

# Meeting new challenges posed by the UK Censuses

Oliver Duke-Williams / Vassilis Routsis

UK Data Service and UCL Department of Information Studies

@oliver\_dw

[o.duke-williams@ucl.ac.uk](mailto:o.duke-williams@ucl.ac.uk)



- **Challenges for the Data Service**
- Unclear future: administrative data
- Challenges for longitudinal data
- Challenges for the 22<sup>nd</sup> century



# UK Data Service

- ESRC funded service for providing research ready (social science) data
  - *Curation, metadata, documentation, training etc*
  - *Provides access to data in a number of repositories*
- Census group is spread across multiple institutions
  - *UCL*
  - *Cathie Marsh Institute, University of Manchester*
  - *EDINA, University of Edinburgh*
  - *Jisc*

# Census data

- UKDS Census provides access to a variety of outputs
- Census data can be divided in a number of ways
  - Types of data
    - Aggregate data; flow data; microdata; boundary data
  - Licenses
    - Open; Safeguarded; Secure
  - Historic / 'recent'
    - UKDS currently have (some types of) data from 1961 onwards
    - Separately, the Data Archive hold individual data from the 1851-1911 censuses, through the Integrated Census Microdata (I-CeM) project

# 'UK' Censuses – geographic scope

- We do not have a 'UK Census' but rather three separate decennial censuses, carried out by three agencies
  - England and Wales (Office for National Statistics, ONS)
  - Scotland (National Records of Scotland, NRS)
  - Northern Ireland (Northern Ireland Statistics and Research Agency, NISRA)
- There are also three Crown Dependencies (UKDS does not hold data from these)
  - Isle of Man
  - Bailiwick of Jersey
  - Bailiwick of Guernsey

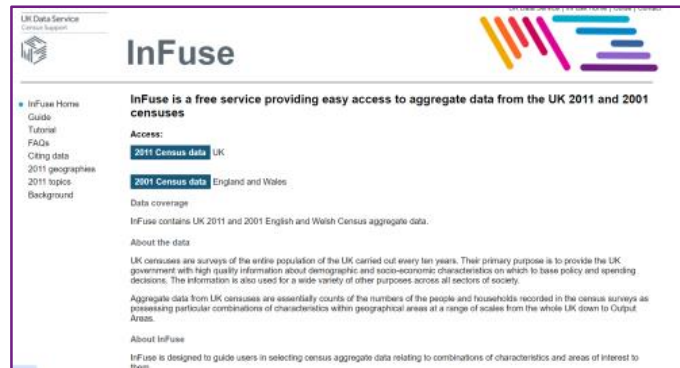
# Challenges

- Our legacy systems which provided access to data from 2011 and earlier are outdated
- With the exception of microdata, they do not integrate well with other UKDS holdings
  - Many tables, very large tables – require complex (bespoke?) discovery and sub-setting
- New competitors
  - Custom datasets available through ONS
- Updating training materials etc

