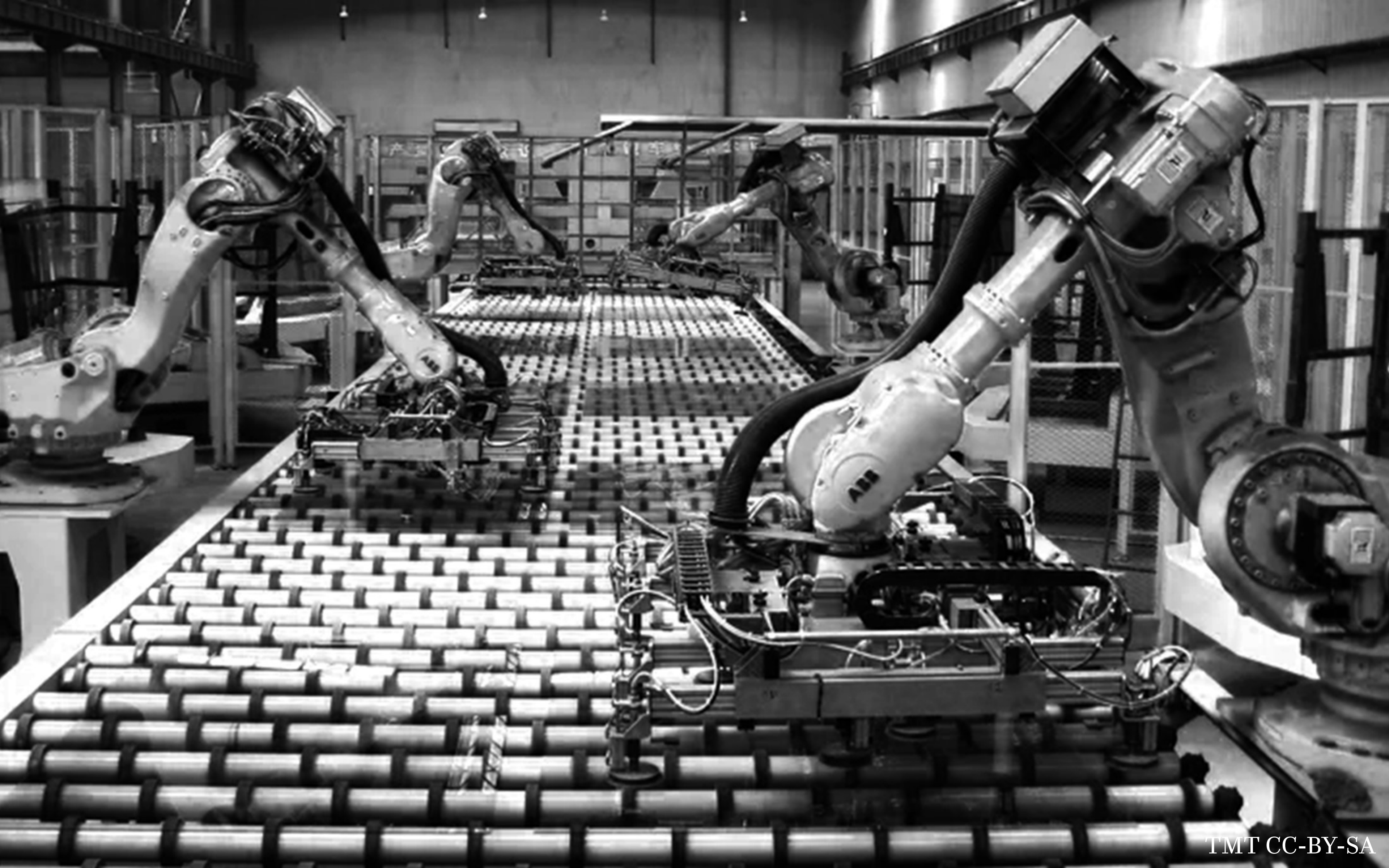
