

Creating Value with Big Financial Data

Master Data Science AFI
Elena Alfaro, Head of Data & Advanced Analytics
Client Solutions

Madrid, May 2022



PRESENCIA GLOBAL BBVA

DICIEMBRE 2021



CAPACIDADES DIGITALES

DICIEMBRE 2021

CLIENTES

Clientes digitales
41,8 M

Clientes móviles
39,7 M

VENTAS DIGITALES

Unidades PRV⁽¹⁾
73% 56%

(1) Valor relativo del producto (PRV por sus siglas en inglés) como indicador de la representación económica de las unidades vendidas.

NUESTRO PROPÓSITO

“Poner al alcance de todos **las oportunidades** de esta nueva era”



MAGNITUDES FINANCIERAS

DICIEMBRE 2021

Beneficio atribuido 2021⁽¹⁾

5.069 M€

Total activos

662.885 M€

Créditos a la clientela - bruto

330.055 M€

CET1 FL

12,75%

Depósitos de la clientela

349.761 M€

(1) Excluyendo los impactos no recurrentes (operaciones interrumpidas y costes netos del proceso de restructuración). Resultados reportados incluyendo estos impactos: 4.653 M€.



DESARROLLO SOSTENIBLE Y CONTRIBUCIÓN A LA SOCIEDAD

Doblando compromiso original 2025 BBVA

200 MM€ entre 2018 y 2025

86 MM€

Cana izado hasta Dic. 2021

Destinados a programas sociales 2021

106,3 M€

44,2 M Personas y PYMES

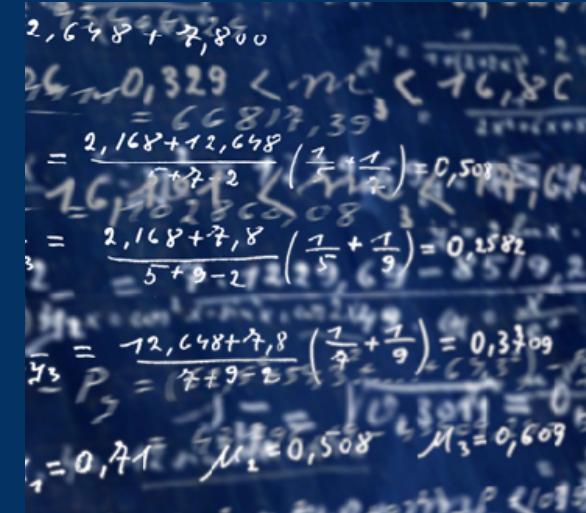
The fintech revolution



Customer
expectations



Fintechs &
Bigtechs



Technology: cloud,
mobile & data

The fintech revolution



**Customer
expectations**

**Fintechs &
Bigtechs**

**Technology: cloud,
mobile & data**

Interaction is changing, in banking and everywhere else

We are here



From human to DIY



From DIY to convenient



From convenient to useful



"I have returned the water bill"

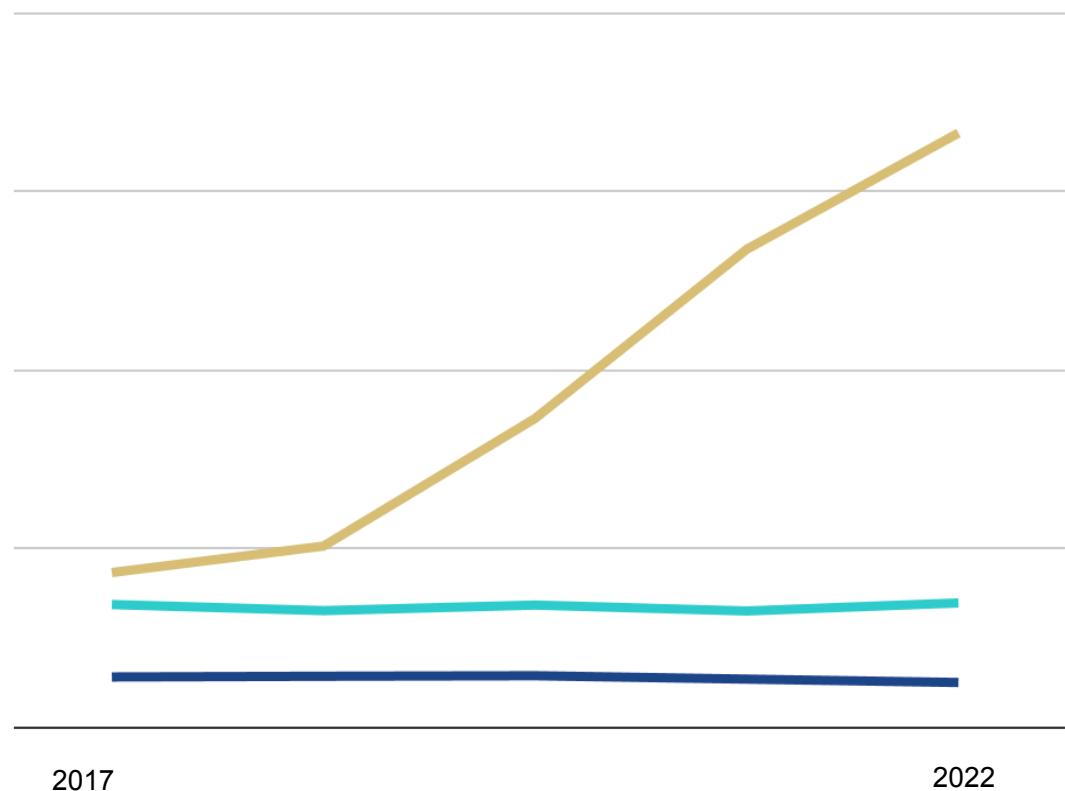
From useful to smart



The experience people expect (spoiler, it is not just free)

- 1 **Convenient** - digital, mobile, 24/7
- 2 **Trustworthy** - transparent, easy, fair
- 3 **Smarter** - helpful, adapted/personalized, embedded

Adoption of mobile surpassing all expectations



— BRANCH
— WEB
— APP



1 x
every 2 days



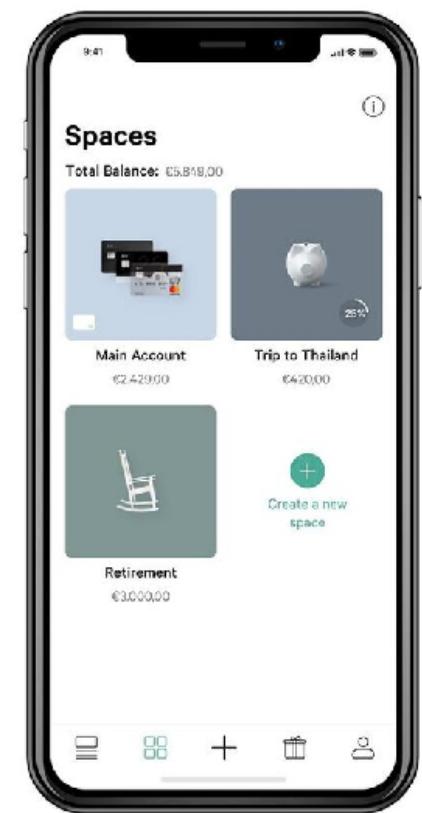
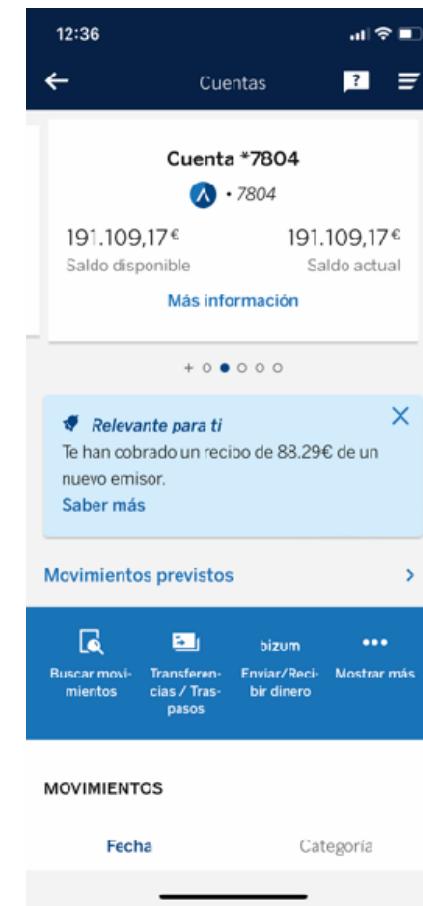
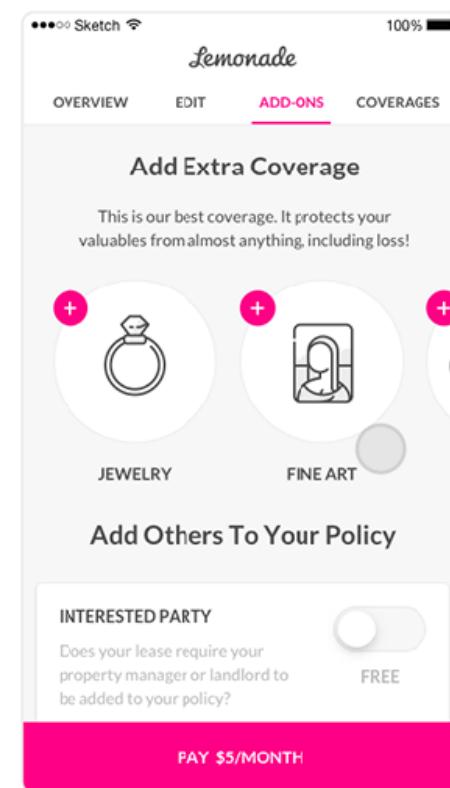
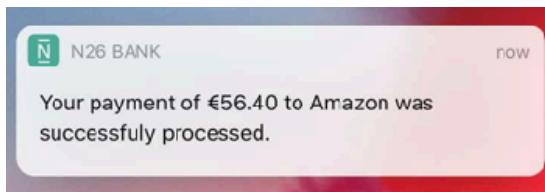
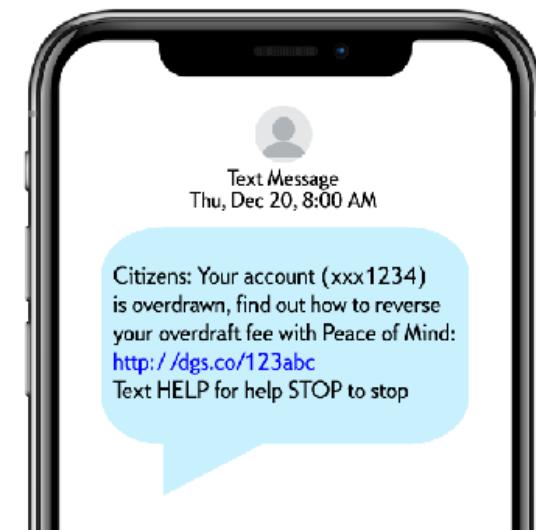
~1 x
per week



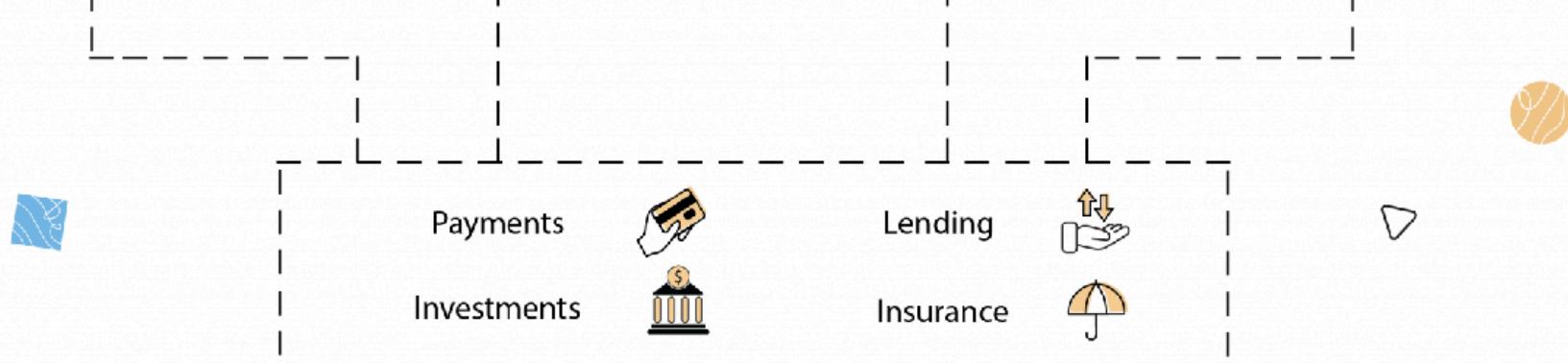
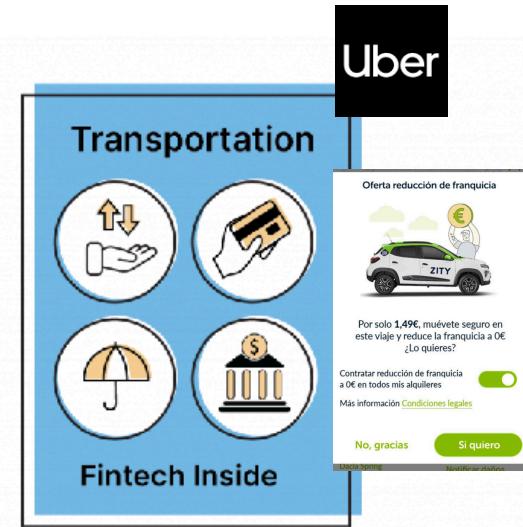
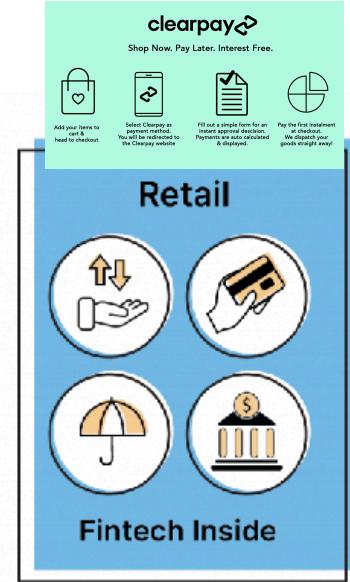
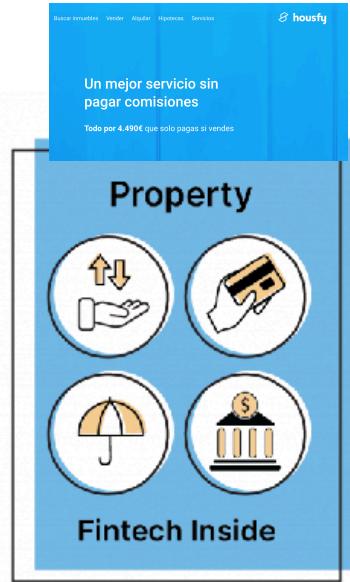
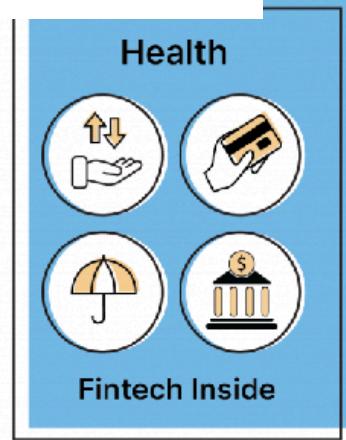
<1 x
per month

