Modern Data Privacy,
or (How Potter Lost his Cloak)

JEFFREY LEUNG

Subscribe

humans saw this and invented data warehouses

2:22 p.m. · 29 Aug 24 · **498K** Views

1,046 Reposts   **53** Quotes   **17.5K** Likes

533 Bookmarks

0. the scale of it all

A DAY IN DATA

The exponential growth of data is undisputed, but the numbers behind this explosion - fuelled by internet of things and the use of connected devcies – are hard to comprehend, particularly when looked at in the context of one day

DEMYSTIFIYING DATA UNITS

From the more familiar 'bit' or 'megabyte', larger units of measurement are more frequently being used to explain the masses of data

Unit		Value	Size
b	bit	0 or 1	1/8 of a byte
B	byte	8 bits	1 byte
KB	kilobyte	1,000 bytes	1,000 bytes
MB	megabyte	$1,000^2$ bytes	1,000,000 bytes
GB	gigabyte	$1,000^3$ bytes	1,000,000,000 bytes
TB	terabyte	$1,000^4$ bytes	1,000,000,000,000 bytes
PB	petabyte	$1,000^5$ bytes	1,000,000,000,000,000 bytes
EB	exabyte	$1,000^6$ bytes	1,000,000,000,000,000,000 bytes
ZB	zettabyte	$1,000^7$ bytes	1,000,000,000,000,000,000,000 bytes
YB	yottabyte	$1,000^8$ bytes	1,000,000,000,000,000,000,000,000 bytes

*A lowercase "b" is used as an abbreviation for bits, while an uppercase "B" represents bytes.

463EB
of data will be created every day by 2025

IDC

500m
tweets are sent every day

Twitter

4PB
of data created by Facebook, including

350m photos

100m hours of video watch time

Facebook Research

95m
photos and videos are shared on Instagram

Instagram Business

320bn
emails to be sent each day by 2021

306bn
emails to be sent each day by 2020

294bn
billion emails are sent

Radicati Group

65bn
messages sent over WhatsApp and two billion minutes of voice and video calls made

Facebook

3.9bn
people use emails

4TB
of data produced by a connected car

Intel

Searches made a day — **5bn**

Searches made a day from Google — **3.5bn**

Smart Insights

28PB
to be generated from wearable devices by 2020

Statista

ACCUMULATED DIGITAL UNIVERSE OF DATA

4.4ZB 2013

44ZB 2020

PwC

RACONTEUR

Monitoring, Observability, Cookies, Logging, etc.
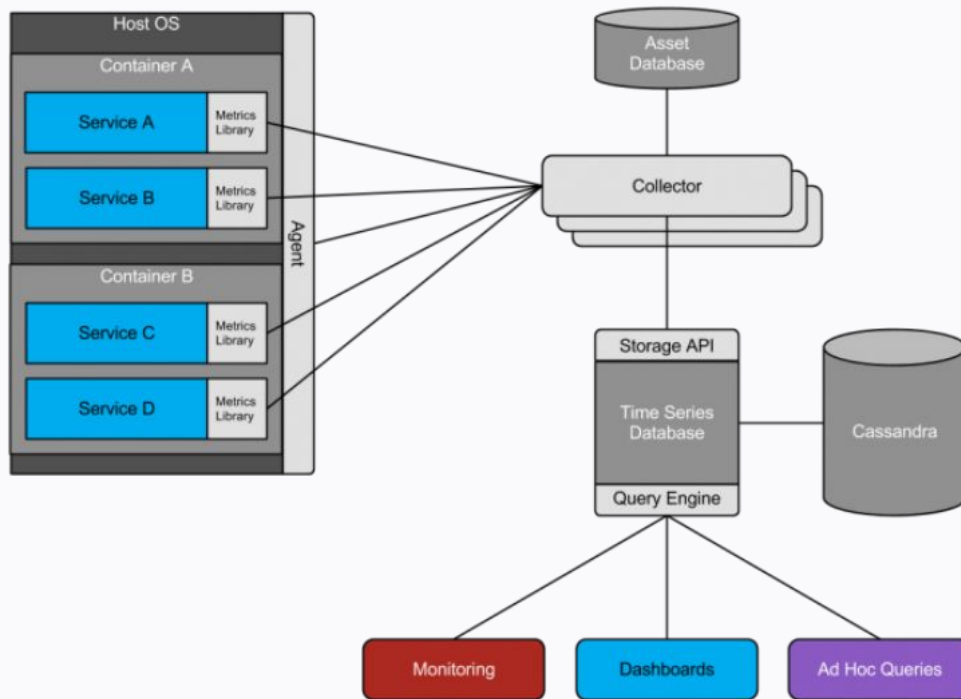
Monitoring, Observability, Cookies, Logging, etc.

