

big data


big responsibility

Trent Yarwood



trent@futurewise.org.au

<https://futurewise.org.au>

 @FutureWiseAU

About **FUTURE**WISE

- Completely politically independent NGO
- Public policy research and advocacy
- Focus on technology, health, education
 - And their impact on society at large
- Australian professionals from tech, law, medicine, academia

Talk outline

- Data Retention (briefly)
 - Problems with Data Retention
 - Implications for healthcare
- Healthcare data breaches
- The future of public health?
- Healthcare workers and IT

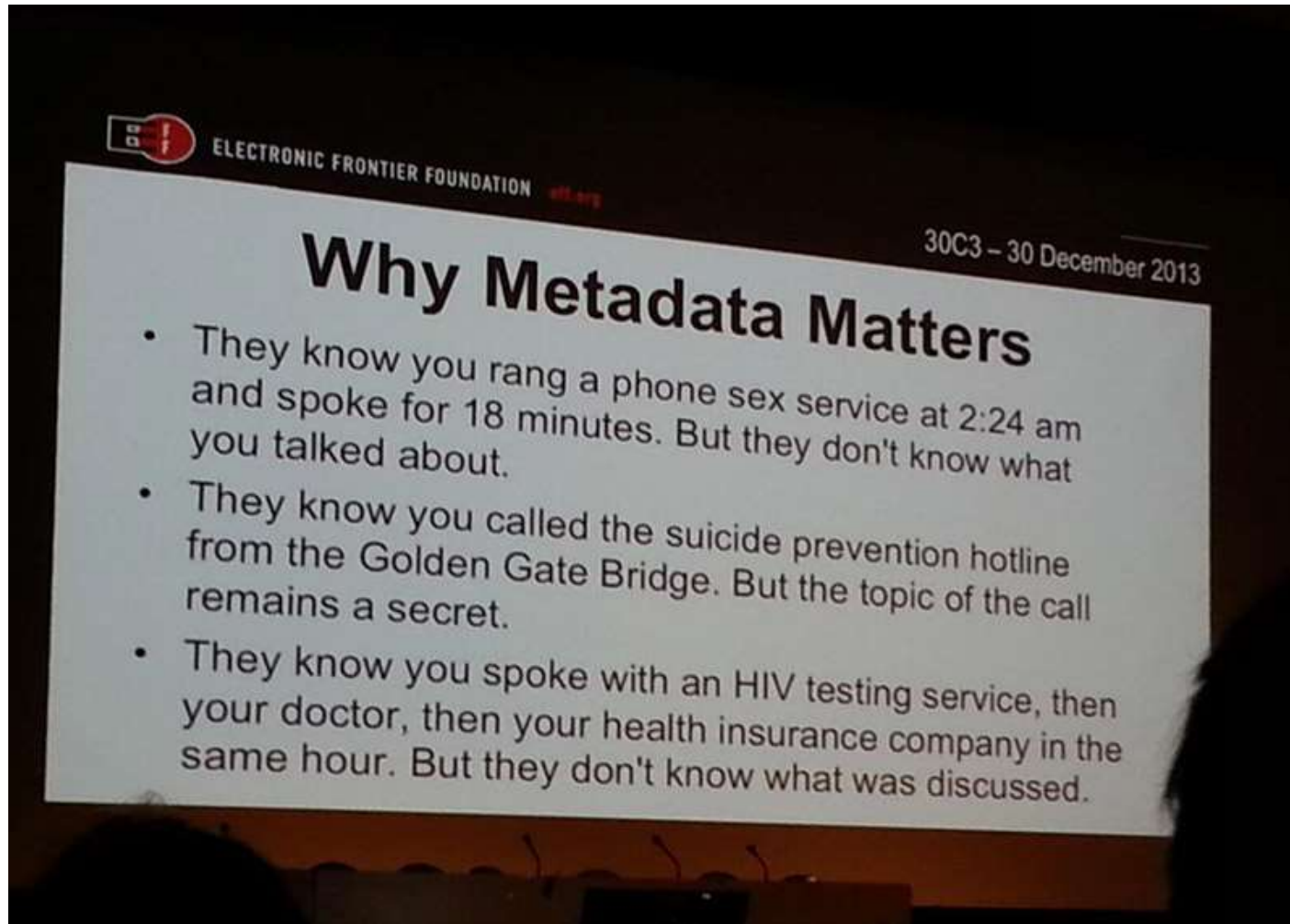
Data Retention

- Requires your telecommunications company to keep records of the communications you make
 - Home phone
 - Mobile telephone
 - Internet (fixed, wireless, mobile)
- Content of communication excluded in Bill
 - Otherwise very broad definition

what is retained?


- Substantial service / account info
- The source and destination of the communication
 - Including username, IP address, phone #
- The type of communication (email, chat)
- The location of the device at the start and end of the communication
- ...but not the content

Is that so bad?



