CALLS-HUB

CENSUS & ADMINISTRATIVE DATA LONGITUDINAL STUDIES HUB

New LS developments

Chris Dibben

THE UNIVERSITY of EDINBURGH

Queen's University Belfast

UCL

University of St Andrews

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

CALLS-HUB
CENSUS & ADMINISTRATIVE DATA LONGITUDINAL STUDIES HUB

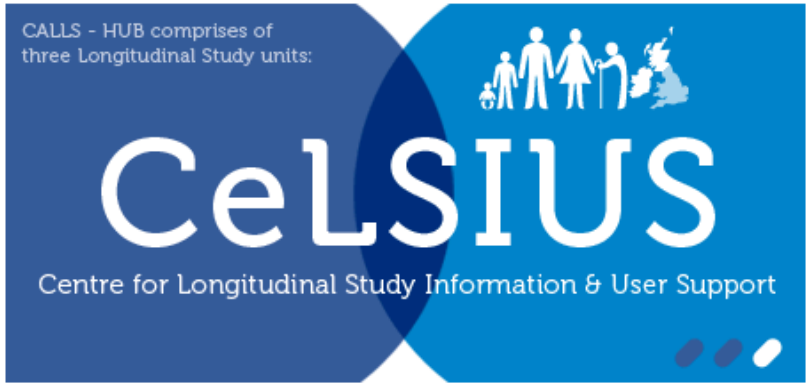
LS developments

- Calls-Hub
- E-dataSHIELD
- Extending the LSs back in time
- Broadening the Scope of the LSs

CALLS-HUB

Census & Administrative data LongitudinaL Studies Hub

Home | About | Support Units | Outputs | Guides & Resources | News | Events | Blog | Variables | More



QUICK DATA DICTIONARY SEARCH

Keyword

SEARCH

Run a refined search to build a list of variables for use in your project application form.

Welcome to the website for the Census & Administrative data LongitudinaL Studies Hub (CALLS-Hub).

Here you can find out more information about the three UK census-based longitudinal studies (LSs) and their support units CeLSIUS, SLS-DSU, NILS-RSU

- What data do the LSs hold?
- How to apply to use LS data in your research
- Using more than one LS dataset for your analysis

Outputs

User Guides

LATEST TWEETS

NINIS @NISRANINIS 7h

Median Sale Price of Residential Properties interactive map with the latest figures out today bit.ly/1h5d2mW pic.twitter.com/ihnM8vBfST

Retweeted by CALLS

Show Photo

ADLS @ADLStweet 18 Feb

Due to high demand extra tickets now available for the UK Longitudinal Studies Linkage Launch. Further info at calls.ac.uk/event-entry/uk...

Retweeted by CALLS

Expand

RECENT NEWS

New email addresses – LSCS and SLS-DSU

Contact details for the Longitudinal Studies Centre Scotland (LSCS) and Scottish Longitudinal Study Development Support Unit (SLS-DSU) have now changed! In addition... Read more...

Speakers announced for UK LS 2011 Census Linkage Launch

Announcing the list of speakers for the Census 2011 launch event, to be held on Thursday 6th March 2014 at... Read more...

Census and Administrative data Longitudinal Studies Hub - CALLS

- One stop shop – information and advice
- Data dictionary
- Examples of uses of the LS's – impact
- Help with running UK level analysis

Using more than one LS for your research

Guides & Resources

Using more than one LS for your research

Census forms

Synthetic LS data

Useful documents

Applying to use the LSs

FAQs

Due to Data Protection concerns, the current agreements in place for the LSs mean that it is not possible to access LS data outwith its respective safe-setting. This means that to-date it has not been possible to fully combine datasets to allow comparisons between regions or to create a UK-wide sample. Instead the only option has been to conduct separate analyses and then compare or combine the results on an ad-hoc basis.

Work in the SLS-DSU by Prof Gillian Raab has resulted in a new methodology which means it is now possible to analyse data from one or more LS *as though they were part of the same dataset*. This is achieved through a process called eDatashield, which is currently able to support all forms of GLM regression analyses.

