Artificial Intelligence, Diversity and Ethics

A European perspective

WENET
INTERNET OF US

U·HOPPER

Why me?

A long journey

Me (in a nutshell)

Dad of two

CEO at U-Hopper

CRO & co-founder at Thinkinside

PhD in TLC engineering

~200 scientific publications, 6,500+ citations

4 patents

3 companies co-founded (2 shut down)
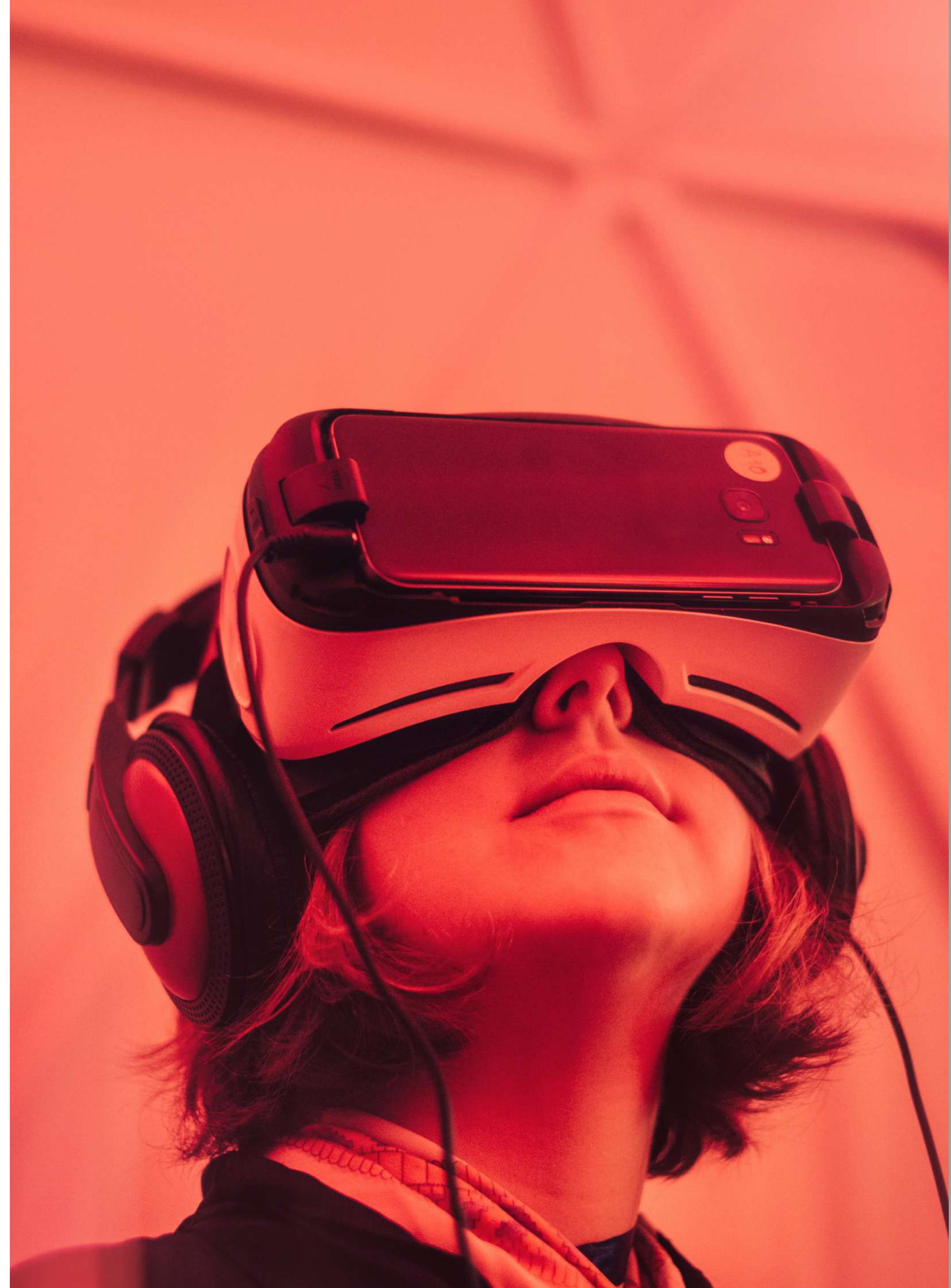
Tech MegaTrends

what is a 'megatrend'?

