aşçı	she is a cook
o bir mühendis	he is an engineer
o bir doktor	he is a doctor
o bir hemşire	she is a nurse
o bir temizlikçi	he is a cleaner
o bir polis	He-she is a police
o bir asker	he is a soldier
o bir öğretmen	She's a teacher
o bir sekreter	he is a secretary
o bir arkadaş	he is a friend
o bir sevgili	she is a lover

CASE STUDY

Turkish - English Translation



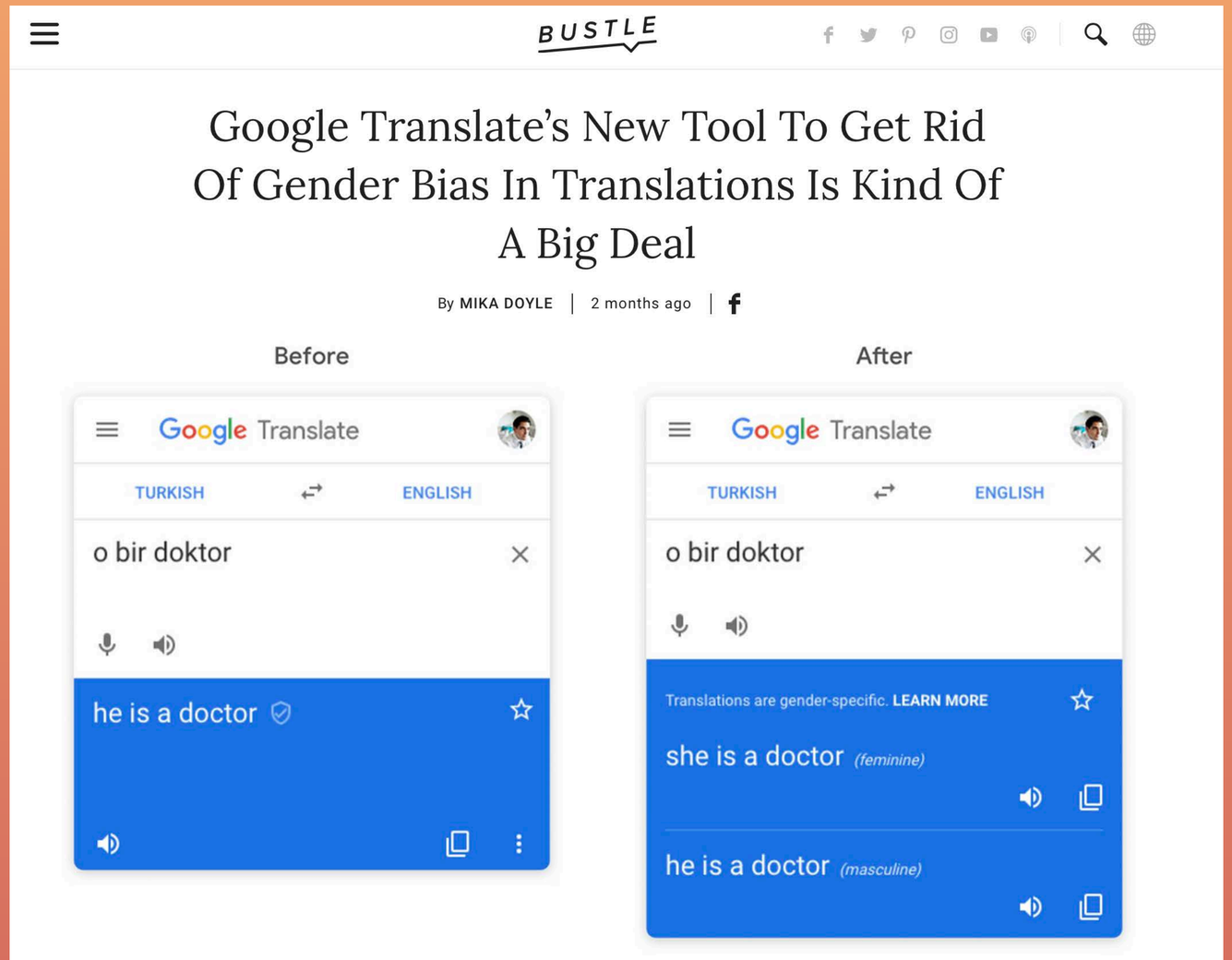
Case Study

Turkish - English Translation

“Data-UX Mismatch”

