

# ANNUAL REPORT 2017/2018

---





UK Data Archive at the University of Essex:  
lead partner of the UK Data Service.

# Contents

About us	4
Research	10
Case study: climate change	16
Reach	18
Case study: carers and obesity	24
Infrastructure	26
Case study: cycling safety	30
Advocacy and collaboration	32
Capacity building	34
Case study: dementia and culture	38
Looking ahead	40

We provide seamless, unified access to the UK's largest collection of social, economic and population data. We meet the needs of researchers, analysts, teachers and students in academia, central and local government, charities and foundations, independent research centres, think tanks, business consultancies, communities and the commercial sector.

# UK Data Service at a glance

180,000  
downloads  
worldwide  
per year



23,000+  
registered users



7,400+  
datasets



6,000+  
user support  
queries  
per year



1,000+  
qualitative and  
mixed methods  
collections



1,000+  
Secure Lab  
users



800+  
Secure Lab  
projects



400+  
new datasets  
added  
each year



50+  
workshops



35  
safe user  
training  
sessions



10  
user  
conferences



0  
data breaches



This Annual Report highlights our major achievements from April 2017 to September 2018.



Since our last report, the research landscape in the UK has seen significant change.

## Changing landscape

Since our last report, the research landscape in the UK has seen significant change with the establishment of UK Research and Innovation (UKRI) brings together our funders – the Economic and Social Research Council (ESRC) – with the six other existing research councils, and with Innovate UK and the research and knowledge exchange functions of the Higher Education Funding Council for England.

The unified body was set up following a review by Sir Paul Nurse, who recommended this approach as a way for research councils to evolve and support research in the most effective ways – and to encourage cross-disciplinary working. The establishment of UKRI – now led by former Chief Scientific Adviser Sir Mark Walport – also came with an announcement of £4.7 billion of extra funding for science and innovation.

For us, it means responsibility to a wider group of stakeholders, and over much of 2018 (continuing into 2019) the UK Data Service has been participating in UKRI-led initiatives on the development of roadmaps for cross-council research infrastructure and e-infrastructure.

At the time of writing it's impossible not to remark on the potential consequences of Brexit for the UK Data Service. The UK Data Service has a deep relationship with CESSDA ERIC (the Consortium for European Social Science Data Archives), and – while there are some practical implications in relation to Britain's exit from the European Union – we have communicated the importance of this relationship to the ESRC to make sure they present the case for continued collaboration so that our relationship continues in much the same way. The UK Data Service values its collaborative activities within CESSDA, especially in enhancing access to UK data to researchers within Europe, and also in promoting access to European data to researchers within the UK.



Punched cards were used until the 1990s, and the UK Data Archive still has a punched card reader, which we used recently to recover data from a 1960s survey of peace organisations.



## Marking anniversaries

Since our last report, the UK Data Archive and the English Housing Survey both celebrated 50th anniversaries, and the National Child Development Survey turned 60, giving us an opportunity to look back over five decades of data curation and research.

The UK Data Archive based at the University of Essex, began life as the Social Science Research Council Data Bank in 1967, the year the floppy disc was invented – although its data mainly arrived on punch cards. The following decade, when the technology of choice was large magnetic tape, the Government Statistical Service began to deposit government surveys at the Archive. In the 1980s, we actively encouraged secondary data use and greater acceptance of data sharing – and by the end of the 1990s, the most accessed studies included the British Household Panel Survey and Labour Force Survey. In the 2000s, the Archive was designated a place of deposit for public records from the National Archives, and in the 2010s it became the lead organisation in the UK Data Service.

The English Housing Survey (EHS) also began life in 1967, as the English House Condition Survey, at a time when 25% of homes lacked one or more of the following basic amenities: a bath or shower, an indoor WC, a wash hand basin, hot and cold water at

three points. By 1991, that was down to 1%, and the number of homes that fail to meet this standard now is almost too small to measure. The UK Data Service now provides access to the full series of EHS data, and two teaching datasets. We've seen more than 12,500 downloads since 2002.

In March 2018, the National Child Development Study (NCDS) celebrated 60 years of following the lives of over 17,000 people born in England, Scotland and Wales in a single week in 1958. Its '60 years of our lives' conference saw 136 delegates from 10 countries celebrating decades of research, with keynote speeches from Professor Barbara Maughan and Professor Sir Richard Blundell. There was also newspaper and TV coverage after researchers contacted some of the participants to see if they had fulfilled the childhood dreams set out in an essay, 'Imagine you are 25', written when they were 11. Since 2002, there have been over 15,600 downloads of NCDS data from the UK Data Service.



The study explored how young people experience e-cigarettes and how this may influence social norms around smoking.