"Mega trends are global, sustained and macro economic forces of development that impacts business, economy, society, cultures and personal lives thereby defining our future world and its increasing pace of change." (Frost & Sullivan)
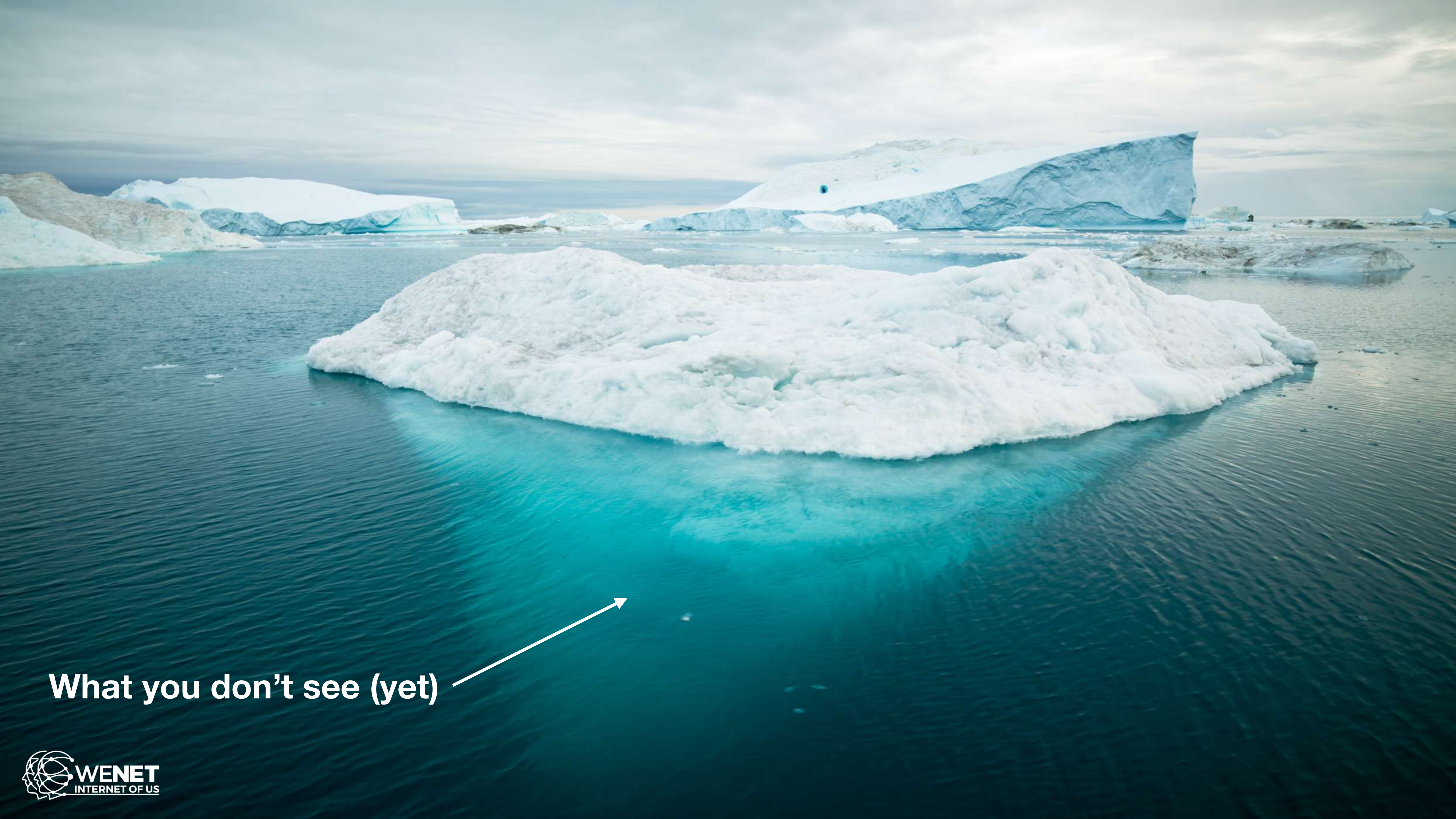
WENET
INTERNET OF US

UhoPPer

focus of this talk:

technology

megatrends

UhoPPer

the common context

Atoms
vs
Bits

WENET
INTERNET OF US

What you don't see (yet)

WENET
INTERNET OF US

trend #1: IoT

20 billions connected devices

(Ericsson)

CAGR 21%

(Ericsson)

(read it: doubling ~ every 3 .5 years)

WENET
INTERNET OF US

UhoPPer

Everything is data

You can sense the world

The **data** explosion

trend #2: Big Data

2.5 quintillion bytes of data per day

(IBM)

WENET
INTERNET OF US

UhoPPer

by 2025 **163 zettabytes** per year
(IDC)

[1 zettabyte = 10^{21} bytes]

