UK Data Service

Trusted data for research

Annual Report

"The UK Data Service has become more than the principal social science digital repository in the UK. Trusted experts in data literacy, data impact, ethics, governance and the curation and active preservation of research data for long-term use, we are passionate about the services we provide and collaborations we nurture.

Facilitators of evidence-led knowledge, we continue to find new ways to enhance our services and provide more value to users. Our pioneering organisation endeavours to enhance learning, research and policy development for a stronger society."

GEMMA HAKINS
Director of Communications and Engagement,
UK Data Service

Contents

Overview

Welcome to the UK Data Service Annual Report for 1 April 2022 to 31 March 2023.

We are proud of our achievements and are pleased to demonstrate the breadth, depth and expertise of our services.

We are at an inflection point as we near the end of our current award and our key funder, the Economic and Social Research Council (ESRC) leads work to define Future Data Services. Increasingly, the rise in user demand for data services and the challenges of securely sharing data at scale are testing traditional data management practices, whilst the scope of our aspirations extends to procure maximum value from data with Artificial Intelligence (AI) and machine learning.

On the European stage, the drive for interoperability, as manifested in the many European Open Science Cloud and Findable, Accessible, Interoperable and Reusable (FAIR) data projects, means it has become necessary to re-think and recalibrate how we curate, manage, disseminate and re-use data.

As we anticipate and position ourselves to deliver quality services in the exciting future, we continue to lead and share best practice across the national and international research landscape as pioneers who have shaped and delivered active data preservation services for nearly 60 years.

> **Together, we have more experience than any other comparable group in the world in delivering quality data services that support and enable economic and social research.**

DARREN BELL
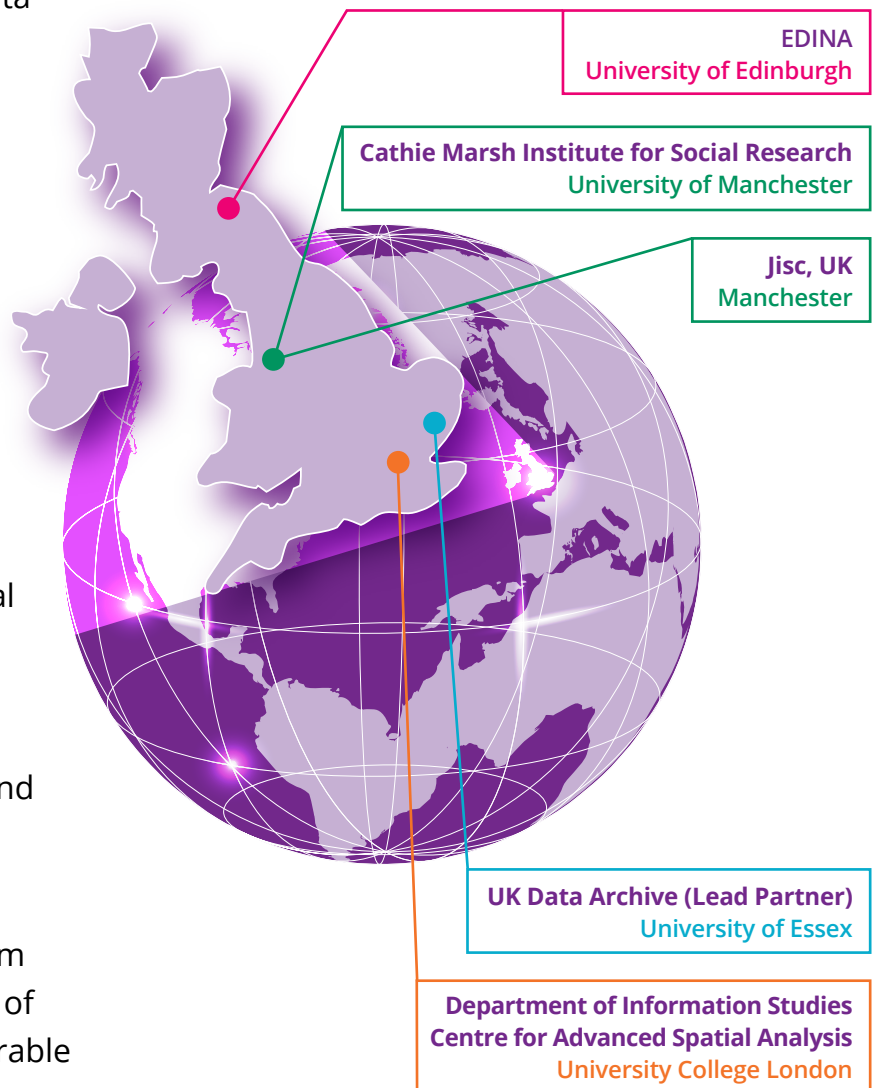Director of Technical Services,
UK Data Service

The AI revolution is in its early stage and will change research in ways which cannot yet be predicted. For example, machine learning is already allowing us to enrich our data, so we can further support both automation and federated access in real-world scenarios. Automation of data application processes will significantly reduce time for researchers to access data. Machines can now create richer, more accurate metadata – and much more quickly – than humans, and this means better discoverability for researchers and better trust and privacy guarantees for our data depositors in the future.

The UK Data Service is a collaborative partnership between the UK Data Archive (UKDA) at the University of Essex, Jisc, the Cathie Marsh Institute (CMI) at the University of Manchester, EDINA at the University of Edinburgh and UCL in London. Together, we have more experience than any other comparable group in the world in delivering quality data services that support and enable economic and social research. We continue to maintain our position as the UK's largest repository of economic and social research data, augmented by world-leading training and support functions.

As we move into the new ESRC award from April 2024 to 2030, the primary aim of the Service will remain the provision of uncomplicated access to easily discoverable and relevant data to enable quality social and economic research. Even after research infrastructures and data toolchains go through a significant step-change, all activities will continue to stem from this core aim.

Now and in the future, the users of our services, whether data consumers or data providers, will always be the beating heart of our operation.

EDINA
University of Edinburgh

Cathie Marsh Institute for Social Research
University of Manchester

Jisc, UK
Manchester

UK Data Archive (Lead Partner)
University of Essex

Department of Information Studies
Centre for Advanced Spatial Analysis
University College London

Vision and strategic objectives

Our vision is for economic, population, and social research data to drive an innovative and thriving research and policy-making ecosystem, leading to enhanced knowledge, better decisions and improved outcomes in society.

Trusted infrastructure

- To be recognised as a trusted national digital repository for data owners and users.
- To innovate and lead in data infrastructure provision, through the development of common standards and agreed strategies for data deposit, preparation processing, documentation and preservation.

Training and user support

To provide training and support services for depositing, finding, accessing, using and managing data.

Our mission is to catalyse impactful social science research by providing access to high-quality, curated and trustworthy datasets of national strategic importance; fostering data literacy; and contributing to the advancement of knowledge through dynamic data infrastructure and strategic partnerships.

Data access

- To provide all service users with a single point of access, and with transparent application and registration processes, to quality-assured economic and social research data.

- To enhance our collection of datasets available for research and impact.
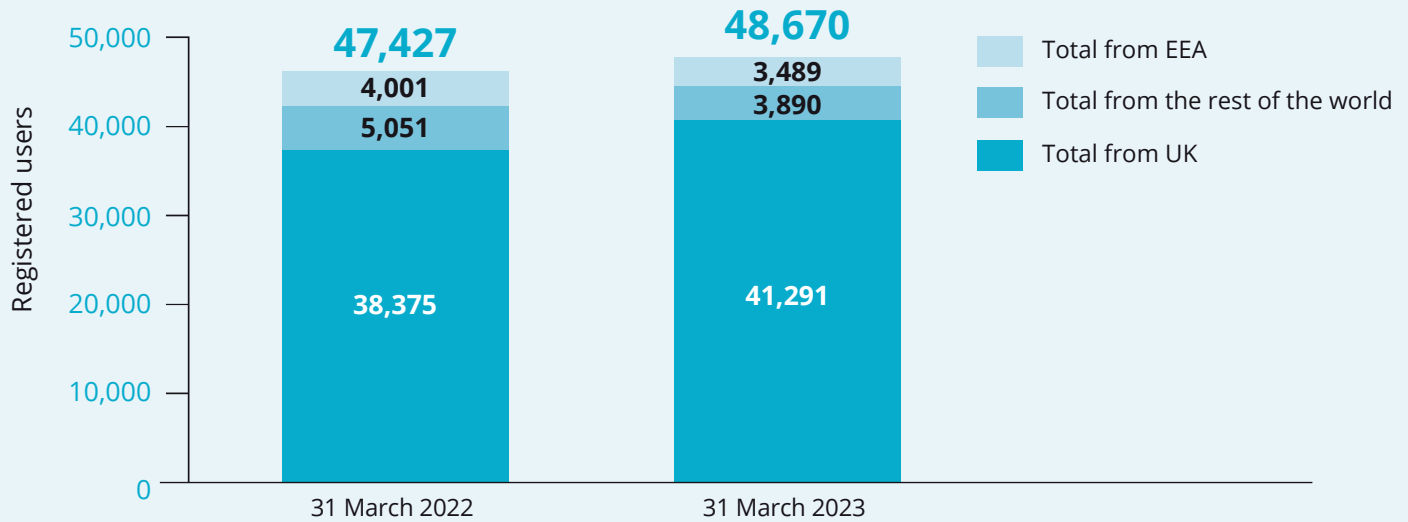
Data impact

- To support research funders in their data-related activities and influence relevant data policy development.

- To lead in the discussions on research integrity and open research, ensuring that data can be used in an ethical and legal way.

- To extend use of our data to the widest possible academic, policy and practitioner communities for generating greater impact.

UK Data Service statistics
1 April 2021 - 31 March 2023

User profile and reach

Registered users

47,427
- Total from EEA: 4,001
- Total from the rest of the world: 5,051
- Total from UK: 38,375

48,670
- Total from EEA: 3,489
- Total from the rest of the world: 3,890
- Total from UK: 41,291

31 March 2022 — 31 March 2023

Downloads across our data services

Total downloads

219,138
- 663
- I-CeM: 16,725
- Nesstar: 16,892
- International: 19,615
- Curated datasets: 80,845
- Census: 84,398

219,862
- 1,091
- I-CeM: 21,731
- Nesstar: 12,073
- International: 17,269
- Curated datasets: 81,166
- Census: 86,532

01 April 2021 – 31 March 2022 — 1 April 2022 – 31 March 2023

I-CeM
The Integrated Census Microdata (I-CeM) is a collection of individual-level census data for Great Britain covering the period 1851 to 1911 – some 185 million person records in total

ReShare
These datasets consist of individual new researcher deposits

Nesstar
The Nesstar data library enables the widest data use without specialist software

International

Curated datasets
(safeguarded, controlled, open)

Census

Source: UK Data Service usage statistics

Curated dataset downloads by region

81,166

Curated datasets downloaded

100,000

75,000

50,000

25,000

0

7,081
7,841
66,244

1 April 2022 – 31 March 2023

Total from the EEA

Total from the rest of the world

Total from the UK
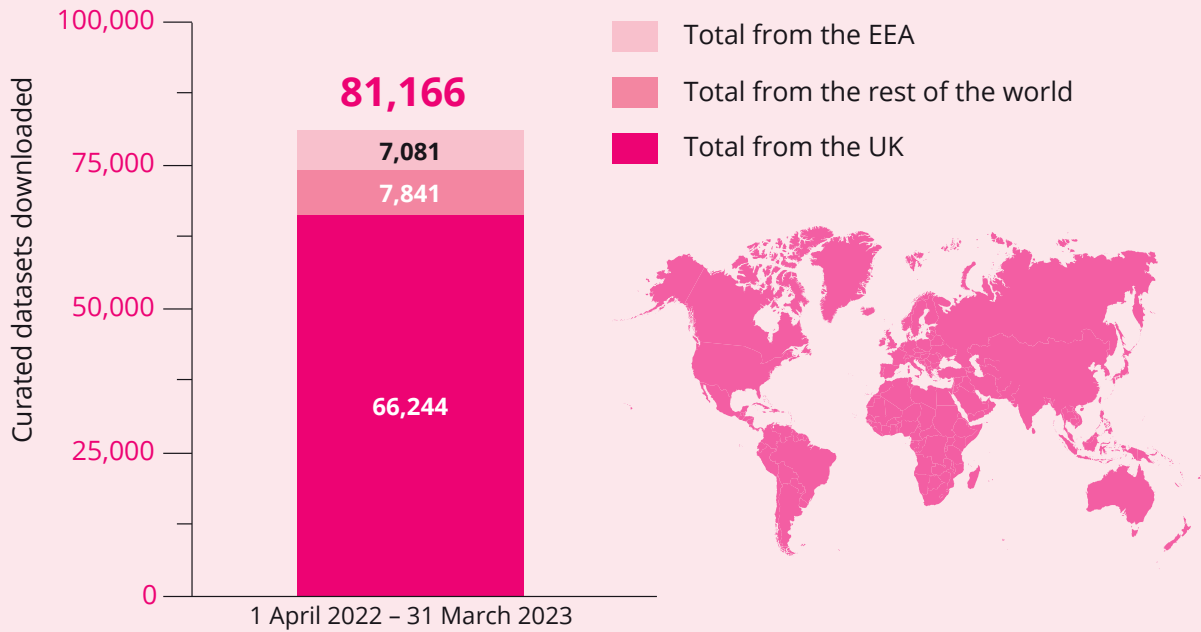
Curated dataset downloads by access type

