



# 500MB

daily data output per person, globally, in 2012

IBM

# 62 GB

daily data output per person, globally, **in 2025**

The Economist, 2017



**The  
Economist**

MAY 6TH–12TH 2017

Crunch time in France

Ten years on: banking after the crisis

South Korea's unfinished revolution

Biology, but without the cells

# The world's most valuable resource



**Data and the new rules  
of competition**



What can we learn from digital footprints?



Michal Kosinski



Michal

Home

20+



Michal Kosinski

Edit Profile

View Activity Log <sup>2</sup>



Timeline

About

Friends 1,081

Photos

More ▾



Likes

Add Likes



All Likes

Movies 4

TV Shows

Music

Books

Sports Teams

Athletes

People

Restaurants

Apps and Games

Search for movies you like



Forrest Gump ✓

Movie

Liked ▾



Kill Bill ✓

Movie

Liked ▾



Fight Club ✓

Movie

Liked ▾



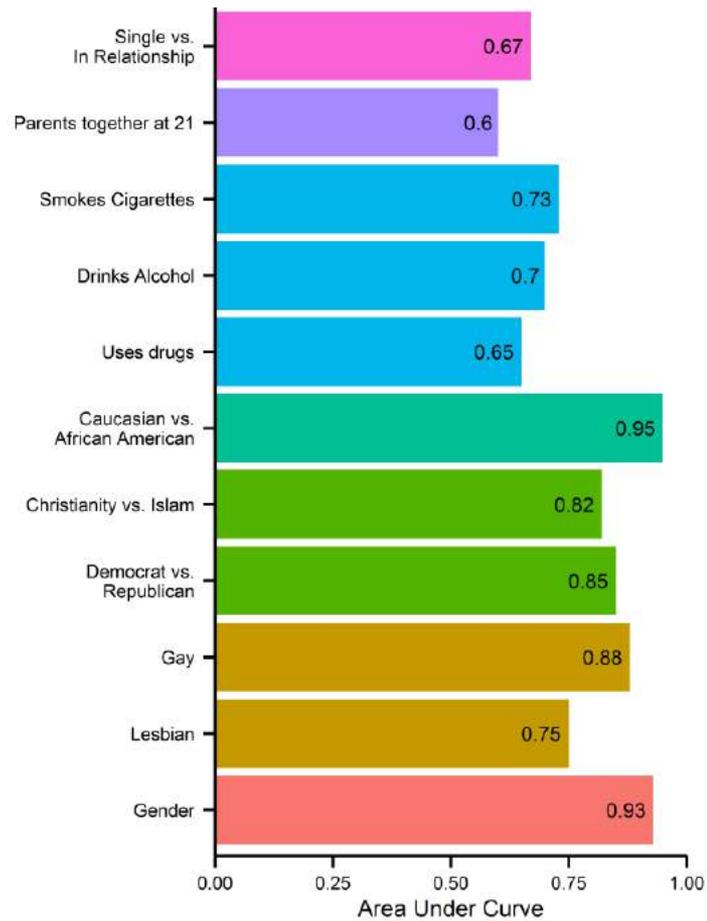
Garden State ✓

Movie

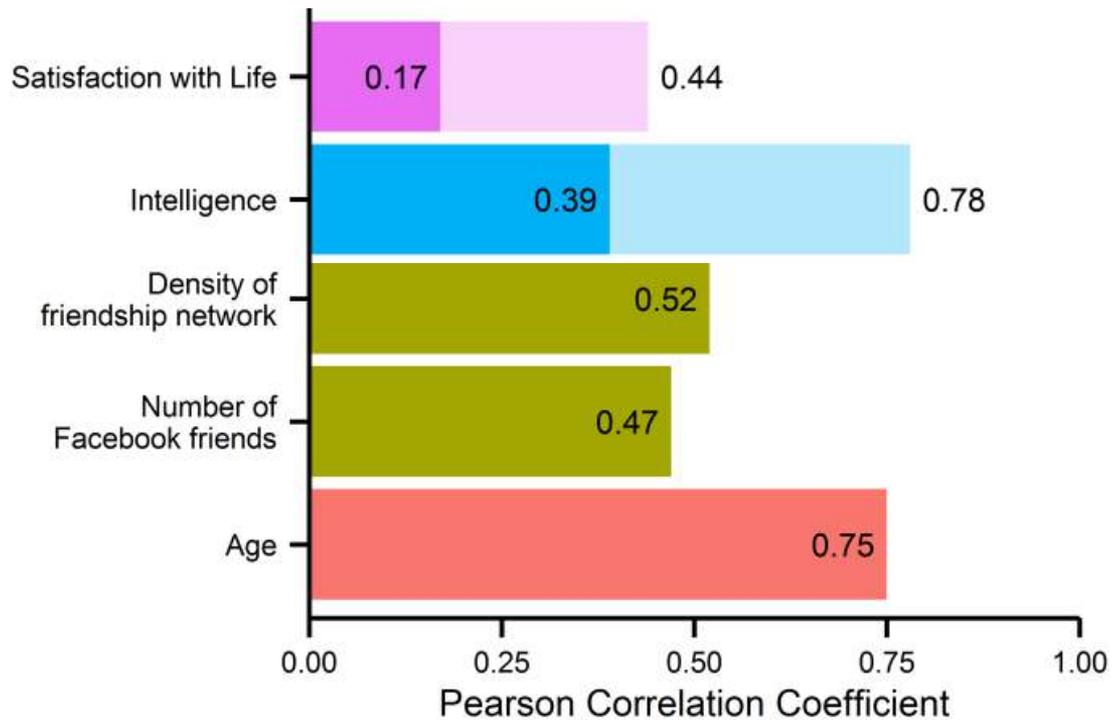
Liked ▾



[YouTube](https://www.youtube.com)



Source: Kosinski, Stillwell & Graepel (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS*.



Source: Kosinski, Stillwell & Graepel (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS*.

How does the prediction model work?

$$\begin{aligned}
\bar{J}^\mu &= \frac{\partial}{\partial \bar{x}^\nu} (\bar{\mathcal{D}}^{\mu\nu}) = \frac{\partial}{\partial \bar{x}^\nu} \left( \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \right) \\
&= \frac{\partial^2 \bar{x}^\mu}{\partial \bar{x}^\nu \partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial^2 \bar{x}^\nu}{\partial \bar{x}^\nu \partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \\
&\quad + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \frac{\partial \mathcal{D}^{\alpha\beta}}{\partial \bar{x}^\nu} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \mathcal{D}^{\alpha\beta} \frac{\partial}{\partial \bar{x}^\nu} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \\
&= \frac{\partial^2 \bar{x}^\mu}{\partial x^\beta \partial x^\alpha} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial^2 \bar{x}^\nu}{\partial \bar{x}^\nu \partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \\
&\quad + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \mathcal{D}^{\alpha\beta}}{\partial x^\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \frac{\partial \bar{x}^\rho}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial \bar{x}^\nu \partial \bar{x}^\rho} \\
&= 0 + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial^2 \bar{x}^\nu}{\partial \bar{x}^\nu \partial x^\beta} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \\
&\quad + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} J^\alpha \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \frac{\partial \bar{x}^\rho}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x^\beta \partial \bar{x}^\rho} \\
&= \frac{\partial \bar{x}^\mu}{\partial x^\alpha} J^\alpha \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] + \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \mathcal{D}^{\alpha\beta} \det \left[ \frac{\partial x^\sigma}{\partial \bar{x}^\rho} \right] \left( \frac{\partial^2 \bar{x}^\nu}{\partial \bar{x}^\nu \partial x^\beta} + \frac{\partial \bar{x}^\rho}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x^\beta \partial \bar{x}^\rho} \right)
\end{aligned}$$

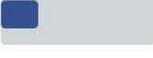
# How does the prediction model work?