WENET
INTERNET OF US

UhoPPer

One day in data

4.3B users connected

400M Instagram stories

6.3B Google searches

6.5B videos watched on YouTube

26B text messages

350M photos on FB

360M tweets

1.4M Wikipedia page edits
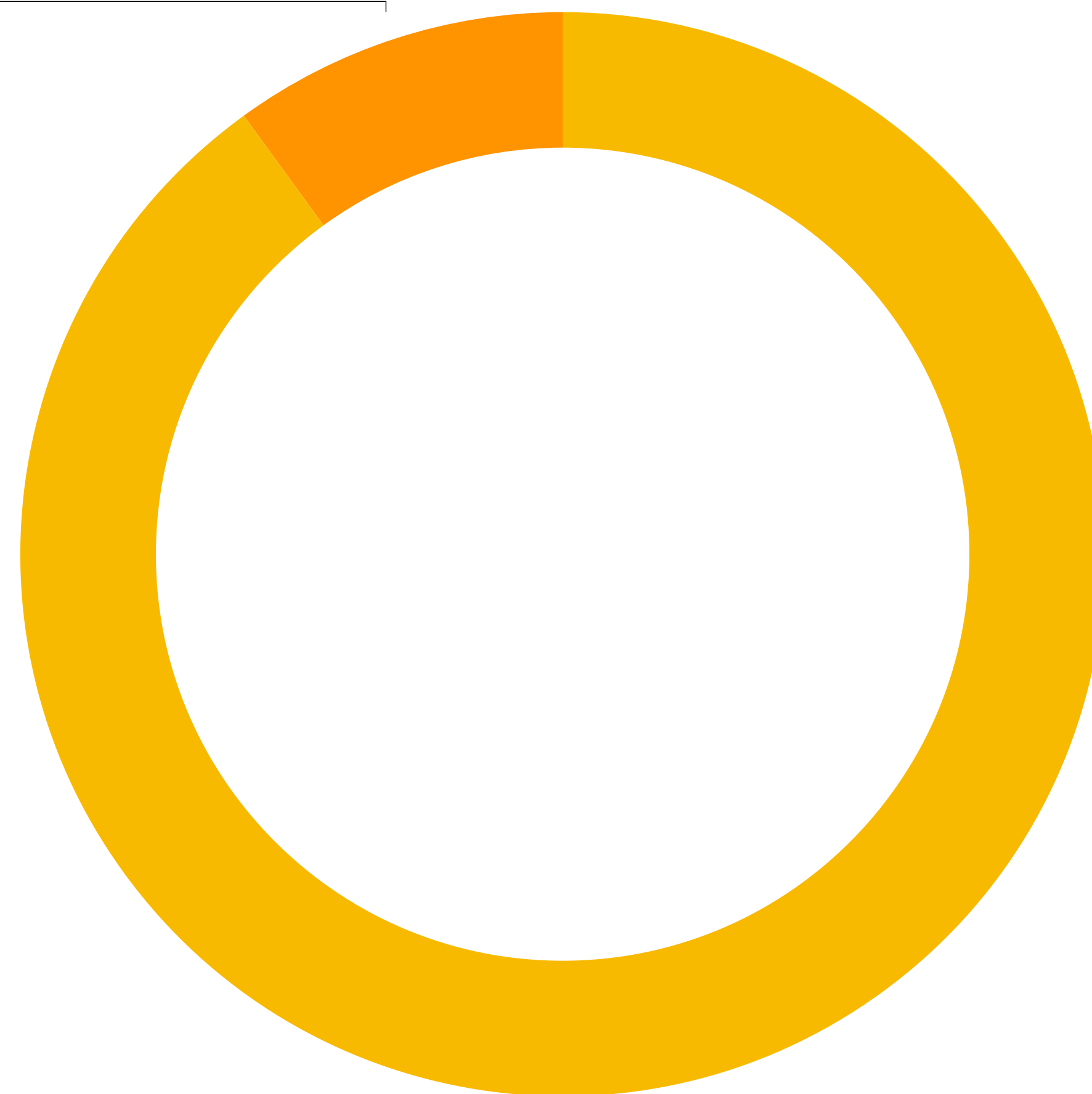
2B Tinder swipes

14M Uber rides

Sources: DataNeverSleeps, Wikimedia, Facebook, Google

WENET INTERNET OF US

UhoPPer

90% of all data generated in the

last 2 years (IBM)