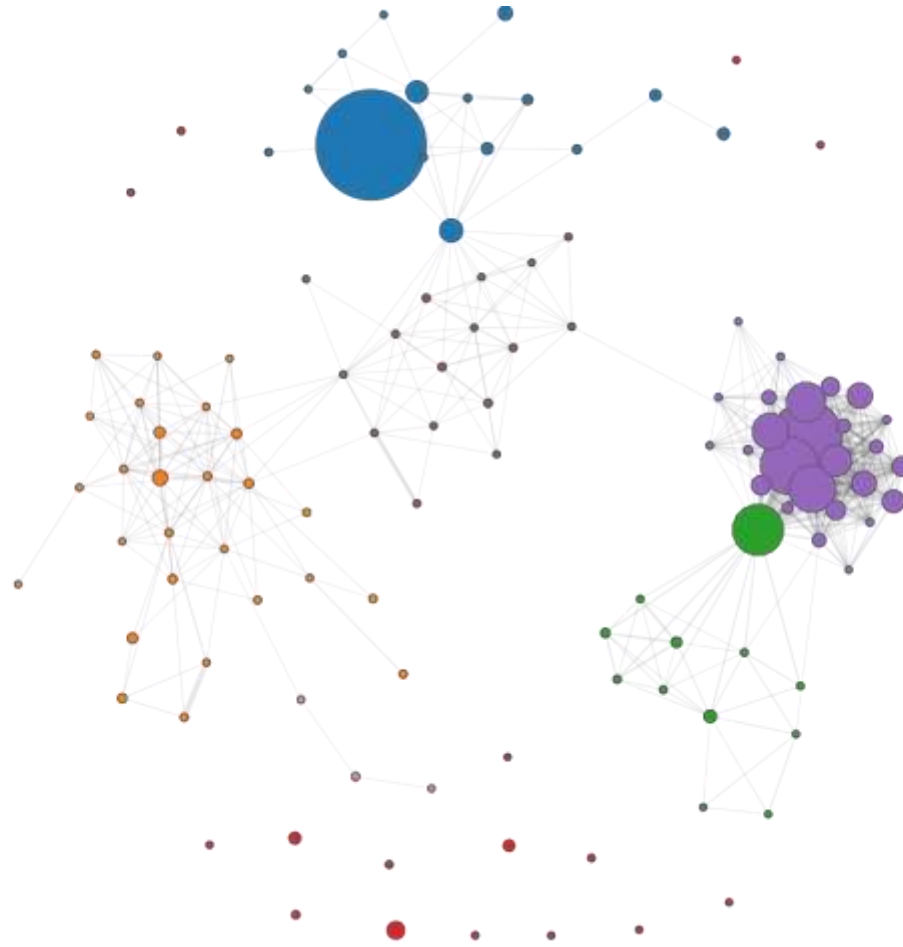
30C3 Conference, December 2013 – Image by Electronic Frontier Foundation via Twitter

<https://futurewise.org.au>


 @FutureWiseAU

My email network

immersion.media.mit.edu

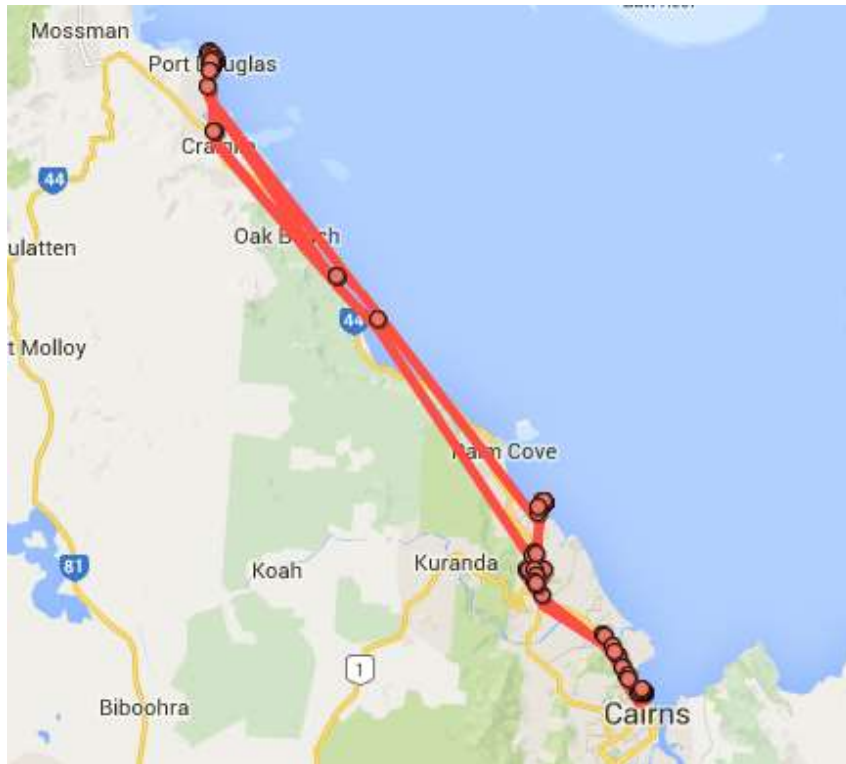


<https://futurewise.org.au>

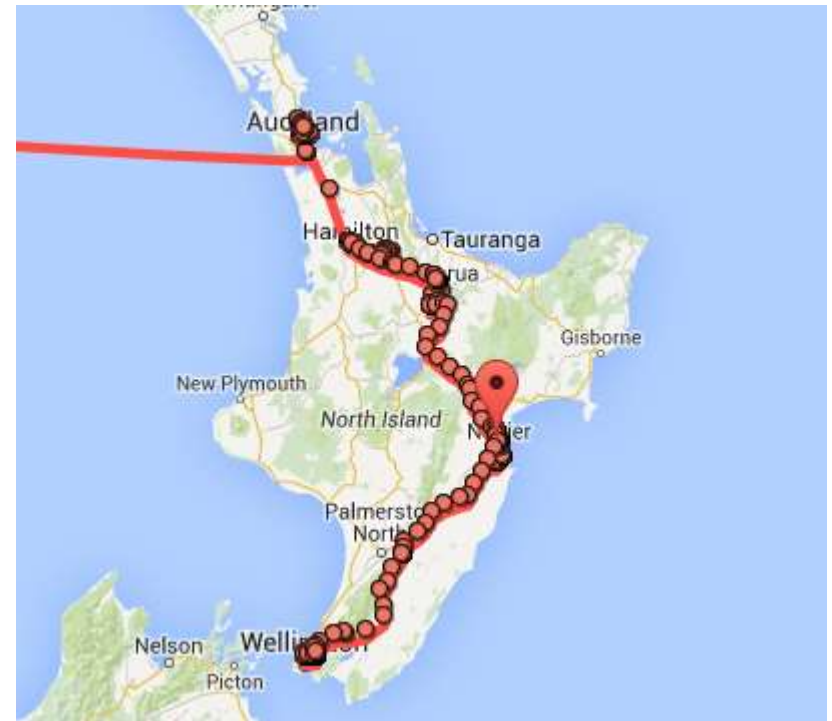
 @FutureWiseAU

Personal tracking

22-23 May 2015




March / April 2015



See your info: <https://www.google.com/maps/timeline/>

<https://futurewise.org.au>

 @FutureWiseAU

SO WHAT'S the problem?

