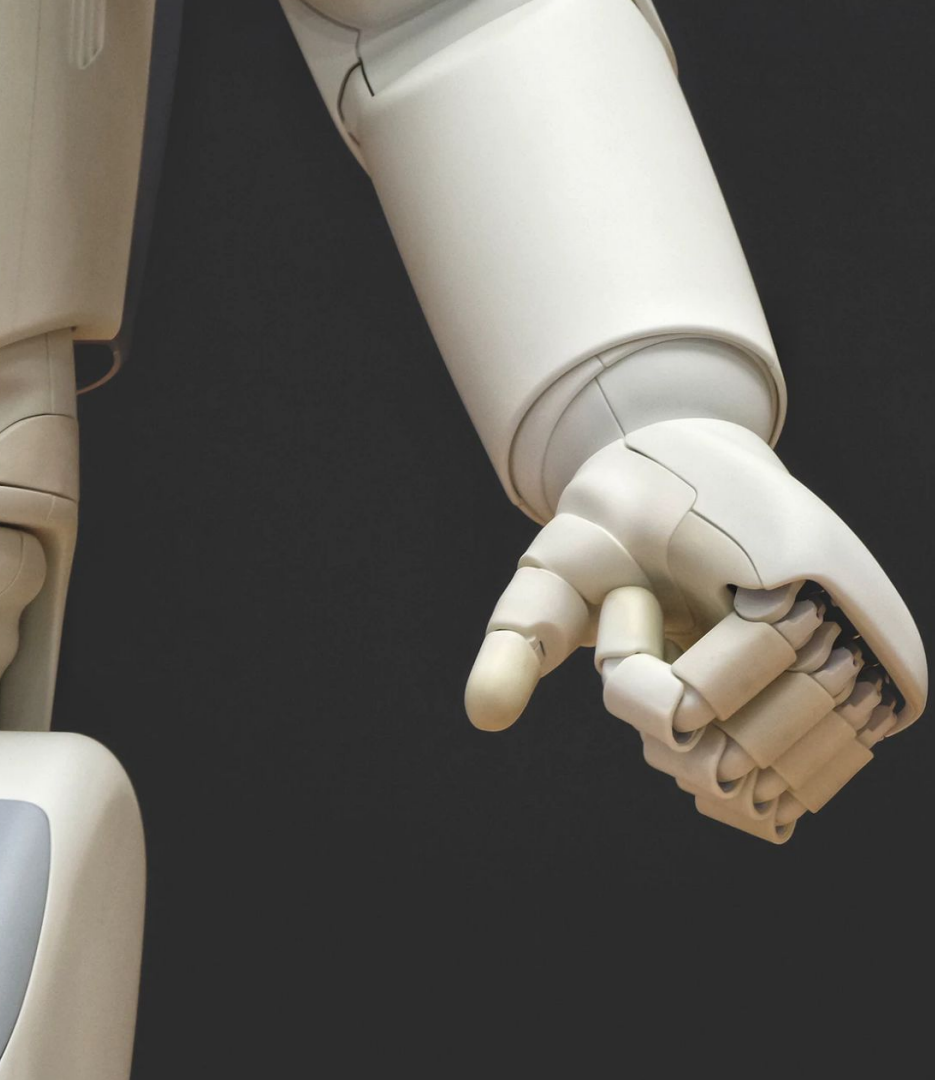


CUANDO LAS MÁQUINAS NOS DEFIENDAN

Javi Ramirez

@rameerez @rameerezcom

GUARD

www.useguard.com

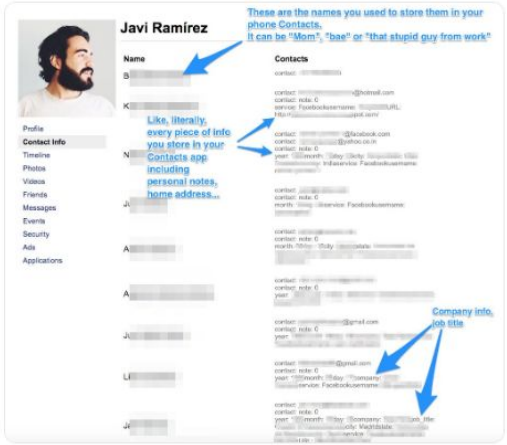
Rameerez
@rameerez

Follow

Don't freak out, but Facebook is keeping an updated copy of your entire phone Contacts info, including every nickname you've used to store them, address info and personal notes.

Okay, maybe freak out a bit.

#DeleteFacebook



12:07 PM - 23 Mar 2018

136 Retweets **165** Likes

15 136 165

Sergio Quintana
@svqjournalist

Follow

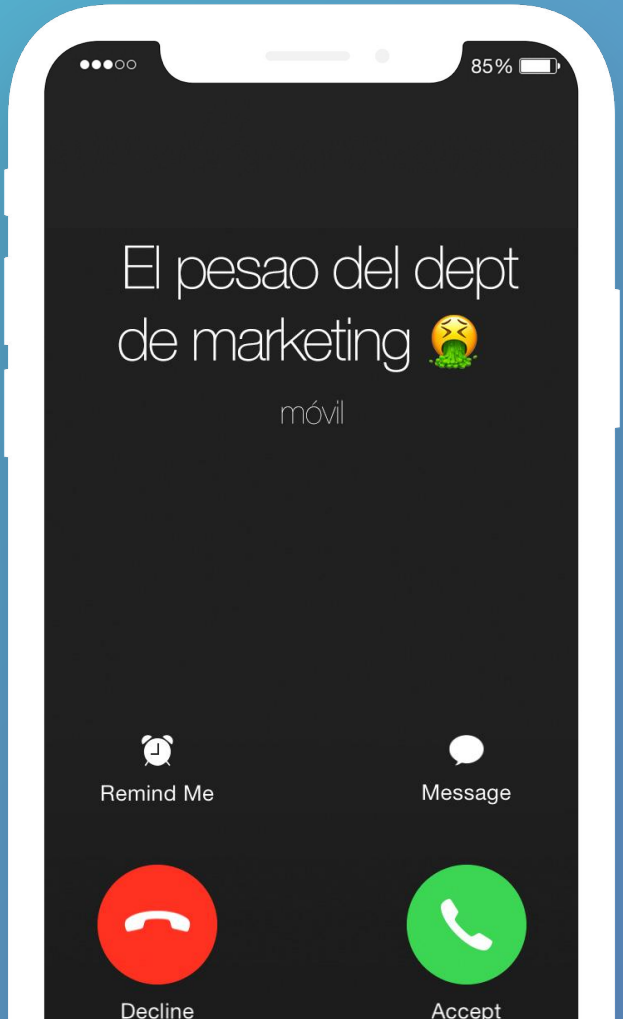
Replying to @rameerez

Javi, I'm a reporter at the NBC Station in San Francisco. May I use your Tweets about Facebook Contacts for my Report tonight? It may appear across all our platforms.

8:47 PM - 25 Mar 2018 from San Francisco, CA

1 Like

1 1


85%
El pesao del dept de marketing 🤮
móvil
Remind Me
Message
Decline
Accept

0.001%

empieza a leer la letra pequeña

Bakos Y. et al.
"Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts"
New York University, 2009

3.040 h

para leer las políticas de privacidad que has aceptado solo en los últimos 5 años

McDonald, A.M. & Cranor, L.F.. (2008). The cost of reading privacy policies. I/S: A Journal of Law and Policy for the Information Society. 4. 540-565.