Curated datasets downloaded

100,000

80,000

60,000

40,000

20,000

0

80,845
3,236
3,481
74,128

01 April 2021 – 31 March 2022

81,166
3,392
3,811
73,963

1 April 2022 – 31 March 2023

Open UKDS
These datasets are open and accessible to all

Controlled
Confidential or sensitive information requiring a high level of security and privacy protection

Safeguarded data downloads
Data licensed for use in the 'safeguarded' category are not 'personal data', but data owners consider to be a risk of disclosure resulting from linkage to other data, such as private databases
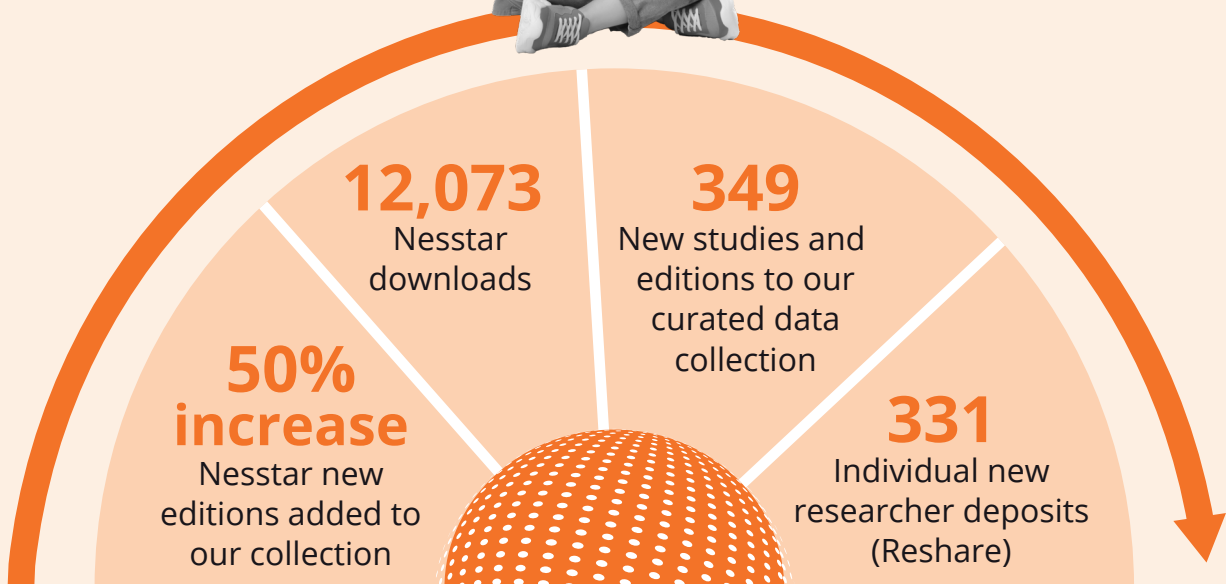
Data curation and acquisition

We continue to add important datasets to our collection all the time, which can be used for public benefit to improve lives.

The Nesstar data library enables the widest data use without specialist software. Researchers search, browse, visualise, explore, analyse and download open and safeguarded datasets; from frequencies to cross tabs and regression.

ReShare is the UK Data Service's online data repository, where researchers can archive, publish and share research data, as open or safeguarded data.

12,073
Nesstar downloads

349
New studies and editions to our curated data collection

50% increase
Nesstar new editions added to our collection

331
Individual new researcher deposits (Reshare)

Source: UK Data Service usage statistics as of 31 March 2023

Pushing the boundaries of infrastructure and technology

Since its inception in 1967 the UK Data Service has developed a strong reputation for building a robust infrastructure for researchers, policy makers and other data users to access data collections from all over the world that drive life-changing research and influence Government policy.

Working with our key partners across the UK Data Service and beyond, we constantly strive to improve the experience of researchers and data users, by leveraging technology further, ensuring work can be carried out even more efficiently.

During this financial year, the team has consulted with partners in Europe and across the world, to ensure the Service is using fit-for-purpose technology and that it continues to provide value and quality.

Many of these plans are still in development, but the discoveries so far will bring long-term benefits to data users, making sure their experience of accessing and using our data collections will be better than ever.

Ambitions for cloud migration

Progress towards our cloud-based approach for all infrastructure has continued well. The Open Clouds for Research Environments (OCRE) framework agreement with Jisc has enabled us to explore new technologies that could make this happen. Transitioning our collections to a cloud-based system will streamline the information retrieval process for researchers, significantly reducing the time required to locate necessary data.

The work the UK Data Service undertook with the Smart Energy Research Lab (SERL) project with UCL between 2017 and 2022 has also helped to speed up the move to a cloud-based system. We continue to work closely with SERL to provide access to these data for research purposes.

Innovative technology

Our new Data Product Builder tool, created by the UK Data Service during this period, is potentially game-changing technology. The tool will significantly improve the way researchers access data by offering a database-driven repository that makes it easier for them to analyse data needed for their project.

We showcased this at the IASSIST 2022 conference (International Association for Social Science Information Service and Technology) in Gothenburg.

Using machine learning to tag data more efficiently and improve the metadata attached to each file, our Product Builder tool significantly enhances the data discovery and retrieval process. This innovative feature ensures a more intuitive and user-friendly experience, allowing researchers to focus more on analysis and less on the preliminary data gathering.

As we continue to refine and expand the tool's capabilities, we are optimistic about its role in significantly advancing our data initiatives and supporting the broader research community in achieving their objectives more efficiently.

Strategic alliances

We now have senior representation on both the Technical Committee and the Scientific Board of the DDI Alliance (Data Documentation Initiative). We are helping to lead the development and implementation of new social science data standards, including DDI-CDI (Cross-Domain Integration).

DDI is an international standard for describing the data produced by surveys and other observational methods in the social, behavioural, economic and health sciences. It is a free standard that can document and manage different stages in the research data lifecycle, such as conceptualisation, collection, processing, distribution, discovery and archiving. Documenting data with DDI facilitates understanding, interpretation and use by people, software systems and computer networks.

Our involvement with this new standard has enabled us to participate in international initiatives like the WorldFAIR's project CODATA (Committee On Data International Science Council).

> "Seeing the data used widely is a key aim – whether through further research that influences policy or demonstrating community benefits through assured evidence. Publishing our datasets via the UK Data Service will provide greater value for researchers and policymakers across the UK."

DR TOM MARSHALL
Data provider and Project Senior Responsible Officer,
People and Nature Survey

DDI-CDI facilitates the combination of data from diverse sources and across disciplines. CODATA exists to promote global collaboration, to improve the usability of data for all areas of research.

We have also designed and built a solution for hosting DDI Controlled Vocabularies to build on our FAIR, interoperable infrastructure.

In other international developments, we have played a key role in the development of the eRImote project, which is an EU-funded scheme that brings together 12 international partners. The project aims to identify solutions for digital and remote service provision across research infrastructure domains, plus identify transferable

practices and new developments for improving accessibility and resilience. We are leading on the delivery of one of the six expert groups being convened by eRImote partners, focusing on "Data sharing, data access and security."

> " Increasing the reusability of data across the digital preservation landscape is not only dependent on the availability of adequate methods, but also their uptake."

HERVÉ L'HOURS
Repository and
Preservation Manager,
UK Data Service
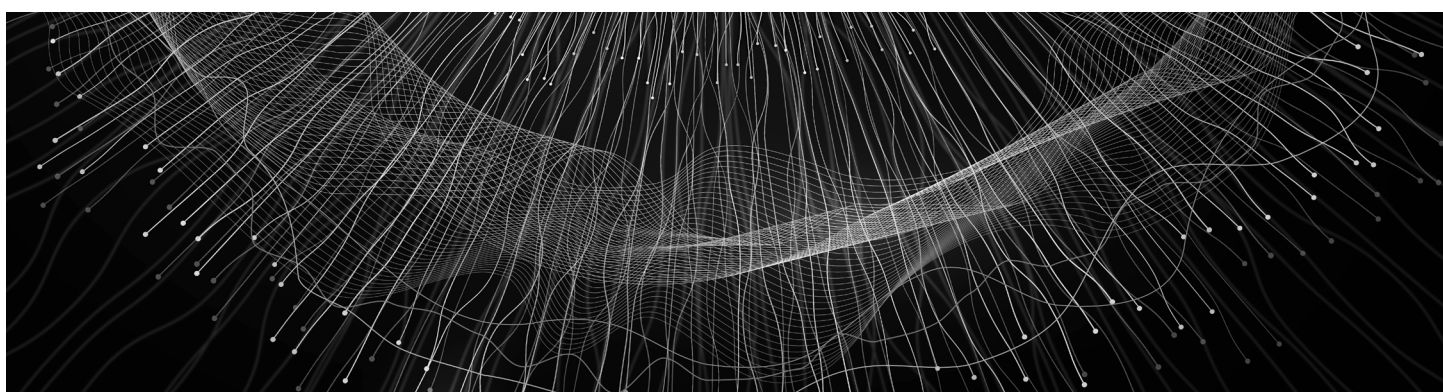
Improved machine learning

Working in collaboration with CLOSER to re-process the entire data collection to DDI Lifecycle, we are now using machine learning to automate the annotation of metadata and significantly reduce human-mediated curation times.

CLOSER, based at the UCL Social Research Institute, is another ESRC funded project, and is the home of longitudinal research. It is also the

interdisciplinary partnership of leading social and biomedical longitudinal population studies, the UK Data Service and the British Library.

DDI Lifecycle is a machine-readable format that will enable further automation of the curation pipeline and support the development of the new UKDS catalogue, which is due to launch in 2024.

User Support and Training continues to grow

We are delighted to report our data skills training attendance has continued to grow long after life has returned to normal following the pandemic.

Interest has been building constantly in our comprehensive training courses that cover a wide variety of foundational level data skills, including "Research Data Management Skills", "The Ethics of Using Data", "How to Become a Computational Social Scientist" and "Getting Started with Secondary Analysis".

Through greater communications and listening to user feedback, we have grown our attendance at these events, particularly by postgraduate students, staff from higher education and people from outside of academia.

Some of the highlights include:

- Training on the new Census 2021 data
- Training on Research Data Management
- Hosting Data User Conferences
- Increased interest in new Computational Social Science topics
- The development of new resources for students and lecturers
- Collaborative training produced with other ESRC investments, as well as government departments

"
There has been so much fantastic work this year. The major highlight for me has been the volume and variety of training events that we have run, as well as all the presentations at conferences and the development of more asynchronous training resources.

We play a very important central role in the development of foundation level data skills; reaching thousands of people each month via our live events and views on YouTube, and receiving consistently positive feedback. "

DR VANESSA HIGGINS
Director User Support and Training,
UK Data Service

"There is international demand for an increase in data skills and these two articles highlight the importance of the UK Data Service data training and other data literacy programmes. The articles provide case studies of success to be shared with other countries and other disciplines.**"**

DR VANESSA HIGGINS
Director User Support and Training, UK Data Service

Dr Vanessa Higgins, Director of User Support and Training at UK Data Service published three papers, one using Health Survey for England and Census data to examine the effects of ethnic density and area deprivation on central obesity in England. Another, published in IASSIST Quarterly that explored how data services and data fellowships play a significant role in creating data-skilled social researchers. The third, published In Mathematics, with collaborators in Colombia, explored how UK models for improving statistical literacy could present opportunities for Colombia.

