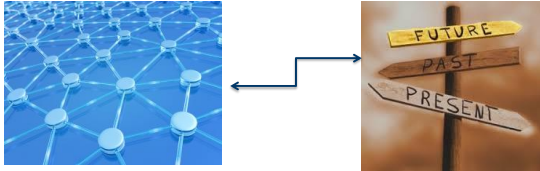


Data linkages



Presentation prepared by: Alison Anlezark
 Title: National Standards Manager, NCVER
 Event: Australian Government Statistical Forum (AGSF)
 Date: 8 October 2015

NCVER administrative collections and surveys

- Apprentice and trainee collection (transactional, collected quarterly, cumulative), in-training n~315,000
- National VET provider collection (annual/quarterly (public), snapshot of training activity, n~1.8M (publicly funded VET only)
- National VET in schools collection (annual), snapshot of training activity n~250,000
- Student Outcomes Survey (annual survey of VET graduates) n~35,000 to n~120,000
- Longitudinal Surveys of Australian Youth (LSAY): 15-25 year olds, 6 cohorts (Y95, Y98, Y03, Y06, Y09, Y15) n~14,000 at wave 1 to n~4,000 at wave 11

Presentation outline

1. Past data linkage (data quality check)
 - Apprentice & trainee collection – VET provider collection
2. Present data linkage
 - VET in schools collection : Census
 - Student Outcomes Survey : Student Journey Dataset
 - Longitudinal Surveys of Australian Youth (LSAY) : National Assessment Program – Literacy and Numeracy (NAPLAN)
3. Future data linkage
 - Unique student identifier (USI) – National VET administrative collections
4. Challenges

Linkage example #1: Apprentice & trainee Collection – VET Provider Collection

Apprentice & Trainee Collection (transactional)
 ~0.315m in-training apprentice and trainees per year

VET Provider Collection (snapshot)
 ~1.8m students per year

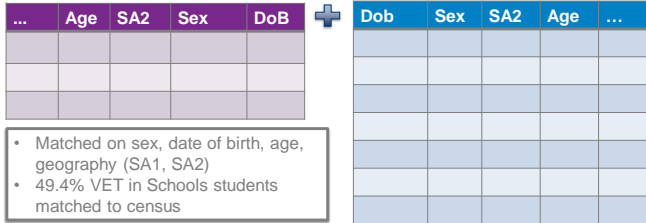
....	Client ID	COT #	+	COT #	Client ID

- Matched on Client ID and Contract of Training (COT) ID
- ~80% of apprentice and trainees in VET Provider Collection matched to the A&T Collection in a given year

Linkage example #2: 2011 VET in Schools – 2011 Census

2011 VET in Schools (snapshot)
n = 236,461 persons (15-20 yrs)

2011 Census (point in time) 14-21 yrs
n= 2.2 million persons



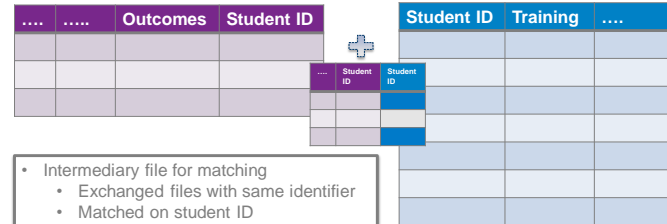
- Matched on sex, date of birth, age, geography (SA1, SA2)
- 49.4% VET in Schools students matched to census

*source: ABS, 2013, 1351.0.55.044 - Research Paper: Assessing the Feasibility of Linking 2011 Vocational Education and Training in Schools Data to 2011 Census Data

Linkage example #3: Student Outcomes Survey – QLD Student Journey Data

Student Outcomes Survey (SOS)
2006 – 2014 QLD only
n~110,000 students

QLD Student Journey Dataset (SJD)
1996 - 2013
n~1.6M subject enrolments

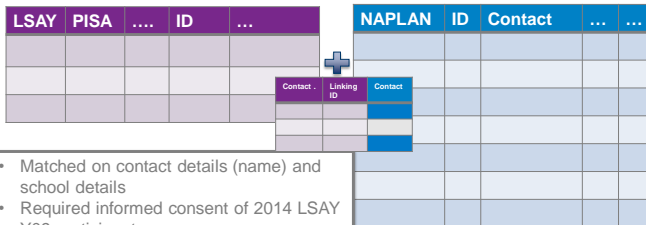


- Intermediary file for matching
 - Exchanged files with same identifier
 - Matched on student ID
- Appended SJD to SOS QLD extract
- 90% eligible SOS data matched SJD

Linkage example #4: LSAY - NAPLAN

LSAY Y09 cohort in 2014
n=5,082

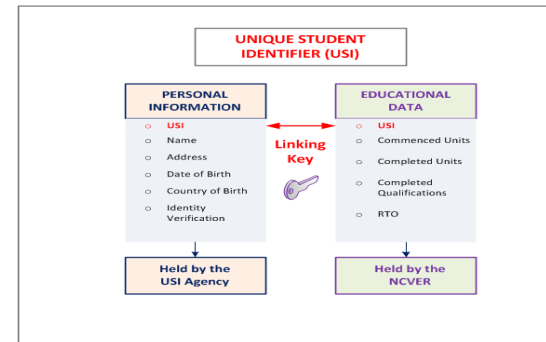
NAPLAN Y9 in 2008
n= 276,613 *



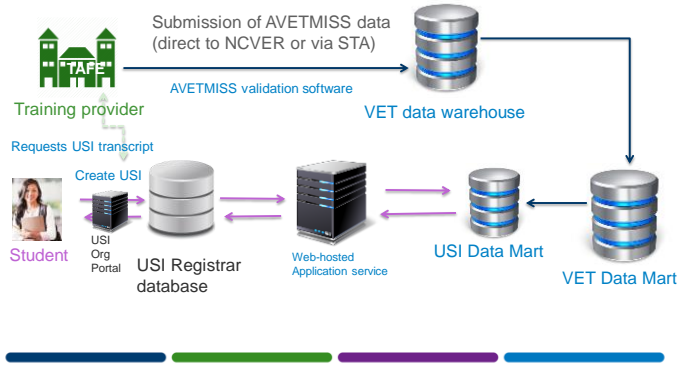
- Matched on contact details (name) and school details
- Required informed consent of 2014 LSAY Y09 participants
- 98.2% who consented matched