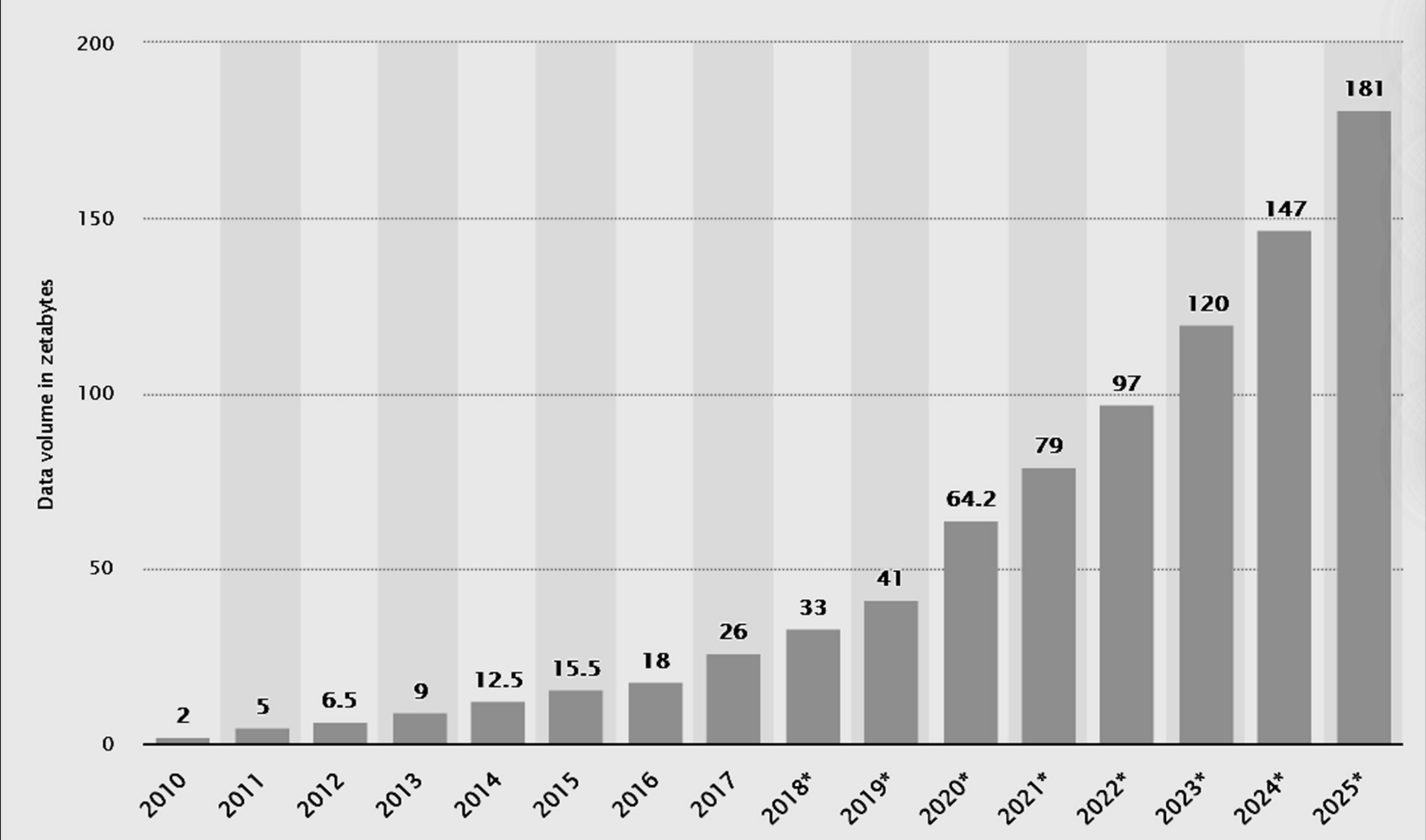
Battling robots