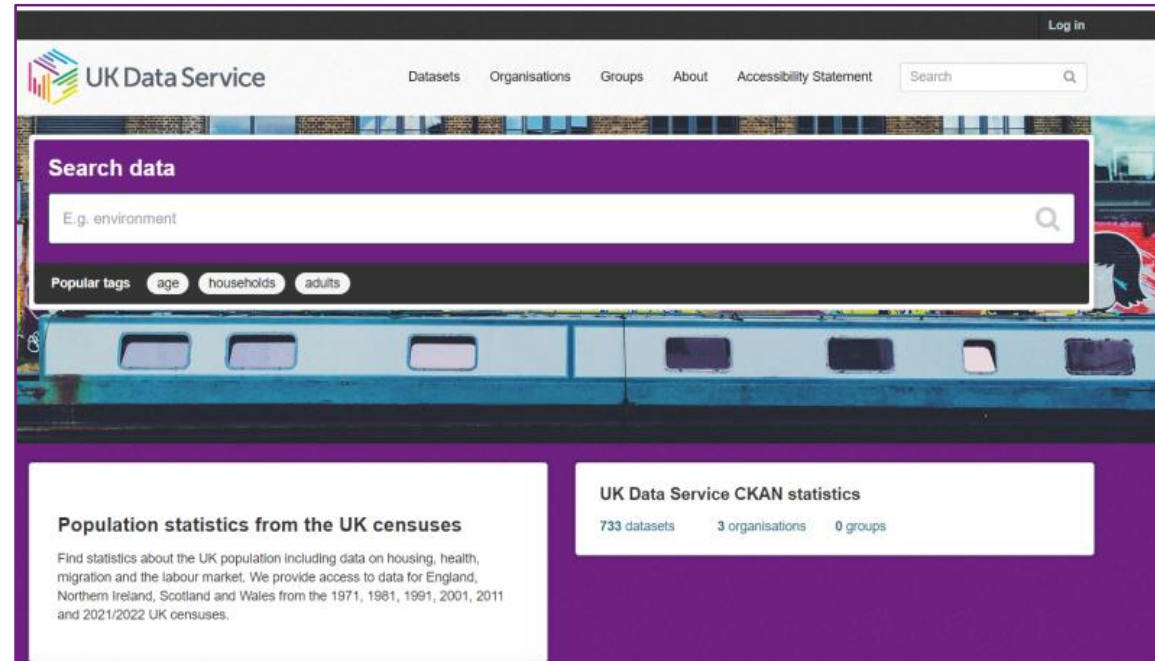
# Challenge: old systems



<https://casweb.ukdataservice.ac.uk/>



<https://infuse.ukdataservice.ac.uk/>



<https://statistics.ukdataservice.ac.uk/>

# Challenge: old systems

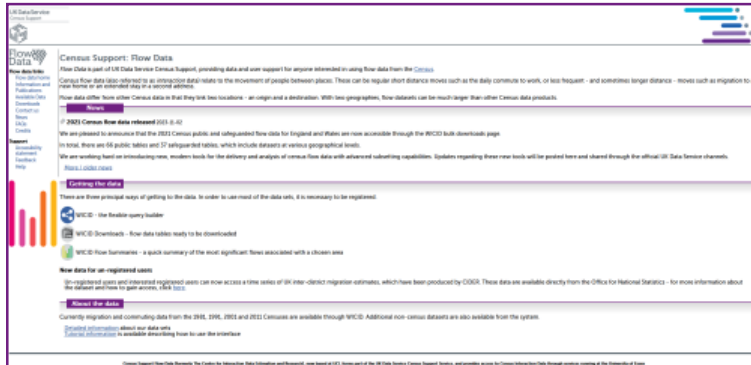
The screenshot displays the UK Data Service website interface. The main navigation bar includes 'UK Data Service', 'Datasets', 'Organisations', 'Groups', 'About', and 'Accessibility Statement'. A search bar is present in the top right. The left sidebar contains filters for 'Organisations' (Office for National Statistics: 709, Flow: 14, Boundary: 10), 'Tags' (UK Open Government: 731), 'Year' (2021: 462, 2011: 268, 2001: 1, 1991: 1, 1981: 1, 1971: 1), and 'Topics' (Demography: 279). The main content area shows a search for '733 datasets found' with an 'Order by: Relevance' dropdown. Three dataset results are visible, each with an 'XLSX' download button:

- Northern Ireland Census 2021 - DT-0003: Economic activity by Religion (8 cate...**  
This table provides Census 2021 estimates that classify people by Economic activity by Religion (8 categories) for Northern Ireland. The table contains 112 counts. The census...
- Northern Ireland Census 2021 - DT-0002: National identity (8 categories) by R...**  
This table provides Census 2021 estimates that classify people by National Identity (8 categories) by Religion or religion brought up in for Northern Ireland. The table contains...
- Northern Ireland Census 2021 - DT-0001: National identity (8 categories) by R...**  
This table provides Census 2021 estimates that classify people by National Identity (8 categories) by Religion (8 categories) for Northern Ireland. The table contains 64 counts....

An inset window in the top right shows a search for 'E.g. environment' with a search button and a 'Popular tags' section. A purple box with the URL <https://statistics.ukdataservice.ac.uk/> is overlaid on the search bar area.



# Challenge: old systems



<https://wicid.ukdataservice.ac.uk/>

JSON Raw Data Headers

Save Copy Collapse All Expand All [slow] Filter JSON

```

data:
  0: {}
  1:
    dataset_id: 2
    table_id: "1"
    sequence: 1
    table_title: "All migrants: age (5 broad age groups) by sex"
    dataset_label: "1991 SMS Set 1"
    dataset_description: "Ward level migration data from the 1991 Census"
    family_id: 1
    series_id: 2
    geographies: {}
    topics: {}
    theme_alias: []
  2:
    dataset_id: 2
    table_id: "2"
    sequence: 2
    table_title: "Wholly Moving Households_ving Households: counts"
    dataset_label: "1991 SMS Set 1"
    dataset_description: "Ward level migration data from the 1991 Census"
    family_id: 1
    series_id: 2
    geographies: {}
    topics: {}
    theme_alias: []
  3: {}
  4: {}
  5: {}
  6: {}
  7: {}
  8: {}
  9: {}
  10: {}
  
