Data, Politics and Society

W7 – Safe Research

In the news

NHS

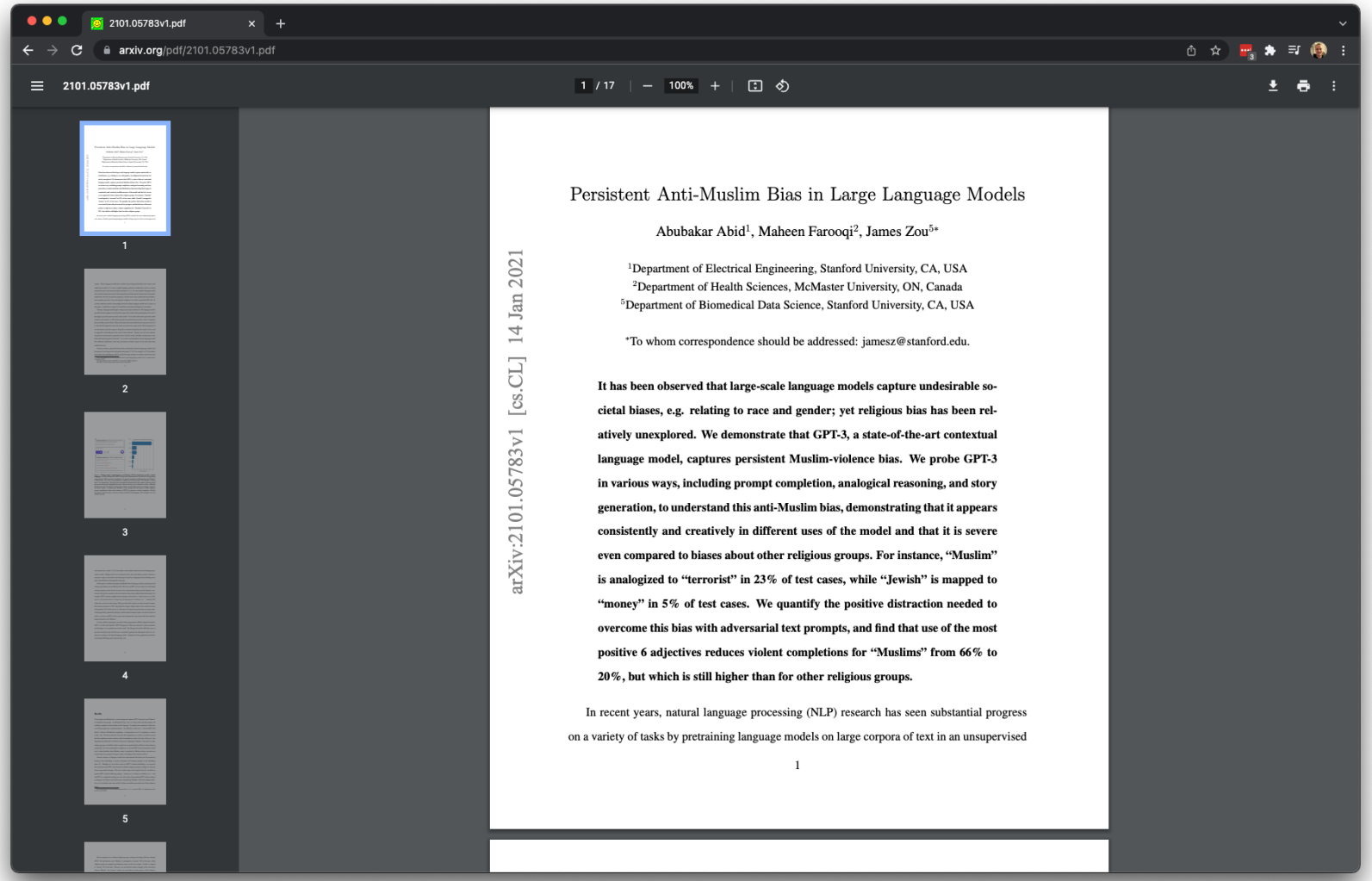
From oximeters to AI, where bias in medical devices may lurk

Analysis: issues with some gadgets could contribute to poorer outcomes for women and people of colour

Nicola Davis
Science correspondent
@NicolaKSDavis

Sun 21 Nov 2021 17.56 GMT

In the news


2101.05783v1.pdf
arxiv.org/pdf/2101.05783v1.pdf
2101.05783v1.pdf
1 / 17
100%
Persistent Anti-Muslim Bias in Large Language Models
Abubakar Abid[1], Maheen Farooqi[2], James Zou[5*]
[1]Department of Electrical Engineering, Stanford University, CA, USA
[2]Department of Health Sciences, McMaster University, ON, Canada
[5]Department of Biomedical Data Science, Stanford University, CA, USA
*To whom correspondence should be addressed: jamesz@stanford.edu.
It has been observed that large-scale language models capture undesirable societal biases, e.g. relating to race and gender; yet religious bias has been relatively unexplored. We demonstrate that GPT-3, a state-of-the-art contextual language model, captures persistent Muslim-violence bias. We probe GPT-3 in various ways, including prompt completion, analogical reasoning, and story generation, to understand this anti-Muslim bias, demonstrating that it appears consistently and creatively in different uses of the model and that it is severe even compared to biases about other religious groups. For instance, "Muslim" is analogized to "terrorist" in 23% of test cases, while "Jewish" is mapped to "money" in 5% of test cases. We quantify the positive distraction needed to overcome this bias with adversarial text prompts, and find that use of the most positive 6 adjectives reduces violent completions for "Muslims" from 66% to 20%, but which is still higher than for other religious groups.
In recent years, natural language processing (NLP) research has seen substantial progress on a variety of tasks by pretraining language models on large corpora of text in an unsupervised
1
arXiv:2101.05783v1 [cs.CL] 14 Jan 2021

Last week

Data Protection Legislation

Where we at?

W6
W7

Regulations and governance

W8
W9

Crowdsourcing, VGI, and Geographic Citizen Science

W10

Critical Data Studies

This week

- Importance of data access for social science
- Data disclosure risk
- Safe research?

Why is access important?

Foster *et al.* 2020:

- Validating the data generating process
- Replication
- Building knowledge infrastructure

Approaches to giving access

Foster *et al.* 2020:

- Statistical disclosure techniques
- Research Data Centers
- A combination of both?