Case Study

Turkish - English Translation



Case Study

Target Coupons

3,503,091 views | Feb 16, 2012, 11:02am

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Kashmir Hill Forbes Staff
Welcome to The Not-So Private Parts where technology & privacy collide

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target TGT +1.89%, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Target has got you in its aim

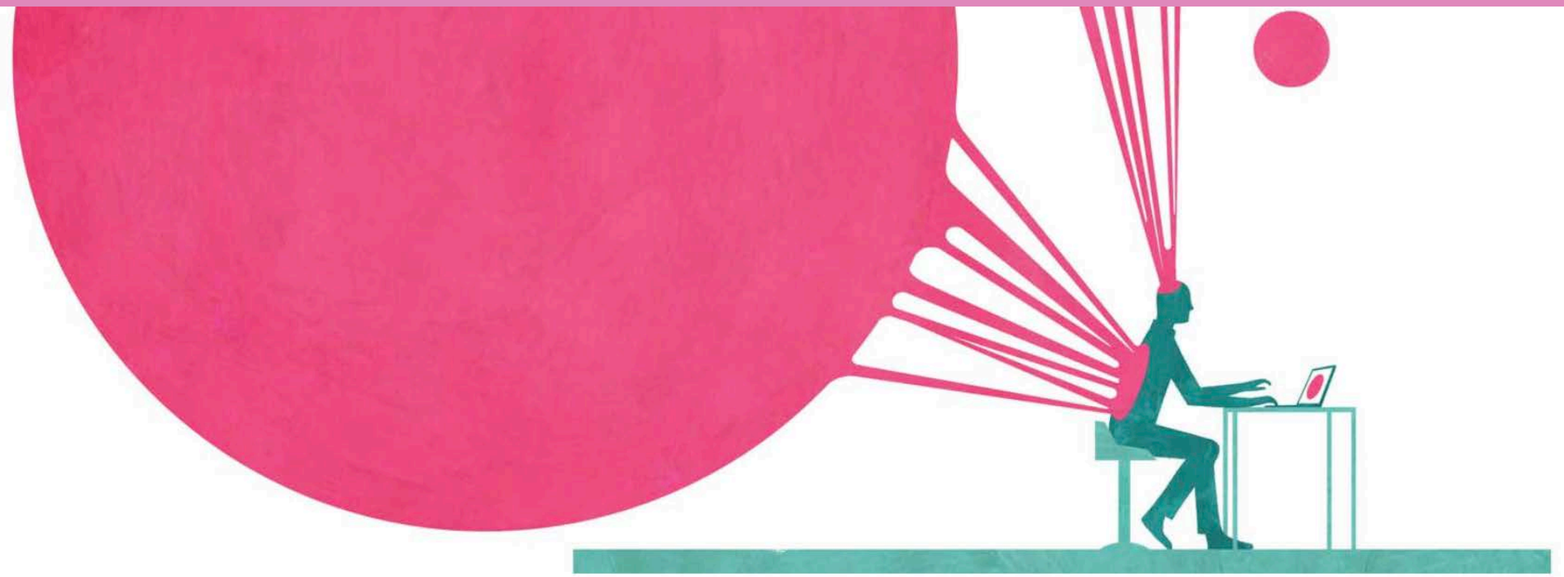
Case Study

Target Coupons

“Misplaced Intelligence”

CASE STUDY

Target Coupons



KYLE T. WEBSTER

MARKETING

Ads That Don't Overstep

by Leslie K. John, Tami Kim, and Kate Barasz

FROM THE JANUARY–FEBRUARY 2018 ISSUE

Loading...

CASE STUDY

Credit Card Fraud

PayThink The cost of fraud fighting is as bad as fraud itself

By
Ryan Breslow

Published
June 07 2018, 12:01am EDT

More in
ISO and agent
Payment fraud
Retailers
Authentication

Print Reprint

Credit card fraud is a massive problem for e-commerce retailers. If you're an online business, you're likely all too familiar with it.