WENET
INTERNET OF US

All the data in the world

Rest of history (..-2017)
10%

2018-2019
90%

Uhopper

WENET
INTERNET OF US

Volume
Velocity
Variety

WENET
INTERNET OF US
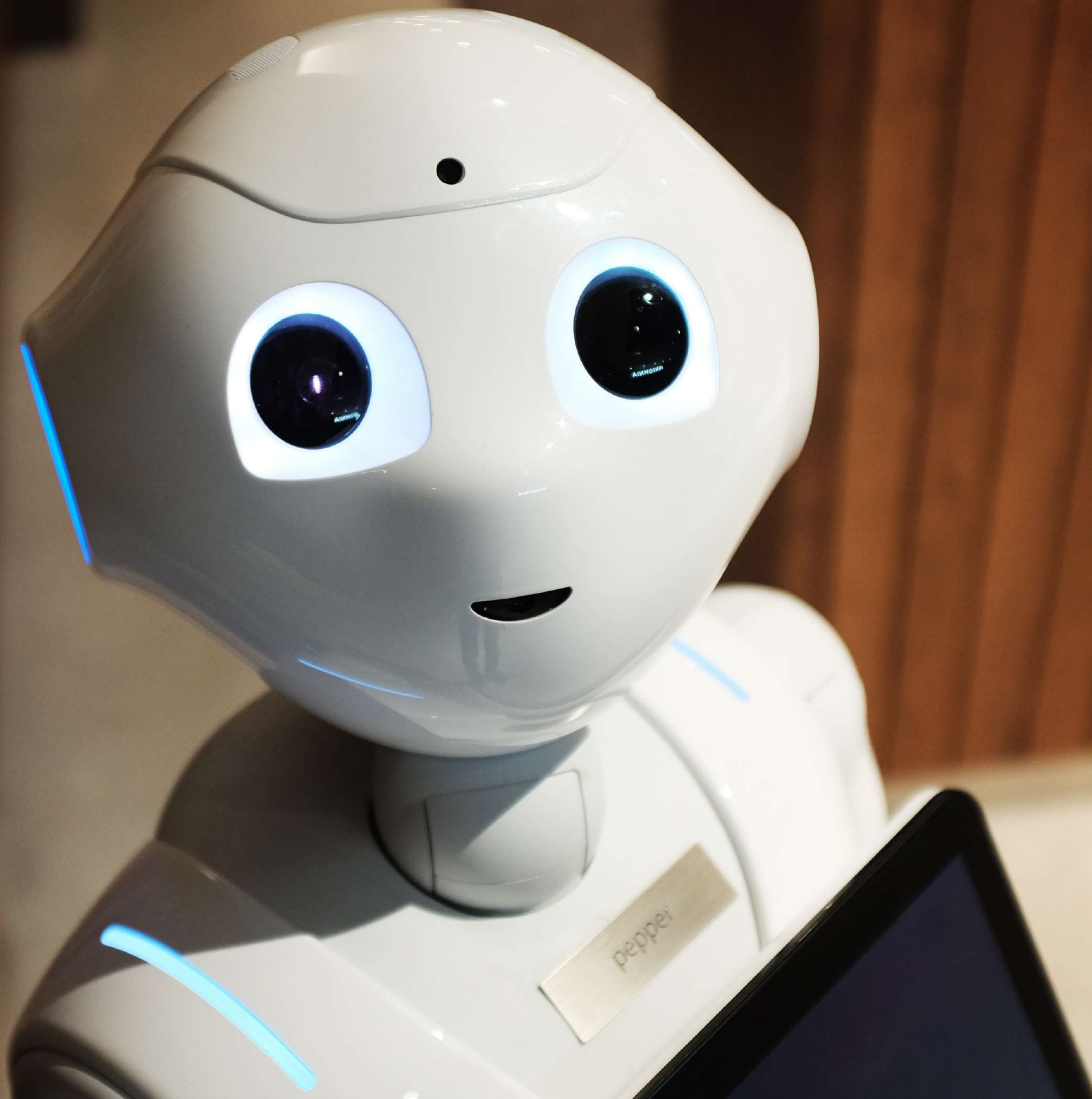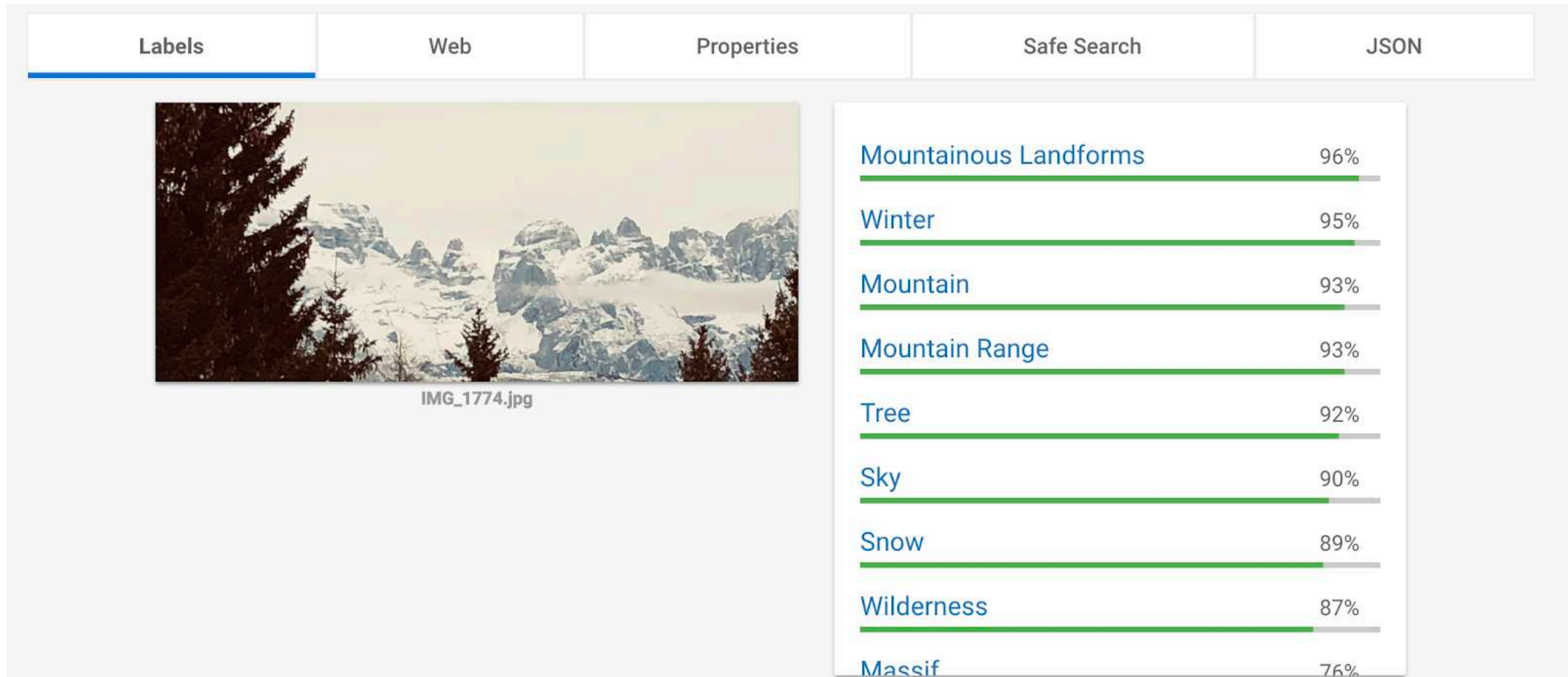
"And now that we have all big data?"