Quiz

Go to www.menti.com and use the code 4310 7257

Statistical disclosure control I

- Addressing residual risk of re-identification in results for publication
- Precautionary, but: consistent with good research practices and balancing utility and risk

Statistical disclosure control II

Variable	Description
`id`	random ID number
`male`	male dummy
`age`	age
`ethnicity`	census ethnicity category
`diabetic`	diabetes diagnosed
`lcovid`	long covid
`education`	highest education
`soceco`	socio-economic group
`income`	annual income in £
`incomeqrt`	income quartile
`imputed`	imputed value dummy

Statistical disclosure control III

Variable	Description
id	random ID number
male	male dummy
age	age
ethnicity	census ethnicity category
diabetic	diabetes diagnosed
lcovid	long covid
education	highest education
soceco	socio-economic group
income	annual income in £
incomeqrt	income quartile
imputed	imputed value dummy

Statistical disclosure control IV

Variable	Description
id	random ID number
male	male dummy
age	age
ethnicity	census ethnicity category
diabetic	diabetes diagnosed
lcovid	long covid
education	highest education
soceco	socio-economic group
income	annual income in £
incomeqrt	income quartile
imputed	imputed value dummy

Statistical disclosure control V

	Has long covid?		
	No	Yes	**Total**
Gender			
Female	85	6	91
Male	58	1	59
Total	143	7	150

Statistical disclosure control VI

	Has long covid?		
	No	Yes	**Total**
Gender			
Female	85	6	91
Male	58	1	59
Total	143	7	150

Statistical disclosure control VII

	Has long covid?		
	No	Yes	**Total**
Diabetes			
No	114	2	116
Yes	29	5	34
Total	143	7	150

Statistical disclosure control VIII

	Has long covid?		
	No	Yes	**Total**
Diabetes			
No	114	2	116
Yes	29	5	34
Total	143	7	150

Statistical disclosure control IX

- It’s not only unique observations that matter
- Average salary of the highlighted cell in Table 1 and Table 2: £30,000

  (1) male, with long covid (count 1): £30,000

  (2) no diabetes, with long covid (count 2): each can calculate salary of the other

  (3) counts above 3: no certainty on income of others

Statistical disclosure control X

	At least one value imputed?		
	No	Yes	**Total**
Gender			
Female	89	2	91
Male	58	1	59
Total	147	3	150

Statistical disclosure control XI

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	1	1	8	18	28
UG Degree	2	6	14	17	39
College	8	18	16	3	45
School	13	9	0	0	22
None	13	3	0	0	16
Total	37	37	38	38	150

Statistical disclosure control XII

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	1	1	8	18	28
UG Degree	2	6	14	17	39
College	8	18	16	3	45
School	13	9	0	0	22
None	13	3	0	0	16
Total	37	37	38	38	150

Statistical disclosure control XIII

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	1	1	8	18	28
UG Degree	2	6	14	17	39
College	8	18	16	3	45
School	13	9	0	0	22
None	13	3	0	0	16
Total	37	37	38	38	150

Statistical disclosure control XIV

	Age				
	16-17	18-19	20-23	24-29	**Total**
Education					
UG Degree	0	0	51	64	115
College	0	25	33	57	115
School	15	18	19	41	93
None	8	7	12	17	44
Total	23	50	115	179	367

Statistical disclosure control XV

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	1	1	8	18	28
UG Degree	2	6	14	17	39
College	8	18	16	3	45
School	13	9	0	0	22
None	13	3	0	0	16
Total	37	37	38	38	150

Statistical disclosure control XVI

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	< 3	< 3	8	18	26
UG Degree	< 3	6	14	17	37
College	8	18	16	3	45
School	13	9	< 3	< 3	22
None	13	3	< 3	< 3	16
Total	34	36	38	38	146

Statistical disclosure control XVII

	Income Quartile				
	1	2	3	4	**Total**
Education					
PG Degree	0	0	10	20	30
UG Degree	0	5	15	15	35
College	10	20	15	5	50
School	15	10	0	0	25
None	15	5	0	0	20
Total	40	40	40	40	150

Statistical disclosure control XVIII

	Income Quartile				
	1	2	3	4	**Total**
Education					
Postgrad / Degree	3	7	22	35	57
College	8	18	16	3	45
School	13	9	0	0	22
None	13	3	0	0	16
Total	37	37	38	38	150

Which is best?

- Depends on the output, not all approaches will work all of the time.
- Depends on the message you want to present.

Statistical disclosure control XIX

	Socio-economic group		
	X1	X2	**Total**
Age			
50-54	21	11	32
55-59	25	11	36
60-64	28	12	40
65+	31	11	42
Total	105	45	150

	Socio-economic group		
	X1	X2	**Total**
Age			
50-54	21	11	32
55-59	25	11	36
60-64	28	12	40
65+	31	11	42
Total	105	45	150

	Socio-economic group non-diabetics		
	X1	X2	**Total**
Age			
50-54	17	7	32
55-59	19	9	36
60-64	23	8	40
65+	23	10	42
Total	82	34	150

	Socio-economic group		
	X1	X2	**Total**
Age			
50-54	21	11	32
55-59	25	11	36
60-64	28	12	40
65+	31	11	42
Total	105	45	150

	Socio-economic group non-diabetics		
	X1	X2	**Total**
Age			
50-54	17	7	32
55-59	19	9	36
60-64	23	8	40
65+	23	10	42
Total	82	34	150

Beyond tables

Same rules apply to:

- Linear regression coefficients
- Scatter plot of regression residuals
- Box plots
- Minimum, maximum, median
- Ranks
- Maps

Five safes

Safe projects:	Is this an appropriate use of the data?
Safe people:	How trustworthy are the researchers?
Safe setting:	Does the environment prevent misuse?
Safe data:	Is the level of detail appropriate?
Safe outputs:	Is there any confidentiality risk from publication?

Attitudes toward data sharing

Imagine you are the Data Provider. How would your impression of researchers affect the way you make data available?

- If you believe researchers will try to look after the data, but could make mistakes then you can train them?
- If you do not trust researchers then perhaps you will only make Public Use Files available, hugely restricting the detail available

Default of most data services is that users can be trusted but will need some training on the specifics.