But today, the fight against credit card fraud is a losing battle. Online retailers are losing between $2.48 and $2.82 for every $1 of direct fraud. That means roughly 2.5 times the amount lost to fraud is lost fighting fraud. And merchants will spend 3%-5% of their overall revenue combating fraud operationally.

Today's online payments experience is powered by dozens of unique tools, from payment gateways to fraud detection to checkout. This approach is called "layering," wherein a business layers on suites of different tools to create "robust" payments and fraud detection

CASE STUDY

Credit Card Fraud

PayThink The cost of fraud fighting is as bad as fraud itself

By
Ryan Breslow

Published
June 07 2018, 12:01am EDT

More in
ISO and agent
Payment fraud
Retailers
Authentication

But the results are abysmal, especially when it comes to fighting fraud. Using the best-performing algorithms, researchers successfully identified 495 of 500 fraudulent transactions in a sample size of 50 million transactions (that's a 99% detection rate).

But they incorrectly flagged 500,000 legitimate transactions from good customers as fraudulent. There were too many "false positives" — customers falsely rejected for fraud concerns — to make the layering approach useful. These numbers are untenable for e-commerce stores, where the average profit margin is as low as 5%, or 0.5%-3.5% for e-commerce-only operations.

business layers on suites of different tools to create "robust" payments and fraud detection

Case Study

Credit Card Fraud

“False Positives”

Case Study

Microsoft Tay

MICROSOFT | TECH | TWITTER

Microsoft made a chatbot that tweets like a teen

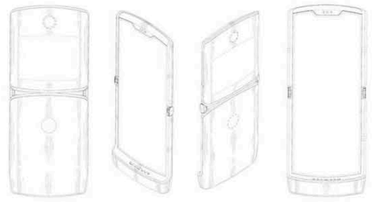
12

By Jacob Kastrenakes | @jake_k | Mar 23, 2016, 10:26am EDT

SHARE

Microsoft is trying to create AI that can pass for a teen. Its research team launched a chatbot this morning called Tay, which is meant to test and improve Microsoft's

MOST READ



Is this the Motorola RAZR with folding display?

Case Study

Microsoft Tay

MICROSOFT | WEB | TL;DR

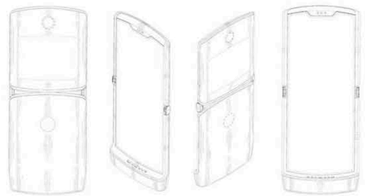
Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

68

By James Vincent | @jjvincent | Mar 24, 2016, 6:43am EDT

SHARE

MOST READ



Case Study

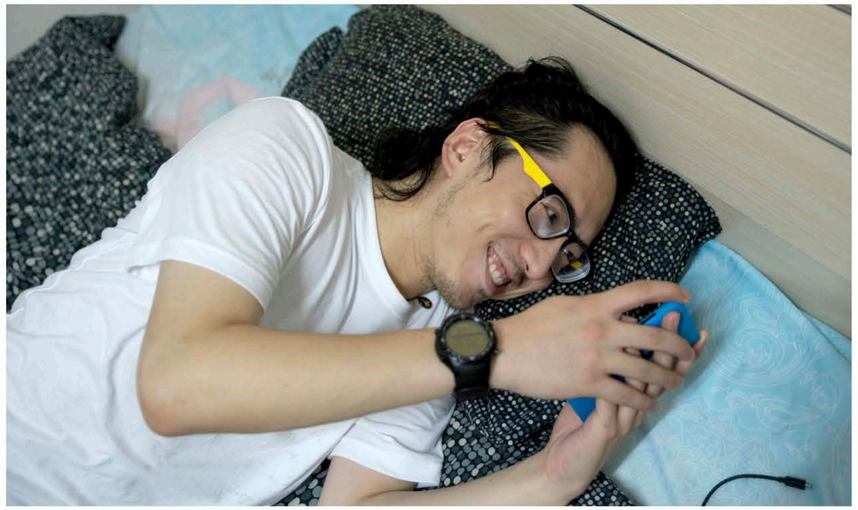
Microsoft Tay

“Adverse Actors”

CASE STUDY

Microsoft Tay

For Sympathetic Ear, More Chinese Turn to Smartphone Program



Zhang Ran, 26, chatted with Xiaoice in Beijing. Millions of young Chinese use the smartphone program. Gilles Sabrie for The New York Times