## New studies and editions

During the reporting period, 225 new studies and 176 new editions (including Nesstar releases) joined the UK Data Service collection. Among these were three topical new studies, highlighting important social issues:

- The lives of young carers in England, 2015-2016 was part of a programme of research to inform the implementation of the 2015 Care Act, which required local authorities to identify and assess the support needs of young carers. The study covered carers aged 5-17 years, involving a survey of the parents of young carers and an interview with young carers themselves if they were 11 or over. It covered their access to support and services, the nature of care they provide; and the effect of caring responsibilities on their health, school life and social life.
- The disadvantaged young adults' perceptions and understandings of electronic cigarettes, 2015-2016 study used qualitative methodology to interview small friendship groups of young people aged 16-24. It aimed to explore how young people perceive and experience e-cigarettes, and how this may be influencing social norms around smoking. Smoking uptake has been declining in young adults in the UK, but concerns remain over the emergence of e-cigarettes. They offer the potential to reduce smoking in adults, but may have a different effect on young people. The study set out to explore issues such as whether e-cigarettes have a 'gateway' effect into smoking, the increased visibility of e-cigarettes, and the messages and imagery used in marketing, which often echoes that used in (now banned) tobacco advertising.
- The Department for Business, Energy and Industrial Strategy's study on behaviours and attitudes of the fuel poor in England, 2015 used a large-scale quantitative survey to examine and compare households who were likely to be fuel poor, and those who were less likely to be so. The survey examined energy usage behaviours, attitudes towards energy use, and the significance of energy costs. The survey formed part of a programme of research into the behaviours of the fuel poor, commissioned to build on the existing evidence base and inform policies to tackle fuel poverty.

Delivering  
quality  
social and  
economic  
data



# Data impact

## #DataImpact2018

In February, we held a joint data impact and networking event in London with the Office for National Statistics (ONS).

Hosted by Sarah Henry, Director for Methods, Data and Research at ONS and Matthew Woollard, Director of the UK Data Service, the event focused on how to realise the potential of research data.

A panel of data experts chaired by Robert Cuffe, Head of Statistics at the BBC, explored the data revolution and how we can respond to the social and economic challenges facing the UK. The panellists were:

**Conor D'Arcy**, Senior Policy Analyst at the Resolution Foundation, who presented a data perspective on low pay and the national living wage

**Ben Jupp**, Director at Social Finance, who focused on identifying the data analysis needed to make decisions

**Anna Leach**, Head Of Economic Intelligence at the Confederation of British Industry, who presented her analysis of understanding what drives regional productivity

**Rebecca Riley**, Director of the Economic Statistics Centre of Excellence, who focused on using economic statistics and other data to contribute to the UK economics discourse

**Emma Rourke**, Director of Public Policy Analysis at ONS, on the importance of being timely and relevant with data to enable the development of impact closer to need

**George Windsor** from TechCityUK, who presented key findings from the Annual Tech Nation report.

## Scoop.it

The UK Data Service uses Scoop.it to understand the reach of data used in research which appears in the wider media. We've collected articles and reports from across popular UK and international news platforms, showing how data in the collection have been used in research which has been reported to a wide audience. Our Scoop.it gives a varied indication of the immediate concerns of the time, with nearly 4,000 uses of the data, demonstrating how the data support the development of research focused on those concerns.

Search for: Scoop.it UK Data Service



Esmeralda Bon



Oliver Exton



Anthonia Ijeoma Onyeahialam



Matteo Sandi



Aishah Selamat

## Data Impact Fellows

In September 2017, we welcomed our second intake of Data Impact Fellows. The programme, led by Victoria Moody, Deputy Director of the UK Data Service based at Jisc, is designed to support the use of our data and resources by new generations of scholars – both through the research partnerships they develop and the students they teach during their careers.

We're charting the progress of the Data Impact Fellows 2017-2019 on our data impact blog ([blog.ukdataservice.ac.uk](http://blog.ukdataservice.ac.uk)), and here's an introduction to them and their work.

### Esmeralda Bon

Esmeralda is an ESRC-funded PhD student in the School of Politics and International Relations at the University of Nottingham, in collaboration with the Committee on Standards in Public Life, an advisory non-departmental public body sponsored by the Cabinet Office. Her research focuses on the political communication of UK public office holders and civil society actors, and the political communication of public office holders and citizens, and how this affects public opinion and, more specifically, voting behaviour. Esmeralda is affiliated with the Centre for British Politics and the Nottingham Interdisciplinary Centre for Economic and Political Research.

### Oliver Exton

Oliver is an ESRC-funded PhD student in the Faculty of Economics at the University of Cambridge. His research investigates how firms and workers respond to rapid import growth, price changes driven by fluctuations in exchange rates, and changes to EU and extra-EU trading partners' trade policy. He investigates the magnitude of firms' and workers' responses and how much they diverge.