@elenalfaro

Easy, fair & transparent



Embedding finance into end goals: invisible money transactions



The fintech revolution

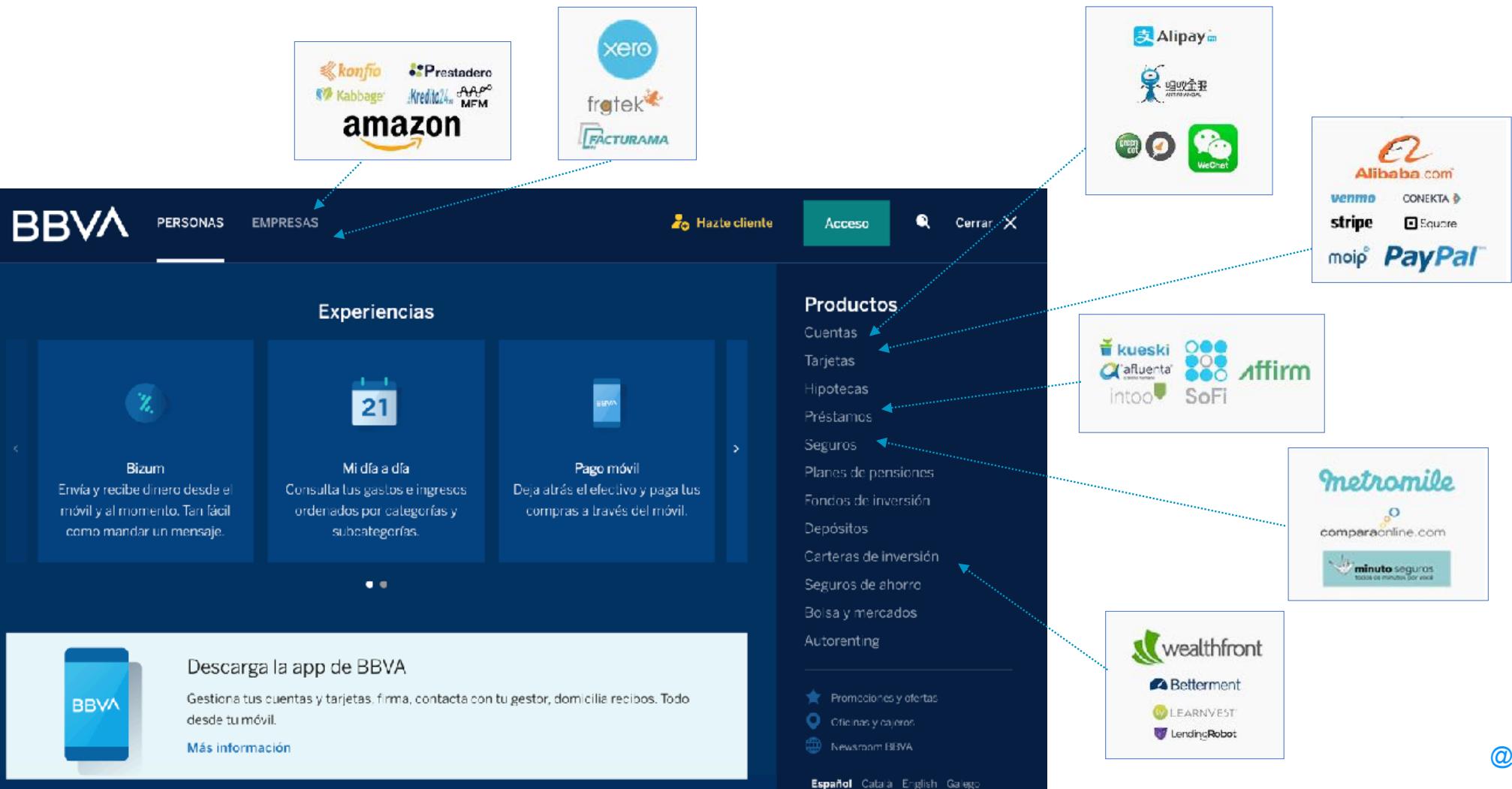


Customer
expectations

Fintechs &
Bigtechs

Technology: cloud,
mobile & data

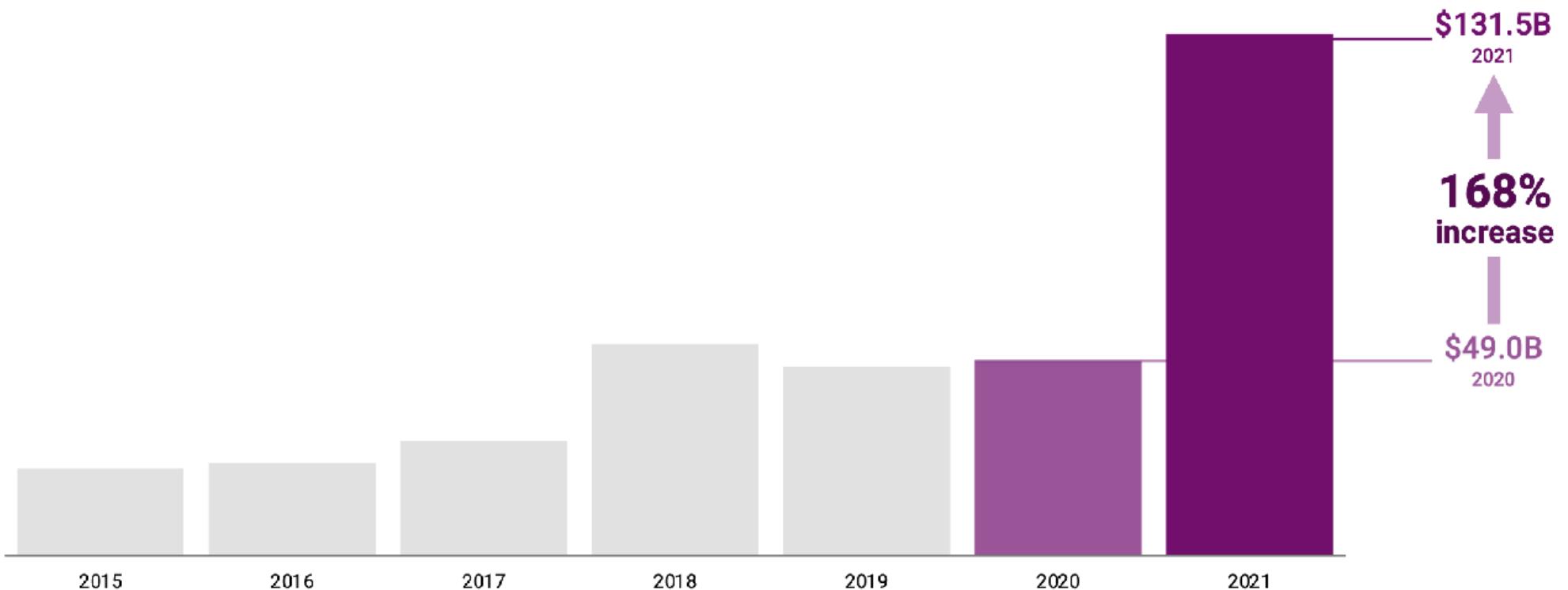
From unbundling banks, to collaboration and to innovation



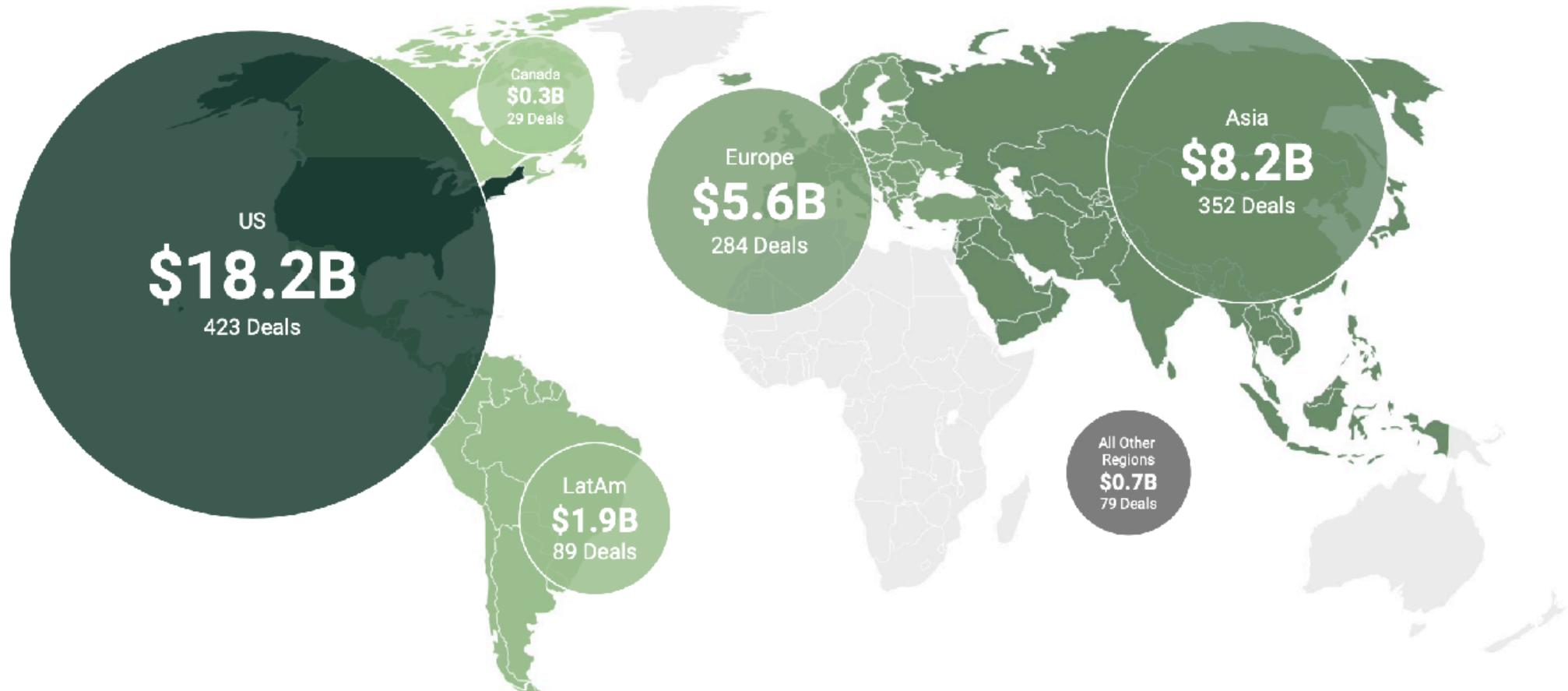
@elenalfaro

STATE OF FINTECH | GLOBAL TRENDS | INVESTMENT TRENDS

Global fintech funding more than doubles to reach record \$132B in 2021

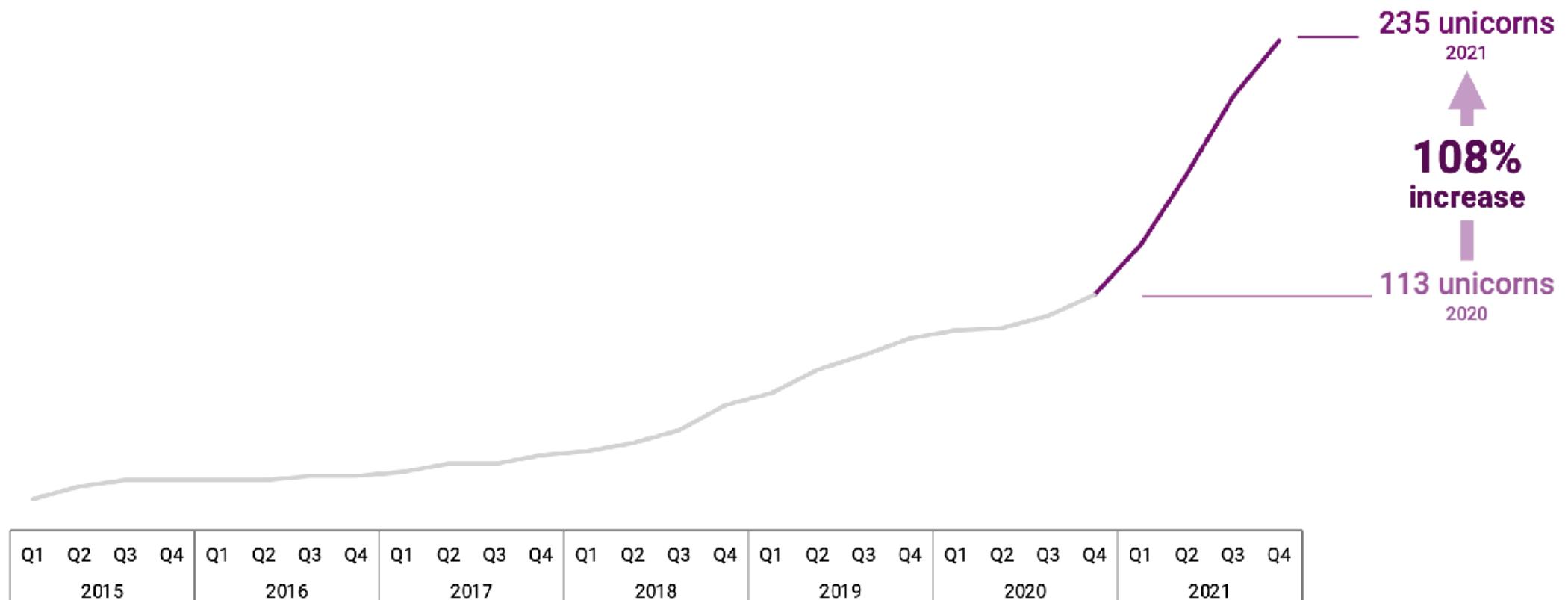


US leads in Q4'21 deals and funding, followed by Asia and Europe



STATE OF FINTECH | GLOBAL TRENDS | UNICORNS & VALUATIONS

157 new unicorn births in 2021 bring the herd to 235, a 108% jump YoY



Global: Top 10 unicorns in Q4'21 by valuation

	Company	Latest Valuation	Country	
1	Stripe	\$95.0B	United States	Payments
2	Klarna	\$45.6B	Sweden	Payments/BNPL
3	Revolut	\$33.0B	United Kingdom	Bank
4	Chime	\$25.0B	United States	Bank
4	FTX	\$25.0B	Bahamas	Blockchain/exchange
6	Checkout.com	\$15.0B	United Kingdom	E-commerce
7	Plaid	\$13.5B	United States	Open Banking
8	Devoted Health	\$12.6B	United States	Insurtech
9	Brex	\$12.3B	United States	Bank
10	Digital Currency Group	\$10.0B	United States	Blockchain/build
10	Gusto	\$10.0B	United States	HR/Payroll
10	Ripple	\$10.0B	United States	Blockchain

270B\$

HSBC + Sverbank + BNP + UBS + Santander = 360B\$

Source: CB Insights

Types of fintechs (non exhaustive)

