UNIVERSITY of TASMANIA

Building business intelligence into health workforce research: a graduate outcome tracking system

Gary Walker and Colleen Cheek

20 September 2018

utas.edu.au

Ad hoc project data – issues and risk

"Now, keep in mind that these numbers are only as accurate as the fictitious data, ludicrous assumptions and wishful thinking they're based upon!"

- Manual data entry and matching prone to error
- Different sources of similar data – what's the 'truth'?
- Data on user-level systems inaccessible
- Process is not explicit – reproducible?
- Time taken to re-run same queries for periodic reporting
- Reliance on individuals
- Lack of data security and governance

Strategic Information Systems Planning

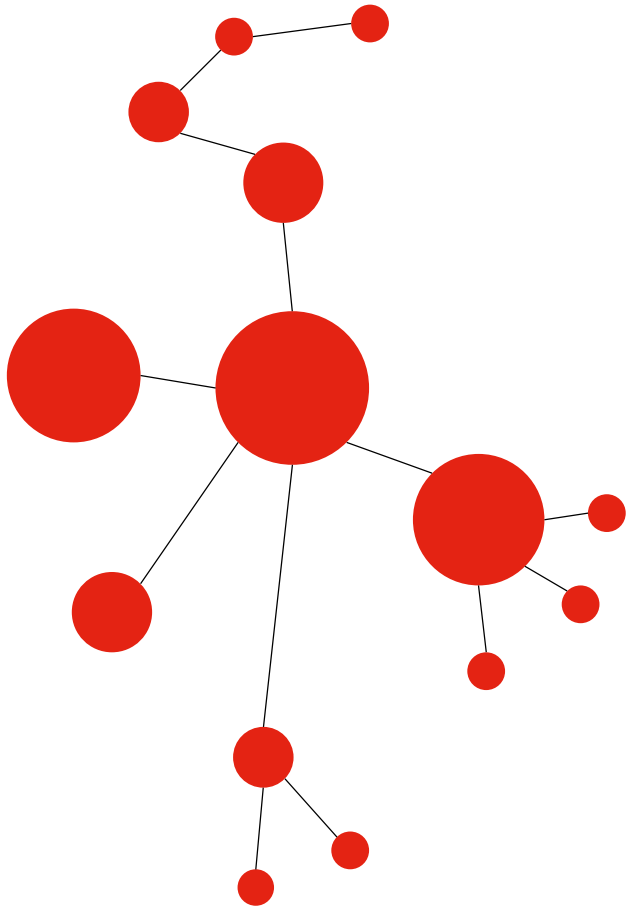
- Explicit process
- Capture tacit knowledge
- Automation – data integrity, time, reproducible
- Enterprise level system - protects organisational knowledge, sustainable, secure
- Governance – ethics, privacy and access
- Opportunities – collaboration, innovation, research, decision support,

GradTrack Establishment

- Vision & scope –
  - Initially medical students
  - But with ability to be extensible to allied health and nursing
  - Secure
  - Sustainable
  - Governance
- Data use:
    - Research
    - Reporting
    - Decision support

GradTrack – Information Gathering

- Understand what's available, what's already been done
- Tacit knowledge
  - Person dependant roles, head knowledge
  - Legacy macros - bespoke systems
  - Manage risk/fear (we are not replacing their job)
- Go beyond the initial points of contact, they don't know everything!