“ACEPTO”

es la mayor mentira del siglo XXI

“VALORAMOS TU PRIVACIDAD”

seguramente sea la segunda

CNBC

TECH

Facebook exodus: Nearly half of young users have deleted the app from their phone in the last year, says study

PUBLISHED WED, SEP 5 2018 • 12:18 PM EDT
UPDATED TUE, SEP 11 2018 • 2:45 PM EDT

Matt Rosoff
@MATTROSOFF

SHARE

KEY POINTS

- A new survey of more than 3,400 U.S. Facebook users finds that 44 percent of users ages 18 to 29 have deleted the app from their phones in the past year.
- Overall, 26 percent have deleted the app, while 42 percent have taken a break of several weeks or more.

Mashable

Facebook sets up holiday pop-up shop in New York to inform users about privacy concerns

Facebook's one-day privacy pop-up shop sits inside New York City's Bryant Park.

IMAGE: MATT BINDER / MASHABLE

BY MATT BINDER

“The digital services we use are constantly gathering information about us, and the mere existence of this data makes us vulnerable in ways we can’t even anticipate yet”

JAVI RAMIREZ

@rameerez @rameerezcom

Tech entrepreneur + AI + design

Ingeniería Informática (URJC)
Administración y Dirección de Empresas (URJC)

Experience on multiple industry Big Data and Data Visualization projects

1 STARTUP A MONTH

Wakefy
Turn your Mac into a Spotify alarm clock.

Decidr
Your own personal decision assistant.

Le French generator
Generate random French text.

HUSTL
Create awesome time-lapse videos of your Mac screen.

MAD RIDES
Beautiful biking maps.

Rameerez
My personal website and online store (3D prints, merch...)

Kibbles
Cable protectors to keep your chargers from breaking.

Gridly
Create beautiful Instagram feeds right from your desktop.

Krowspot
Discover and share the best places to work from.

The Pixel Challenge
Build impactful daily habits and transform your life.

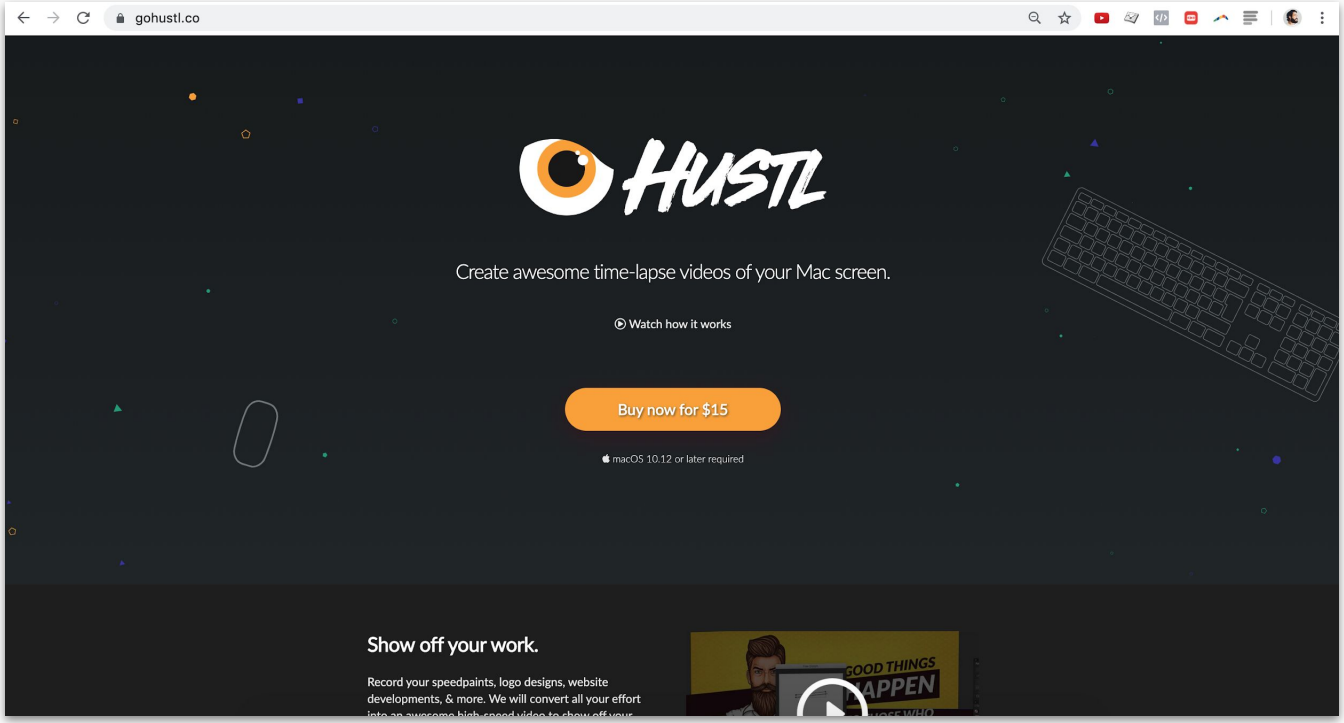
MINDLESS
Streetwear for creative people

Momoise
Block the distracting parts of time wasting websites.