If you are interested in finding out more about eDatashield or would like to discuss using more than one LS for your study, please contact us.

QUICK DATA DICTIONARY SEARCH
Keyword

SEARCH

RECENT NEWS

New email addresses – LSCS and SLS-DSU

Speakers announced for UK LS 2011 Census Linkage Launch

UPCOMING EVENTS

UK LS Census Linkage Launch Event – registration now open

March 6, 2014 at Harvey Goodwin Suite, Church House, Westminster, SW1P 3NZ

CALLS-HUB
Census & Administrative data LongitudinaL Studies Hub

E-dataSHIELD

- Method for running analysis across horizontally split datasets
- Can break down the iterative step within Generalised Linear Models to separate cases
- Software developed in R

CALLS-HUB
CENSUS & ADMINISTRATIVE DATA LONGITUDINAL STUDIES HUB

Iteratively reweighted least squares

Score vector

Information matrix

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \boldsymbol{I}(\boldsymbol{\beta}_t)^{-1}\boldsymbol{s}(\boldsymbol{\beta}_t)$$

CALLS-HUB
Census & Administrative data LongitudinaL Studies Hub

$$\boldsymbol{I}(\boldsymbol{\beta}_t) = \sum_{i=1}^{N} w_{ii}(t)\, \boldsymbol{x}_i \boldsymbol{x}_i^T,$$

and

$$\boldsymbol{s}(\boldsymbol{\beta}_t) = \sum_{i=1}^{N} (y_i - \mu_i(t)) g'\big(\mu_i(t)\big) w_{ii}(t) \boldsymbol{x}_i.$$

Norsk Epidemiologi 2012; **21** (2): 231-239

DataSHIELD – shared individual-level analysis without sharing the data: a biostatistical perspective

E.M. Jones[1], N.A. Sheehan[1], N. Masca[1], S.E. Wallace[1], M.J. Murtagh[1] and P.R. Burton[1,2]

1) Department of Health Sciences, University of Leicester, UK
2) Public Population Project in Genomics (P^3G), Montreal, QC, Canada

Correspondence: Paul Burton, Department of Health Sciences, University of Leicester, Room 317 Adrian Building, University Road, Leicester LE1 7RH, United Kingdom
E-mail : pb51@le.ac.uk Telephone: +44 (0)116 229 7251 Telefax: +44 (0)116 229 7250