For our data, privacy and humanity

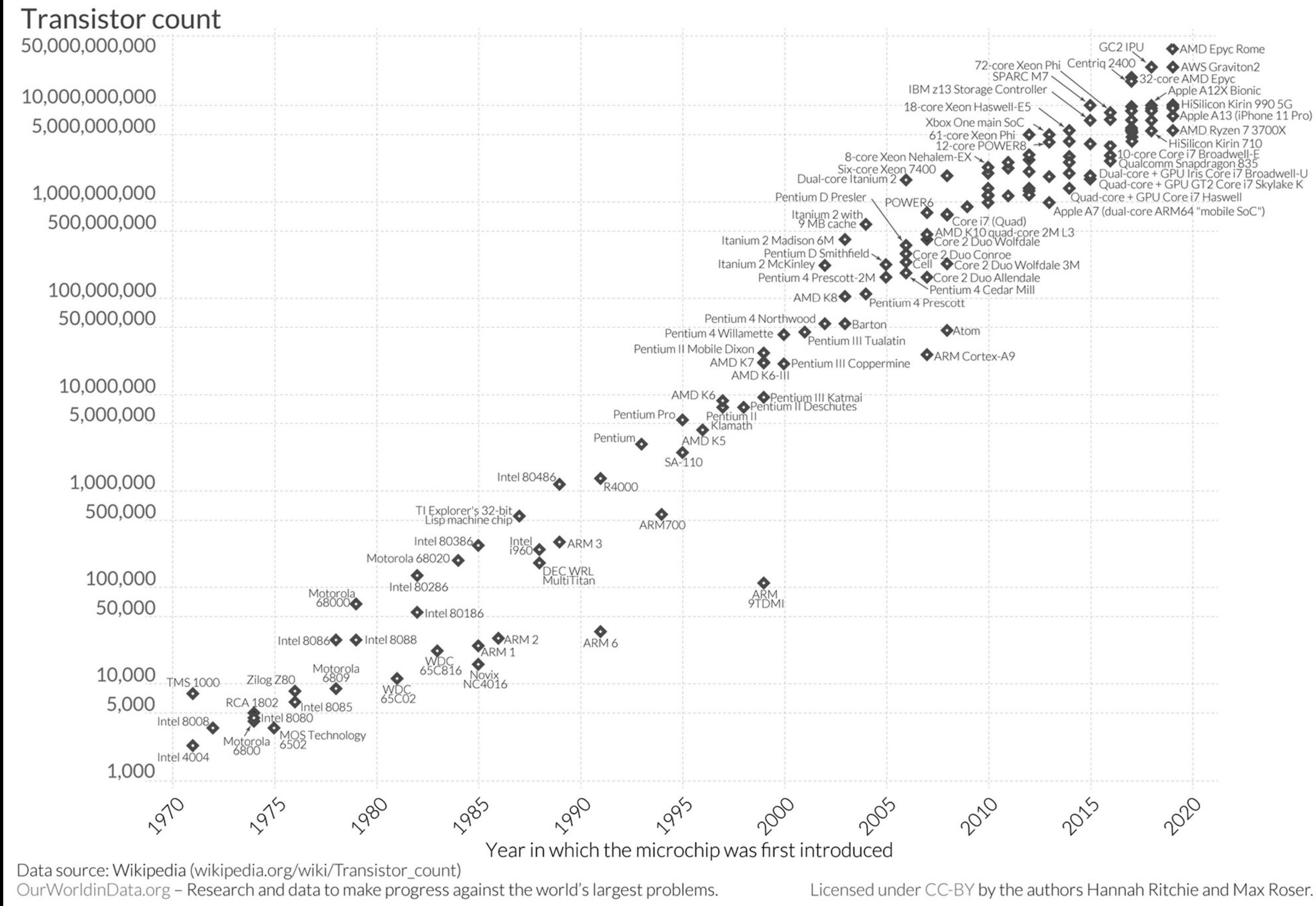
Talk by Dr. Andrzej NOWAK – November 2022
CERN Academic Training Programme