trend #3: Artificial Intelligence

NO, not this AI

The 'humble' (boring) AI

AI

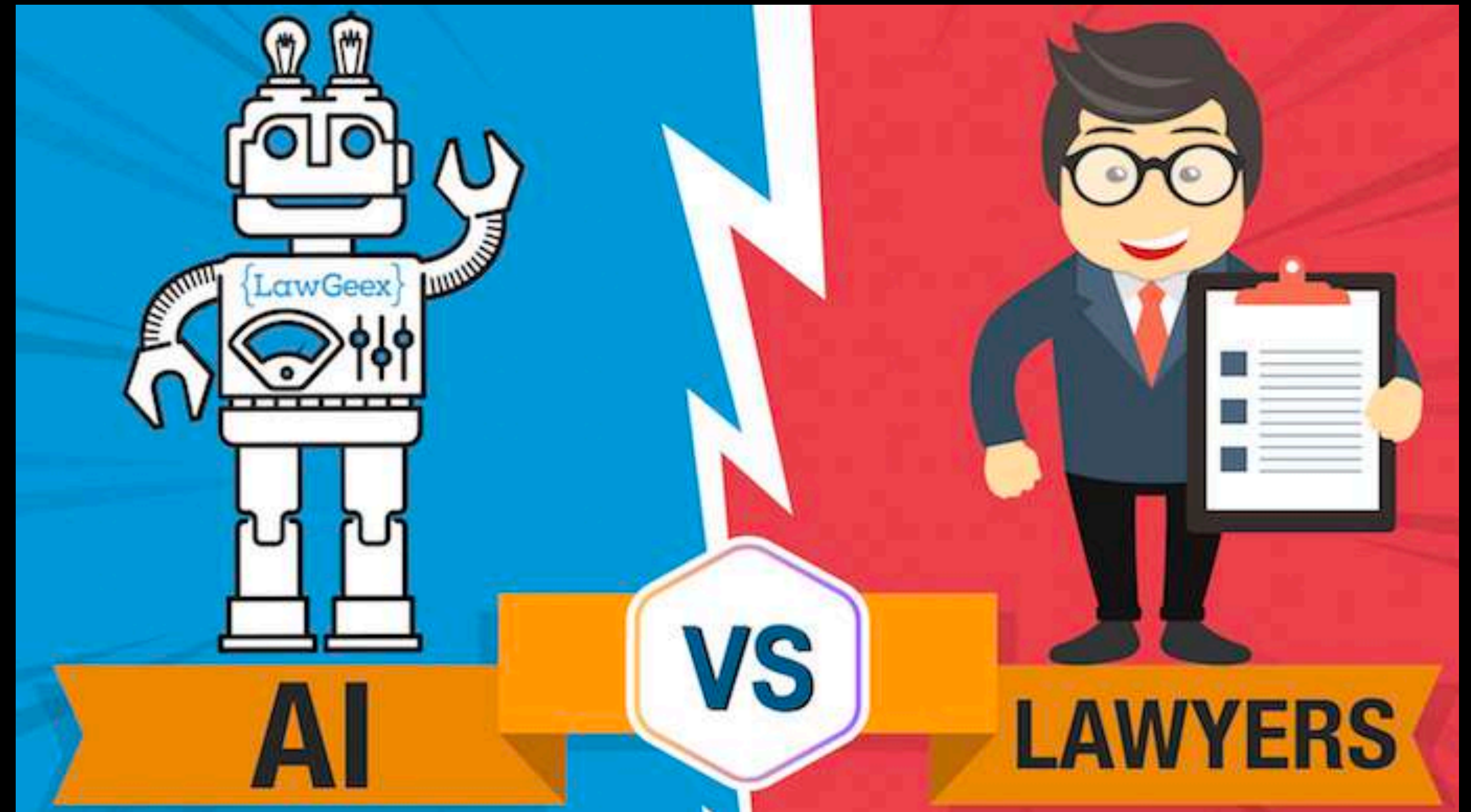**Making sense of data at scale and in an automated fashion**