 @rameerez

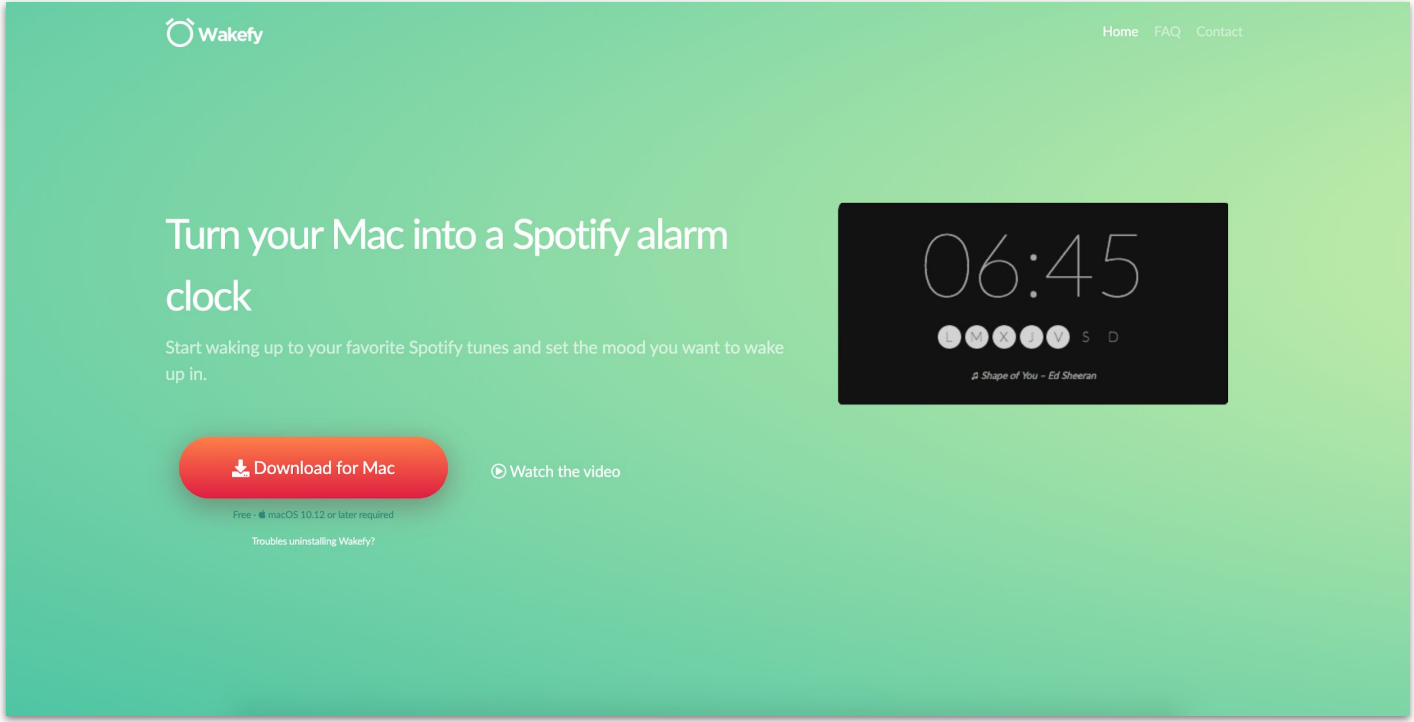
 @rameerezcom

1 STARTUP A MONTH



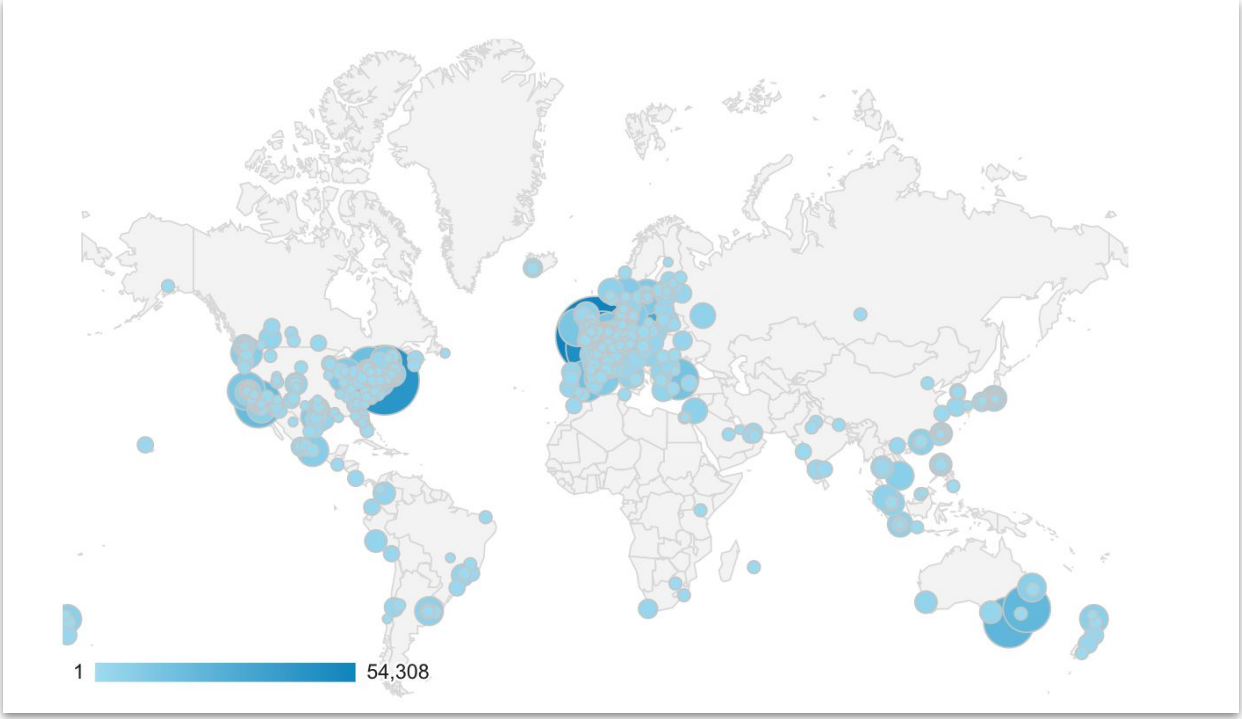
@rameerez @rameerezcom

1 STARTUP A MONTH



@rameerez @rameerezcom

50k+ USERS IN 167 COUNTRIES



@rameerez @rameerezcom

TRUSTED BY

Uber

Adobe

And many more

@rameerez

@rameerezcom


GUARD

DEMO

¿CÓMO LE ENSEÑARÍAS PRIVACIDAD A UNA MÁQUINA?

OPCIÓN 1:
REGLAS SIMPLES

Si tus datos son privados = bien

Si tus datos son públicos = mal

PROBLEMA 1:
¿ESO ES TODO?

No todo es blanco o negro, hay escalas de grises amplísimas.

PROBLEMA 2:
¿DE DÓNDE SACAMOS ESA INFO?

Las políticas de privacidad son la única interfaz entre servicio y usuario.

OPCIÓN 2:
QUE ALGUIEN SE LAS LEA Y SAQUE LAS REGLAS

PROBLEMA 3:
LOS HUMANOS FALLAN Y TARDAN

Tienen sesgos subjetivos, despistes, errores y son difíciles de escalar.

IDEA:

OBTENER LO QUE TODA LA HUMANIDAD ENTIENDE COMO PRIVADO Y ENSEÑÁRSELO A UNA MÁQUINA PARA QUE LO HAGA AUTOMÁTICAMENTE

PROBLEMA:
¿QUÉ ES LA PRIVACIDAD?

PROBLEMA:

¿CÓMO MEDIMOS DE MANERA OBJETIVA PERCEPCIONES SUBJETIVAS?

PROBLEMA:

¿CÓMO OBTENEMOS UN AGREGADO MASIVO DE PERCEPCIONES SUBJETIVAS?

PROBLEMA:

¿CÓMO LE ENSEÑAMOS ESO A UNA MÁQUINA?

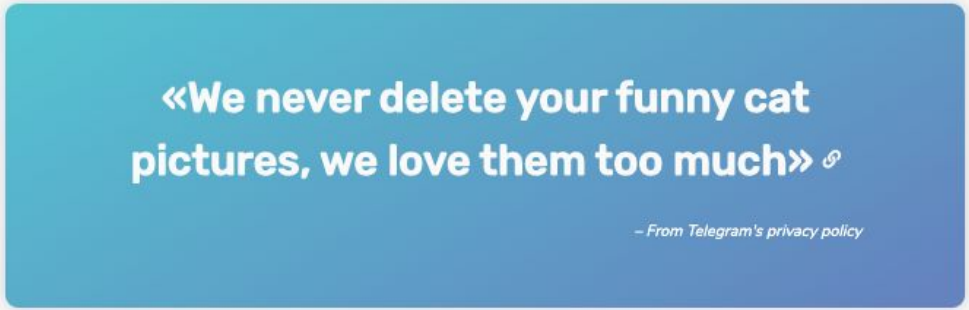
TRES PASOS:

1. Definir lo que es privacidad y cómo medirla

2. Construir un dataset de lo que el ser humano considera “privacy friendly” y lo que no

3. Construir una IA que sepa leer textos y usar esos datos para enseñarle a diferenciar lo “privacy-friendly” de lo potencialmente dañino

1 ¿QUÉ ES PRIVACIDAD?