Increased training reach

5,290

5,792

9% increase

Attendees April 2021 – March 2022

Attendees April 2022 – March 2023
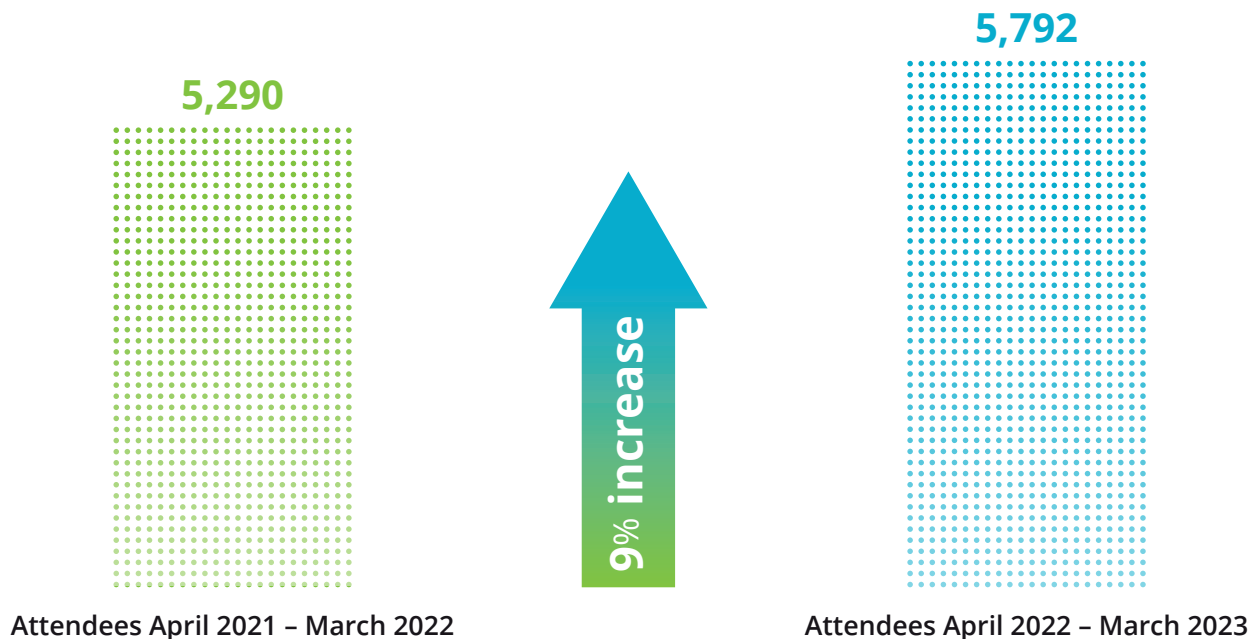
Source: User Support and Training team, UK Data Service

Training event highlights

107 training events

5,792 attended training events

88,042 online YouTube training visitors

Source: User Support and Training team, UK Data Service

In total, from 1 April 2022 to 31 March 2023, we delivered 107 training events, with 5792 recorded attendees. We gave a further 53 presentations at training events/conferences hosted by other organisations and partners.

This includes our annual series of online foundation level workshops, which have been very popular. These sessions cover a wide range of areas, such as:

- What the UK Data Service offers to users
- How to find and access data
- Secondary analysis

Over the last year, we ran 67 workshops in total with 4079 attendees.

New resources and guidance

During this period, our popular data skills modules provided free training and guidance for anyone looking to find, understand and analyse good quality data.

As a result, we are working on the production of more data skills modules focused on several government surveys.

One of our priorities over this time period has been to improve the resources we offer to students. As a result of this work, we have established a strategic collaboration with three doctoral training partnerships to deliver a training course for ESRC-funded postgraduate students.

The course provides the foundational data skills needed for students to use data for their own research. We set this up as a pilot course and our intention is to eventually roll it out to a wider network of users.

Following a consultation with lecturers in 2021, we also launched new online content aimed at students. The new pages bring all our student resources together, offering clear and accessible information about the UK Data Service and our data collections.

These materials contain resources for dissertation students on secondary analysis and survey data, as well as advice on how students can find the right data for their projects.

"We have acted on the results of the previous year's consultation with lecturers and produced fantastic new resources to encourage use of UK Data Service data collections in teaching. They will also help students use our data collections in their dissertations, in order to train up future generations of social researchers."

DR JENNIFER BUCKLEY
Research Associate,
UK Data Service

Census training courses

With data from the 2021 Census being released, we launched a new programme of training events to help researchers get the most from the data. These events included webinars on some of the changes between the 2011 Census and 2021 Census, what to expect from future data releases and how to develop plans for using census data.

Additionally, we provided 25 training courses and materials on working with census data for researchers to the Greater London Authority (consisting of London Councils, as well as health and housing departments).

We have also given presentations on using data from the Census at several conferences during this period, including events organised by:

- British Society for Population Studies
- Volunteering and Voluntary Sector Research
- The Runnymede Trust
- Centre on the Dynamics of Ethnicity.

Future plans include holding technical sessions with the ONS (Office for National Statistics) on the quality assurance process, statistical disclosure controls and the impact of the COVID-19 pandemic on the 2021 Census.

New access to census data

The UK Data Service has redeveloped the Integrated Census Microdata platform as a modern cloud-based system (a project sponsored by Kevin Schurer from the University of Cambridge). We now have plans to add the 1921 Census to the existing 1851-1911 corpus.

Our team at Jisc has also developed a new interface to archive and disseminate bulk outputs (Dataset - UK Data Service CKAN). Integration of outputs across the three UK statistical agencies has developed considerably since the last report, making it a unique offering from the Service.

Dissertation Award 2022 winners

The **UK Data Service Dissertation Award** annually celebrates outstanding use of data from the UK Data Service in undergraduate dissertations.

This competition has been running since 2016 and every year our panel of external judges are impressed with both the quality and variety of work accomplished by students using data through the UK Data Service. This year entries came from a breadth of disciplines, including politics, sociology, business and public health.

Jacqui Evans
School of Geography and Sustainable Development, University of St Andrews

The transition to adulthood: Investigating changes in experiences of the transition from 1980 to 2015, using a temporal cohort comparison and an analysis of contextual precursors.

Using data from a longitudinal study of adults born in 1989, Next Steps: Sweeps 1-8, 2004-2016, Jacqui Evans examined how these transitions are shaped by individual and social factors, such as gender, socio-economic background and geographical context.

The transition to adulthood among millennials, and how their experience differs from older generations.
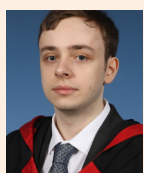
Jan van Heese
School of Education and Social Sciences, University of the West of Scotland

An investigation into Labour's electoral performance in Scotland and Wales using constitutional issues and valence voting.

Jan van Heese's analysis of data from Scottish Election Study, 2016 and the Welsh Election Study 2016 showed that both constitutional issues and valence voting were important for explaining Labour's performance.

Jan's dissertation investigated Labour's different electoral performance in Scotland and Wales between 2013 and 2016.

Christian Wessels
Department of Sociology,
University of Cambridge

Who ate my lunch? A statistical investigation of national income inequality and the social work environment in Europe.

Using the [European Working Conditions Survey from 2015 (EWCS)](), Christian was able to explore the connection between national income inequality and the social environment in the workplace, which had not been looked at in so much detail previously.

Christian's dissertation explored the impact of income inequality on workplace environments.

" Creating this new teaching resource makes this data accessible for students, too – the social science researchers of tomorrow – and it comes with a user guide to make the experience easier. This introduces students to the structure of the dataset, and the topics covered – and helps them find other training and support from Understanding Society. "

DR PIOTR MARZEC
Senior Data Analyst and Training Officer, Understanding Society, Institute for Social and Economic Research

Understanding Society:
COVID-19 Study – new teaching dataset

We have collaborated with the Institute for Social and Economic Research and Professor Sin Yi Cheung of Cardiff University to launch the Understanding Society: COVID-19 Study teaching dataset, 2020-2021.

We created this new resource following a teaching request by Professor Cheung to produce a new resource for her undergraduate students – the future generations of social researchers and policy makers.

Jennifer Buckley said that as the UK went into the first lockdown, Understanding Society developed a way to capture these experiences by asking participants to take part in the COVID-19 Study - initially monthly.

"Although the dataset was designed to meet the needs of a lecturer and their students, it is now freely available for others to use. We are proud to have been able to meet the need for an engaging dataset to support students developing vital data analysis skills.

"The dataset is useful because it is designed to provide new students with the data skills they need to analyse longitudinal data."

"I'm very happy that my suggestion of a Covid teaching data set was adopted. I never thought it would be possible to use the complex Understanding Society data for UG teaching, but the UK Data Service and the Understanding Society teams made it happen.

In particular, I was impressed by how timely their responses were and their collaborative approach in developing teaching resources with lecturers. The dataset helps engage undergraduate students who are just developing data skills as they find it particularly interesting because they too experienced many months of lockdown and some of the awful consequences of the pandemic."

**PROFESSOR
SIN YI CHEUNG
Cardiff University**

https://ukdataservice.ac.uk/2022/11/03/
new-covid-19-teaching-dataset-launched/

Reproducibility Bootcamp

We worked with the ESRC National Centre for Research Methods (NCRM) to deliver a five-week Reproducibility Bootcamp, featuring live and interactive instruction sessions.

Reproducibility is an essential part of scientific study and discovery and is a topic of importance to the social science community. We were early providers of training and the feedback from our attendees, ranging from early career researchers to more established academics, was very positive. More training events on this topic are planned.

Alongside the live sessions, we also asked participants to complete out-of-hours assignments, such as conducting additional research on a topic using the skills that had been presented in the session.

> "Reproducibility looks set to grow more important as society demands more transparency and evidence when making policy decisions and plans. I hope to take forward what I have learned from running this five-week Bootcamp to run more targeted training sessions that drill down into specific concepts, and encourage researchers to take time to learn and explore relevant skills and tools."

DR JULIA KASMIRE
Research Associate,
UK Data Service

Collaborative training events

To amplify the impact of our Training and User Support Services, we continue to collaborate with a range of partners. For example, our Research Data Management team has held meetings, webinars and workshops with:

- The Digital Futures at Work Research Centre
- The Centre national de la recherche scientifique
- University of Glasgow
- Northumbria University
- The Greater Manchester Centre for Voluntary Organisation (GMCVO)
- The Centre for Research into Energy Demand Solutions (CREDS)

Leading international best practice

In 2022, we worked with the Austrian Social Science Data Archive and the Czech Social Science Data Archive to deliver a train-the-trainer workshop with advice for practitioners and support staff on how to encourage researchers to make their data available for re-use. This was provided as part of our work with the Consortium of European Social Science Data Archives (CESSDA).

> "The workshops, which covered Data Management Planning, Ethical Considerations, and Anonymisation, have not only been informative but also helped in fostering a rich exchange of ideas among trainers. We are particularly encouraged by the positive feedback received from both participants and panellists, which underscores the effectiveness and relevance of the workshop in meeting its three main objectives: raising awareness of key tools and resources, enabling a platform for idea exchange, and providing reusable training materials."

CRISTINA MAGDER
Data Collections Development Manager, UK Data Service

Additionally, we have delivered four collaborative training events to help users get the most from the cross-sectional survey and longitudinal data we hold. These events covered:

- Mental health and wellbeing research in the UK, in collaboration with the Data Resources Training Network
- The Family Resources Survey, in collaboration with the Department for Work and Pensions (DWP)
- Accessing and using "real-world" study data, in collaboration with CLOSER (Cohort and Longitudinal Studies Enhancement Resources)
- Quantitative time-diary analysis, in collaboration with University College London and the National Centre for Research Methods (NCRM).

The Family Resources Survey is an important source of data on incomes and living circumstances in the UK. In 2022, we collaborated with the Department for Work and Pensions to deliver a workshop to introduce new users to the data, which was attended by 62 people from academic, government and third sector organisations. We continue to collaborate with the DWP to develop online learning materials to further support users.

Newcastle

Durham

York

Lancaster

Manchester

Liverpool

Leeds

Sheffield

Five collaborative training sessions were also run with the **N8 Centre of Excellence in Computationally Intensive Research,** which is a partnership between the eight most research-intensive universities in the North of England:

Durham • Lancaster • Leeds Liverpool • Manchester Newcastle • Sheffield • York

The sessions covered text-mining for digital health, managing digital health data and the use of synthetic data in research.

Census 2021/2022 explained

Professor Oliver Duke-Williams, the UK Data Service's Director for Census, produced several Census explainers with his team to provide insights around the newly released Census 2021/22 data.

