UK Strategy for Data Resources for Social and Economic Research 2013-2018

A five-year plan to inform and guide the development and utilisation of data and related resources for social and economic research

Contents

PREFACE

This view, which the UK Data Forum strongly endorses, has underpinned the UK Strategy for Data Resources for Social and Economic Research since its first days and remains as strong as ever as we here present the third version of the strategy, covering the period 2013-2018.

High quality empirical social science needs high quality data sources and associated infrastructure. The strategy presented here sets out a vision and practical actions for what must be done to ensure that these are maintained and improved. The signatories to the strategy come from key producers of social and economic data and funders of data resources and the strategy itself is a statement of intent for these organisations to work together to achieve the aims. Given the necessarily fractured nature of strategic decision making across a range of Government departments, Research Councils and Charitable Foundations that support empirical social science research, this commitment to work together is vital.

This is a good time to produce a new strategy. The environment for data availability is changing rapidly and the next few years have exciting potential. There is growing interest in the interactions between social, economic and medical phenomena for example and the Research Councils, in particular, need to develop programmes to promote more inter-disciplinary work.

Recent Government approaches to open data have had a major impact on data availability and there is optimism that much better access to administrative data sources will yield major benefits. A number of reports cited in this document contain common themes and all are pushing to achieve a significant shift in the range of data available. The report from the Administrative Data Taskforce makes recommendations that have the potential to overcome some of the severe barriers that exist. The social science community has put in place a variety of data access mechanisms to ensure that sensitive data may be accessed in secure settings. If access to administrative data through mechanisms like these can be achieved, it will yield significant public benefits while protecting the confidentiality of the data subjects.

In general the UK has a strong array of data sources and in some areas, such as longitudinal data, we are international leaders. However empirical research which makes use of data about organisations and their customers is a relatively undeveloped area. We see it as a strategic aim to strengthen the availability of data sources to support more research in this area. Also there are rich sources of novel data that are

" We are committed to the principle of the Government basing its policy on evidence and social science based evidence needs to be just as robust as that from any other science.

Parliamentary Select Committee on Science and Technology, September 2012

becoming available, arising from the growth of electronic communication, from internet usage, the growth of web-based services and from the expansion of the digital economy. Such data have potential research value and will become ever more pervasive. We need a strategic approach to the development and exploitation of these data.

The UK Strategy for Data Resources for Social and Economic Research is owned by the members of the UK Data Forum and we would like to thank all those who have contributed to it. In particular we are indebted to Peter Elias and John Hobcraft who have been the driving force behind its development. This new version of the strategy builds on the two that have gone before. As we looked back at each of these we saw that much of what we committed to was achieved. I very much hope that the same is true when we look back in 2018 and I look forward to working with our members to ensure that this is so.

