THE ART OF THE POSSIBLE: NEXT STEPS FOR DATA LINKAGE IN AUSTRALIA

AHHA DATA DAY 27th June 2017 Suzanne Robinson and James Boyd

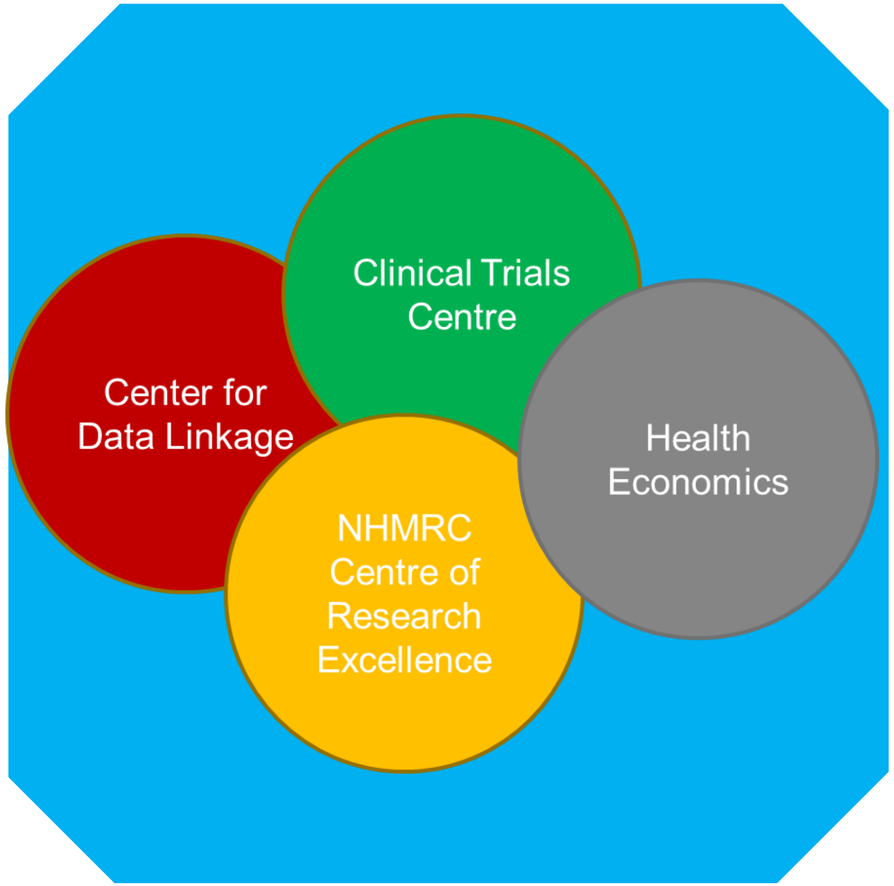
School of Public Health, Faculty of Health Sciences

Bentley Campus, Perth

Session Outline

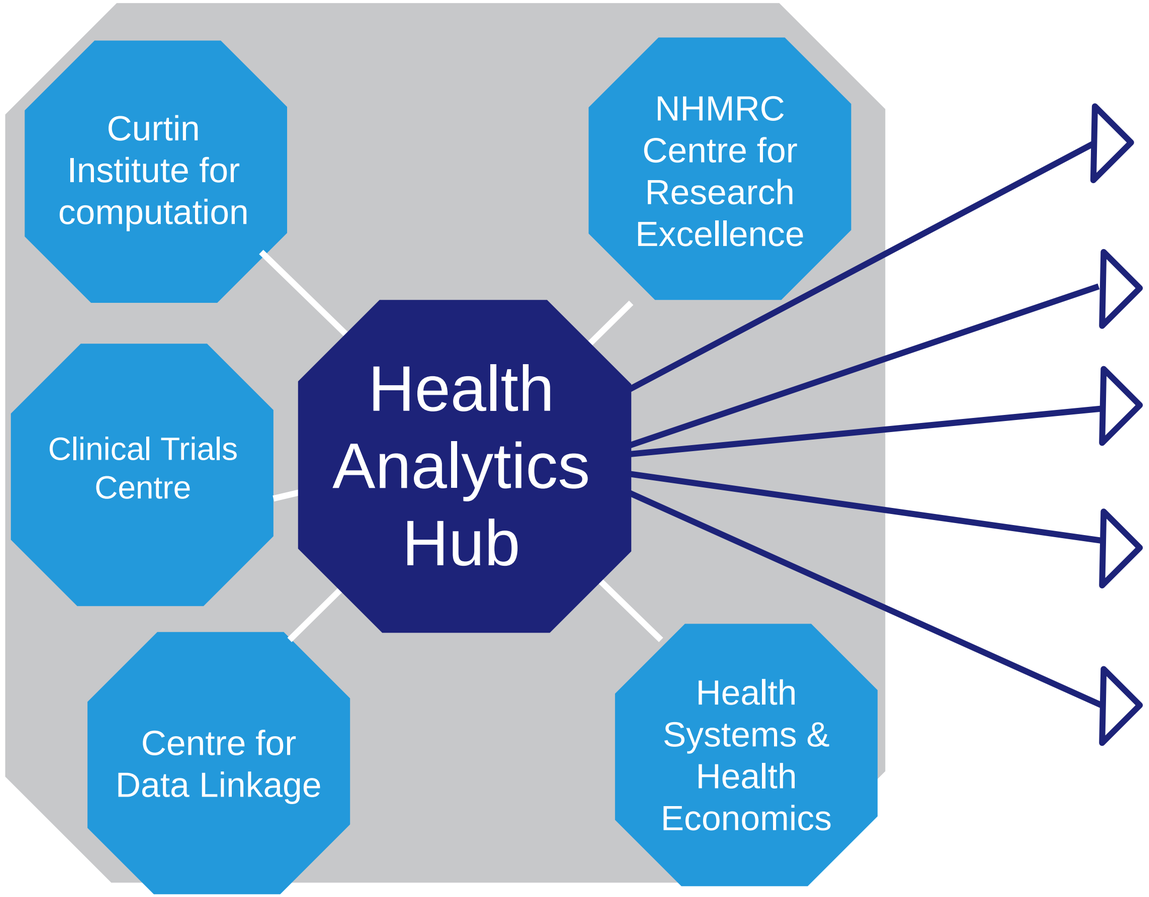
- Context setting – developments in research and analytics at Curtin (WA)
- Data Linkage focus on
  - Benefits of data sharing
  - Data governance
  - Case study data linkage projects

Health Research and Analytics Hub



- Increasing ability to collect large volumes of data within IT systems
- Analyse and interpreting “big health data” - a “revolution” in healthcare.
- Increased requirement to provide evidence of the effectiveness and long term safety of medical & surgical interventions
- A growing emphasis and financial incentives to hospitals, primary care and general practice with a shift in funding focus on prevention of illness