These explainers covered such areas as:

- Understanding race and migration
- Internal migration
- National identity
- Ethnicity
- Veteran status
- Religious preferences
- Sexual orientation
- Gender identity

The explanations help researchers to understand how they can use the data comprehensively to analyse changes to our society over the years.
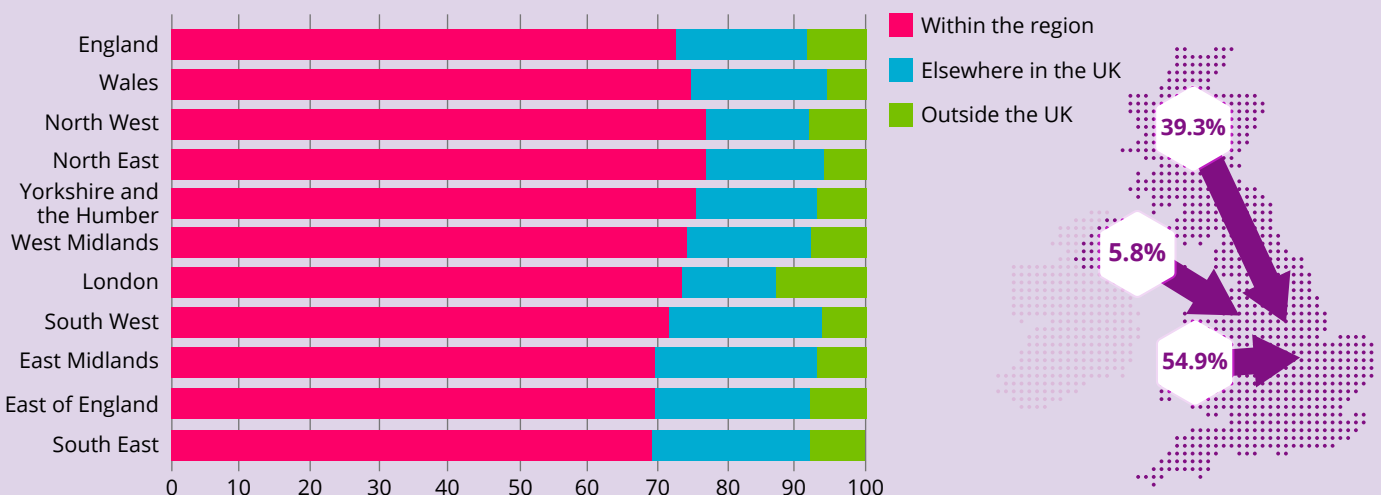
For example, the Census' internal migration data provides insights into the impact the COVID-19 pandemic has had on people moving to different places within the UK. Professor Oliver Duke-Williams said. **"As a result of the pandemic, there is some evidence of changing patterns of internal migration, particularly out of urban areas."**

The data also reveals that **54.9%** of the total moves to England from another UK country were from Wales. **39.3%** were from Scotland and **5.8%** were from Northern Ireland, as you can see in the graph below.

Additionally, across the English regions, the North West recorded the highest proportion of residents who moved within their own region, one year before Census 2021.

Origin and destination of usual residents for those aged one year and over, who had a different address one year before Census 2021, England and Wales.

Source: UK Census 2021

TRUSTED RESEARCH DATA FOR A STRONGER SOCIETY

Total migration from outside the UK was lower than that observed in 2011, probably due in part to movement restrictions during the pandemic. Looking more deeply into international migration data, the Census 2021/22 found that across the English region, the most common country of origin was India, except for in the East of England and South West. In these regions, most people who migrated from abroad were from the United States and the "other Middle East" group (which includes all countries in the Middle East, except Iran and Iraq), respectively.

The top three local authorities in England which received the highest number of people outside of the UK, as a proportion of usual residents aged one year and over, were the City of London (6.4%), Westminster (4.9%) and Cambridge (4.3%). The United States was reported as the major contributor to the migration in each of these local authorities.

All of this information is useful for understanding the current status quo. It could also be used by national and regional leaders to compare the statistics with their aims of encouraging current populations to stay in their own area; attracting more workers from other parts of the country; or attracting people from other areas of the world.

> **The UK Data Service is a very unique proposition. No one else provides direct, interactive access to post-war census data in the way we do.**

PROFESSOR OLIVER DUKE-WILLIAMS
Director for the Census, UK Data Service

Analysing geographical variations

Much of the data that is published from the census consists of aggregate data that covers an entire area, such as England and Wales, Northern Ireland or Scotland. Yet there are also key geographic variations across the UK that users may wish to examine.

The UK Data Service team at EDINA (a world class centre for data and digital expertise), based at the University of Edinburgh, helps to make this possible by bringing together digital boundary data derived from the census. This work covers a variety of different boundaries across the UK, such as administrative boundaries, electoral boundaries and postcode areas.

We combine boundary data from England and Wales, Northern Ireland and Scotland into a harmonised UK version that we then make available in several Geographic Information System (GIS) formats, for researchers to use in their work.

We have also created a **Boundary Data Selector** that allows users to make sub-selections from any boundary dataset held within the UK Data Service boundary database.

Additionally, we have been applying our expertise to help the census agencies, such as the ONS and the NISRA (Northern Ireland Statistics and Research Agency), with future data releases of census microdata and flow data.

Census release highlights

The census data released so far offers invaluable information about life across the UK. The <ins>first release for England and Wales</ins> in June 2022 registered the highest ever population in a census of the two countries and showed that a higher proportion of the population than ever before **(18.6%) were aged 65 or over**.

In November 2022, <ins>a new release provided data on religion, ethnicity and identity in England and Wales</ins>. The key takeaways included a **decline in the percentage of people indicating their ethnic group as "White"** from 86% in the 2011 Census to 81.7% in the 2021 Census. The second largest ethnic group was recorded as "Asian, Asian British or Asian Welsh", accounting for 9.3% of the population. The release also showed that for the first time ever in a census of England and Wales, fewer than half of the population (46.2%) described themselves as "Christian".

In December 2022, <ins>data was released on housing and accommodation, health, disability and unpaid care was released for Northern Ireland</ins>. The release indicated that 34.7% of people in Northern Ireland had **one or more long-term health conditions** and that around 1 in every 8 people aged 5 or older provided **unpaid care** to a relative or friend with a health condition. The 2021 Census represented the first time a census had included "**Autism or Asperger syndrome**" as a listed health condition, with around 35,000 people in Northern Ireland being recorded with the condition, representing 1.9% of the population.

The 2021 Census was also <ins>the first to include questions on sexual orientation and gender identity</ins>. Data from these questions was released for England and Wales in January 2023, showing that **89.4% of people identified as being straight or heterosexual**, with 3.2% identifying as "Gay or Lesbian", "Bisexual" or "Other sexual orientation" and 7.5% choosing not to answer the question.

The data indicated that 94.0% of people in England and Wales reported their **gender identity** as being the same as their sex at birth, with 0.55% of those in England and 0.40% of those in Wales indicating their gender identity was different from their sex at birth. Around 6.0% of people did not answer the question.

Professor Oliver Duke-Williams noted on the release: "The importance of the census is that it presents information on the whole of the population, allowing detail of diverse groups to be explored. **It is a great step forward for the census that diversity now extends to understanding more about gender identity and sexual orientation,** aspects of our lives that are vitally important, but which we've previously only really had data on through much smaller surveys."

Dr Nigel de Noronha, Research Associate at the UK Data Service, offered this comment: "**Census 2021 provides our first comprehensive picture of sexuality and gender identity in England and Wales.** We know there's more to come about the lived experiences of those within the LGBTQ+ community, as further breakdowns of the data by age, sex, living arrangements and work are published."

Data acquisitions and data management support

As a key data infrastructure in the UK for the social sciences, our focus continues to be on expanding our collection, negotiating and acquiring more high-profile data series from government departments and academic institutions. We are committed to maintaining a diverse, balanced, and representative collection that facilitates interdisciplinary research within and beyond the social sciences.

From exploring societal well-being and health to economic dynamics and cybersecurity, over 120 nationally representative data collections have been acquired this year. Highlights of new data include:

- **UCL COVID-19 Social Study, 2020-2022 –** a longitudinal study that delves into the psychological and social effects of the COVID-19 pandemic on the UK population.

- **Gemini COVID-19 Study: Home Environment Interviews and Twin Questionnaires, 2019-2021**– with the sample drawn from the Gemini Study, an established birth cohort of twins, the study examines the impact of the COVID-19 crisis on UK family home environments, including children's health, psychological wellbeing, eating, physical activity and sedentary behaviours.

- **COVID Social Mobility and Opportunities Study: Wave 1, 2021-2022 –** a longitudinal cohort study supporting research into how the pandemic affected the life chances of pupils with different characteristics, in terms of short-term effects on educational attainment, and long-term educational and career outcomes.

- **UK Business Data Survey, 2021** – a time series gathering evidence on the importance and awareness of digital data use for UK businesses and seeking to understand the role of digital data in UK businesses.

- **Cyber Security Longitudinal Survey: Wave 1, 2021** and **Wave 2, 2022** – part of the National Cyber Strategy 2022, explores evolving cybersecurity practices in UK organisations, a vital area in the era of digital transformation.

- **Public Attitudes to Data and Artificial Intelligence Tracker Survey: Wave 1, December 2021 –** assesses public perceptions of data use and AI, highlighting societal views on digital advancements.

- **European Working Conditions Telephone Survey, 2021** – explores diverse aspects of employment across Europe, from working hours to health and safety.

- **Electrification of Heat Demonstration Project: Heat Pump Performance Cleansed Data, 2020-2022** – assesses the potential for heat pump installations across diverse UK housing, relevant in the context of environmental sustainability.
- **HMIP Prisoner Survey: Adults in England and Wales, 2000-2023** – these surveys provide invaluable insights into the conditions and experiences of detainees, comprising a key resource for informed policy recommendations and potential prison reforms.
- **The People and Nature Survey, 2022** – investigates the relationship between people and the natural environment in the UK, offering insights into environmental engagement and attitudes.

In addition to introducing new data collections, we continue to dedicate substantial efforts to existing studies and series. By working closely with data owners and producers and continually updating our collection, we ensure that researchers have access to the most current and relevant data, enhancing the validity and impact of their work.

Updates include key longitudinal data such as Understanding Society, Centre for Longitudinal Studies cohorts and the English Longitudinal Study of Ageing, cross-sectional health surveys including the Health Survey for England and the Scottish Health Survey, longstanding governmental survey series such as Continuous Recording of Social Housing Lettings and Sales, and the Longitudinal Small Business Survey.

Each study in our collection is not just a repository of data but a narrative in itself, offering unique insights into our society. The practical application of these datasets truly demonstrates their value.

> By making these data available to a wider audience via the UK Data Service we hope to derive even greater value. Seeing the data used widely is a key aim – whether through further research that influences policy or demonstrates community benefits through assured evidence. Publishing our datasets via the UK Data Service will provide greater value for researchers and policymakers across the UK.

DR TOM MARSHALL
Data provider and Project Senior Responsible Officer, People and Nature Survey

Empowering research frameworks

To support our mission in enhancing knowledge sharing and data accessibility we have continued to support research frameworks and data repository quality assurance. Most notably, we engaged in a consultation process for the Department of Culture, Media & Sport Video Games Research Framework. Our contributions were geared towards providing researchers with additional resources, including guidance on making FAIR data available and overcoming key challenges in data sharing, hence enhancing its wider relevance.

Furthermore, advocating for harmonised standards of data management we have worked closely with partners at re3data COREF and CoreTrustSeal, work which culminated in an open-access report.

Cristina Magder, Collections Development Manager at UK Data Service said:

"Our emphasis on open access is not limited to data, but extends to sharing our policies, procedures, materials and resources. Under Open Licences we actively advocate their use, recognising the transformative power of open-source knowledge in fostering innovation and promoting transparency. Overall, we continue to be committed to driving excellence in research infrastructures and research data management."

Advancing net zero goals

Recognising the need for a synergised approach to foster FAIR data practices in the role that social research plays in meeting the net zero goals, we have worked collaboratively with the Centre for Research into Energy Demand Solutions (CREDS) and the UK Energy Research Centre (UKERC).

Our training and guidance to streamline processes ensured that data generated by CREDS was seamlessly made available through our repository. Concurrently, working on its associated metadata enabled it to be indexed and discoverable via UKERC's repository.

> ❝ I have been impressed at the preparedness to go the extra mile to offer bespoke support services, the training on offer, the fact training is recorded as well as on the website and the overall responsiveness of staff.❞

DR SARAH HIGGINSON
Research Knowledge Exchange Manager, CREDS

As part of our drive to continually increase and enhance our data collections for future use, we actively target the acquisition of data that is most valuable for teaching and research.

We are currently engaging with the Office for National Statistics (ONS), who are digitising the 1961 Census. In conjunction with our UK Data Service Census partners at UCL, The University of Manchester, Jisc and EDINA, we hope to build the digital collection of all census data from 1851 onwards.

UK Data Service ReShare news

30% increase
in usage compared to the previous year: **16725** from 1 April 2021 to 31 March 2022 compared to **21731** from 1 April 2022 to 31 March 2023

331 data collections
were deposited, reviewed and made available through ReShare

Data hosting
Originally created to host ESRC-funded data only, we now host data funded by a variety of UKRI research councils, including **MRC**, **EPSRC** and **NERC**, plus collections funded by other funding bodies, such as the European Commission, Leverhulme Trust and the British Academy.