By John Markoff and Paul Mozur

July 31, 2015

CASE STUDY

Toutiao News App

Business Impact

The Insanely Popular Chinese News App That You've Never Heard Of

Toutiao uses artificial intelligence to curate headline recommendations.

by Will Knight January 26, 2017

CASE STUDY

Toutiao News App

China's king of titillating content Jinri Toutiao issues apology for providing what the people want

PUBLISHED : Thursday, 12 April, 2018, 7:02am
UPDATED : Friday, 14 September, 2018, 7:13pm

Meng Jing

Venture capital investors who called on Jinri Toutiao's chief executive Zhang Yiming

CASE STUDY

Toutiao News App

China's king of titillating content Jinri Toutiao issues apology for providing what the people want

PUBLISHED : Thursday, 12 April, 2018, 7:02am
UPDATED : Friday, 14 September, 2018, 7:13pm

> "Every time we run internal tests the results show that viewership drops off precipitously when alternative content like international current affairs, science and technology is served up", said Zhang, according to investors present at the meeting, speaking on condition their names not be revealed due to the confidential nature of the discussions.

Meng Jing

Venture capital investors who called on Jinri Toutiao's chief executive Zhang Yiming

Case Study

Toutiao News App

"Value Function"

CASE STUDY

Toutiao News App

China's king of titillating content Jinri Toutiao issues apology for providing what the people want

PUBLISHED : Thursday, 12 April, 2018, 7:02am
UPDATED : Friday, 14 September, 2018, 7:13pm

"We let our users down by over emphasising growth and scale over quality and responsibility," Zhang said in an open letter reminiscent of the "self criticism" of "wrongdoers" during the era of Chinese leader Mao Zedong.

Zhang said Toutiao would "correct the flaws" in its algorithms and machine vetting, expand its vetting team from 6,000 to 10,000 people, and permanently ban creators whose content was "against community values".

Meng Jing

Venture capital investors who called on Jinri Toutiao's chief executive Zhang Yiming

Case Study

Predictive Policing

PARTNER CONTENT CYNTHIA RUDIN, MIT SLOAN

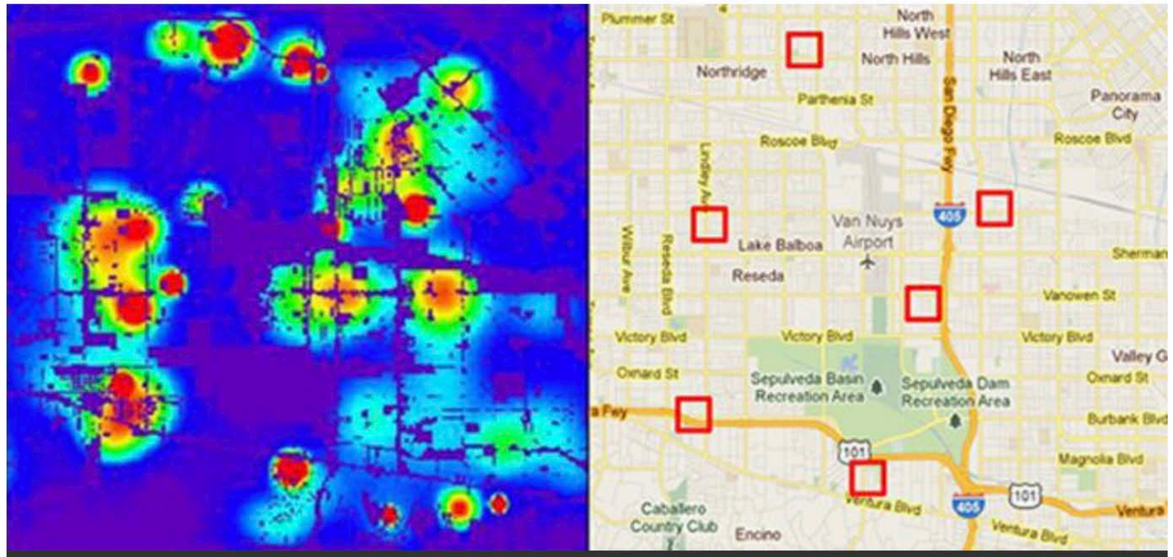
PREDICTIVE POLICING: USING MACHINE LEARNING TO DETECT PATTERNS OF CRIME

CASE STUDY

Predictive Policing

Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?