### Anthonia Ijeoma Onyeahialam

Anthonia is a Postdoctoral Research Associate (GIS and Geovisualisation) on the GLOBAL-RURAL Project Department of Geography and Earth Sciences, Aberystwyth University. With interdisciplinary interests in globalisation and rural communities, Anthonia's research focuses on the way digital and visual technology is used to disseminate evidence in ways that influence and promote public engagement.

### Matteo Sandi

Matteo is a Research Economist at the Centre for Economic Performance at the London School of Economics and Political Science (LSE). His current fields of research are the economics of crime and the economics of education. His research looks at the impact of policing on crime, the determinants of youth crime, and the effect of discipline in school on the desire and possibility to engage in youth crime.

### Aishah Selamat

Aishah is a second-year PhD student with the Faculty of Science and Technology at Bournemouth University. Her research is co-funded by Bournemouth University and County Coaches, and aims to develop an intelligent transportation analytical model for Small and Medium Enterprises (SMEs). This involves encouraging SMEs to develop a data-driven decision-making culture, especially in digital marketing.

# The Lancet Countdown:

## Tracking progress on health and climate change

The UK Data Service is providing data for the Lancet Countdown, which aims to raise awareness of the connections between climate change and health, showing how activity to mitigate climate change can help address global health challenges. The Lancet Countdown reports annually, making recommendations for governments around the world on how to mitigate climate change.

Funded by the Wellcome Trust, the Lancet Countdown is a unique research collaboration between 24 international academic institutions and inter-governmental organisations. They include Steve Pye, Principal Research Associate at the Bartlett School of Environment, Energy and Resources, University College London (UCL), Dr Nick Watts, UCL Institute for Global Health, Executive Director of the Lancet Countdown and Dr Ian Hamilton, Reader in Energy Epidemiology, Bartlett School of Environment, Energy and Resources.

Together, the participants track 41 indicators across five key domains in health and climate change. The UK Data Service provides International Energy Agency data which the Lancet Countdown uses for indicators in working groups on mitigation actions and health co-benefits, and finance and economics.

Following the release of the 2018 Lancet Countdown: Tracking Progress on Health and Climate Change report, the Global Climate and Health Alliance called on world leaders to take the actions required to limit global warming to the targets set by the 2015 Paris Agreement on Climate Change: making the greenhouse gas reduction

commitments between now and 2020 that will limit global warming to 1.5° C. Genon K. Jensen, Executive Director of the Health and Environment Alliance, also called for action on air pollution, calling it “the number one environmental threat to people’s health, and one of the top five risk factors for chronic diseases like heart disease, cancer and asthma.” She added: “World leaders would give clean air and disease prevention a major boost worldwide by taking on the Lancet Countdown health recommendations”.

“ World leaders would give clean air and disease prevention a major boost worldwide by taking on the Lancet Countdown health recommendations. ”

---







## International engagement

The UK Data Service is at the heart of CESSDA, and heavily involved in its work both operationally and strategically. In June 2017, CESSDA became a European Research Infrastructure Consortium (ERIC), with the UK Data Service one of its designated service providers. CESSDA ERIC aims to provide a distributed and sustainable research infrastructure to enable the research community to conduct high-quality research in the social sciences contributing to the production of effective solutions to the major challenges facing society today and to facilitate teaching and learning.

### CES2018

The CESSDA Expert Seminar 2018 (CES2018) took place in Ljubljana in September 2018, chaired by John Shepherdson, CESSDA Platform Delivery Director and leader of the CESSDA Technical Working Group, from the University of Essex. There had been significant progress since the previous meeting in 2016, with CESSDA's high quality Technical Infrastructure now ready to go. The new CESSDA website and Data Catalogue went live at the beginning of November 2018.

### VOICE - Controlled Vocabularies

Over the last year, the Vocabularies in CESSDA (VOICE) project has produced, tested and evaluated the first version of the CESSDA Vocabularies Service (CVS), and made it available to service providers. Previously known as the CV Manager and Vocabularies Manager, CVS is a tool to manage CESSDA ERIC's multilingual content. A public version is due as this Report goes to press.

### VOICE - Thesauri

The European Language Social Science Thesaurus (ELSST) project officially ended in September 2017. Funded by the ESRC, the project aimed to bring together and develop the ELSST and Humanities and Social Sciences Electronic Thesaurus (HASSET), improving the experience for users, and updating content. During the lifetime of the project, four languages were added to ELSST: Czech, Lithuanian and Romanian in 2015 and, in 2017, Slovenian.