2 ¿CÓMO MEDIR ESTO A ESCALA?

FRASE A

FRASE B

GUARD

346 votes

WHICH SENTENCE LOOKS MORE PRIVACY FRIENDLY?

SENTENCE A

In addition, users giving feedback may be individually contacted for follow-up due to concerns raised during the course of such evaluation. Demographic information and Web log data may be stored for future research and evaluation. 6.

SENTENCE B

Statistical information Aggregate statistical information provided by us to our advertisers or others regarding sales or website usage will not include personally identifying information. Security Your personal data will be held on a secure server. Our sites are equipped with the latest security devices/firewalls.

Guard's "Ethical Privacy" project quantitatively measures privacy perception by crowdsourcing human ethics surveys to users around the globe.

VIEW HELP

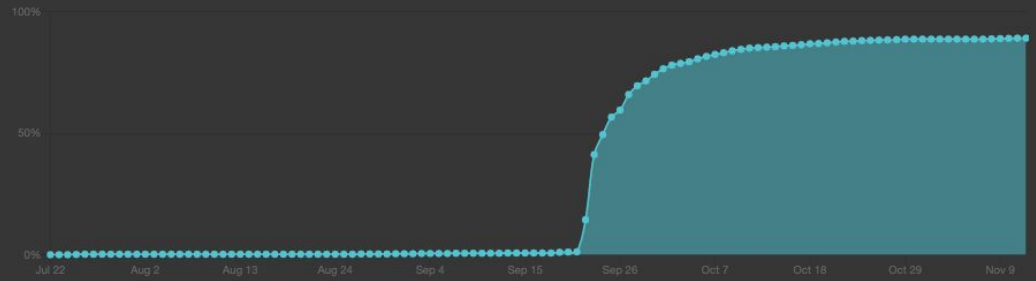
USEGUARD.COM

AI STATUS - OPEN DATA

Guard is an Artificial Intelligence that reads privacy policies for you – and warns you of the biggest risks it finds. Right now it's learning, this is the open report of its status, in real time.
PLAY NOW to teach the AI.

AI status: gathering data

AI training progress



Want to help? You can teach the AI just by playing a simple game.

PLAY NOW

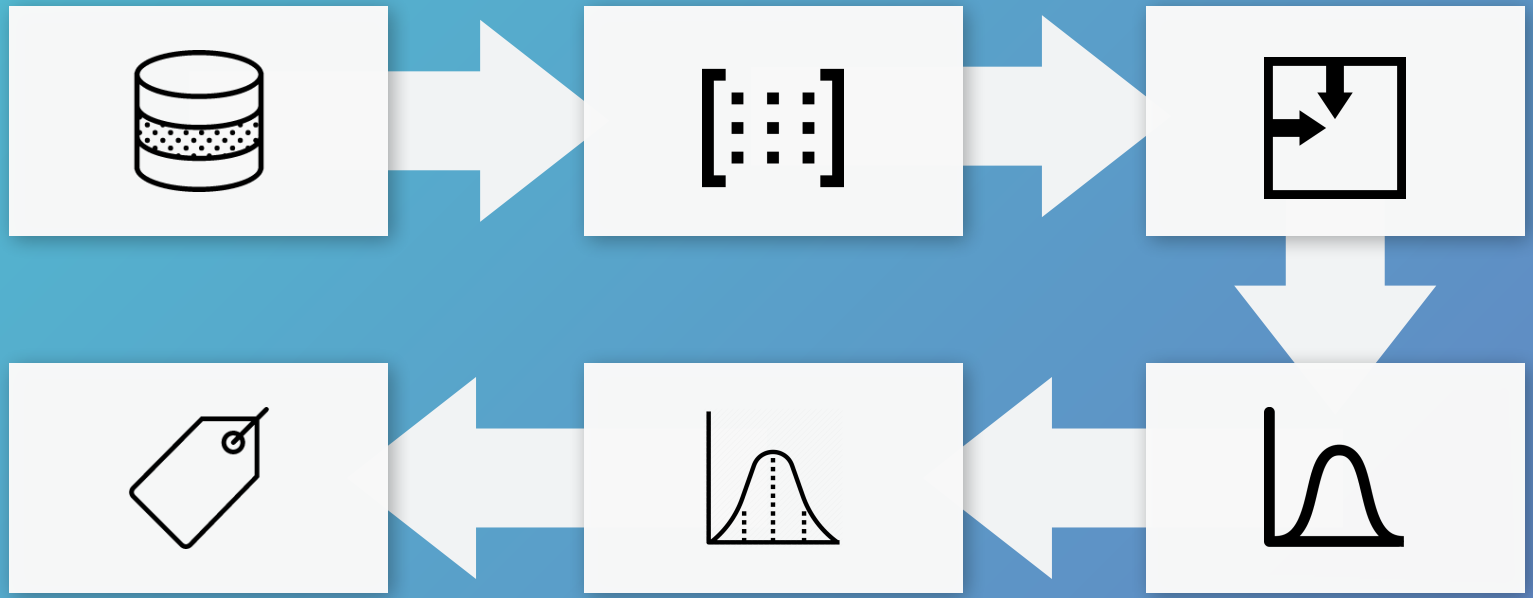
89.3%
TRAINED

89264
DATA POINTS

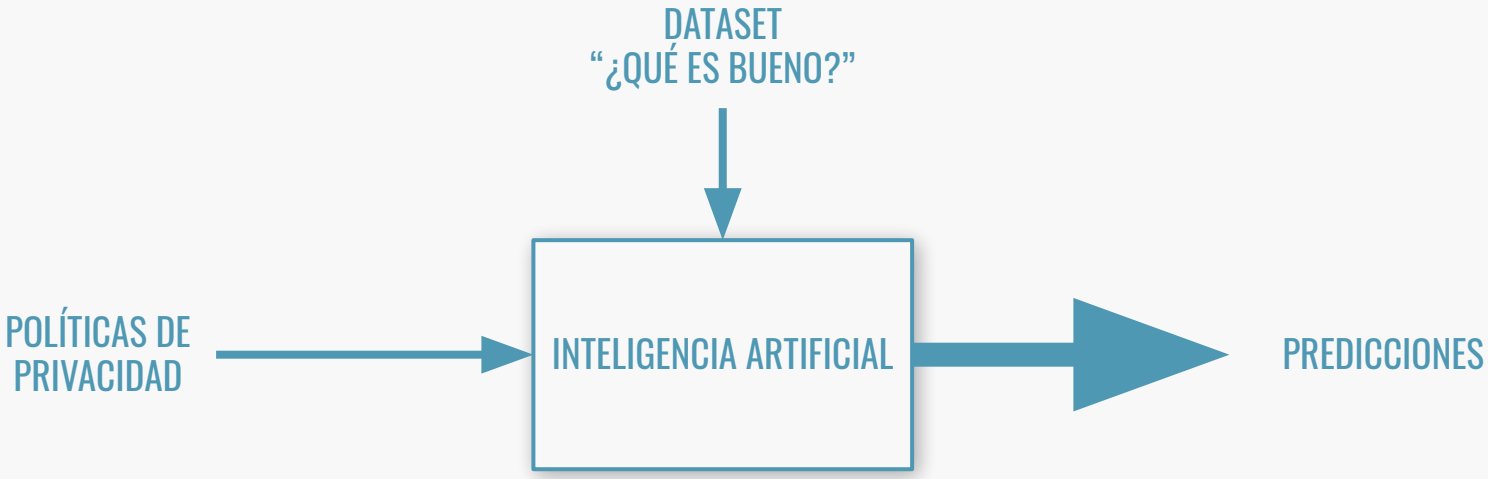
2443
AI INSTRUCTORS

MY APPROACH: «MORALSCORE»