GradTrack – Information Gathering - Other Medical Schools

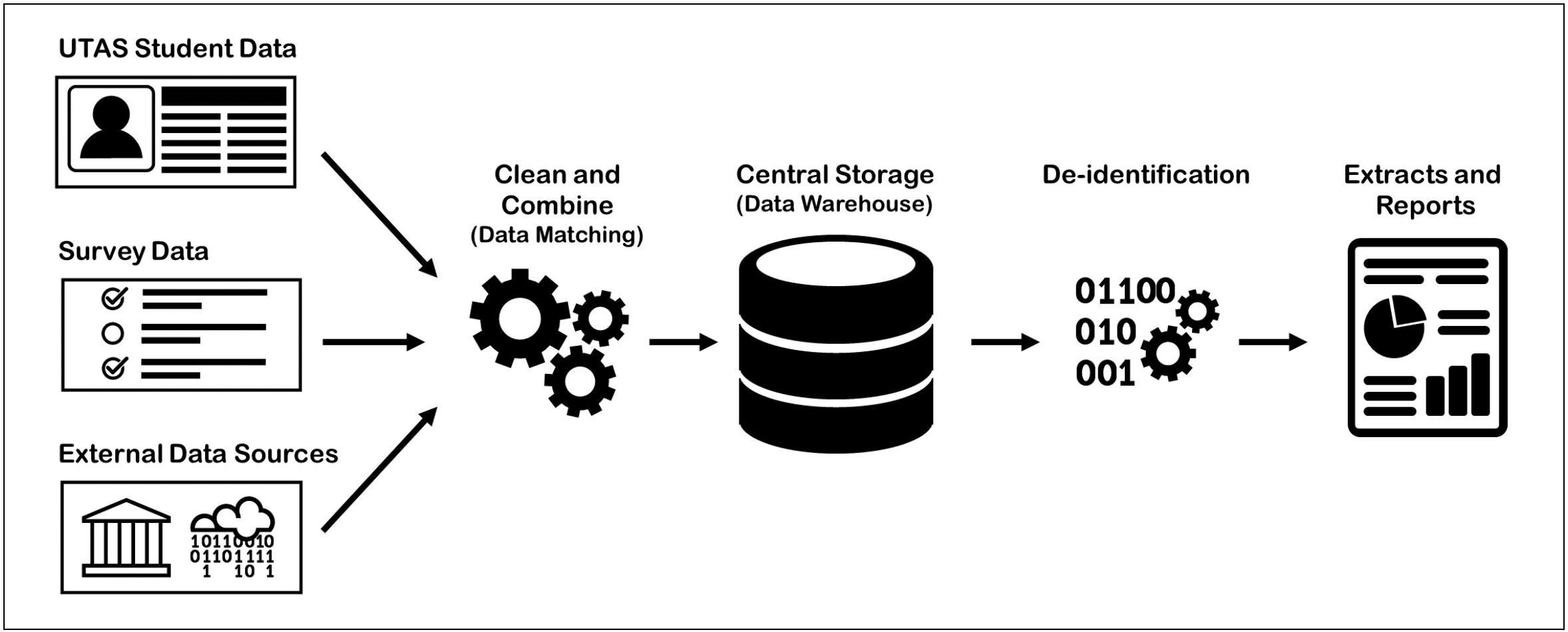
- Admin staff driven, but not resourced
- Alumni tracking off the side of desk
- Only time to get bare minimum data
- Manual and sporadic matching
  "It's quite a process I'd have to say"
- Manual cleaning of data
- Linking RA to postcode
- Stored in Excel or Access Databases

"Data that is loved tends to survive"

Kurt Bollacker, Computer scientist

There is a better way

The GradTrack Data Warehouse



What type of data is in GradTrack?