* Source : ABS, 42210DO015_2008 Schools, Australia, 2008 (Reissue)

Linkage example #5: USI - National VET administrative collections



USI transcript service: Conceptual data flow model



USI transcript

Australian Government			
Student Identifiers Registrar (name/logo)			
Authenticated Vocational Education and Training Transcript			
The records held in the National VET Provider Collector from 1 January 2015 to 20 December 2016 show that			
Adam John SMITH			
Completed the requirements for the following nationally accredited or recognised VET qualification(s):			
2011			
Certificate IV in Information Technology (ICA40305)			
From the CHISHOLM INSTITUTE (0260)			
2013			
Certificate III in Children's Services (CHC30708)			
From the CHISHOLM INSTITUTE (0260)			
Undertook the following nationally accredited VET units and modules:			
Provider: CHISHOLM INSTITUTE (0260)	Unit/Module (Identifier) - (PT)	Outcome	Enrolment Period
	HANDLE MAIL (BSBNM202A) - (12)	Competent	3/2/11-23/12/11
	PROCESS AND MAINTAIN WORKPLACE INFORMATION (BSBNM203A) - (11)	Not yet comp.	23/2/11-10/1/13
	PRODUCE SIMPLE WORD PROCESSED DOCUMENTS (BSBITU201A) - (13)	RPL	9/3/11-1/10/11
Provider: NSW TAFE (1111)	Unit/Module (Identifier) - (PT)	Outcome	Enrolment Period
	PROCESS CUSTOMER COMPLAINTS (BSBMM3001A) - (13)	Withdrawn	24/3/11-30/9/11
	DELIVER A SERVICE TO CUSTOMERS (BSBQJ201A) - (20)	Credit transfer	5/5/11-2/6/11
	WORK EFFECTIVELY IN A BUSINESS ENVIRONMENT (BSBMD201A) - (15)	RPLN	5/5/11-2/7/11
	BASIC COMPUTER SKILLS (BSBICM112) - (15)	NA-SC	1/1/10-2/1/10

USI transcript: Draft screen design

Undertook the following nationally accredited VET unit(s) and module(s):			
Provider: NSW TAFE (1111)	Unit/Module (Identifier) - (PT)	Outcome	Enrolment Period
	<input type="checkbox"/> DELIVER A SERVICE TO CUSTOMERS (BSBQJ201A) - (2)	A	05/05/2011 - 03/08/2014
Provider: CHISHOLM INSTITUTE (0260)	Unit/Module (Identifier) - (PT)	Outcome	Enrolment Period
	<input type="checkbox"/> HANDLE MAIL (BSBNM202A) - (12)	A	03/02/2011 - 23/12/2011
	<input type="checkbox"/> PROCESS AND MAINTAIN WORKPLACE INFORMATION (BSBNM203A) - (11)	F	23/02/2011 - 10/1/2011
	<input type="checkbox"/> PRODUCE SIMPLE WORD PROCESSED DOCUMENTS (BSBITU201A) - (13)	RPL	06/03/2011 - 01/10/2011
Provider: CHISHOLM INSTITUTE (0260)	Unit/Module (Identifier) - (PT)	Outcome	Commenced
	<input type="checkbox"/> HANDLE MAIL LEVEL 2 (BSBNM202A) - (12)	Cont	03/01/2015
Provider: ROYAL LIFE SAVING (0075)	Unit/Module (Identifier) - (PT)	Outcome	Commenced
	<input type="checkbox"/> FIRST AID LEVEL 2 (087574) - (12)	NVE	23/04/2014
	<input type="checkbox"/> FIRST AID LEVEL 1 (087573) - (12)	Cont	23/02/2014
Provider: NSW TAFE (1111)	Unit/Module (Identifier) - (PT)	Outcome	Commenced
	<input type="checkbox"/>		

Potential of the USI

- Better able to understand VET pathways
- Better able to count students (reduces duplications)
- Understand and measure completion rates
- Manage entitlement funding models
- Manage student training
- Expandable to other education sectors



INFORMING POLICY AND PRACTICE
IN AUSTRALIA'S TRAINING SYSTEM

Data linkage challenges

- Cooperation, knowledge sharing & goodwill between agencies
- Privacy, informed consent
- Approvals (ethics, data custodians, etc.)
- Data access
- Data security and longevity
- Timing, reporting lags
- Limited opportunity for re-use



INFORMING POLICY AND PRACTICE
IN AUSTRALIA'S TRAINING SYSTEM

Data linkage challenges (Continued)

- Resources, costs
- Correction or updates to linked datasets
- Data quality (missing values)
- Suitability of matching variables
- One: many, many: many, many : one
- Technical skills
- Appropriate software



INFORMING POLICY AND PRACTICE
IN AUSTRALIA'S TRAINING SYSTEM

Questions?

Stay connected

Subscribe

at www.ncver.edu.au

Follow



twitter.com/ncver or [@ncver](https://twitter.com/ncver)



[linkedin.com/company/ncver](https://www.linkedin.com/company/ncver)