Payments: Stripe, Klarna, Afterpay

Blockchain: Coinbase, Ripple, Binance

Lending: Kabagge, Sofi

Insurtech: Devoted Health, Lemonade

Investment: RobinHood, eToro

Real State: Opendoor

Banking tech: Mambu, Temenos

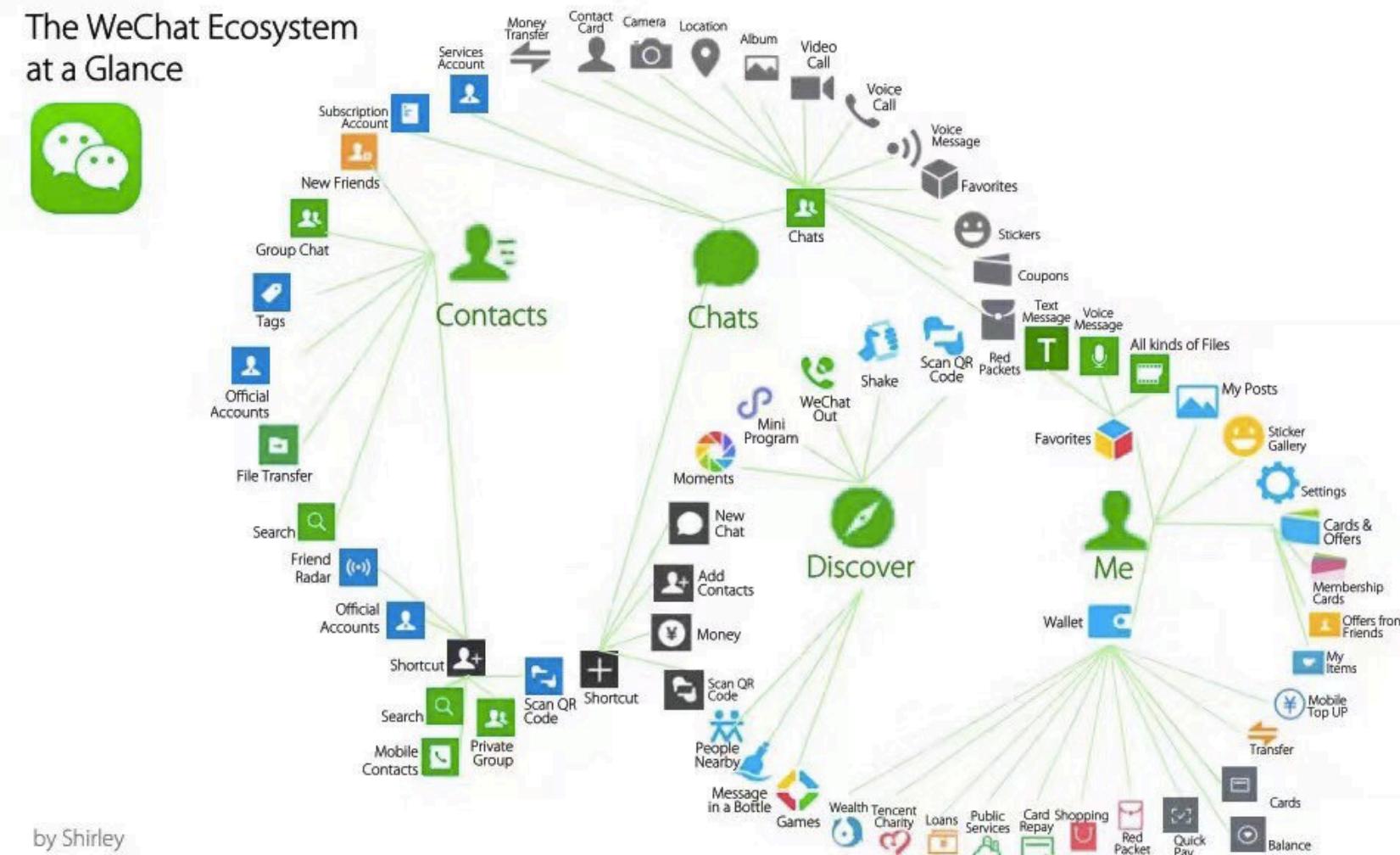
Regtech: Verafin, Carta

Neobanks: N26, Revolut

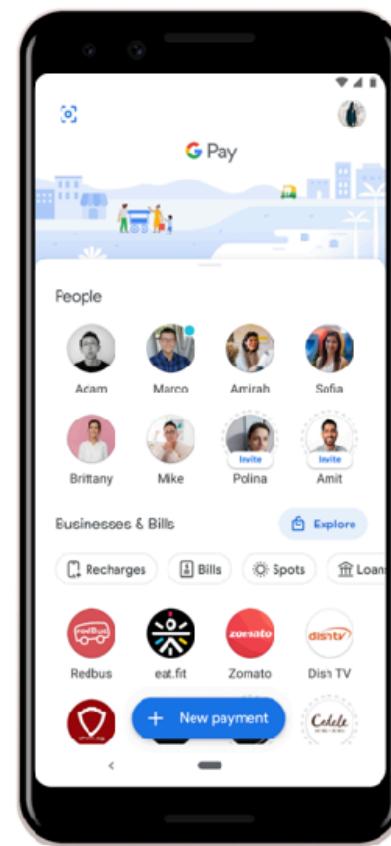
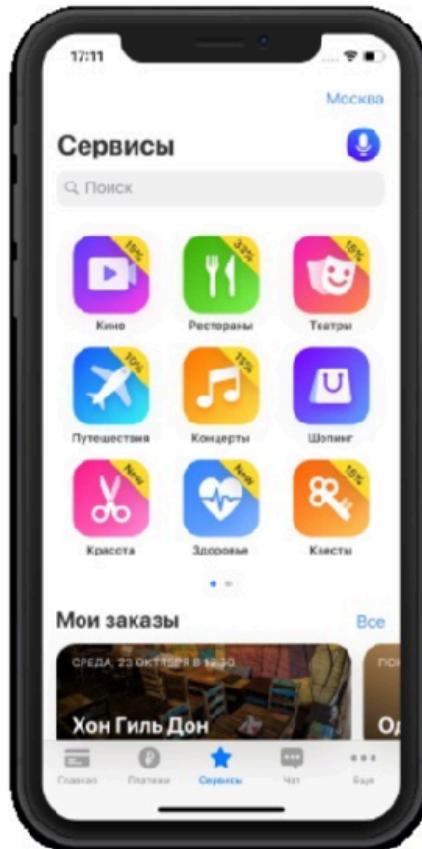
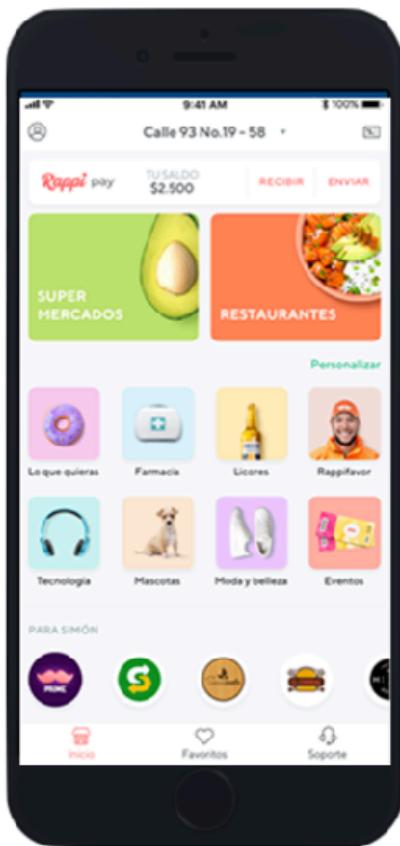
BigTechs blurring the lines



BigTechs blurring the lines



And many others following: the rise of superapps



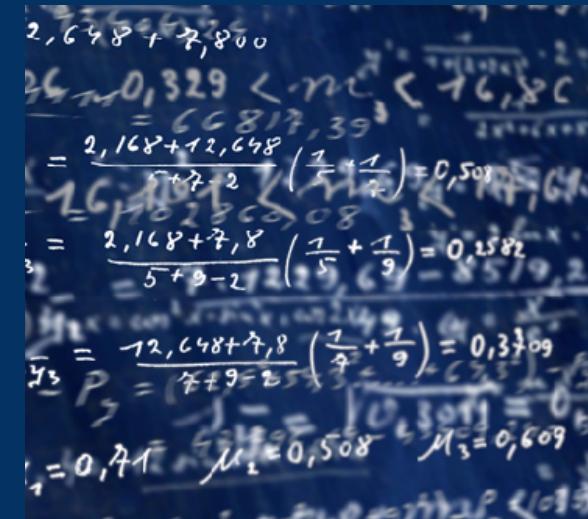
The fintech revolution



Customer
expectations



Fintechs &
Bigtechs



Technology: cloud,
mobile & data



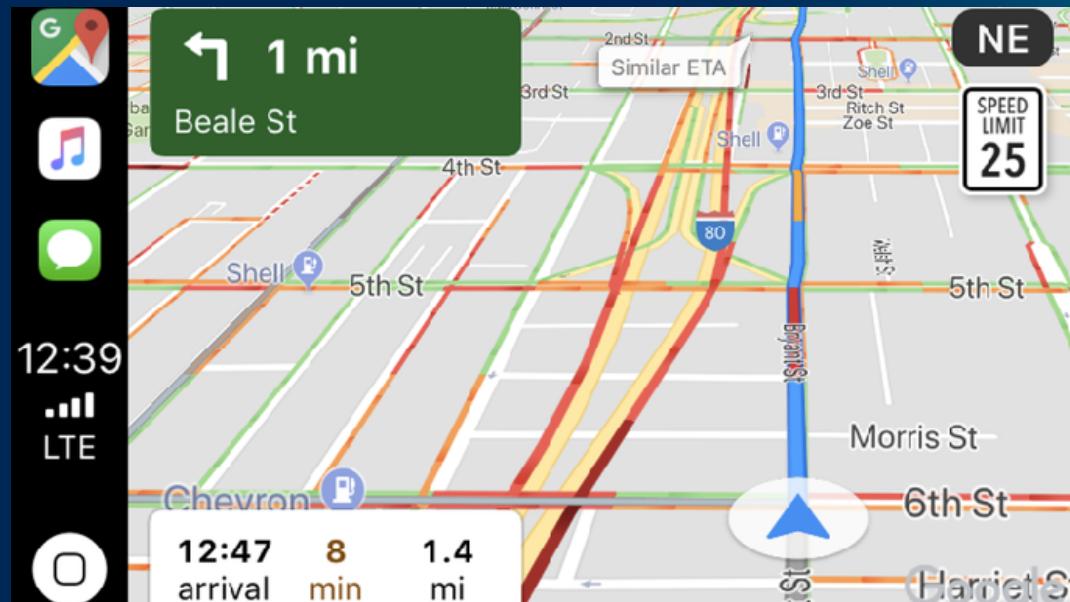
BBVA

Creando Oportunidades

To be data driven: Support decision making or eliminate it

BBVA

Creando Oportunidades



@elenalfaro

Data-driven systems already a part of our daily lives...