ABSTRACT

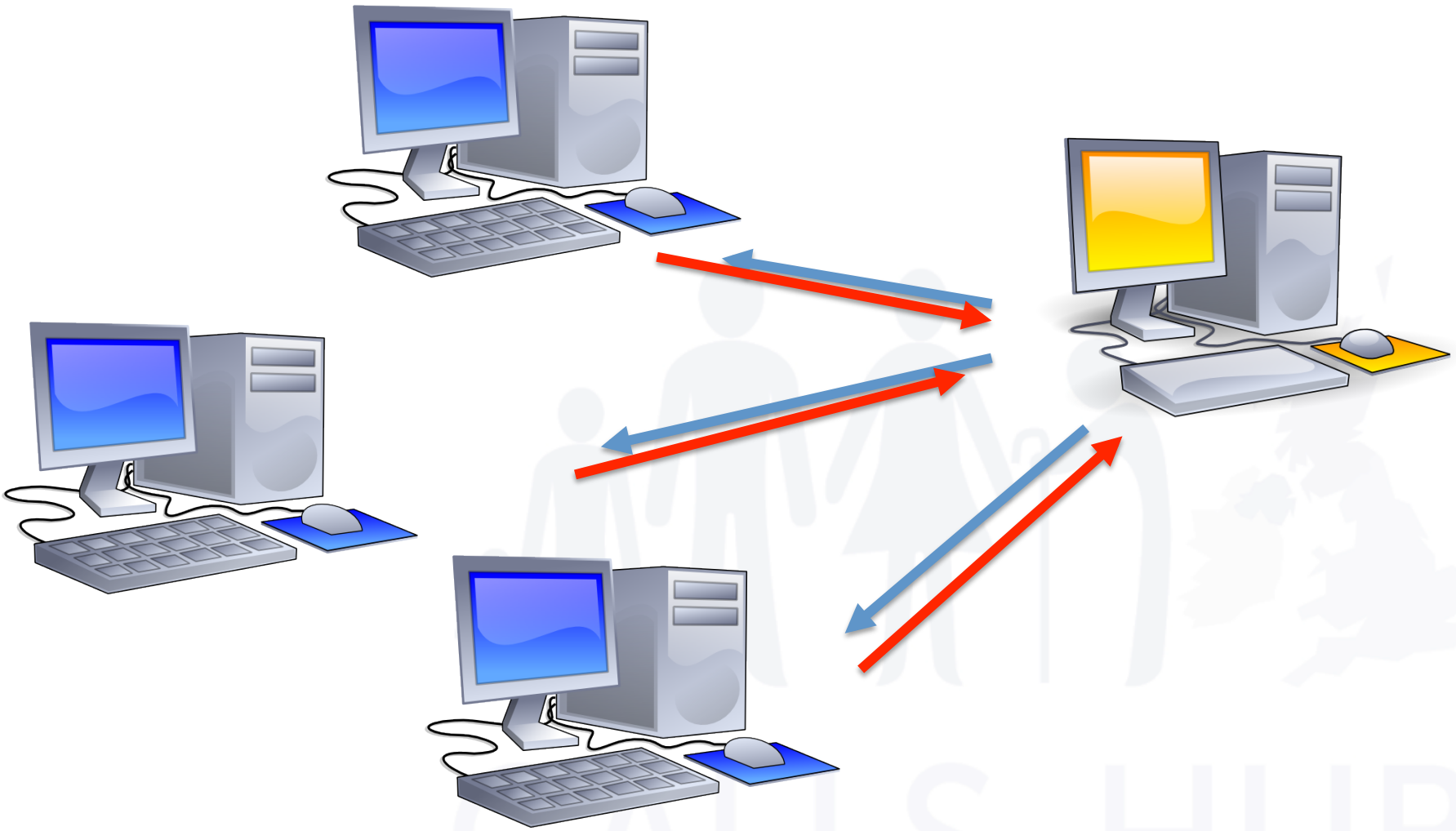
Very large sample sizes are required for estimating effects which are known to be small, and for addressing intricate or complex statistical questions. This is often only achievable by pooling data from multiple studies, especially in genetic epidemiology where associations between individual genetic variants and phenotypes of interest are generally weak. However, the physical pooling of experimental data across a consortium is frequently prohibited by the ethico-legal constraints that govern agreements and consents for individual studies.

Study level meta-analyses are frequently used so that data from multiple studies need not be pooled to conduct an analysis, though the resulting analysis is necessarily restricted by the available summary statistics. The idea of maintaining data security is also of importance in other areas and approaches to carrying out 'secure analyses' that do not require sharing of data from different sources have been proposed in the technometrics literature. Crucially, the algorithms for fitting certain statistical models can be manipulated so that an individual level meta-analysis can essentially be performed *without* the need for pooling individual-level data by combining particular summary statistics obtained individually from each study. DataSHIELD (Data Aggregation Through Anonymous Summary-statistics from Harmonised Individual levEL Databases) is a tool to coordinate analyses of data that cannot be pooled.

In this paper, we focus on explaining why a DataSHIELD approach yields identical results to an individual level meta-analysis in the case of a generalised linear model, by simply using summary statistics from each study. It is also an efficient approach to carrying out a study level meta-analysis when this is appropriate and when the analysis can be pre-planned. We briefly comment on the IT requirements, together with the ethical and legal challenges which must be addressed.