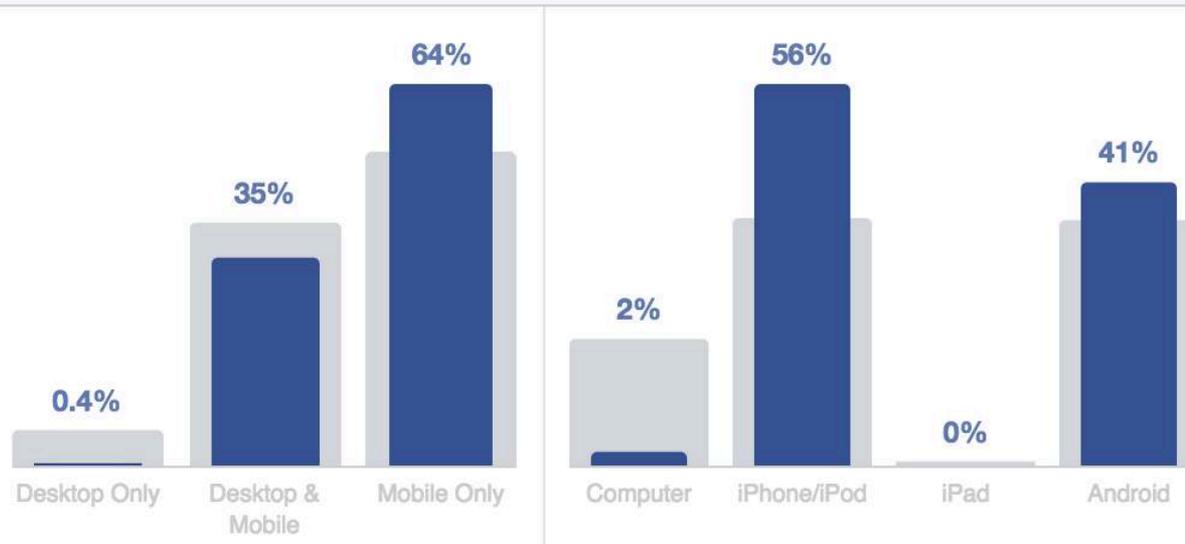
## Job Title

Likely industries where people work based on self-reported data on Facebook.

Job Title	Selected Audience	Compare <span>▼</span>
Sales	45% 	+69%
Food Preparation and Services	21% 	+67%
Personal Care	21% 	+55%
Construction and Extraction	3% 	-62%
Architecture and Engineering	2% 	-73%
Military	0.5% 	-77%
Veterans (US)	0.9% 	-79%

## Device Users

How the selected audience accessed Facebook in the last 30 days, based on user activity and environmental data.





# most strongly related to Gender:

## Male:

- Modern Warfare 2
- ESPN
- Sportscenter
- Band Of Brothers
- Starcraft
- Deadliest Warrior
- Dos Equis
- Red Vs Blue
- X Games
- Bruce Lee

## Female:

- Tv Fanatic
- Chiq
- Gillette Venus
- Shoedazzle
- Bebe
- Proud To Be A Mom
- Covergirl
- Wet Seal
- Aerie By American
- Eagle Mall World



# BIG5 Personality Traits:

**O**penness

**C**onscientiousness

**E**xtraversion

**A**greeableness

**N**euroticism



# most strongly related to Openness:

## Conservative:

- NASCAR
- Austin Collie
- Monster-In-Law
- I don't read
- Justin Moore
- ESPN2
- Farmlandia
- The Bachelor
- Oklahoma State University
- Teen Mom 2

## Liberal:

- Oscar Wilde
- Charles Bukowski
- Sylvia Plath
- Leonardo Da Vinci
- Bauhaus
- DMT: The Spirit Molecule
- American Gods
- John Waters
- Plato
- Leonard Cohen

# Agreeableness

## Disagreeable:

- I Hate Everyone
- I Hate You
- I Hate Police
- Friedrich Nietzsche
- Timmy South Park
- Atheism / Satanism
- Prada
- Sun Tzu
- Julius Caesar
- Knives

## Agreeable:

- Compassion International
- Logan Utah
- Jon Foreman
- Redeeming Love
- Pornography Harms
- The Book Of Mormon
- Circles Of Prayer
- Go To Church
- Christianity
- Marianne Williamson

# Extroversion

## Reserved:

- RPGs
- fanfiction.net
- Programming
- Anime
- Manga
- Video Games
- Role Playing Games
- Minecraft
- Voltaire
- Terry Pratchett

## Outgoing:

- Beerpong
- Michael Jordan
- Dancing
- Socializing
- Chris Tucker
- I Feel Better Tan
- Modeling
- Cheerleading
- Theatre
- Flip Cup

# Neuroticism

## Relaxed:

- Business Administration
- Getting Money
- Parkour
- Track & Field
- Skydiving
- Mountain Biking
- Soccer
- Climbing
- Physics / Engineering
- 48 Laws Of Power

## Neurotic:

- Sometimes I Hate Myself
- Emo
- Girl Interrupted
- So So Happy
- The Addams Family
- Vocaloid
- Sixbillionsecrets.com
- Vampires Everywhere
- Kurt Donald Cobain
- Dot Dot Curve