ReShare is the UK Data Service's online data repository, where researchers can archive, publish and share research data, as open or safeguarded data.

Working with the Office for National Statistics

The UK Data Service and its predecessors have worked closely with the Office for National Statistics (ONS) since the 1970s. A Concordat between the ONS and the UK Data Service was established originally in 2012, with the main aim of ensuring the streamlined dissemination of data created by the ONS at appropriate access levels. Considering the changing data landscape and researchers' needs the Concordat has been reshaped and re-signed in 2022.

The agreement reaffirms our commitment to provide access to essential microdata, but also ensures the continuation of a relationship that has been instrumental in advancing research and informing policy. Under the agreement, the UK Data Service makes available key governmental surveys such as the Labour Force Survey, the Crime Survey for England and Wales, and the Living Costs and Food Survey, to name a few.

The agreement recognises and promotes the importance of a robust licence and access framework, allowing data collections to be made available at appropriate access levels, including Open Data under the Open Government Licence and Controlled Data via the Digital Economy Act framework. The UK Data Service ensures the exclusive provision of effectively anonymised data under Safeguarded access, ensuring that data are handled responsibly while remaining accessible to researchers.

NHS administrative data and National Longitudinal Cohort Studies

Since 2015, the UK Data Service has worked with the Centre for Longitudinal Studies (CLS) and NHS Digital to create an innovative linked data series for our users. This series combines information from NHS administrative records with data from four national longitudinal cohort studies administered by CLS:

- **1958 National Child Development Study**
- **1970 British Cohort Study**
- **Next Steps**
- **Millennium Cohort Study**

These four studies follow large, nationally representative groups of people born in a particular year. They capture the trajectories of people's lives such as their education, employment and health. We have combined the data from these studies with a set of four datasets known as Hospital Episode Statistics. These datasets cover attendance at Accident and Emergency care facilities, attendance at admitted patient care, attendance at critical care and attendance at outpatient facilities.

Combining these statistics with the data from the national longitudinal cohort studies allows us to create a novel resource for analysing the factors that influence health and the way health interacts with other elements of people's lives. This means that researchers can benefit from enriched data for research.

> "One of the cohorts focuses on adolescence including the generation of the millennium, which is a very interesting generation to study, and extremely significant to researchers. The linkage is amazing as it not only gives information related to households including the parents and the children and the circumstances they are in, but also on the healthcare utilisation."

DR MARIO MARTINEZ-JIMENEZ
**PHD student,
Lancaster University**

HMIP Prisoner Survey

Over the last 20 years, His Majesty's Inspectorate of Prisons for England and Wales (HMI Prisons) has conducted a systematic survey of prisoners in England and Wales to better understand detainees' experiences in custody. This survey, known as the HMIP Prisoner Survey, has taken place every year since 2000, with the exception of 2020-21 due to COVID-19 restrictions.

The UK Data Service in collaboration with Royal Holloway, University of London has engaged with HMI Prisons to make data from the Prisoner Survey available to our users for the first time under a special licence. The data offers a large and representative sample of the adult prisoner population in England and Wales, covering topics such as prisoner demographics, experiences of prison conditions, support needs, safety and education.

The survey, which has been conducted with a view to making data as comparable as possible over time and between custodial institutions, represents an invaluable resource for anyone researching issues related to prisoners in England and Wales. We are delighted to be able to bring this resource to our users.

"This is more than a database - it is the painstakingly recorded memories and perceptions of more than 100,000 prisoners surveyed over the last twenty years in every prison in England and Wales. I have long understood its potential for helping us make prisons safer and more rehabilitative places, but it is only now that the unique partnership between UK Data Service, HM Inspectorate of Prisons and Royal Holloway has enabled us to realise that potential which I know will be of real value to prison researchers, policy makers and practitioners for a long time to come."

PROFESSOR NICK HARDWICK
HM Chief Inspector of Prisons 2010-2016
(now Royal Holloway University of London)

Hardwick adds that having the data hosted within the UK Data Service is invaluable.

"The UK Data Service provides a trusted and accessible platform for researchers, educators, policymakers, and practitioners to access and make use of these critical datasets. It ensures the long-term preservation, curation, and dissemination of the data, making it readily available to a wide audience. This not only enhances its visibility but also maximises its potential to drive informed decisions and reforms, which we hope will benefit prisoners, the prison system, those involved in criminal justice policy and practice, and the broader public.

"Throughout the process of depositing the data with the UK Data Service, we found the team to be exceptionally professional, supportive, and responsive. They guided us through the depositing process and addressed our questions promptly. Their expertise and dedication to data management were more than evident, making the process efficient and ensuring the data's quality and accessibility."

Researchers, educators, policymakers, and practitioners can now use the datasets via the UK Data Service to inform various aspects of government policy and public benefit.

The data can also support evidence-based policymaking in the criminal justice system, enabling policymakers to make informed decisions about prison conditions. Additionally, it can aid researchers and educators in conducting in-depth studies on the impact of prison conditions on prisoners' well-being, safety, healthcare, rehabilitation, and education programmes, living conditions, family connections, and the effectiveness of governance and leadership within prisons.

Ultimately, the availability of this data through the UK Data Service facilitates collaboration and data-driven research, which can drive positive changes in policy and benefit everyone by contributing to a more effective and humane criminal justice system.
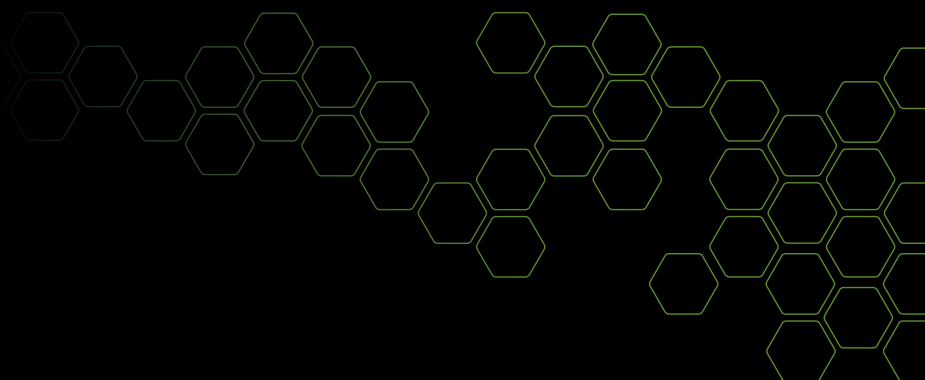
Improving access to controlled data

We provided Safe Researcher Training to 155 people, enabling users of the UKDS SecureLab, as well as continuing to provide training for other services (including users of CDRC, ONS SRS, HMRC DataLab).

Safe Researcher Training

As a result of our expertise in delivering Safe Researcher Training, we were approached by BT to provide training for some staff around the use of controlled data. We also delivered one-off training sessions in October 2022 instructing users on how to import external data to their UKDS SecureLab project spaces.

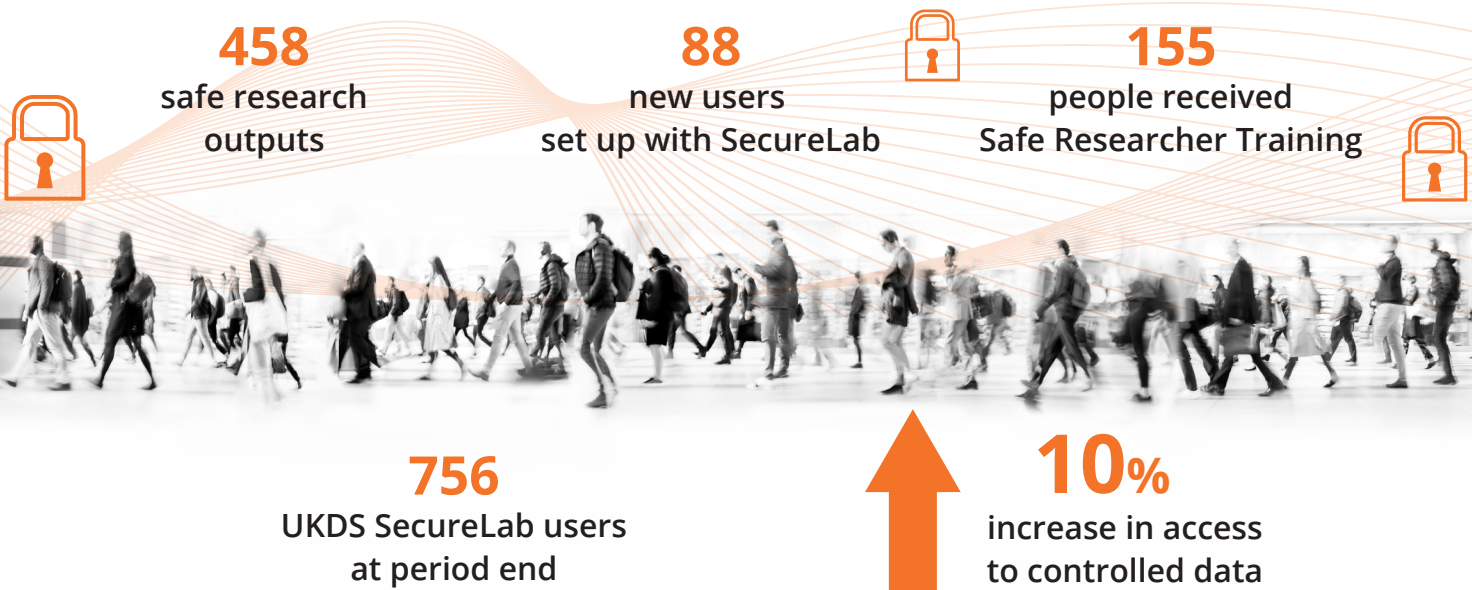From our SecureLab, we supported over 300 research teams to use our most detailed, sensitive data for their projects. We facilitated the release of 458 outputs from SecureLab during April 2022 and March 2023.

Improving the access experience

The UK Data Service is committed to a continuous improvement approach to how we support researchers to utilise our rich data collection.

From Spring 2022 onwards we have made several changes to the application process for researchers applying to use controlled data in the UKDS SecureLab. We simplified the requirements to provide information about ethical approval, and behind the scenes made process changes to allow us to speed up/ streamline the administrative steps.

Enabling safe access for more users

458
safe research outputs

88
new users set up with SecureLab

155
people received Safe Researcher Training

756
UKDS SecureLab users at period end

10%
increase in access to controlled data

"The Access Team have worked collaboratively with users and other stakeholders, to identify and deliver tangible improvements to the services we provide.

We are always striving to enhance our systems, through informed research of our own, for the purpose of helping researchers carry out important work that can have a hugely beneficial impact on our society.**"**

LEIGH TATE
Data Access and Helpdesk Manager, UK Data Service

In August 2022, the UKDS Safe Room at the University of Essex reopened, having been closed since 23 March 2020 due to the Covid-19 pandemic. As well as allowing UK Data Service users to access SecureLab through this route, this enabled wider international access through our membership of the IDAN network.

UK Data Archive, University of Essex

During 2022/2023, we streamlined the website interface to improve the registration experience for our users.

John Sanderson, the Data Access Services Manager for the UK Data Service, said these alterations have made a big difference.

> " We're enabling our users to get to the data they want faster, and improvements have led to a marked reduction in unnecessary registration requests."

JOHN SANDERSON
Data Access Services Manager, UK Data Service

In February 2023 we restructured our Access team to provide greater resilience to our capacity to process applications for data.

In Spring 2023 we successfully completed the necessary steps to maintain our Digital Economy Act Processor Accreditation, allowing researchers to use our SecureLab to access 150+ datasets which rely on DEA as the legal basis for access.

As part of the ESRC's plans to develop the Future Data Services (FDS) programme, to map out the requirements for the next generation of data services beyond 2024, we have contributed to discussions on how to improve access to vital datasets held in our collection.

In August 2022, in response to user feedback, we improved the initial registration experience through re-design of the 'Login' section of the UK Data Service website. The continued integration of our user account processes with the **UK Access Management Federation** enables us to provide single-sign-on access to hundreds of UK organisations for the smooth registration of 87% of our users.

Championing international secure access between nations

We have taken the lead in setting up the **International Secure Data Facility Professionals Network (ISDFPN)**, which was formed in March 2022, as part of the Social Sciences and Humanities Open Cloud project (SSHOC). The aim of the ISDFPN has been to bring together international colleagues working in Trusted Research Environments (TREs) or working towards their organisations becoming TREs.

