Prof. Paula Swatman

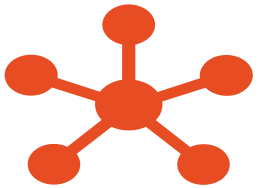
Swinburne University & Bellberry Ltd.

Data, consent and privacy: navigating the increasingly tricky world of data access

Planning & managing data in a rapidly-changing environment

- Our current research environment encompasses so many new and emerging technologies – all generating huge amounts of data:
  - Social media / social networks / social enterprises
  - Big Data / medical data :
    - Registries / repositories / databanks / data warehouses
    - Genomic / epidemiological research
    - Dataset proliferation including: medical & pharmaceutical records / statistical data matching / natural language processing
    - Diagnostic imaging
    - Biometric / longitudinal data
  - Mobile technology / mobile enterprise / BYOD
  - Internet of Things
  - Blockchain
  - Cloud computing / cloud storage
  - Virtual Reality / Augmented Reality / Virtual Environments
  - Artificial Intelligence / predictive analytics

Social media & the loss of privacy



Cavazza (2017) summarises the evolving social media ecosystem as including:

Publishing (blog platforms / wikis / Tumblr / etc.)

Sharing (videos / music / photos / etc.)

Discussing (Facebook / Disqus / Reddit / etc.)

Collaborating (Dropbox / Yammer / Quora / etc.)

Messaging (Snapchat / Facebook Messenger / etc.); and

Networking (LinkedIn / Tinder / Eventbrite / etc.)

Gaitho (2018) identifies 7 ways social media affects individuals & groups:

Politics

Society

Commerce

World of work

Training & development

(Im)morality

Personal relationships

Clinical Registries

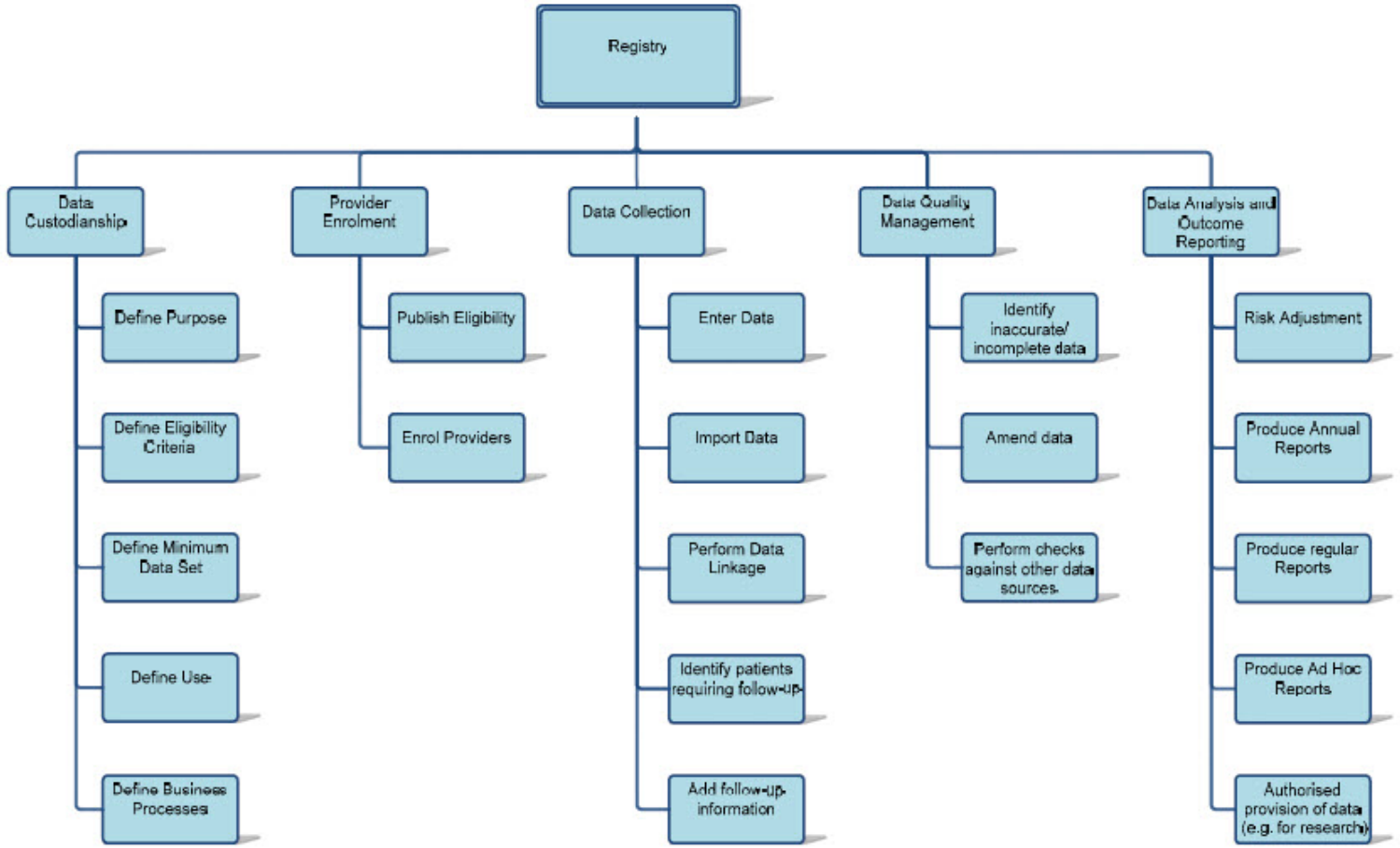
Clinical Trial Registries

- An organisation or website which:
  - Lists clinical trials (being) conducted
  - Allows potential participants to register interest in clinical trial involvement
- Major examples:
  - ANZCTR (Australasia)
  - ClinicalTrials.gov (USA)

Clinical Quality Registries

- Provide complementary data to that obtained from clinical trials:
  - Collect data on the quality of health care provided to patients
  - Enable monitoring/benchmarking of patient outcomes by, e.g.
    - Specific condition/s
    - Treatment approaches or devices
    - Effectiveness
  - Information can be fed back to clinicians & decision makers to improve outcomes

Functional overview of Australian clinical quality registries (ACSQHC, 2014)



Clinical Registry FAQ

- Data entry: via secure online portals or EHRs (electronic health records)
- Consent: opt-out (higher data capture rate & increasingly common: NS 2.3.6) opt-in (NS 2.2.2-6); extended (NS 2.2.14); withdrawal (NS 2.2.19)
- Data custodian (NS 3.1.55, 3.1.57):
  - Can be: individual researcher / agency collecting the data; or an intermediary managing multiple data sources
  - enables access by researchers / participants within limitations on data access or confidentiality
- Data sharing & disclosure must consider:
  - Security / confidentiality / privacy (NS 3.1.56, 3.1.58)
  - Parties to the disclosure & consent requirements (3.1.59, 3.1.60)
  - Value of the information & participant rights (NS 3.1.60)
  - Potential identifiability of individuals (NS 3.1.61, 3.1.62)
  - Future use of data (NS 2.2.14-16, 3.1.60)
- HREC approval required:
  - Before participant recruitment (NS 3.1.36)
  - When registry added to existing trial / study (NS 2.2.14)

Increasingly complex & sensitive health data: Implications for consent or identifiability

- Genomic data: who are the participants?
- Single gene tests are being replaced by next generation sequencing, including whole genome sequencing ➔
- Incidental findings which may also be significant for participants' relatives, because they can:
  - Reveal predispositions to disease
  - Have implications for access to employment, financial services, etc.
  - Be used to stigmatise or discriminate against people
  - Bring to light information about previously unrealised paternity or familial relationships
- Relatives may thus become participants in their own right

- Emerging medical technologies: implications for consent & privacy
- New and emerging technologies have significant implications for 'standard' ethical considerations of risk / privacy / consent, additional to those included in NS Chapter 3.1 – for example:
- Xenotransplantion:
  - Possibility of (epidemic) xenozoonosis ➔ limitation/s on ability to withdraw consent
  - Need for lifelong monitoring for safety ➔ limitations on privacy & confidentiality
- Mitochondrial replacement therapy & human genome editing:
  - Legal and moral issues still to be addressed
- HRECs must balance the interests of:
  - Participants themselves
  - Close contacts / blood relatives

Ethical issues affecting future humans

Heritable genome editing extends these potential effects beyond a single generation into the future

With JK He’s announcement in Nov 2018 of CRISPR-Cas9 editing of human embryos (Lovell-Badge, 2019) the legal and ethical issues of heritable genome editing have become an immediate reality

(Inter)national regulations are still in their infancy – HRECs may well be asked to review applications for human germline editing studies

National Statement Section 3 Revisions (2018)

Section 3: Ethical Considerations in the Design, Development, Review and Conduct of Research

- Ch. 3.1: The Elements of Research
  - Introduction
  - Element 1: Research scope, aims, themes, questions and methods
  - Element 2: Recruitment
  - Element 3: Consent
  - Element 4: Collection, use and management of data and information
    - What is data and what is information?
    - Identifiability of information
    - Data management
    - Secondary use of data or information
    - Sharing of data or information
  - Element 5: Communication of research findings or results to participants
    - Disclosure to third parties of findings or results
  - Element 6: Dissemination of project outputs and outcomes
  - Element 7: After the project
- Ch. 3.2: Human biospecimens in laboratory-based research
- Ch. 3.3: Genomic Research
- Ch. 3.4: Animal-to-Human xenotransplantation

For all research, researchers should develop a data management plan that addresses their intentions related to ***generation, collection, access, use, analysis, disclosure, storage, retention, disposal, sharing and re-use of data and information, the risks associated with these activities and any strategies for minimising those risks***.

The plan should be developed as early as possible in the research process and should include, but not be limited to, details regarding:

- (a) physical, network, system security and any other technological security measures;
- (b) policies and procedures;
- (c) contractual and licensing arrangements and confidentiality agreements;
- (d) training for members of the project team and others, as appropriate;
- (e) the form in which the data or information will be stored;
- (f) the purposes for which the data or information will be used and/ or disclosed;
- (g) the conditions under which access to the data or information may be granted to others; and
- (h) what information from the data management plan, if any, needs to be communicated to potential participants.

Researchers should also clarify whether they will seek:

- (i) extended or unspecified consent for future research (see paragraphs 2.2.14 to 2.2.16); or
- (j) permission from a review body to waive the requirement for consent (see paragraphs 2.3.9 and 2.3.10).

Data management review encapsulated NS 3.1.45

Storing & securing data in the Cloud

Cloud data storage may be the norm today, but HRECs must consider two possible areas of risk where the cloud is concerned

Privacy requirements

- Since March 2014 both private and federal public sector organisations must comply with 13 Australian Privacy Principles (APPs) under the Privacy Act, which regulate the collection, holding, use and disclosure of "personal information"
- The Privacy Act now defines personal information VERY broadly:

  "information or an opinion about an identified individual, or an individual who is reasonably identifiable, whether the information or opinion is true or not and whether recorded in a material form or not"
- Two APPs are especially relevant to cloud storage:
  - APP8 (cross–border disclosure of personal information) regulates the disclosure/transfer of personal information to an entity (including a parent company) offshore
  - APP11.1 (Security of personal information): an organisation must *"take reasonable steps to protect the personal information it holds…"*

Security requirements

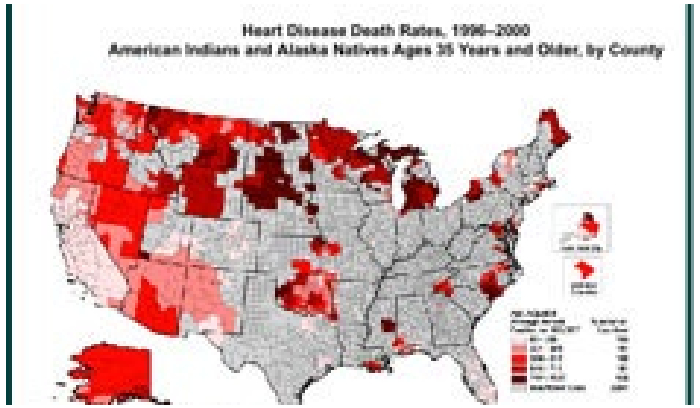
- How secure are your data in a public cloud?
- Can you depend on privacy / reliability claims from your Cloud Service Provider (CSP)?
- Can US and EU governments access your data stored in (mostly US-based) public clouds? Spoiler alert: yes
- And what about government cloud storage – especially those containing sensitive and/or medical databases?
  - July 2018: Singapore suffers a catastrophic breach of its health record database (Cheok, 2018)
  - Feb 2019: 2.5 million Australian families opt out of My Health Record (Barbaschow, 2019)
  - Aug 2019: Oxford researcher demonstrates that GDPR offers a *'heaven for identity thieves'* (Kellon, 2019)

Big Data's implications for research ethics: is a waiver of consent sufficient?

- Mobile apps collect a **LOT** of data, e.g.
  - '… the device's GPS location;
  - which other apps are running and how much network traffic they use;
  - what type of wireless network the device is using;
  - the device manufacturer, model, and OS version;
  - which mobile carrier the device uses;
  - the device's battery level;
  - … the current cell tower ID …
  - a complete network packet trace
  - … complete information on which websites and apps have been used' (Welsh, 2013)
- All this information comes in addition to the data the investigators may actually be seeking!

And there are many potential ethical issues for research involving Big Data (Sula, 2016)

- Invasiveness: the large collections of posts or datasets involved can reveal additional, unrelated information & timestamps / geolocation data embedded in posts and images can identify individual participants
- Informed consent: participants in 'big data' studies are rarely even aware the study is taking place!
- Privacy/anonymity: online datasets never disappear and are increasingly interoperable
- Exploratory research: big data researchers may not know exactly what they are looking for in time to predict all risks
- Researchers have recently demonstrated how to re-identify individuals, even when anonymised and aggregated datasets are incomplete (Nature, 2019)



Big Data example: Blood pressure research

- CDC (2015) offers resources for blood pressure research – both in the USA and globally – including:
  - Interactive atlas of heart disease & stroke
  - Data trends and maps
  - Chronic disease GIS exchange
  - Social determinants of health maps
  - Behavioral risk factor surveillance system
  - Downloadable software for epidemiological analysis

Data Sharing: funding agencies, governments & journals support this concept:
But how do participants feel about data-sharing?

Researchers are increasingly being pressured to share their data:

- Many journals ask for raw datasets
- Funding bodies & governments encourage data-sharing
- Universities & research institutes have policies supporting data-sharing
- And huge medical datasets are becoming evermore accessible…

Researchers & participants, however, have mixed feelings about issues like:

- Ownership of data
- Access by 'hostile' parties with vested interests in findings
- Participant anxiety about where data might end up – and who might see it! (Pearce & Smith, 2011)

EDITORIAL · 31 JULY 2019

Time to discuss consent in digital-data studies

Anonymized data sets are growing and it is becoming easier to identify individuals. Research-consent procedures must be updated to protect people from being targeted.

Refugees, migrants, religious minorities and political dissidents are at risk of being targeted from studies that use anonymized call records. Credit: Petros Giannakouris/AP/Shutterstock

PDF version

RELATED ARTICLES

Can tracking people through phone-call records improve lives?

Estimating the succe re-identifications i incomplete datase generative mode

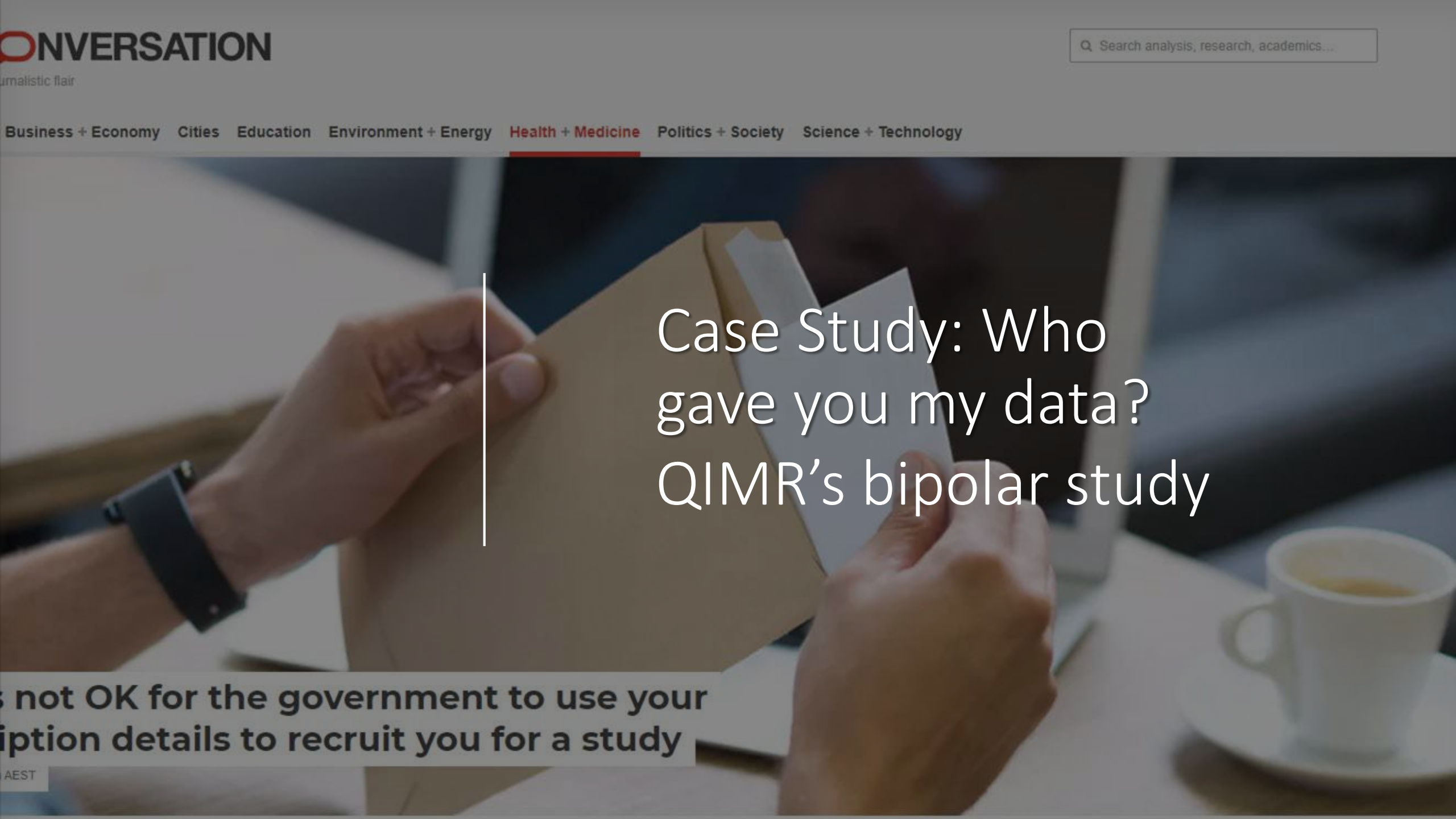
Ethics of research

People today shed data wherever they go. Data flow from their financial transactions, social-media platforms, wearable health monitors, smartphone apps and phone calls.

By tapping massive digital data sets collected by phone providers, technology companies and government agencies, researchers hope to reveal patterns in th data and ultimately to improve lives. Such studies range from an analysis of records in Nepal that showed where people moved to following an eart that aid could be delivered; to estimates of pollution exposure bas data from the Google Maps smartphone app. But relatively lit given to the ethics of how this research is conducted an who supply their data should consent to taking par

Consent for digital data use example:

"Right now, the decision on whether the benefits of digital-data studies outweigh the risks largely falls to the researchers who collect and analyse the data — and not to the people who are unwittingly taking part" (Nature, 2019)



Case Study: Who gave you my data? QIMR's bipolar study

QIMR's Australian 'Genetics of Bipolar Disorder' Study
Opt-in consent ➔ waiver of consent ➔ controversy

- Initially, QIMR Berfhofer sought 5,000 adults treated for bipolar with lithium (QIMR, 2018), using:
  - Opt-in consent with dedicated website + support from Fed. Health Minister
- BUT … media recruitment campaign primarily attracted young, female, capital city respondents
- Researchers then obtained a waiver of consent to gain access to Medicare / PBS records of patients receiving lithium (Aubusson, 2019)
  - After the mail-out the study successfully attracted responses from male and regional participants
  - 6,000 people consented to participate and 4,000 actually participated
- Many outraged invitation recipients believed their psychiatrists had breached their privacy…

Ethical issues :

- PBS data did not include personal medical details, but the mail-out identified those on the list as having a specific mental illness (Doggett et al., 2019)
  - Potential NS 2.1 harms include: psychological, devaluation of personal worth, social & economic
  - NS 4.5 vulnerable participants
- Arnold & Bonython (2019) claim this data release was contrary to DHS' privacy policy
- DHS have clarified their privacy policy in a formal statement, outlining opt-out processes (DHS, 2019)
  - Is it really possible for patients to ensure their medical data are not used for studies of this kind?
  - What is the likely effect of the Data Sharing and Release discussion paper? (DPM, 2019)
- Is a waiver of consent ethical under such circumstances?
  - Is this a breach of trust and/or coercion?
- Would a tactful invitation approach have made any difference to the hostile response?

References – General

- ACSQHC (2014) 'Framework for Australian clinical quality registries', Australian Commission on Safety and Quality in Health Care'; accessible: https://www.safetyandquality.gov.au/sites/default/files/migrated/Framework-for-Australian-Clinical-Quality-Registries.pdf
- Barbaschow, A. (2019) 'Over 30,000 Australians cancelled their My Health Record in under two months', ZDNet, 5 August; accessible: https://www.zdnet.com/article/over-30000-australians-cancelled-their-my-health-record-in-under-two-months/
- Cavazza, F. (2017) 'Social Media Landscape 2017', FredCavazza.net blog, Wordpress,19 April: https://fredcavazza.net/2017/04/19/social-media-landscape-2017/
- CDC (2015) High Blood Pressure Statistics and Maps, Centers for Disease Control and Prevention, December 8; accessible: https://www.cdc.gov/bloodpressure/maps_data.htm
- Cheok, J. (2018) 'SingHealth hacked: records of 1.5m patients, including PM Lee Hsien Loong, stolen', The Business Times, 20 Jul; accessible: https://www.businesstimes.com.sg/government-economy/singhealth-hacked-records-of-15m-patients-including-pm-lee-hsien-loong-stolen
- Gaitho, M. (2018) What is the real impact of social media?, Simplilearn.com resources, 20 July: https://www.simplilearn.com/real-impact-social-media-article
- Kellon, L. (2019) 'Black Hat: GDPR privacy law exploited to reveal personal data', BBC News 8 August; accessible: https://www.bbc.com/news/technology-49252501
- Lovell-Badge R. (2019) CRISPR babies: a view from the centre of the storm. *Development.* 2019;146. DOI: 10.1242/dev.175778
- Nature (2019) 'Editorial: Time to discuss consent in digital-data studies', vol 572, 31 July; accessible: https://www.nature.com/articles/d41586-019-02322-z?sf216722087=1
- Pearce & Smith (2011) 'Data sharing: not as simple as it seems' EnvironHealth: 10: https://core.ac.uk/download/pdf/13114975.pdf
- Sula, C.A. (2016) 'Research Ethics in an Age of Big Data', Bulletin of the Assoc. for Science & Technology, January: http://onlinelibrary.wiley.com/doi/10.1002/bul2.2016.1720420207/full
- Welsh, M. (2013) 'The ethics of mobile data collection', Blog post, 22 January: http://matt-welsh.blogspot.com/2013/01/the-ethics-of-mobile-data-collection.html

References – Case Study

- Arnold, B.B. & Bonython, W. (2019) 'No, it's not OK for the government to use your prescription details to recruit you for a study', The Conversation, 31 July; accessible: https://theconversation.com/no-its-not-ok-for-the-government-to-use-your-prescription-details-to-recruit-you-for-a-study-121122
- Aubusson, K. (2019) 'Medicare data used to recruit people with bipolar for research', The Sydney Morning Herald, 29 July, accessible: https://www.smh.com.au/healthcare/medicare-data-used-to-recruit-people-with-bipolar-for-research-20190722-p529k9.html
- DHS (2019) 'Medicare data collection for medical research purposes', Dept of Human Services, 2 August; accessible: http://mediahub.humanservices.gov.au/ontherecord/2-august-2019-medicare-data-collection-for-medical-research-purposes/
- DPM (2019) 'Data Sharing and Release Legislative Reforms Discussion Paper', Dept of Prime Minister & Cabinet, September; accessible: https://www.datacommissioner.gov.au/data-sharing/discussion-paper-PIA
- Doggett, J. et al. (2019) 'Debate over release of Medicare data continues', croakey, 1 August; accessible: https://croakey.org/debate-over-release-of-medicare-data-continues/
- QIMR (2018) 'Cracking the genetic code of bipolar disorder', Australian Genetics of Bipolar Disorder Study website, 20 November; accessible: https://www.geneticsofbipolar.org.au/cracking-the-genetic-code-of-bipolar-disorder/