# Conscientiousness

## Spontaneous:

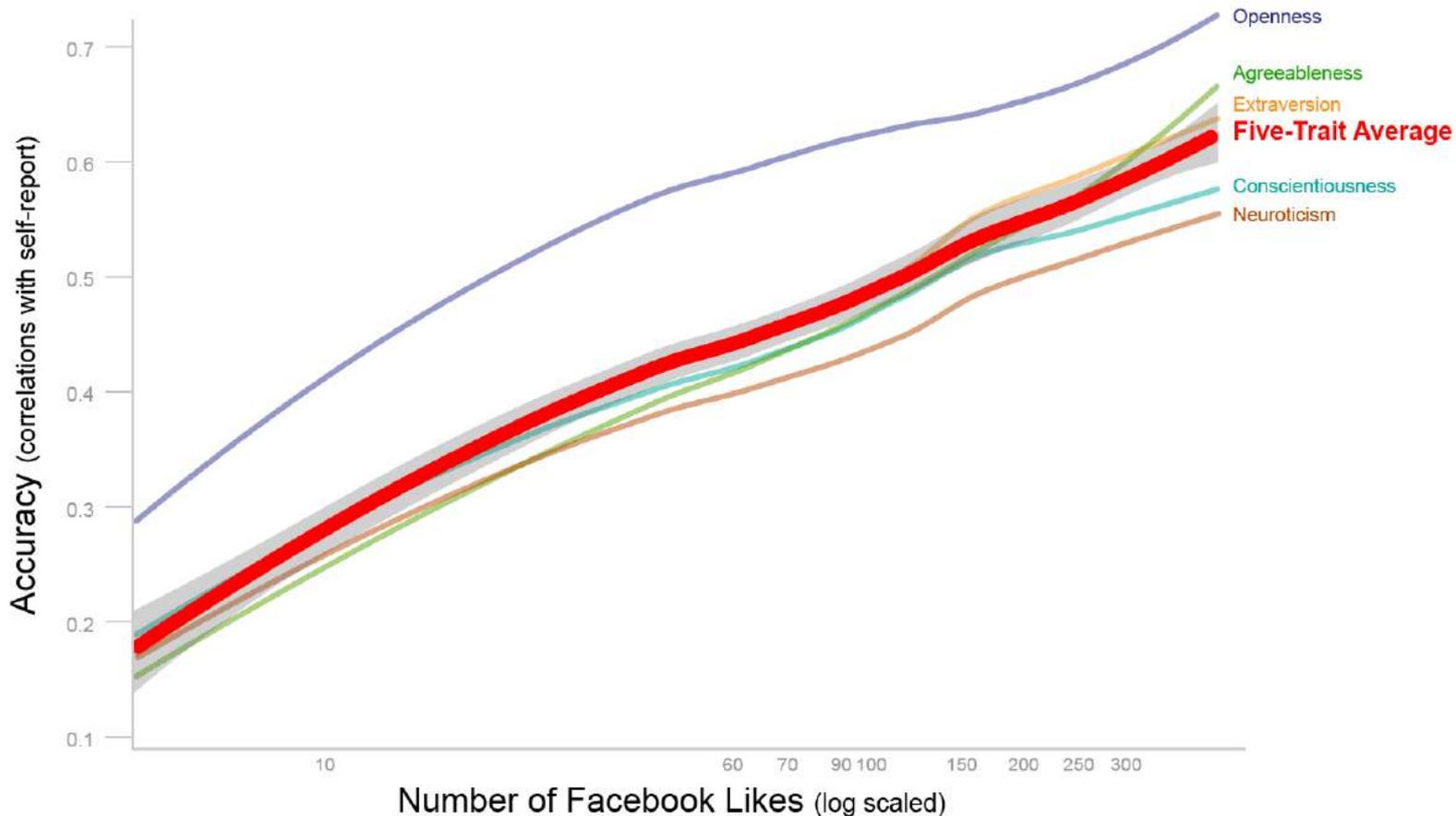
- Wes Anderson
- Bandit Nation
- Omegle
- Vocaloid
- Serial Killer
- Screamo
- Anime
- Vamplets
- Join If Ur Fat
- Not Dying

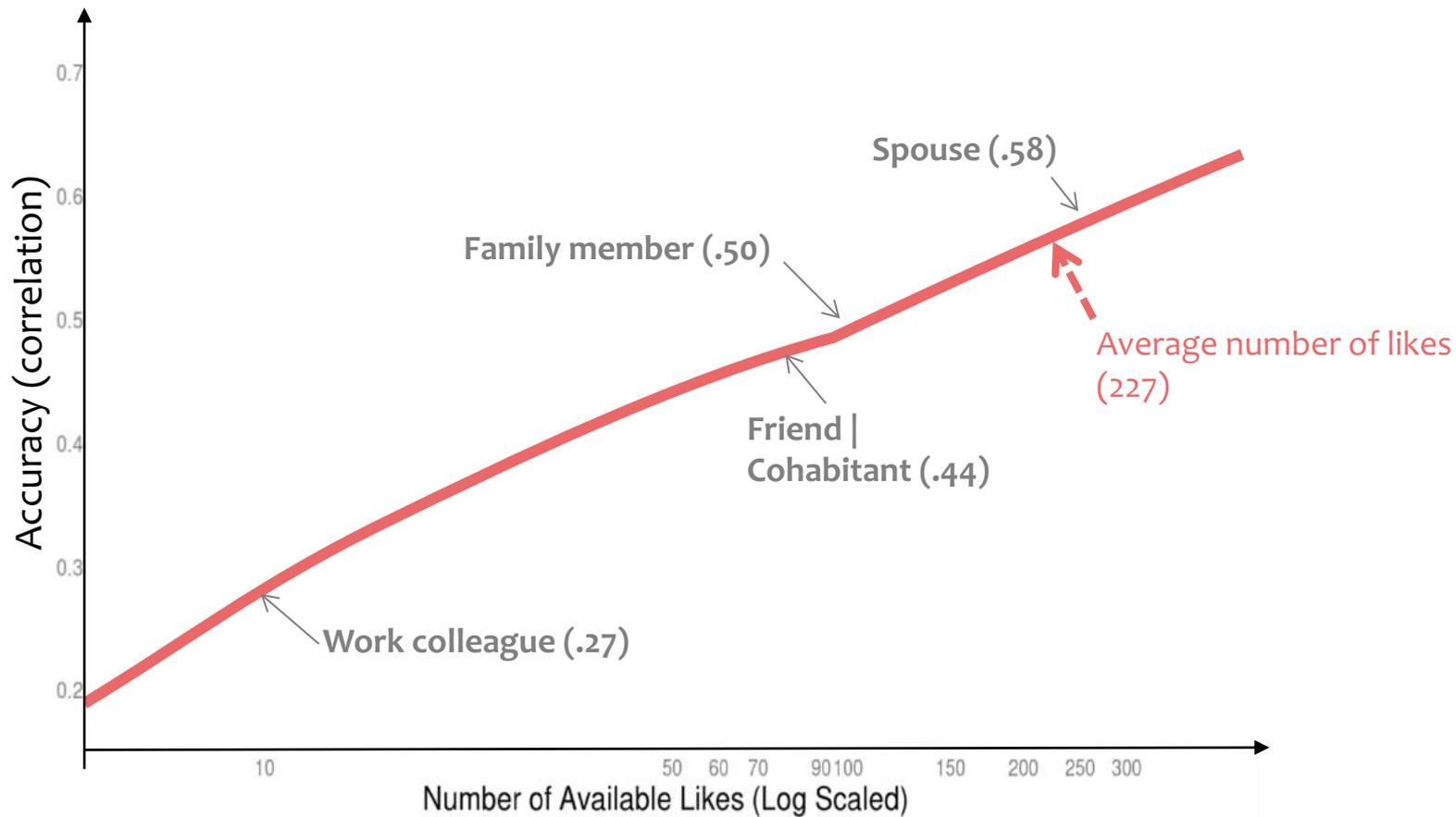
## Well organized:

- Law Officer
- National Law Enforcement  
lowfares.com
- Accounting
- Foursquare
- Emergency Medical Services
- Sunday Best
- Kaplan University
- Glock Inc.
- Mycalendar 2010

How accurate are psychological predictions?