After the end of the SSHOC project in April 2022, the network was taken forward as a collaboration between the UK Data Service (led by Beate Lichtwardt) and GESIS Leibniz Institute for the Social Sciences (led by Deborah Wiltshire).

Beate Lichtwardt, Senior Training and Support Officer for the UK Data Service, said the network currently involves 25 institutions from nine countries.

"The purpose of the ISDFPN is to share expertise, experiences, discuss relevant areas of their work, spark collaboration and nurture new ideas. It has expanded to different disciplines, including the social sciences, the health sector, and humanities.

> **"** ISDFPN is unique in discussing issues in relation to quantitative secure data and also by actively working on, for the first time, generating options for making qualitative secure data available via Secure Data Facilities in the future.**"**

BEATE LICHTWARDT
Senior Training and Support Officer, UK Data Service

Data Impact: developing a broader picture

7

Data Impact is a growing area in which the UK Data Service leads the way. As part of developing our understanding of the impact of data from the collection, we have begun exploring a broader view of impact in terms of research, policymaking and beyond.

This could be through the wider use of datasets across research, a larger body of research work or how data enhances research in particular areas.

Our initial foci for this 'broader picture' approach are three themes which are extremely relevant to society:

- **Poverty, deprivation and the cost-of-living crisis**
- **Housing and homelessness**
- **Mental health and wellbeing**

Our aim is that these themes will build a narrative of the impactful work being done in the area that is linked to the UK Data Service data and beyond. As we build our understanding of these broader areas, we will continue to publish case studies, add to discussions through the Data Impact blog and provide a point where the joined-up narrative is easily available.

Data Impact blog

The **UK Data Service Data Impact blog**, established in 2015, continues to attract a wide audience. It received almost 22,000 pageviews by nearly 20,000 visitors during the period covered by this report. Readers have visited the blog from 199 countries around the world.

A key theme on the blog during this period has been the release of data from the 2021 Census for England and Wales and the 2021 Census for Northern Ireland. Given the importance of this data, **the blog has published a dedicated series of articles on new releases**, as well as some early analysis of what we can learn from the data.

Full Fact

We worked with the fact-checking charity Full Fact to publish a **blog post on how data from the UK Data Service supports its work**.

It is vital that claims in the public domain can be checked quickly and accurately using reliable information. Full Fact regularly uses data from the UK Data Service to carry out its work. Our data has been used to assess issues such as **regional housing tenure**, the reasons for **young people being in NEET status** and to carry out comparisons of people with **zero hours contracts who are members of a trade union**.

Members of Full Fact have also attended several of our online training sessions and webinars, where we were pleased to give advice on the services we offer and the best methods for accessing our data. This underlines the important role the UK Data Service plays in enabling research and the wider impact we continue to have on society.

COSMO

In January 2023, the UK Data Service was pleased to highlight the work of one of our data depositors, the **COVID Social Mobility and Opportunities study (COSMO)**. The study, which is a collaboration between the **UCL Centre for Education Policy & Equalising Opportunities (CEPEO)**, the **UCL Centre for Longitudinal Studies (CLS)** and **the Sutton Trust**, provides detailed evidence of the unequal impact the COVID-19 pandemic had on citizens across the UK.

Jake Anders, Principal Investigator of the project, wrote a post for the Data Impact blog highlighting **key early findings that over half of young people experienced changes to their education and career plans as a result of the pandemic.**

The Homelessness Monitor and Homelessness Projections

We worked with Professor Glen Bramley and Professor Suzanna Fitzpatrick of Heriot-Watt University, who used datasets in our collection to produce a key monitor of homelessness and a model for homelessness projections.

The work offers vital evidence on the scale of the UK's homelessness problem. Official measures of homelessness can present a misleading picture as not all core homeless people apply to councils for assistance. Their research has had a major policy impact by informing the Homelessness Reduction Act (2017), helping to improve official measures of homelessness and leading to policy changes on issues such as affordable housing.

More than 200,000 households in the UK experienced "core homelessness" in 2020. Core homelessness is a concept developed jointly by the researchers and the homeless charity Crisis in 2017. Based on this model the number of those experiencing core homelessness was set to increase in line with the current policies of the time. This has proved to be the case. As of the end of 2021 it had risen to 227,000 households and could further rise to 300,000 by the end of 2023.

Bramley's work for the National Housing Federation and Crisis fed heavily into a report by the cross-party House of Common Select Committee on Housing, Communities and Local Government in July 2020. A key section of this report highlights and quotes extensively from Bramley's NHF/Crisis study and endorses it as its recommended target basis. Its key recommendation is to assert in bold that 'There is compelling evidence that England needs at least 90,000 net additional social rent homes a year'.

In November 2022, Bramley, Fitzpatrick and other researchers at Heriot-Watt University produced an additional report on the impact of homelessness on black and minority ethnic communities. The report, which was produced in partnership with Race on the Agenda and Oak Foundation, found that black people in England are over three times more likely than white British people to experience homelessness.

Data Accessed:

- Labour Force Survey
- English Housing Survey
- Understanding Society
- British Household Panel Survey
- Scottish Household Survey
- CORE (Continuous Recording of Social housing lettings data)
- 1970 British Cohort Study

" The micro individual/ household level datasets have been very valuable for modelling key housing decisions and outcomes.

The ability to attach housing and labour market information at local or (sub-)regional levels has been crucial to developing effective models of this kind, facilitating more effective and relevant model structures."

PROFESSOR GLEN BRAMLEY
Heriot-Watt University

PROFESSOR SUZANNE FITZPATRICK
Heriot-Watt University

How can we tackle work insecurity in the UK?

We worked with Joe Richardson from The Living Wage Foundation (LWF), who used data from the Labour Force Survey (LFS) and the Family Resource Survey (FRS) to promote a **Living Hours accreditation system for employers**.

Building on its earlier work introducing the **'real Living Wage'**, the **Living Wage Foundation** launched the **'Living Hours'** accreditation scheme. The LWF used data from the LFS and FRS to demonstrate that the Living Hours scheme can play a crucial role in solving the problem of work insecurity and consequently help increase the number of accredited employers.

Their research on the **'insecurity complex'** has sharpened understanding of the scale and severity of insecure work in the UK, especially during the Covid-19 pandemic. They helped improve accuracy of low pay figures through developing a re-weighting methodology that pegs LFS data to Annual Survey of Hours and Earnings (ASHE) data.

Living Hours is an accreditation scheme designed to provide workers with secure hours, predictable shifts and working contracts which reflect actual hours worked, alongside a real Living Wage. The scheme was developed over an 18-month period

of consultation with workers, Living Wage Employers, trade unions and experts, which culminated in 3 key measures:

- Decent notice periods for shifts: of at least 4 weeks' notice, with guaranteed payment if shifts are cancelled within this notice period.

- The right to a contract that reflects accurate hours worked.

- A guaranteed minimum of 16 hours a week (unless the worker requests otherwise).

By enhancing awareness among businesses and policymakers of the "Living Hours" accreditation scheme as a solution to the problem of work insecurity, the research has contributed to lasting benefits for employers and vulnerable employees in some of the most insecure regions of the UK.

Over 120 employers across the UK have signed up to becoming 'Living Hours' employers, with more expected.

Enhancing understanding of the socioeconomic distribution of alcohol-related violence and anti-social behaviour in England and Wales

Lucy Bryant and Dr Carly Lightowlers from The Institute of Alcohol Studies at the University of Liverpool used the Crime Survey for England and Wales to investigate **alcohol-related violence and anti-social behaviour**.

Alcohol-related violence and anti-social behaviour (ASB) place a significant burden on the public and emergency services. 2 of every 5 (42%) violent crimes are committed under the influence of alcohol each year **(figures taken from 2019/20)**, and the majority of 13,000 Penalty Notices for Disorder issued for non-notifiable offences in England and Wales **in 2017** related to drunk and disorderly behaviour.

Those belonging to the lowest socioeconomic groups may be more likely to experience these incidents. Socioeconomic inequalities have been repeatedly identified in other alcohol-related harms, such as health outcomes.

Despite this, the socioeconomic distribution of alcohol-related violence and ASB remains under examination, particularly analysis drawing out patterns in subtypes of alcohol-related violence, such as alcohol-related domestic violence.

The research presented in this case study made extensive use of the Crime Survey for England and Wales to address this challenge.

The researchers showed how the experiences of alcohol-related violence and anti-social behaviour are not evenly distributed across society. Lower socioeconomic groups experience higher prevalence rates of alcohol-related violence overall, higher incidence and prevalence rates for alcohol-related domestic and acquaintance violence, and are more likely to experience alcohol-related ASB weekly or more often.

These research findings, which have influenced parliamentary debate on the Domestic Abuse Bill, demonstrate that the provision of publicly funded domestic violence services must be urgently revisited, alongside interventions related to alcohol pricing and availability.

"Research has repeatedly linked alcohol's **price** and **availability** with rates of violence at the local and population level. Others have also identified through modelling that alcohol-related health inequalities can be reduced through alcohol **pricing interventions**. Given this, our findings of similar inequalities in experiences of victimisation suggest alcohol pricing and availability interventions might disproportionately benefit lower SES groups. This should be investigated by researchers and policymakers."

Data Impact Fellows

Ben Brindle, one of our 2019 Data Impact Fellows, ran an exciting competition for 16-19 year olds with support from the UK Data Service impact team, which he called the "**Changing Perceptions**" challenge. Opinion polls in recent years have consistently found that the UK public believes immigration has a negative impact on the economy, their job opportunities, and their wages.

At the same time, media presentations of immigrants are overwhelmingly negative. Economic research, on the other hand, shows that, if anything, immigration tends to have a positive impact on local economies.

The challenge of the competition was how to help change public attitudes by communicating these findings in an exciting and engaging way.

The competition asked participants to compare public and media attitudes to immigration with the latest findings of data-driven academic research, and then to design a campaign to improve public understanding of the economic impacts of immigration.

In May 2022 our **panel of judges** were delighted to announce Lily from Tiffin Girl's School as the winner. You can read the **winning entry here**.

It was great to have two of our previous Fellows share on the Data Impact blog over the year.

Dr Sarah Knight, an ESRC Postdoctoral Fellow in the Department of Geography at the University of York, **shared her experience** using a SafePod. SafePods are a new way of accessing secure data. Sarah discussed **how to book and use the pods** as well as the potential benefits of using them and why they're a **positive innovation for researchers**.

We also had Dr Claudia Zucca, Assistant Professor at the Jheronimus Academy of Data Science, Tilburg University, and several of her colleagues, **share their dynamic approach to teaching data-driven entrepreneurship** in the Netherlands. This course allows students to get hands-on experience in data analytics when looking at real-life business situations.

> **"** It is great to read about Sarah's positive experience of using a SafePod to access her research data. I hope the SafePod Network will be of benefit to many more researchers across the UK to provide equitable access to data and help vital research for the public good.**"**

DARREN LIGHTFOOT
SafePod Network Manager, University of St Andrews

Knowledge exchange and international collaborations

The UK Data Service prides itself on working with partners across the UK, Europe and the rest of the world, to share our knowledge on best practice in a number of different disciplines, within and beyond the social sciences and humanities.

These areas of expertise include research data management, training in how to use key datasets, the implementation of global curation and preservation standards, and the development of new technology.

CESSDA

During this period, we continued our close collaboration with CESSDA, providing advice on how to improve data processing and management standards in Europe and throughout the world, working closely with our European partners.

This included working on a number of CESSDA-internal and European funded projects including CoreTrustSeal support, monitoring the trusted repository landscape, participating in the emerging SSHOC (Social Sciences and Humanities Open Cloud) cluster and the FAIR IMPACT project.

These activities involved developing new processes for implementing metrics; driving automation; increasing the reuse of data; aligning principles; better preserving data collaboratively; expanding standards across the data landscape; implementing data quality assurances; and supporting of FAIR-enabling and certified Trustworthy Digital Repositories.

Additionally, we continued to produce the CESSDA Newsletter.

Highlights from CESSDA projects

Leading on the FAIR-IMPACT programme

Funded through the European Commission's Horizon scheme, this project aims to enable the "FAIRification" of different research objects, such as datasets, software and semantic artefacts, including controlled vocabularies and ontologies, originating from a large range of scientific disciplines. FAIR[1] means: Findable, Accessible, Interoperable and Reusable.

Repositories that enable FAIR data can use a range of FAIR assessment tools to evaluate the FAIRness of digital objects against the 15 FAIR Principles. This can include measures to address data, metadata, documentation and technical and human infrastructure.

The FAIRification process consists of the following steps:

- Retrieving non-FAIR data: gain access to the data to be FAIRified.
- Analysing the retrieved data: inspect the content of the data. Then we ask – which concepts are represented? What is the structure of the data? What are the relations between the data elements?