A data processing algorithm to transform votes into morally labelled data



3 ¿CÓMO ENSEÑARLE ESTO A UNA MÁQUINA?



“We collect information about you when you use our products or services. We may collect information and use it as required by law (for example, how we use our service and your choices about receiving future marketing communications from us). We also collect information you provide.”

GUARD

100% AI-generated text.
No human input or modifications whatsoever.

malicious actors on our services, or to explain why we have removed content or accounts from our services; to address fraud, security, or technical issues; or to protect our rights or property or the rights or property of those who use our services. However, nothing in this Privacy Policy is intended to limit any legal defenses or objections that you may have to a third party's, including a government's, request to disclose your personal data.

3.4

Affiliates and Change of Ownership

In the event that we are involved in a bankruptcy, merger, acquisition, reorganization, or sale of assets, your personal data may be sold or transferred as part of that transaction. This Privacy Policy will apply to your personal data as transferred to the new entity. We may also disclose personal data about you to our corporate affiliates in order to help operate our services and our affiliates' services, including the delivery of ads.

3.5

Non-Personal Information

We share or disclose non-personal data, such

POTENTIAL RISK DETECTED

In sentence: 326
Labeled as: «very_dangerous»
Confidence: 96.7%

from Twitter's privacy policy

We sell your data to third parties.

94.9%

We don’t sell your data to third parties.

1.6%

GUARD

DEMO

GUARD

USEGUARD.COM

¿QUÉ ES UNA IA?

¿PUEDE UNA IA SER BUENA O MALA?

LO QUE NO ES IA

Que los ordenadores piensen

Que las máquinas nos conquisten

Ordenadores “creativos”

Magia

LO QUE NO ES IA

Que los ordenadores piensen

Las máquinas nos conquisten

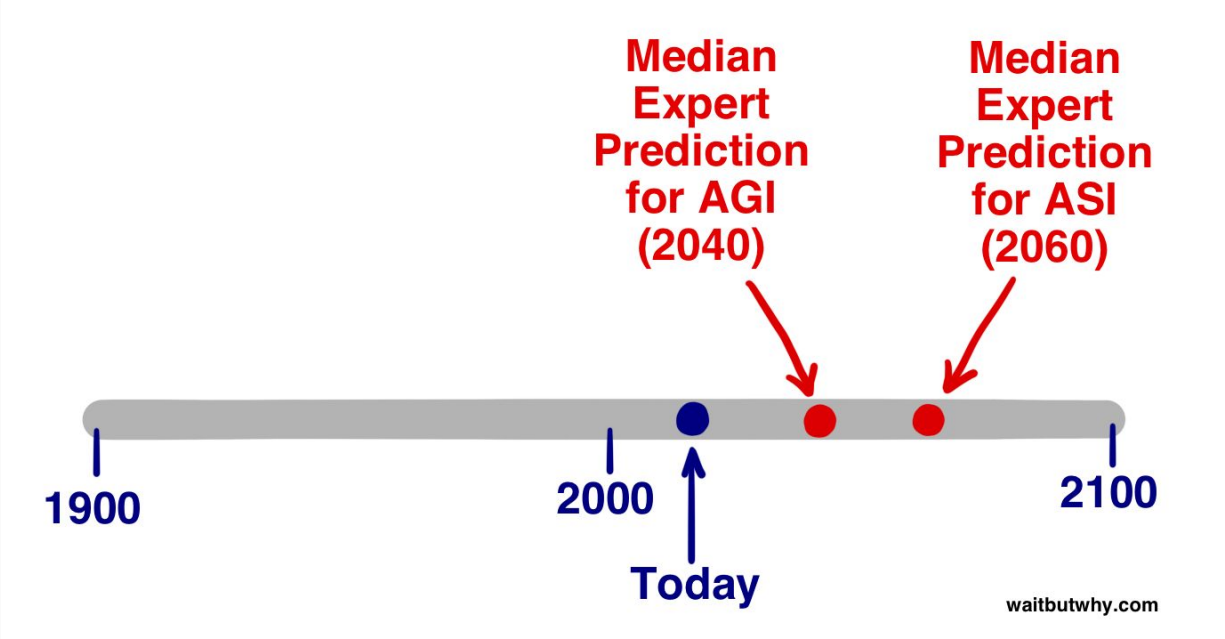
Ordenadores “creativos”

Magia

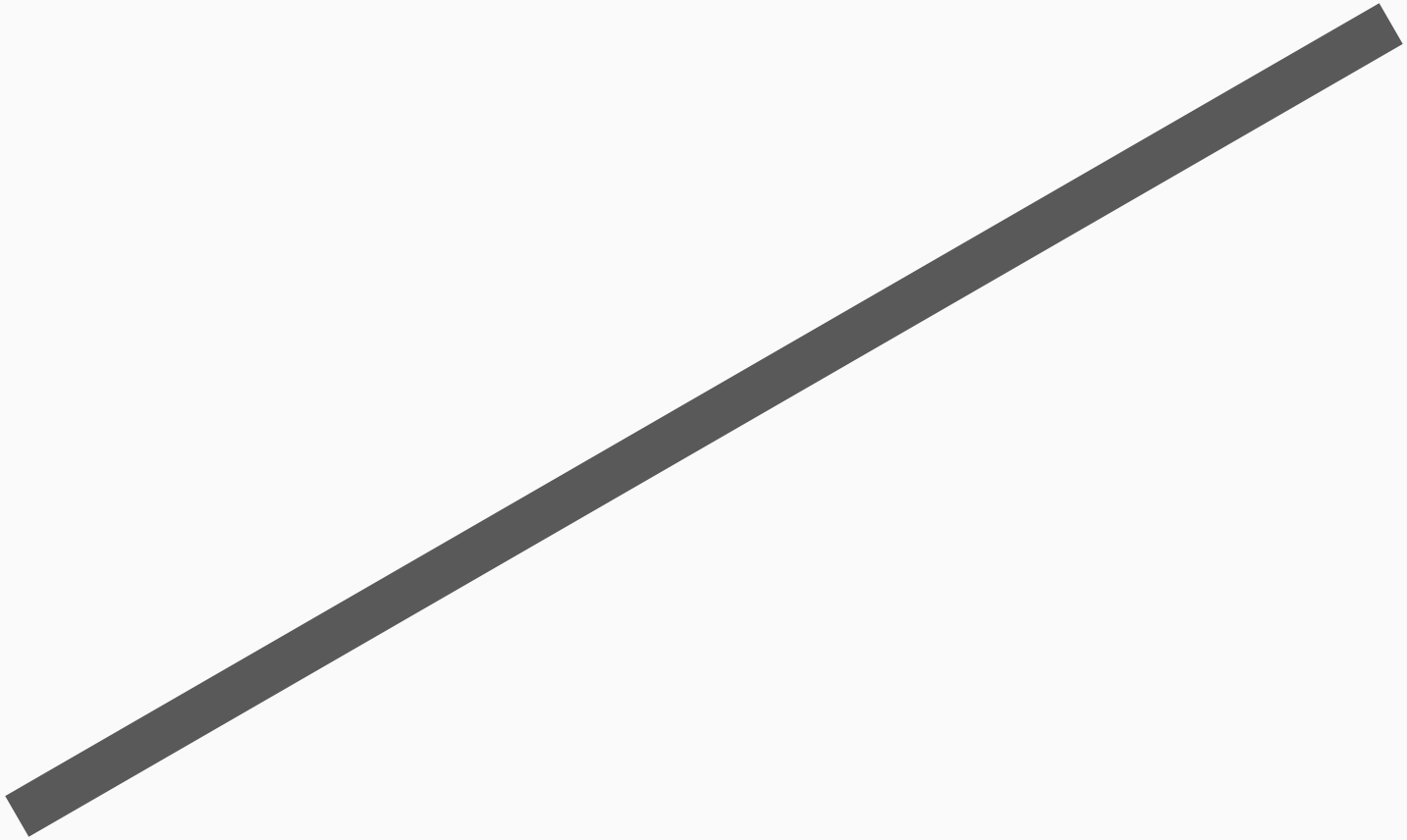
People with no idea about AI saying it will take over the world:

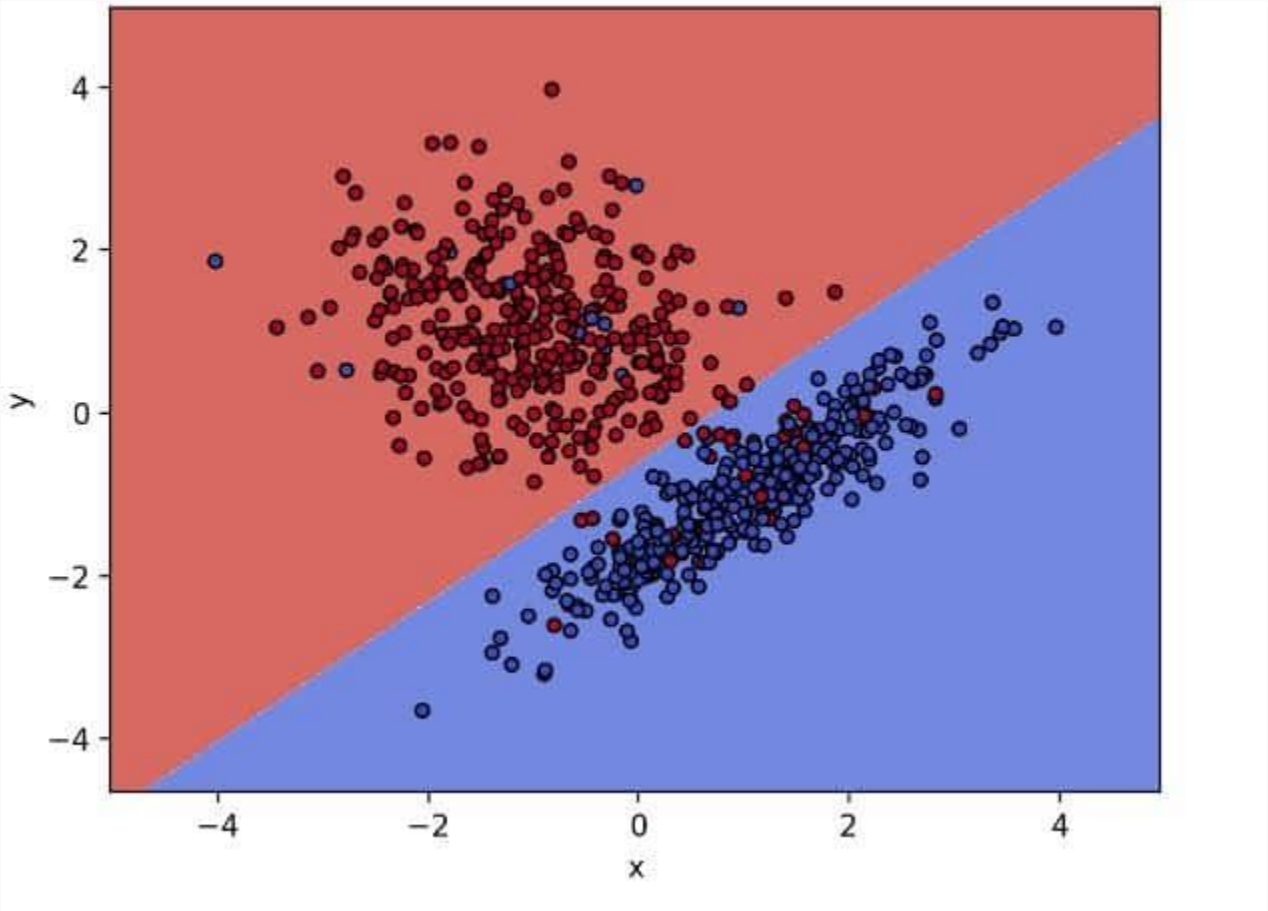
My Neural Network:

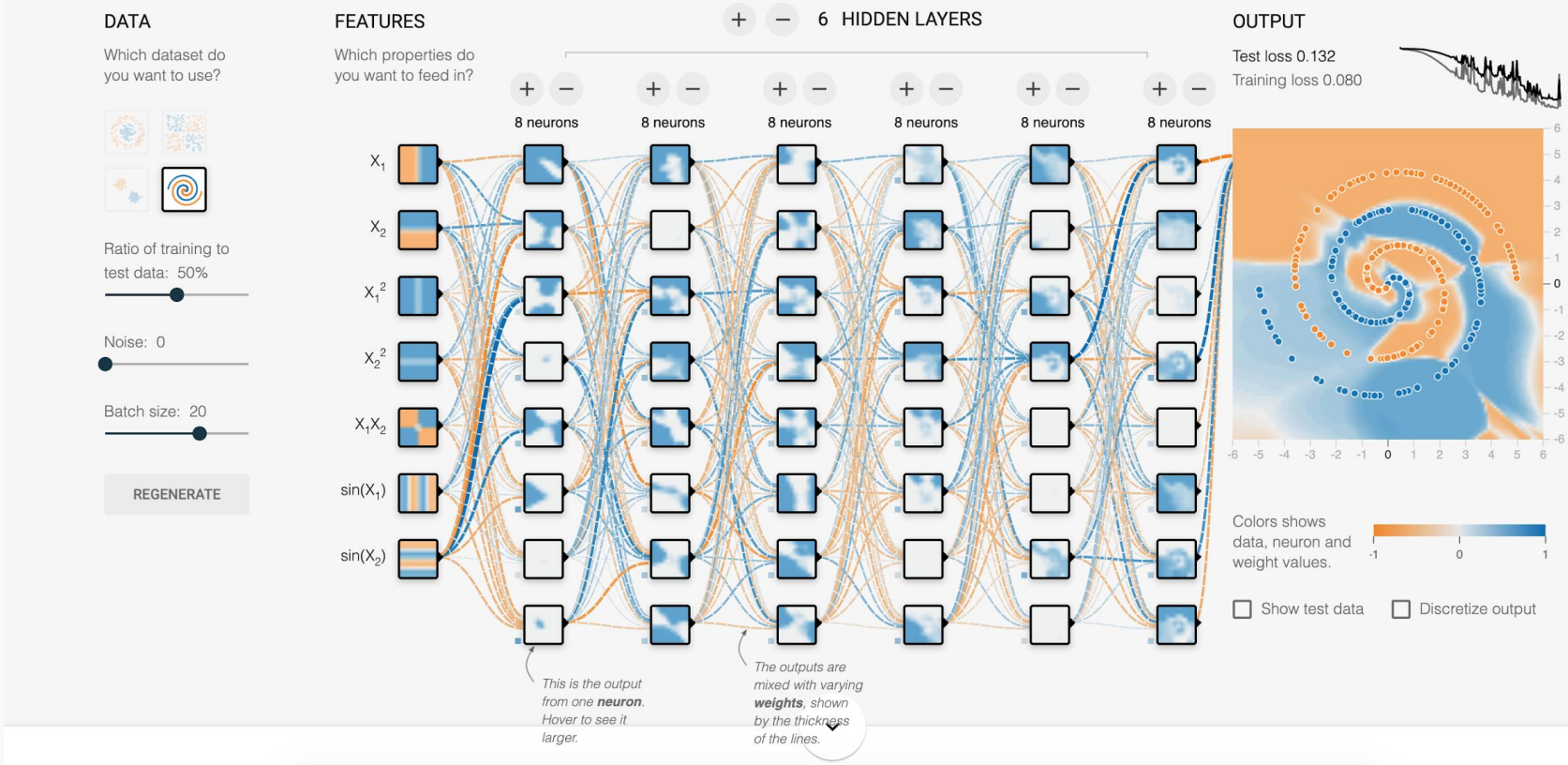
@rameerez @rameerezcom


Median Expert Prediction for AGI (2040)
Median Expert Prediction for ASI (2060)
1900
2000
2100
Today
waitbutwhy.com

LA IA MÁS SENCILLA DEL MUNDO




4
2
0
−2
−4
y
−4
−2
0
2
4
x



@rameerez @rameerezcom

$$i_t = \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right) \quad (7)$$

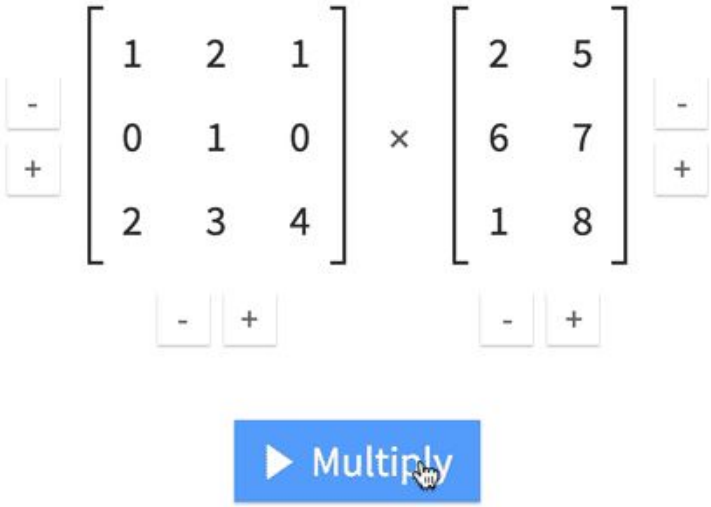
$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \quad (8)$$