# Predicting Personality from Facebook Likes





Youyou, Kosinski, Stillwell (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings Of The National Academy Of Sciences*.

# Apply Magic Sauce

**PredictionAPI**

Apply Magic Sauce translates individuals' digital footprints into detailed psychological profiles

- for research
- for business
- for you



## Digital footprints

- Facebook LikeIDs >
- Facebook statuses >
- [Redacted]



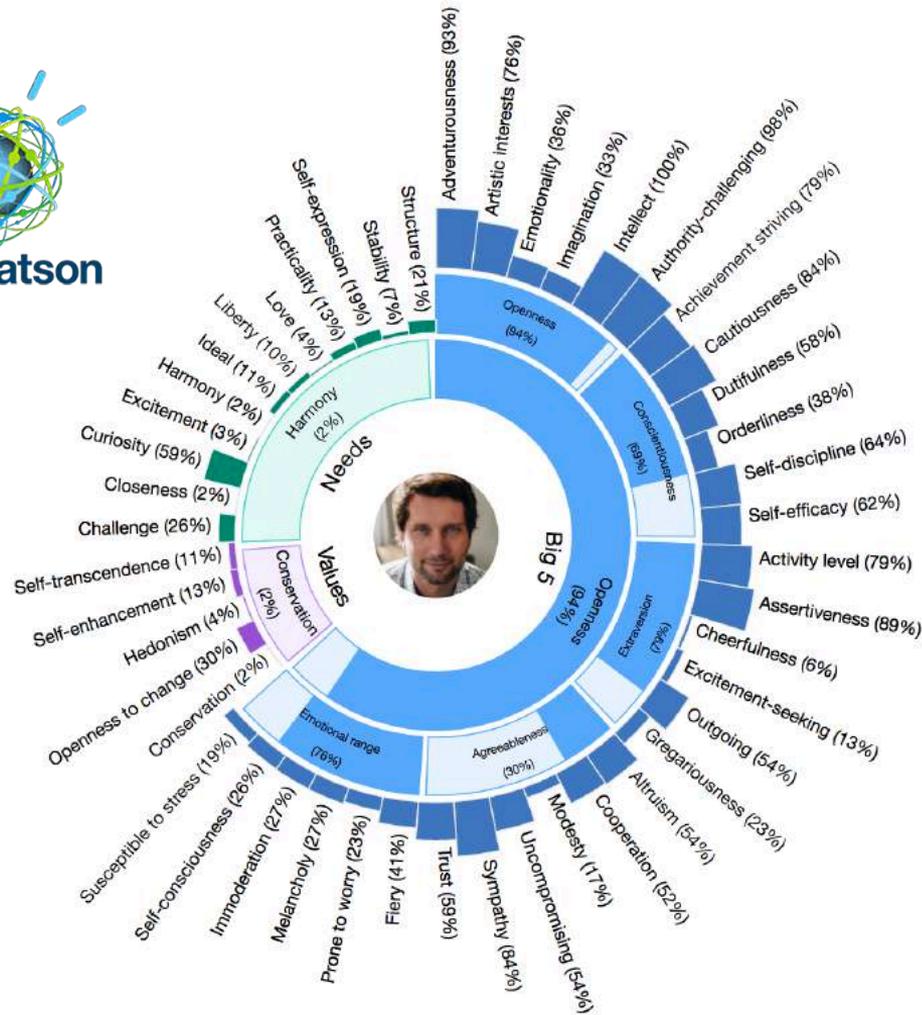
## Individual profiles

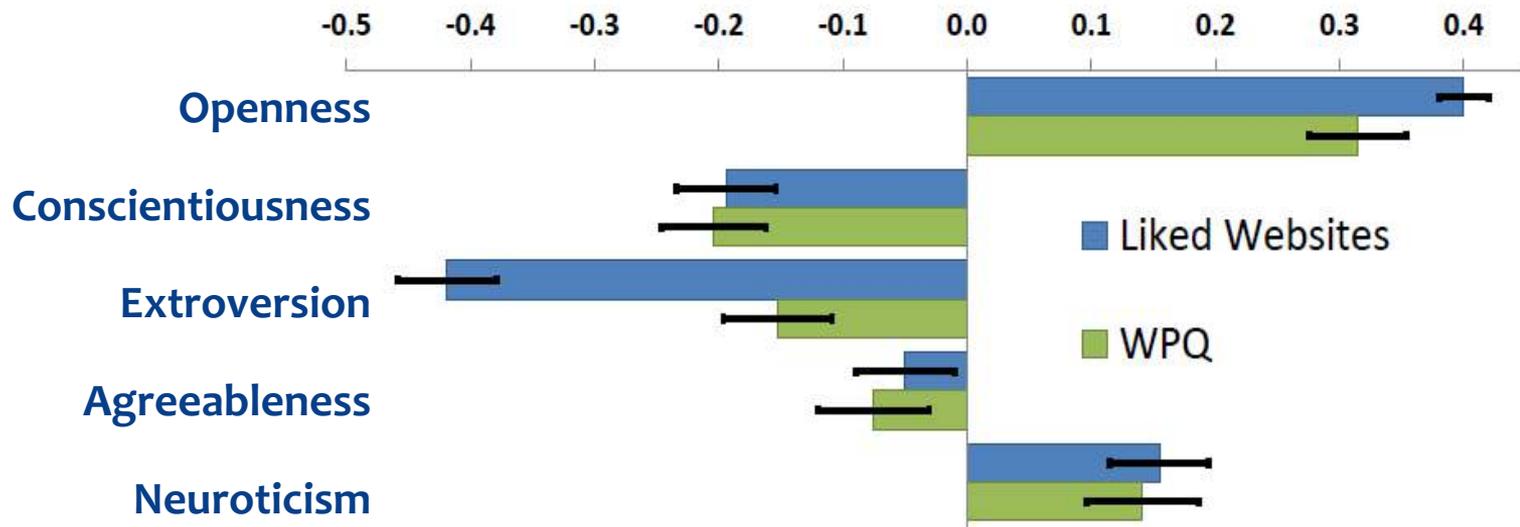
- Psychographics
  - BIG5 Personality
  - Intelligence
  - Life Satisfaction



# Beyond Facebook Likes



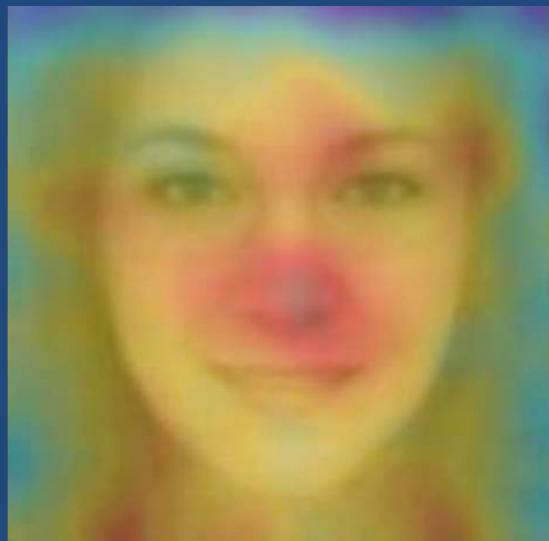




Personality profile of [www.deviantart.com](http://www.deviantart.com) audience

Kosinski et al. (2013). Manifestations of User Personality In Website Choice And Behaviour On Online Social Networks. *Machine Learning*.