- Because this data is “not privacy sensitive”, it is available to LEAs *without a warrant**
- LEAs made 550,000 requests in 2013
- “Metadata” allows:
 - Continuous motion tracking
 - Extensive network analysis

Practical considerations

- Your ISP will need to store this data for 2 years
 - Cost
 - ID-theft honeypot datacenters
- The Federal Police have already had at least ~~two~~ five data breaches
 - And there is no requirement for them to delete the data after using it

Health implications

- Obviously we all have a requirement for confidentiality
- Some services have more than others
 - Sexual health
 - O/G & Paeds
- If your patient requests absolute confidentiality – what are you going to do?

Family Court

- Family court lawyer requests disclosure of metadata
- Several SMS to unknown mobile number
- Phone connected to tower between 5pm-7pm at location not matching work or home address
- Home internet search leading to IP belonging to sexual health advice website
- Phone contact with Cairns SHS
- Phone connects to Cairns North cell tower
- Following day, SMS from CSHS results number
- Internet session connecting to IP for www.letthemknow.org.au

Facebook is evil (1)

2013

- “Emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness”

2017

- FB can “can monitor users' posts and determine when they are feeling “stressed, defeated, overwhelmed, anxious, nervous, stupid, silly, useless, or a failure.” and detect “moments when young people need a confidence boost.”

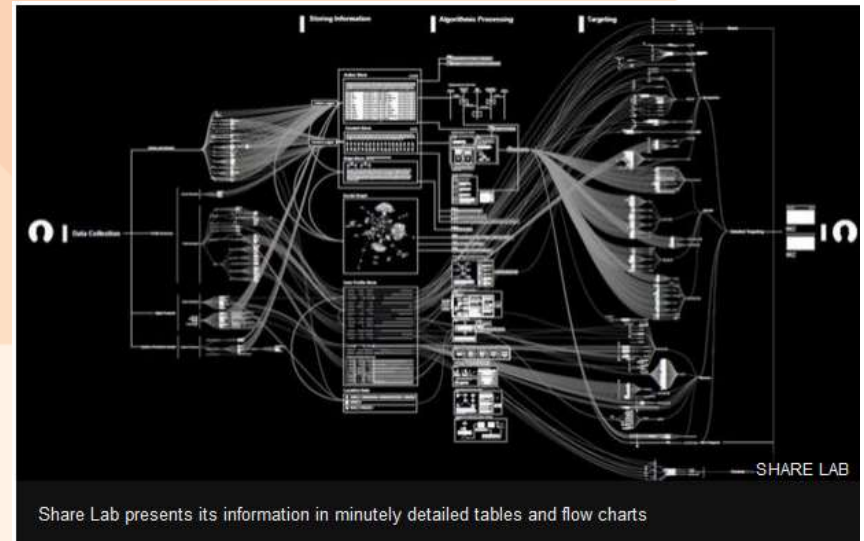
Facebook is evil (2)

How Facebook's tentacles reach further than you think

By Joe Miller
Business reporter

🕒 26 May 2017 | Business


     Share



<http://www.bbc.com/news/business-39947942>

<https://labs.rs/en/>

<https://futurewise.org.au>

 @FutureWiseAU

And bad for privacy

PEOPLE YOU SHOULD NOT KNOW 8:20 '14 4:27 PM

Facebook recommended that this psychiatrist's patients friend each other




- Mobile numbers on patients' phones?
 - Or on doctor's phone?
- Browsing history?
- Common location?

- Collateral collection of other personal data

<http://fusion.net/story/339018/facebook-psychiatrist-privacy-problems/>


<https://futurewise.org.au>

 @FutureWiseAU



HEALTHCARE DATA LINKAGE

<https://futurewise.org.au>

 @FutureWiseAU

new FRontier

- Open data is the movement to making scientific or civic data freely accessible
- Allows researchers, citizen scientists or the curious to discover new things
- Increases your research's citations

Open Data

https://data.gov.au/organisation/queensland-health?q=cairns

HCY HIV Hospital Apps Infection Control

Data

- > Datasets
- > Categories
- > Organisations
- > Tags
- > Maps and geospatial
- > Other data types

Queensland Health

4 datasets found for "cairns"

[Consultancy spending - Cairns and Hinterland Hospital and Health Service \(HHS\)](#)
1 recent view 890 total views
Consultancies - Cairns and Hinterland Hospital and Health Service

[Overseas travel - Cairns and Hinterland Hospital and Health Service \(HHS\)](#)
1 recent view 217 total views
Overseas travel undertaken by officers of Cairns and Hinterland Hospital and Health Service

data.gov.au

Datasets Organisations About Site Statistics

Request Data Use Cases National Map Toolkit

Search data

Popular tags: [Earth Sciences](#) [Oceans](#) [GA Publication](#)

The new way to find open data

Visit [search.data.gov.au](#) to try our new open data search.

As part of the National Innovation and Science Agenda we are working with Data61 to develop the next generation of data.gov.au. Part of this work includes making our prototypes

data.gov.au statistics

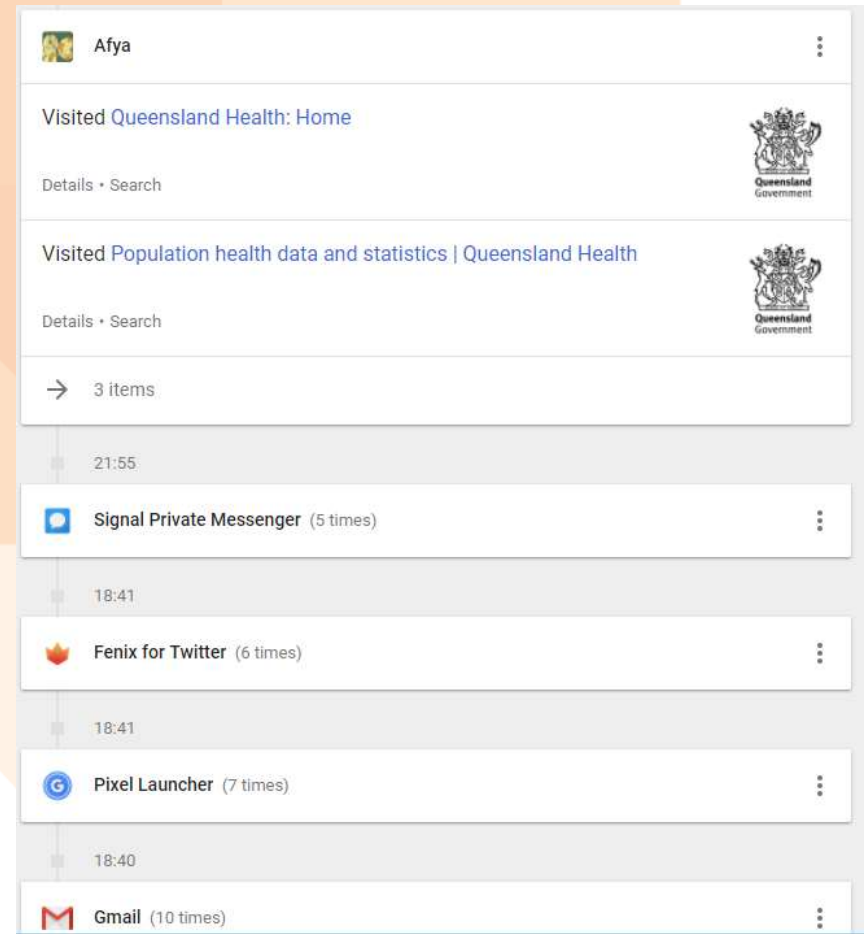
49.9k	9k	36.9k	24
discoverable datasets	API-enabled resources	openly licensed datasets	unpublished datasets

<https://futurewise.org.au>

@FutureWiseAU

Big Data

- Everything we do these days generates data
 - Lots of data – 2.5×10^{18} bytes per day (IBM)
 - 25 million terabytes
 - Think about the phone tracking picture, Facebook, your Fitbit, etc etc
- Mass analysis of data
- Much of this data is in “walled gardens”



Google deepmind

- Google's subsidiary artificial intelligence research branch
- Entered partnership with Royal Free NHS
- Patient records run through machine learning computers

New app helping to improve patient care

Print | Share




27 February 2017

A new app which detects early signs of kidney failure is now being used to improve care for some of the Royal Free's most vulnerable patients.

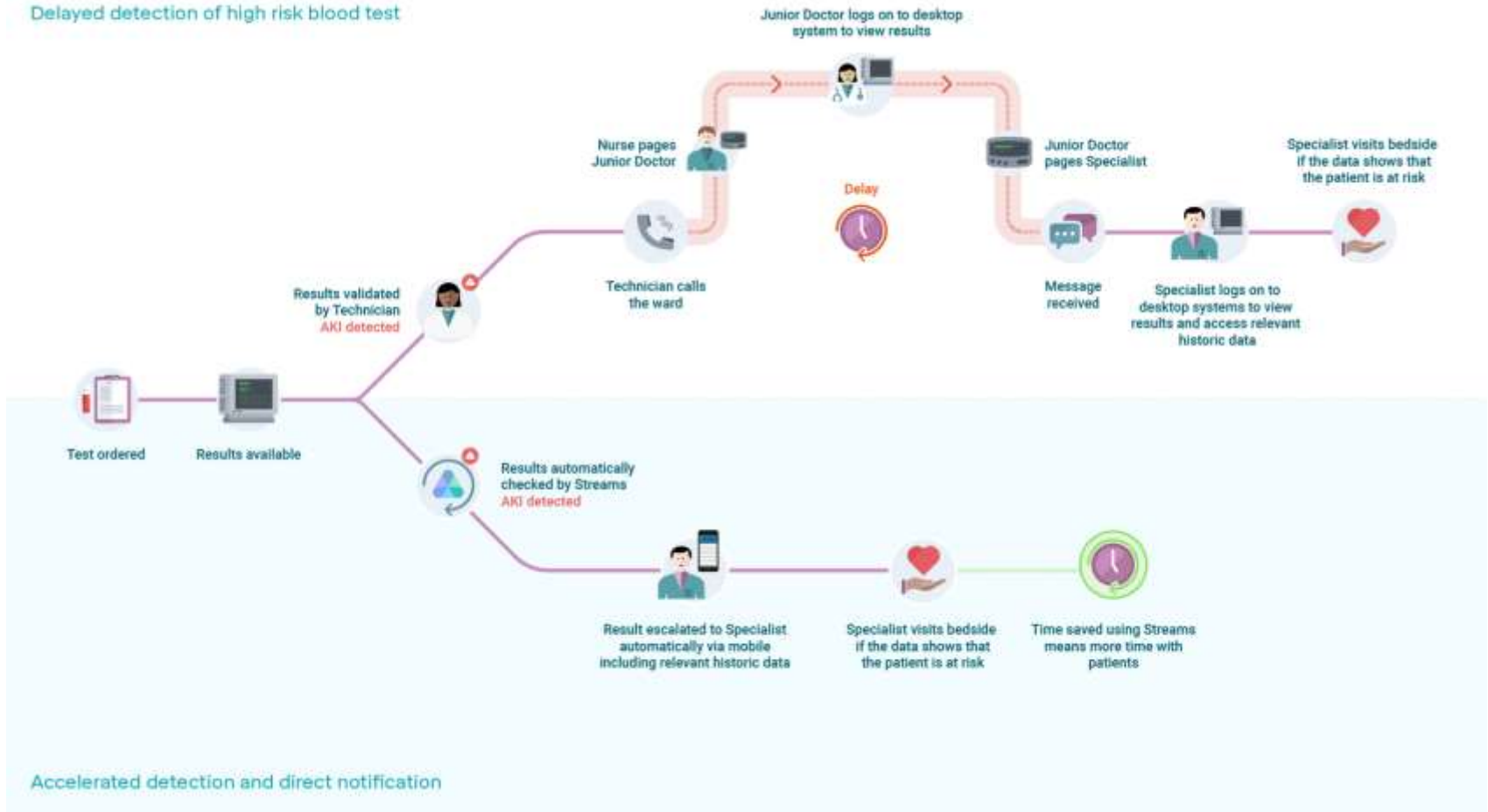
Using an instant alert system, the app, known as Streams, directs clinicians to patients who are at risk of or who have developed a serious condition called acute kidney injury (AKI).

<https://futurewise.org.au>

 @FutureWiseAU

Streams workflow

Delayed detection of high risk blood test



Improved outcomes

A day in the life of Streams at the Royal Free
February 2017



2,211

Blood tests analysed

66

Changes in kidney function

23

AKI Alerts issued

11

Cases requiring action

11


Cases requiring action

⚠ Including potentially fatal sepsis and rise in potassium

🕒 > 1 min 60% reviewed in less than a minute

📱 2 critical cases reviewed remotely

<https://futurewise.org.au>

 @FutureWiseAU

What cost?

Revealed: Google AI has access to huge haul of NHS patient data

A data-sharing agreement obtained by **New Scientist** shows that Google DeepMind's collaboration with the NHS goes far beyond what it has publicly announced



New Scientist – April 2016

Google DeepMind NHS data deal was 'legally inappropriate'




Streams is designed to help monitor people with kidney disease
DeepMind

By Timothy Revell

Google DeepMind received 1.6 million identifiable personal medical records on an "inappropriate legal basis", according to a letter written by Fiona Caldicott at the UK's National Data Guardian, a government watchdog that monitors patient confidentiality. The letter obtained by Sky News was sent to the medical director of the Royal Free NHS Trust in London on 20 February. The data sharing agreement between the trust and DeepMind was first revealed by a *New Scientist* investigation last year.

New Scientist – May 2017

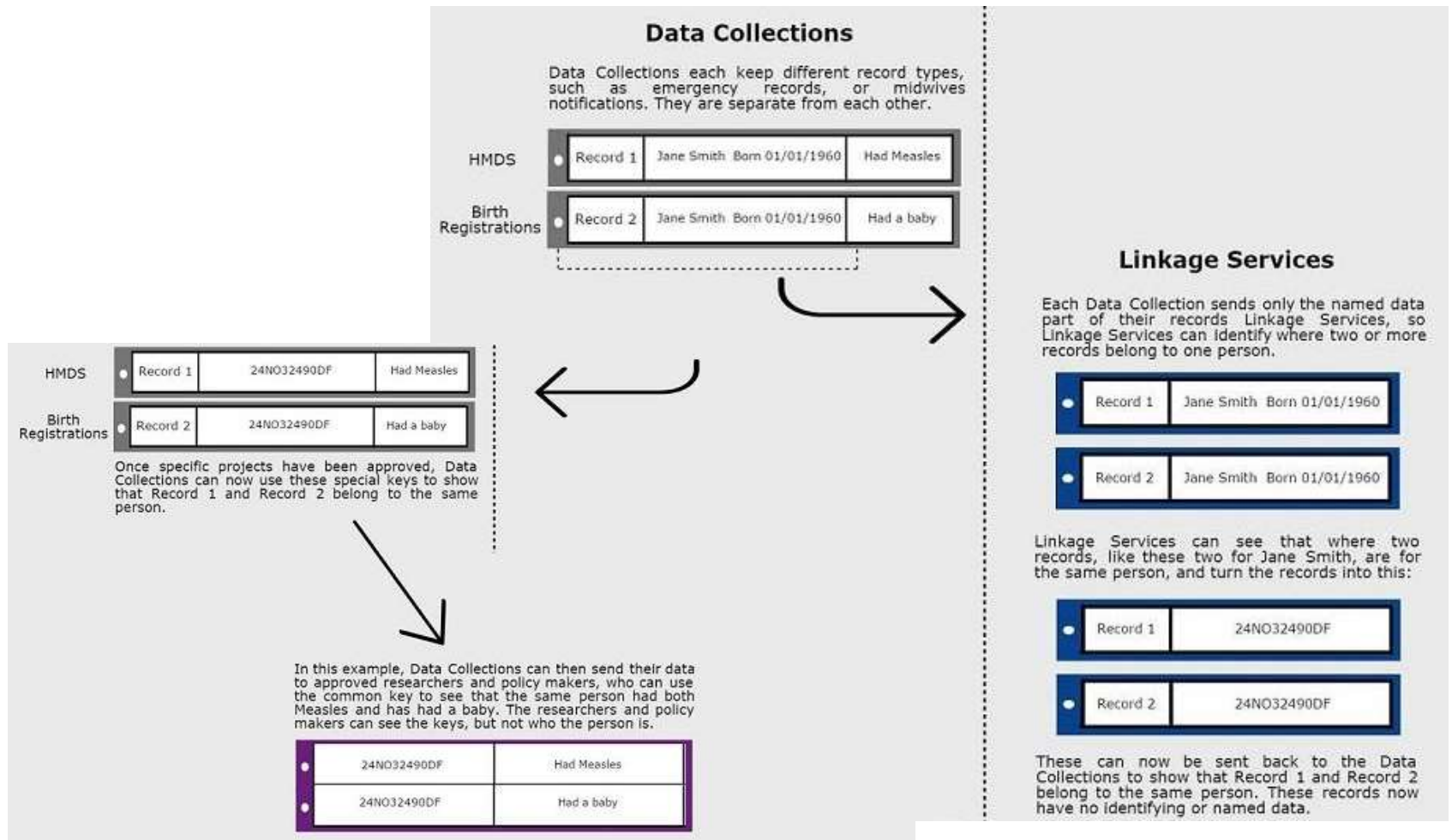
<https://futurewise.org.au>

 @FutureWiseAU

Data Linkage

- The process that is used to combine information relating to an individual from multiple sources
- Absolutely has the potential to generate public health benefits
- But is unequivocally a risk to privacy

How data is linked



Medical Opinion

- “In our increasingly complex world, members of the public should be reassured that, *without risk to privacy*, the personal information collected from them can be aggregated and used in such ways to inform and protect them, and to increase the efficiency and effectiveness of public services.”
 - Fiona Stanley – Epidemiologist, Telethon Kids

Trust the government


- Details of asylum seeker detainees
 - Including reasons for seeking asylum
- Personal mobile phone numbers of members of federal parliament
- Detailed demographic details of 50,000 Australians and 5,000 APS staff Finance
- Robodebt, Censusfail etc
 - And the deliberate 'doxxing' of Andi Fox

Red Cross data breach

- 550,000 donors 2010-2016 affected
- Donor questionnaire database posted to publicly accessible webpage
 - Name, address, DoB
 - Answers to donor questions
- Any risky behaviours
 - Male-male sex, sex with sex workers, injecting drug use

<https://www.troyhunt.com/the-red-cross-blood-service-australias-largest-ever-leak-of-personal-data/>

<https://futurewise.org.au>

 @FutureWiseAU

Medicare Data Leak

- Medicare published a sample of “deidentified” MBS/PBS data on data.gov.au
- Random 10% of Medicare Nos
- All claims 1984-2014
- Two datasets:
 - MBS (services)
 - PBS (prescriptions)

ABS Linkage


- ABS definition of statistical linkage key
 - XXXXXXDDMMYYYYYN
 - Published online
 - ARWRE010119011
- Partial details of the algorithm used to obscure data was published with the datasets

Unimel b Research (1)

- Using only publicly available data
- Decrypted the provider ID numbers
- Implications
 - Breached privacy of doctors
 - Patient IDs were likely also obtainable
 - Linkage of provider info
- Hasty Govt legislation making reID illegal

<https://pursuit.unimelb.edu.au/articles/understanding-the-maths-is-crucial-for-protecting-privacy>

<https://futurewise.org.au>

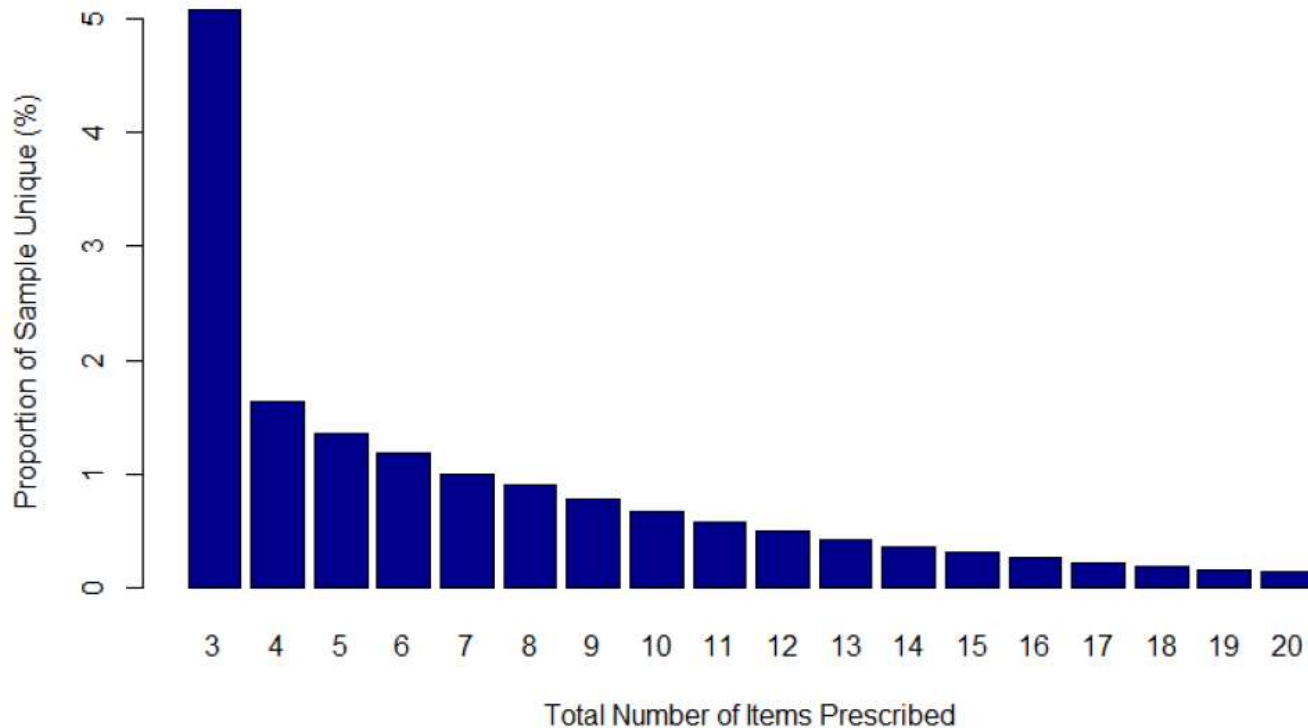
 @FutureWiseAU

Make crumbs into bread

- The same group were able to link publicly available data with the dataset to positively identify individuals including sportspeople and politicians
- “Decryption” not necessarily required
- Eg
 - Sportsperson with appendicitis
 - Demographics from Wikipedia
 - Date of MBS item for appendicectomy

Uniqueness - open data


Percentage of 2003 Patients with a Unique Set of 3 Items



- Multiple strengths and brands reduce the entropy of a combination

Culnane, Rubinstein & Teague. Health data in an Open World. arXiv cs - <https://arxiv.org/abs/1712.05627>

<https://futurewise.org.au>

 @FutureWiseAU

Trent CRUMBS

- JETACAR the year I started med school
- Children aged 8 & almost 5

Health Practitioner: Dr Trent Phillip Yarwood

Personal Details

Sex: Male

Languages (in addition to English):

Qualifications:

- Bachelor of Medicine / Bachelor of Surgery, University of Queensland, Australia, 2002
- Fellowship of the Royal Australasian College of Physicians, The Royal Australasian College of Physicians, Australia, 2011

Principal Place of Practice

Suburb: Cairns

State: QLD

Postcode: 4870

Trent CRUMBS

The use of antibiotics in
food-producing animals:
antibiotic-resistant bacteria in
animals and humans

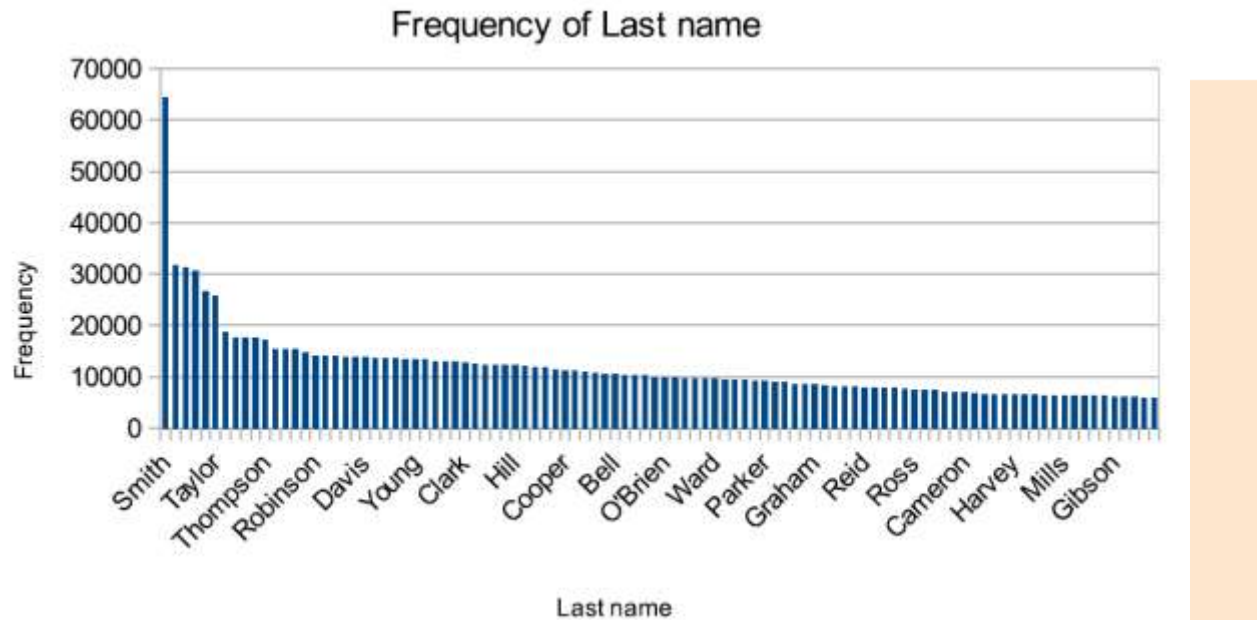
Report of the
JOINT EXPERT ADVISORY COMMITTEE
ON ANTIBIOTIC RESISTANCE
(JETACAR)

© Commonwealth of Australia 1999

ISBN 1 86496 061 2

- 4 year degree >2002
 - Postgrad UQ MBBS
- Finished school
1993-1995
- Link 3 records with
same Surname ID &
Postcode 487X
 - Born 1976-1978
 - Born 2008-2010
 - Born 2013-2015


Yarwood v SMITH



- “Frequency attack”
- Search for the modal frequency encrypted hash
 - Reduces the number of possible records

Culnane C, Rubinstein BIP & Teague V. arXiv cs.CR: <https://arxiv.org/abs/1802.07975>

<https://futurewise.org.au>

 @FutureWiseAU

Govt plans

- AMA National Conference – Melbourne, 26th May 2017
- Health Minister Greg Hunt
“[Opt-out myHealthRecord] means that we will have the capacity for some of the best health data analytics in the world, and that means that we’ll be able to do preventive health on a national basis in an utterly informed way”

Secondary use




HealthConsult, as leader of a consortium consisting of two other firms and eight subject matter experts, was engaged by the Department of Health (the 'Department') to:

“develop a Framework for the secondary use of data held in the My Health Record system for research, policy, system use, quality improvement and evaluation activities”

For the purposes of this project 'secondary use' is defined as using the information in the My Health Record system for purposes other than the provision of direct healthcare to the individual person, which is considered to be the primary use. The use of data solely for commercial and non-health related purposes is considered out of scope.

<https://www.myhealthrecorddata.healthconsult.com.au>

<https://futurewise.org.au>

 @FutureWiseAU

WAIT...what??

The questions that arise when considering the secondary use of health data are the same questions that arise in any attempt to achieve the public good. Secondary use of health data has the potential to enhance future healthcare experiences for patients by enabling the expansion of knowledge about disease and appropriate treatments, strengthening the understanding about effectiveness and efficiency of service delivery, supporting public health and security goals, and assisting providers in meeting consumer needs.¹¹ It allows a range of organisations to conduct research and innovate to improve health and healthcare outcomes, which can in turn improve well-being, productivity and the efficient use of resources.

Garbage in garbage out

- System Performance and Reporting
 - SPR Dashboards
- “Bloodstream infections”
 - SABs are OK
 - What about the rest?
- Commission list of HAC?



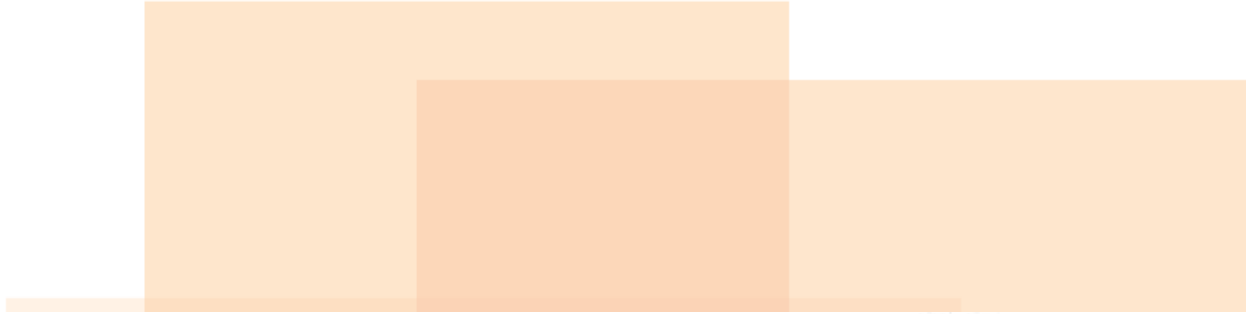
Local AUDIT

- 21 patients tagged by clinical coding as having “bloodstream infection”
 - 13 had no positive BCs during episode
 - 4 had CNSBCs considered NS by ID/CM
 - 1 had Community-onset BSI
 - 3 did actually have haBSIs
- $3/21 = 14\%$
- This is what our funding penalties (and our big data predictions) are based on

HERE'S the issue

- Healthcare workers are not really all that great at confidentiality
 - See also: The curtain of silence
 - “Charts in cupboards” mentality
- Almost all healthcare workers are completely terrible at computering
- Almost everyone everywhere is terrible at computer security
- Technical solutions / sanctions don't prevent “authorised” but improper access

Doctors CAN'T COMPUTER




The field of record linkage has a long history, dating back nearly 50 years [25]. Privacy Preserving Record Linkage (PPRL) dates back over a decade [22]. A significant proportion of the literature has been published outside of Computer Science and Information Security venues, with a particular prevalence for publication within the medical domain. This has resulted in proposals not being subjected to the normal level of rigour and analysis associated with information security. A number of schemes do not achieve their claimed privacy properties. Cryptographic primitives such as hashing have been used without a clear understanding of the security properties they provide and the attacker model they defend against. Persistent mistakes leading to security problems are repeated in many papers, even when the security issues have been identified.



Culnane C, Rubinstein BIP & Teague V. arXiv cs.CR: <https://arxiv.org/abs/1802.07975>

<https://futurewise.org.au>


 @FutureWiseAU

More siloing

- Huge disconnect between the needs of healthcare workers and IT departments
- “They just need to double the Gigabits”
 - staff specialist in this hospital
- “You just need to patch your OS”
 - geeks everywhere in response to WannaCry

<https://futurewise.org.au/2016/08/10/never-the-twain/>

<https://futurewise.org.au>

 @FutureWiseAU

What can we do?

- Consider the implications of your communications
 - Patients, colleagues and self
- Information not collected can't be leaked
- Get more engaged in eHealth
 - Without clinicians, the geeks will do it for us

📌 Pinned Tweet



Trent   @trentyarwood · 10 Apr 2016

I'm a doctor and a geek - I've opted out of a myHealthRecord. So should you.

Before we finish, could you share with us why you don't want a digital health record

Please select at least one reason:

- I have no use for a digital health record
- I prefer to manage my medical records on my own
- I prefer that my doctor manages my medical records
- I am concerned about others having access to my private medical information
- I am concerned about the security of my medical information stored online
- I do not trust what the Government or others will do with my medical information


💬 19

↻ 189

❤️ 184



<https://futurewise.org.au>

 @FutureWiseAU


Visit our website for more information: <https://futurewise.org.au>

Join our discussion forum: <https://forum.futurewise.org.au>

Follow us on Twitter: @FutureWiseAU

THANK YOU

<https://futurewise.org.au>

 @FutureWiseAU