Social Science Research Data 2023 and Beyond

Simon Parker

UK Data Service

UK Data Service

About the UK Data Service

- We are the UK's largest collection of social, economic and political data

- A large number of Secure Access datasets

- A majority of our data is available for registered users to download

- Some data can be analysed online, using tools such as Nesstar and UKDS.stat

- A number of key longitudinal studies including the UK Household Longitudinal Study (Understanding Society)

UK Data Service

Understanding Society

- Around 40,000 households have been involved in the study

- Continues and expands upon the British Household Panel Survey

- Data collection for UKHLS began in 2009

- BHPS Cohort - over 25 years worth of data collected

- Variables collected focused on people's social and economic circumstances, attitudes, lifestyle, health, family relationships and employment

- Secure Access versions of the data have lower levels of geography, full dates of birth, and are linked to the National Pupil Database

- Biomarkers and genetic data

UK Data Service

UKHLS and biomarker research

- Genomics of social support, personality and cognition and their relation to mental health and cognitive ageing

- Assortative mating and genetics

- Investigating the genetic relationships between anxiety, depression, stressful life outcomes, and cardiovascular risk factors and disease

- Understanding the genetics of neurodevelopmental disorders

UK Data Service

Linked data

- The Big Data era – more data collected in 2017 than previous 7,000 years

- Volume, Variety, Velocity, Veracity, Value

- Will surveys become bigger? Maybe…

- Will data become bigger? Yes

- Data will grow through linkage to other data – administrative, health, NNFD

- This increases the utility of the data and can reduce costs

UK Data Service

Linked data – the risks

- By increasing the number of variables, we increase the risk of disclosure.

UK Data Service

Linked data – the risks

ID	Region of residence	Age	Sex	Marital status	Employment status	Country of Birth	Tenure	Ethnic Group	Hours worked last week	Region of work (main job)
1	South East	1	Female	Single	N/A	UK	Being bought with mortgage	White	N/A	N/A
2	West Midlands	4	Female	Single	N/A	UK	Being bought with mortgage	White	N/A	N/A
3	Eastern	5	Female	Single	N/A	UK	Being bought with mortgage	White	N/A	N/A
4	South East	7	Female	Single	N/A	UK	Rented Private	White	N/A	N/A
5	Northern Ireland	10	Male	Single	N/A	UK	Rented LA	White	N/A	N/A
6	South East	10	Female	Single	N/A	UK	Being bought with mortgage	Indian	N/A	N/A
7	North East	12	Male	Single	N/A	UK	Rented Private	Indian	N/A	N/A
8	Esat Midlands	15	Male	Single	N/A	UK	Rented LA	White	N/A	N/A
9	London	16	Male	Widowed	Employed PT	UK	Being bought with mortgage	Indian	15	London
10	North West	18	Male	Single	Student	UK	Rented Private	White	0	N/A
11	South East	20	Female	Single	Student	UK	Rented Private	White	10	South East

Id	Label	timeset	twitter_type	created_at	lang	description	email	friends_c	followers	real_name	location
@privateguy	Tough year, birthday next week though #17	<[2018-05-10T11:47:03.495Z]>	Tweet	Thu May 10 12:47:02 BST 2018	en	Widow :-(		362	125	Simon Parker	London
@privateguy	@CR_UK not great news #fighting	<[2018-05-10T11:47:03.528Z]>	Tweet	Thu May 10 12:52:33 BST 2018	en	Widow :-(		362	125	Simon Parker	London
@privateguy	@widowsupport missing H today	<[2018-05-10T11:47:03.584Z]>	Tweet	Thu May 10 13:01:18 BST 2018	en	Widow :-(		362	125	Simon Parker	London

Healthcare in East London				
ID	Age	Sex	Marital status	Has Cancer?
1	35	Female	Married	Yes
2	28	Female	Single	No
3	63	Male	Married	No
4	42	Female	Divorced	No
5	55	Male	Single	Yes
6	70	Female	Widowed	No
7	16	Male	Widowed	Yes

UK Data Service

Linked data – the risks

- By increasing the number of variables, we increase the risk of disclosure.

- Data may be linked in ways not predicted by the data owners

- Risks can be mitigated by controlling access, or by *ex post* techniques such as the creation of synthetic data or differential privacy

- Archives have a role to aid researchers to overcome these challenges

UK Data Service

BREAKING NEWS

PARKER TO RETIRE

16:16 RESEARCH DATA STILL AVAILABLE ONLINE | ESOTERIC FILE STRUCTURE TO REMAIN

Preservation and sustainability

- Continued usability of data

- Mediated safe use of data

- OAIS digital repositories

- Support for users to maximise the research potential of data

- Expert data stewardship

UK Data Service

The End

The actual end this time!

ukdataservice.ac.uk/help/

Follow us at:

- ukdataservice@jiscmail.ac.uk
- twitter.com/ukdataservice
- facebook.com/ukdataservice

UK Data Service