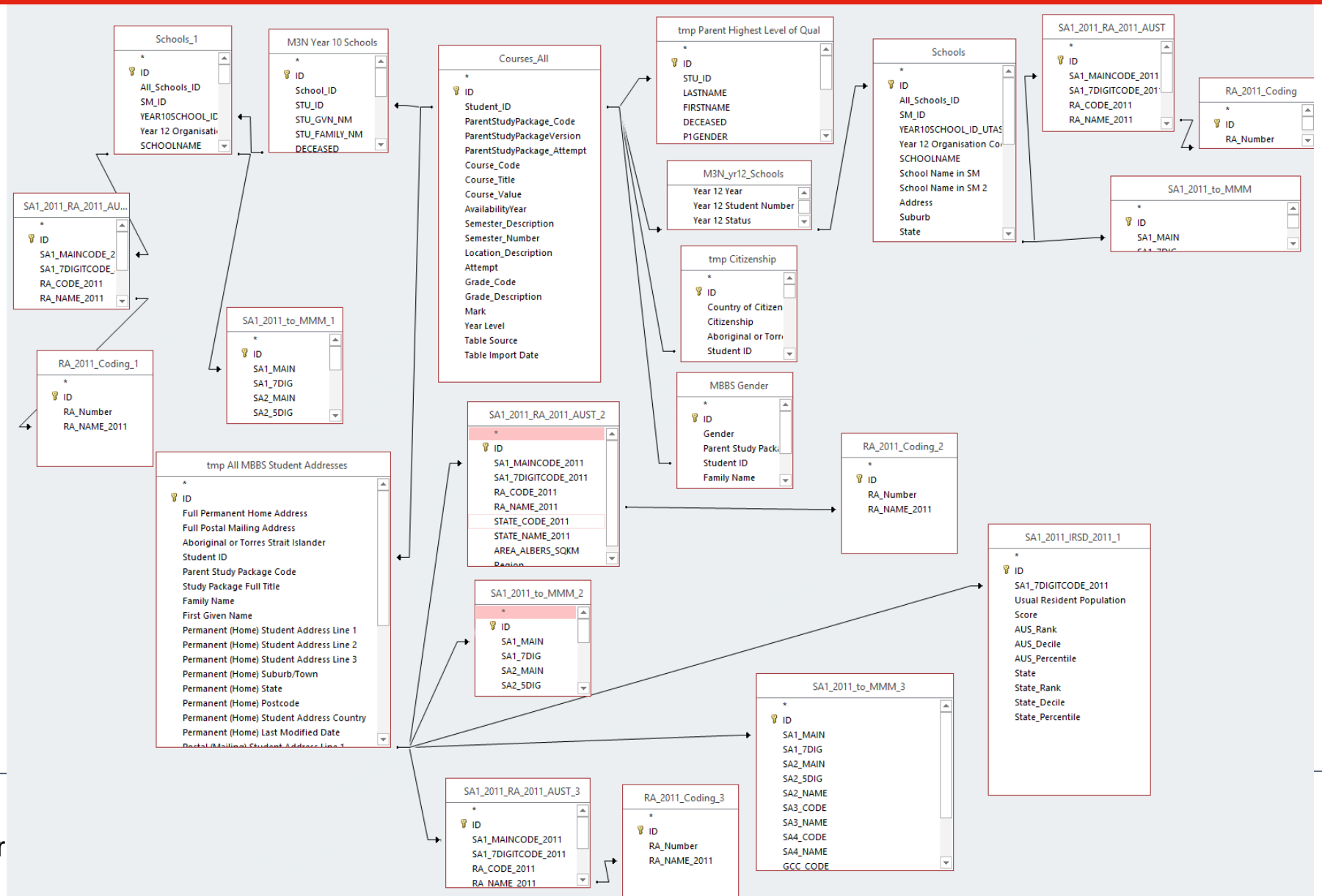
INTERNAL DATA

- Student Management
  - Demographics
  - Intake data
  - Course progression
- Student Placements (inPlace)
- Exams/OSCE results
- Student Awards
- Scholarships
- Survey Data

EXTERNAL DATA