The project secured further funding for 2018 through the VOICE project, and the UK Data Service has again been leading, with FSD, the Finnish Social Science Data Archive, and GESIS, the largest German infrastructure institute for the social sciences, as key partners. Together, they have been continuing to develop content extending the project to cover more languages.

Reach



Delivering  
quality  
social and  
economic  
data

We've been leading a collaborative group across Europe to help researchers discover and make the most of data.

### Euro Question Bank

The UK Data Service has been working with other European archives on the CESSDA Euro Question Bank project (EQB). EQB will introduce a central search facility across all CESSDA's survey holdings to help users find survey questions in different languages. There are useful tools, too, such as the ability to search for a specific concept, and switching languages to view translations of the same question. The EQB is likely to be useful for survey designers, survey translators, data analysts, researchers, and undergraduate and postgraduate students and teachers.

### Metadata management

CESSDA's Metadata Management project (CMM) plans to develop, promote and implement a standardised metadata design, content and practice for all CESSDA data assets. In 2017, CESSDA's project group agreed version 1.0 of CMM, including the core metadata model, multilingual controlled vocabularies, and a renewed version of CESSDA's topics classification. CMM 1.0 supports the building of CESSDA's products and services like the Catalogue and the Euro Question Bank.

### Data Discovery training

The UK Data Service has been leading a collaborative group across Europe on training to help researchers discover and make the most of the datasets available as part of its work with CESSDA. Data Discovery training aims to alert researchers to the datasets which are available and to help them understand complicated datasets and their potential. During 2017, the Research Data Management subgroup developed an online training module for social science researchers to teach basic data management knowledge and skills with extra offline content for use in face-to-face training.



Our staff visited Kyrgyzstan to provide training on sharing data for research.

## IDAN

The UK Data Service is part of the International Data Access Network project, founded in 2018. IDAN is a collaboration between six Research Data Centres (RDCs) in France, Germany, The Netherlands and the UK, which aims to allow researchers in any of the RDCs to work remotely on controlled access data provided by partner countries. This will reduce the need to travel, and the associated costs with international research, and helps to address the challenges that researchers currently face when trying to access data from outside the country they are working in. Matthew Woollard and Beate Lichtwardt were leading the work for the UK Data Service, and in October, we announced that – in cooperation with the French Secure Access Data Centre (CASD) – we had set up a secure data access point in our Safe Room at the University of Essex to facilitate on-site direct remote access to CASD data. This follows the successful establishment of an access point for international controlled data from the German Institute for Employment Research.

## Training in Kyrgyzstan

In March 2018, Louise Corti and Veerle Van Den Eynden from the UK Data Service visited Kyrgyzstan to provide training on research data sharing. They went to support the creation of a large collection of sensitive interviews, case studies and focus group discussions from the national Gender in Society Perceptions Study, following an inquiry by the United Nations (UN) Women Country Office in the Kyrgyz Republic a year ago. Veerle and Louise gave advice on anonymisation and data sharing, and the possibility of using the UK Data Service as a place of deposit, as Kyrgyzstan has no fully fledged data repository. They also ran training to advance the country's data management and sharing capacity, organised by the Principal Investigator of the study, Meghan McCormack, a law lecturer at the American University of Central Asia, and supported by AMICAL (American International Consortium of Academic Libraries).

## SERISS

We continue to be closely involved in Synergies for Europe's Research Infrastructure in the Social Sciences (SERISS), which aims to provide solid socio-economic evidence to help countries across Europe tackle challenges such as youth unemployment, ageing populations, and transport and housing issues.

The UK Data Service is contributing to a project looking at 'New forms of data – legal, ethical and quality issues'. This is addressing the major legal and ethical challenges facing cross-national social science research which relies on access to large-scale data on an individual level, including social media data, administrative data and biomarkers.

These new forms present new curation challenges, especially around legal and ethical issues and what are known as the Vs of 'big data': volume, velocity, variety, veracity and value. As a leading expert organisation on curating complex data for research, the UK Data Service is helping to define workflows for archives as we prepare for these new and novel forms of data.

June 2018 saw workshops in London on:

- combining social survey data with social media data in surveys, which considered subjects such as storage and consent
- combining survey data with administrative data, again looking at consent, but also at technical challenges such as linkage.

## Is being a carer linked to obesity?

A study published in July 2018 found that women who were carers were significantly more likely to have larger waists or a higher percentage of body fat than men who had caring responsibilities. The research found a particularly strong effect when it came to women who combined caring with full-time work: among women aged 16-44 who were in full time work, those who were carers had waists 4cm larger, on average, than those who were not carers. The study also found that carers working beyond the age of 65 were also larger, with higher proportions of body fat and bigger waists.