???

today's talk

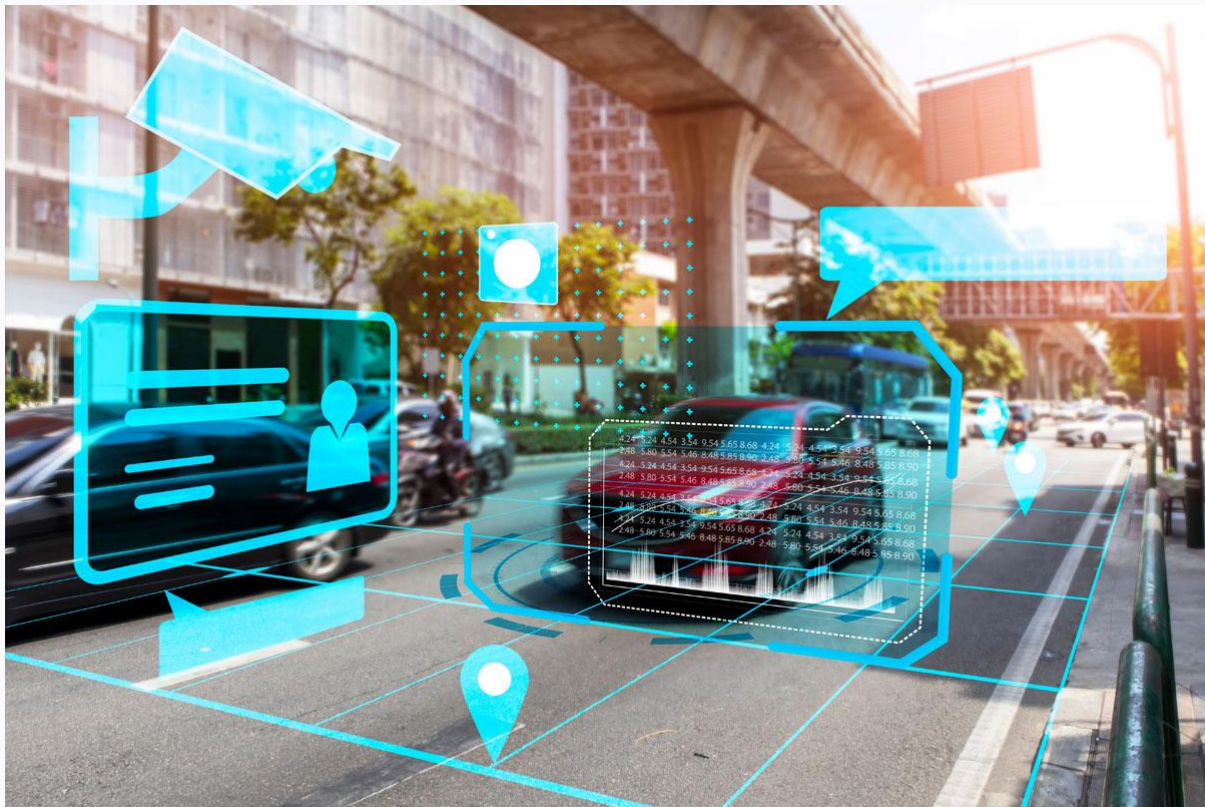# 1. data consequences

Where are you generating data?