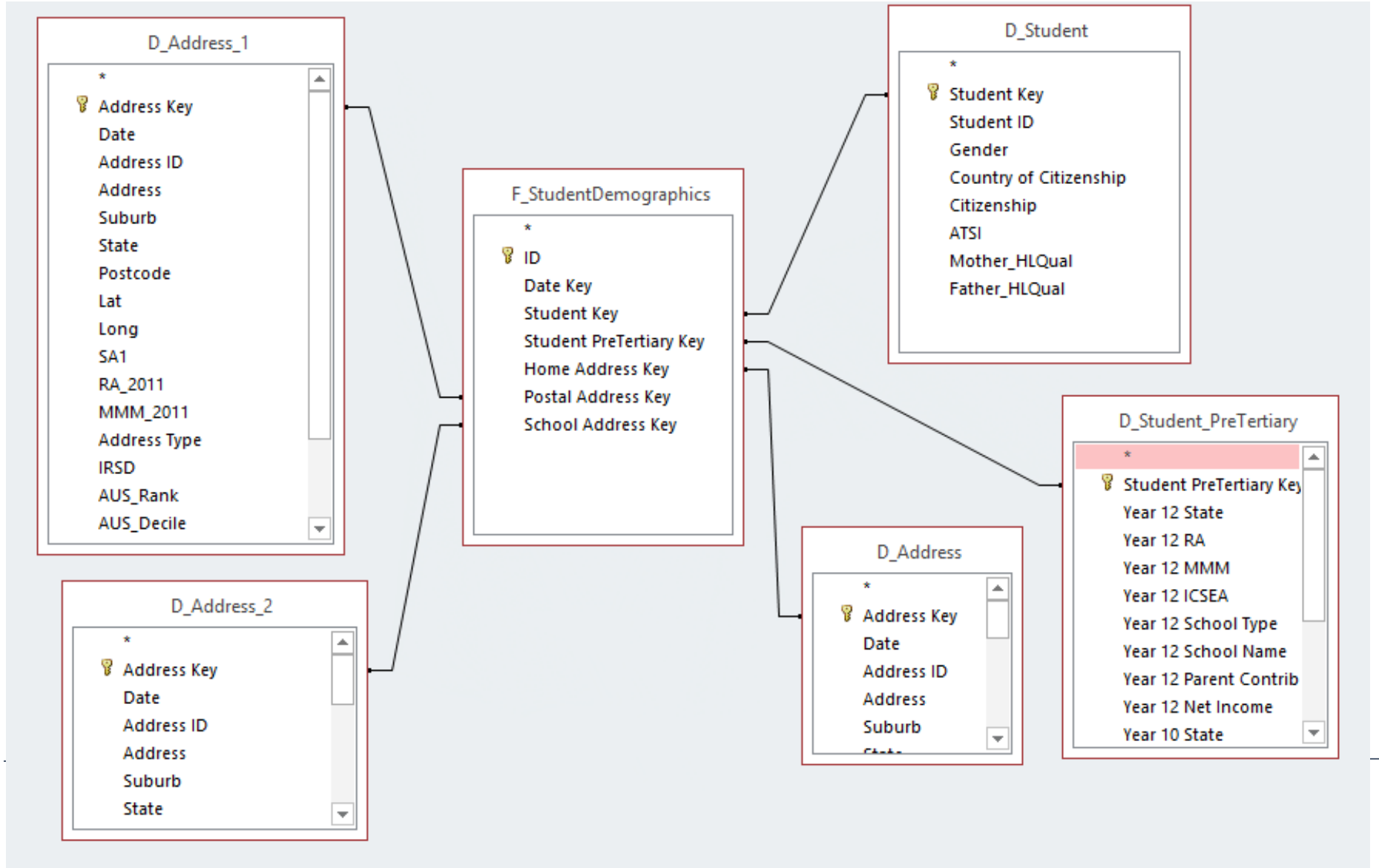
- Student/School Placement Addresses Geocoded
- Link to ABS Indices eg IRSD, SEIFA
- Doctor Connect (RA, MMM)
- AHPRA (full national extracts)
- Longitudinal surveys
  - Medical Deans
  - FRAME
- MySchool data linked to student