This FAIR-IMPACT work will include extending the use of FAIR assessment metrics to a variety of research communities, including the social sciences.

This is the first time these tests and metrics have been developed in cooperation with FAIR-IMPACT Social Sciences and Humanities (SSH) partners.

[1] https://doi.org/10.1038/sdata.2016.18

Review of the CoreTrustSeal for future use

The UK Data Service prepared a working paper about the potential application of the CoreTrustSeal to data and metadata services, beyond Trustworthy Digital Repositories.

The Core TrustSeal is an international certification that promotes sustainable and trustworthy data infrastructures by setting out requirements for Trusted Digital Repositories.

Some concluding remarks from this paper provided an initial outline of the functional components required of new trustworthy data services. With the exceptions of a "Preservation Plan" and a "Reuse" strategy, our assessment is that all the CoreTrustSeal Requirements are applicable to any (meta) data service that seeks to demonstrate trustworthiness to its depositors, funders, users and peers.

This paper was a first step in an internal planning and analysis process conducted by the UK Data Service. It was shared in the interests of transparency and to provide a reference point for some engagement with peers on this topic.

L'Hours said the main takeaways included:

> **"** Trust and FAIR assessments continue to be useful in helping to clarify the added value of trustworthy digital repositories offering preservation services, in contrast to more technically-driven repository systems that cannot guarantee the accessibility and usability of data in the long-term.**"**

Exploring the Trust Landscape with CESSDA

Our Repository and Preservation Manager for the UK Data Archive, Hervé L'Hours, led a presentation at CESSDA during this time period, which involved discussing Trust activities as an essential part of open science, FAIR, and the European Open Science Cloud (EOSC).

The webinar provided an overview of the recent developments in Trust activities, focusing on the SSH domain and Europe. Topics covered included repository certification, curation and preservation levels discussion, community principles, metrics, and current initiatives within and beyond Europe. The webinar's main target group was CESSDA organisations, but it was open to all.
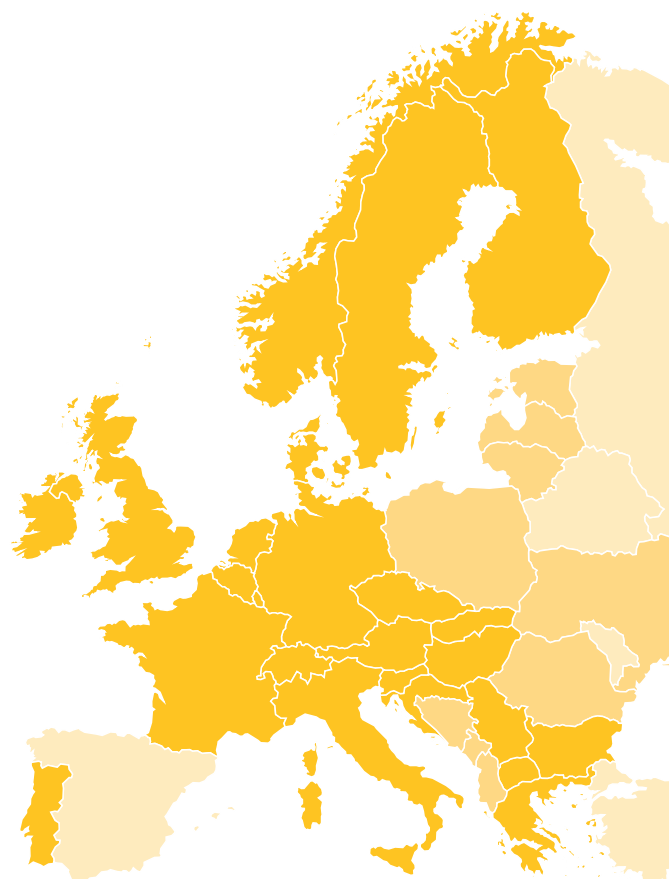
Increasing the reuse of data internationally

The UK Data Service is leading on the idea of setting up a new network of FAIR-enabling Trusted Digital Repositories (TDR), to help bring about the possibility of achieving community consensus on the best way to increase the reusability of data, and the uptake of the FAIR data principles with more organisations.

This idea was initially proposed by the FAIRsFAIR Coordination Plan for a Sustainable Network of FAIR-enabling Trusted Digital Repositories.

A collaborative working paper, FAIR+Time: Preservation for a Designated Community, was published by members of the CESSDA-Trust, SSHOC, EOSC-Nordic and FAIRsFAIR projects, which highlighted the need for an interoperable ecosystem of FAIR data preserved by FAIR-enabling Trusted Digital Repositories, along with a statement of support towards such a network.

Hervé L'Hours said: "Increasing the reusability of data across the digital preservation landscape is not only dependent on the availability of adequate methods, but also their uptake.

"This requires a community-wide commitment to develop best practices to increase the reuse of data. Once reached, the best practices can be incorporated within the requirements of the different TDR certification bodies such as the CoreTrustSeal.

"This network could help ensure that specific and minimal measures are in place to provide evidence that the relevant designated community (as a subset of the wider community of users) are consulted on the necessary requirements to ensure the availability of data for reuse."
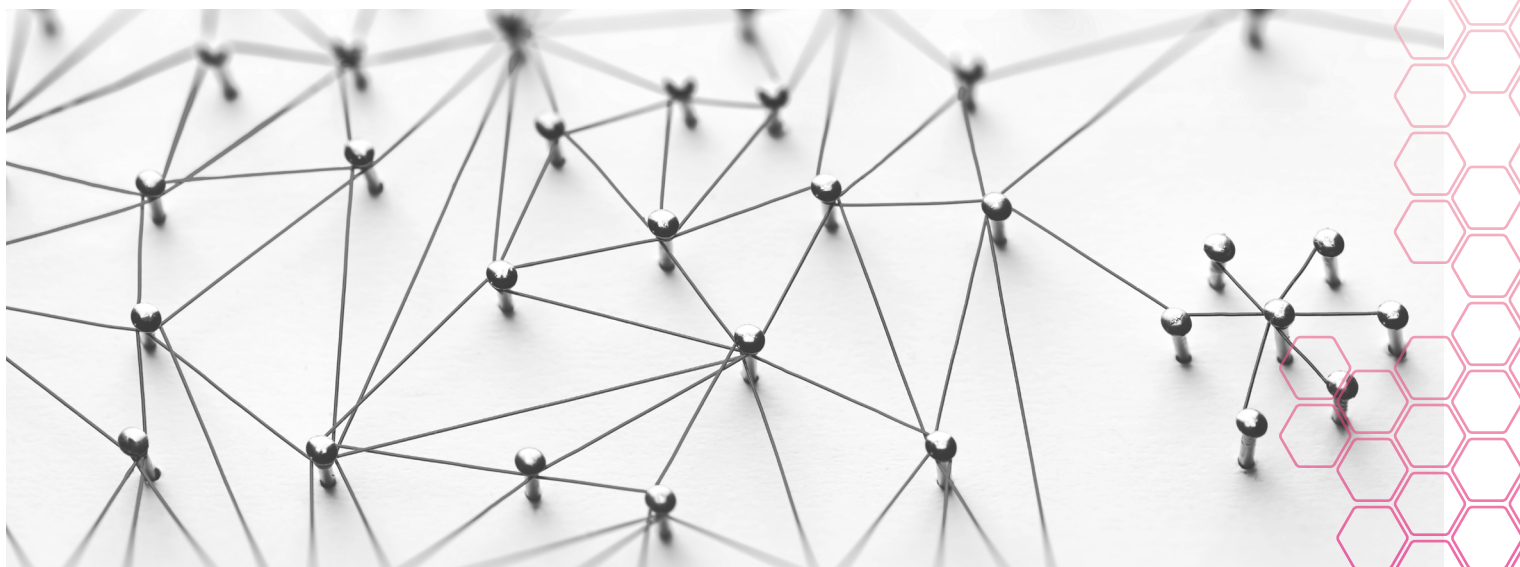
# New developments with the European Language Social Science Thesaurus

CESSDA's European Language Social Science Thesaurus (ELSST) is viewed internationally as an exemplar of best practice of semantic web data and FAIR data. We were pleased to have worked closely with CESSDA partners in developing and making available the third version of ELSST, which was released in September 2022. The fourth version is due to be released in September 2023.

Two new languages, Hungarian and Icelandic, were included in the latest release. Staff from the **Centre for Social Sciences Research Documentation Centre (TK KDK)** in Hungary and **DATICE (The Icelandic Social Science Data Service)** worked with the ELSST team to produce the new translations.

Existing language versions have also been updated where possible, including the addition of Definitions and Scope notes for the Lithuanian translation.

ELSST content work this year includes completion of the revision of concepts related to sexuality, and an update of terms related to migration. This release also sees a number of technical innovations designed to improve ELSST compliance with the **FAIR** principles, including more defined versioning (this release of ELSST is Version 3); improved visibility of deprecated concepts; and a simpler, more streamlined **SKOS** concept scheme.

> ELSST is now available in 16 languages including English: Czech • Danish • Dutch • Finnish • French • German • Greek • Hungarian • Icelandic • Lithuanian • Norwegian • Romanian • Slovenian • Spanish • Swedish
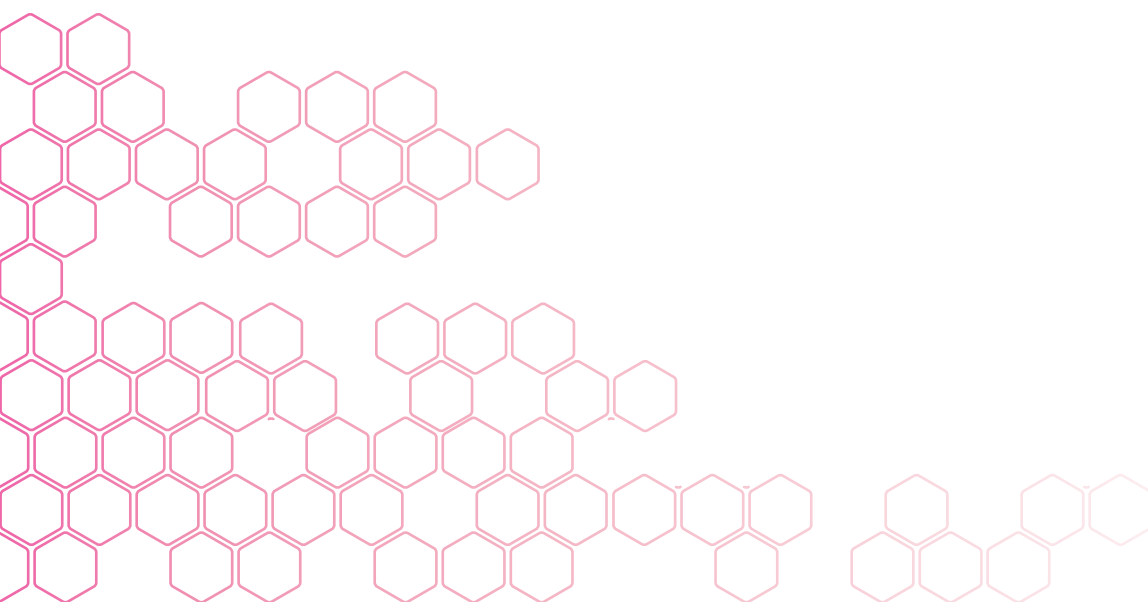
These are the key pieces of work that have been covered during this period in Europe:

Implementing metrics for automated FAIR digital objects assessment in a disciplinary context

Review of CoreTrustSeal Applicability to non-Preservation (Trustworthy) Data Services

How a European network of FAIR-enabling Trustworthy Data Repositories can align to the vision of EOSC

CoreTrustSeal+FAIRenabling: Alignment between the FAIR Principles and CoreTrustSeal 2023-25

EOSC Preservation: Overview Discussion Paper

CESSDA Trust Landscape Event 2022 [Webinar]

Data Quality Assurance from the Perspective of CoreTrustSeal

Trust Standards, Support and FAIR Enabling Trustworthy Repositories

Increasing the ReUse of Data through FAIR enabling Certification of Trustworthy Digital Repositories

Information Architecture For Secure Trustworthy Data Services

FAIRsFAIR Data Object Assessment Metrics

International Data Access Network (IDAN)

As founding members of IDAN since 2018, we remain committed to enabling international access to controlled data across Europe and internationally.

The UK Data Service now has an agreement and the technical set-up in place with our French IDAN partners from CASD to safely enable research access to secure data. This follows the first transnational secure access via safe room access points in Germany (IAB FDZ and GESIS) a year previously.

At present, a French researcher is looking forward to analysing Understanding Society data made available via this new UKDS access point in the CASD Safe Room in Paris.