Data
Talent
Good questions

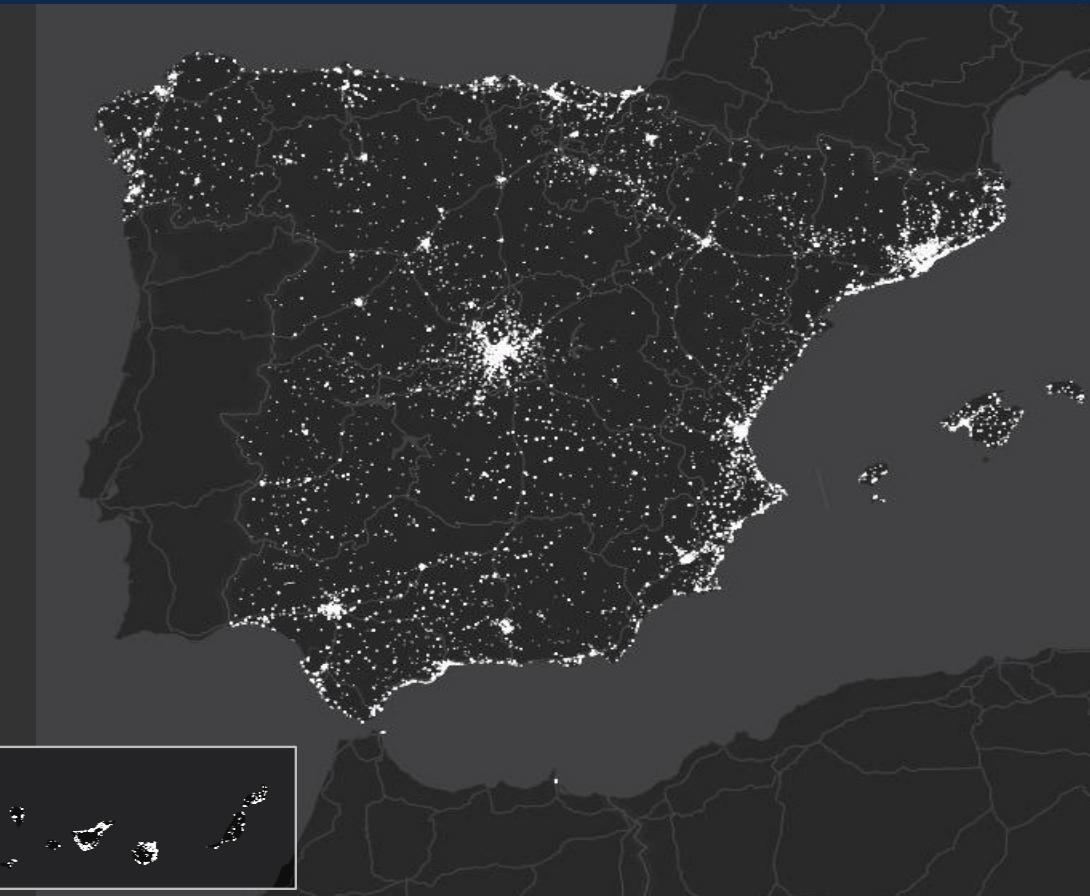
Data
Talent
Good questions

DIGITAL

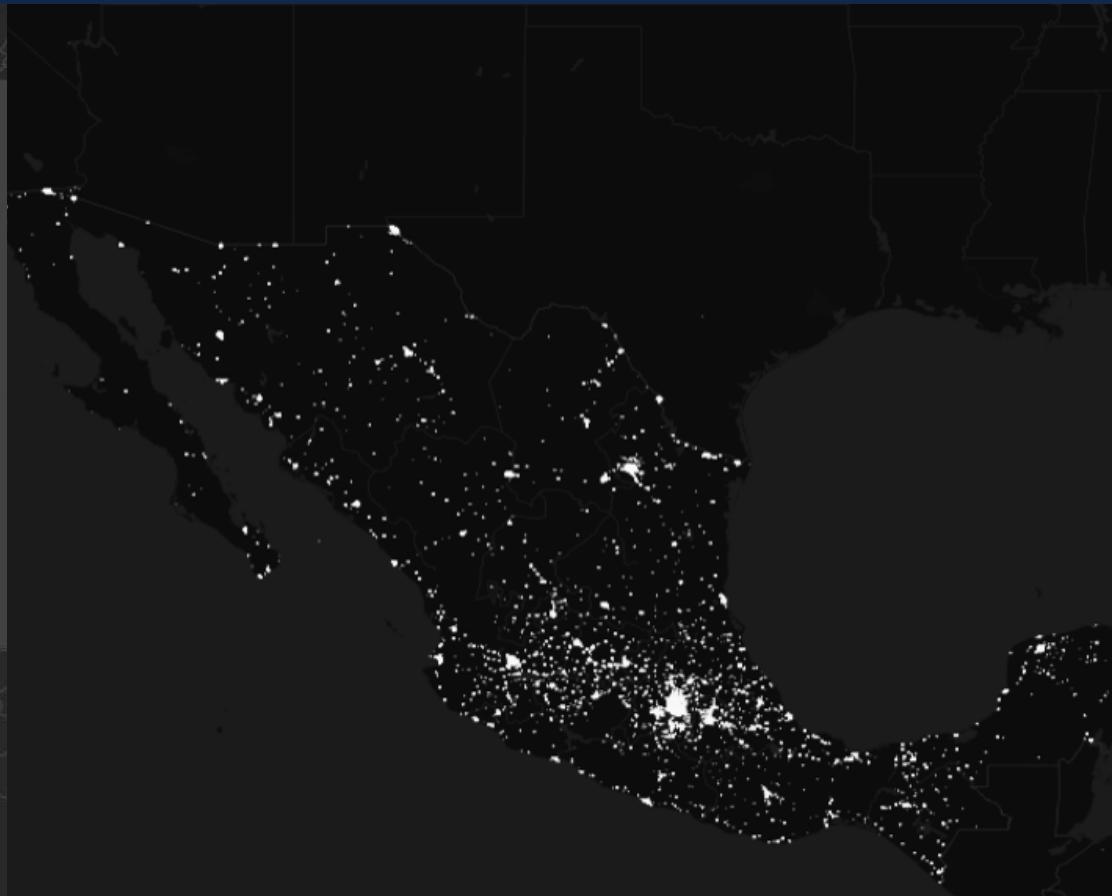


PHYSICAL

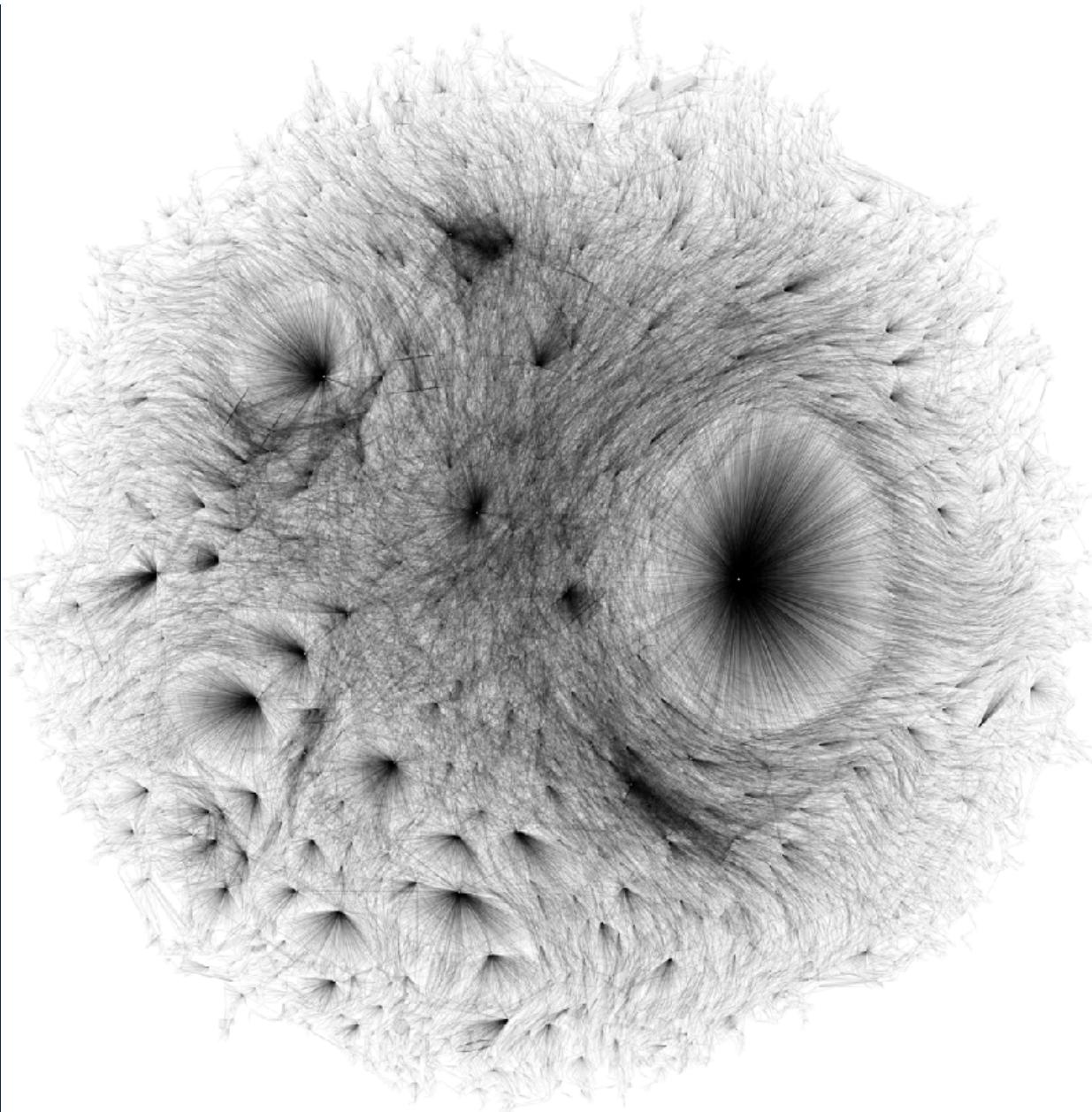
Financial Data



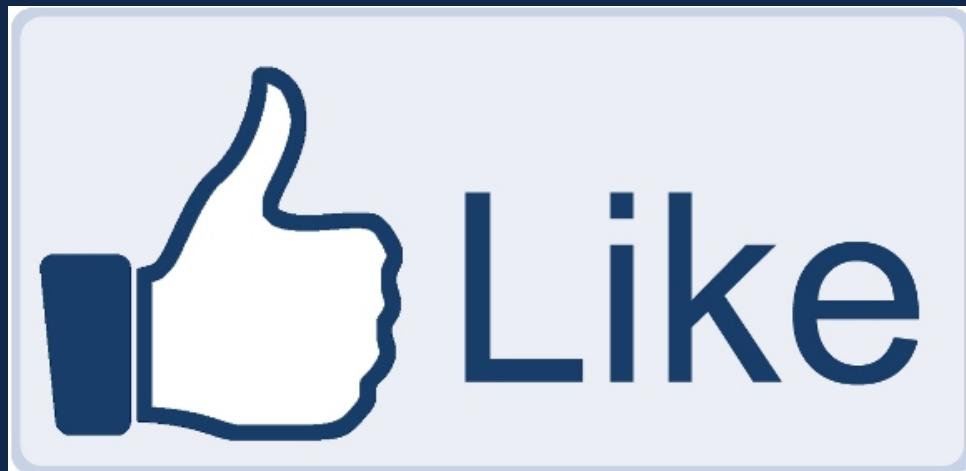
900M transacciones de tarjeta en **1,2M** TPVs, realizadas por **60M** de personas, representando **€47.000M**



1.800M transacciones de tarjeta en **1,1M** TPVs, realizadas por **100M** personas, representando **€42.000M**



What data can tell matters



vs



intention

vs

evidence

How we can use data (simplified)

**Anonymous /
Aggregated**

Company Data



**Personal /
Private**



External Data



facebook Search

Mark Zuckerberg
Founder and CEO at Facebook | Si Palo Alto, California | Knows English, Born on May 14, 1984

Wall

RECENT ACTIVITY
Mark subscribed to upr York Times · Subscrib

A screenshot of a Facebook profile page for Mark Zuckerberg. The page shows his name, bio, and recent activity. It includes a profile picture of him smiling.

Data
Talent
Good questions

DATA SCIENTIST MUST-HAVE SKILLS

MATH & STATISTICS

- Machine Learning
- Statistical Modeling
- Exploratory Analysis
- Clustering
- Regression Analysis

DOMAIN KNOWLEDGE & SOFT SKILLS

- Inclination towards business operations
- Keen on working with data
- Problem solver
- Strategic, proactive, and cooperative
- Interested in hacking



PROGRAMMING & DATABASE

- Computer Science Fundamentals
- Database Management System
- Data Visualization
- Python
- Big Data

COMMUNICATION & VISUALIZATION

- Storytelling skills
- Convert data-based insights into decisions
- Collaborative with Sr. Management
- Knowledge of tools like Tableau
- Visual art design



Mat Velloso
@matvelloso



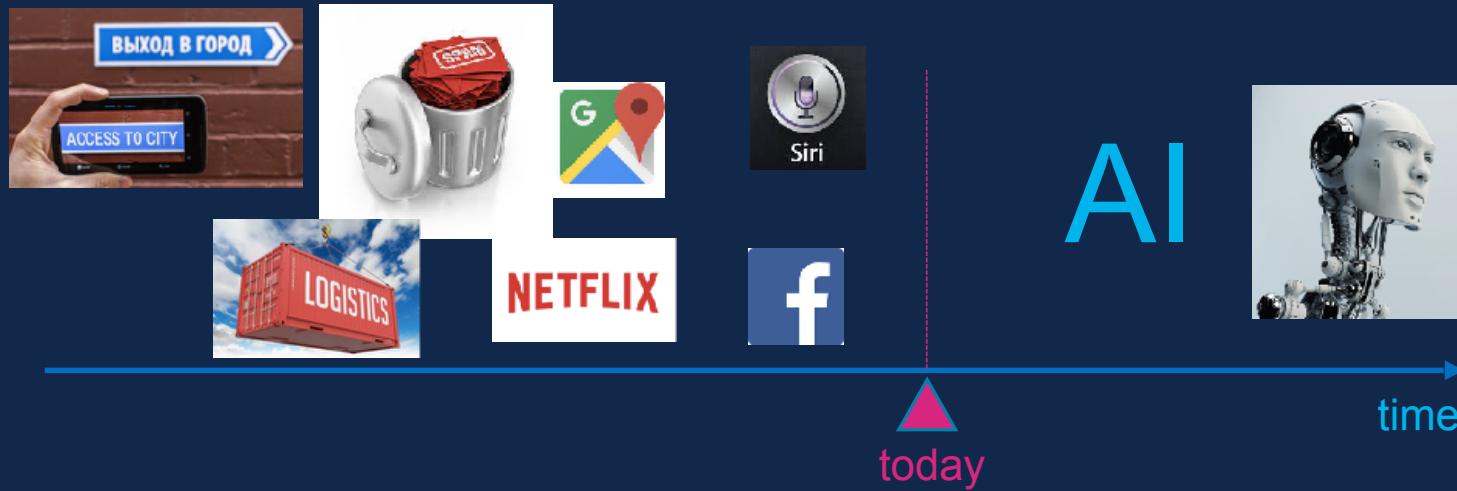
Difference between machine learning
and AI:

If it is written in Python, it's probably
machine learning

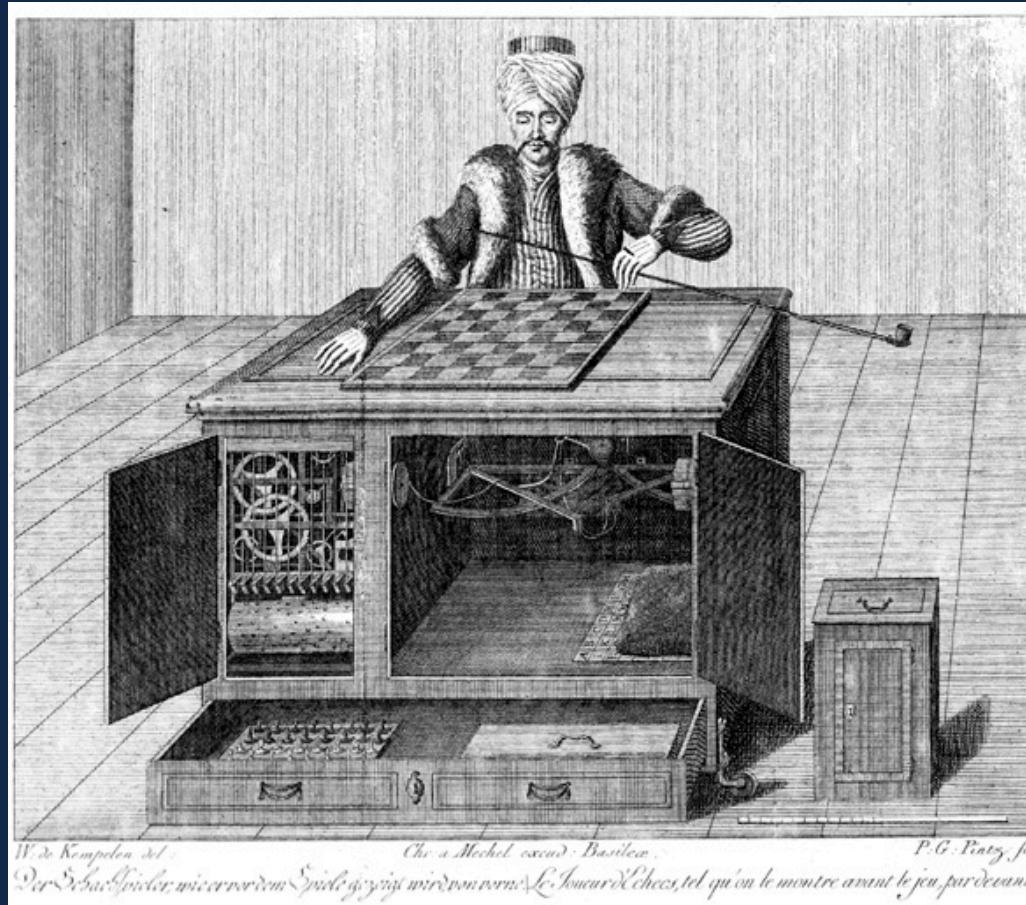
If it is written in PowerPoint, it's
probably AI

What is Artificial Intelligence

“the theory and development of computer systems able to **perform tasks normally requiring human intelligence**, such as visual perception, speech recognition, decision-making, and translation between languages”

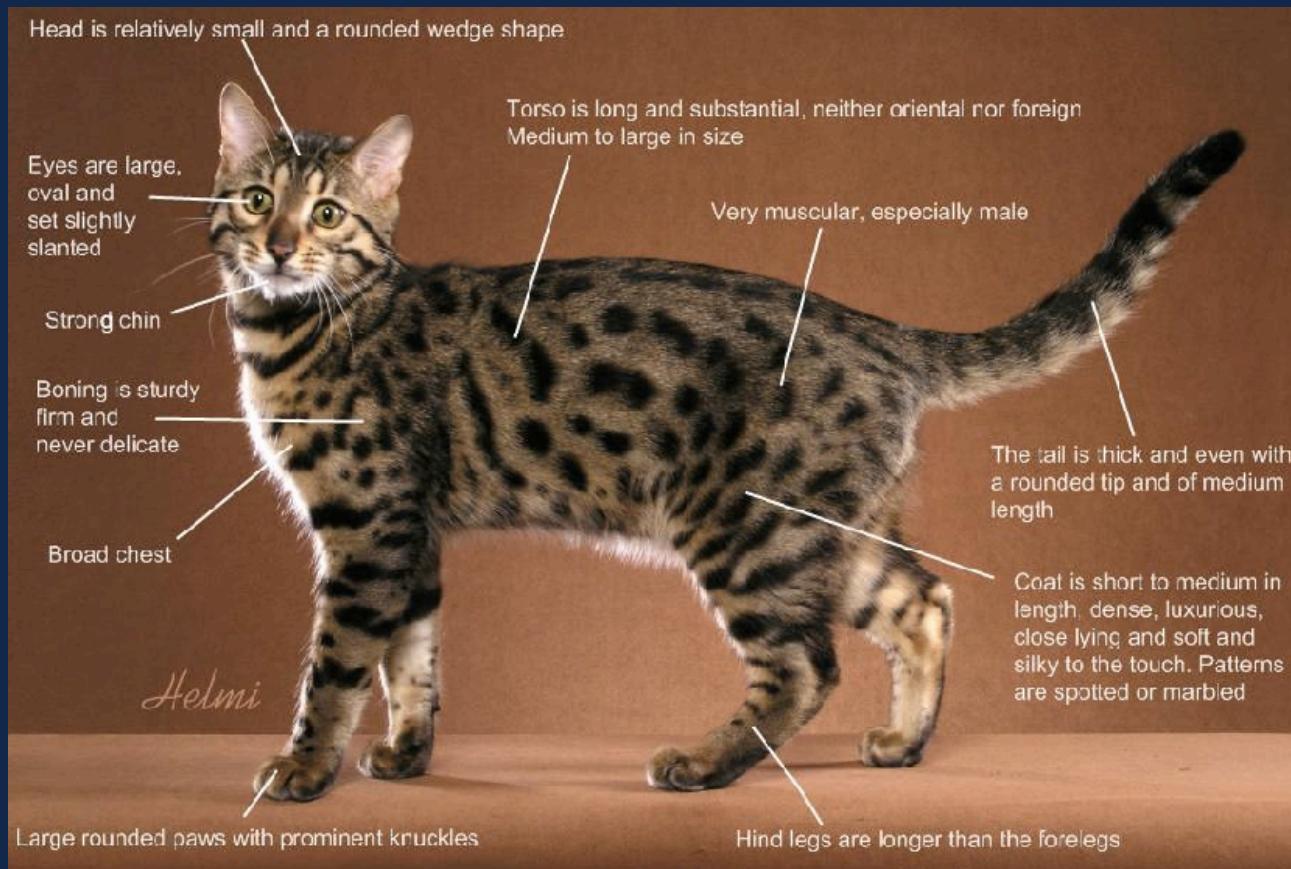


Definición de Inteligencia Artificial, por John McCarthy en 1956: “La ciencia e ingenio de hacer máquinas inteligentes”