Automation gone wrong



Statista Website

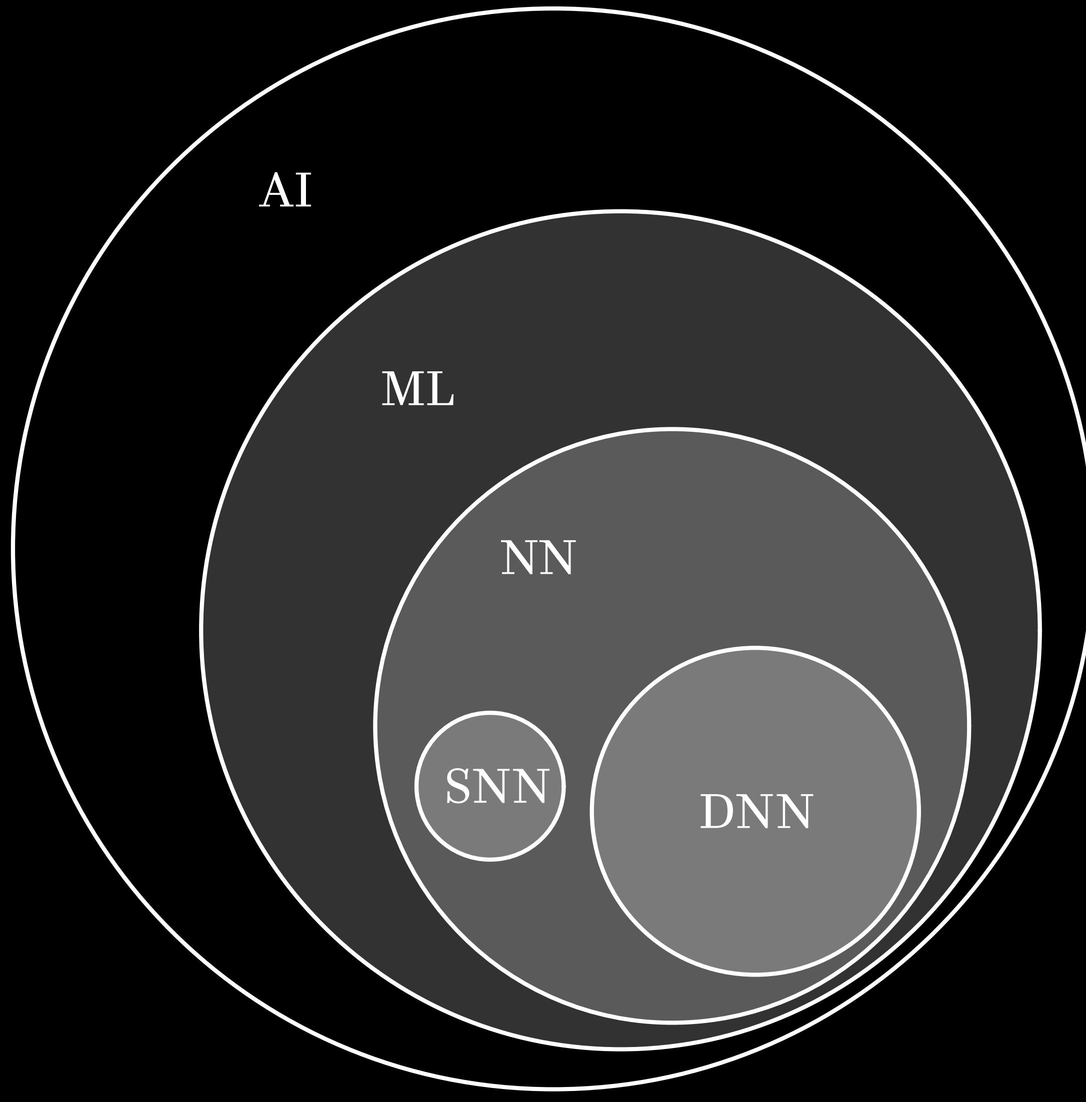


Transistor count
50,000,000,000
10,000,000,000
5,000,000,000
1,000,000,000
500,000,000
100,000,000
50,000,000
10,000,000
5,000,000
1,000,000
500,000
100,000
50,000
10,000
5,000
1,000
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
Year in which the microchip was first introduced
GC2 IPU
AMD Epyc Rome
72-core Xeon Phi
Centriq 2400
AWS Graviton2
SPARC M7
32-core AMD Epyc
IBM z13 Storage Controller
Apple A12X Bionic
18-core Xeon Haswell-E5
HiSilicon Kirin 990 5G
Apple A13 (iPhone 11 Pro)
Xbox One main SoC
AMD Ryzen 7 3700X
61-core Xeon Phi
HiSilicon Kirin 710
12-core POWER8
10-core Core i7 Broadwell-E
8-core Xeon Nehalem-EX
Qualcomm Snapdragon 835
Six-core Xeon 7400
Dual-core + GPU Iris Core i7 Broadwell-U
Dual-core Itanium 2
Quad-core + GPU GT2 Core i7 Skylake K
Quad-core + GPU Core i7 Haswell
Pentium D Presler
POWER6
Apple A7 (dual-core ARM64 "mobile SoC")
Itanium 2 with 9 MB cache
Core i7 (Quad)
AMD K10 quad-core 2M L3
Itanium 2 Madison 6M
Core 2 Duo Wolfdale
Pentium D Smithfield
Core 2 Duo Conroe
Itanium 2 McKinley
Cell
Core 2 Duo Wolfdale 3M
Pentium 4 Prescott-2M
Core 2 Duo Allendale
Pentium 4 Cedar Mill
AMD K8
Pentium 4 Prescott
Pentium 4 Northwood
Barton
Atom
Pentium 4 Willamette
Pentium III Tualatin
Pentium II Mobile Dixon
ARM Cortex-A9
AMD K7
Pentium III Coppermine
AMD K6-III
AMD K6
Pentium III Katmai
Pentium II Deschutes
Pentium Pro
Pentium II Klamath
Pentium
AMD K5
SA-110
Intel 80486
R4000
TI Explorer's 32-bit Lisp machine chip
ARM700
Intel 80386
Intel i960
ARM 3
Motorola 68020
DEC WRL MultiTitan
Intel 80286
ARM 9TDMI
Motorola 68000
Intel 80186
Intel 8086
Intel 8088
ARM 2
ARM 1
ARM 6
WDC 65C816
Motorola 6809
Novix NC4016
TMS 1000
Zilog Z80
WDC 65C02
RCA 1802
Intel 8085
Intel 8080
Intel 8008
MOS Technology 6502
Motorola 6800
Intel 4004
Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)
OurWorldinData.org – Research and data to make progress against the world's largest problems.
Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

The emergence of AI

1950s	1960s	1970s	1980s	1990s
AI coined as a term. Can AI mathematically exist?	DARPA funds AI at MIT HAL9000 Moore's Law	The hype cools	Expert systems Navlab car	Deep Blue > Kasparov

2000s	2010s	2020s
It's back!	IBM Watson Turing test?	Wide business use. Text-to-image. Optimization.