Tim Holt
Chair, UK Data Forum

1 INTRODUCTION
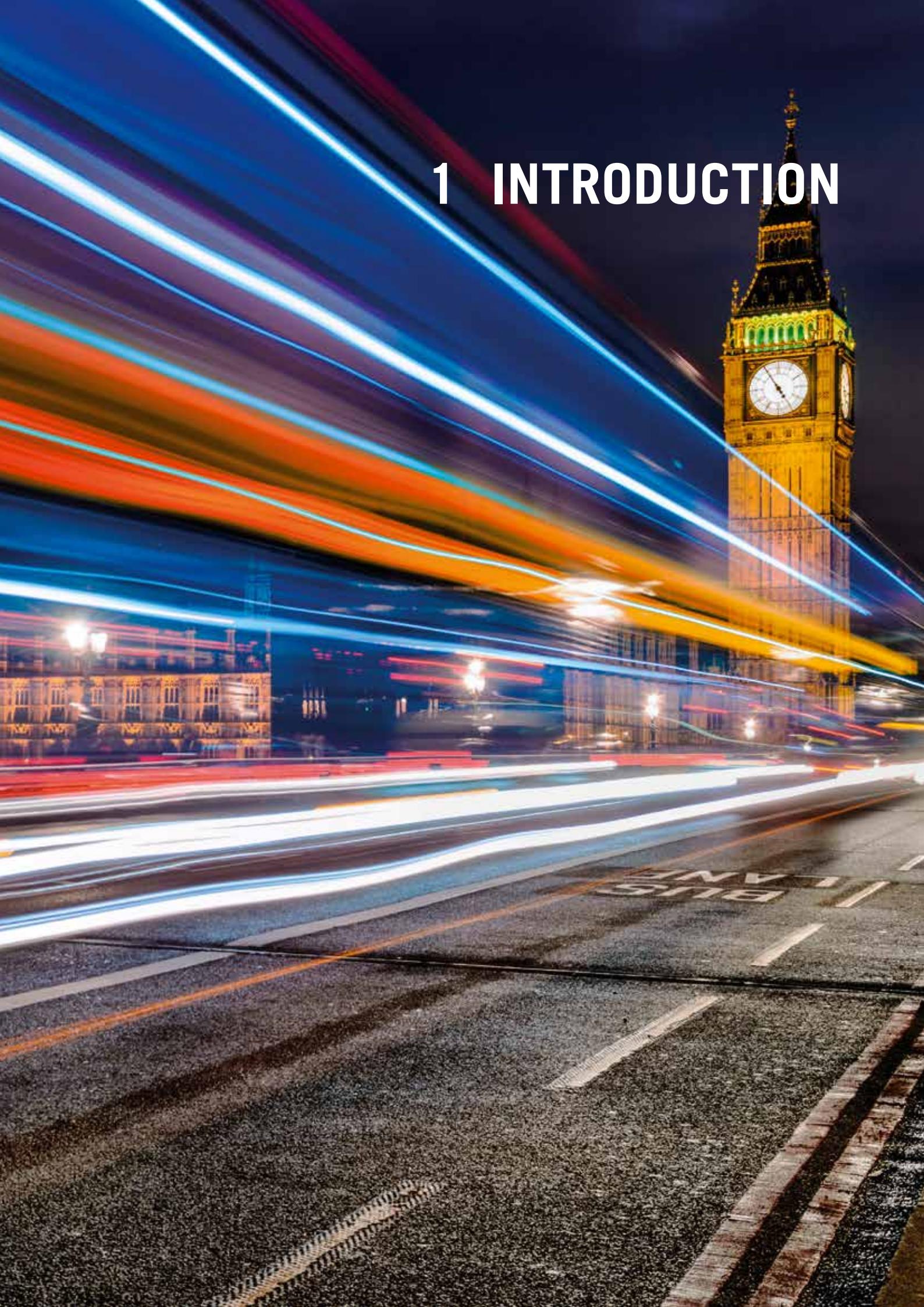
1.1 Aim of the strategy

The UK Strategy for Data Resources for Social and Economic Research is a plan designed to facilitate and coordinate the development of data resources for research across the social and economic sciences and at their interface with other areas of scientific enquiry. This is the third version of the strategy, looking ahead over the period 2013 to 2018. As with the earlier versions, it seeks to identify those areas where developments in the research data infrastructure need to be prioritised in order to meet future research needs. Unlike its predecessors, this version makes recommendations which relate to the capacity of the social science research community to take advantage of the recent and planned future developments in data resources. It reflects the need to provide data resources for research in the most efficient manner possible, whilst maintaining and enhancing the well-deserved reputation UK has for facilitating world class social and economic research.

1.2 Revising the strategy

The UK Data Forum[1] has oversight of the strategy. The forum draws its membership from most of the public sector organisations in the UK involved in the production of data with value for social and economic research. It includes representatives from bodies which fund the development and maintenance of such research data resources. It has links with private sector organisations which hold large data collections and it includes representatives responsible for data service infrastructure (providing access and value added services to data). Through its continued monitoring of new developments both in the UK and more widely, members of the UK Data Forum have helped to identify those areas where further collaborative efforts should be made between data producers and research funding agencies to gain the maximum benefit from existing data. Additionally, and where appropriate, the forum works to ensure that new types of data are available for research in a timely and cost-effective manner.
Other sources that have informed the development of the UK Strategy for Data Resources for Social and Economic Research include the public consultation on research infrastructure needs held by RCUK[2] and similar consultations with researchers and data producers held in Germany and the USA[3][4].

1.3 The changing landscape for social and economic research

Over the past eight years the scope, scale and importance of social and economic research has been transformed. In many research fields such as epigenetics, biomedical and the environmental sciences, social and economic events and circumstances are recognised as causal, consequential or as moderating factors. New and large-scale research data collections have been established, some as longitudinal resources with a life course perspective. Improvements in data access have been achieved and will continue to develop. Major investments have been made in the development of tools and methods for research and associated training.

Alongside these developments, new challenges have arisen over the past few years, and these will continue to shape the research landscape. The global economic downturn highlights the interdependence of economic activity. The blossoming of social media as new forms of communicating and networking continues, and their behavioural impact is, as yet, largely unexplored. Technological progress is reshaping the human/computer interface in ways which are proceeding at an unprecedented pace. Transparency in both government and private sector organisations, revealing data which help shape decision-making, is rapidly replacing the opacity within which many previously conducted their business.

