Key issues in reusing data

Patty Doran
UK Data Service
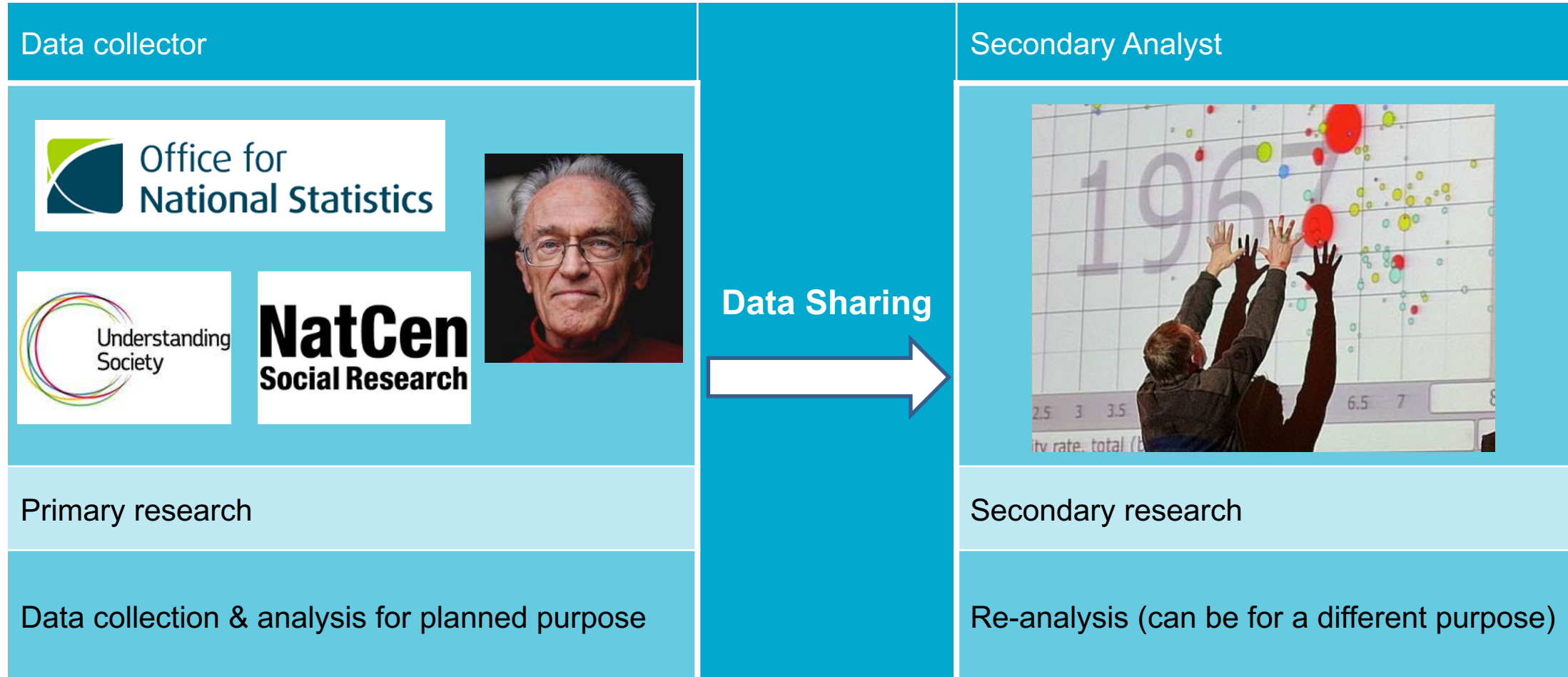patty.doran@manchester.ac.uk
@Patty_NZ

Overview

- Data available for reuse
- Issues with reusing data
- Quantitative
  - Case study
  - What's available
  - Key issues
- Qualitative
  - Case study
  - What's available
  - Key issues
- Talk for ~35 min
- Questions

type questions as they occur to you

UK Data Service

Reusing data

Data collector		Secondary Analyst
Office for National Statistics		
Understanding Society NatCen Social Research	**Data Sharing** →	
Primary research		Secondary research
Data collection & analysis for planned purpose		Re-analysis (can be for a different purpose)

UK Data Service

Pros and cons of reusing data

Pros
datasets impossible to create
cost effective
ethical issues about data collection dealt with
do not need to re-contact data subjects
reuse data used by others to make claims

| Cons |
|---|
| do not have insider understanding of data and data collection |
| effort to get to know the data |
| ethical issues about data use still apply (limit access to sensitive data) |
| data may not match research question |
| cannot extend studies |

Researchers reusing data need to:

1. Make effort to understand the data
2. Be pragmatic about whether data are good enough for your purpose

Research plan
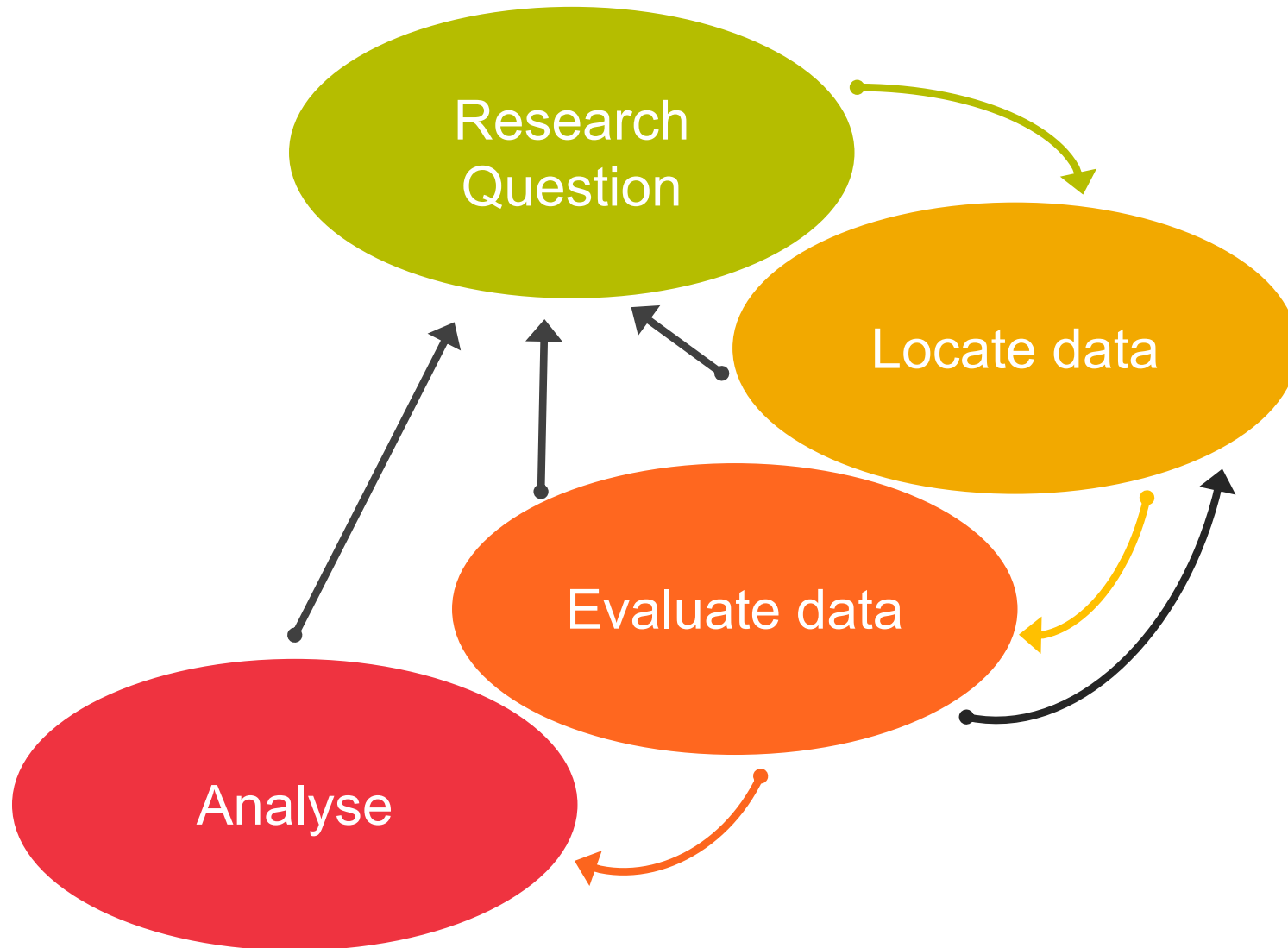
Making sense of your data

To understand data and it's suitability for your research you must understand…

- what information was collected
- from whom
- when and where
- any changes to the raw data before being archived

Use documentation
- user guides
- questionnaires
- interview schedules

UK Data Service

Quantitative data

- Case study
- What data is available?
- What sorts of documentation are available?
- Key issues

UK Data Service

Case Study: Crime Survey for England and Wales (formerly the British Crime Survey)

- Example of a repeated cross-sectional survey
- Important source of information about crime (Crime statistics independent from police records)
- Annual survey (interviews throughout the year)
- n≈35,000 individuals aged 16+ and 3,000 aged 10-15
- Victim of a crime in the previous 12 months
- Questions covering demographics and crime-related subjects such as attitudes to the police and the criminal justice system (CJS)

UK Data Service

Case Study: Crime Survey for England and Wales

Data stored as individual anonymised records

	rowlabel	split	sex	yrsarea	work2	tenure1	livharm1	agegrp7	ethgrp2a	educat3
1	137068050	A (Experiences o...	Female	20 years o...	Yes	Buying it ...	Single	45–54	White	Degree or ...
2	147461190	C (Crime preven...	Female	10 years b...	No	Own it out...	Married/C...	55–64	White	Degree or ...
3	137116250	A (Experiences o...	Female	20 years o...	No	Rent it	Widowed	55–64	White	Degree or ...
4	147354190	C (Crime preven...	Female	20 years o...	Yes	Buying it ...	Married/C...	55–64	White	O level/G...
5	137061230	C (Crime preven...	Female	20 years o...	No	Rent it	Widowed	65–74	White	None
6	136898230	C (Crime preven...	Female	20 years o...	No	Own it out...	Married/C...	65–74	White	O level/G...
7	135507330	A (Experiences o...	Male	10 years b...	Yes	Rent it	Married/C...	45–54	White	None
8	136450220	B (Attitudes to t...	Female	5 years bu...	Yes	Own it out...	Married/C...	55–64	White	Degree or ...
9	136111200	D (Online Security)	Male	20 years o...	No	Own it out...	Married/C...	55–64	White	Degree or ...
10	136599250	A (Experiences o...	Male	20 years o...	No	Own it out...	Married/C...	75+	White	Apprentic...
11	136229130	A (Experiences o...	Male	3 years bu...	Yes	Buying it ...	Single	25–34	White	Degree or ...
12	136947260	B (Attitudes to t...	Male	5 years bu...	No	Own it out...	Married/C...	75+	White	Apprentic...

Reference: University of Manchester, Cathie Marsh Institute for Social Research (CMIST), UK Data Service, Office for National Statistics. (2019). *Crime Survey for England and Wales, 2013-2014: Unrestricted Access Teaching Dataset*. [data collection]. Office for National Statistics, [original data producer(s)]. Office for National Statistics. SN: 8011, http://doi.org/10.5255/UKDA-SN-8011-1.

Different levels of access

- A standard End User Licence (download after registering)
- Secure Access (for sensitive information)

Case Study: Crime Survey for England and Wales

- Hind Khalifeh et al (2013) examined violence against people with disability in England & Wales

British Crime Survey 2009/10
introduction of disability measures
(a special licence version of the data)
n = 46,398 adults 16+
9,037 had at least one limiting disability

Findings :

Adjusting for age, sex, socio-economic characteristics…

- disability increases risk of experiencing violence
- levels of victimisation highest amongst those with mental health problems
- estimated 116,000 victims of violence attributable to disability

http://www.plosone.org/article/info:doi/10.1371/journal.pone.0055952

UK Data Service

Quantitative data by type of analysis

Type of analysis	Type of data
Individuals, families, households, businesses : one point in time	Cross-sectional surveys (e.g. Time Use Survey) and census microdata
Individuals, families, households, businesses : multiple points in time	Repeated cross-sections (e.g. different years of Crime Survey for England and Wales, Labour Force Survey)
Individuals, families, households: over time	Longitudinal data (e.g. Understanding Society)
Small geographic areas	Census aggregate data, flow data
Comparing countries (over time)	International time series (e.g. World Bank Indicators)

UK Data Service

Where, when and why?

- Start with catalogue record

Access to documentation including user guide or technical report

Date & Geographic coverage is in the catalogue record for most studies

Who was asked what…

FLEX10   **UK**
Some people have special working hours arrangements that vary daily or weekly.   **W1/EQ**
In your (main) job is your agreed working arrangement any of the following…   **MAIN**
Code up to 3   **AW/AJ and OD**
     RUNNING PROMPT

1. flexitime (flexible working hours),
2. an annualised hours contract,
3. term-time working,
4. job sharing,
5. a nine-day fortnight,
6. a four-and-a-half day week,
7. zero hours contract,
8. on-call working, or
9. none of these?

ONS ✔   **GOV** ✔   **SPL** ✔   **EUL** ✔

Applies if in work during reference week
IF (Wrking=Yes) OR (JbAway=Yes) OR (OwnBus=Yes) OR (RelBus=Yes)
OR (TYPSCH12=1,2,3,5.8) On a government training scheme considered as employment
OR ((TYPSCH12=9) AND (YTETJB=Yes)) OR (YTETJB=Yes) Doing paid work in addition to this particular
government training scheme

- Computer Aided Interviewing (CAI) makes it easy to send respondents through the questionnaire by different routes
- Many questions may only be applicable to some of the sample

UK Data Service

And what was done with data afterwards?

FLEXW7 - Whether respondent works zero hours contract

START → ALL FLEX10 (1-3) = -9 —NO→ ANY FLEX10 (1-3) = 7 —NO→ FLEXW7 = 2 DOES NOT WORK A ZERO HOURS CONTRACT

ALL FLEX10 (1-3) = -9 —YES→ FLEXW7 = -9

ANY FLEX10 (1-3) = 7 —YES→ FLEXW7 = 1 WORKS A ZERO HOURS CONTRACT

Uses:

FLEX10
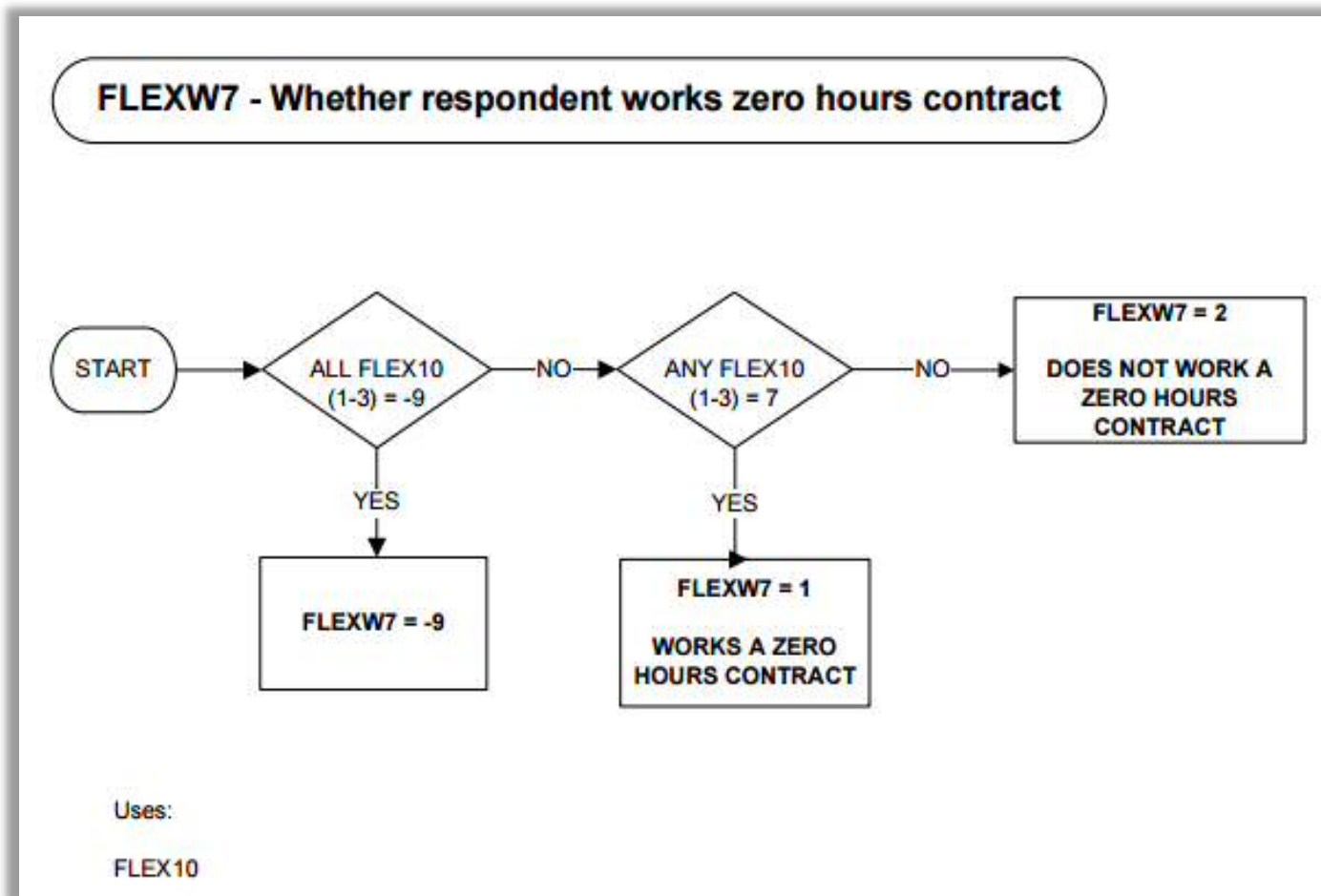
- derived variables are created from the 'raw' data

UK Data Service

Sampling considerations

- Surveys are based on samples
- Is the sample representative?
  - Who is included (adults, those in private address)?
  - Response rate and bias
  - Do you need to apply a weight to make the data representative?
- Do I have enough cases to make a precise estimate?
  - Important for small sub-populations

For more information, see UK Data Service guides to Survey weights and Complex sampling

UK Data Service

Cite the Data

- Citations provided for for every record and dataset
- Use the UK Data Service **citation tool** to copy and paste the correct citation

MAKE THE DATA COUNT
CITE THE DATA

British Social Attitudes Survey, 2013

| Details | Documentation | Resources | Access data |

Details ⌄

Title:	British Social Attitudes Survey, 2013
Alternative title:	BSA
Study number (SN):	7500
Access:	These data are safeguarded
Persistent identifier (DOI):	10.5255/UKDA-SN-7500-1
Series:	British Social Attitudes Survey
Principal investigator(s):	NatCen Social Research

Sponsors and contributors ⌄

Citation and copyright ⌄

The citation for this study is:

NatCen Social Research. (2014). *British Social Attitudes Survey, 2013*. [data collection]. UK Data Service. SN: 7500, http://doi.org/10.5255/UKDA-SN-7500-1

UK Data Service

References and resources

General discussions:

- Dale, A., Wathan, J. and Higgins, V. (2008) 'Secondary Analysis of Quantitative Data Sources' in The Sage Handbook of Social Research Methods eds: P. Alasuutaryi, L. Bickman and J. Brannen Sage:London

- Smith, E. (2006) Using Secondary Data in Educational and Social Research OUP: Berks

UK Data Service resources:

- Data Skills Modules: https://www.ukdataservice.ac.uk/use-data/data-skills-modules

- Secondary analysis web pages: https://www.ukdataservice.ac.uk/use-data/secondary-analysis

- Using Survey Data: https://www.ukdataservice.ac.uk/media/455510/using-survey-data.pdf

- What is weighting?: http://ukdataservice.ac.uk/media/285227/weighting.pdf

- What is complex sample design: http://ukdataservice.ac.uk/media/398743/complexsampleguide.pdf

Reusing qualitative data

- Types of qualitative data reuse

- Case study of reusing UK Data Service data: School Leavers Study

- Getting started with reusing qualitative data
  - Key issue: documentation and sampling
  - Key issue: searching

- Finding data – browse, search, QualiBank

UK Data Service

Many ways of re-using data

- Description
  - Literature review with data
  - See more data, not limited to published extracts
- Research design and methodological advancement
  - Study sampling, data collection methods, topics guides
  - Compare original interview schedule questions actually asked in the interview
- Reanalysis –a research question different to the original research
  - Very different – comparative keyword analysis of illness narratives Seale and Charteris-Black (2008)
  - Similar topics, but with different focus (Bornat et al. 2012)
- Learning and Teaching (Haynes and Jones 2012)

UK Data Service

SN 4867: **School Leavers Study**

Principal investigator: Ray Pahl

In 1978, teachers at a comprehensive school on the Isle of Sheppey were asked to set an essay about 10 days before pupils were due to leave school. The essay asked students to imagine that they were nearing the end of their life, and that something had made them think back to the time when they left school. They were then asked to write an imaginary account of their life over the next 30 or 40 years.

UK Data Service

School Leavers Re-Study: Imagining the Future

Re-users: Graham Crow and Dawn Lyon, University of Kent and University of Southampton

UK Data Service
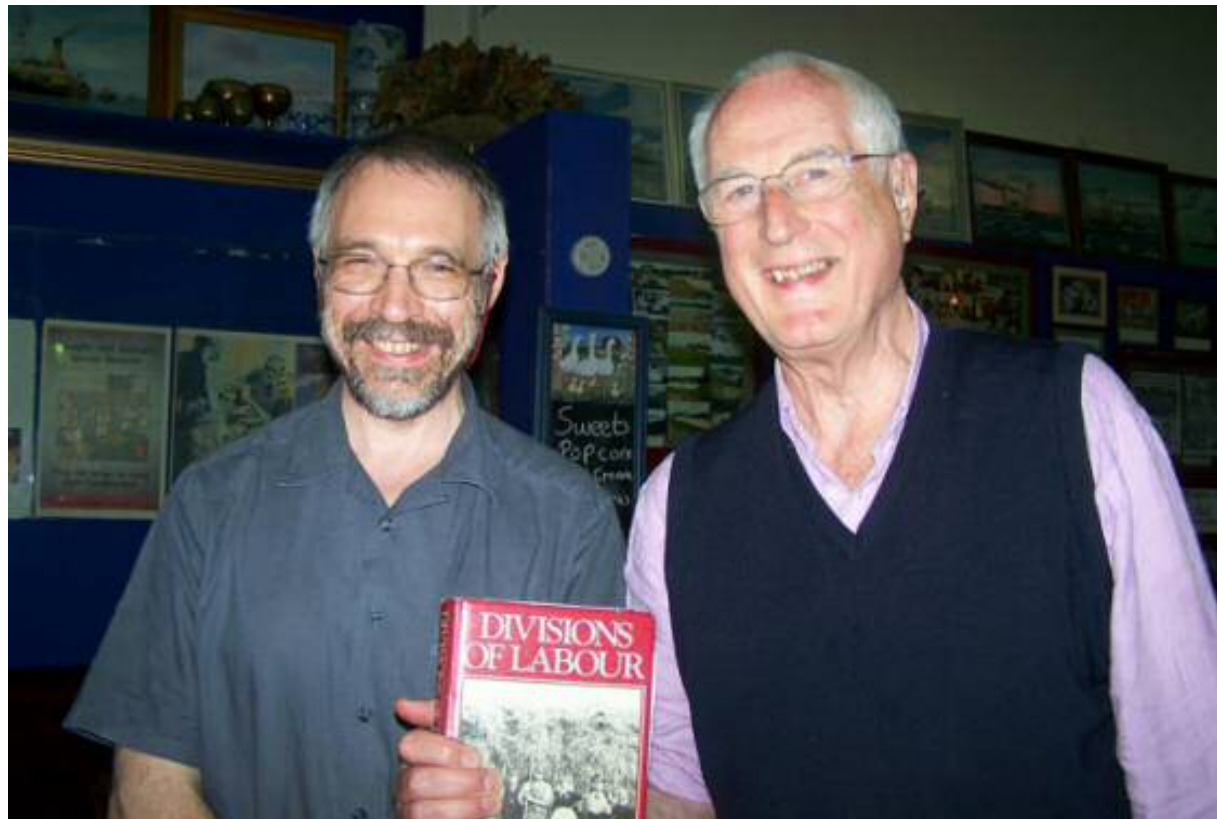
School Leavers Re-Study

Essay instructions 2010: Imagining the Future:

I want you to imagine that you are towards the end of your life. Look back over your life and say what happened to you. Don't write a very exaggerated story, just tell the straightforward story of your life as it might really be. Of course you cannot know what is going to happen to you, but you can describe the sort of thing that could happen if things go as you expect or hope. Spread your story over your whole life from the time of leaving school. Continue on another sheet as necessary.

Sheppey: Essay 30
Group: Boys

My Past

I was sitting in front of the fire smoking my pipe, just thinking of the past. I was nearly seventeen when I left school, I had an apprenticeship in the Tudor Glass, an electrical apprenticeship lasting four years, I served my four years and got my diploma for an electrical engineer. I was now twenty and courting a rearly lovely girl. I had to work in London in my next job and came back home every weekend. Jobs were scarce so you had to take what you can get, I didn't want to go to London,

UK Data Service

Living and Working on Sheppey

1978: **141 essays** (89 boys and 52 girls)
'Living without a job: how school leavers see the future' published in *New Society* in 1978 (2 November, pp. 259-62).

2009-10: The exercise was repeated by the Living and Working on Sheppey project and **110 essays** (55 boys and 55 girls) were gathered from school pupils and members of youth groups on the Isle of Sheppey to compare to the earlier ones.

Together, these two sets of essays shed light on the aspirations of Sheppey's young people (and young people more generally) and cover a range of topics including

• Health, Education, Career, Family and Leisure

UK Data Service

1978 Mundane & grounded jobs; gradual career progression; periods of unemployment.	2010 Well-paid, instantaneous jobs; choice but uncertainty; influence of celebrity culture.
"It was hard finding a job, I failed a few chances, but eventually got what I wanted locally, a craft apprenticeship." (Essay 27, male)	"I could not decide what to do. Then it came across my mind that I should be a body builder, and be good enough to enter the Mr Olympia contest. After that I went to an athletics club, within a few months I was scouted for the Olympics for the 200 metre sprint." (Essay 10, male)
"I was on the dole for six months after leaving school, until I got a job in a garage." (Essay 42, male)	"I was 20 now living the dream I had a amazing band...I had toured the world 3 times sold 4 million records." (Essay 30, male)
"When we found a house it was a semi-detached in Sittingbourne. I wanted to live in Italy but that was asking for too much." (Essay 56, male)	"I arrive at my 3-bedroom luxury villa; I land my helicopter on my own heli-pad and walk inside. I grab my keys and jump in my Bentley Continental GTS." (Essay 40, male)
"I longed for something exciting and challenging. But yet again I had to settle for second best. I began working in a large clothes factory." (Essay 104, female)	"In my future I want to become either: a dance teacher, hairdresser, or a Professional Show Jumper/horse rider. If I do become a dancer my dream would be to dance for Beyoncé or some really famous." (Essay 61, female)

School Leavers Study: publications
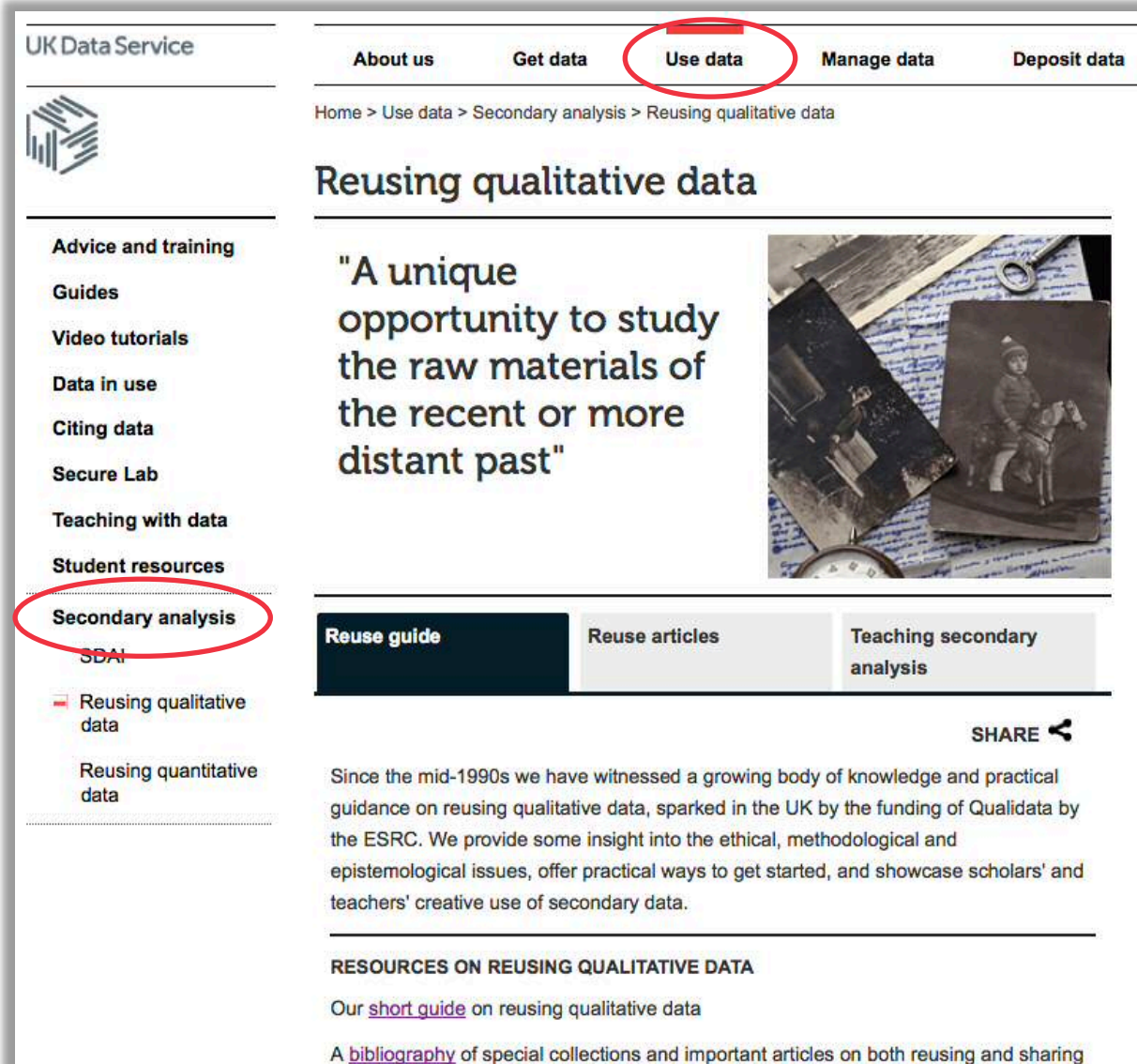
- Lyon, Dawn, and Graham Crow (2012) 'The challenges and opportunities of re-studying community on Sheppey: young people's imagined futures', *Sociological Review* 60: 498-517.

- Lyon, Dawn, Bethany Morgan, and Graham Crow (2012) 'Working with material from the Sheppey archive', *International Journal of Social Research Methodology 15(4): 301-309.*

- Weddell, Emma, Graham Crow and Dawn Lyon (2012) 'Imagining the Future, What can the aspirations of school-leavers in 1978 and 2010 tell us about the changing nature of society?' *Sociology Review*

- Crow, G. and N. Takeda (2011) 'Ray Pahl's Sociological Career: Fifty Years of Impact', Sociological Research Online, 16 (3) 11: http://www.socresonline.org.uk/16/3/11.html

- Crow, G., Hatton, P., Lyon, D. and Strangleman, T. (2009) 'New divisions of labour?: Comparative thoughts on the current recession', Sociological Research Online vol14 issue 2/3 http://www.socresonlline.org.uk/14/2/10.html

UK Data Service

Getting started

- Orient yourself to the original research project
  - Documentation and metadata

- Understand the structure of the original data
  - Contexts – at 3 levels
  - Sampling and recruitment

- Become familiar with the project data as a whole
  - Logic of subsampling if volume prohibits reading it all
  - Both inductive and deductive strategies can work
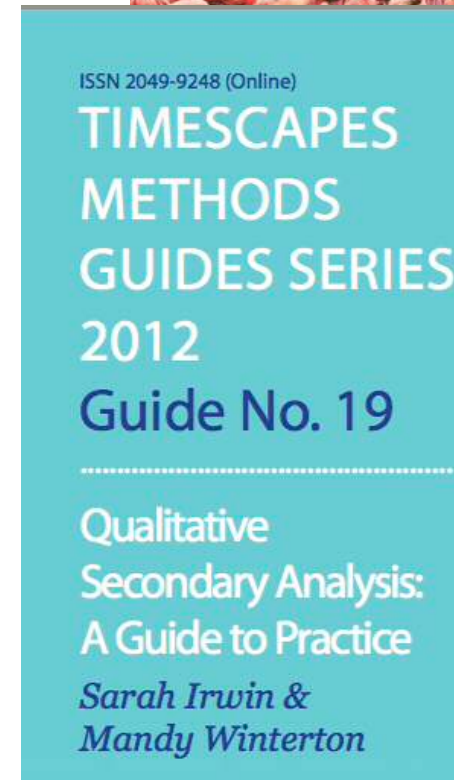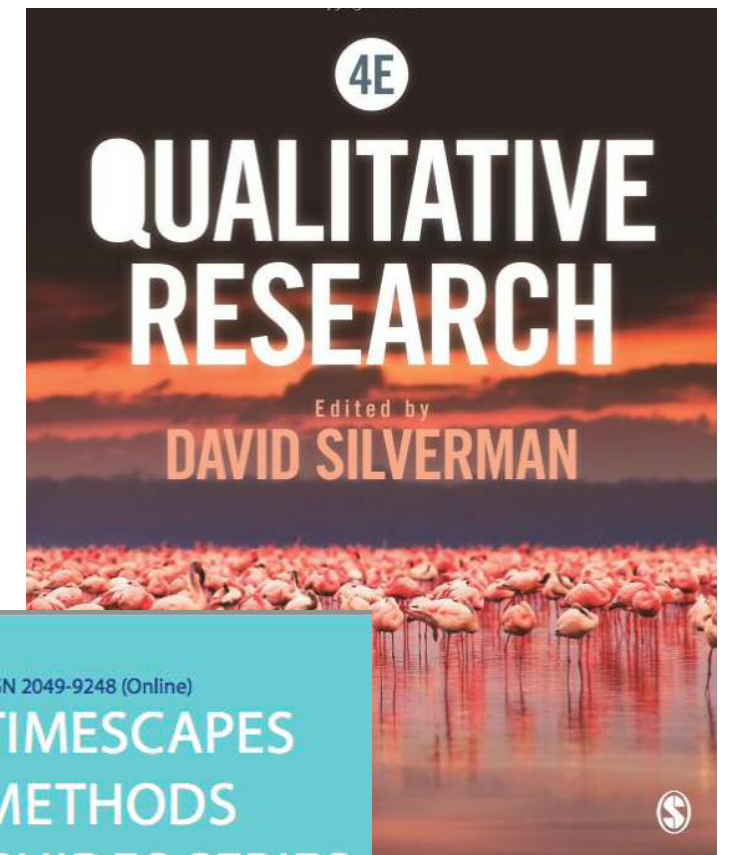
Publications and other resources for reusing data