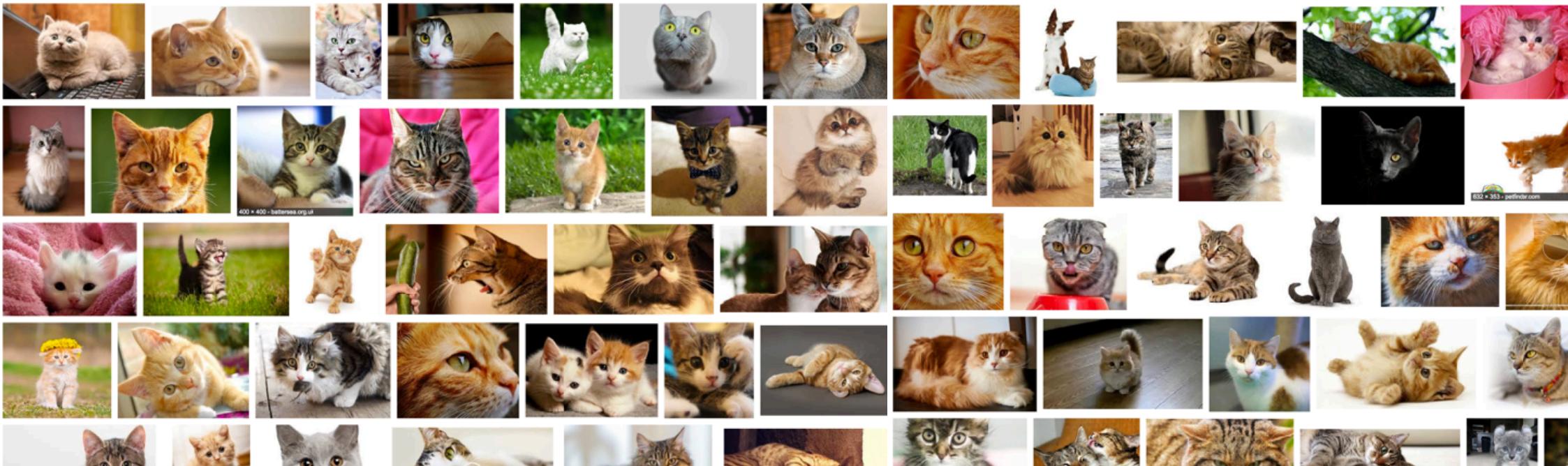
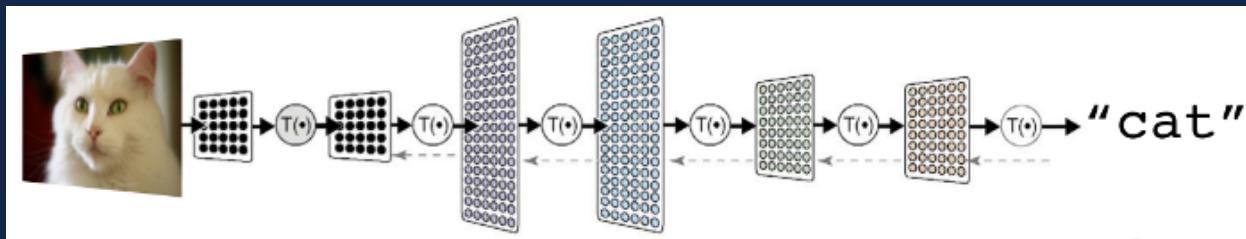


Autómata El Turco, construido por Wolfgang von Kempelen (Hungría 1734-1803) en 1769, fascinó a los europeos durante más de 80 años ganando a jugadores y famosos como Napoleón y Benjamin Franklin

From programming rules to programming with examples



From programming rules to programming with examples



Today, AI is created through learning from data



BUT WITH THE RIGHT COMPETENCES

- Curiosity, creativity & “explorative mind”
- Scientific mindset – data skeptic (cause vs correlation mistake)
- Focus on value & the right questions
- Empathy & communication
- Be aware of “data superpowers” and potential damages (privacy, discrimination, dehumanization, manipulation...)

Data
Talent
Good questions

AI is used across Google products



Assistant
Query understanding
Conversation



Search
Query understanding
Search ranking



YouTube
Video recommendations
Better thumbnails



Cloud
Cloud ML APIs
TPU



Home
Speech recognition
Conversation



Translate
text and speech
translations



Clips
smart image capture



Gmail
Smart Reply
Spam classification



Photos
Photos / video search
Auto-smile montage



Android
Keyboard input
(also in iOS)



Cardboard
Image stitching



Drive
Quick Access



Maps
Street View images
Parsing Local Search



Ads
Richer Text Ads
Video summarization



Play
App clustering
Music recommendations

The main tasks of Machine Learning

	Generic	In digital companies	In financial services
CLASSIFY	Putting similar things together?	Amazon's customers-also-bought	Classify customers on their probability of churn/default
COMPARE	Figuring out what kind of thing something is?	Gmail spam/not spam and FB photos (detecting faces)	Compare companies to infer financial situation and offer products
DETECT	Finding uncommon things?	Most “trending” products (Foursquare, Twitter, Facebook)	Emerging clusters of customers
OPTIMIZE	What is the right price for a product and a customer?	Amazon dynamic pricing engines	Optimize advertising budget or product pricing
PREDICT	Predicting a numerical value of a thing?	Kayak predicting how much a flight will cost in two hours	Predict future transactions/cash flows in a client account and send alerts
RECOMMEND	Giving users things they may be interested in, without them explicitly searching?	Netflix, Spotify, who-to-follow on Twitter	Recommend an investment product depending on clients saving goals
SCORE	Helping users find the right thing when they search?	Google, Bing, Twitter search	Credit scoring

Data/ML application @BBVA

Above the glass

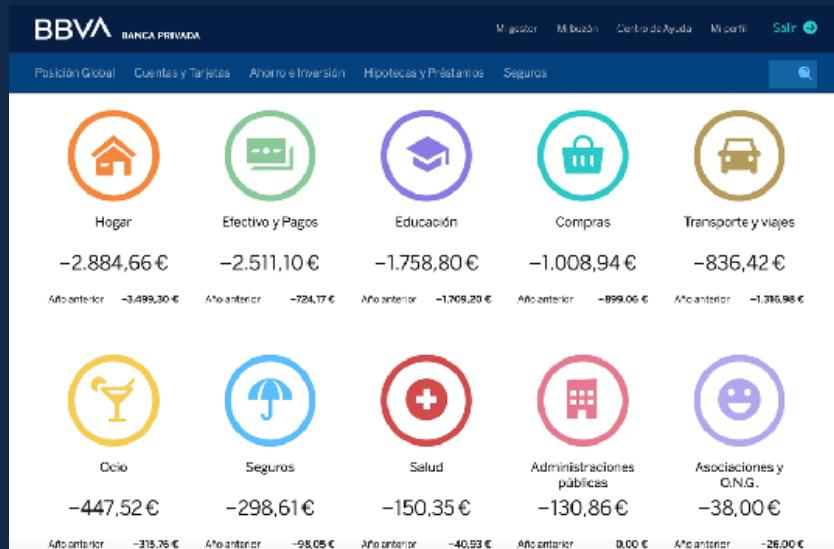
**Experience, Personalization,
Scale, Relevance**



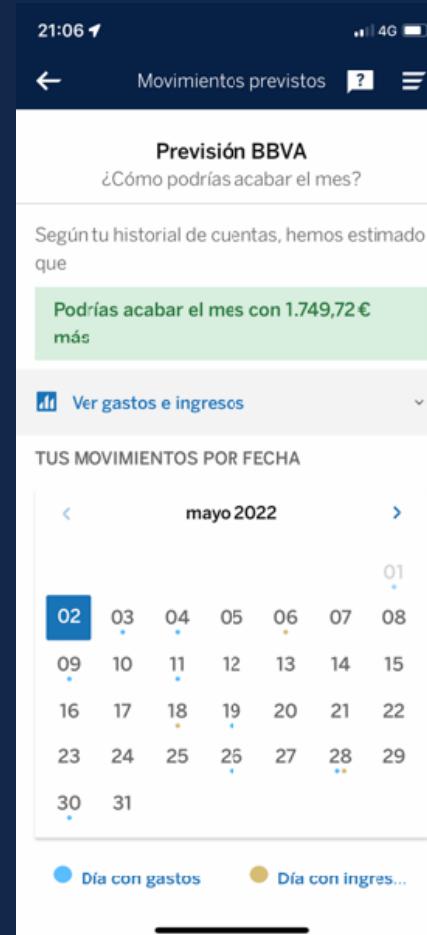
Below the glass

**Automation, Efficiency, Internal
decision making**

PFM/Advice: from information to action and financial wellbeing



Categorization



Prediction

Automatic Transaction Categorizer: How does it work?

Movimiento de cuenta

Código	Cantidad
232	- 60.30 €

Fecha
23/10/2015

Match nombre
compañía

Match CIF

Texto

RECIBO GAS NATURAL CIF C456XXXX

Otros campos
...

Taxonomía con
reglas

Clasificación de
texto
(machine learning)

Otras reglas



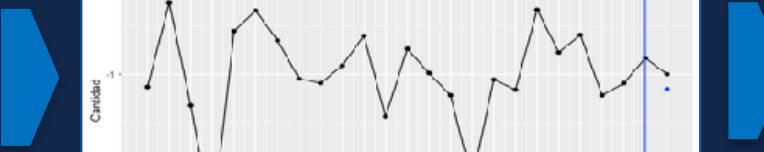
+7 mil transactions processed per day

Predictive Calendar, how does it work?

Operaciones de un cliente

04/07/16	REPSOL	56,72
05/07/16	RESTAURANT FORN	10,00
05/07/16	RETIRADA EFECT.	40,00
06/07/16	PAGO COMUNIDAD	30,00
...		

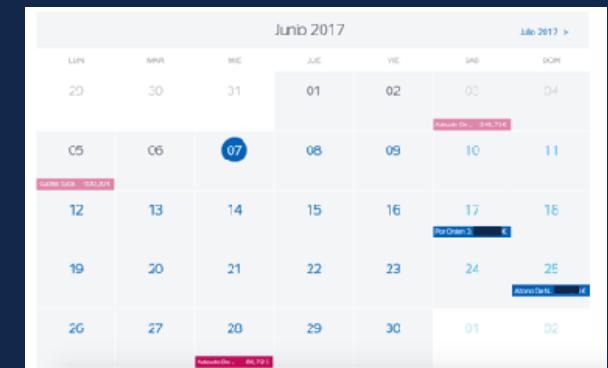
(1600 mill. de transacciones como input)



>100M de series analizadas
cada mes



(totales fin de mes)



(movimientos con fecha)

Financial Health & Smart Events



Creando Oportunidades

The image displays three screenshots of the BBVA mobile application interface, illustrating various financial health and smart event features.

Screenshot 1: Destacados para ti (8:27)

This screen shows a notification for an "Ingreso no habitual recibido" (Unusual income received) of 1,784,45 €. It includes a call-to-action button "Aportar a fondo". Below the main message, there's a section asking if the user wants to invest part of their money, with a "Saber más" link.

Screenshot 2: Cuentas (9:20)

This screen displays a list of transactions from February 16, 2021, including a transferencia recibida (Received transfer) of -6,30 € and a transferencia realizada (Made transfer) of -217,80 €. A blue callout box highlights a "Relevante para ti" (Relevant for you) message about a recent unusual payment of 217,80 €.

Screenshot 3: Destacados para ti (8:47)

This screen shows a forecast for the account balance by the end of the month. It indicates a "Saldo previsto" (Expected balance) of 551,25 €. A blue callout box asks if the user wants to invest part of their money, with a "Saber más" link and a "Aportar a fondo" button.

Supplies, subscriptions & brands

16:30 🔍

Suscripciones y Suministros

Suscripciones	Suministros
6.148,36 € Gasto medio anual estimado	512,37 € Gasto medio mensual estimado
 IBERDROLA 2.716,13 €/año aprox	
 AQUALIA 257,28 €/año aprox	
 CANAL DE ISABEL II 1.115,17 €/año aprox	
 NATURGY 234,78 €/año aprox	
 TELEFONICA 962,99 €/año aprox	
 TELEFONICA 337,62 €/año aprox	
 ENDESA 524,39 €/año aprox	

16:30 🔍

Suministros

 **Telefonica**
962,99 €/año aprox

[Ver histórico](#)

Compárate

Gasto octubre 2020	Gasto octubre 2021
68,80€	105,20€

ÚLTIMOS MOVIMIENTOS

04/10/2021
ADEUDO DE TELEFONIA FIJA -105,20€
ADEUDO DE TELEFONIA FIJA

06/09/2021
ADEUDO DE TELEFONIA FIJA -120,39€
ADEUDO DE TELEFONIA FIJA

04/08/2021
ADEUDO DE TELEFONIA FIJA -120,92€

16:29 🔍

Suscripciones y Suministros

Suscripciones	Suministros
503,88 € Gasto medio anual estimado	41,99 € Gasto medio mensual estimado
Simular ahorro anual <input checked="" type="checkbox"/>	
 SPOTIFY 191,88 €/año aprox	
 UNICEF 312,00 €/año aprox	