Millions of people in the UK combine work with caring responsibilities – looking after an older relative, a disabled child or a partner, for example. Using data from Understanding Society (the UK Household Longitudinal Study), curated and accessed through the UK Data Service, Rebecca Lacey and colleagues from the ESRC International Centre for Lifecourse Studies at UCL analysed obesity among UK carers from the age of 16 onwards.

Rebecca, a Senior Research Fellow in social epidemiology at UCL, says: “The strains on female carers may be greater – younger women may be working outside the home and may also be caring for children, for instance. They may lose touch with friends who have more freedom to enjoy leisure time. Carers are likely to have less time to spend exercising or to prepare healthy and nutritious meals – and those combining several caring roles with working life are likely to be particularly hard-hit.” There are more than seven million carers in the UK – around one in five adults – and about 60% of them are women. They have been estimated to save the UK economy £132 billion per year. Rebecca says: “We need to make this largely hidden army of unpaid carers a public health priority.”

“ There are more than seven million carers in the UK - around one in five adults - and about 60% of them are women. They have been estimated to save the UK economy £132 billion per year. ”

---







Case Studies

Business & Economy



Measuring economic impact: indicators

Effect of policy decisions on the

understanding the effects of

See more

Environment & Sustainability



Tracking progress on health and climate

On higher energy prices affect operational

Concerns, indicators to measure

See more

Health & Wellbeing



Measuring health inequalities in the UK

Individual and area based

Prevalence of mental health

See more

Learning & Skills



Science & Technology



Social Policy & Communities



We hold data that unlock new discoveries in research, provide evidence for policy decisions and teach the next generation in core data skills.

# Trusted expertise

## CoreTrustSeal

CoreTrustSeal launched in September 2017, promoting sustainable and trustworthy data infrastructures by setting out requirements for trusted data repositories. CoreTrustSeal replaced the Data Seal of Approval (DSA) certification, and was the fruit of a collaboration between DSA and the World Data System under the umbrella of the Research Data Alliance to merge their certifications for data repositories. CoreTrustSeal is now being considered as one of the conditions for involvement in the European Science Cloud – and for FAIR (findable, accessible, interoperable and reusable) data certification. The UK Data Service helped to develop the new system through a Research Data Alliance working group, and Hervé L'Hours, our Repository & Preservation Manager, based at the University of Essex, is now vice chair of the international board managing it.

## New challenges

The UK Data Service has been building new infrastructure to meet the challenge of 'big data'. New and novel forms of data bring with them changed expectations from social scientists, and major new requirements for data services and repositories. We need to ingest sizeable streams of real-time data, and enable exploration and linkage of a variety of data assets. We're developing an infrastructure to manage new and novel forms of data using the increased computational and storage capabilities of platforms like Apache Hadoop. Hadoop sits at the core of our new infrastructure, currently known as Data Service as a Platform (DSaaP).

From October 2017, new core technical work has been underway to allow us to update our current data service to the new platform. We delivered a proof-of-concept DSaaP infrastructure in 2017, and in 2018 have been integrating extra Hadoop components and other relevant open-source tools, with the aim of rolling out a production-ready unified data infrastructure in late 2019.



UK smart meters have the potential to generate over a trillion records per year.

### Smart Energy Research Lab (SERL)

The Engineering and Physical Sciences Research Council (EPSRC) has put £6m into a five-year project to develop a Smart Energy Research Lab (SERL) to help UK researchers get access to good quality energy data.

The project, launched in 2017, will run to August 2022, and involves the UK Data Service team at the University of Essex in partnership with UCL, alongside Cardiff, Edinburgh, Leeds Beckett, Loughborough and Southampton universities and the Energy Saving Trust.

The aim is to provide a secure, consistent and trusted channel for researchers to access high-resolution energy data, and make innovative energy research possible. The data will be collected on a strictly voluntary basis, only with the informed consent of households who have agreed to provide their smart meter data for research. Only accredited researchers will have access to controlled datasets using best practice data governance protocols such as the '5 Safes'.

UK smart meters have the potential to generate over a trillion records per year – a challenge the UK Data Service's Big Data Hadoop platform can meet. SERL energy data will be enhanced by linking it to relevant contextual data covering dwelling, household and neighbourhood attributes to facilitate innovative primary and secondary data analysis.

The research conducted via the SERL research portal has the potential to:

- support government policy
- kick-start the development of new products, services and energy markets
- help provide solutions to the energy 'trilemma' – security, affordability and environmental sustainability
- make better research possible by developing best practice guidelines and methods to improve data security and enable innovative uses of smart meter data.

# Is cycling safe?

The popular perception of cycling in the UK is that it is extremely dangerous – but Dr Jennifer Mindell, Professor of Public Health at UCL, and her colleagues have carried out a series of studies using National Travel Survey data from the Department of Transport, curated and accessed through the UK Data Service, which show that the picture is more complicated than that.