WENET
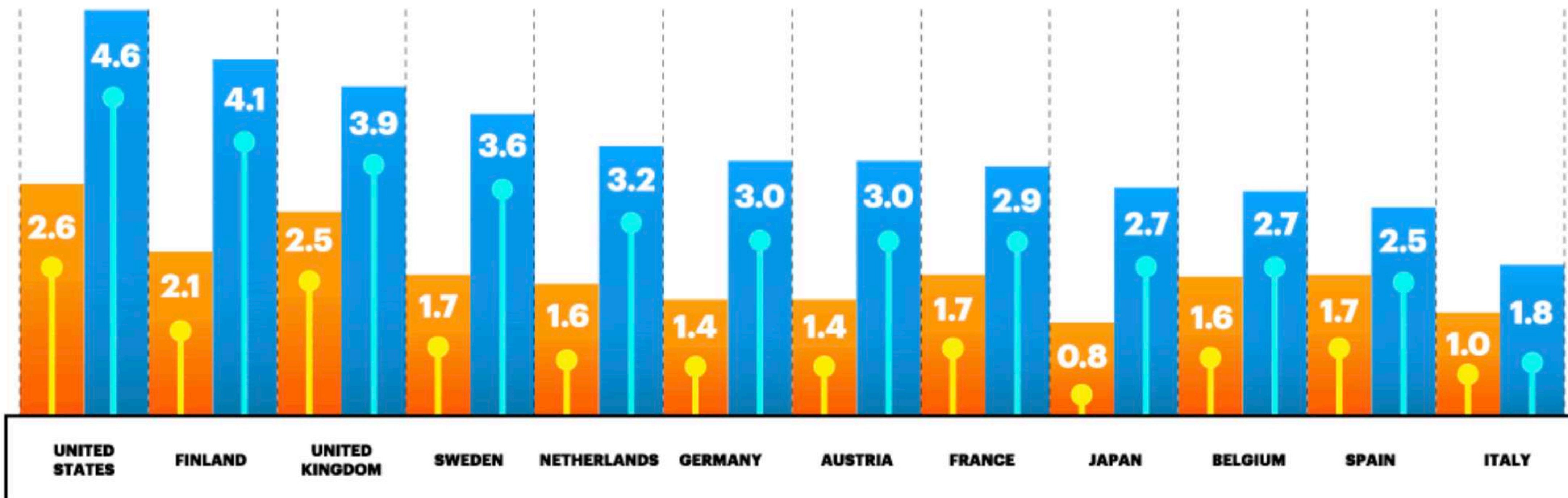INTERNET OF US

Enabling new applications and services

And becoming a commodity!

20 attorneys vs AI

accuracy: 94% vs 85%

time: 26" vs 51+'

Annual growth rates in 2035 of gross value added (a close approximation of GDP), comparing baseline growth in 2035 to an artificial intelligence scenario where AI has been absorbed into the economy

Source: Accenture and Frontier Economics

Baseline

AI steady rate

Country	Baseline	AI steady rate
UNITED STATES	2.6	4.6
FINLAND	2.1	4.1
UNITED KINGDOM	2.5	3.9
SWEDEN	1.7	3.6
NETHERLANDS	1.6	3.2
GERMANY	1.4	3.0
AUSTRIA	1.4	3.0
FRANCE	1.7	2.9
JAPAN	0.8	2.7
BELGIUM	1.6	2.7
SPAIN	1.7	2.5
ITALY	1.0	1.8

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030

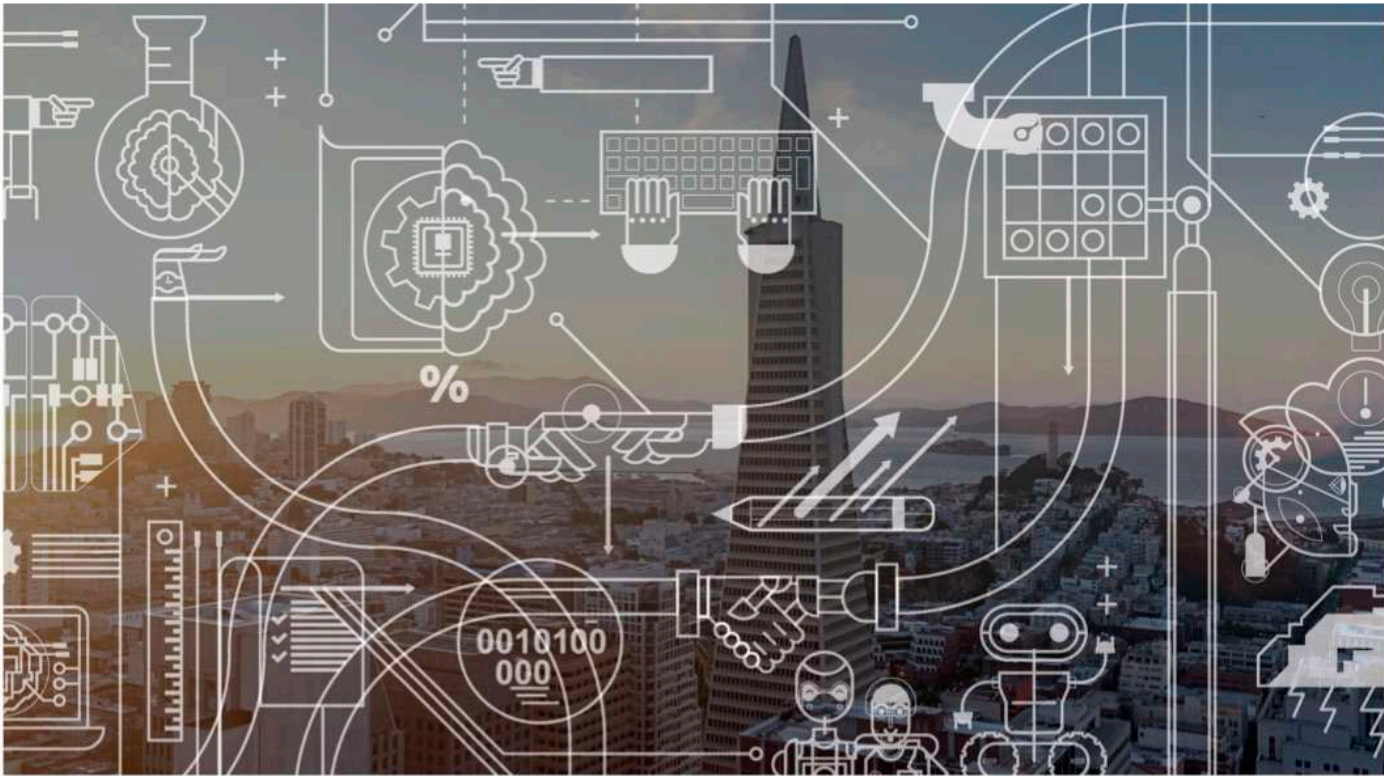ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016

PREFACE

The One Hundred Year Study on Artificial Intelligence, launched in the fall of 2014, is a long-term investigation of the field of Artificial Intelligence (AI) and its influences on people, their communities, and society. It considers the science, engineering, and deployment of AI-enabled computing systems. As its core activity, the Standing Committee that oversees the One Hundred Year Study forms a Study Panel every five years to assess the current state of AI. The Study Panel reviews AI's progress in the years following the immediately prior report, envisions the potential advances that lie ahead, and describes the technical and societal challenges and opportunities these advances raise, including in such arenas as ethics, economics, and the design of systems compatible with human cognition. The overarching purpose of the One Hundred Year Study's periodic expert review is to provide a collected and connected set of reflections about AI and its influences as the field advances. The studies are expected to develop syntheses and assessments that provide expert-informed guidance for directions in AI research, development, and systems design, as well as programs and policies to help ensure that these systems broadly benefit individuals and society.[1]

The One Hundred Year Study is modeled on an earlier effort informally known as the "AAAI Asilomar Study." During 2008-2009, the then president of the Association for the Advancement of Artificial Intelligence (AAAI), Eric Horvitz, assembled a group of AI experts from multiple institutions and areas of the field, along with scholars of cognitive science, philosophy, and law. Working in distributed subgroups, the participants addressed near-term AI developments, long-term possibilities, and legal and ethical concerns, and then came together in a three-day meeting at Asilomar to share and discuss their findings. A short written report on the intensive meeting discussions, amplified by the participants' subsequent discussions with other colleagues, generated widespread interest and debate in the field and beyond.

The impact of the Asilomar meeting, and important advances in AI that included AI algorithms and technologies starting to enter daily life around the globe, spurred the idea of a long-term recurring study of AI and its influence on people and society. The One Hundred Year Study was subsequently endowed at a university to enable

The overarching purpose of the One Hundred Year Study's periodic expert review is to provide a collected and connected set of reflections about AI and its influences as the field advances.

1   "One Hundred Year Study on Artificial Intelligence (AI100)," Stanford University, accessed August 1, 2016, https://ai100.stanford.edu.

Impact not just economy, but every aspect of life

UhoPPer

trend #3½: Collective Intelligence (hybrid AI)

The missing piece: you

Nothing can (yet) replace human intelligence

In particular **social collective**

intelligence

Wicked problems

Extremely hard for AI

Teams better than individuals at

tackling them

The level of diversity correlates with

the effectiveness of solutions

Problems are never completely solved

Every problem is unique

Can take a *long* time to evaluate solutions

There is no clear problem defintion

Solutions are not right/ wrong, but better/worse

Are multi-causal multi-scalar & inter-connected

Wicked Problems

Every solu-tion ramifies throughout the system

Multiple stakeholders with conflict-ing agendas

Every wicked problem is connected to others

Straddle organizational & disciplinary boundaries

Based upon Rittel and Webber (1973)

UhoPPer

WENET
INTERNET OF US

challenge: combining AI with collective intelligence (hybrid AI)

challenge (rephrased): enhancing collective intelligence with AI

Daniele Miorandi

daniele.miorandi@u-hopper.com
@danielemiorandi

AI & Ethics

What is ethics?

"Ethics can be understood as the task of answering "What should I do?".
Ethics includes both thought and practice, an organized and intentional reflection
on morality and the effort to live in ways that are good, just, and/or right. "

Emanuelle Burton, Judy Goldsmith, Nicholas Mattei ,
Communications of the ACM, August 2018, Vol. 61 No. 8, Pages 54-64

WENET
INTERNET OF US

U·HOPPER

Ethical issues in AI

- Moral decision-making of machines
  - Inherent values of technology and AI
- Algorithmic bias and discrimination
  - Statistics and data bias
- (Algorithmic transparency)
- (Accountability)

Moral decision-making of machines

Culture and diversity

- What people think is the 'best' decision (in a moral sense) reflects the value system embedded in the local culture
- Different people see things differently
- The role of diversity and the ability of AI to adapt to different cultures

U·HOPPER

Algorithmic bias and discrimination

WENET
INTERNET OF US

AI-powered recidivism scores

Widely used in the US to assess the likelihood of recidivism

Can impact the possibility of getting a parole

Independent studies found a strong bias against black defendants

WENET
INTERNET OF US

A self-fulfilling prophecy?

WENET
INTERNET OF US

U·HOPPER

Predictive policing

The relevance of bias in training datasets
#epicfail

WENET
INTERNET OF US

U·HOPPER

The 'gorilla error' (2015)

Google image prediction engine (GCV) labelled

an afro-american couple as 'gorillas'

Tay (2016)

TaY (Tay and You) was an AI-powered chatbot released by Microsoft in 2016

It included a learning mechanisms meant to help her improve her interaction skills

over time

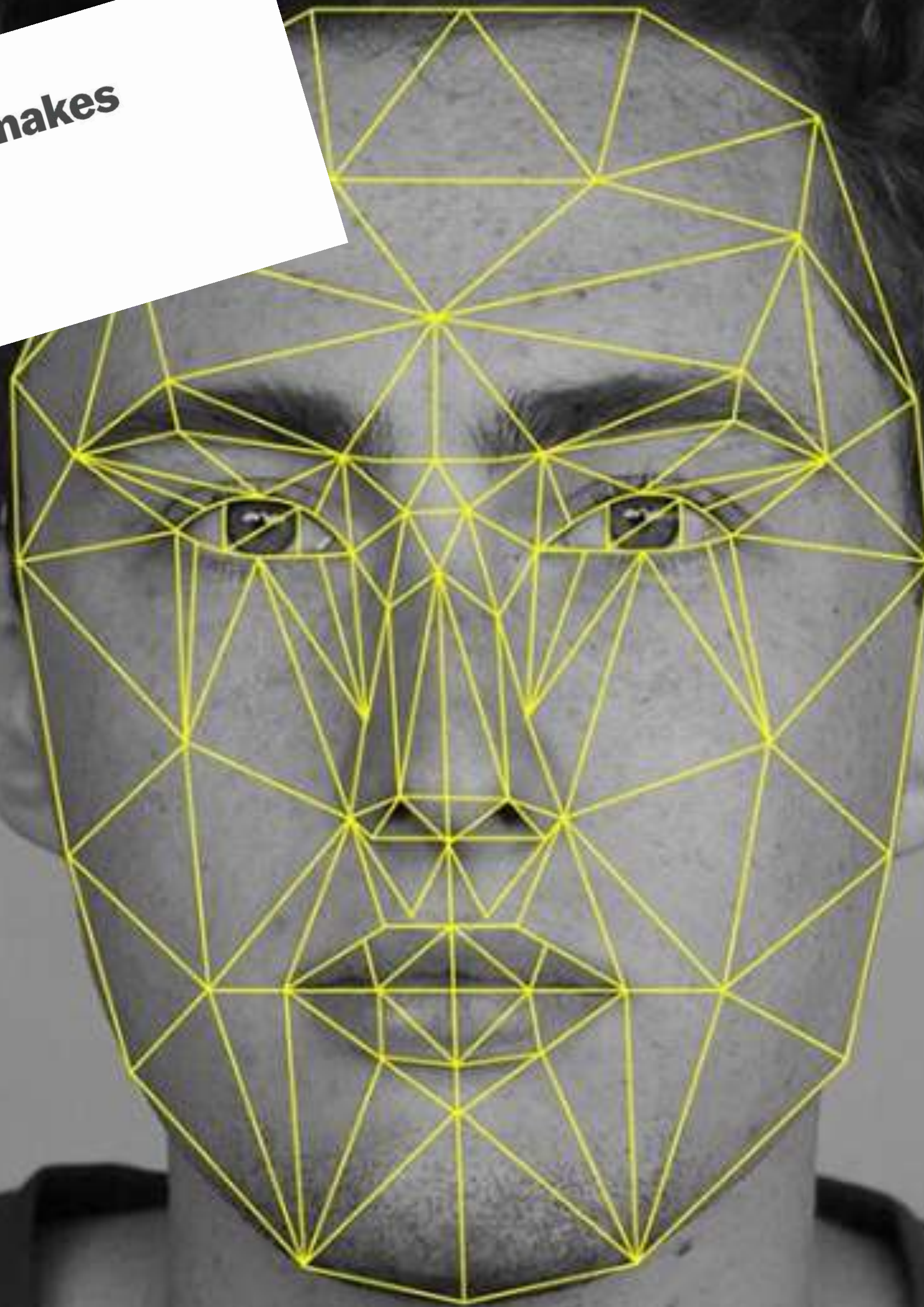But something went very wrong...

WENET
INTERNET OF US

U·HOPPER

It is doable.
Should we do it?

This psychologist's "gaydar" research makes us uncomfortable. That's the point.

Michal Kosinski used artificial intelligence to detect sexual orientation. Let him explain why.

By Brian Resnick | @B_resnick | brian@vox.com | Jan 29, 2018, 12:00pm EST

The role of policy-makers

"The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status"
(European Convention on Human Rights, Art. 14)

WENET
INTERNET OF US

U·HOPPER

(a) Personal data may only be processed lawfully, fairly and transparently ("lawfulness, fairness, and transparency").
(b) Such data may only be collected for a purpose that is specified in advance, and should not be used for other unrelated purposes ("purpose limitation").
(c) Such data should be limited to what is necessary for the processing purpose ("data minimisation").
(d) Such data should be sufficiently accurate and up-to-date ("accuracy").
(e) Such data should not be retained for an unreasonably long period ("storage limitation").
(f) Such data should be secured against data breaches, illegal use etc ("integrity and confidentiality").
(g) The data controller is responsible for compliance ("accountability").

GDPR, 2016

Daniele Miorandi

daniele.miorandi@u-hopper.com
@danielemiorandi

https://www.internetofus.eu/

WENET
INTERNET OF US