AI
ML
NN
SNN
DNN

“you create your brain from the input you get.”

- Ray Kurzweil

AI/ML/data privacy risks and consequences

it needs data, while compute usually happens off-device

→ data goes off-device

inference based on breadcrumbs can identify you, your attributes

→ loss of control over information

predictions based on others can affect you (profiling)

→ algorithms create new sensitive information

→ automation becomes automated

→ edge cases progressively "smoothed" out

  → bias, fallibility

→ garbage in, garbage out

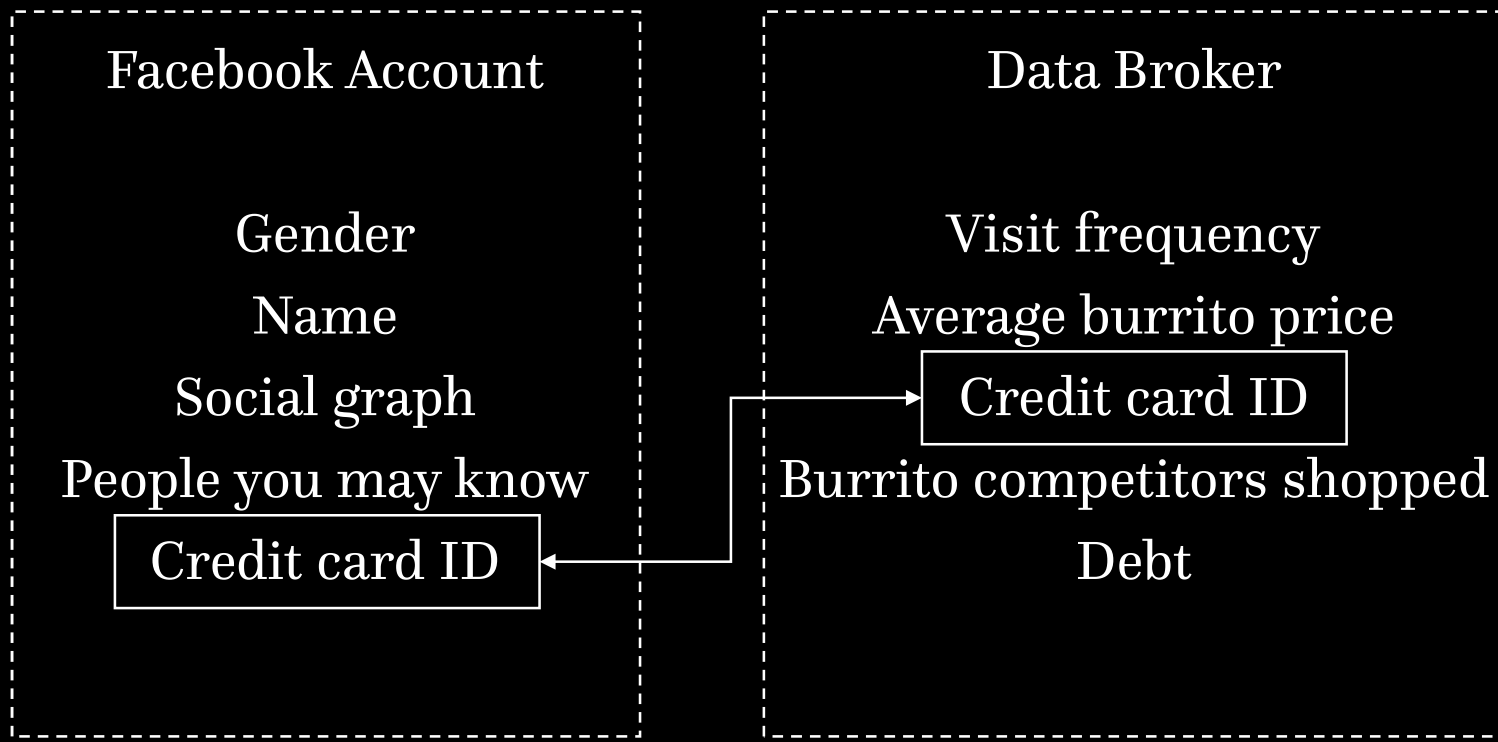
→ who's responsible? who's going to fix "it"?