$$c_t = f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \quad (9)$$

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right) \quad (10)$$

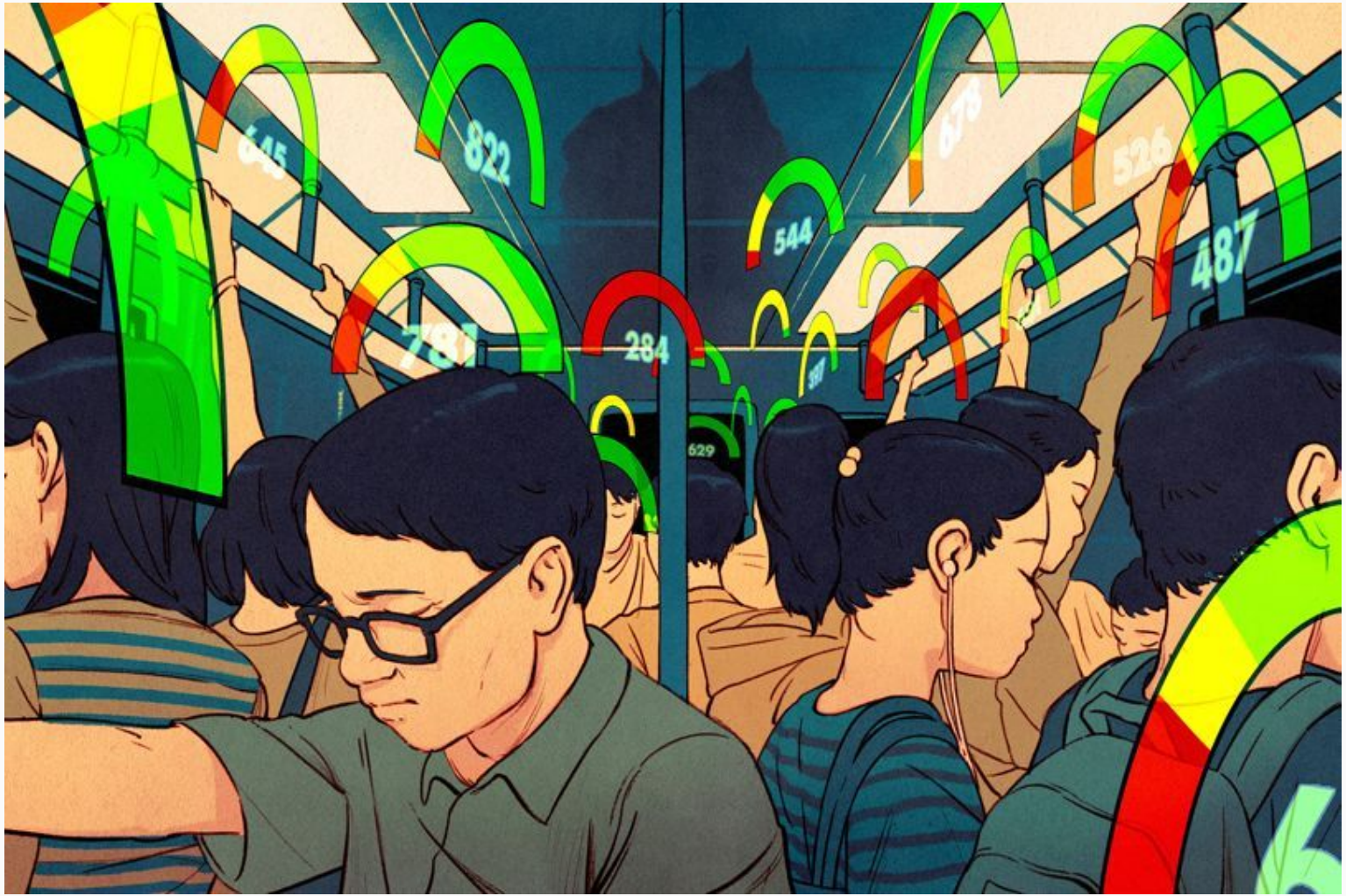
$$h_t = o_t \tanh(c_t) \quad (11)$$

Matrix Multiplication



@rameerez @rameerezcom

EL PROBLEMA ES QUE ESTAS MULTIPLICACIONES SE PUEDEN USAR PARA MUCHAS COSAS


545
822
544
526
487
781
284
629

"People with low ratings will have slower internet speeds; restricted access to restaurants and the removal of the right to travel"

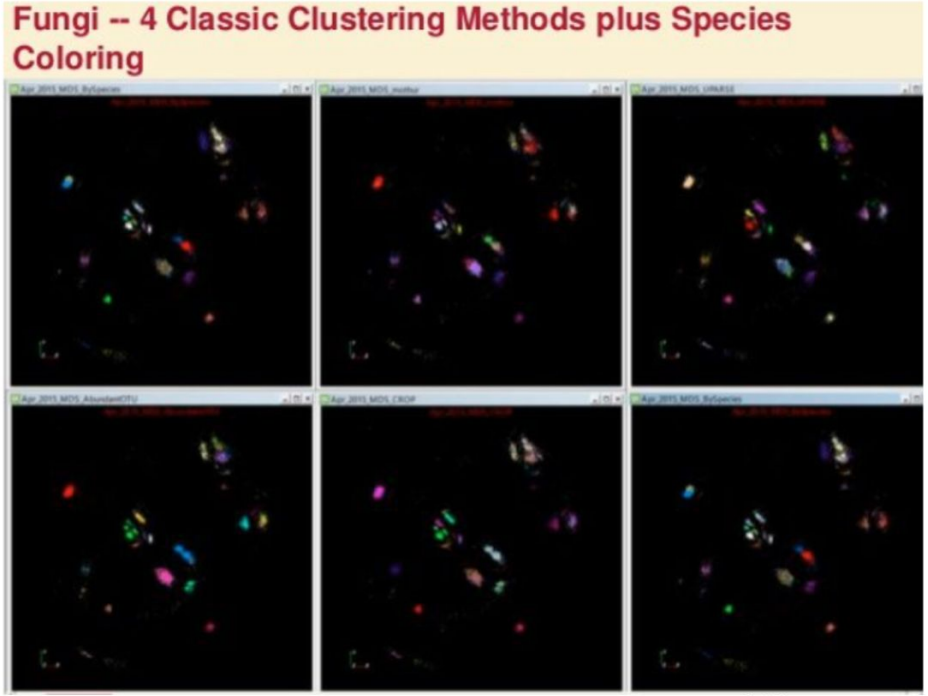
Rachel Botsman, author of 'Who Can You Trust?'

http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion

Detectar fraudes, redes criminales...

(los fraudes necesitan relaciones entre personas)

Investigación / Ciencia



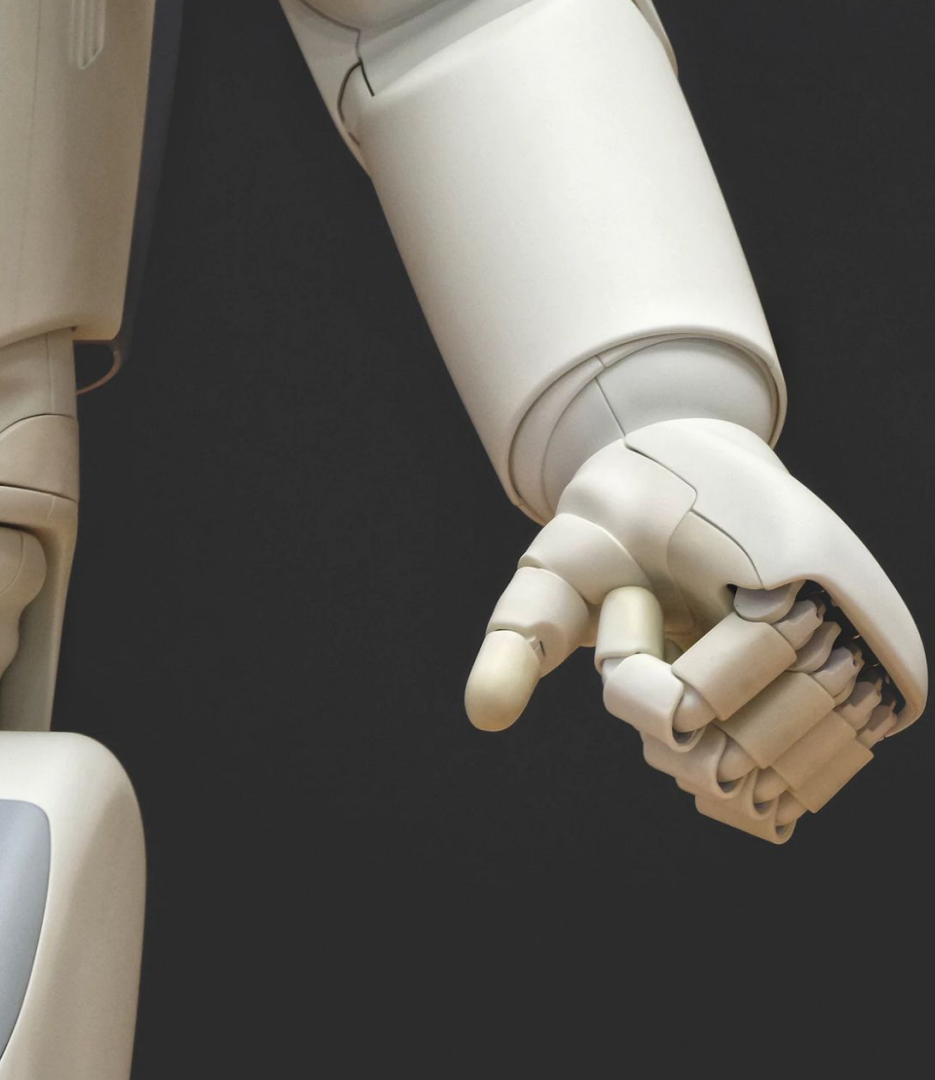
Rameerez
www.rameerez.com

Optimización de ciudades

www.rameerez.com

LA IA EN 2019 NO ES MÁS QUE MATEMÁTICAS. SOLO DEPENDE DE NOSOTROS CÓMO APLICARLAS.

OVERFITTING



CUANDO LAS MÁQUINAS NOS DEFIENDAN

Javi Ramirez

@rameerez @rameerezcom

GUARD

www.useguard.com