GradTrack – Data linkage



Before

GradTrack – Data transformation



After

UNIVERSITY of TASMANIA

Why not Excel/Access?

Example – Medical Students Outcome (MSoD) survey data

- 1 row per student, containing all survey questions from all surveys administered
- Received in CSV format
- Contained 717 columns (fields) of data
- MS Access can only handle 256 fields in a table.

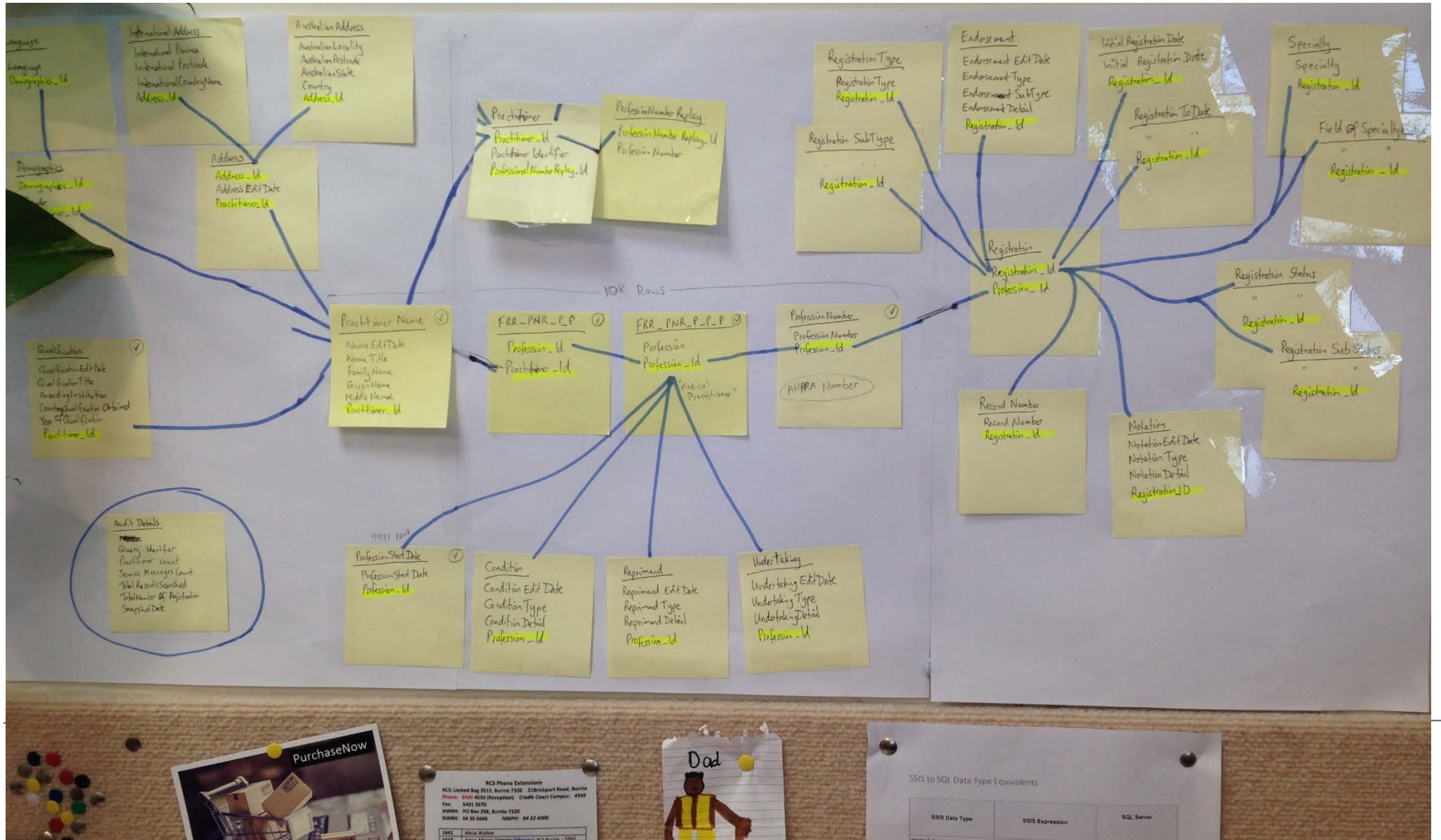
RESULT: Imported into SQL table

Why not Excel/Access?

Example - AHPRA Full National Extract

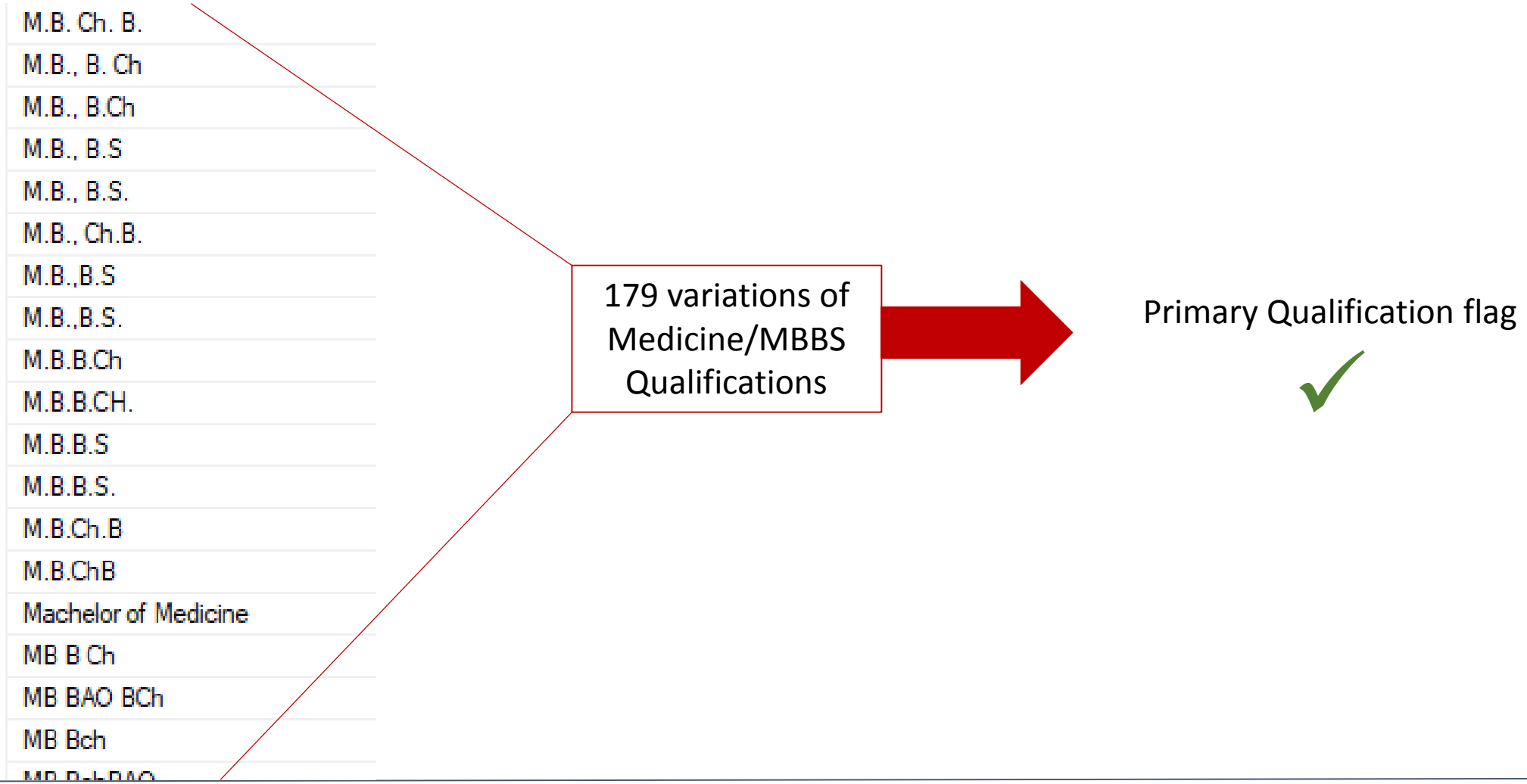
- 115000 Medical Practitioners
  - Over 2 million rows of data (per year)
- 320Mb, per year (Medical only)
- Provided in 12 separate XML files
- Nested data structure