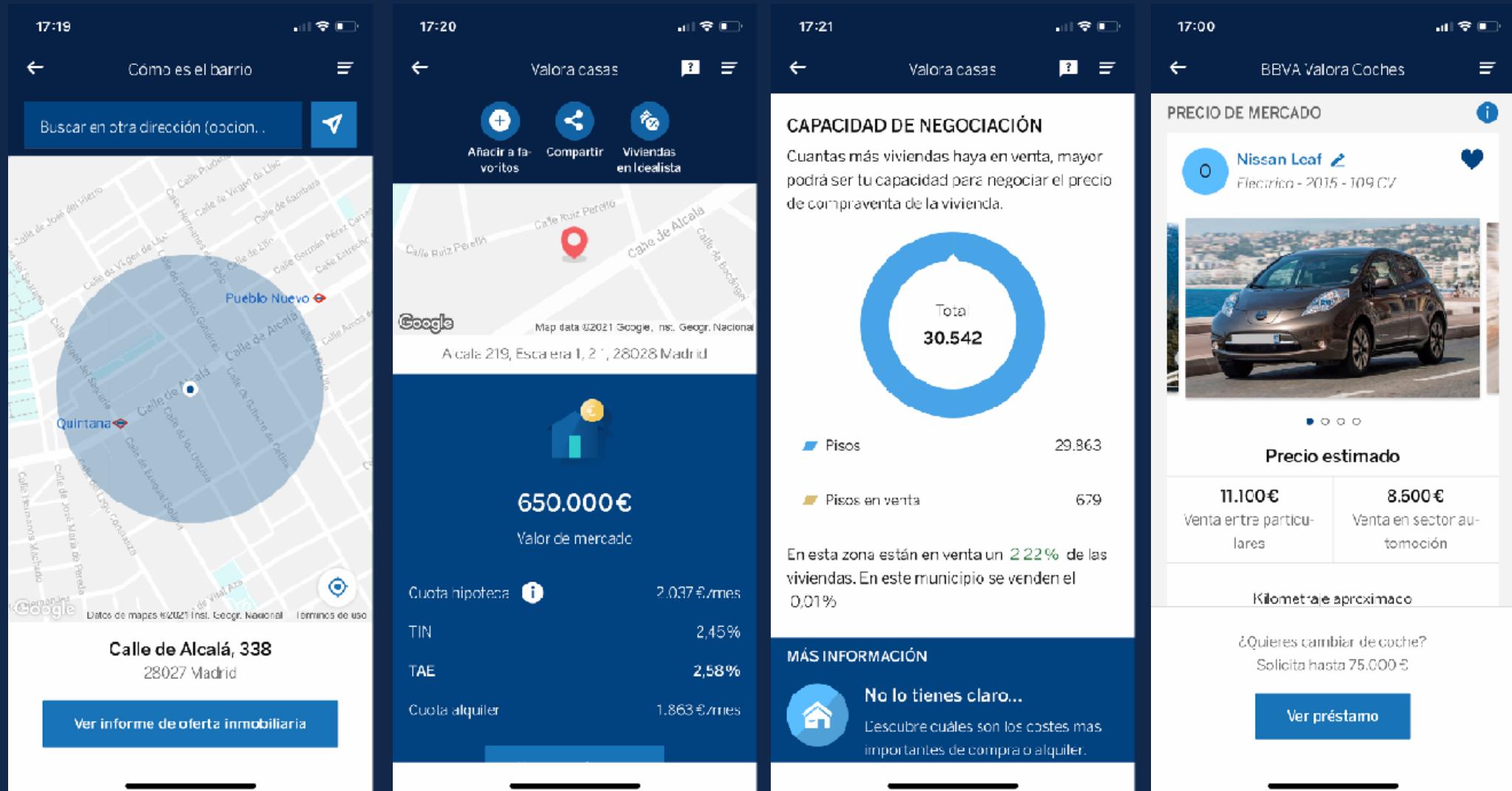
21:09 🌙

Gasto total

Mis presupuestos **Predicción de gasto**

Categorías	Marcas
01 feb. - 28 feb.	
 Mercadona 8 movimientos	-695,20€
 Amazon 7 movimientos	-226,82€
 Allianz 1 movimiento	-194,18€
 Telefonica 3 movimientos	-142,71€
 Nike 1 movimiento	69,99€
 Securitas Direct 1 movimiento	-68,57€
 Uber 2 movimientos	-35,18€

BBVA Valora

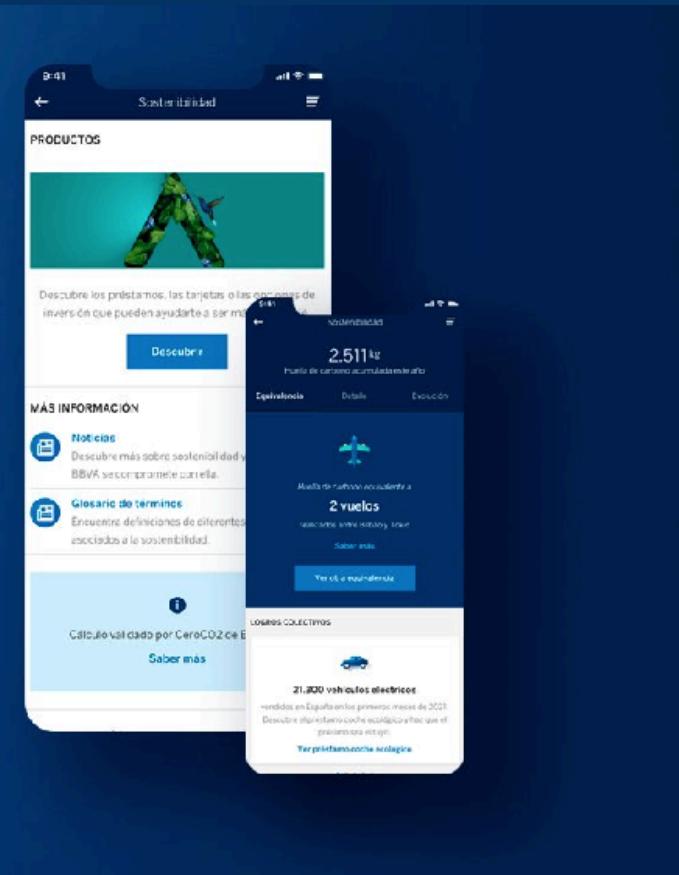


@elenalfaro

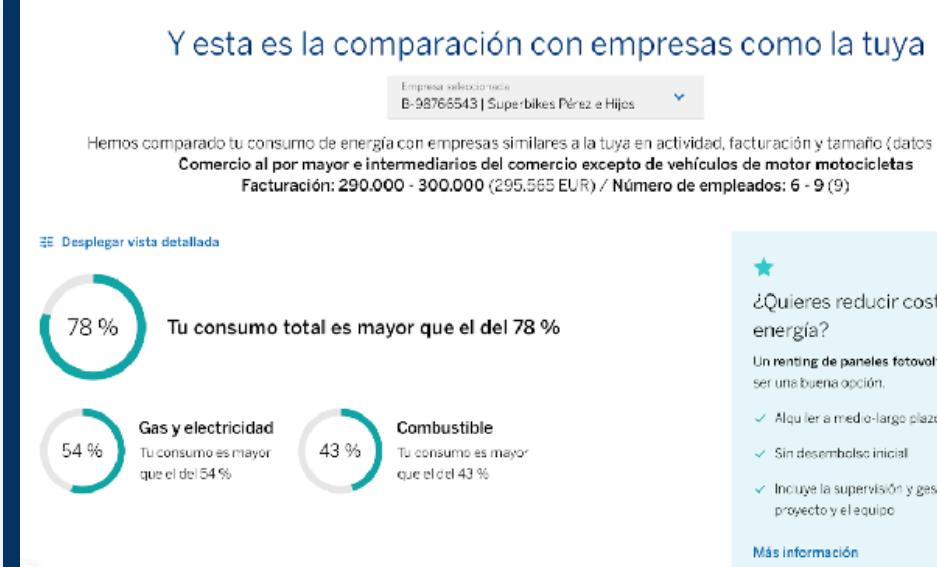
Carbon footprint for retail clients

Nueva calculadora de Huella de Carbono

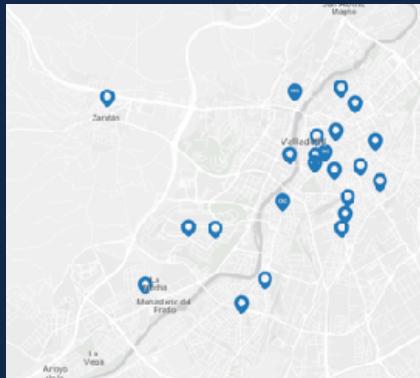
en la #app de BBVA.



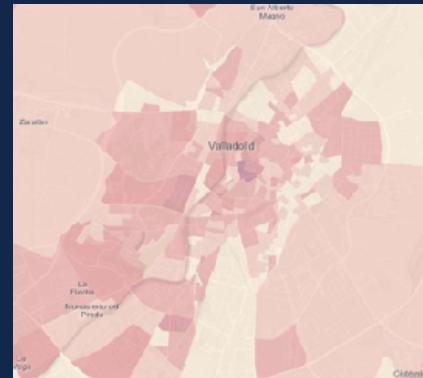
Carbon footprint for enterprise clients



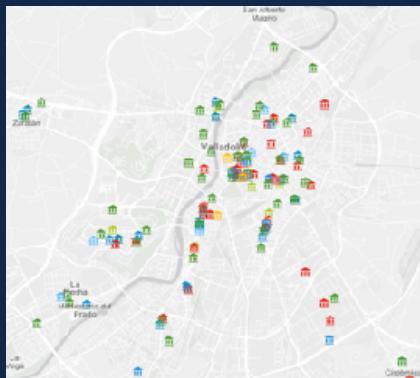
Geoblue: optimal branch location



BBVA footprint



Area, client, business & mobility data



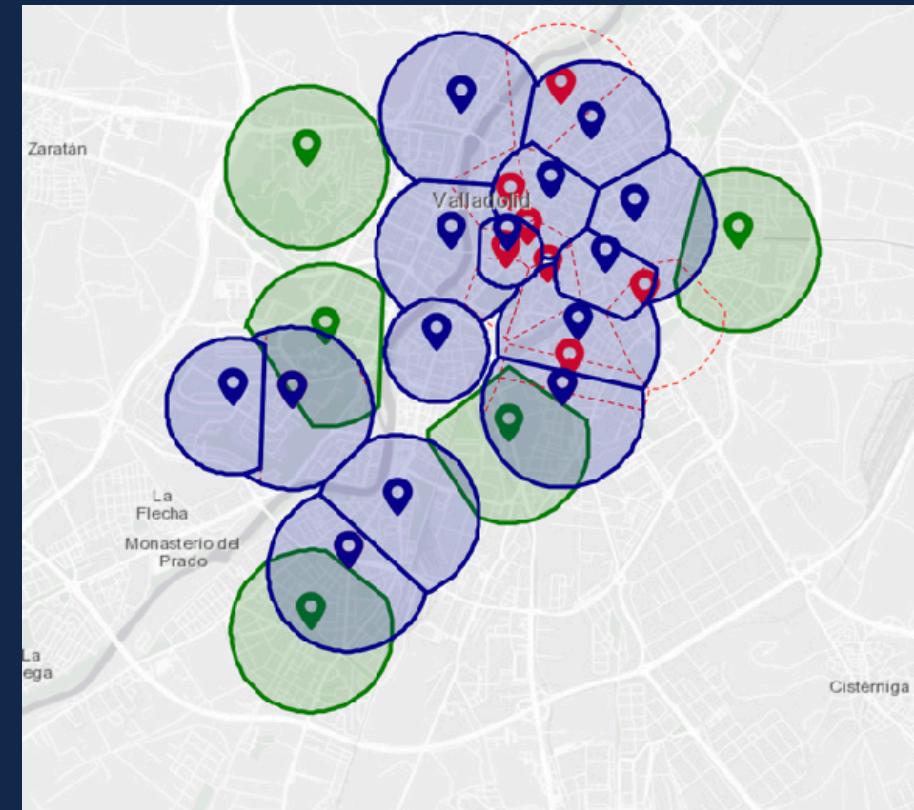
Competitors footprint



Digitality of customers



Optimized branch map



Business restrictions: Final number, to keep, to close...

Automatic Document Reading for Investment Funds



Select KIID: KID_ES_en_BE0058182792

1 de 2 Tamaño automático

management, an exposure to equity securities of eurozone companies.

Investment policy

- The fund mainly invests (without any sectorial restrictions) in shares and/or other equity securities of companies which have their registered office and/or a significant part of their assets, activities, profit centres or decision-making centres in a eurozone EU Member State.

Subscription/Redemption

- Subscription or redemption requests may be made every business day before 3 p.m. (Belgian time) to Banque Degroof Petercam SA or Caceis Belgium SA.

Risk and reward profile

Lower risk Higher risk

Typically lower rewards Typically higher rewards

1 2 3 4 5 6 7

As any investment is risky by nature, the return (performance) thereof is potential only. The risk and reward profile of your investment is reflected by a synthetic indicator which ranks the fund over a risk/return scale from the lowest (1) to the highest (7) risk category.

Certain risks are not (or not adequately) reflected in the risk and reward profile mentioned herein. Amongst those risks, the following could have a material impact on the risk and reward profile of the fund:

- Concentration risk: since the fund's portfolio mainly consists of equity securities issued by euro-zone companies, it is likely to be more specifically exposed to the economical development of this area.
- Risks linked to derivatives: the derivatives in the fund's portfolio are likely to be more volatile than the underlying assets thereof.

SRRI	Ongoing Charges(%)	Entry Charge(%)	Exit Charge(%)	Performance Fee(%)	Accurate Date
6	1.72	2	-	-	18/03/2019

- Allows parsing +500 daily files
- Operational costs & risk reduction

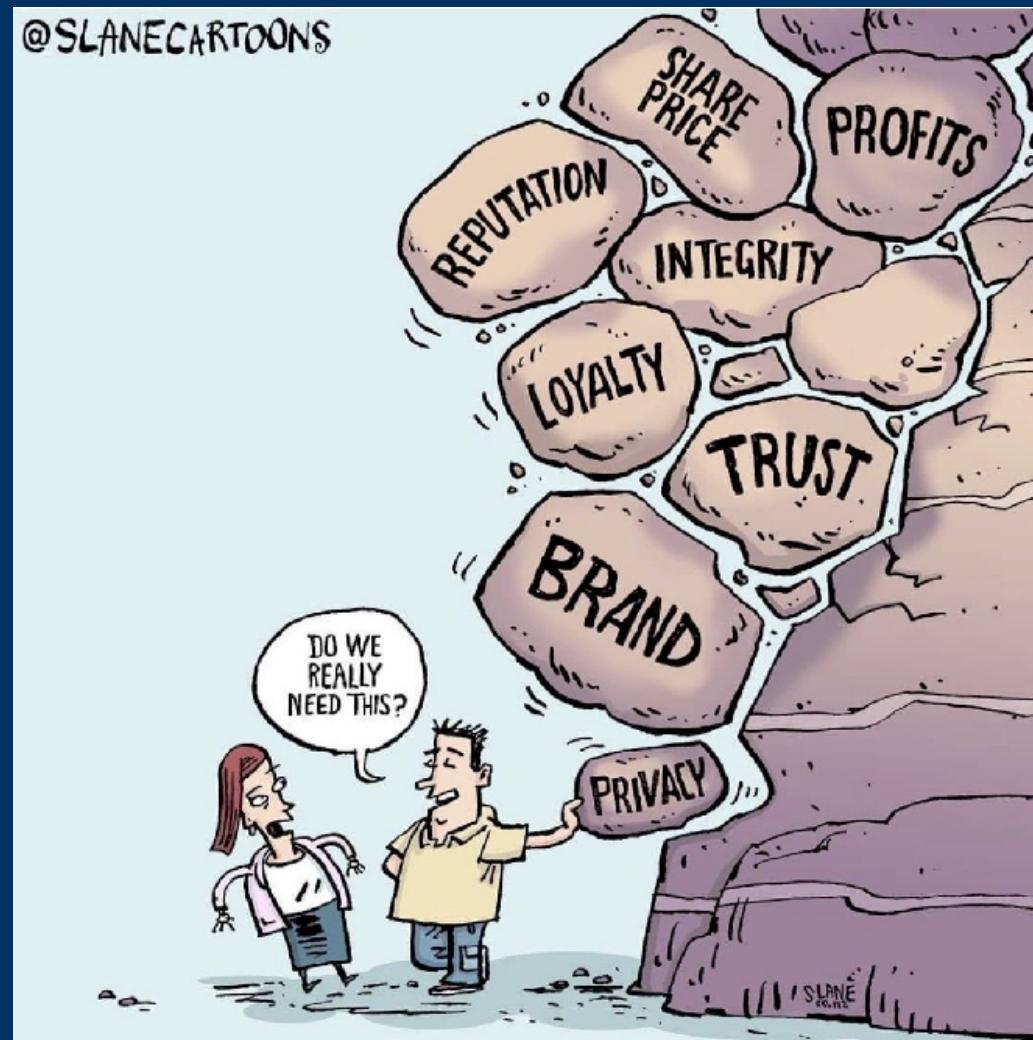
Design for Trust: limitations and risks of data driven systems