- Your laptop
- Your browser
- Your phone
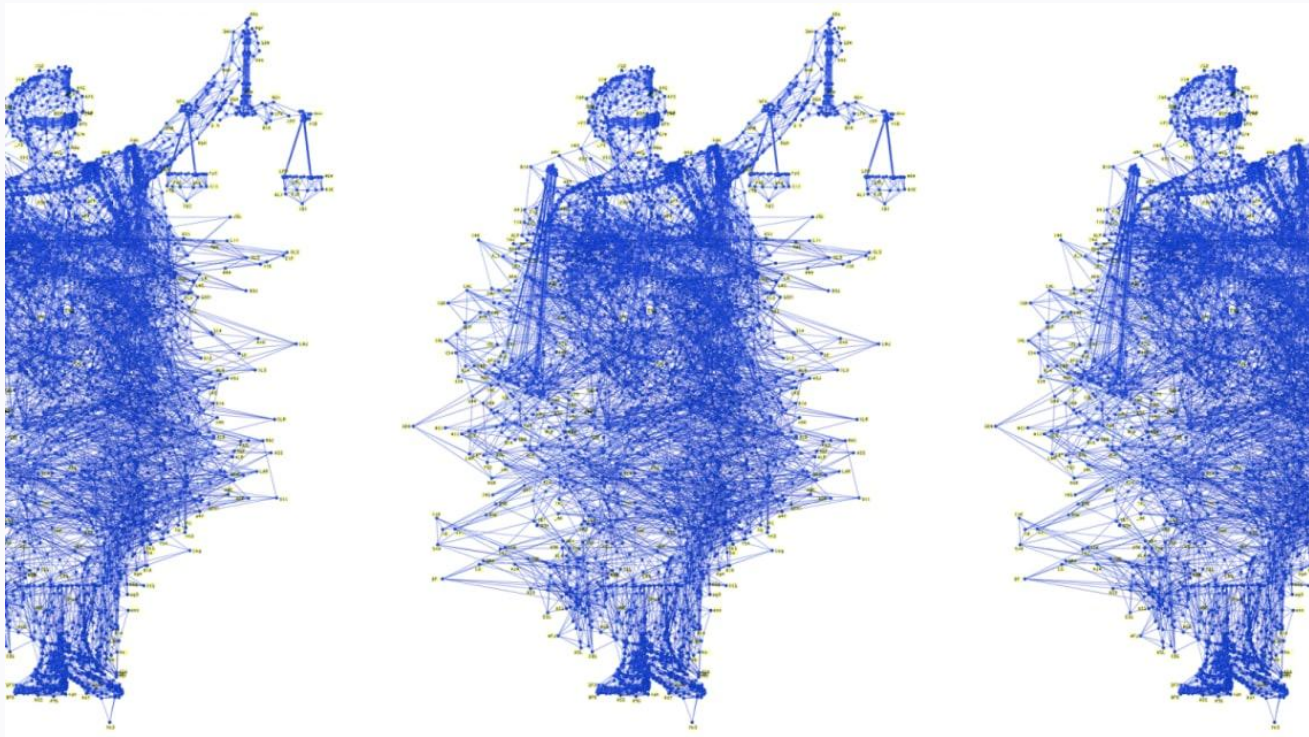- Your apps
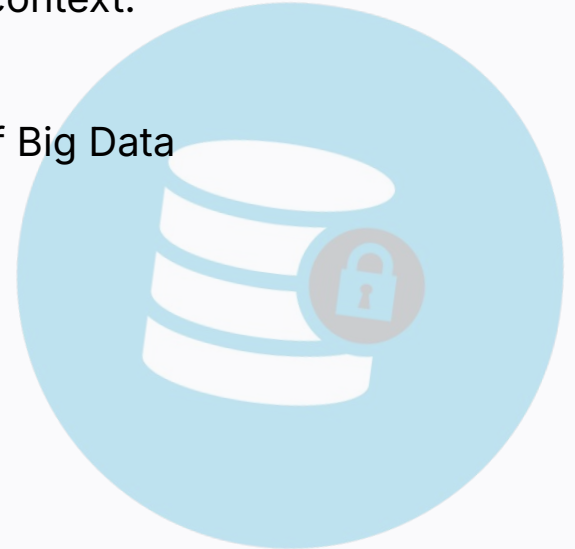- Your purchases
- Your car
- Your face

Have I Been Pwned?

"[Privacy is] **having control** over a situation. It's about understanding the audience and knowing **how far information will flow**.
It's about **trusting** the people, the situation, and the context."

- Danah Boyd, Privacy and Publicity in the Context of Big Data

Who owns your data?

CRUD

CELECT

INSERT

UPDATE

DELETE

2. reactive compliance

Data Protection Regulations

I wish data had never come to me.

I wish none of this had happened.

Why do regulations exist?

3. our responsibility
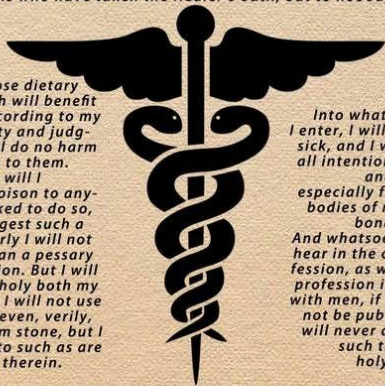
The Hippocratic Oath

I swear by Apollo Healer, by Asclepius, by Hygieia, by Panacea, and by all the gods and goddesses, making them my witnesses, that I will carry out, according to my ability and judgment, this oath and this indenture.

To hold my teacher in this art equal to my own parents; to make him partner in my livelihood; when he is in need of money to share mine with him; to consider his family as my own brothers, and to teach them this art, if they want to learn it, without fee or indenture; to impart precept, oral instruction, and all other instruction to my own sons, the sons of my teacher, and to indentured pupils who have taken the Healer's oath, but to nobody else.

I will use those dietary regimens which will benefit my patients according to my greatest ability and judgment, and I will do no harm or injustice to them. Neither will I administer a poison to anybody when asked to do so, nor will I suggest such a course. Similarly I will not give to a woman a pessary to cause abortion. But I will keep pure and holy both my life and my art. I will not use the knife, not even, verily, on sufferers from stone, but I will give place to such as are craftsmen therein.

Into whatsoever houses I enter, I will enter to help the sick, and I will abstain from all intentional wrong-doing and harm, especially from abusing the bodies of man or woman, bond or free. And whatsoever I shall see or hear in the course of my profession, as well as outside my profession in my intercourse with men, if it be what should not be published abroad, I will never divulge, holding such things to be holy secrets.

Now if I carry out this oath, and break it not, may I gain for ever reputation among all men for my life and for my art; but if I break it and forswear myself, may the opposite befall me.

As a user

- Password manager (Bitwarden, 1Password)
- Signal messaging app
- HTTPS encryption
- VPN
- Multi-Factor Authentication
- Data collection settings

Read these

As a developer or company

- Get a Chief Information Security Officer (CISO)

- Build systems privacy-first

  - Enable SSL

  - Avoid logging PII (personally identifiable information)

  - Don't roll your own encryption

  - Implement MFA

- **Audit your systems**

As a developer or company

- Increase privacy accessibility

- Implement compliance for GDPR, DSAR, right to be forgotten, etc.

- Prepare an incident response plan
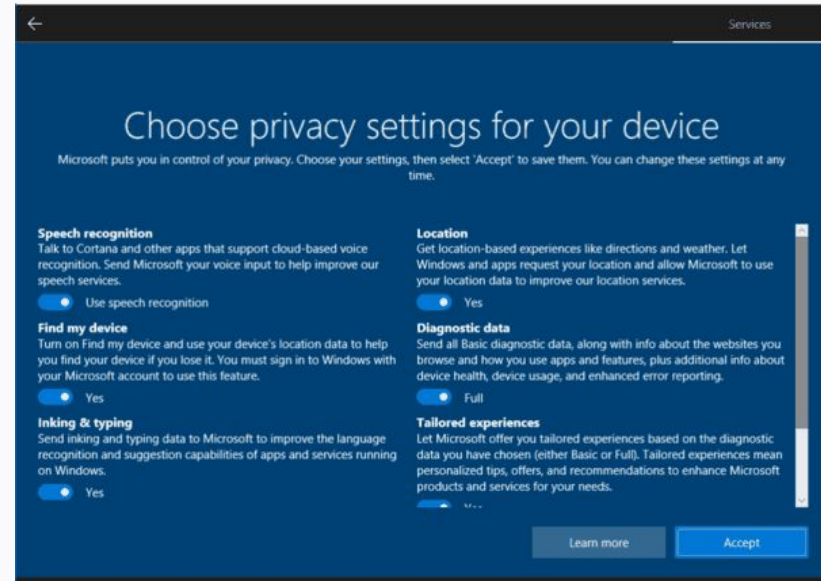
As a developer or company

- Minimize attack surface:

  - Set standards for retention/deletion

  - Collect as anonymous data as possible

    - Anonymize

    - Pseudonymize

    - Aggregate

  - Secure your data stores and channels

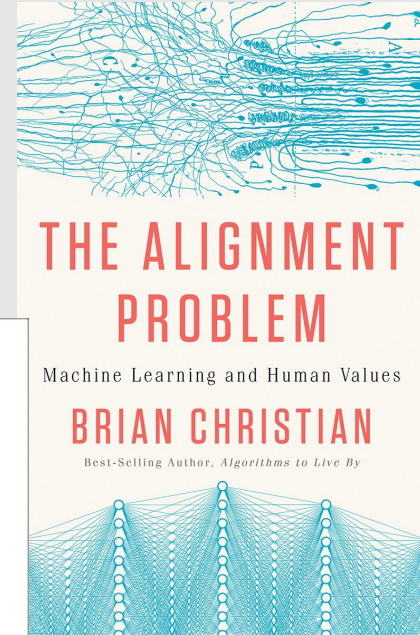 - Data governance: Who owns the data? Who can access/copy it?
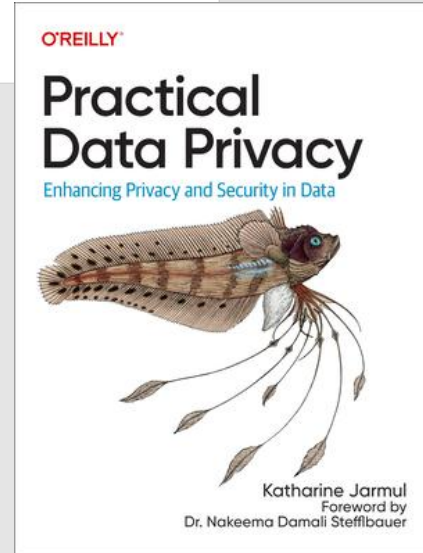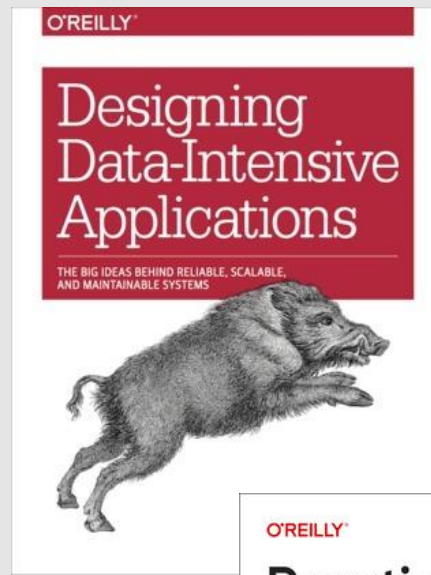
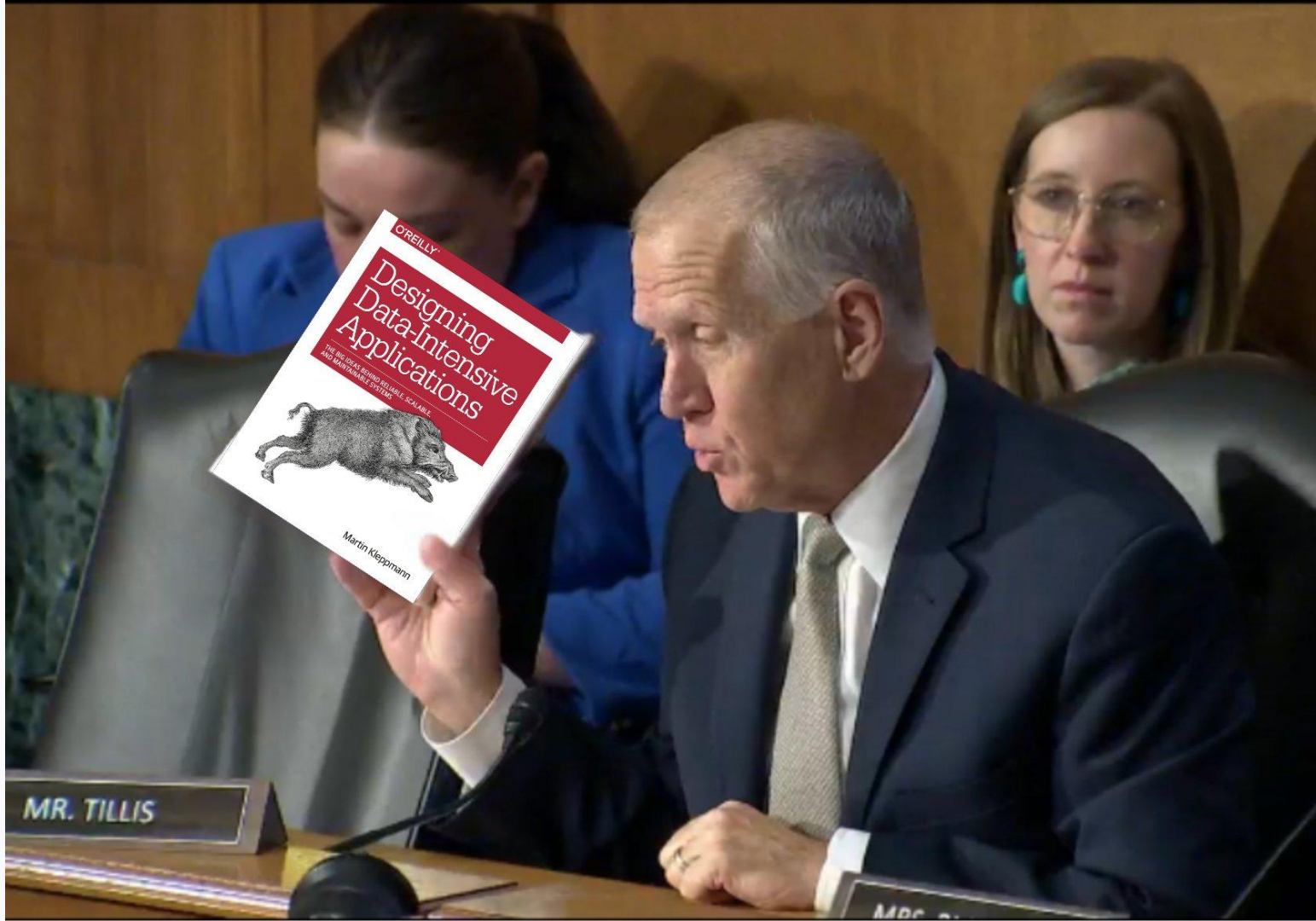HACKERMAN

As a developer or company

- Implementing a new feature or additional data collection:

    - Ask for the purpose

    - Collect only what is needed

    - Follow known regulations

    - Maximize privacy

- Be accountable

GDPR

Protecting Children's Data   Legalities of Processing Data   Data Protection Officer (DPO)   Communication Privacy   Compliance

Read these

Designing Data-Intensive Applications

Practical Data Privacy

The Alignment Problem

MR. TILLIS

in summary

0. we collect a lot of data

1. we have to be accountable

2. reduce, anonymize, and secure our data to protect our users (ourselves)

"Tech people rarely, if ever, have a sense of the **broader applications** and **policy implications** of the projects to which they're assigned.
And the work that consumes them tends to require such specialized knowledge that to bring it up at a barbecue would get them **disinvited** from the next one, because nobody cared."

"In retrospect, maybe **that's what got us here**."

- Edward Snowden, Permanent Record

Questions?

Jeffrey Leung

GitHub: **jleung51**