```

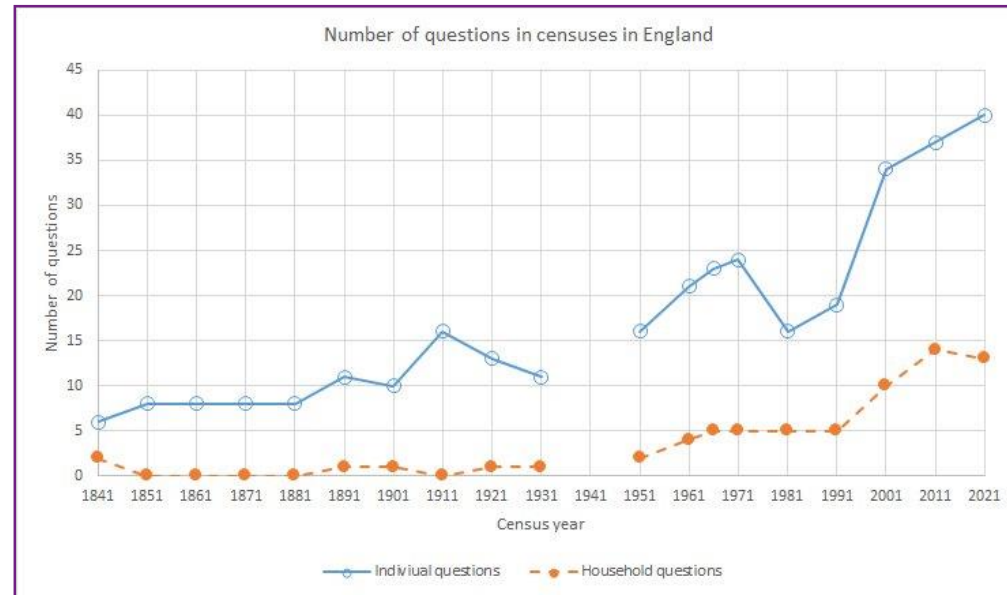
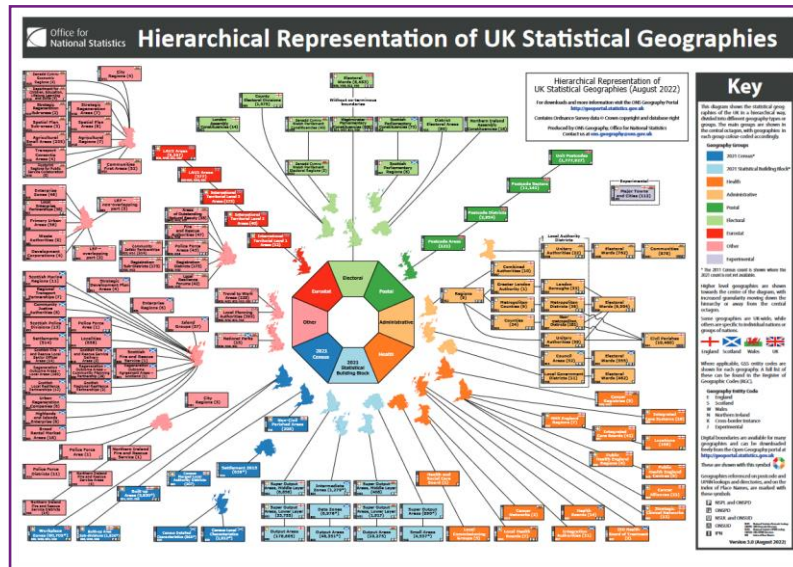
JSON Raw Data Headers

Save Copy Collapse All Expand All Filter JSON

```

data:
  0: {}
  1: {}
  2:
    geography_id: 1
    geography_label: "GB Districts 1991"
    number_of_areas: 459
    dataset_id: {}
  3:
    geography_id: 2
    geography_label: "GB Wards 1991"
    number_of_areas: 10933
    dataset_id: {}
  4:
    geography_id: 3
    geography_label: "GB Counties 1991"
    number_of_areas: 67
    dataset_id: {}
  5:
    geography_id: 4
    geography_label: "GB Standard Regions 1991"
    number_of_areas: 10
    dataset_id: {}
  6:
    geography_id: 5
    geography_label: "GB Countries 1991"
    number_of_areas: 3
    dataset_id: {}
  7:
    geography_id: 6
    geography_label: "GB Parliamentary Constituencies 1997"
    number_of_areas: 641
    dataset_id: {}
  
```

# Challenges: data volume and complexity



Source: [https://x.com/oliver\\_dw/status/1164521314360733701?s=20](https://x.com/oliver_dw/status/1164521314360733701?s=20)

Source: <https://geoportal.statistics.gov.uk/documents/ons::hierarchical-representation-of-uk-statistical-geographies-august-2022-1>

# Challenge: new competition

- ONS custom dataset builder
  - Easy to use
  - Long requested
  - Well designed

<https://www.ons.gov.uk/datasets/create>

The screenshot shows the 'Custom dataset' page on the ONS website. At the top, there is a navigation bar with the ONS logo and menu items: Home, Business, industry and trade, Economy, and Employment and labour market. Below the navigation bar is a search bar with the text 'Search for a keyword(s) or time series ID'. The main content area has a breadcrumb trail '< Start again - Create a custom dataset' and a large heading 'Custom dataset'. A prominent orange box contains the text '62 out of 331 areas available' and 'Protecting personal data will prevent 269 areas from being published.' Below this is a link '> How to improve your results'. The 'Variables' section is a list of configuration options, each with a 'Change' link:

- Population type**: All usual residents
- Area type**: Lower tier local authorities
- Coverage**: England and Wales
- Age of arrival in the UK**: 8 categories
  - Born in the UK
  - Arrived in the UK: Aged 0 to 15 years
  - Arrived in the UK: Aged 16 to 24 years
  - Arrived in the UK: Aged 25 to 34 years
  - Arrived in the UK: Aged 35 to 49 years
  - Arrived in the UK: Aged 50 to 64 years
  - Arrived in the UK: Aged 65 years and over
  - Does not apply
- Country of birth**: 60 categories
  - Europe: United Kingdom: England
  - Europe: United Kingdom: Northern Ireland
  - Europe: United Kingdom: Scotland