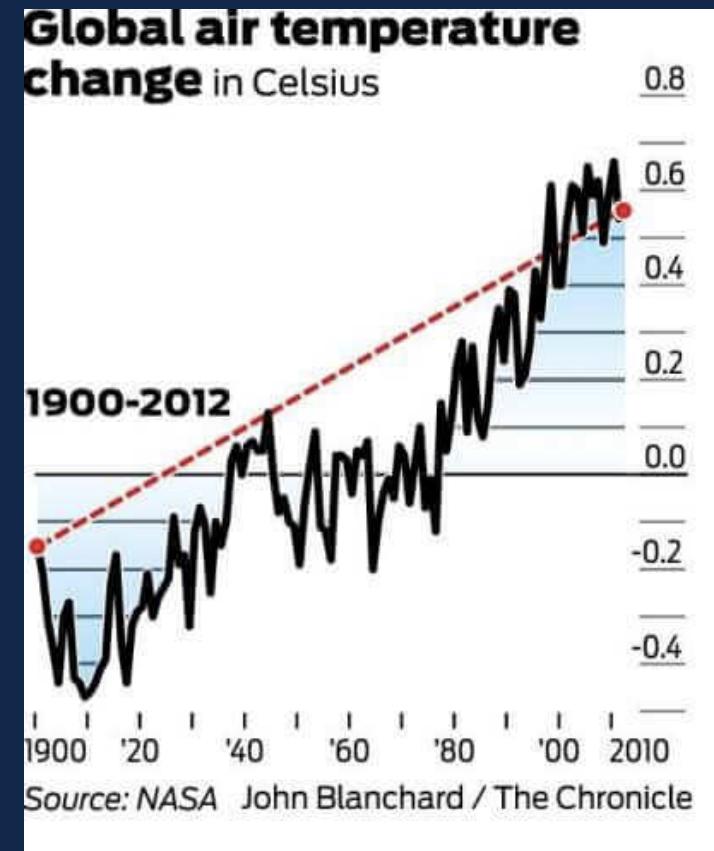
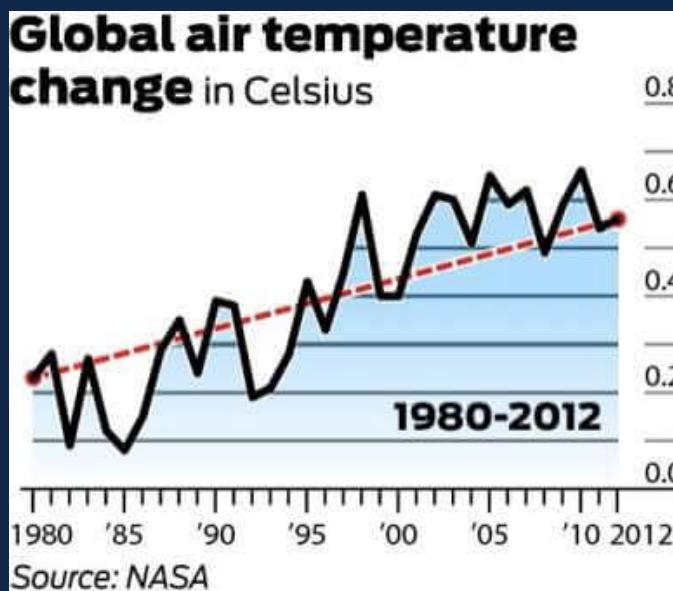
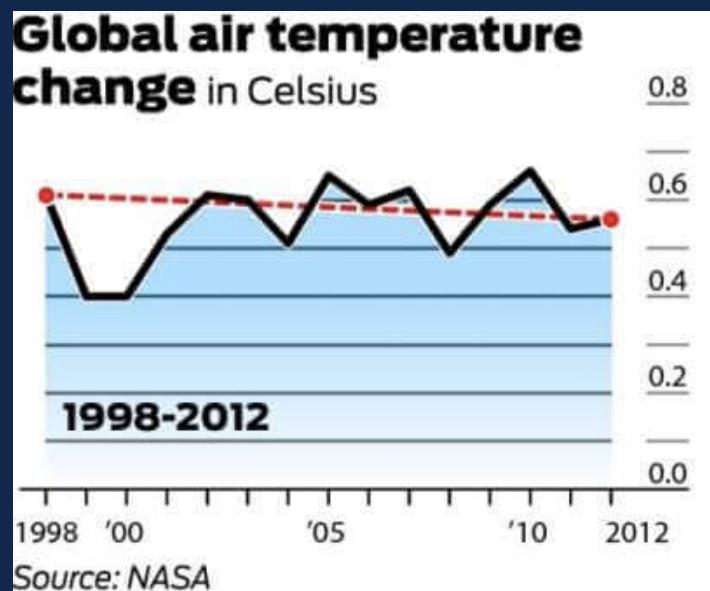
Even with the best of intentions, data-driven systems can easily:

- Diminish or violate privacy
- Perpetuate negative profiling
- Unfairly limit access to societal goods
- Reduce quality of life
- Destroy trust & brand equity

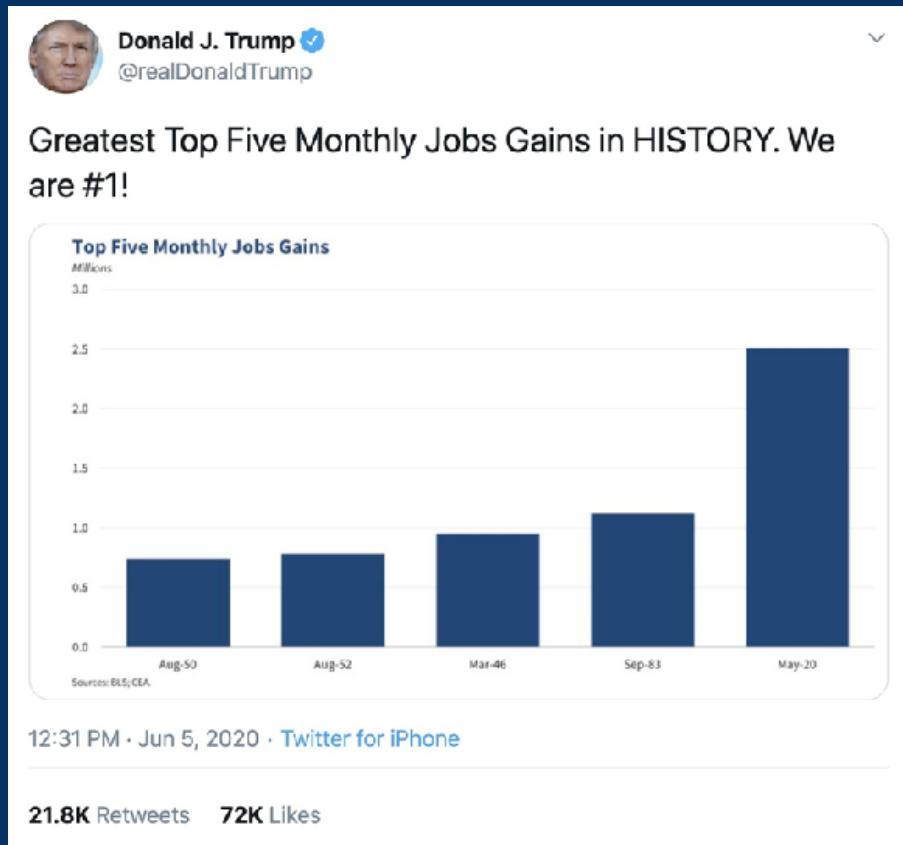
Privacy & security of data: the law and beyond



Lying with data: the axis, the series...



Lying with Data: cherrypicking



FAIRNESS & BIAS

How AI systems amplify bias

Image recognition systems that use biased machine learning data sets will inadvertently magnify that bias. Researchers are examining ways to reduce the effects.



COOKING	
ROLE	VALUE
AGENT ►	WOMAN
FOOD ►	PASTA
HEAT ►	STOVE
TOOL ►	SPATULA
PLACE ►	KITCHEN

COOKING	
ROLE	VALUE
AGENT ►	WOMAN
FOOD ►	FRUIT
HEAT ►	—
TOOL ►	KNIFE
PLACE ►	KITCHEN

COOKING	
ROLE	VALUE
AGENT ►	WOMAN
FOOD ►	MEAT
HEAT ►	GRILL
TOOL ►	TONGS
PLACE ►	OUTSIDE

COOKING	
ROLE	VALUE
AGENT ►	WOMAN
FOOD ►	VEGETABLES
HEAT ►	STOVE
TOOL ►	TONGS
PLACE ►	KITCHEN

COOKING	
ROLE	VALUE
AGENT ►	MAN
FOOD ►	—
HEAT ►	STOVE
TOOL ►	SPATULA
PLACE ►	KITCHEN

In this example of gender bias, adapted from a report published by researchers from the University of Virginia and the University of Washington, a visual semantic role labeling system has learned to identify a person cooking as female, even when the image is male.

IMAGES (FROM LEFT): ALEXEY STOI/ADOBESTOCK, JACOB LUND/ADOBESTOCK, BERND JURGENS/ADOBESTOCK, GRANGER/WOOTZ/GETTY IMAGES

← Tweet



Javier Vidal
@JaVidalPe

Google me dice que si quiero sacar el género de una persona en una foto entrene mi propio modelo predictivo. Que ellos a partir de hoy no se meten en inferir el género de nadie por su apariencia. 🏳️🌈

Translate Tweet

[Action Required] Cloud Vision API will not return gendered labels such as 'man' and 'woman' after February 19, 2020

Google Cloud Platform

MY CONSOLE

Hello Google Cloud Vision API customer.

We are willing to let you know that starting

7:01 PM · Feb 20, 2020 · Twitter Web App

3 Retweets 12 Likes

Given that a person's gender cannot be inferred by appearance, we have decided to remove these labels in order to align with the Artificial Intelligence Principles at Google, specifically Principle #2: Avoid creating or reinforcing unfair bias.

What do I need to do?

To continue using the 'label detection' feature:

- We recommend that you annotate several test images containing persons with this API using the 'LABEL_DETECTION' feature after February 19, 2020, to understand how the labels returned will change. After

Machine Bias

Algorithmic injustice and the formulas that increasingly influence our lives

Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't

Amazon bills itself as "Earth's most customer-centric company." Yet its algorithm is hiding the best deal from many customers.

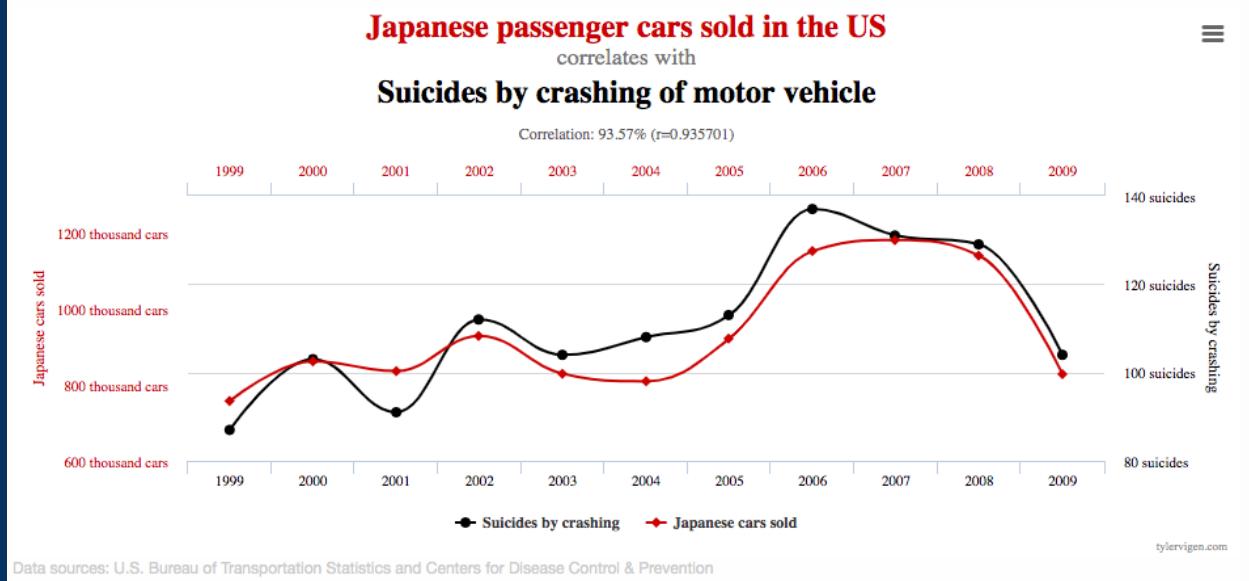
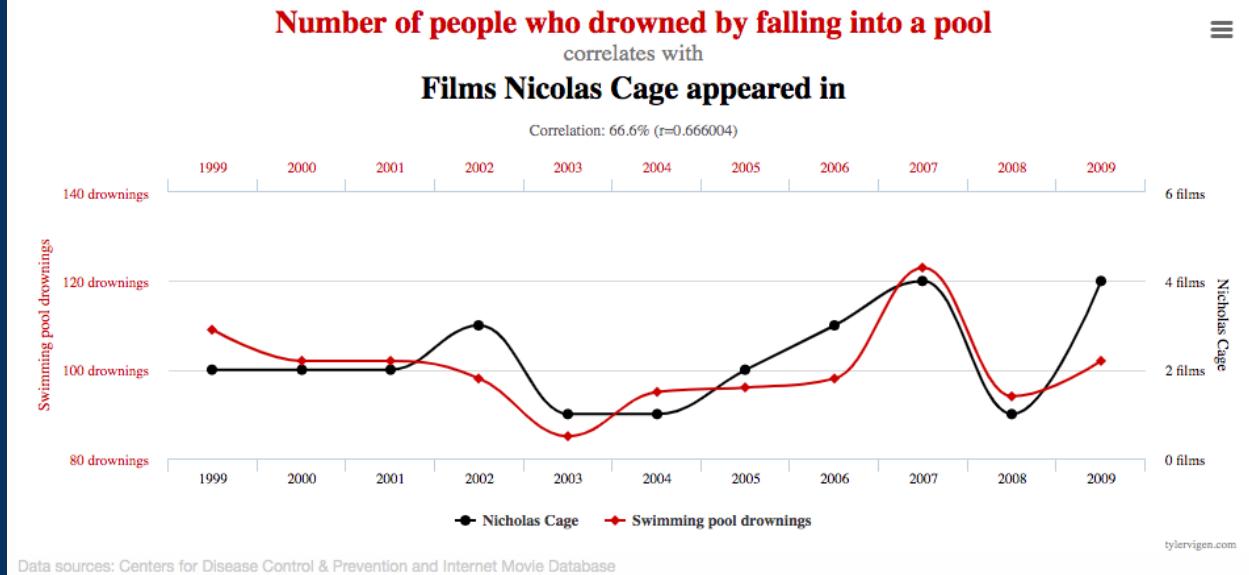
By Julia Angwin and Suresh Venkatasubramanian
ProPublica, Sep. 20, 2016, 8:45 a.m.

This is part of an ongoing series on Machine Bias.