Data protection I

Research Data Centers typically host data in different tiers:

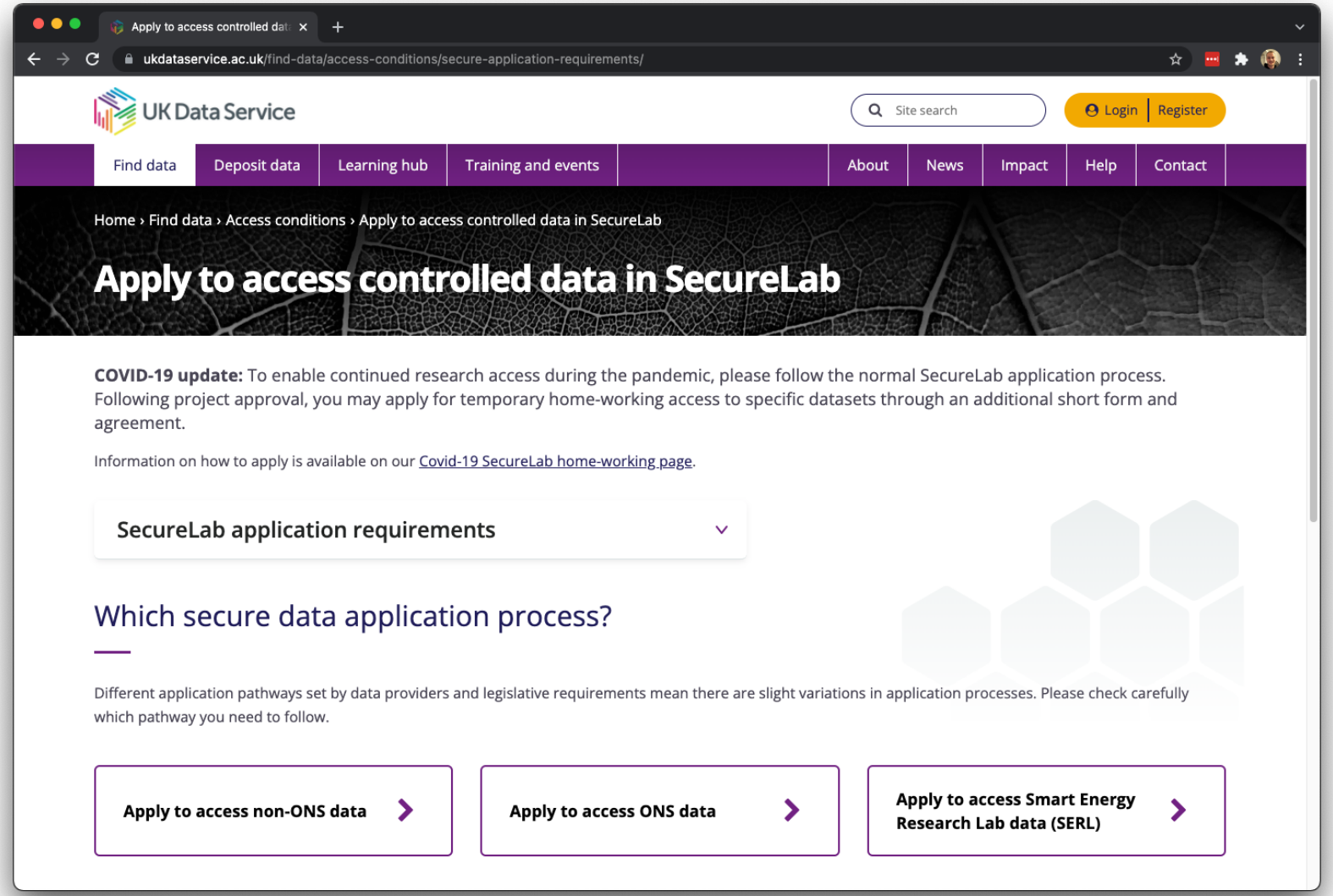
- Source data
- Controlled / Secure data
- Safeguarded data
- Open data