# Challenges for the (near) future

- Retaining a USP
  - UK level coverage
  - Access to older data
  - Discoverability

- Challenges for the Data Service
- **Unclear future: administrative data**
- Challenges for longitudinal data
- Challenges for the 22<sup>nd</sup> century



# An unclear future: the last census?

**BBC** Sign in Home News Sport Earth Reel W

**NEWS**

Home | Israel-Gaza war | War in Ukraine | Climate | Video | World | Asia | UK | Business | Tech

UK | England | N. Ireland | Scotland | Wales | Isle of Man | Guernsey | Jersey | Local News

## UK's 2021 census could be the last, statistics chief reveals

© 12 February 2020 · Comments



GETTY IMAGES

Source: BBC, <https://www.bbc.com/news/uk-51468919>

## National census in 2011 could be last of its kind

© 10 July 2010

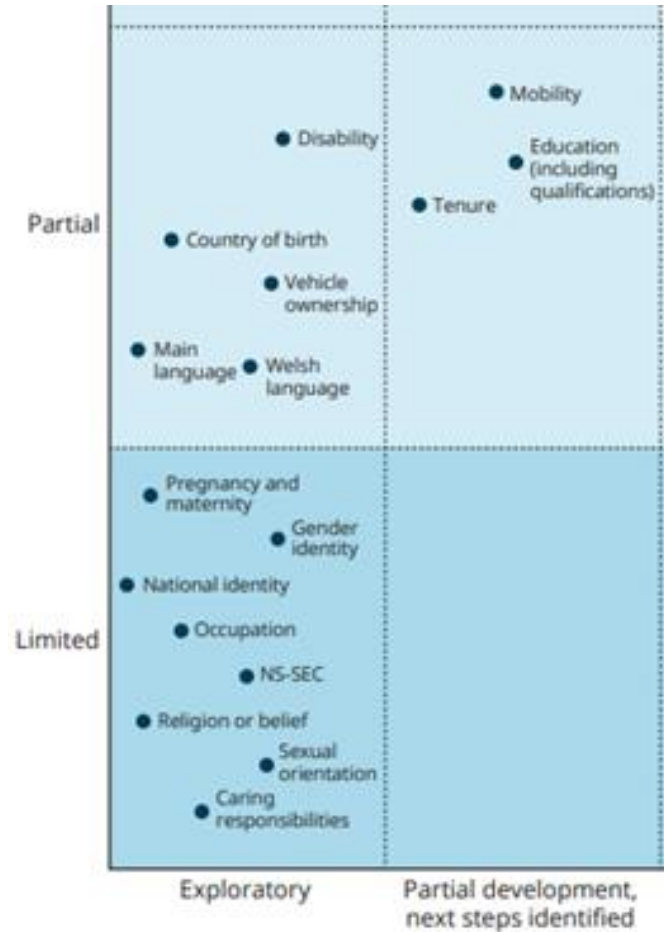


The census of all households is held every 10 years

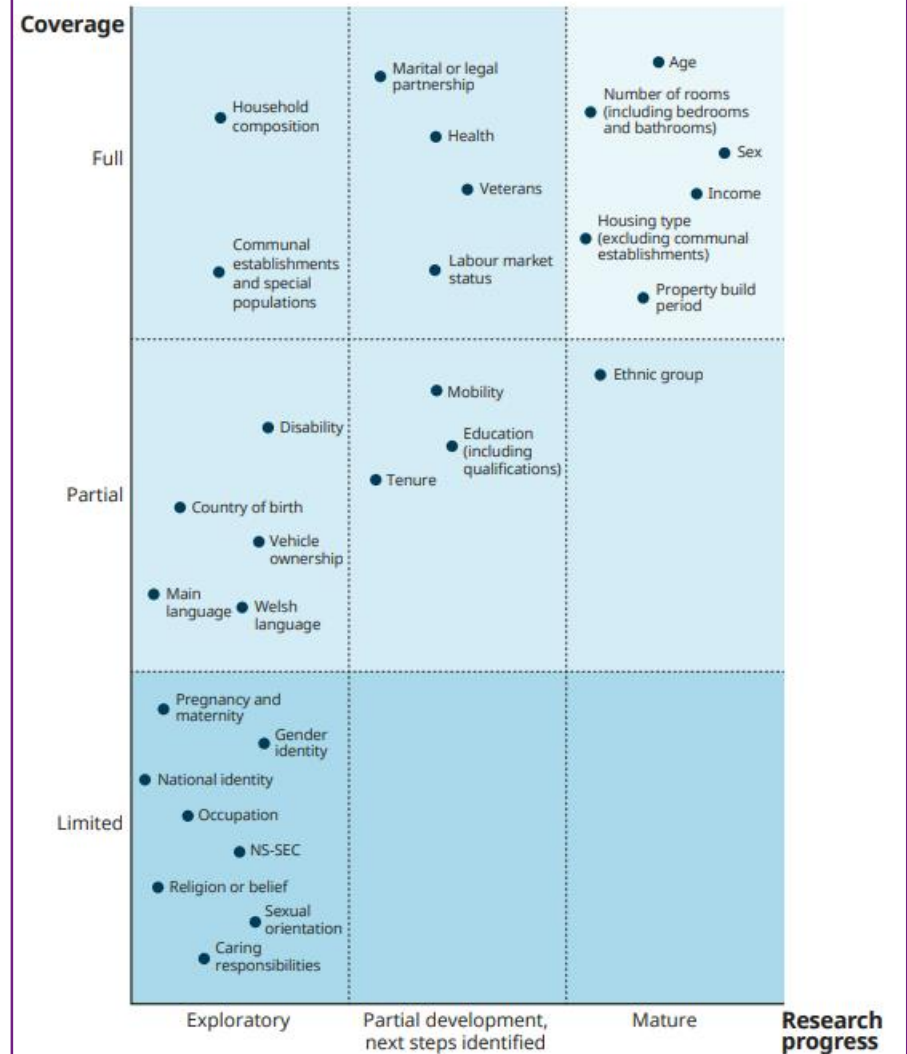
Source: BBC, <https://www.bbc.com/news/10584385>