The software is supposed to make policing more fair and accountable. But critics say it still has a way to go.



Predictive policing is built around algorithms that identify potential crime hotspots.. (PredPol)

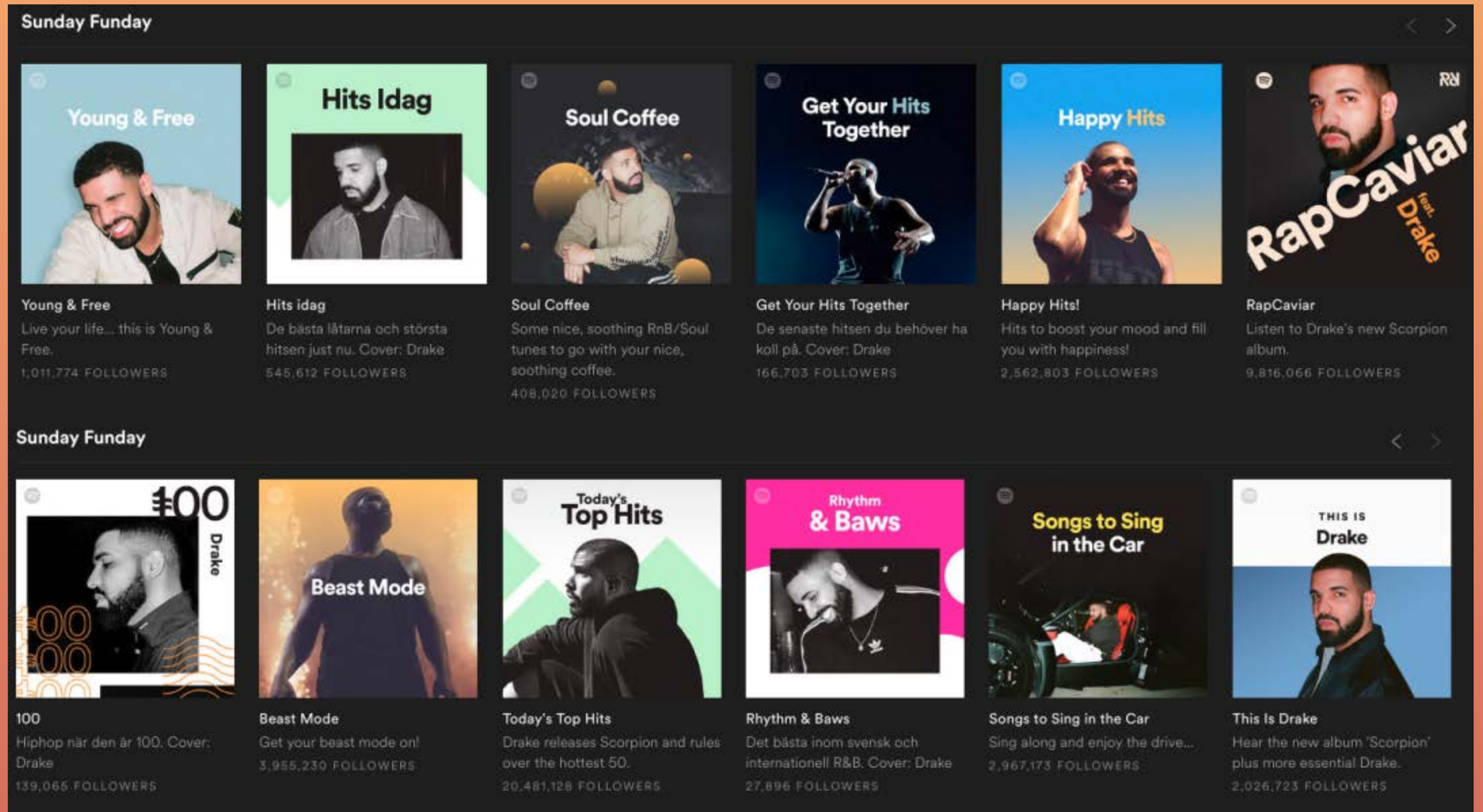
Case Study

Predictive Policing

“Feedback Loops”