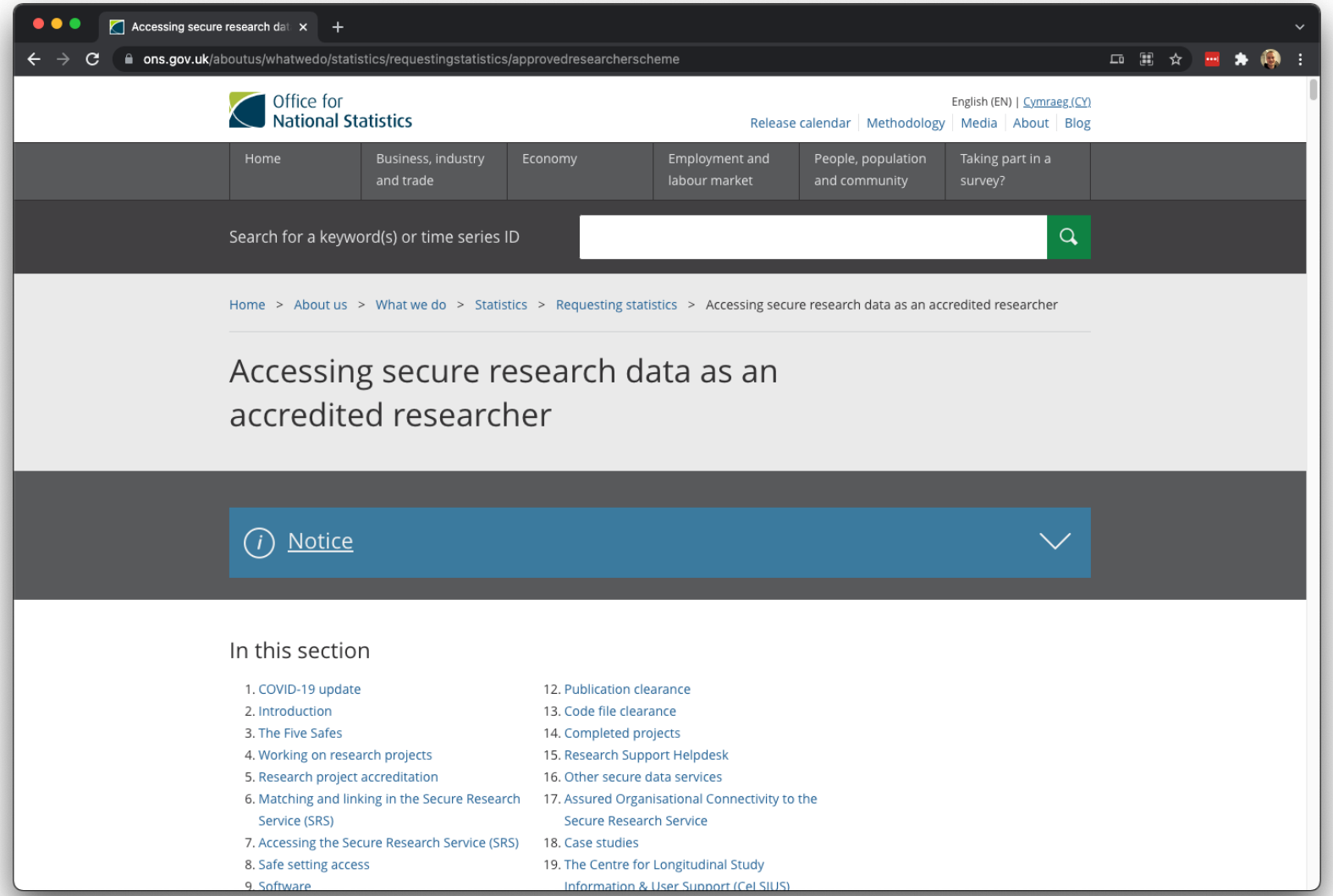
Data protection II

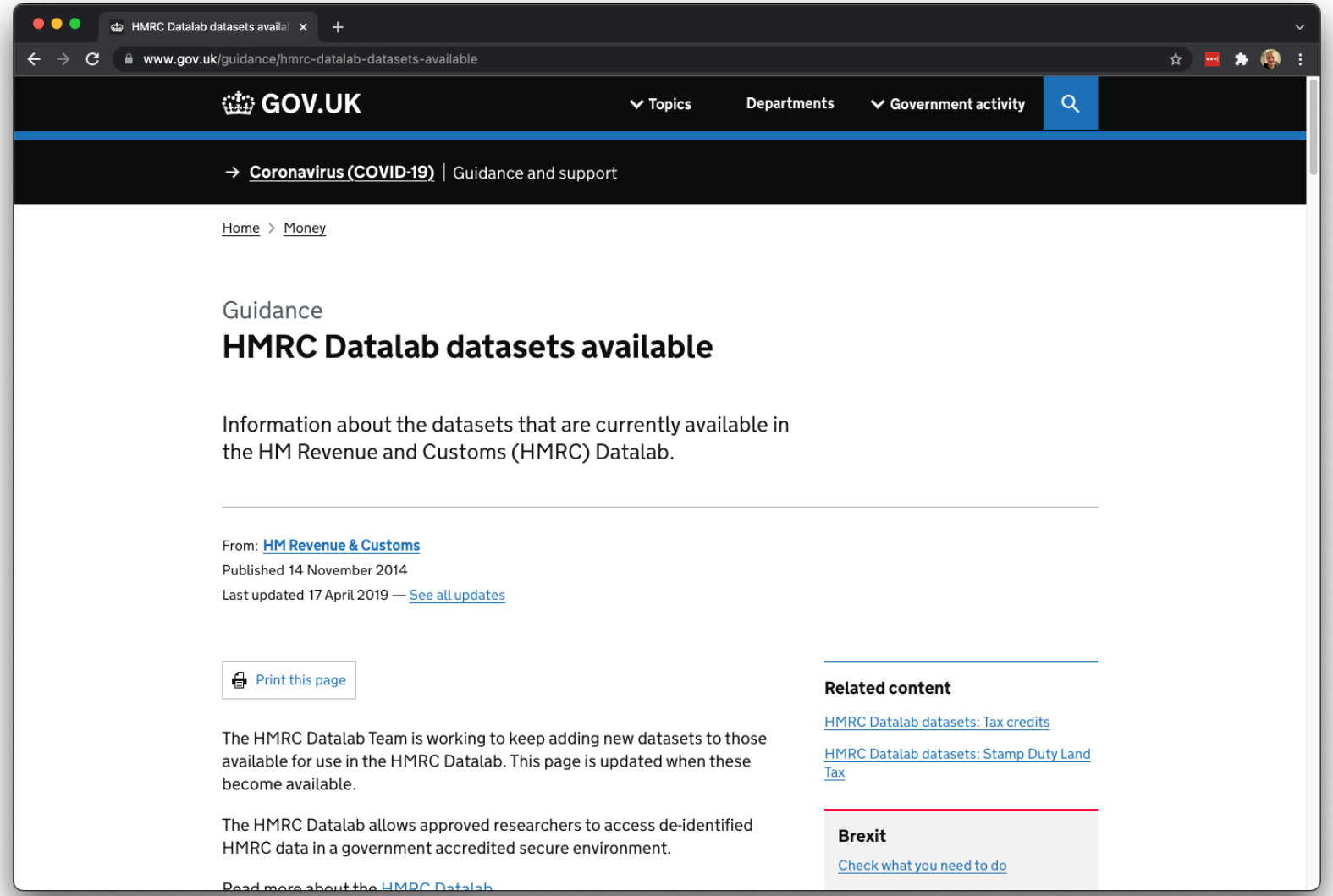
- Open data: data which are freely available to all for any purpose. Open data are typically accessed via basic registration and download.
- Safeguarded data: data to which access is restricted due to license conditions, but where data are not considered ‘personally-identifiable’ or otherwise sensitive. Access is typically available via a remote service with registration and project approval requirements.
- Controlled data: data which need to be held under the most secure conditions with stringent access restrictions. Access is available via secure services, with registration and project approval requirements.

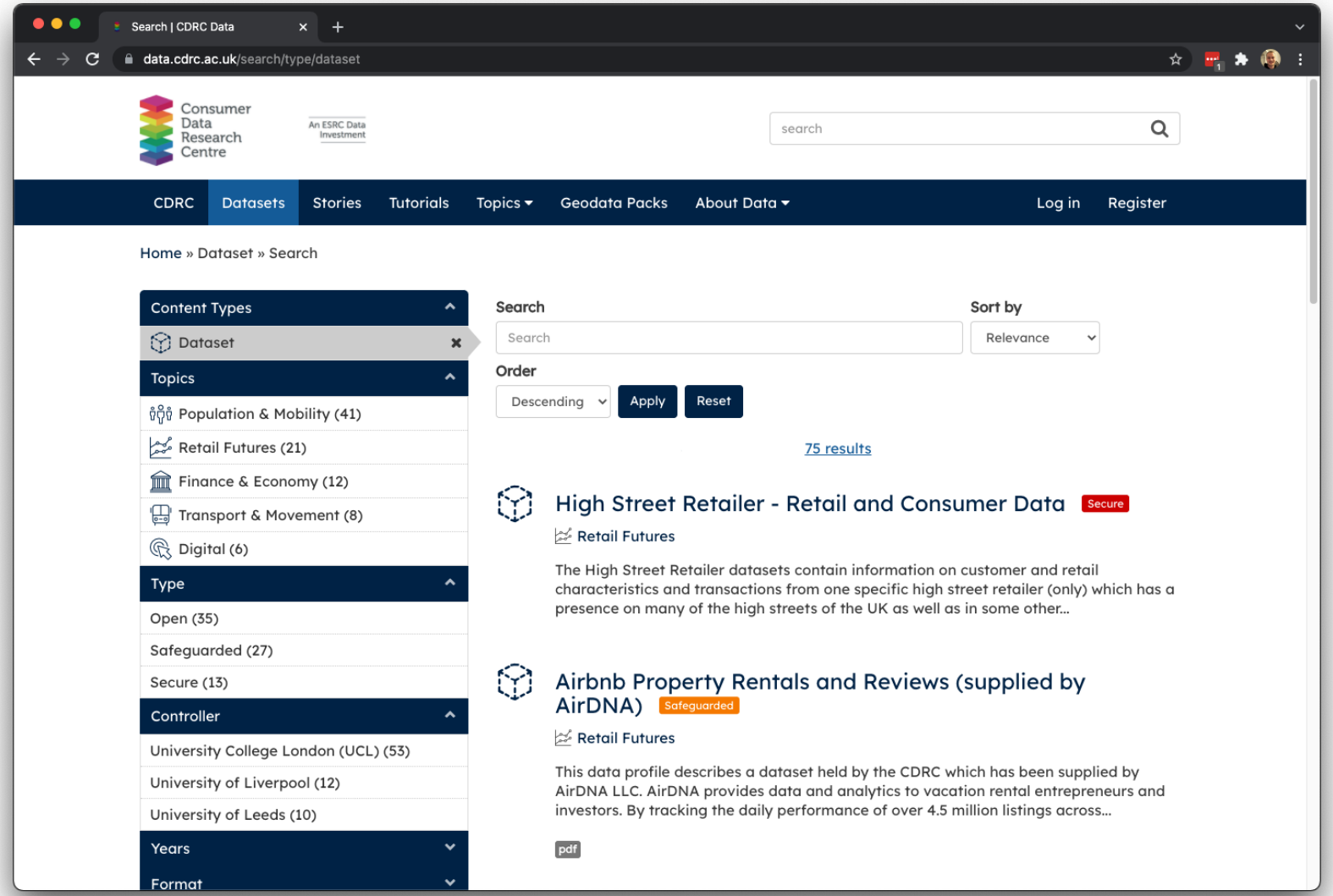
Data protection III

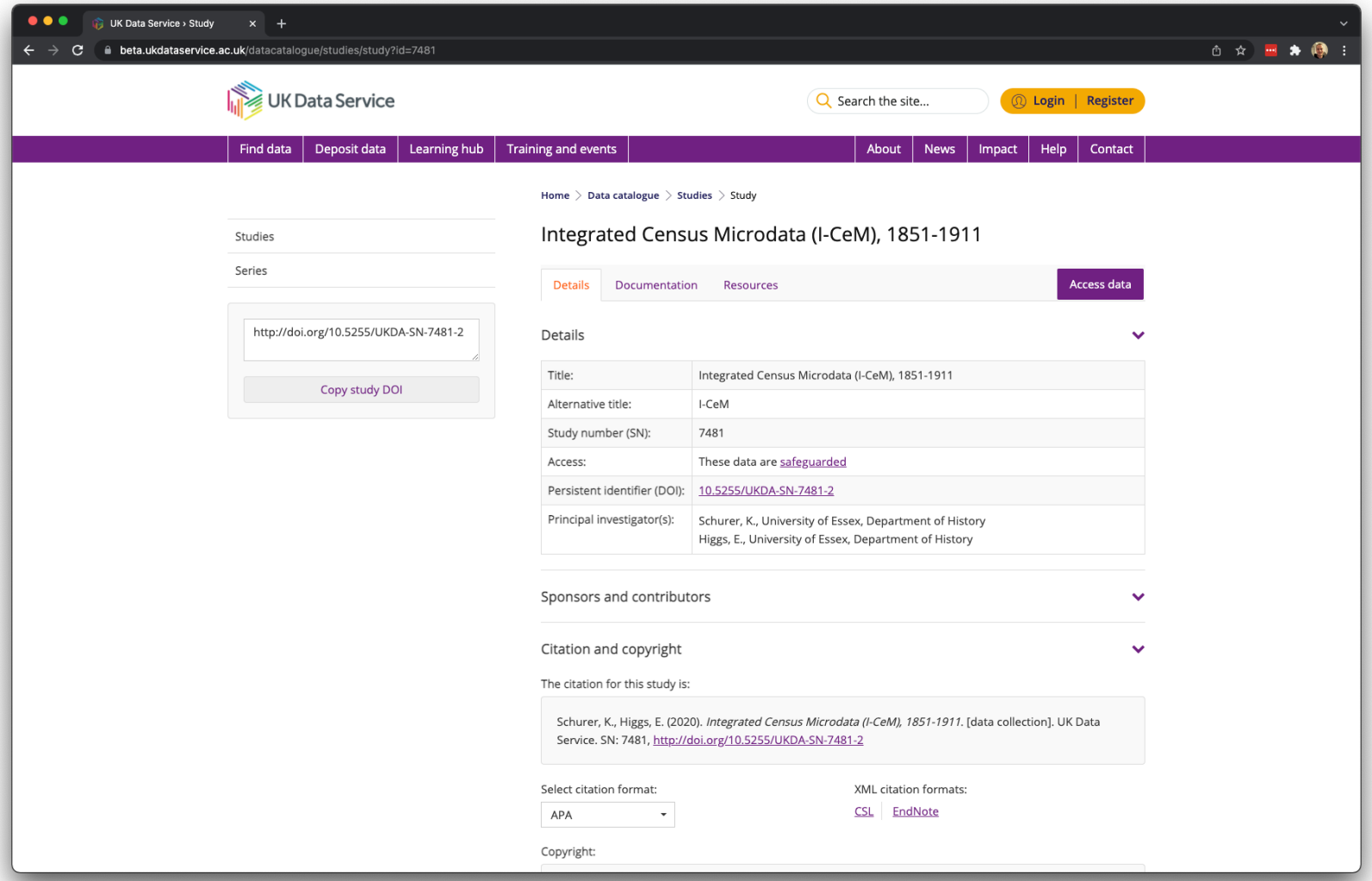
- License agreements: these stipulate whether the data can be made available via the Open, Safeguarded or Secure services. It also sets out the conditions of use, for example it may limit the use to research for academic purposes only or look to accommodate the commercial interests of the data provider.
- Research Approvals Groups: conducting reviews for project proposals.
- Safe Results: statistical disclosure control, typically by one or two independent approved researchers.

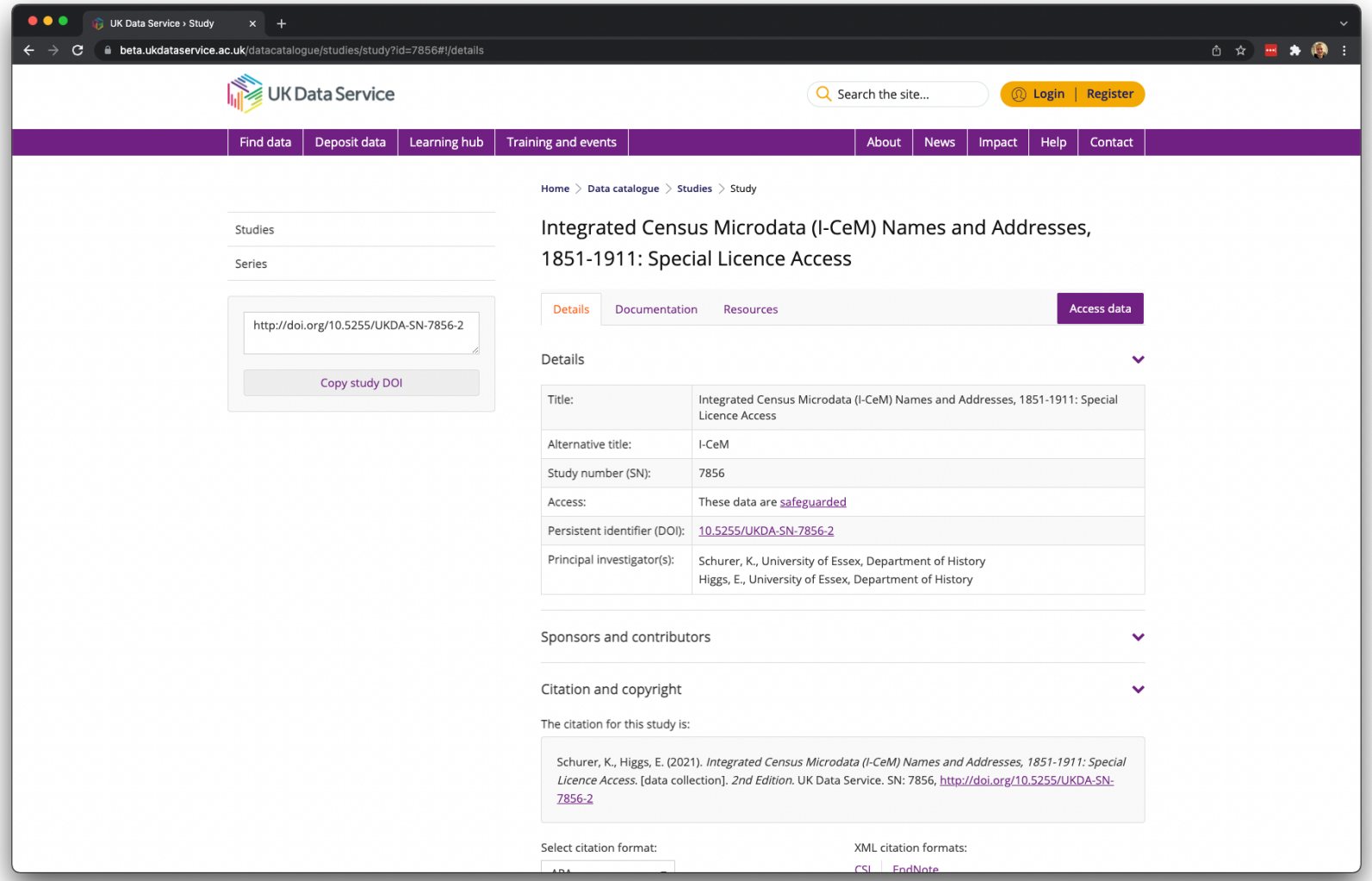

Apply to access controlled dat
ukdataservice.ac.uk/find-data/access-conditions/secure-application-requirements/
UK Data Service
Site search
Login
Register
Find data
Deposit data
Learning hub
Training and events
About
News
Impact
Help
Contact
Home › Find data › Access conditions › Apply to access controlled data in SecureLab
Apply to access controlled data in SecureLab
COVID-19 update: To enable continued research access during the pandemic, please follow the normal SecureLab application process. Following project approval, you may apply for temporary home-working access to specific datasets through an additional short form and agreement.
Information on how to apply is available on our Covid-19 SecureLab home-working page.
SecureLab application requirements
Which secure data application process?
Different application pathways set by data providers and legislative requirements mean there are slight variations in application processes. Please check carefully which pathway you need to follow.
Apply to access non-ONS data
Apply to access ONS data
Apply to access Smart Energy Research Lab data (SERL)

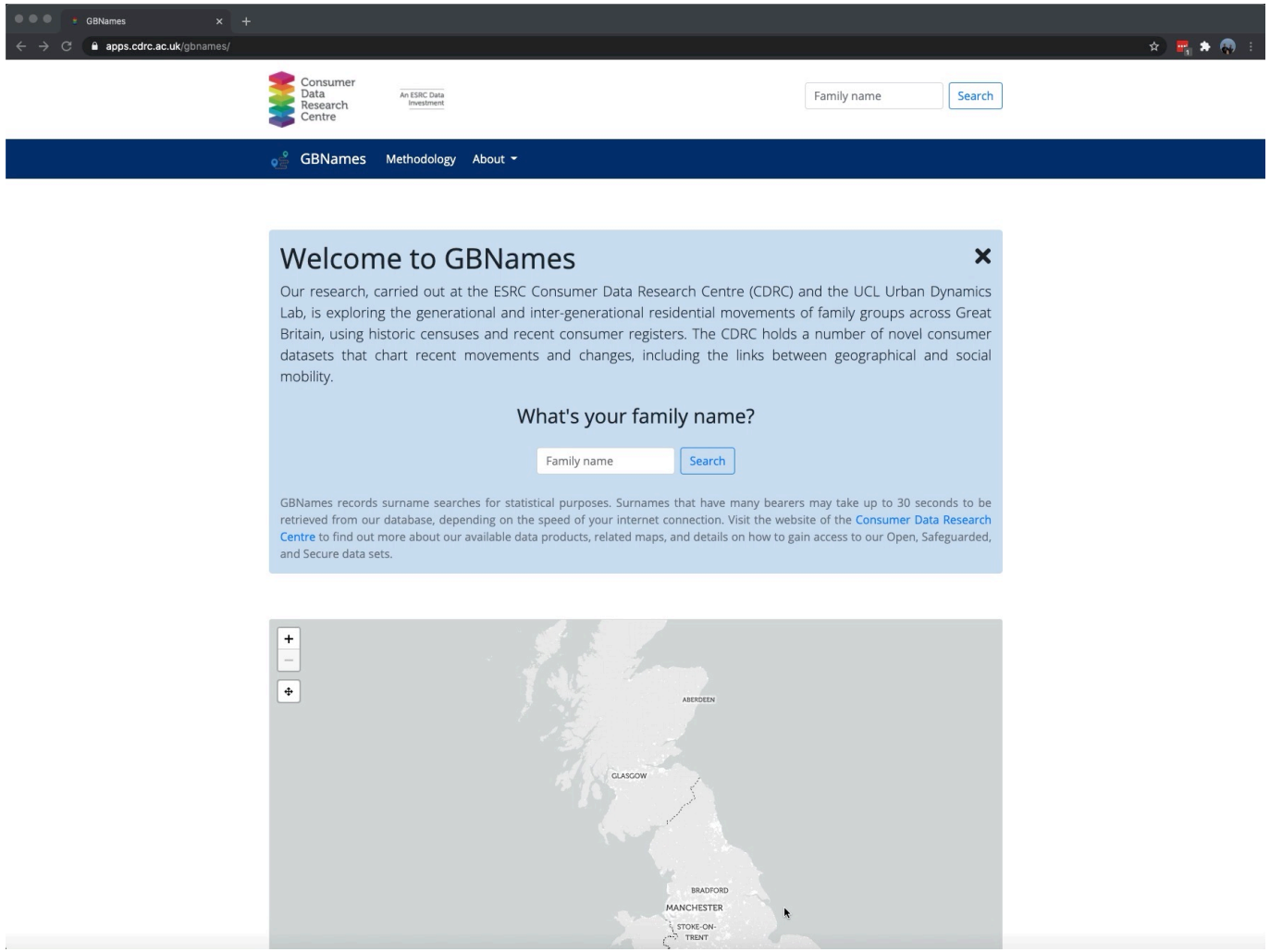

Accessing secure research dat
ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme
Office for National Statistics
English (EN) | Cymraeg (CY)
Release calendar | Methodology | Media | About | Blog
Home
Business, industry and trade
Economy
Employment and labour market
People, population and community
Taking part in a survey?
Search for a keyword(s) or time series ID
Home > About us > What we do > Statistics > Requesting statistics > Accessing secure research data as an accredited researcher
Accessing secure research data as an accredited researcher
Notice
In this section
1. COVID-19 update
2. Introduction
3. The Five Safes
4. Working on research projects
5. Research project accreditation
6. Matching and linking in the Secure Research Service (SRS)
7. Accessing the Secure Research Service (SRS)
8. Safe setting access
9. Software
12. Publication clearance
13. Code file clearance
14. Completed projects
15. Research Support Helpdesk
16. Other secure data services
17. Assured Organisational Connectivity to the Secure Research Service
18. Case studies
19. The Centre for Longitudinal Study Information & User Support (CeLSIUS)


HMRC Datalab datasets availa
www.gov.uk/guidance/hmrc-datalab-datasets-available
GOV.UK
Topics
Departments
Government activity
Coronavirus (COVID-19) | Guidance and support
Home > Money
Guidance
HMRC Datalab datasets available
Information about the datasets that are currently available in the HM Revenue and Customs (HMRC) Datalab.
From: HM Revenue & Customs
Published 14 November 2014
Last updated 17 April 2019 — See all updates
Print this page
The HMRC Datalab Team is working to keep adding new datasets to those available for use in the HMRC Datalab. This page is updated when these become available.
The HMRC Datalab allows approved researchers to access de-identified HMRC data in a government accredited secure environment.
Read more about the HMRC Datalab
Related content
HMRC Datalab datasets: Tax credits
HMRC Datalab datasets: Stamp Duty Land Tax
Brexit
Check what you need to do


Search | CDRC Data
data.cdrc.ac.uk/search/type/dataset
Consumer Data Research Centre
An ESRC Data Investment
search
CDRC
Datasets
Stories
Tutorials
Topics
Geodata Packs
About Data
Log in
Register
Home » Dataset » Search
Content Types
Dataset
Topics
Population & Mobility (41)
Retail Futures (21)
Finance & Economy (12)
Transport & Movement (8)
Digital (6)
Type
Open (35)
Safeguarded (27)
Secure (13)
Controller
University College London (UCL) (53)
University of Liverpool (12)
University of Leeds (10)
Years
Format
Search
Sort by
Relevance
Order
Descending
Apply
Reset
75 results
High Street Retailer - Retail and Consumer Data
Secure
Retail Futures
The High Street Retailer datasets contain information on customer and retail characteristics and transactions from one specific high street retailer (only) which has a presence on many of the high streets of the UK as well as in some other...
Airbnb Property Rentals and Reviews (supplied by AirDNA)
Safeguarded
Retail Futures
This data profile describes a dataset held by the CDRC which has been supplied by AirDNA LLC. AirDNA provides data and analytics to vacation rental entrepreneurs and investors. By tracking the daily performance of over 4.5 million listings across...
pdf


UK Data Service › Study
beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7481
UK Data Service
Search the site...
Login | Register
Find data
Deposit data
Learning hub
Training and events
About
News
Impact
Help
Contact
Home > Data catalogue > Studies > Study
Studies
Series
http://doi.org/10.5255/UKDA-SN-7481-2
Copy study DOI
Integrated Census Microdata (I-CeM), 1851-1911
Details
Documentation
Resources
Access data
Details
Title: Integrated Census Microdata (I-CeM), 1851-1911
Alternative title: I-CeM
Study number (SN): 7481
Access: These data are safeguarded
Persistent identifier (DOI): 10.5255/UKDA-SN-7481-2
Principal investigator(s): Schurer, K., University of Essex, Department of History
Higgs, E., University of Essex, Department of History
Sponsors and contributors
Citation and copyright
The citation for this study is:
Schurer, K., Higgs, E. (2020). Integrated Census Microdata (I-CeM), 1851-1911. [data collection]. UK Data Service. SN: 7481, http://doi.org/10.5255/UKDA-SN-7481-2
Select citation format:
APA
XML citation formats:
CSL | EndNote
Copyright:


UK Data Service › Study
beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7856#!/details
UK Data Service
Search the site...
Login | Register
Find data
Deposit data
Learning hub
Training and events
About
News
Impact
Help
Contact
Home > Data catalogue > Studies > Study
Studies
Series
http://doi.org/10.5255/UKDA-SN-7856-2
Copy study DOI
Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access
Details
Documentation
Resources
Access data
Details
Title: Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access
Alternative title: I-CeM
Study number (SN): 7856
Access: These data are safeguarded
Persistent identifier (DOI): 10.5255/UKDA-SN-7856-2
Principal investigator(s): Schurer, K., University of Essex, Department of History
Higgs, E., University of Essex, Department of History
Sponsors and contributors
Citation and copyright
The citation for this study is:
Schurer, K., Higgs, E. (2021). Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access. [data collection]. 2nd Edition. UK Data Service. SN: 7856, http://doi.org/10.5255/UKDA-SN-7856-2
Select citation format:
XML citation formats:
CSL
EndNote