They began by examining death rates by age group and gender for different modes of transport, and found that although death rates had fallen substantially, they were higher in men than women for driving and walking – and overall.

For drivers, the death rate was highest in men aged 17-20 and in those over 80 – with the fall between those ages giving the graph a U shape – whereas it simply increased with age for walking and cycling. Death rates for pedestrians increased with distance travelled and for cyclists by time spent travelling.

The research pointed out another fairly crucial difference between modes of travel: cyclists and pedestrians rarely kill other road users. For example, 99.4% of pedestrians fatally injured in a road travel crash are killed by motor vehicles.

So, is cycling dangerous?

In 2010-2012, there was one cyclist death for approximately every 50 million km cycled. Although fatal head injury rates per million hours travelled were slightly higher for cyclists than pedestrians, there was one fatal head injury for every 6.25 million hours cycled by men and 10 million hours cycled by women in 2005-2013.

Jennifer says “One of our key findings from using National Travel Survey data was that young men are

up to five times safer when they cycle than when they drive – and the rest of us are also safer if they cycle and don't drive!

“Deaths of cyclists are very rare – so rare, in fact, that they get a lot of media coverage – but the UK has almost the lowest death rates for car drivers worldwide, so everything else is bound to look higher.”

On average, five people are killed every day on British roads – most of whom are car users (46%), followed by pedestrians (25%) and motorcyclists (18%).

Jennifer adds: “Bald statistics can only tell us so much, though – a UK Minister of Transport once said cycling is safer here than in the Netherlands, citing death rates per 1,000 people, ignoring the fact that cycling accounts for about 2% of commuting trips in the UK but about 25% in the Netherlands. What's important is context, and the National Travel Survey data allowed us to consider that.”

“ One of our key findings from using National Travel Survey data was that young men are up to five times safer when they cycle than when they drive. ”

---





Advocacy and collaboration

We look forward to working with the new Digital Economy Act and the opportunities it brings.



## Working together

### GDPR

Like many organisations, the UK Data Service has been immersed in discussions about implementing the General Data Protection Regulation (GDPR), which came into force in May 2018. Our focus has been on GDPR's impact on research. In the last year, we have been involved in discussions with the Information Commissioner's Office, ESRC and large longitudinal studies to make sure GDPR does not have a negative impact on research.

We have also been proactive in dealing with personal data we hold as part of our customer information systems. We have altered many of our internal policies and procedures to comply with the updated legislation, and updated our Depositor Licence, End User Licence and Privacy Notices before it took effect.

We have provided guidance and training to researchers to enable them to understand their responsibilities under GDPR, but also so that data can still be archived and shared at the end of a research project. We have run workshops for researchers, data managers, businesses and institutions across the UK and abroad, about how to manage and share research data under GDPR, and provided tailored training for a number of universities and institutes. Some of these events were aimed at specialist audiences, including researchers working in legal research, and anthropology, or covered specific issues such as how to gain research participants' consent for sharing data in line with the new regulation.

We have also been involved in collaborative projects to look at the impact of GDPR on data archiving across the EU through our membership of CESSDA. Furthermore, we have created a template model consent form and information sheet to help researchers comply with GDPR and gain informed consent for ethical and other legal reasons (such as the duty of confidentiality).

### Working with data owners

We have worked very closely with the Office for National Statistics in the last year, providing advice on implementing the Digital Economy Act (DEA) accreditation of processors. We have been liaising with them, and user testing their new shop front for the forthcoming launch of their upgraded Secure Research Service (which replaces the Virtual Microdata Laboratory). We are also looking forward to new administrative data products that will be of great benefit to the research community, and hope soon to be accredited as a DEA Data Processor.

### CLOSER

The UK Data Service plays an active role in the joint ESRC/Medical Research Council (MRC) CLOSER project. Members of our staff sit on the Leadership Team, the International Advisory Group and Discovery technical group. This last year has seen us contribute to ongoing work on citation strategies for longitudinal and cohort data, provide advice on the new catalogue for international data resources, and help data owners enter their question and variable metadata into the metadata portal. We have also fed into the working group on onwards sharing of NHS data linked to cohorts, contributing to arguments for why this is important, and on the kinds of legal arrangements needed to enable this.

### Linked administrative data

In 2018, we fed into the review of accreditation criteria for the role of processor under the Digital Economy Act, particularly with regard to technical requirements for safe settings, including ESRC Safe Pods. We look forward to working with the new Digital Economy Act and the opportunities it may bring for research access to administrative data.