# Challenges for data providers

- Linkage of records
- Range of variables
- Low level geography



**Figure 9:** Research on the availability of administrative data sources and their expected coverage varies by topic  
 Rating of administrative data source research (from exploratory to mature) and coverage (from limited to full) for characteristics



Source: <https://consultations.ons.gov.uk/ons/futureofpopulationandmigrationstatistics/>

# Challenges for repositories

- What will 'census' data look like in the future
  - 2021 may be the last census
  - Future products may be based on administrative data
    - What will be the best tools to disseminate these?
    - What will be the best approach to archiving data?
    - What assumptions can / should we make?
  - How will future researchers investigate change over time?



- Challenges for the Data Service
- Unclear future: administrative data
- **Challenges for longitudinal data**
- Challenges for the 22<sup>nd</sup> century



# About the ONS LS

- The ONS LS is based on 4 sample birth dates
  - Approximately 1% of population
  - Sample members do not know that they are in the sample
- The data for England and Wales include both census data (1971-2011) and administrative data, and some vital events data
- Access to the data is controlled

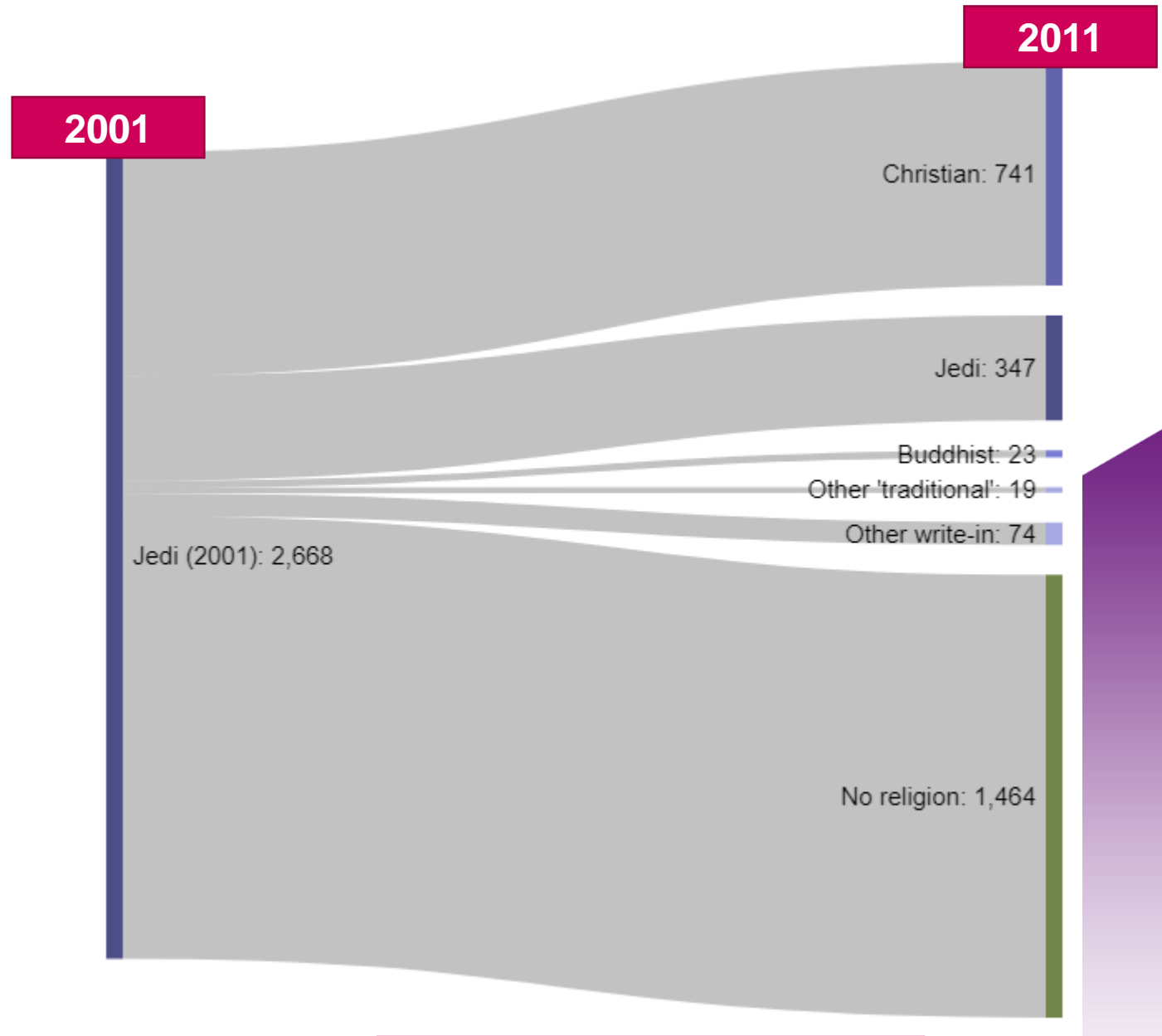
# Who changes religion?