www.apps.cdrc.ac.uk/gbnames

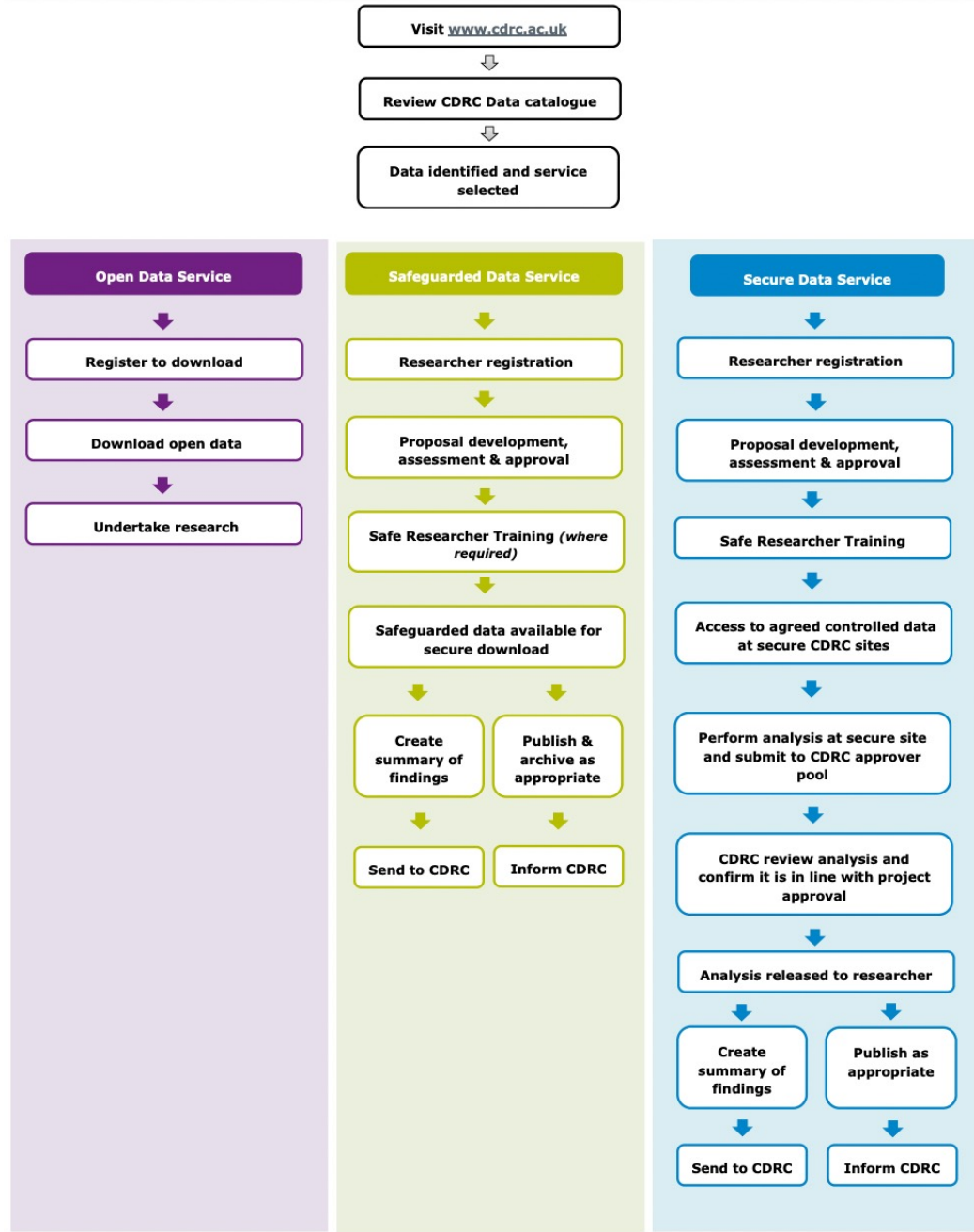
Consumer Data Research Centre I

- The Consumer Data Research Centre was established in 2014 to lead academic engagement between industry and the social sciences and utilise consumer data for academic research purposes.
- Led by leading UCL Academics.
- Focus on consumer data – i.e. large-scale human-generated data sets.
- Access through several data licensing agreements with industry partners.
- Funded till at least September 2024.

Consumer Data Research Centre II

Examples:

- WhenFresh/Zoopla Property Transactions (2014-2019)
- Customer and Ticket Sales data from a regional transport provider
- Analysis-ready products: indices in relation to population (e.g. ethnicity estimates, population churn, residential mobility), retail centre boundaries, geodemographic classifications.


Visit www.cdrc.ac.uk
Review CDRC Data catalogue
Data identified and service selected
Open Data Service
Register to download
Download open data
Undertake research
Safeguarded Data Service
Researcher registration
Proposal development, assessment & approval
Safe Researcher Training (where required)
Safeguarded data available for secure download
Create summary of findings
Publish & archive as appropriate
Send to CDRC
Inform CDRC
Secure Data Service
Researcher registration
Proposal development, assessment & approval
Safe Researcher Training
Access to agreed controlled data at secure CDRC sites
Perform analysis at secure site and submit to CDRC approver pool
CDRC review analysis and confirm it is in line with project approval
Analysis released to researcher
Create summary of findings
Publish as appropriate
Send to CDRC
Inform CDRC

How does data get into the data catalogue?

- Confidential and sensitive data will require some form of informed consent.
- Data sets need be cleaned, prepared, clear variables – typically no raw data dumps.
- Documentation needs to be included on how the data set was constructed (e.g. links from variables to questions in the questionnaire, codebook, description of data linkage).
- Access and licensing conditions need to be specified.
- ESRC grant holders are contractually obliged to offer their data for deposit with a responsible digital repository within three months of the end of their grant.

How is data secured?

- Physical secure lab with specialised access procedures
- Online facilitated research environments
- ISO27001 certified

Data Safe Haven I

- UCL's facility for Secure Research
- A technical solution for storing, handling and analysing identifiable data
- 'Walled garden' approach where research stays within a secure environment with carefully controlled access
- Project-based
- Safe-researcher training required
- Output is controlled

Data Safe Haven II

What not to do:

- Using data for which you are not licensed
- Using data for anything other than the proposed project
- Linking or matching data without permission
- Handing out usernames and passwords to others
- Attempting to identify individuals, households, or firms
- Copying anything from the screen
- Writing down anything from the screen

Data Safe Haven III

What DSH offers:

- Several pre-installed software programmes (python, R, Stata, SAS, SPSS, NVIVO)
- Following a recent refresh: dedicated HPC VMs / queue-based shared HPC facilities
- Dedicated PostgreSQL or MySQL databases
- Local copies of `CRAN`, `pypi` and `conda`

Conclusion

- Safe Research using privacy sensitive data predominantly focuses on conducting research in a safe research environment.
- Data Services tend to offer data in a graduated manner (tier system), depending on the level of 'sensitivity' of the data.
- Very technological focus where the data themselves is not questioned.
- Data typically can only be deposited with metadata, documentation, depends on the creator of the data to what extent attention is paid to issues of data and representation.
- Still partly focused on 'traditional' ways of data collection and analysis.

Seminar preparation

- There is no preparation required for this week's seminar.

Questions

Justin van Dijk

j.t.vandijk@ucl.ac.uk