# The science of face reading

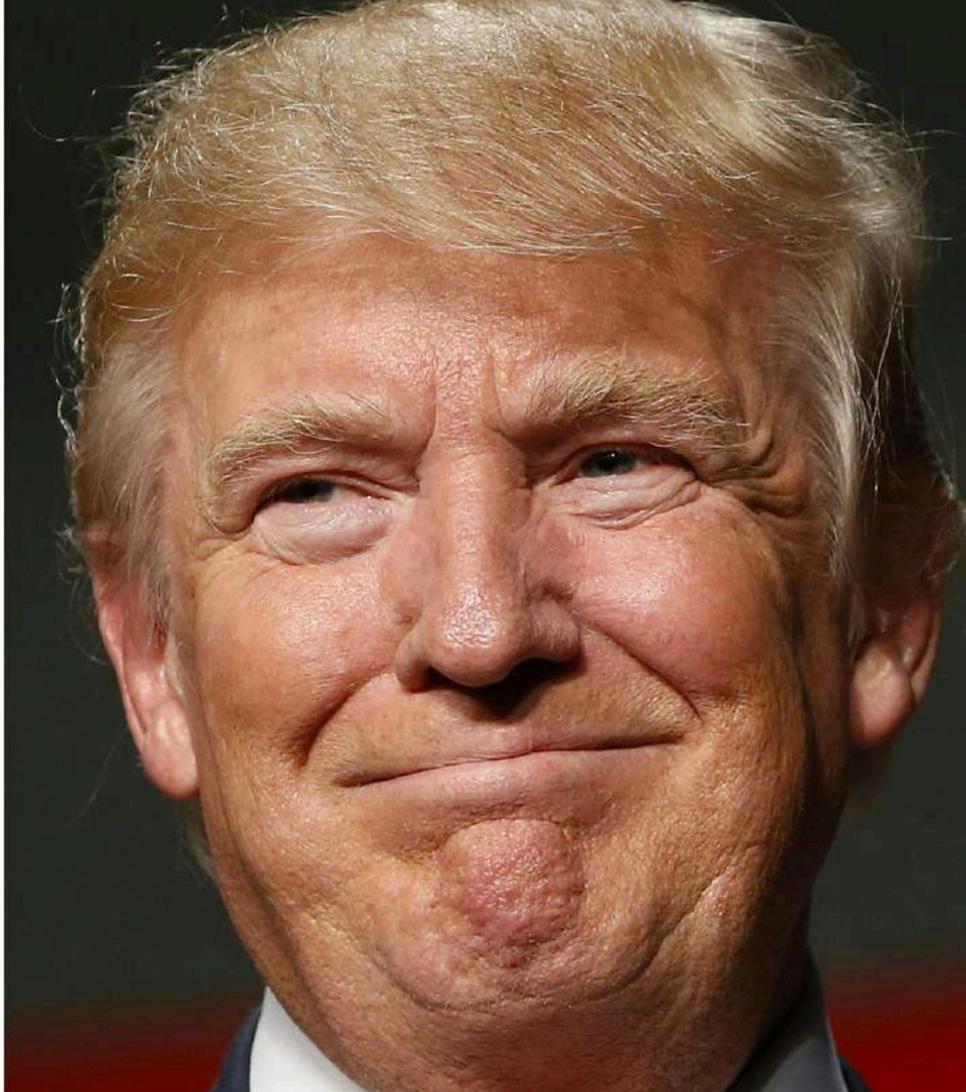








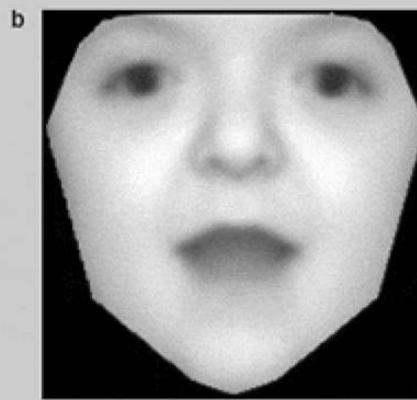




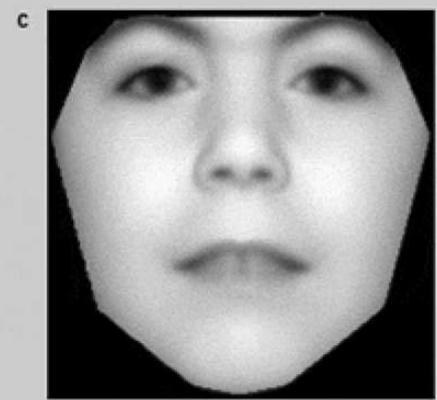




Angelman



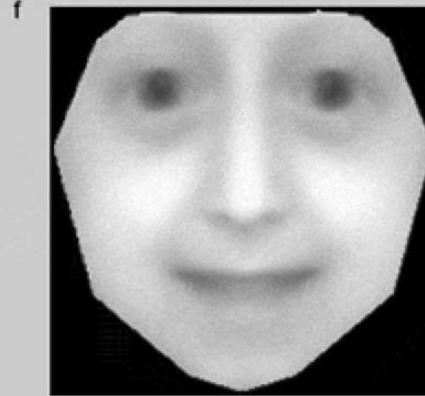
Apert



Cornelia de Lange



Fragile X



Progeria



Treacher-Collins

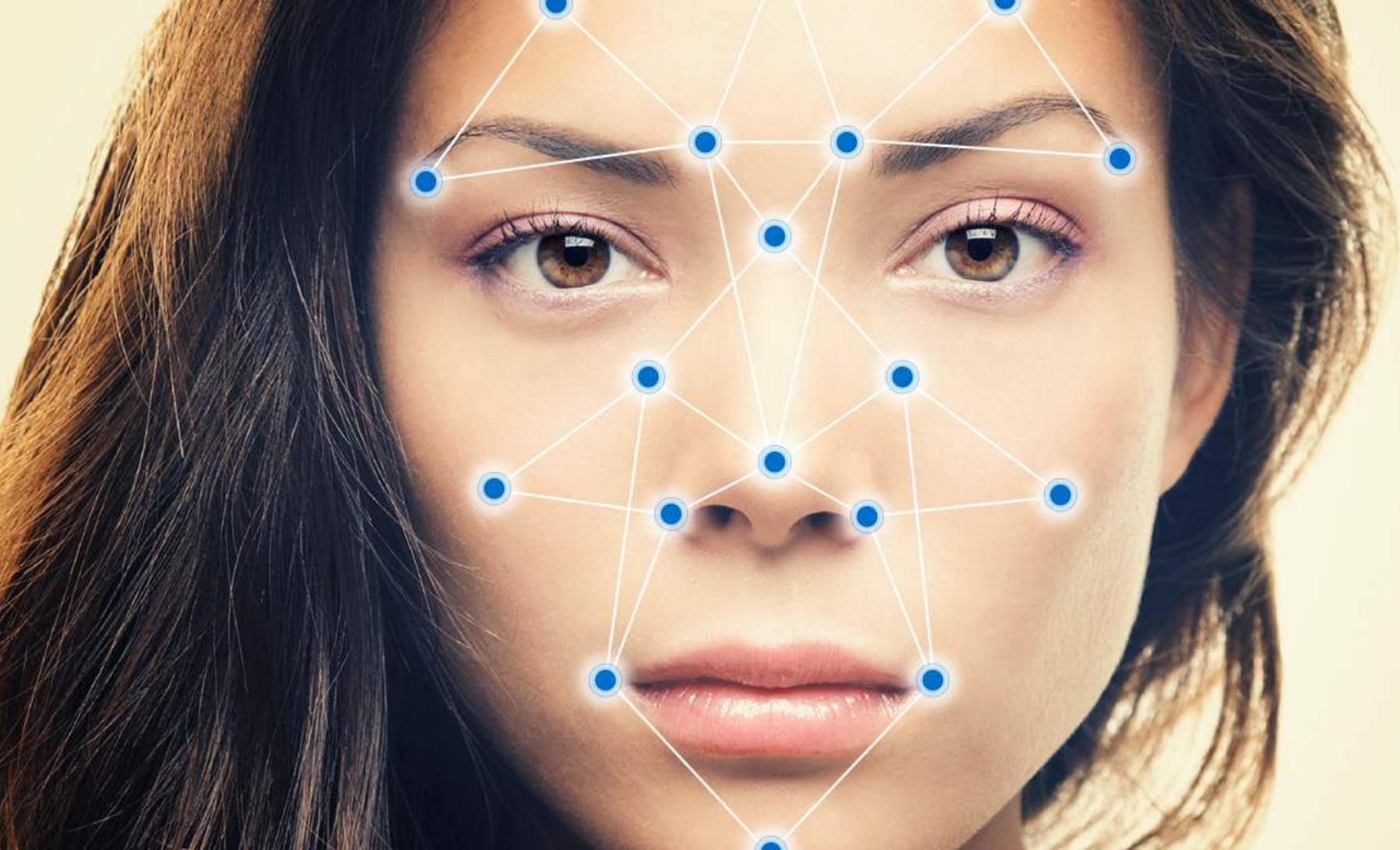
Ferry, Q., Steinberg, J., Webber, C., FitzPatrick, D. R., Ponting, C. P., Zisserman, A., & Nellåker, C. (2014). Diagnostically relevant facial gestalt information from ordinary photos. *eLife*, 2014(3).



Samochowiec, J., Wanke, M., & Fiedler, K. (2010). Political ideology at face value. *SPPS*, 1(3), 206–213.

## Good proxy for:

- Genetic factors
- Hormonal factors
- Developmental/environmental factors
- Cultural factors & SES

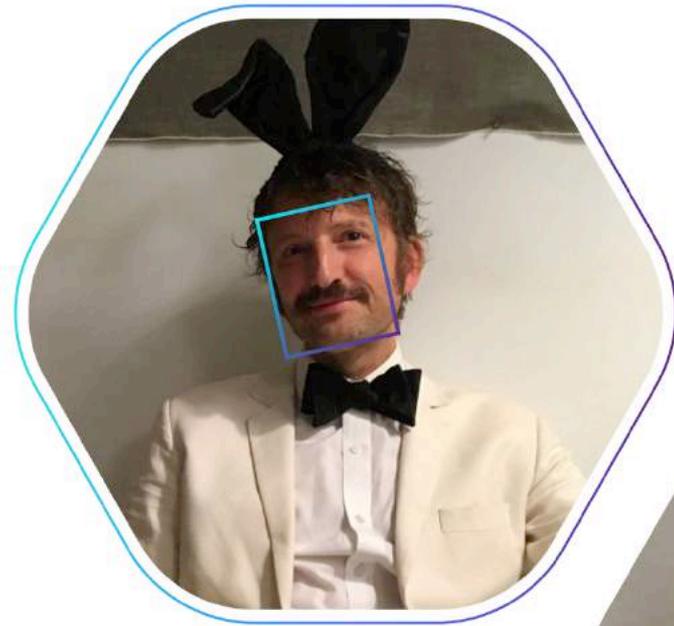




2. Images are cropped and resized



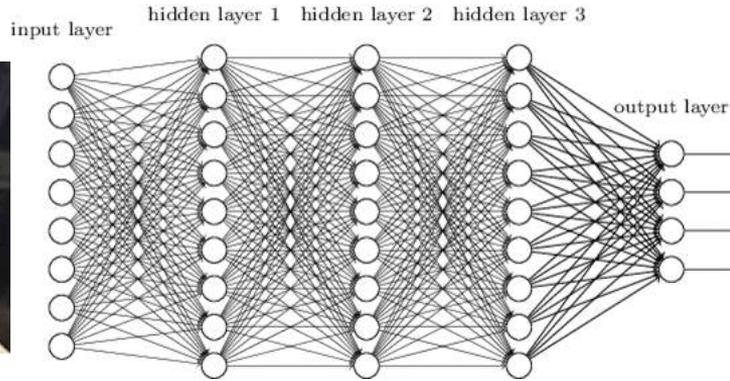
3. VGG-Face is used to turn facial images into face-vectors.



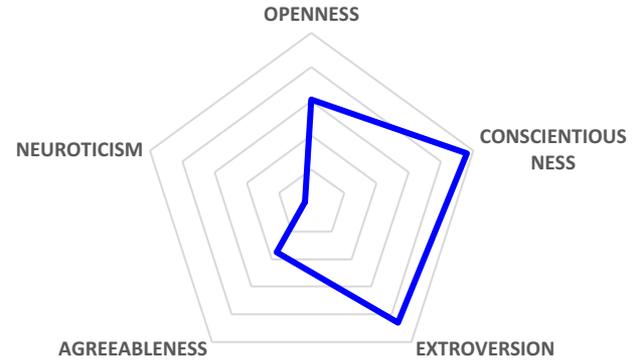
1. Faces are detected on the images using F++ algorithm



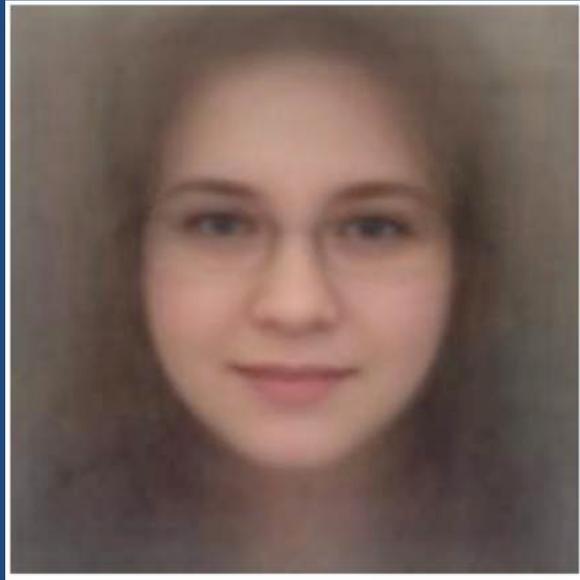
## Deep neural network



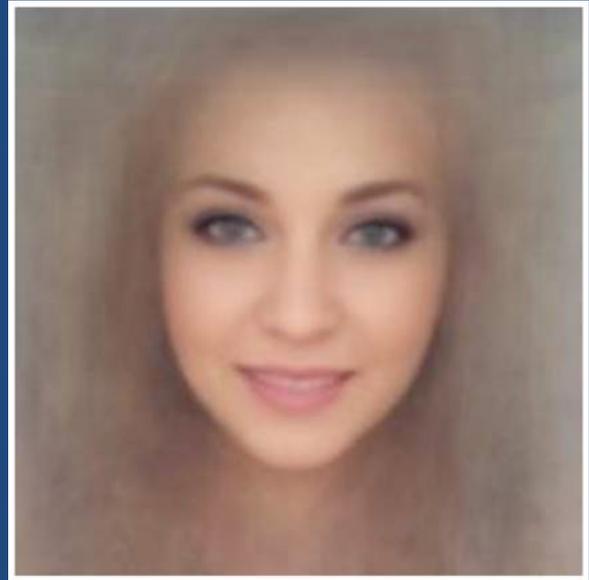
## Personality profile



AI “dreaming” about the faces of introverts and extroverts



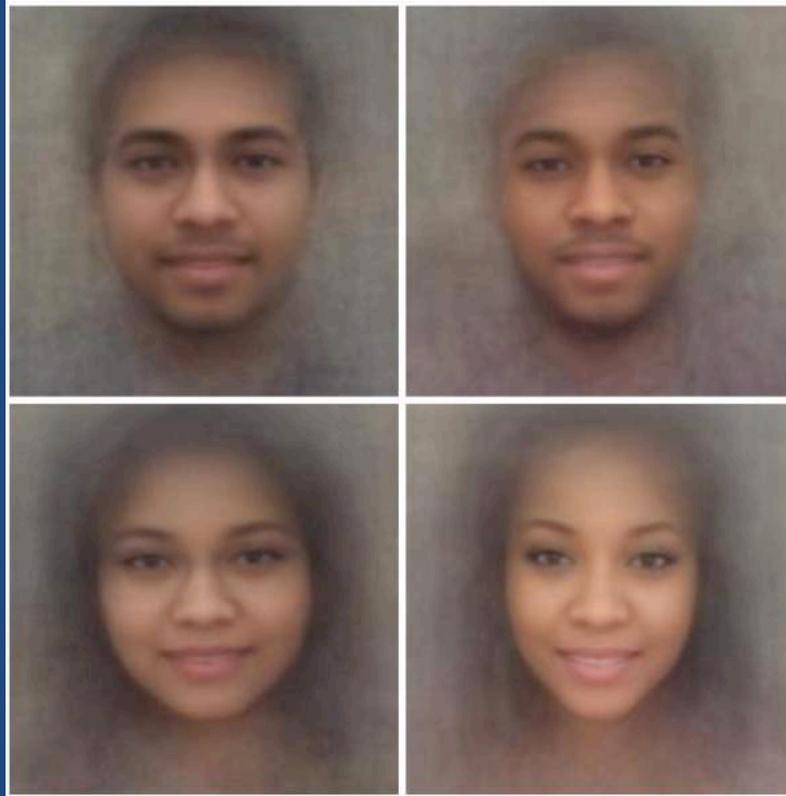
**Most  
Introverted**



**Most  
Extroverted**

*African American*

**Introverted**



**Extroverted**

*Asian American*

**Introverted**



**Extroverted**

*Openness? Conscientiousness? Agreeableness? Neuroticism?*

**Disagreeable**



**Agreeable**

*Openness? Conscientiousness? Neuroticism?*

**Spontaneous**



**Conscientious**

*Openness? Neuroticism?*

**Conservative**



**Liberal**

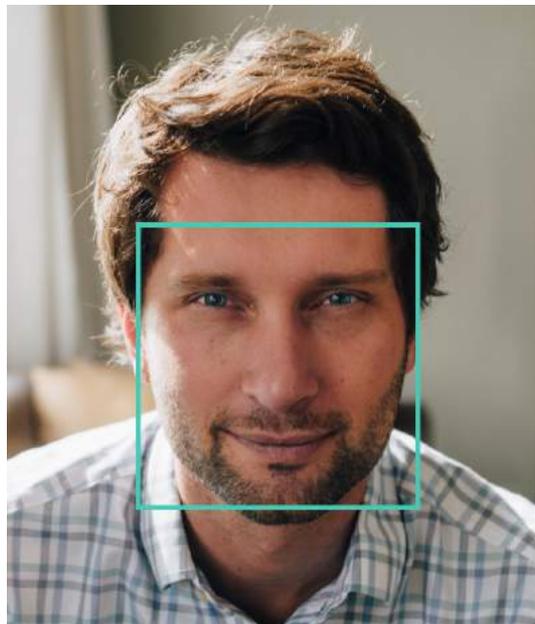
**Emotionally Stable**



**Neurotic**



# Cognitive Services



Detection Result:  
JSON:

```
[
  {
    "faceRectangle": {
      "left": 558,
      "top": 686,
      "width": 899,
      "height": 899
    },
    "scores": {
      "anger": 0.00009627592,
      "contempt": 0.00526537234,
      "disgust": 0.00002383919,
      "fear": 0.00000227332748,
      "happiness": 0.6947198,
      "neutral": 0.299442232,
      "sadness": 0.0004173915,
      "surprise": 0.00003286575
    }
  }
]
```



# The End of Privacy





TWEETS

234

FOLLOWING

175

FOLLOWERS

4,559

LIKES

69

MOMENTS

0

Edit profile

## Michal Kosinski

@michalkosinski

Professor at Stanford University  
Graduate School of Business.  
Computational Psychologist and Big  
Data Scientist.