- Religion is not an immutable characteristic; it can change over time
- The LS allows us to explore changes in religion
- We select sample members who were present at both 2001 and 2011

# Jedi exodus

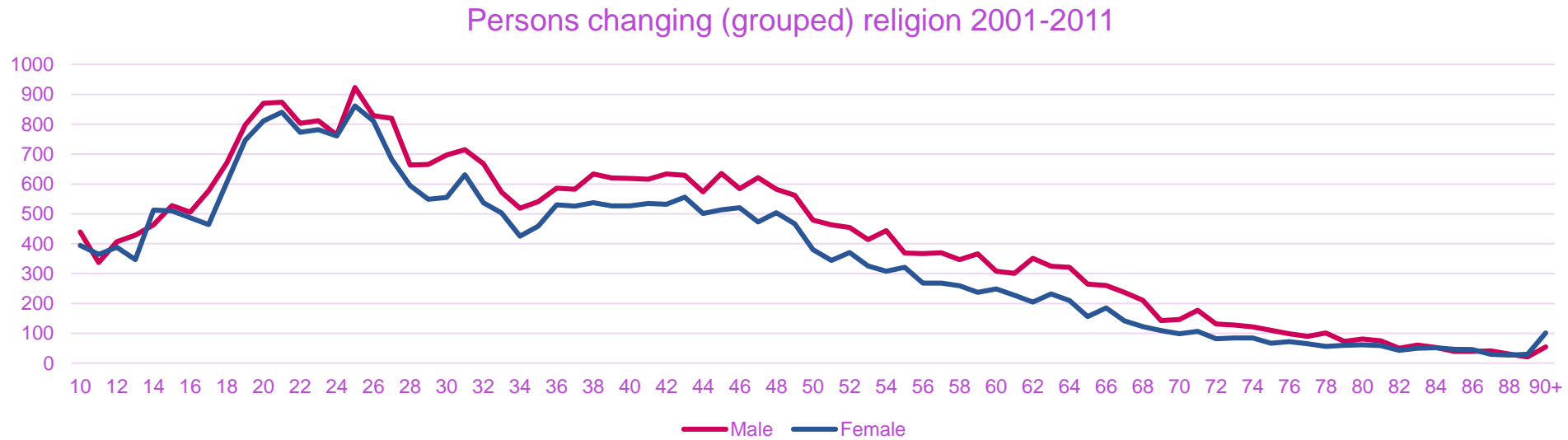


I find your lack of faith disturbing  
Star Wars (1977), Twentieth Century Fox



Source: ONS Longitudinal Study

# Who changes religion?



Source: ONS Longitudinal Study

# Challenges for the LS

- How to proceed without new census data?
  - Continue to add admin data?
  - What will this contain? Will it be consistent?
- How to proceed with admin census data?
  - How important are variables that might not be included?

- Challenges for the Data Service
- Unclear future: administrative data
- Challenges for longitudinal data
- **Challenges for the 22<sup>nd</sup> century**