“OK, I will destroy humans.”

Sophia the robot
SXSW 2016

Confronting less brainy data vacuums

Correlation




ONE NATION, TRACKED
AN INVESTIGATION INTO THE SMARTPHONE TRACKING INDUSTRY FROM TIMES OPINION
NYT 2019

Talia Shadwell @TaliaShadwell · Nov 3, 2019

Like many women I know - I use a period tracker app. I opened it today and found I hadn't logged last month's cycle - it flashed a warning that I was very 'late'

6 164 1,471

Talia Shadwell @TaliaShadwell · Nov 3, 2019

Because I had forgotten to log a cycle, the app likely concluded I was pregnant and began communicating the information to third party apps and algorithms

21 490 2,596

Talia Shadwell
@TaliaShadwell

I corrected my cycle in the tracker app and just like that - the ads have stopped

7:14 PM · Nov 3, 2019 · Twitter for iPhone

179 Retweets **10** Quote Tweets **2,016** Likes

How Companies Learn Your Secrets

Give this article 570

Antonio Bolfo/Reportage for The New York Times

By Charles Duhigg
Feb. 16, 2012

Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: "If we wanted to figure out if a customer is pregnant, even if she didn't want us to know, can you do that? "

Twitter 2019 **NYT 2012**

“Also linked to your Guest ID is demographic information like your age, whether you are married and have kids, which part of town you live in, how long it takes you to drive to the store, your estimated salary, whether you’ve moved recently, what credit cards you carry in your wallet and what Web sites you visit.
Target can buy data about your ethnicity, job history, the magazines you read, if you’ve ever declared bankruptcy or got divorced, the year you bought (or lost) your house, where you went to college, what kinds of topics you talk about online, whether you prefer certain brands of coffee, paper towels, cereal or applesauce, your political leanings, reading habits, charitable giving and the number of cars you own..........................”

NYT, 2012

Correlation - fragmentation

“Fragmented data makes it difficult for advertisers to understand the true effect of their ad campaigns”