AHPRA Full National Extract – XML map



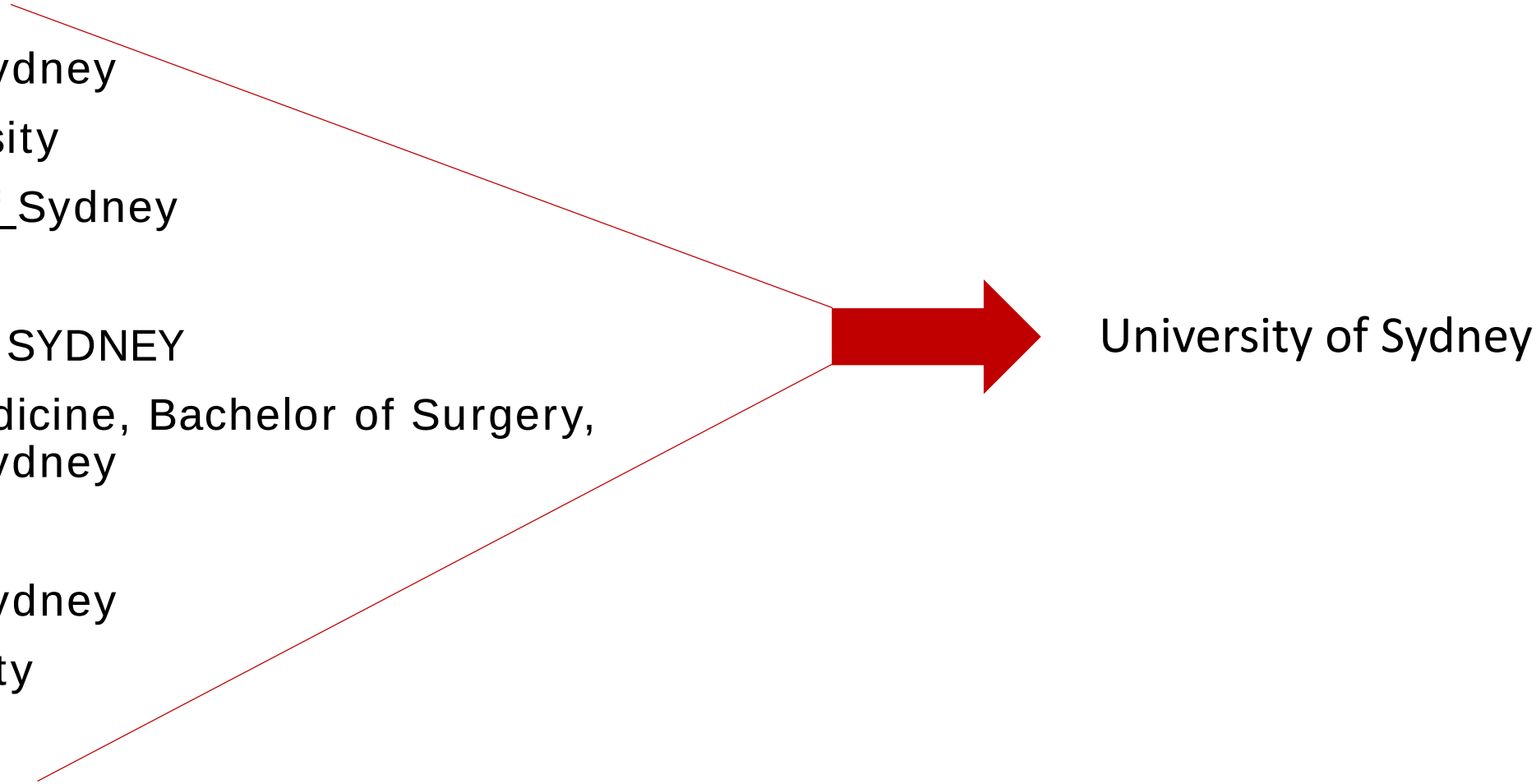
UNIVERSITY of TASMANIA

AHPRA Full National Extract – Qualification



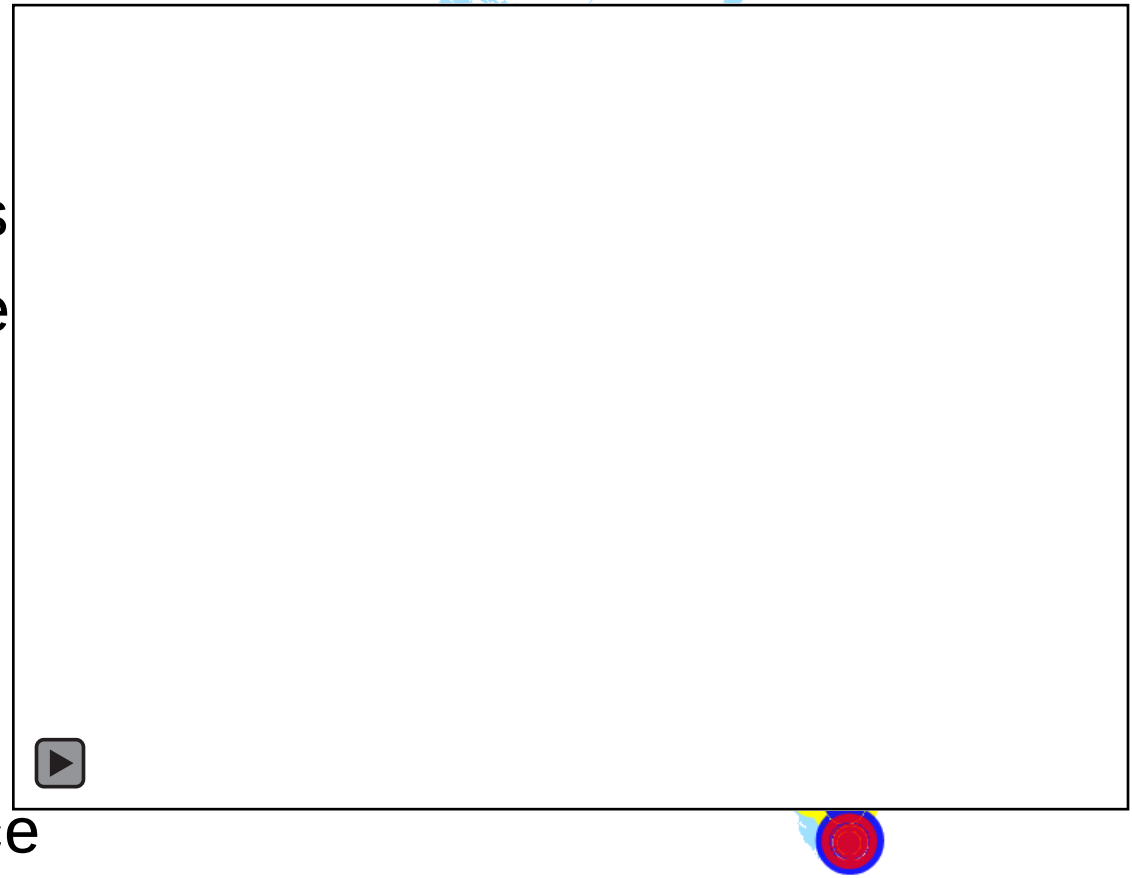
AHPRA Full National Extract – Awarding Institution