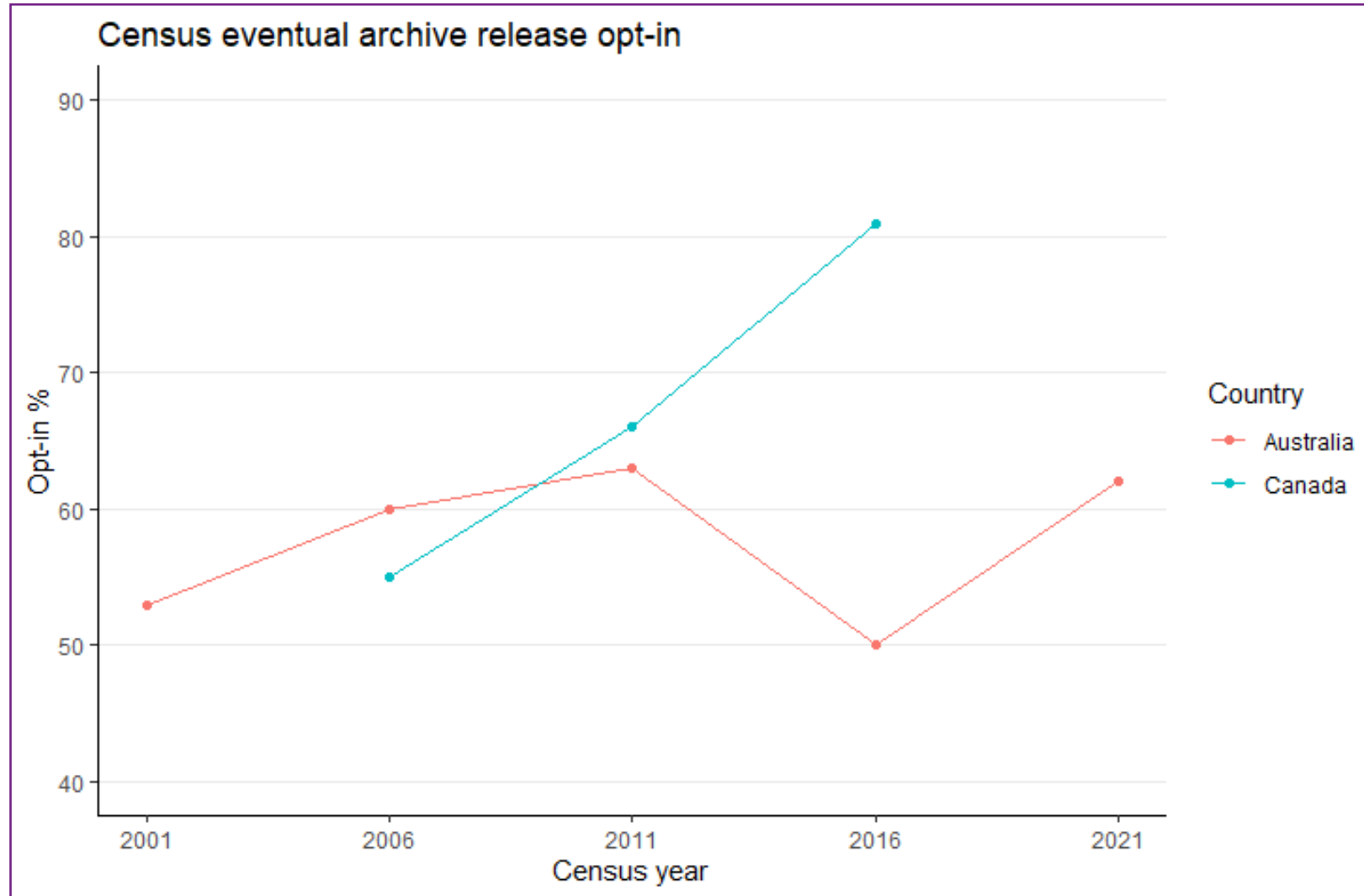


# Challenges for the (far) future

- In the UK: all data retained, released after 100 years
  - What data should be retained for future historians, historical demographers, genealogists etc?
  - Census schedules make it fairly obvious (most of the time) what to keep
  - How do we select administrative data to retain?
  - How important is emotional engagement with the retained data? How excited will our descendants be about transactional records of routine occurrences?



# Privacy/trust and data retention



# Census form of Constance Tite

**CENSUS OF ENGLAND AND WALES, 1911.**

*Before writing on this Schedule please read the Examples and the Instructions given on the other side of the paper, as well as the headings of the Columns. The entries should be written in Ink.*

*Should with regard to individual persons. The returns are not to be used for proof of age, as in connection with Old Age Pensions, or for any other purpose than the preparation of Statistical Tables.*

Number of Schedules  
(To be filled up by the Enumerator  
after enumeration.)