These developments are creating exciting new opportunities for path-breaking research in the social and economic sciences, and at their boundaries with other scientific disciplines. But they also present challenges for research funders, data infrastructure providers, researchers and the policy community to ensure that these opportunities are grasped.

1.4 Structure of the strategy

The UK Strategy for Data Resources for Social and Economic Research is presented in six short sections. Following this introduction, section two sets the scene for this version of the strategy. It reviews progress to date, reflects upon the extent to which the aims and ambitions expressed in the previous version have been achieved and outlines the moves towards a more explicit approach to data sharing, both within the UK and internationally. Section three brings together information on research needs with developments in the data infrastructure, identifying areas where further coordination and new initiatives will benefit both data producers and users. Section four considers a critical dimension of the strategy – the need to ensure that there is sufficient capacity within the social science research environment to take advantage of the new and proposed developments in data resources. Section five scans the horizon of research in the social sciences to identify issues that are likely to shape future data developments. Finally, the strategy concludes with a broad action plan; identifying those activities that should be achieved within the next two years and those which will move forward on a longer timescale.

---

[1] See www.ukdataforum.org
[2] See RCUK (2012) *Investing for Growth: Capital infrastructure for the 21st Century*.
[3] German Data Forum (ed.) (2011) *Building on Progress: Expanding the Research Infrastructure for the Social, Economic and Behavioural Sciences*. Opladen, Germany: Budrich University Press.
[4] National Science Foundation, Directorate for Social, Behavioural and Economic Sciences (2011) *Rebuilding the Mosaic: Fostering Research in the Social, Behavioural, and Economic Sciences at the National Science Foundation in the Next Decade*. Arlington, VA: National Science Foundation.
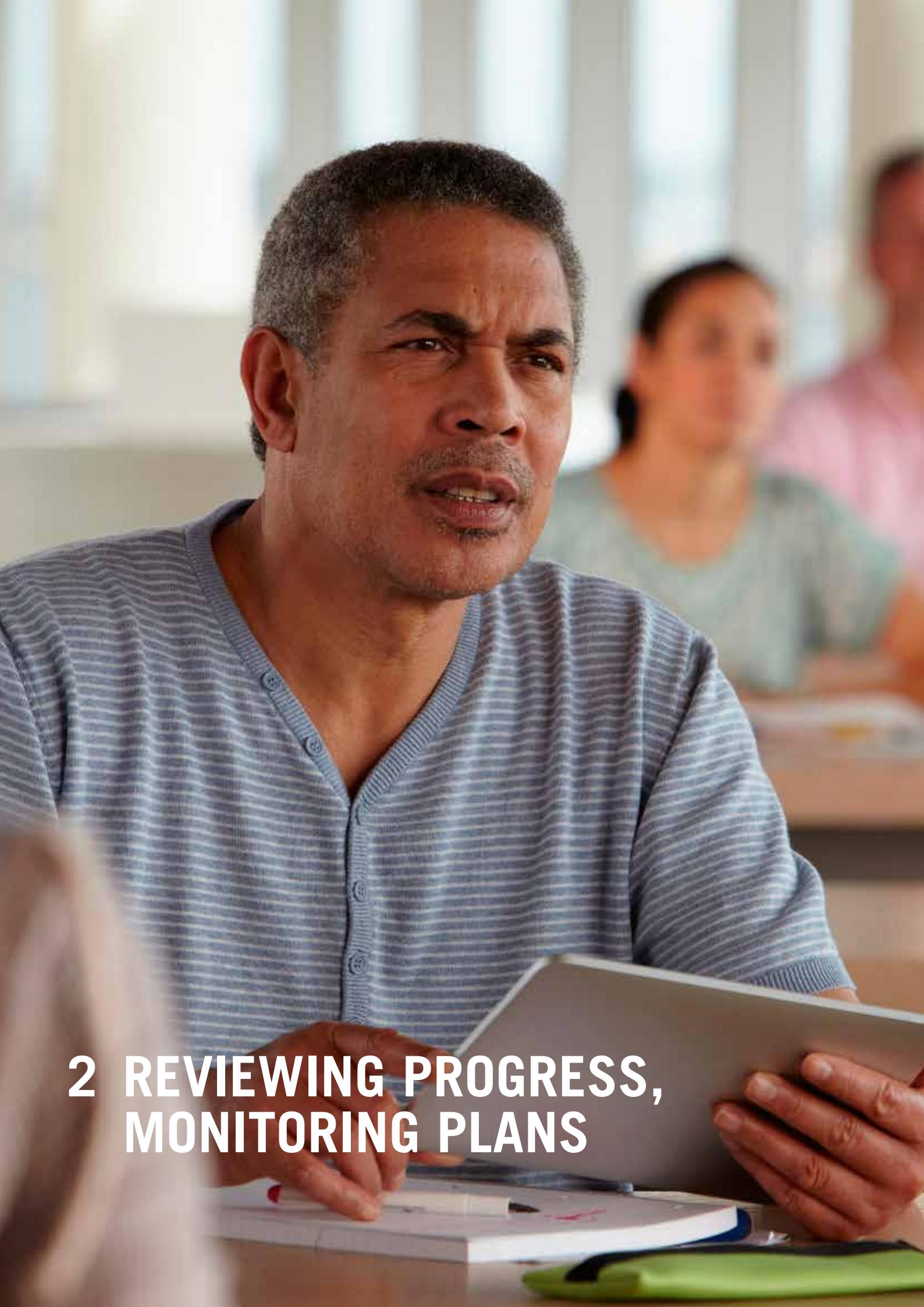
2 REVIEWING PROGRESS, MONITORING PLANS