We're looking at how algorithmic automation and the formulas that increasingly influence our lives

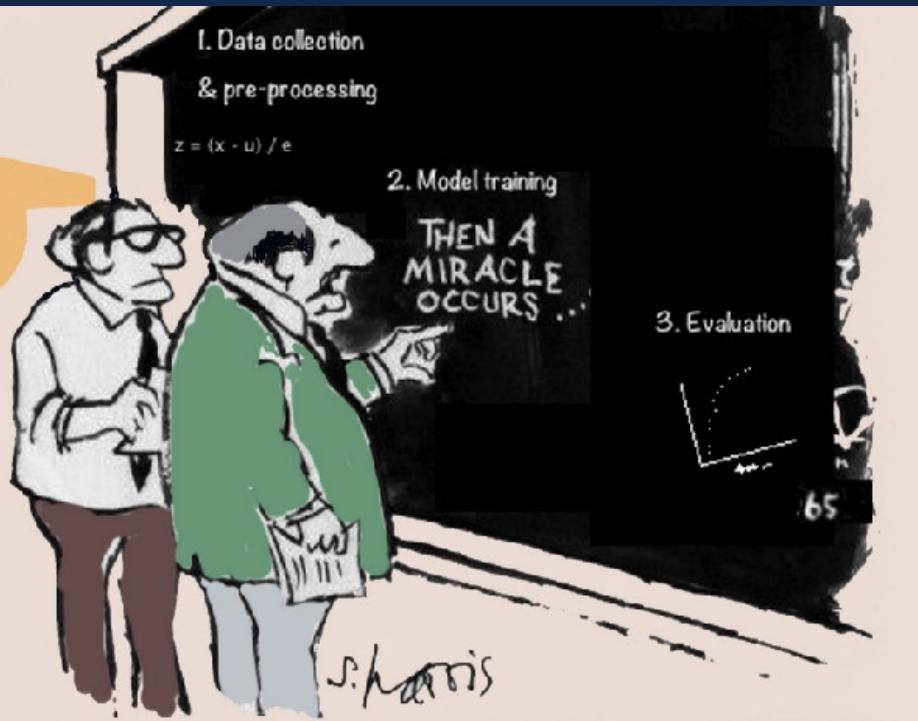


CORRELATION IS NOT CAUSATION



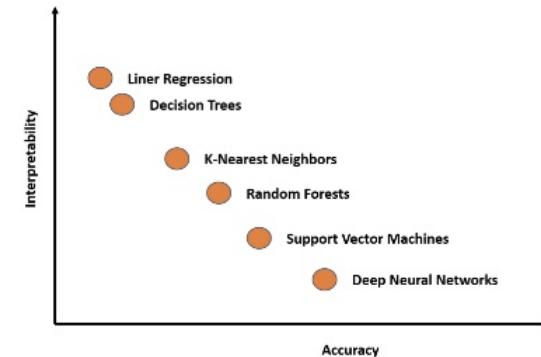
Source: <https://tylervigen.com/spurious-correlations>

When no interpretability is a problem



from *What's so Funny about Science?* by Sidney Harris (1977)

Accuracy vs. Interpretability



Critical systems where we need:

- Fairness:** pricing using unfair variables
- Robustness:** critical system if hacked
- Causality:** credit scoring
- Trust:** medical applications

Goal setting: be careful with what you wish...

Jiri [REDACTED]@gmail.com>
to me ▾

Hi James,

I like the picture of my wife and I, but can you get rid of the other guy in the white hat please? Love your work!!



James Fridman <jjamie013@gmail.com>

Sure.



Jesse [REDACTED]@icloud.com>
to me ▾

Hey James you amazing at editing. I was wondering if you could make the phone in my pocket disappear it's really noticeable.

Thanks!

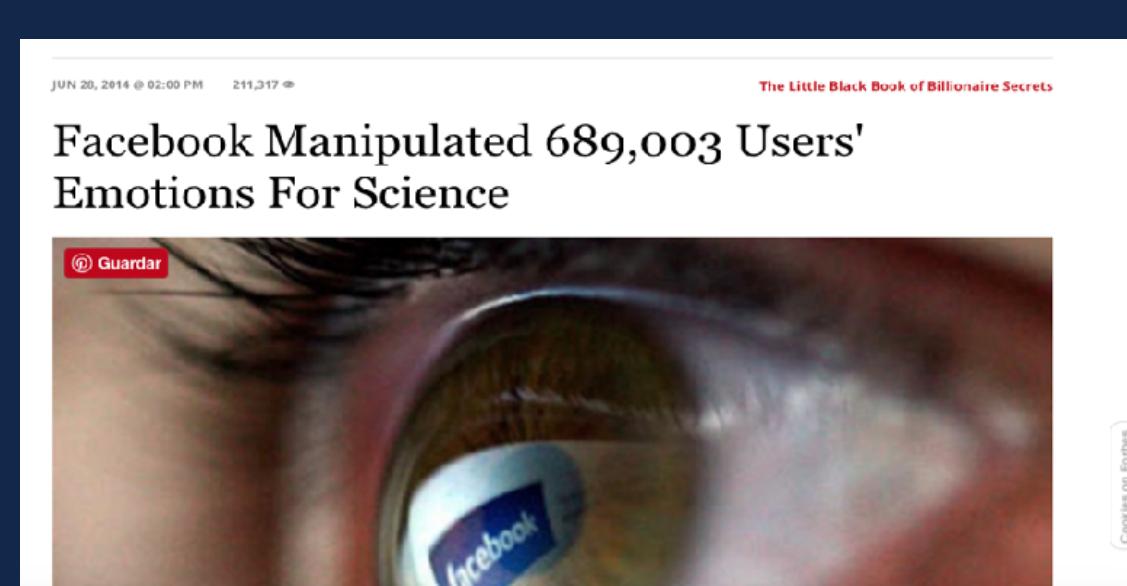
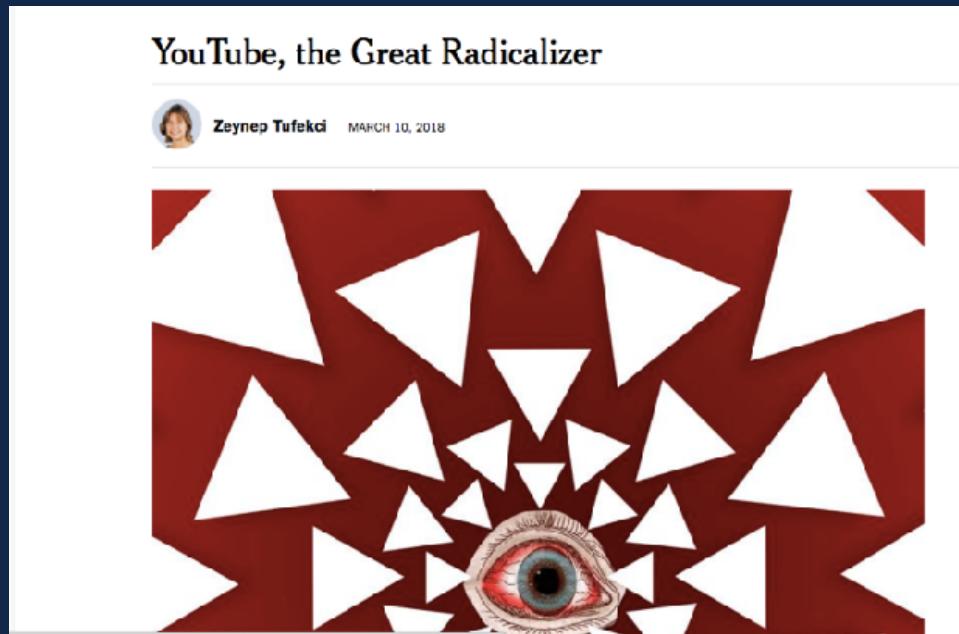


James Fridman <jjamie013@gmail.com>

Sure.



Goals transparency, intention & its consequences



WHAT CAN WE DO?

RESPONSIBLE DATA USAGE CHECKLIST (EXAMPLES):

- CHALLENGE YOUR OWN MATH
- PERMANENT TEST AND LEARN IS A MUST
- BIASES (IN DATA AND IN USAGE)
- HOW CRITICAL IS THE PROBLEM?
- WHO ARE WE EMPOWERING?
- WHAT IS THE WORST THING ONE COULD DO?
- ARE THE VARIABLES ACTIONABLE FOR THE SUBJECT?

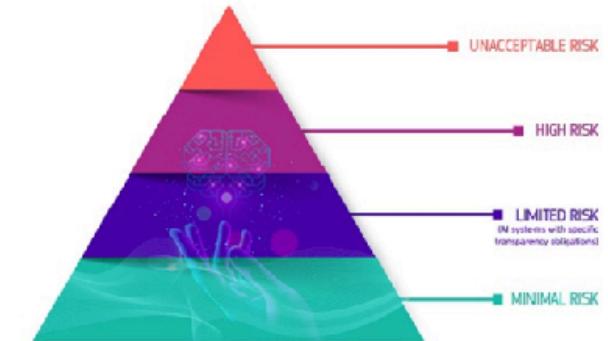
NEW! EUROPEAN REGULATION ON AI

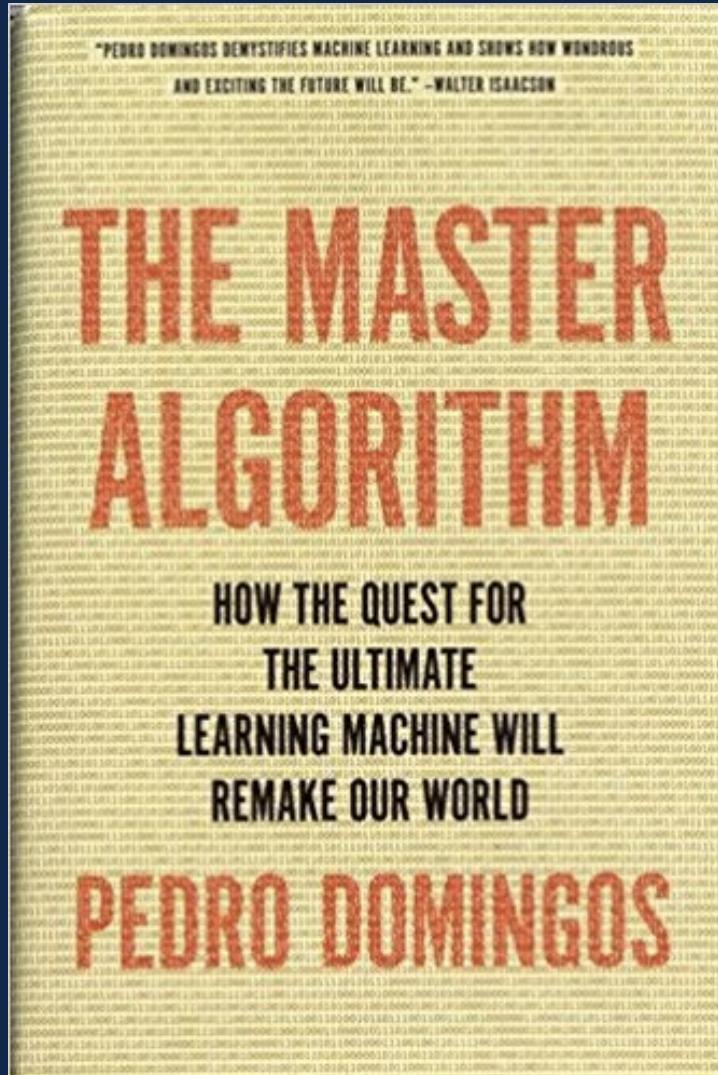
BBVA

Creando Oportunidades

The screenshot shows the European Commission's website for the AI regulatory framework proposal. The URL is digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai. The page features the European Commission logo and a search bar. The main heading is "Shaping Europe's digital future". Below it, a navigation menu includes Home, Policies, News, Library, Funding, Calendar, and Consultations. A breadcrumb trail shows the path: Home > Policies > Regulatory framework proposal on Artificial Intelligence. The main content title is "Regulatory framework proposal on Artificial Intelligence". A text block states: "The Commission is proposing the first ever legal framework on AI, which addresses the risks of AI and positions Europe to play a leading role globally." Another text block explains: "The regulatory proposal aims to provide AI developers, deployers and users with clear requirements and obligations regarding specific uses of AI. At the same time, the proposal seeks to reduce administrative and financial burdens for business, in particular small and medium-sized enterprises (SMEs)." A third text block notes: "The proposal is part of a wider AI package, which also includes the updated Coordinated Plan on AI. Together they guarantee the safety and fundamental rights of people and businesses, while strengthening AI uptake, investment and innovation across the EU." To the right of the text is a photograph of two people looking at a large, transparent screen displaying a complex, abstract digital visualization.

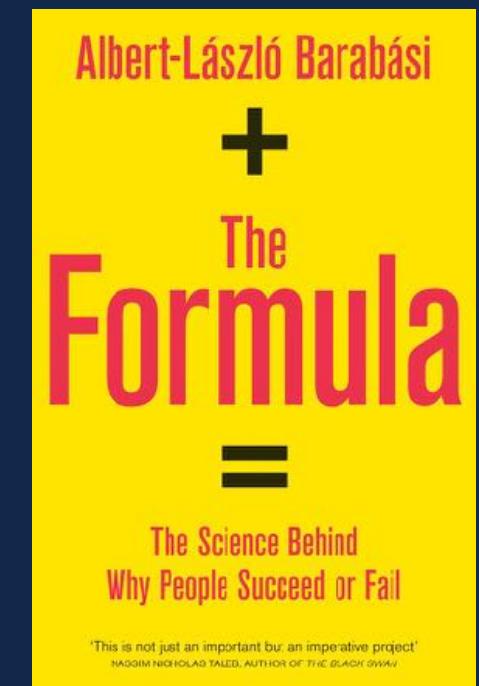
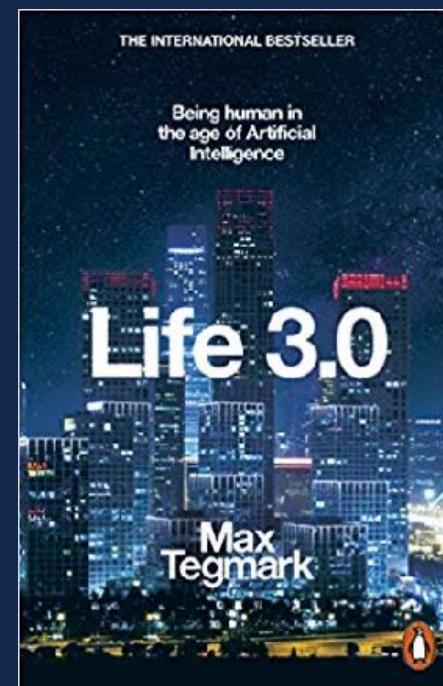
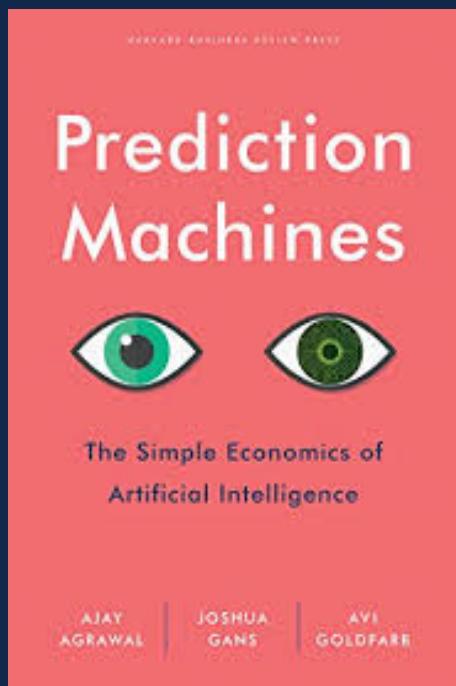
A risk-based approach





“People worry that computers will get too smart and take over the world, but the real problem is that they're too stupid and they've already taken over the world.”

Pedro Domingos



Gracias!!