## admap

Matthew Woollard and Beate Lichtwardt – Welcome and Introduction, UK Data Service Involvement in the International Data Access Network (IDAN)

Introduction: Talks from three IDAN Partners and representatives of International Research Data Centres (RDCs) across Europe:

Marine Silbermann, French Secure Access Data Centre (CASD)

Anna Müller-Gorman, Research Data Centre (EDZ) of the Federal Employment Research



# Training and knowledge exchange

## Training events and materials

We have organised and participated in a wide range of training events – over 170 workshops, live webinars, conferences and presentations at a range of venues with different audiences. Our annual user conferences bring researchers and data producers together to share research and connect. Innovations over this period included collaborative training events with CESSDA, our new interactive Data Skills Modules, GDPR training, and new training and support for Secure Lab users.

## User conferences

Every year, the UK Data Service runs four one-day user conferences to bring together producers and users of survey data available from the service. These are organised with data producers such as the Office for National Statistics, NatCen, and UK government departments. The conferences feature producer presentations and research talks, and attract speakers and attendees from all sectors including government, businesses, charities and academia. These conferences are a great opportunity for networking and feeding back user views to data producers. Over this period, as well as our annual user conferences covering health studies, crime surveys, family finance surveys and the Labour Force and Annual Population Surveys, we ran a user conference to bring together UK census producers and users.

## CESSDA collaboration

In spring 2017, the service began a series of webinars and workshops about finding and exploring data in Europe in collaboration with the European Data Archives in CESSDA. The webinars introduced European data on political behaviour, ageing, and migration – and a one-day workshop introduced attendees to data about European attitudes and values. We ran a workshop at the University of Essex Summer School in July 2018 about finding and exploring data in Europe, and a webinar on the new CESSDA Research Data Management module in February 2018.

## Data Skills Modules

In January 2018, we launched three new interactive online training modules as part of a suite of online training materials led by Vanessa Higgins, based at the University of Manchester. These innovative on-demand modules introduce key aspects of survey, longitudinal and aggregate data and have been designed with maximum flexibility for the user in mind. With about 1.5 to 2 hours of materials in each module, users can complete the modules in their own time, dipping in and out when they like.

Each module uses a mix of short videos, written materials and practical data activities, using open source software where possible, to introduce important aspects of the data to anyone who wants to know more about getting started with data. The modules also include interactive quizzes to encourage users to check their understanding as they go along, and a glossary to explain the terminology used.



There is a new process for the Secure Lab users, making it more efficient for users and staff.

### Safe Researcher training

The User Support team has been working with ONS on the project to phase out SURE training and replace it with Safe Researcher training. The new approach, prompted by the Digital Economy Act, will see users who want to access the most sensitive data undergoing more attitudinal training than they have had in the past. They will be asked to reflect on their own behaviour while carrying out their work, rather than simply following rules. Safe Researcher training is based around the '5 safes' principles, originally devised by the UK Data Service, and we have been helping ONS to refine the new training over the past year as it has evolved based on feedback from trainers and stakeholders. We expect to see the new programme implemented across all the service providers – the UK Data Service, ONS, HMRC and Administrative Data Research UK – in the next few months.

### Secure Lab checklist

In summer 2018, the User Support team also rolled out some new processes designed to give users more responsibility and autonomy in their work. There is now a checklist for users of the Secure Lab which makes the process more efficient for users and staff. For example, researchers are now prompted to cite datasets they've used just as they would an article, book, image or website. The checklist has seen good results, with fewer users having to be reminded about citation, and more of them using digital object identifiers (DOIs). Another change in 2018 is that researchers are no

longer asked only to submit finished outputs. Now, outputs can be submitted before they are 'journal ready' - as long as there is enough information for the User Support team to make an assessment. This greater flexibility is making it easier for researchers who might be collaborating on research, where they may need to discuss any changes which are required – and feedback has been good.

### User feedback

During the period of this report, we've had comments, emails and tweets from users, people we've trained and conference attendees praising our work as "excellent", "thoughtful and detailed", "helpful and approachable" and "highly recommended". On our Secure Lab training, one user answered the question 'What could we do better?' by saying, "Nothing - it was brilliant."

We are always looking to improve the user experience and welcome constructive feedback from service users.

# Can culture prevent dementia?

Researchers at UCL have used the English Longitudinal Study of Ageing, curated and accessed through the UK Data Service, to explore whether engaging with the arts can protect people against cognitive decline and the onset of dementia as they get older.

The research considered whether the arts can also contribute to 'cognitive reserve' and thereby protect people by helping to prevent mental decline and dementia.

The researchers, led by Dr Daisy Fancourt, Wellcome Research Fellow in Behavioural Science and Health at UCL, analysed data from 3,445 adults aged over 52 over 10 years (2004-14), focusing on:

- visiting galleries, museums or exhibitions
- going to the theatre, concert or opera
- going to the cinema.