Health Analytics Hub: working with our partners and broader stakeholder groups



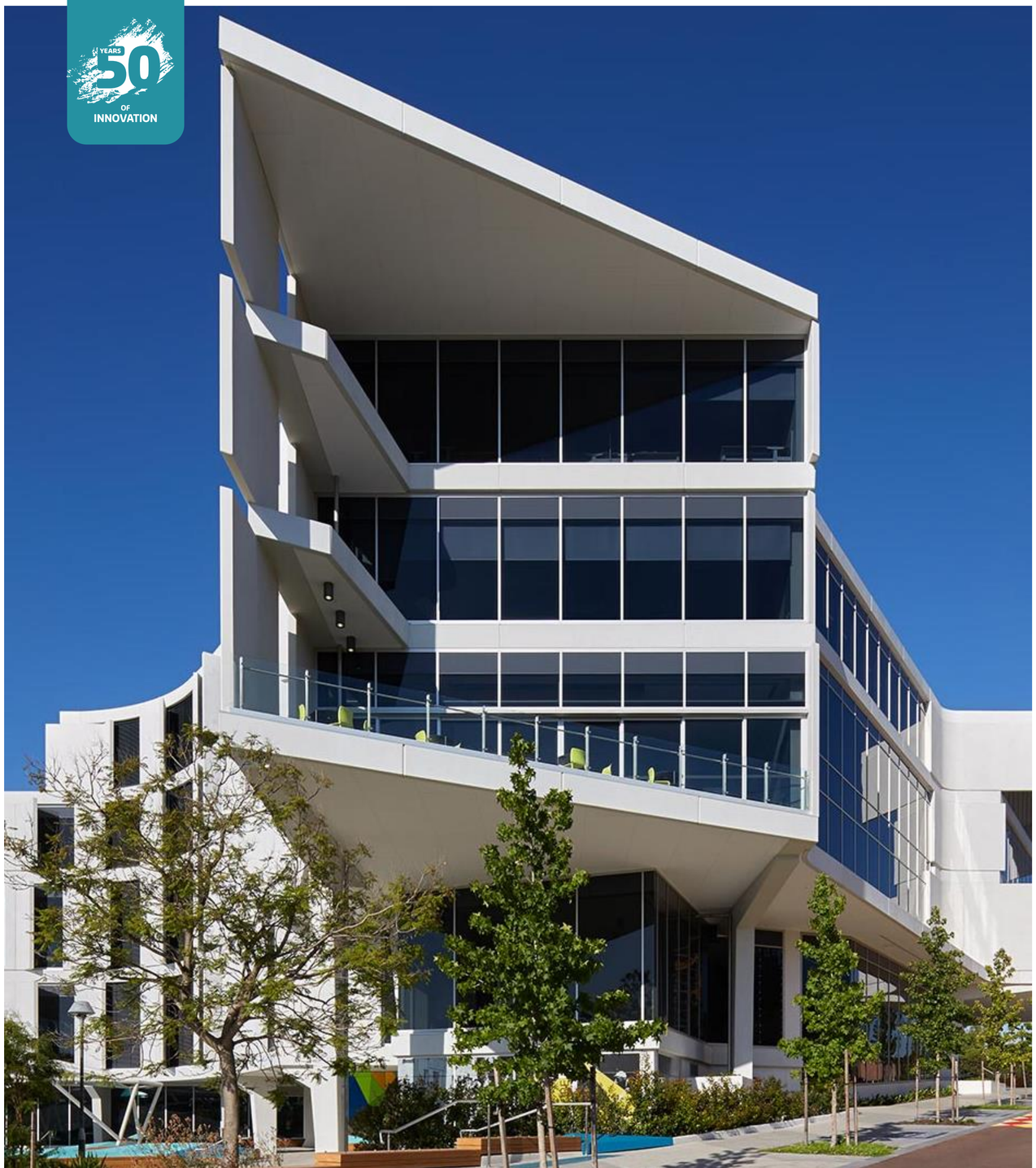
Lead clinical trials and outcomes research capabilities for Western Australia

Lead health systems and population health research for Western Australia

Develop capacity and capabilities in health economics

Research translation & solutions development through access and use of large data sets

Translation & Commercialisation



Curtin Collaboration

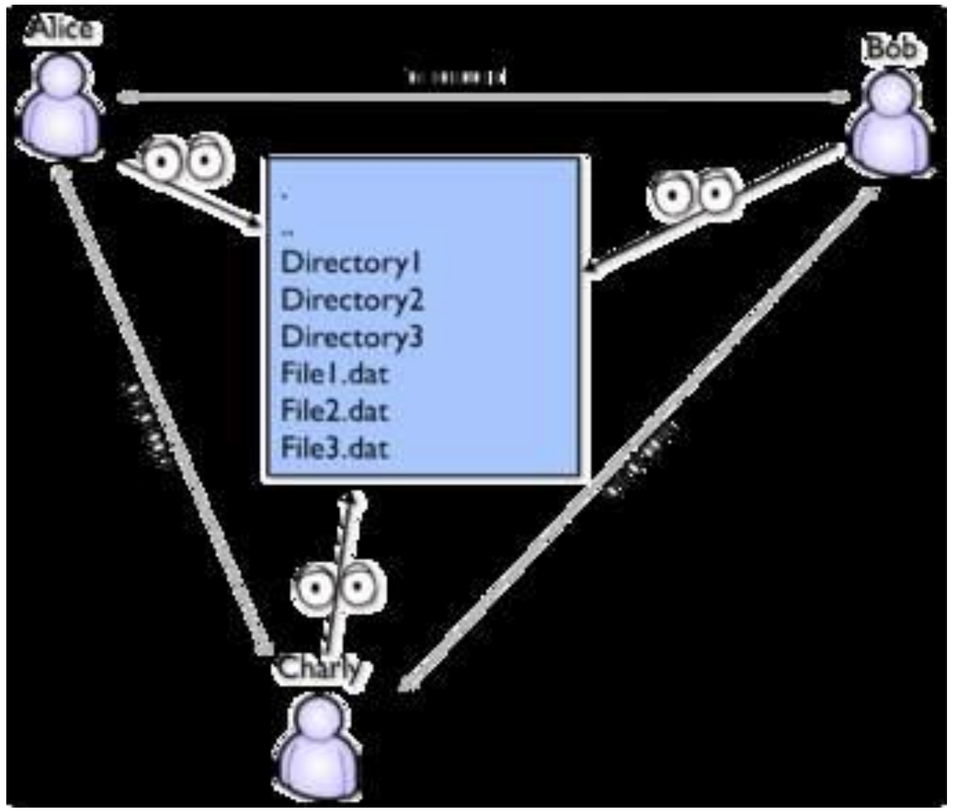
Our strengths in data linkage, health research and data analytics, come together to improve health outcomes and health service delivery in Western Australia.

Infrastructure + Analytics + Domain Expertise

Curtin University

Data Linkage in Australia

Benefits of Data Sharing



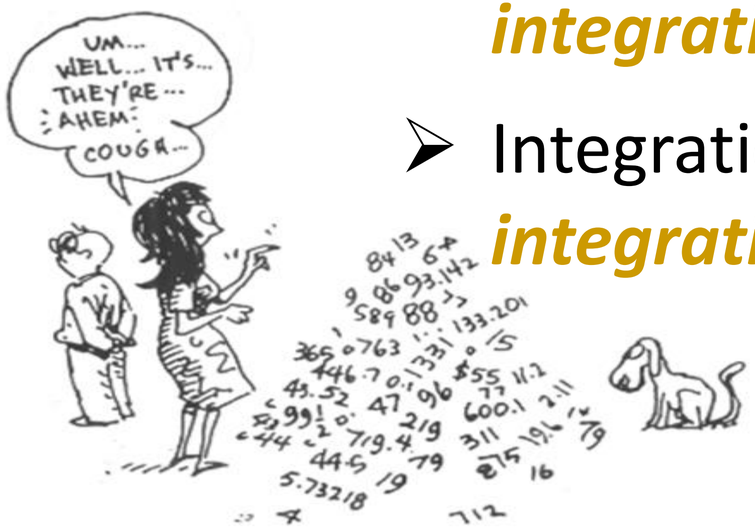
Ability to share the same data resource with multiple applications or users:

- Data and information **used/reused** to inform significant decisions
- Bring together key elements to **exploit data** for public benefit
- **Enhance the value** of information gained from a single source

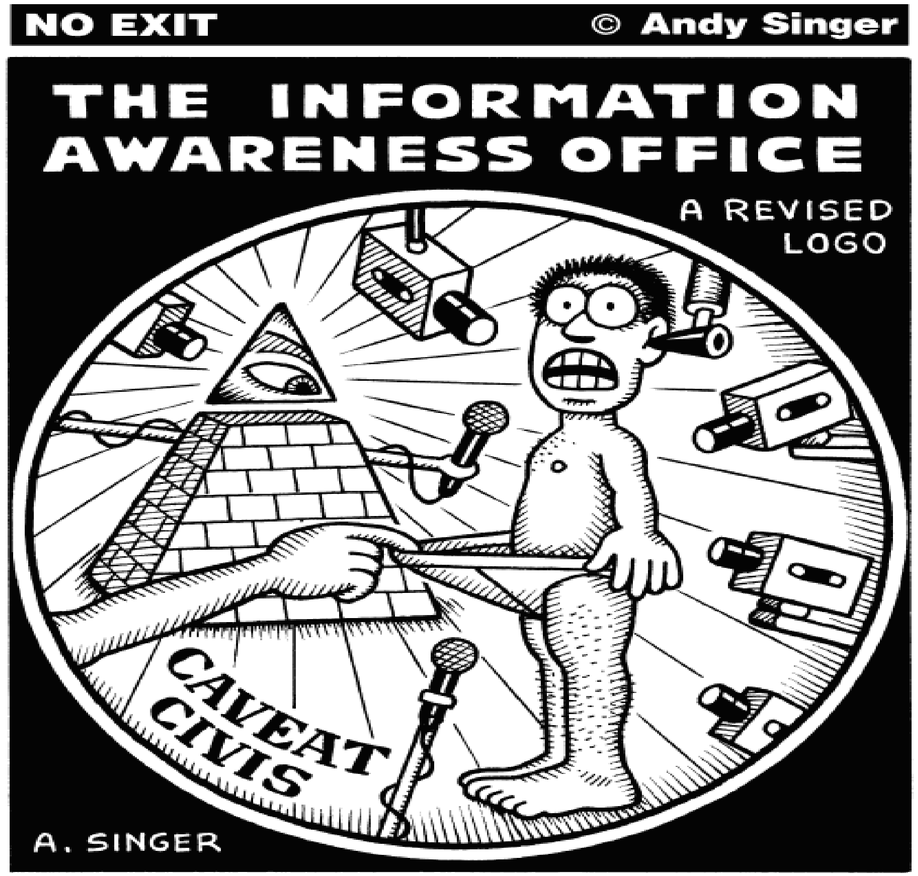
Developing Care Pathways

Developing and extending information on care pathways through ***data linkage***:

- Integration across an individual level of care (e.g. across Primary Care) (***horizontal integration***)
- Integration between primary and secondary care (***vertical integration***)
- Integration across health and social services (***inter-sectoral integration***)



Barriers to Record Linkage



- Organisational barriers
- ***Legal constraints***
  - specific datasets excluded due to legislation
- Administrative issues
  - ***privacy*** concerns
- Release of ***named identifiers*** is the issue

Curtin University

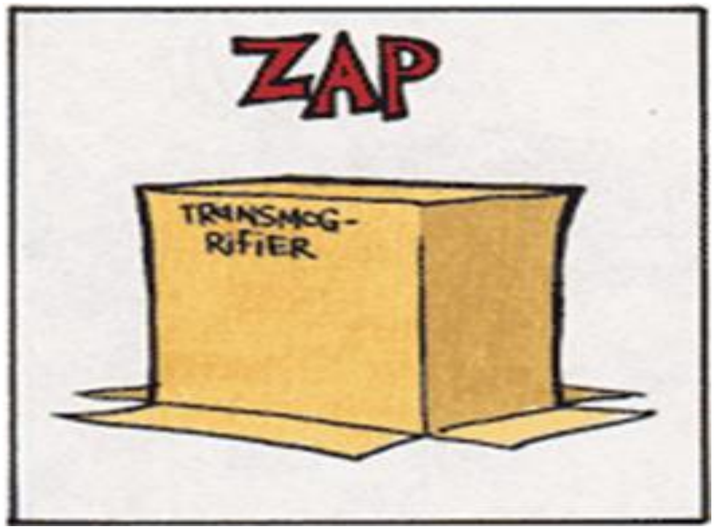
Privacy-preserving linkage

What is “privacy-preserving”?

Also known as blind-folded record linkage

Named identifiers ***are not required***

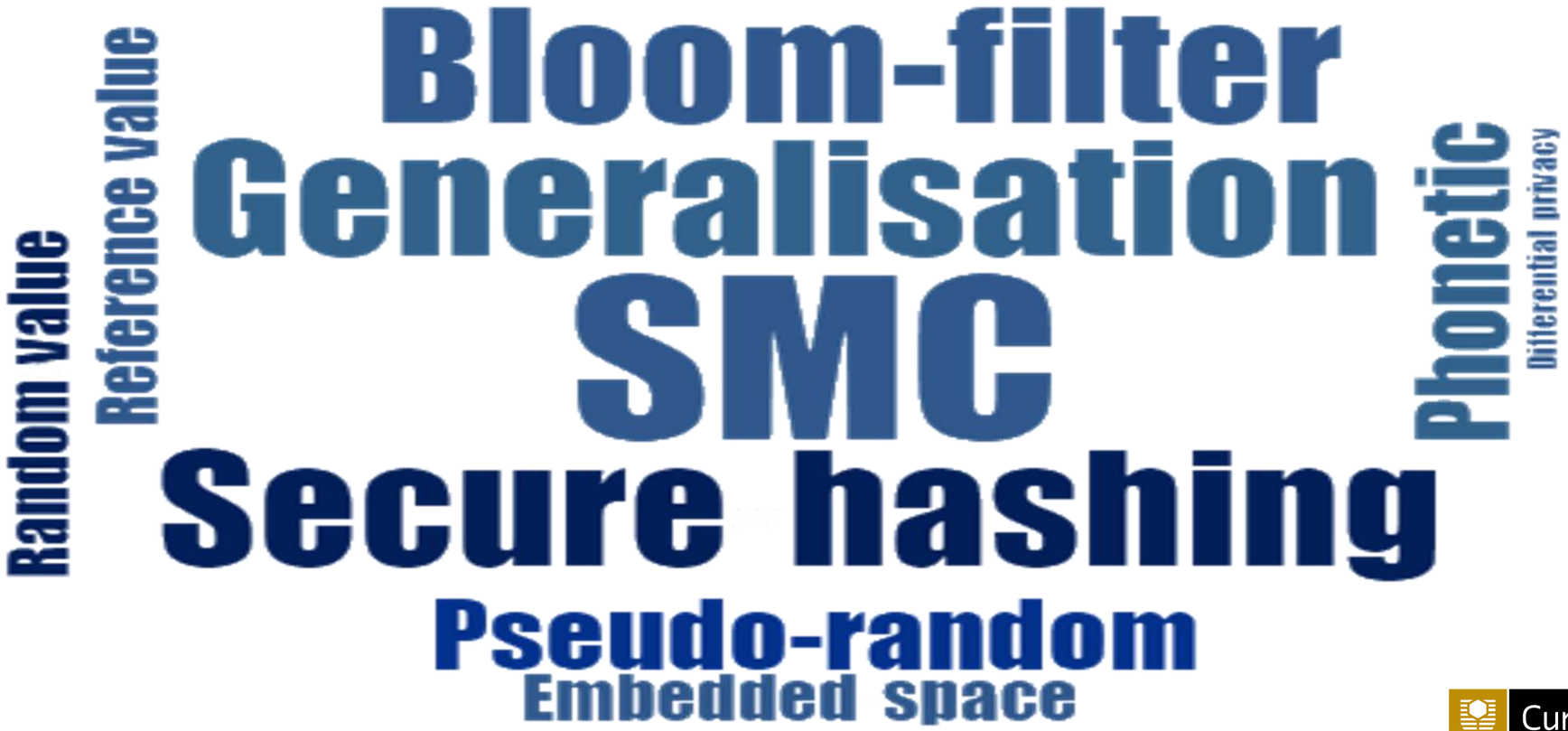
Data is transformed into an ***unidentifiable*** state at each data source



- ***Cryptographic*** techniques
- Deterministic or ***probabilistic***

Privacy Preserving Options

Methodologies for the ***integration*** of large datasets using Privacy Preserving Record Linkage:



Curtin University

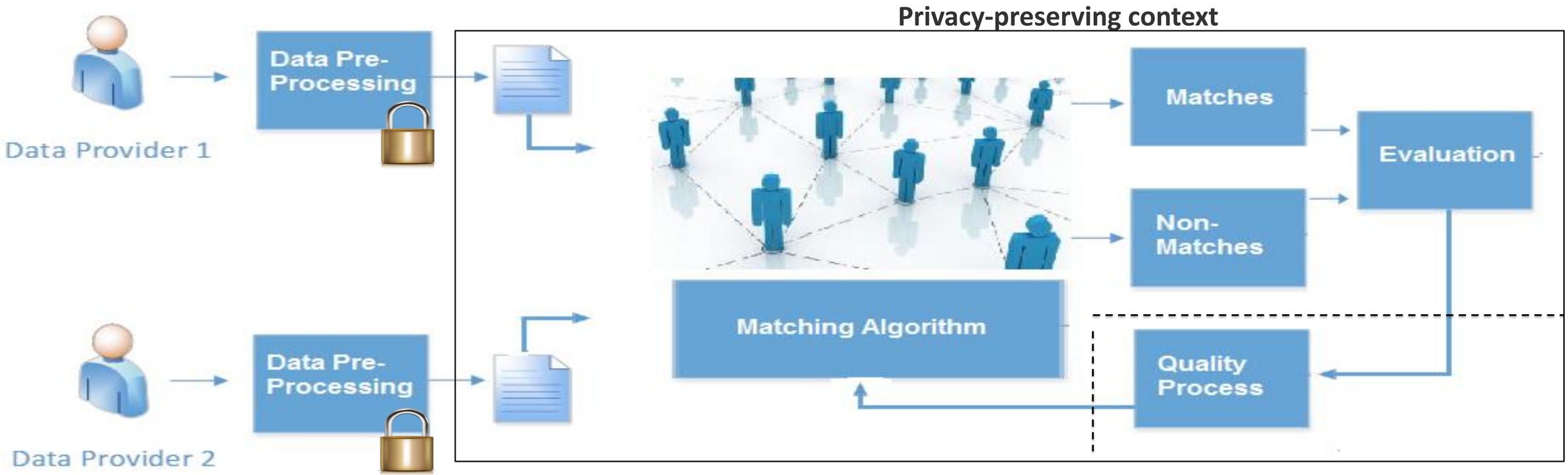
Population Health Research Network (PHRN) Strategic Project

Probabilistic Linkage Using Privacy Preserved Linkage Variables

Project aims:

- Supplement existing PHRN linkage systems
- Develop an agreed operation model
- Develop standard operating procedures
- Evaluate performance and assess linkage quality
- Recommendations on operational aspects

Privacy preserving linkage model



Privacy Preserving Process

Transforming data to a Privacy Preserved State:

- Cryptographic hash functions used to convert data into ***irreversibly de-identified*** code
- Transformation of *de-identified* code into a ***Bloom Filter*** for matching
- Bloom Filter does not contain any sensitive or personal information.

PPRL Evaluation Process

Is it possible to link individual level care data across care settings?

- Data from different parts of the health system in Western Australia and NSW (including GP data)
- ***Accuracy and efficiency*** of Privacy Preserving Record Linkage assessed
- Processing requirements and linkage quality
- internationally transferable technology, standards and applications (***interoperability***)

PPRL Evaluation Summary

Strengths

- Large datasets processed without disclosure of identifiers
- Similar linkage quality as unencrypted matching

Challenges

- Optimal linkage strategy and threshold setting
- Assessing the impact of poor quality data
- Clerical assessment and intervention difficult

Conclusion

PPRL techniques can be used to ***supplement*** traditional linkage

- Balance of ***efficiency***, ***accuracy*** and ***security*** required
- Stay tuned for more…
  - International studies

Curtin University

Next steps…

Possible Benefits

Information on the care pathway across primary and secondary care can support:

- Monitoring ***outcomes*** e.g. measuring the effects of intervention
- Monitoring ***care pathways*** e.g. activity and cost for a specific type of pathway.
- Comparative data on Practice performance - ***Benchmarking***
- Improved ***performance*** monitoring e.g. identifying referral rates across practices
- Improving ***Data Quality***