CALLS-HUB

Census & Administrative data LongitudinaL Studies Hub



CALLS-HUB
Census & Administrative data LongitudinaL Studies Hub

Extending the LSs back in time

- The NILS team have been linking the 1991 census to the study - 20 year period of observation
- The E&W LS team are starting an exploratory study into linking the 1961 census – 50 years of observation may be possible
- SLS – have linked sample members born in 1936 to their 1939 register and 1947 SMS

Scottish Mental Survey

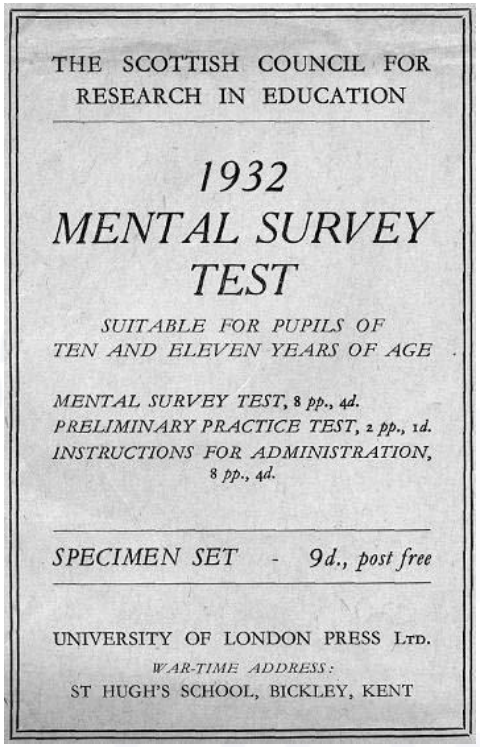
(2) Write the three letters between A and E and cross out the middle one ... ... ( )

(3) Finger is to hand as toe is to what? The answer is one of the five words in the bracket. Underline the right word ... (foot, knee, arm, shoe, nail)

You have nothing to write, only UNDERLINE what you think is the right answer.

(4) Man is to clothes as what is to fur?

(coat, animal, bird, skin, cloth)



THE SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION

1932

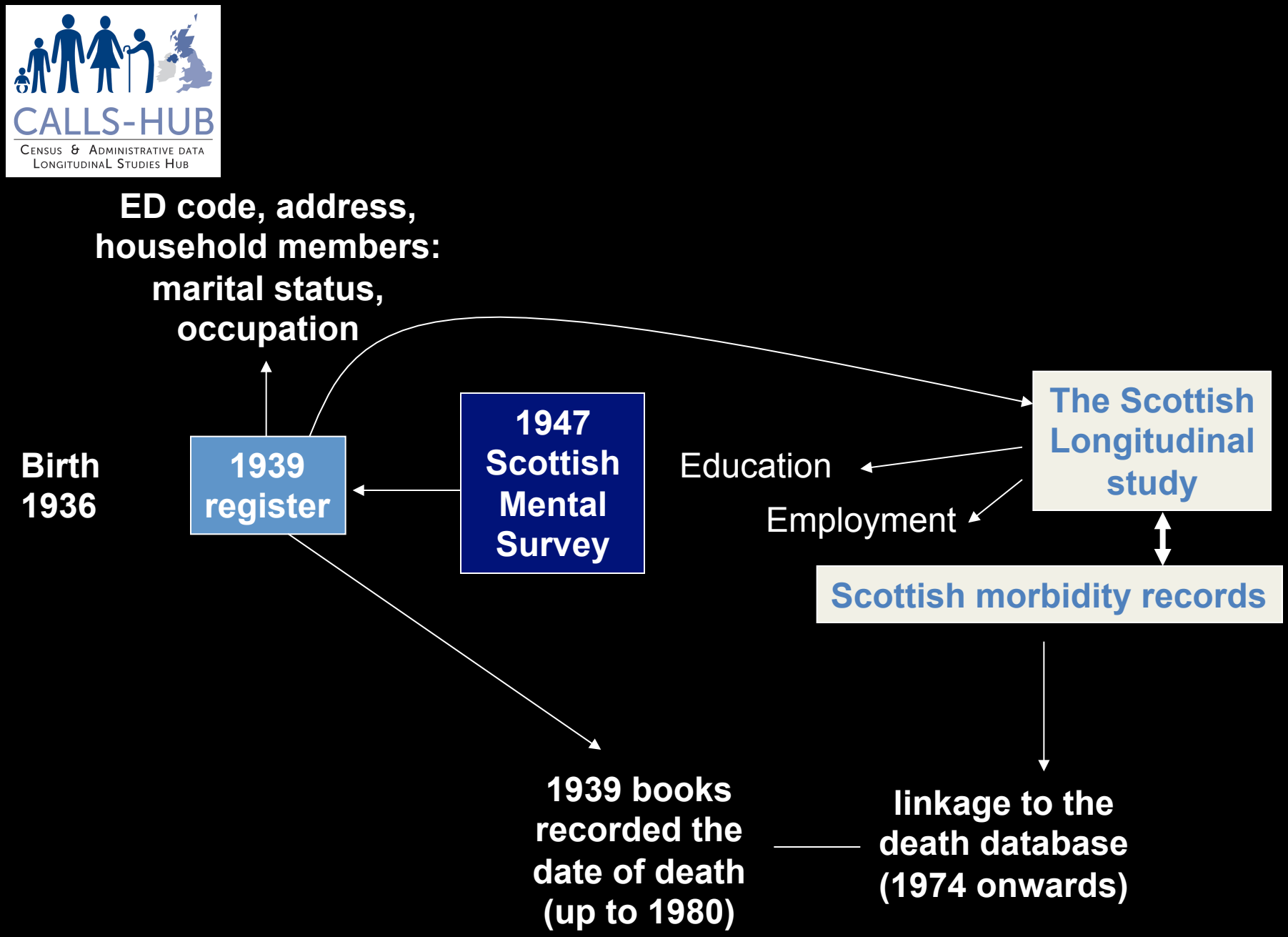
MENTAL SURVEY TEST

SUITABLE FOR PUPILS OF TEN AND ELEVEN YEARS OF AGE

MENTAL SURVEY TEST, 8 *pp.*, 4*d.*
PRELIMINARY PRACTICE TEST, 2 *pp.*, 1*d.*
INSTRUCTIONS FOR ADMINISTRATION, 8 *pp.*, 4*d.*

SPECIMEN SET - 9*d.*, *post free*

UNIVERSITY OF LONDON PRESS LTD.
WAR-TIME ADDRESS:
ST HUGH'S SCHOOL, BICKLEY, KENT


CALLS-HUB
Census & Administrative data LongitudinaL Studies Hub
ED code, address, household members: marital status, occupation
Birth 1936
1939 register
1947 Scottish Mental Survey
Education
Employment
The Scottish Longitudinal study
Scottish morbidity records
1939 books recorded the date of death (up to 1980)
linkage to the death database (1974 onwards)

Age	0	11	34	55	65	75
Year		1947	1970	1991	2001	2011

Mortality

Hospitalisation

Birth 1936

Occupation (estimated)

Mental ability

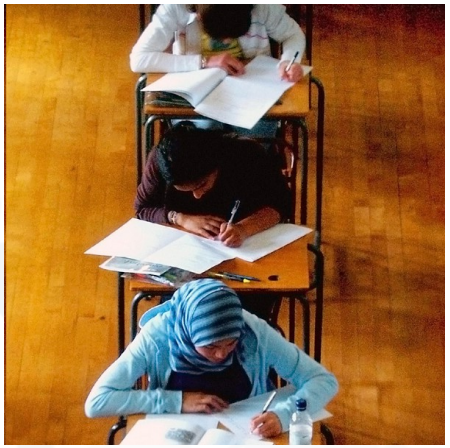
Early life environment

School Achievement (time estimated)

Detailed household/ individual information

Broadening the Scope of the LSs

- NILS has linkage to Property Service (LPS) data
- NILS makes one-off links to Health & Social Care data (e.g. prescription data)
- The SLS-DSU has linked annual school and attainment data to the SLS and is exploring higher education data



This will only increase over the coming years