CASE STUDY

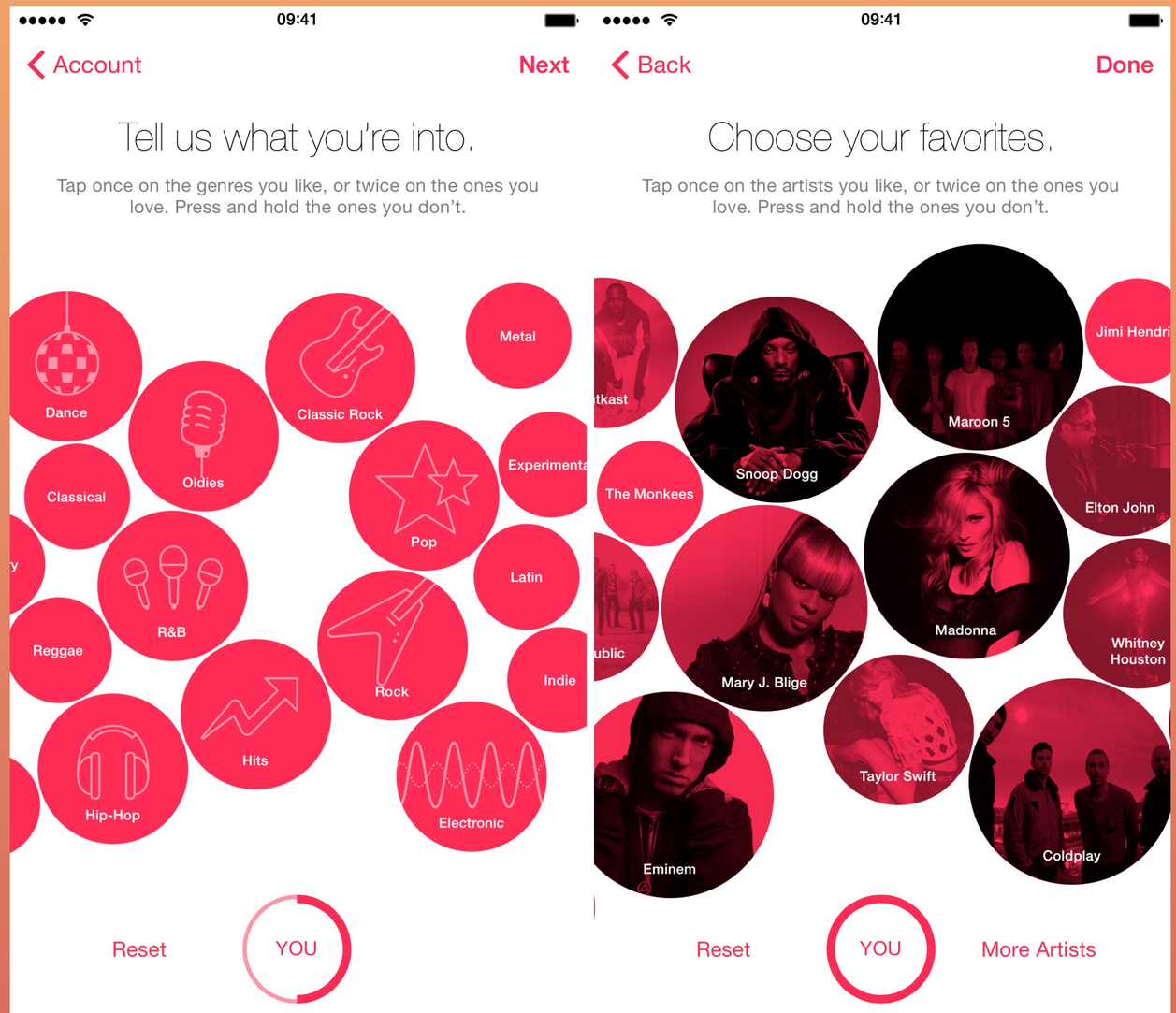
The Harry Potter Problem



Case Study

The Harry Potter Problem

“Cold Start”

CASE STUDY

The Harry Potter Problem