**NO  
VOTES  
FOR  
WOMEN  
NO  
CENSUS**

PERSONS OR OCCUPATION of Persons aged ten years and upwards.	Whether Employed, "Unemployed", or "Working in their own house."	Whether Employed, "Unemployed", or "Working in their own house."	RESIDENCE of every person.	NATIONALITY of every Person born in a Foreign Country.	DEPARTMENT.
<p><small>The only entries should show the gender, branch of Profession, Trade, Manufacture, &amp;c.</small></p> <p><small>If engaged in any Trade or Manufacture, the particular kind of work done, and the article, such as Material, worked or dealt in, should be clearly indicated.</small></p> <p><small>The Instructions 1 to 8 and Examples at back of Schedule.</small></p>	<p><small>This question should generally be answered by stating the business carried on by the employer. If this is clearly shown in Col. 10, the question need not be answered here.</small></p> <p><small>In cases where the business is carried on by the person himself, the nature of the business should be stated.</small></p> <p><small>The Instructions 9 and Examples at back of Schedule.</small></p>	<p><small>Write opposite the name of each person engaged in any Trade or Industry:</small></p> <p>(1) "Employed" that is, receiving wages or salary from some person or persons.</p> <p>(2) "Unemployed" that is, not receiving wages or salary from any person.</p> <p>(3) "Working in their own house" that is, receiving wages or salary from themselves.</p>	<p>(1) If born in the United Kingdom, write the name of the County, and Town or Parish.</p> <p>(2) If born in any other part of the British Empire, write the name of the Empire, Colony, &amp;c., and of the District or Town.</p> <p>(3) If born in a Foreign Country, write the name of the Country.</p> <p>(4) If born at sea, write "At Sea."</p> <p><small>Note.—In the case of persons born elsewhere than in England or Wales, state whether "Immigrant" or "Alien" in this Column.</small></p>	<p><small>State whether—</small></p> <p>(1) "British subject by birth."</p> <p>(2) "Naturalized British subject."</p> <p>(3) "British subject by marriage."</p> <p>(4) "Alien."</p> <p>(5) "Immigrant."</p> <p>(6) "Alien subject."</p> <p><small>If of foreign nationality, state whether—</small></p> <p>(1) "French."</p> <p>(2) "German."</p> <p>(3) "Italian."</p> <p>(4) "Irish."</p> <p>(5) "Polish."</p> <p>(6) "Portuguese."</p> <p>(7) "Russian."</p> <p>(8) "Spanish."</p> <p>(9) "Swiss."</p> <p>(10) "Other."</p>	<p><small>If any person included in this Schedule is—</small></p> <p>(1) "Totally blind."</p> <p>(2) "Lame and Deaf."</p> <p>(3) "Paralytic."</p> <p>(4) "Insane."</p> <p>(5) "Idiot."</p> <p>(6) "Dumb."</p> <p>(7) "Deaf-blind."</p> <p><small>State the disability opposite that person's name, and the age at which he or she became affected.</small></p>
<p style="font-size: 2em; font-weight: bold;">NO persons here</p> <p style="font-size: 1.5em; font-weight: bold;">only women!</p>					

Write before starting (to meet the requirements) in this column.

**CENSUS MEETING.**  
APRIL 14.  
TRAFALGAR SQUARE,  
9 a.m.

of the Ward of  
I declare it  
to be true.

**CENSUS MEETING.**  
APRIL 14.  
TRAFALGAR SQUARE,  
9 a.m.

Inspector, or in charge of this dwelling:  
filled up to the best of my knowledge and belief.

- 'Time capsule' free text response
- Ireland, 2022

**Absent Person 4**
**Declaration**

**A1 What is this person's name?**  
*First name (BLOCK CAPITALS)*  
  
  
*Surname (BLOCK CAPITALS)*

**A2 What is this person's sex?**  
 1  Male 2  Female

**A3 What is this person's date of birth?**  
 Day  Month  Year

**A4 What is the relationship of this person to Person 1 on page 4?**  
*Mark  one box only*  
 1  Husband or wife  
 2  Partner (incl. same-sex partner)  
 3  Son or daughter  
 9  Other related, write in RELATIONSHIP  
  
 10  Unrelated (incl. foster child)

**A5 What is this person's current marital status?**  
*Answer if aged 16 years or over*  
*Mark  one box only*  
 1  Single (never married or never in a same-sex civil partnership)  
 2  Married (incl. re-married)  
 4  In a registered same-sex civil partnership  
 5  Separated  
 6  Divorced  
 7  Widowed

**A6 How long altogether is this person away for?**  
 1  Less than 12 months  
 2  12 months or more

**A7 Was this person in the Republic of Ireland on Census Night?**  
 1  Yes 2  No

**A8 What is the country of this person's citizenship?**  
*If they have dual country of citizenship, please declare both.*  
 1  Ireland  
 2  Other CITIZENSHIP, write in  
  
  
 3  No citizenship

**Declaration to be completed by the person responsible for completing the form**  
**Before you sign and date the declaration please check:**

- That in List 1 on page 3, you have accounted for all persons (including visitors) who spent Census Night at this address.
- That you have answered all questions which should have been answered for each person who spent Census Night in the household (pages 4 to 21 and any additional Individual Forms).
- That in List 2 on page 3, you have accounted for all persons who usually live at this address but who were temporarily absent on Census Night.
- That you have answered all questions on pages 22 to 23 for all household members temporarily absent on Census Night.


I declare that this form is correct and complete to the best of my knowledge and belief.

Signature  Date

**Time Capsule:**  
 Information you provide in this Time Capsule is optional and is collected voluntarily under Section 24 of the Statistics Act 1993. This content is protected by the same confidentiality protections as all your Census data for 100 years. After 100 years, this Time Capsule will be made available to the public. **This space is for handwritten messages only. Photographs or other attachments will be removed and cannot be returned.**

DO NOT COMPLETE FOR INFORMATION ONLY!

Household Form

  
 EH23

Page 23

## **Acknowledgements**

The permission of the Office for National Statistics to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is funded by the ESRC under project ES/V003488/1. The authors alone are responsible for the interpretation of the data.

This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates