

NIHR Mindtech 2017 National Symposium

Derek McAuley

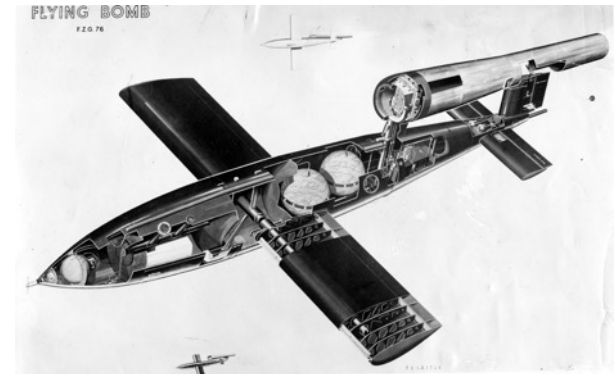
December 2017



I'm not a lawyer but... (1 of 2)

'Data, what data, who owns it, why does it matter?'

- In Law, there is no property right in data
 - There are obligations and limitations based on what the data represents
 - Copyright material
 - Personal Data
 - Plans for bombs
 - ...
- Why do we say “own” so often?



“...much has to do with the psychology of ownership. <...> I mean ownership as signaled by a compulsion to use possessive adjectives — my face, my body, my information, our beliefs, my space, our house, my name. [1]”

[1] Keiran O'Hara "The Seven Veils of Privacy" <https://doi.org/10.1109/MIC.2016.34>

I'm not a lawyer but... (2 of 2)

- 'There is a right in collections of data (in the EU) where:
 - "... there has been a substantial investment in obtaining, verifying or presenting the contents of the database"
- For now we primarily concern ourselves with *Personal Data*
 - Hence EU General Data Protection Regulation (GDPR)
 - On statute books since May 2016
 - Enforceable May 2018 (in the UK #brexit or not...)
 - But also have regard for other specific regulations (e.g. financial, NIS directive, NHS governance, ...)

Observation...

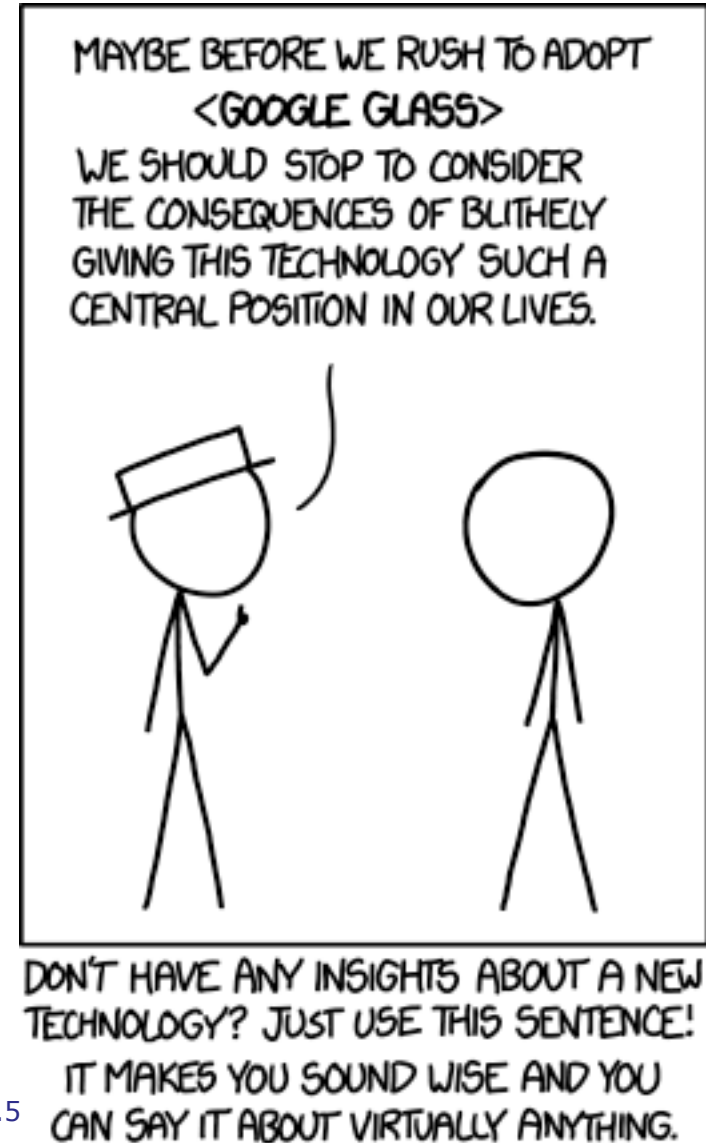
Why is a technologist interested?

Challenge:

- Legal and *ethical* development...
- ...to increase value from personal data

Observation:

- What can we do without sharing?



TRUST IN PERSONAL DATA: A UK REVIEW

Following and assessing the
UK's journey to becoming a
data-driven nation

KEY STATISTICS:

96% of respondents claim to understand the term 'personal data' but less than two-thirds (64%) picked the correct definition

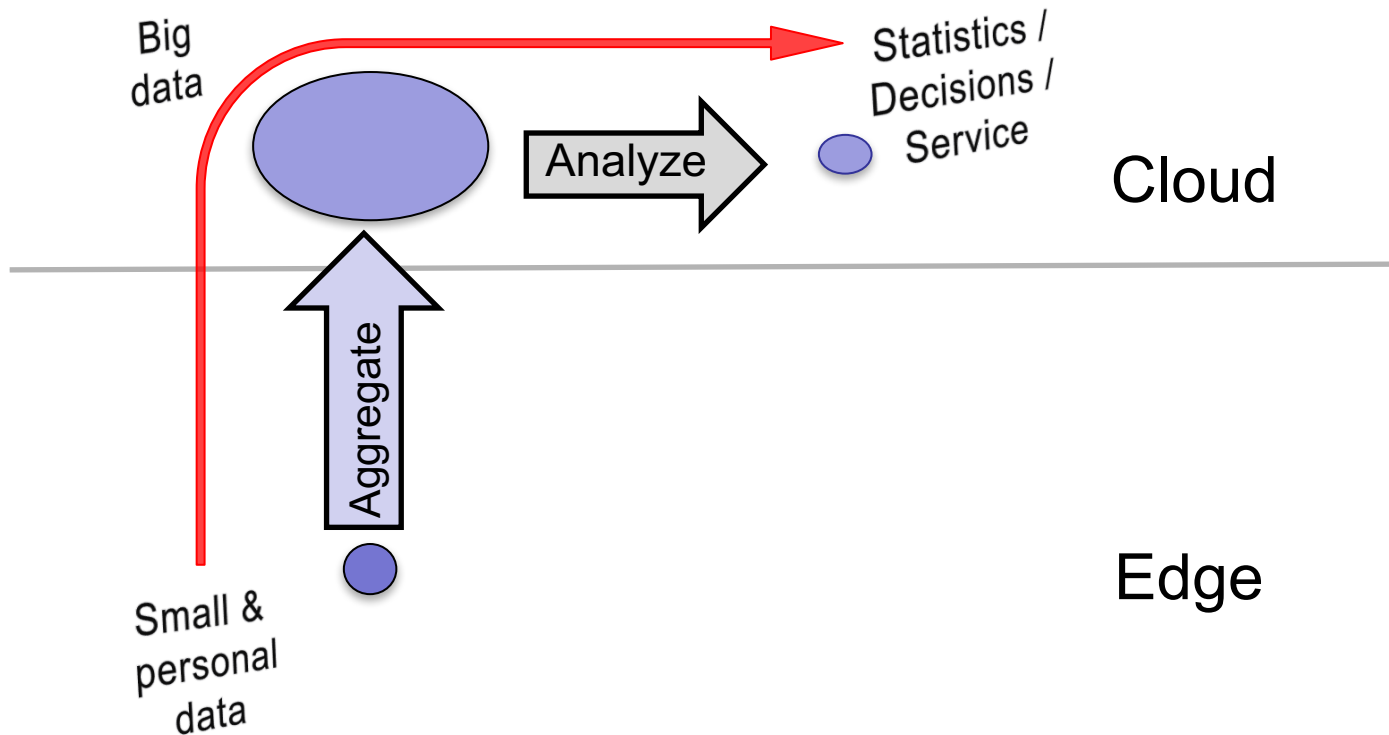
65% of consumers are unsure if data is being shared without their consent

30% of respondents believe they are responsible for educating themselves on the use and protection of data

32% believe this responsibility lies with the Government

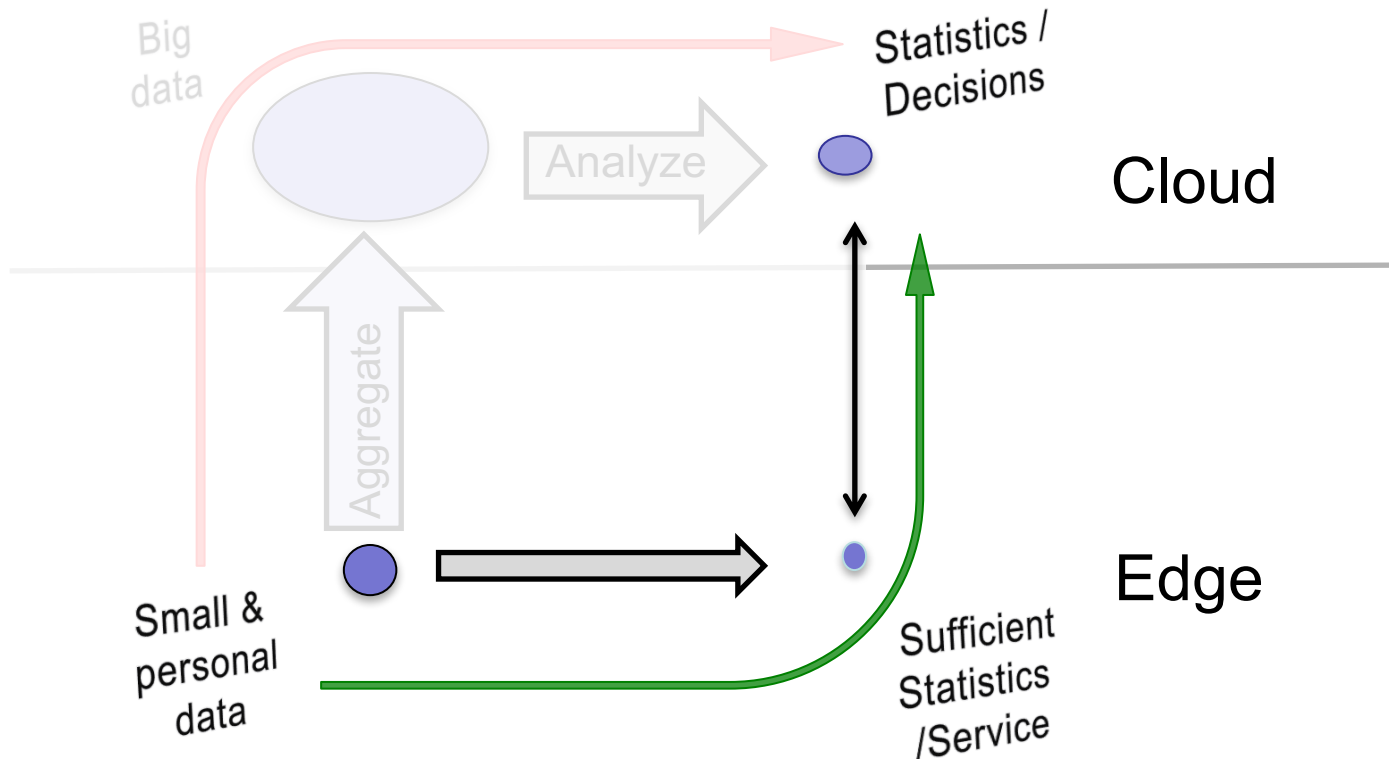
30% interested in data management service

The way of big data



Traditionally: aggregate (into cloud) then analyze

The way of small data



The way of small data: analyze at edge and then aggregate

- Privacy & Performance & Scalability & Resilience

What's your data problem?

- Private computation for compliance and acceptability
e.g. activity / affect interference from in home video
- Derived statistics can be much less sensitive
e.g. per day records? or generative model?
- What classes of "big data" algorithm are amenable to "small data" approaches?
- Where can a reduction in accuracy per individual be compensated for by greater participation?
- What are the stability criteria for iterative algorithms in such a world? (e.g. Machine Learning^[*])

[*] do not get me started on why this is NOT AI

Data portability under GDPR...



Are
you
ready?



<https://www.flickr.com/photos/catjannik/4297548747> CC BY-NC 2.0

US public domain, UK inconclusive

Personal Data Ecosystem

Data about us:

- Digital footprint, inferences



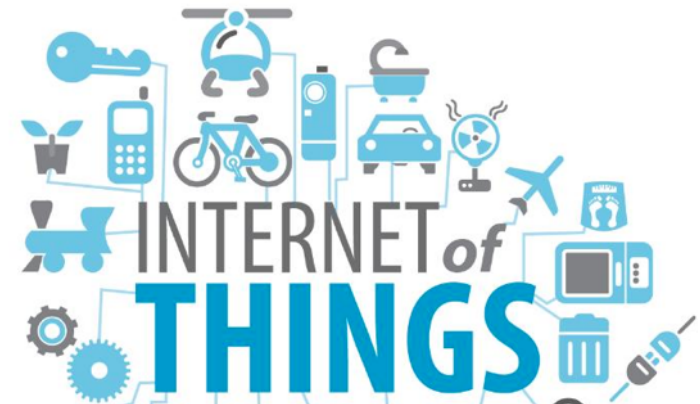
Data generated by us:

- Social media, wearables, smartphones

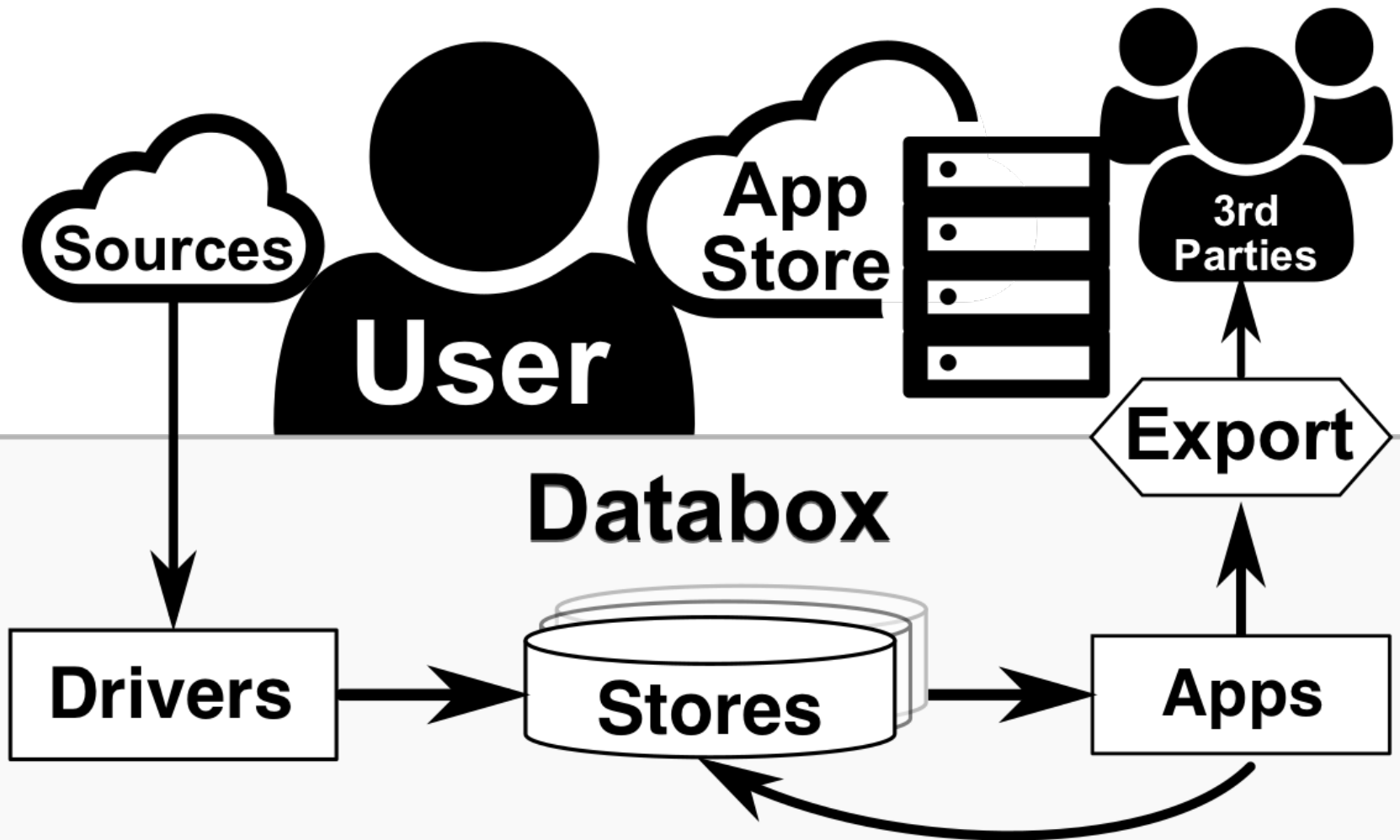


Data around us:

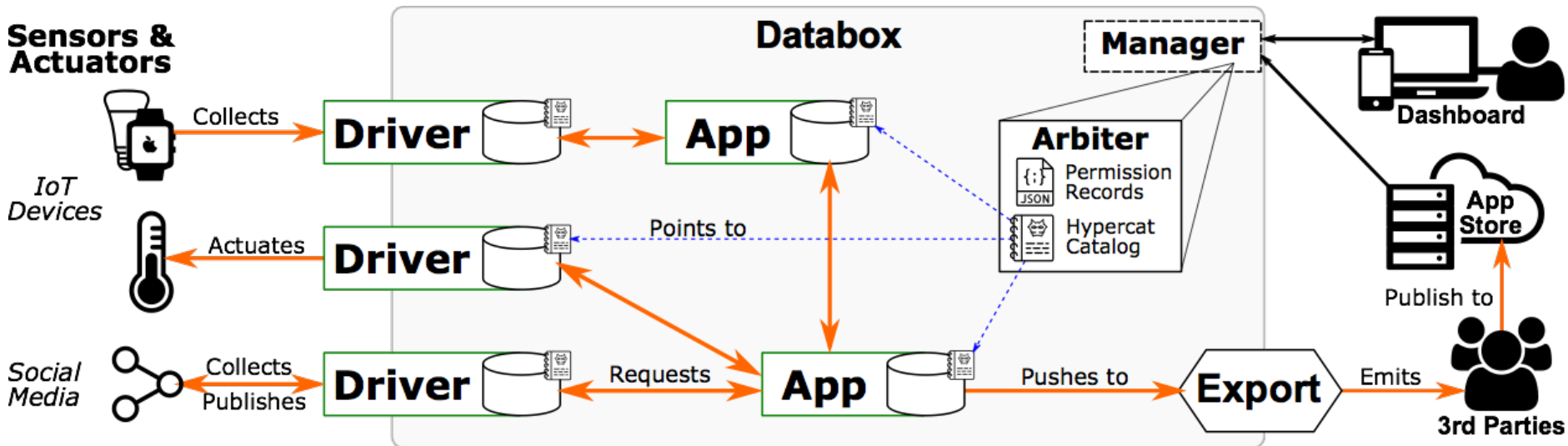
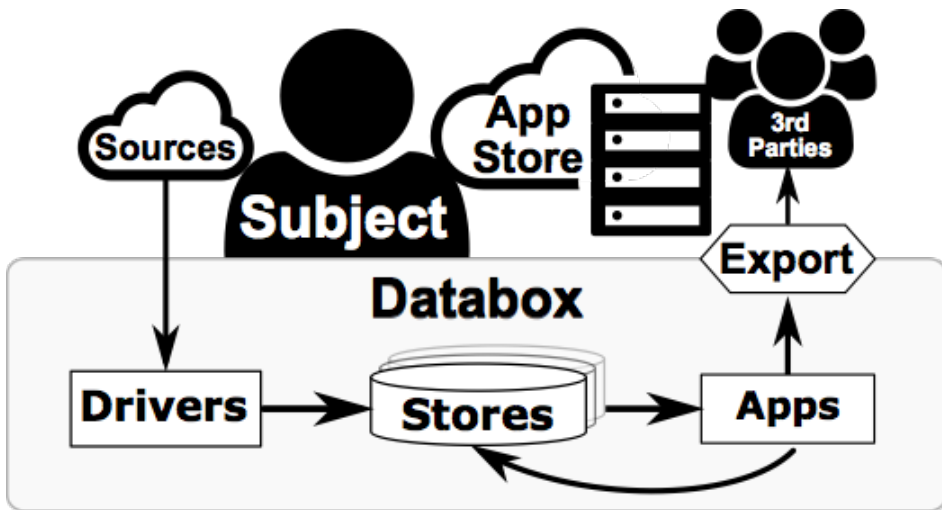
- IoT devices, EHRs



Databox components



Databox components



And in conclusion...

- Local processing will be essential for some data types
 - Databox (come join us!) or similar will be needed
- Embrace such approaches with an overall architecture of distributed analytics across a heterogeneous infrastructure
- Stay ethical and legal and encourage participation!

derek.mcauley@nottingham.ac.uk

@drdrmc

<https://www.horizon.ac.uk>

<https://www.databoxproject.uk>

Abstract

- I would like you to present your perspective on the issues of personal data and approaches to manage public concerns and build trust. In the mental health technology area we are just beginning to see the emergence of personal health records and huge amounts of personal health data generated by wearables and collected by AI personal (healthcare) assistants. Traditionally, at our National Symposium, we have sought challenging perspectives on digital technology from leading figures outside the mental health field. I thought your interests would fit perfectly and a working title for your talk could be something along the lines of: *'Data, what data, who owns it, why does it matter?'*