Stanford

[michalkosinski.com](http://michalkosinski.com)

Joined June 2009

Tweets

Tweets & replies

Media



Pinned Tweet



**Michal Kosinski** @michalkosinski · Mar 25

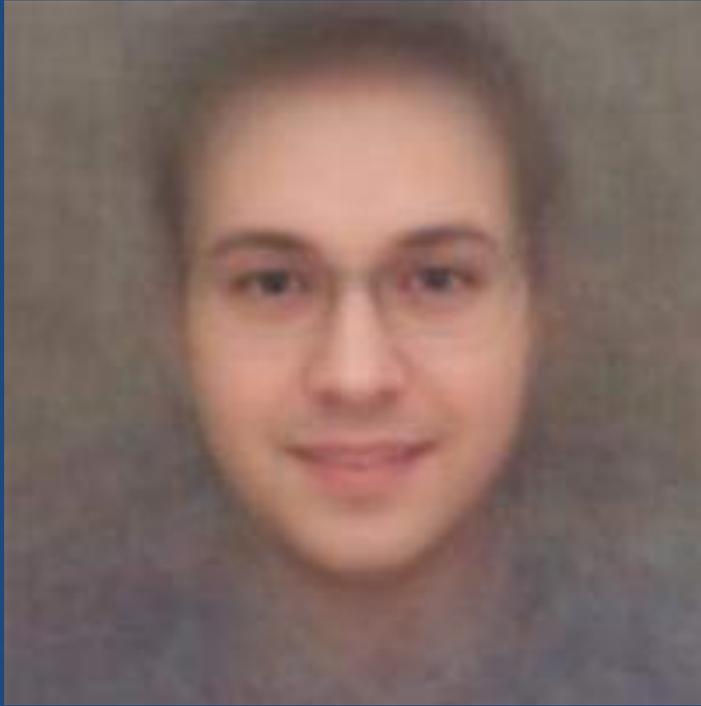
The End o Privacy @cebit



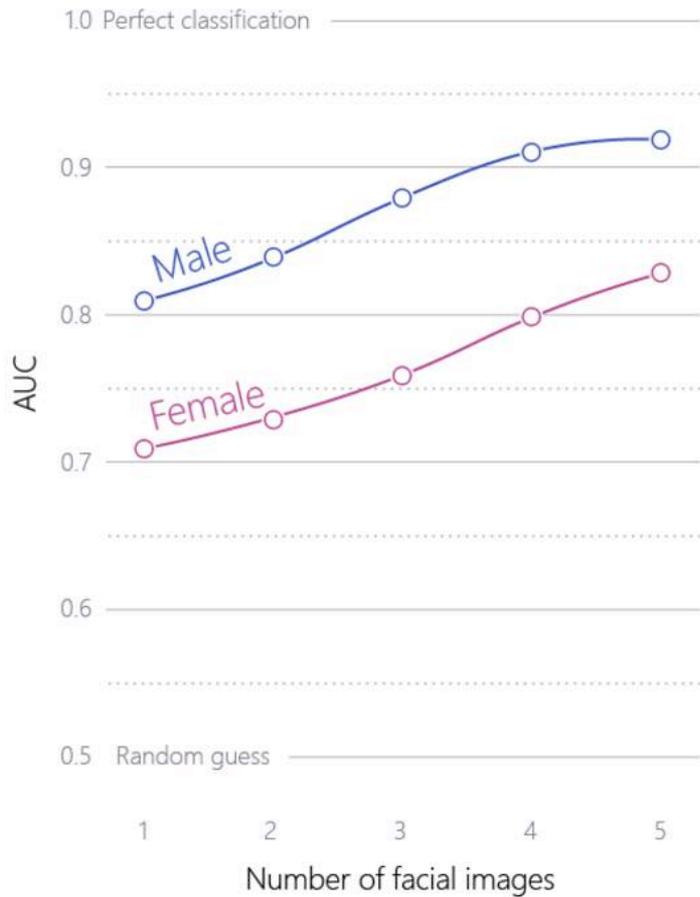
**Keynote "The End of Privacy", Dr. Michal Kosinski**

CeBIT Global Conferences - 23 March 2017: Keynote  
"The End of Privacy" / Dr. Michal Kosinski, Stanford  
University, United States (USA) <http://bit.ly/2nqZrDv>

[youtube.com](http://youtube.com)



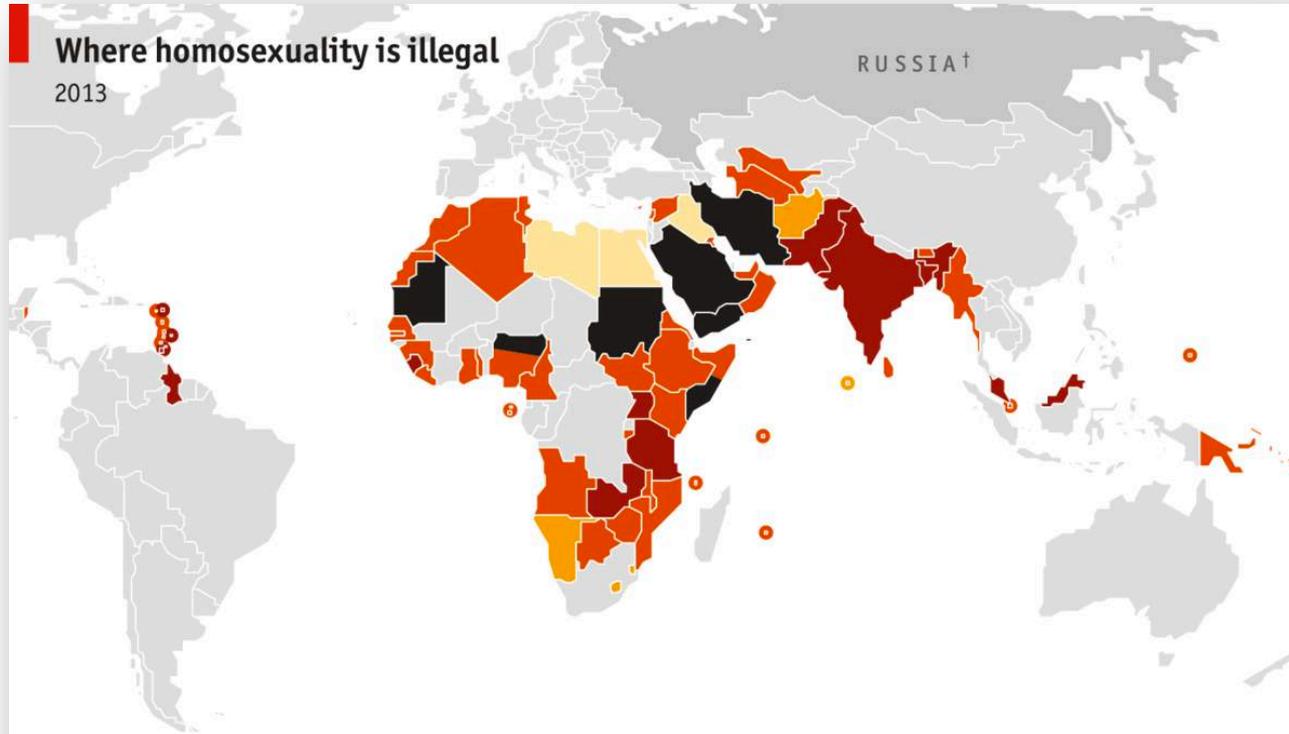
AI 'dreaming' about faces of homosexual and heterosexual males



Sexual orientation: Classification accuracy vs. the number of facial images per person.

# Where homosexuality is illegal

2013



Death penalty	14 years to life	Up to 14 years	No indication of length	Unclear*
---------------	------------------	----------------	-------------------------	----------

Source: ILGA

\*Legislation not specifically homophobic but can be used as such/persecution by religious courts  
†"Propaganda law" restricting freedom of expression and association





**Thank you!**

[michalk@stanford.edu](mailto:michalk@stanford.edu)