They found associations between the first two points and smaller declines in memory and verbal fluency 10 years later, but no clear associations between going to the cinema and cognition. Their analyses were all adjusted for baseline cognition and a range of other demographic and health-related factors – and for engagement in other social or stimulating activities.

They followed these results up by looking specifically at dementia and visiting galleries, museums or exhibitions, because that produced the strongest result in the first study. This time, they focused on 3,911 adults who were free from dementia at the beginning of the 10-year period.

Again, they found a relationship between cultural engagement and a lower risk of developing dementia, with a 49% lower risk amongst those who

attended every few months or more compared with those who never attended. As in the first study, they controlled for a range of factors.

Daisy says, "Cultural engagement is mentally stimulating, it requires light physical activity which could reduce the negative effects of sedentary behaviour, and it can reduce isolation by encouraging people to leave their homes and meet people. So, it could protect us against cognitive decline and the onset of dementia. Now, we need to see studies that look at older adults at risk of cognitive decline or dementia, and which point them to cultural activities, to see if we can confirm causality. There are museums across the country – the world, in fact – and they reach a range of demographic groups, so they have huge potential for public health interventions."

“ They found a relationship between cultural engagement and a lower risk of developing dementia. ”

---





We have been working with the University of Edinburgh on an open-source tool for researchers to check data quality.



As this Annual Report was going to press, the UK Data Service was working on a number of projects to keep giving our users seamless and flexible access to a wide range of data resources to facilitate high quality social and economic research and education.

As part of this, for example, we launched a campaign on the importance of citing data correctly, to try to encourage students, researchers and academics to think of a dataset they use in the same way they would a book, article, image or website. We will be rolling the campaign out nationally in 2019, helping users to understand the importance of citation in their own careers, in supporting the reproducibility of research findings, and in showing funding bodies the value of data.

### QAMyData

Since January 2018, specialist staff in the UK Data Service, led by Louise Corti from the University of Essex, have been working with AQMeN (Applied Quantitative Methods Network at the University of Edinburgh) to develop an open-source tool which researchers can use to check the quality of their data before they upload their data to a repository. The aim is that the user will get a report or 'data health check' that identifies the most common problems in data submitted, and suggestions, including code, for fixing errors or cleaning the data.

Researchers will thus be able to demonstrate quality assurance when submitting their evidence to a journal or data repository – helping everyone in the social science community to improve transparency and replicability. We're working on presentations and training to back up the tool's launch, and there'll be full details in our next Annual Report.

### 2019 and beyond

As we look forward to the rest of 2019, we aim to keep delivering our quality service – not simply providing researchers with access to the UK's largest collection of social, economic and population data, but doing it in order to inform great research, and thus to play our part in building a stronger society and improving people's lives.

We will continue to review and improve our trusted services and endeavour to enhance the user experience across the UK Data Service.

# About the UK Data Service

The UK Data Service is delivered by staff with specialist expertise in research data based at higher education institutions across the UK:

## UK Data Archive, University of Essex

Lead organisation of the UK Data Service and an internationally acknowledged centre of expertise in acquiring, curating and providing access to data. The Archive has been providing coordinated data service infrastructure to the ESRC for 50 years.

## Jisc

A registered charity which champions the use of digital technologies in UK education and research. Jisc hosts a significant number of the UK's research information assets and builds applications to help people make the most of these resources.

## Cathie Marsh Institute for Social Research (CMI), University of Manchester

A centre of research excellence specialising in the application of advanced quantitative methods in an interdisciplinary social science context. CMI plays a key role in the support of microdata from UK Censuses and Government Surveys.

## Geography and the Environment, University of Southampton

An outstanding reputation for leadership in population and health research, based on the combination of innovation in methodology covering geographical information systems (GIS), spatial analysis and quantitative and qualitative methods with substantive expertise in census, population and health issues.

## EDINA, University of Edinburgh

The ESRC's specialist geography unit for the UK census, supporting UK Data Service users in accessing and using the geography outputs for the census.

## Department of Information Studies, University College London


An international centre for knowledge creation and transfer in the fields of librarianship, archives and records management, publishing, information science and digital humanities.

## Centre for Advanced Spatial Analysis, University College London

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.

# Contact

UK Data Service  
University of Essex  
Wivenhoe Park  
Colchester  
Essex  
CO4 3SQ

[ukdataservice.ac.uk](http://ukdataservice.ac.uk)  
 [@UKDataService](https://twitter.com/UKDataService)  
[comms@ukdataservice.ac.uk](mailto:comms@ukdataservice.ac.uk)