Furthermore, an insightful 'IDAN Expert Online Workshop' was held on 22 May 2023. The future strategy of this collaboration between European data providers aims to enable and improve access to controlled data across borders, which is dependent crucially on the demands and needs of the research community.

Researchers, who have used or are currently using international controlled data, were invited to participate in this Expert Workshop. The outcomes of this workshop will help us to better understand researchers' needs which will inform IDAN's future work.

The GESIS building in Germany

Annual User Conferences

These unique forums provide attendees with an opportunity to hear from the producers of national survey and longitudinal datasets about the latest developments in data, and about the research findings from researchers who use these data.

In February 2023, we held a successful Labour Force and Annual Population Surveys User Conference, with 141 attendees.

Jennifer Buckley, a Research Associate at the UK Data Service, said they have been running the conference now for about twenty years and it continues to go from strength to strength.

" As always, it was incredibly popular with attendees from academia, government and many other organisations outside of academia. The keynote address from Professor Richard Blundell, University College London and Co-Director at the Institute for Fiscal Studies, highlighted his impactful research on inequality, redistribution and wage progression data.

The Office for National Statistics also gave a presentation on the highly anticipated Transformed Labour Force Survey. Since the Labour Force Survey is one of the most well-used datasets available from the UK Data Service, we will be collaborating with the ONS to provide comprehensive user engagement and support during the transition to the Transformed Labour Force Survey. Planned activities include an introductory webinar, a user conference and a workshop.**"**

DR JENNIFER BUCKLEY
Research Associate, UK Data Service

Awareness, reach and communications

Nearly 5,000 people read our quarterly newsletters during the reporting period where we continued to raise greater awareness of our latest curated data collections, training and data impact.

One of our main stories highlighted that **in contrast to national media, data revealed that the quality of family relationships are more important than their composition**, whilst the 2021 British Social Attitudes Survey was cited for confirming the majority of UK public have socially liberal attitudes.

Dr Vanessa Higgins published two journal articles that provided some important insights into our approach to developing data literacy and how UK models for improving statistical literacy could present opportunities for Colombia.

Two new podcasts were added to our successful series, **The Experts' Expert**, highlighting UK Data Service expertise. One from a User Support and Training perspective, the other from a research perspective, with Dr Neli Demireva. A thought-provoking series that helps experts showcase their work to a wider audience, we welcome curated data collection providers and researchers who wish to be interviewed.
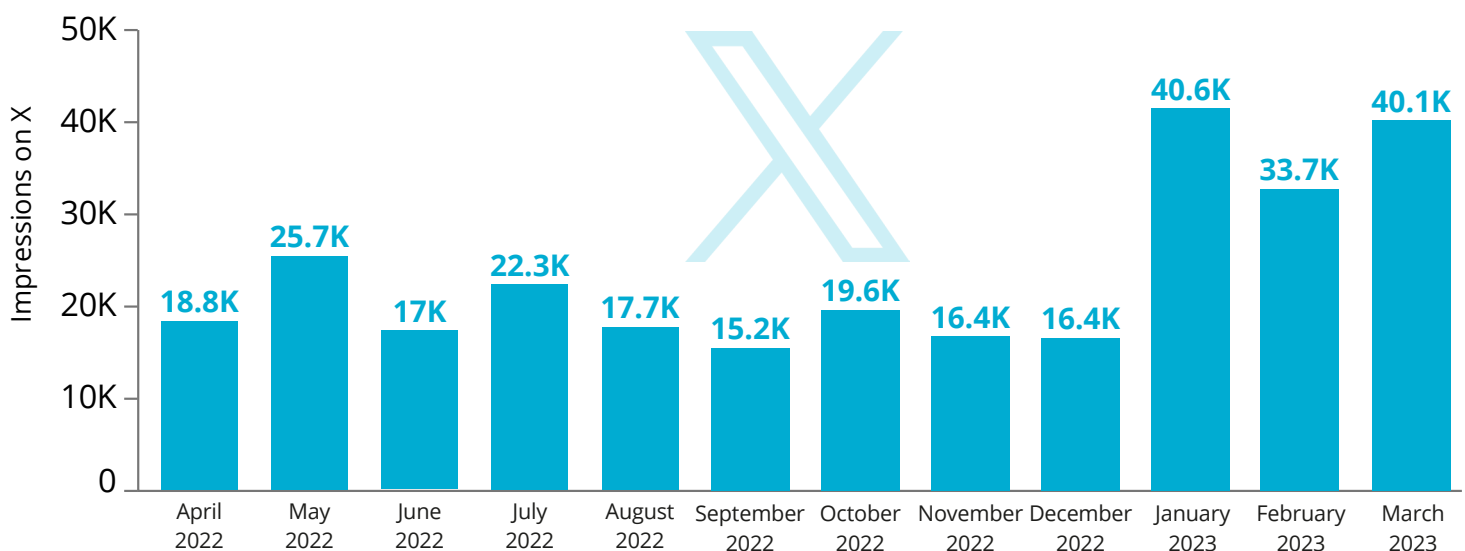
The spike between January and March 2023 is due to our promotions around the new Census dataset releases, Love Data Week, International Women's Day, the Dissertation Award and our Annual Report.

We presented to the CLOSER communications group about how effective our training promotional campaign was and continued to produce the CESSDA Newsletters and share best practice with CESSDA and CLOSER colleagues.

Our Data Impact themed newsletter became the most read during this period, whilst our focus on census data and services led to a significant increase in social media engagement.

In January 2023 we received 40.6K impressions on X (formerly known as Twitter), compared to 16.4K impressions in December and 16.4K impressions in November.

In news to celebrate International Women's Day, we promoted datasets in our collection that can help to further equality for women and help break down traditional taboo subjects, including data and research around menstruation, the menopause, miscarriages, and mental health related to motherhood.

As with all cross-party themes and societal challenges, data and research in our collection can provide invaluable insights to drive positive change and inclusion in the future.

Becoming more user-centric

We work closely with users to improve the customer experience and accessibility across everything we do.

Frequent user testing throughout the UK Data Service is an important way we continue to improve and develop services that better meet customer needs. We listened to lecturers and are proud to have developed new student web pages in our Learning Hub and free online events to help new researchers and undergraduates use our data collections and dissertation resources more effectively.

If you use the UK Data Service in some capacity, we want to hear from you. Please contact comms@ukdataservice.ac.uk if you want to participate in user testing or want to share feedback on our services.

" The UK Data Service is a cornerstone of British social science research. It continues to find ways to push the boundaries of what the UK social science research community can achieve. By equipping researchers with better data and better skills, the UK Data Service facilitates better answers to better questions, paving the way for excellent social science research to measurably improve lives."

PROFESSOR SUZY MOAT
Professor of Behavioural Science at Warwick Business School and UKDS Advisory Board member

Looking ahead

In 2023 the UK Data Service secured £37.5 million UK Research and Innovation (UKRI) funding through Economic and Social Research Council (ESRC) to continue to meet the evolving needs of social science research and active data preservation (2024-2030).

This ensures we can continue our pioneering work in the technical development of digital standards and further innovation in data curation, research data management, data literacy and impact.

User requirements will continue to drive improvements to our services as we start to set the strategic direction of the UK Data Service for the new award and support the ESRC in their development of a robust Future Data Service.

Host organisations and expertise

- **UK Data Archive – University of Essex:** lead partner of the UK Data Service.
The UK Data Archive is curator of the largest collection of digital social science data in the United Kingdom. Founded in 1967, the Archive is an acknowledged centre of expertise in the areas of acquiring, curating, preserving and providing access to data. In 2010 it became the first academic department in a UK university to earn ISO27001 certification, an international standard of information security. In 2020 it was accredited to continue providing controlled access to secure research data under the Digital Economy Act 2017, and gained international certification from CoreTrustSeal as a trusted digital repository.

- **Cathie Marsh Institute for Social Research - University of Manchester**: the Cathie Marsh Institute for Social Research (CMI) is a centre of research excellence, specialising in the application of advanced quantitative methods in an interdisciplinary social science context. CMI is named after Cathie Marsh who founded the Census Microdata Unit in 1993.

- **Jisc**: Jisc is a registered charity and champions the use of digital technologies in UK education and research. As part of the Digital Resources Directorate of Jisc, UK Data Service staff provide access to, and specialist support for, the databanks of inter-governmental organisations such as the International Monetary Fund, the World Bank and OECDA, as well as aggregate statistics from the 1971 to 2021/2022 UK censuses.

- **Department of Information Studies - University College London**: the UCL Department of Information Studies is an international centre for knowledge creation and transfer in the fields of librarianship, archives and records management, publishing, information science and digital humanities. The department brings together academics and practitioners in these fields, with research aiming to develop the understanding and insights needed to shape the emerging information environment, while elucidating and building on the historical developments that have created this environment.

- **Centre for Advanced Spatial Analysis - University College London:** The UCL Centre for Advanced Spatial Analysis (CASA) is one of the leading forces in the science of cities, generating new knowledge and insights for use in city planning, policy and design, and drawing on the latest geospatial methods and ideas in computer-based visualisation and modelling. CASA is part of The Bartlett, UCL's global faculty of the built environment.

- **EDINA - University of Edinburgh**: EDINA is a national academic data centre operating since 1995. It is designed by Jisc to support the activities of universities, colleges and research institutes across the UK. Its mission is to develop and deliver cost effective shared services and infrastructure for research and education, founded on innovation, research and evidence-based knowledge. EDINA also continues to act as the ESRC specialist geography unit for the census, supporting UK Data Service users in accessing and using the geography outputs of the 2021/2022 and previous censuses.

Appendix

Top 10 most downloaded studies

April 2022 to March 2023

	Study number	Title	Number of downloads
1	6614	Understanding Society: Waves 1-12, 2009-2021 and Harmonised BHPS: Waves 1-18, 1991-2009	8202
2	8644	Understanding Society: COVID-19 Study, 2020-2021	2765
3	5050	English Longitudinal Study of Ageing: Waves 0-9, 1998-2019	1914
4	8011	Crime Survey for England and Wales, 2013-2014: Unrestricted Access Teaching Dataset	1577
5	8485	Quarterly Labour Force Survey, January - March, 2019	1236
6	8812	Crime Survey for England and Wales, 2019-2020	1052
7	8682	Millennium Cohort Study: Age 17, Sweep 7, 2018	943
8	7483	Young Lives: an International Study of Childhood Poverty: Rounds 1-5 Constructed Files, 2002-2016	929
9	8860	Health Survey for England, 2019	897
10	5151	British Household Panel Survey: Waves 1-18, 1991-2009	787
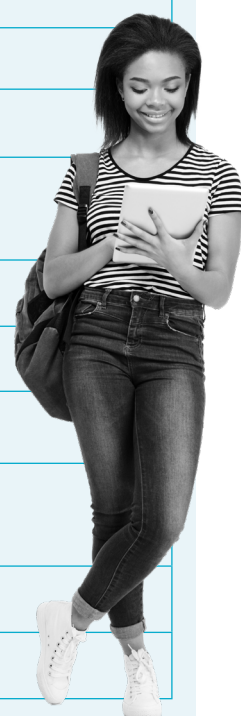
Top 20 countries accessing UK Data Service data collections globally

Country	No. of times accessed	Country	No. of times accessed
United Kingdom	67050	Canada	429
United States of America	3054	Switzerland	384
Germany	1636	India	336
China	1418	Belgium	319
Italy	901	Korea (the Republic of)	256
Spain	876	Poland	224
Netherlands	764	Denmark	200
France	483	Sweden	197
Peru	477	Australia	191
Ireland	430	Israel	189

Top 20 Institutions accessing UK Data Service data collections

Institution	No. of times accessed	Institution	No. of times accessed
University of Warwick	6488	University of Essex	1251
UCL	4776	University of Oxford	1199
London School of Economics and Political Science (LSE), University of London	3846	King's College London	1115
University of Glasgow	3318	University of Leeds	1110
University of Manchester	2743	University of Kent	1064
University of Edinburgh	2139	University of Cambridge	974
University of Sheffield	1939	Queen Mary, University of London	928
University of Bristol	1799	University of Strathclyde	876
University of Birmingham	1785	University of Southampton	874
University of Nottingham	1521	University of Bath	740

Find out more about how you can access valuable datasets from our collection:

www.ukdataservice.ac.uk

UK Data Service

This report summarises the work of the UK Data Service, covering the period 1 April 2022 to 31 March 2023. The UK Data Service has made every effort to ensure the information and data were accurate at the time of printing. Should you notice an error please notify: comms@ukdataservice.ac.uk.

UK Data Service

in **LinkedIn**

𝕏 **X, formerly known as Twitter**

🌐 **www.ukdataservice.ac.uk**

Economic and Social Research Council