- University of Sydney
- Sydney University
- University___of Sydney
- MB BS SYDNEY
- MB BS (HONS) SYDNEY
- Bachelor of Medicine, Bachelor of Surgery, University of Sydney
- MD Sydney
- Universtiy of Sydney
- Syndy University
- etc



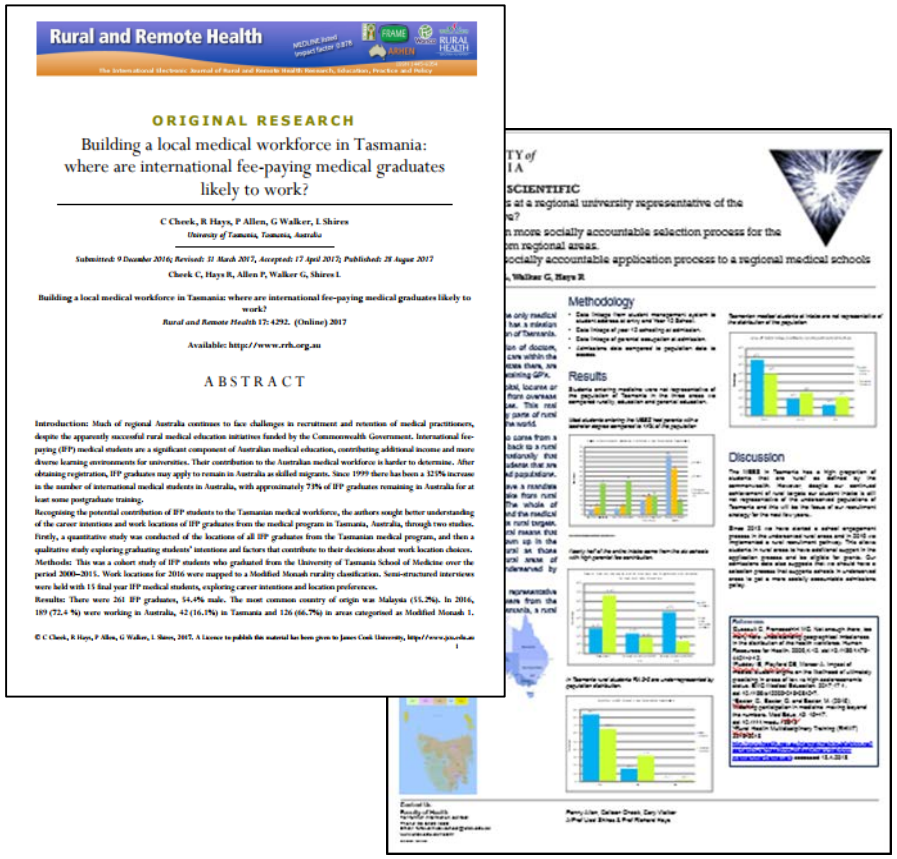
GradTrack - Other features

- Granularity of data
  - Addresses geocoded to SA1 level
  - Accurate MMM and RA coding
  - Mapping of graduate work locations
- QA - provided clean data back to source
- AHPRA match rate: from 68% to 95%
- Used existing IT infrastructure
  - No hardware/software costs
  - Centrally located
  - Centralised security policies
  - Backup, replication and DRP in place
  - The IT people know it exists



Results

- Data from multiple sources (937 fields)
- Automated collection, cleaning and linking
- Data protection
  - Governance, 5 safes
  - Encryption of outputs
  - Ethical approval, consents
- Contribute to research
- Faster reporting
- Foundational work done
- Scale to All health disciplines
- Analytics and decision support



UNIVERSITY of TASMANIA

What we've learned

- GradTrack uses standard Data Warehouse principles
- Compelling need
- Strong sponsor
- Build rapport
- Organisational readiness for a DW
- Dedicated personnel
  - With the right mix of skills
- Defined processes quality data

- Still need to frame the right questions

"The idea is to go from numbers to information to understanding"

Hans Rosling

Opportunities and challenges

Challenge

- Student consent to tracking means we may never get 100% participation
  - onus on us to champion and demonstrate value and integrity

Opportunity

- Collaboration with other disciplines eg investigate use of statistical techniques to contrast the usefulness of different modelling approaches

Cheek C, Walker G, Hays R, Allen P, Shires L. **"Building business intelligence into health workforce research – a graduate outcome tracking system"**, *Under review.*