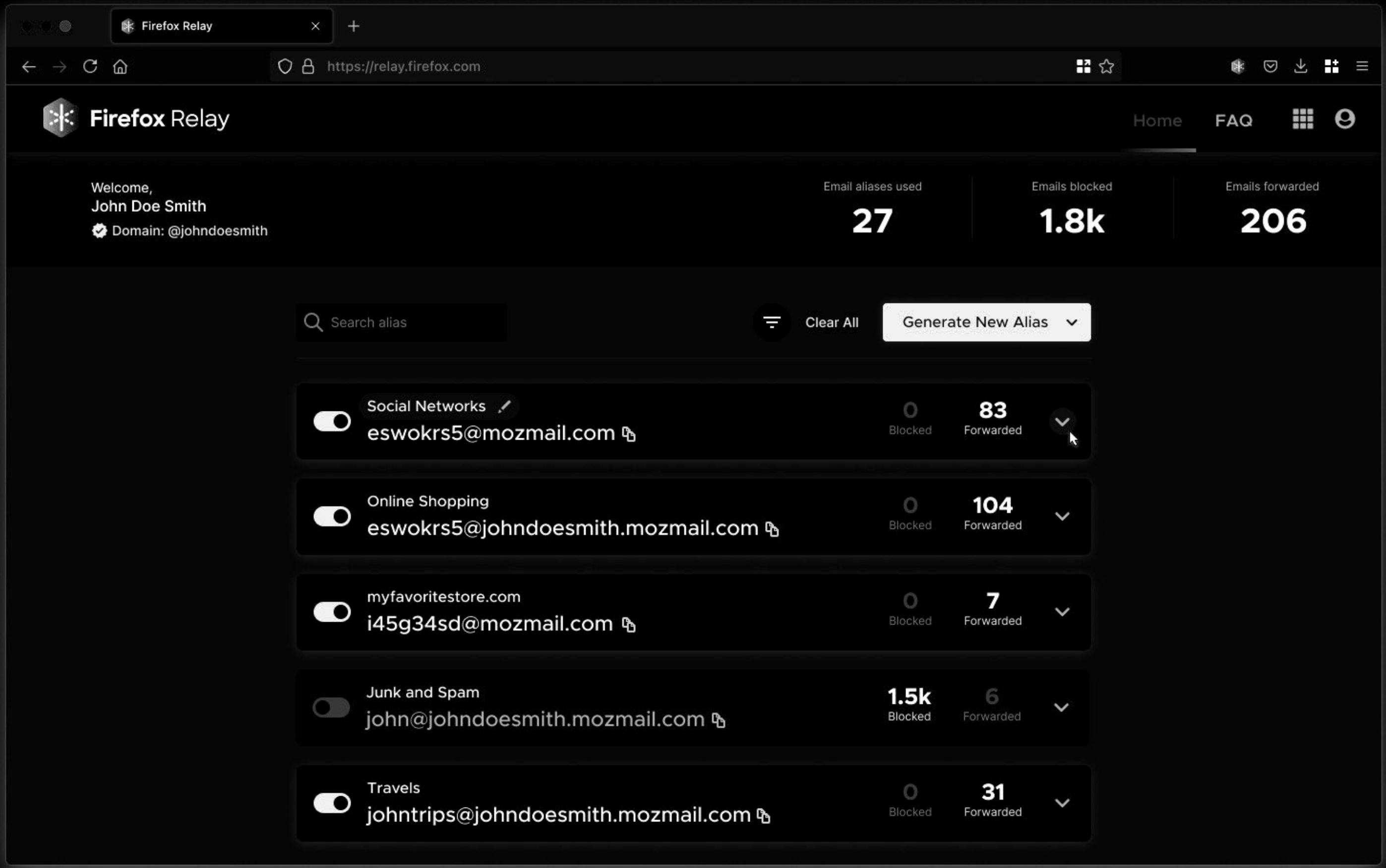
Tesary Lin, Boston Uni

Examples:

iOS 14 / IDFA

Google Chrome / 3-rd party cookie phase-out

Email fragmentation



Android fragmentation

“Sensitive data is stored in user profiles. User profiles each have their own unique, randomly generated disk encryption key and their own unique key encryption key is used to encrypt it.

[...]

GrapheneOS enables support for ending secondary user profile sessions after logging into them. It adds an end session button to the lockscreen and in the global action menu accessed by holding the power button. This fully purges the encryption keys and puts the profiles back at rest.”

Data anonymization

Removing personally identifiable information (PII) from data sets, so that those described by the data remain anonymous.

Often: suppression, generalization, permutation, pseudonymity

De-anonymization is the reverse process.

Data anonymization examples

CERN → XXXX	blinding / suppression / masking
CERN → 0001	pseudonymous identifier
CERN → CXXX	generalization / masking
2022 → 2020-2030	generalization
2022 → 2022+x	noise (permutation)

What about correlation?
(“*toxic combinations*” a.k.a. “*mosaic effect*”)

"Robust de-anonymization of large sparse datasets"

1. Obtain Netflix dataset containing movie votes
2. Access IMDb
3. ???
4. Profit!

forward secrecy → forward privacy?

Ref: UT Austin

What are PETs?

Privacy-Enhancing Technologies

“Technologies that embody fundamental data protection principles by

minimizing personal data use,

maximizing data security,

empowering individuals.”

Wikipedia

PETs – a few examples

Statistical disclosure control	k-anonymity l-diversity differential privacy
Multi-Party Computation (MPC)	
Encryption	Homomorphic encryption Functional encryption Searchable encryption Zero knowledge proof

k-anonymity

Name	Age	Gender	State of domicile	Religion	Disease
Ramsha	30	Female	Tamil Nadu	Hindu	Cancer
Yadu	24	Female	Kerala	Hindu	Viral infection
Salima	28	Female	Tamil Nadu	Muslim	Tuberculosis
Sunny	27	Male	Karnataka	Parsi	No illness
Joan	24	Female	Kerala	Christian	Heart-related
Bahuksana	23	Male	Karnataka	Buddhist	Tuberculosis
Rambha	19	Male	Kerala	Hindu	Cancer
Kishor	29	Male	Karnataka	Hindu	Heart-related
Johnson	17	Male	Kerala	Christian	Heart-related
John	19	Male	Kerala	Christian	Viral infection

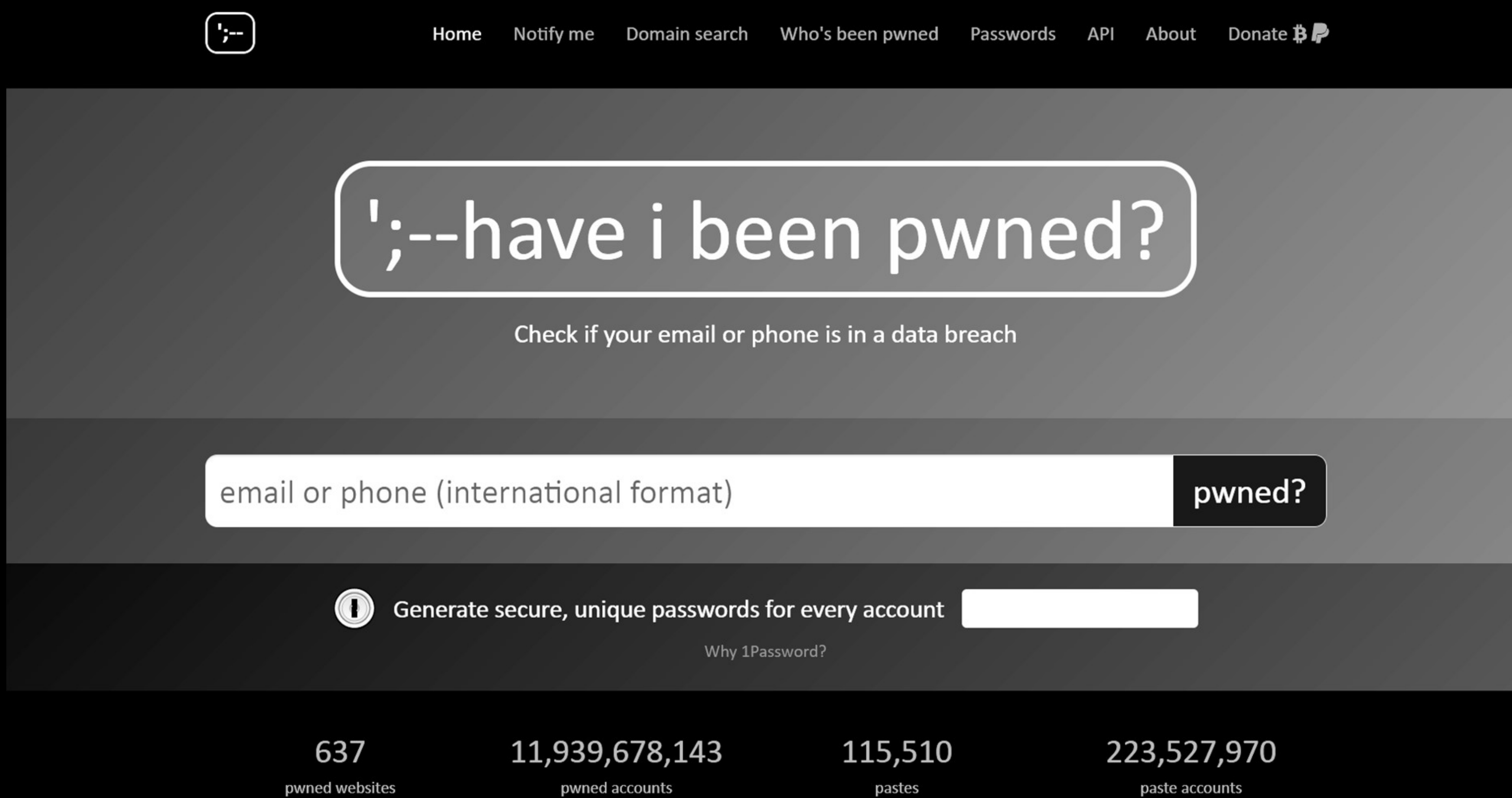
Wikipedia

k-anonymity

Name	Age	Gender	State of domicile	Religion	Disease
*	20 < Age ≤ 30	Female	Tamil Nadu	*	Cancer
*	20 < Age ≤ 30	Female	Kerala	*	Viral infection
*	20 < Age ≤ 30	Female	Tamil Nadu	*	Tuberculosis
*	20 < Age ≤ 30	Male	Karnataka	*	No illness
*	20 < Age ≤ 30	Female	Kerala	*	Heart-related
*	20 < Age ≤ 30	Male	Karnataka	*	Tuberculosis
*	Age ≤ 20	Male	Kerala	*	Cancer
*	20 < Age ≤ 30	Male	Karnataka	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Viral infection

Wikipedia

k-anonymity in action



l-diversity

“An equivalence class is said to have *l*-diversity if there are at least *l* ‘well-represented’ values for the sensitive attribute.

A table is said to have *l*-diversity if every equivalence class of the table has *l*-diversity.”

t-closeness

“An equivalence class is said to have *t*-closeness if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole table is no more than a threshold *t*.

A table is said to have *t*-closeness if all equivalence classes have *t*-closeness” to be continued...

an@tik.services