http://ukdataservice.ac.uk/use-data/secondary-analysis/reusing-qualitative-data.aspx

Find data: Key data

Find data: Search data catalogue

QualiBank

Essay by 5

Discover

Variable and question bank

• QualiBank

External resources | Images   Access this collection from Discover

DETAILS

Create citation

Collection ID (SN):	4867
Title:	School Leavers Study, 1978
Principal investigator:	Pahl, R.E., University of Kent at Canterbury. Faculty of Social Sciences
Sex:	Male
Age group:	16-24
Socio-economic status:	Not classified
Region:	South East
Licence:	Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence.

RELATED RESOURCES

Related external resources

XML TEI Transcript

XML QuDEx metadata

Related images

Image relating to Essay by 5

...sics and history I **narrowly** scraped passes and by
...prolonged stay away from home but I wasn't to
...common with this is how my friendships have been
...ge I had been **fascinated** with the past and **since**
the age of fourteen I had regularly attended archaeological digs and so it followed that I got a job as **assistant** site manager the pay
wasn't much but I was **really** happy the odd girlfriend came and went but as I spent a great deal of time in a tent or **digging** holes I **didn't**
get too attached but one day I made the mistake and decided to settle and have a home so I tried to get a post somewhere where I could
continue poking in things that have **happened** but there were no opening I then I felt I had wasted my education on learning something I
liked when I should have **learnt** **something** to make money so I had to decided to live as a probably **bachelor** doing what I liked or as a
someone doing a **menial** task but with **security** and a **permanent** home on one side was the trap of **depression** and **poverty** and on the
other a high chance of not having a family I chose to stay as I was a hoop and if the **opportunity** came to have a job I liked where I could
still go out and dig but still have a **permanent** base I would take it and so when my chance came to take a job in a **museum** came I **leapt**
at it and apart for changing **museums** once or twice I stayed that way till I retired last month. I did get married to someone who I met on a
archaeological site who been in a similar kind of position we had very similar tastes in most things and we both shared nearly the same
views of life.

UK Data Service

Cite the Data

UK Data Service

Key issues in reusing data - summary

- **Many benefits**

  - large scale research, ethics completed, cost effective

- **Iterative**

  - research questions, locate and evaluate data, analysis

- **Context**

  - consider methodology and prior research

- **Plan and prepare**

  - understand data through study documentation

- **Support**

  - more information available in the UK Data Service website

- **Cite the data**

UK Data Service

Forthcoming Events

- **Webinars**
  - 26 May, 11.30-12.30pm: Dissertation projects: Introduction to secondary analysis for qualitative and quantitative data
  - 28 May, 3-4pm: Data management basics
  - 2 June, 1-2pm: Introduction to Text-Mining
- **Online Conferences**
  - 30 June: Health Studies User Conference
  - 8 July: Family Finance Surveys User Conference
- **UK Data Service Dissertation Award**
  - Deadline: 3 June
  - The Award recognises outstanding undergraduate student dissertations that include excellent reuse of data from the UK Data Service.

UK Data Service

Questions

Website > Use Data > Secondary Analysis
https://www.ukdataservice.ac.uk/use-data/secondary-analysis.aspx

email
help@ukdataservice.ac.uk

Video guides and webinars:
https://www.youtube.com/user/UKDATASERVICE

To follow UK Data Service on Twitter:
@UKDataService