The previous *UK Strategy for Data Resources for Social and Economic Research*, covering the period 2009-2012, put forward a six point action plan for new or renewed efforts to carry forward the strategic development of data resources across the social sciences. Progress on the implementation of this plan is reviewed, indicating where further monitoring by the UK Data Forum is required.

2.1 Sustaining a strategic approach to the development of data resources

Work has been undertaken throughout the period 2009-2012, supervised by the UK Data Forum, to inform this revision. Government departments and research funding agencies have established mechanisms through which their long run plans can be discussed and, where necessary, coordinated to ensure maximum benefit for the research community.

The important role played by the *UK Strategy for Data Resources for Social and Economic Research* was recognised in a recent report from the House of Commons Science and Technology Committee[5] in its consideration of the use of information generated in the decennial Censuses of Population and any future alternative form of data collection. The Committee indicated its view that the strategy should play a prominent part in the long term provision of data for the social sciences. In taking forward the views of the Science and Technology Committee, the UK Data Forum will consider this issue, particularly with respect to the manner in which it should respond at the strategic level to government consultations and Parliamentary Committee calls for evidence.

A number of reports were published in 2012 which have an important bearing upon the shape and direction of the strategy presented in this document. These are shown in the box opposite. These reports all point in the same direction, towards an environment where the sharing of data for research purposes is regarded as an essential part of the research process, and where resources are directed towards this end. In line with these reports, the Cabinet Office published a White Paper in 2012[6]. Among its many recommendations, this position paper states:

'The Government Social Research Service is currently working on an initiative to put policy and practice in place that will enable routine archiving of research data, and its release to bona fide researchers'. (para. 2.67)

The UK Data Forum welcomes this initiative which will secure a vital link between the UK Data Service and those government departments and agencies commissioning

Recent reports with relevance for the UK Strategy for Data Resources for Social and Economic Research

Science as an Open Enterprise – Royal Society 2012
The main thrust of this support was that transparency and openness should characterise all scientific research. As a major part of this, data sharing should be regarded as the norm and researchers, their funders and research institutions should adopt this stance in all their research activities. An important recommendation relates to situations where data hold personal information. In such cases, appropriate safeguards should be put in place to prevent disclosure of such details whilst facilitating data sharing.

New data for Understanding the Human Condition: International perspectives – OECD Global Science Forum Expert group 2012
The focus of this report was on the need for global collaboration over data sharing. This will require improved incentives for researchers who agree to share data, and the adoption of agreed standards and protocols for data description. Additionally, the report calls for an international approach to the use of 'Big Data' for research, covering collaboration over the exploration of the research value of new formats of data, the development of tools for their analysis and improved access to administrative datasets on a cross-national basis.
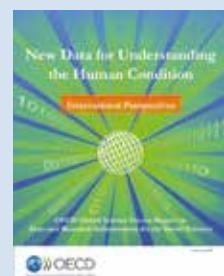
Report of the Administrative Data Taskforce 2012
This cross-departmental Taskforce proposes a major boost to the resources available for linkage and sharing across administrative datasets with the establishment of Administrative Data Research Centres in the countries in the UK. Additionally, all taskforce members are agreed that new legislation is required in order to overcome current legal obstacles to record-level linkage between data held by different administrative bodies.

---

5 House of Commons Science and Technology Committee (2012). *The Census and Social Sciences*, paras. 51 and 70-72.
6 Cabinet Office (2012) *Open Data White Paper: Unleashing the Potential.* June. Cm 8353.

research that generates valuable data and/or provide data from their administrative processes.

In summary, the social science community has done well so far to progress strategic plans which position it at the forefront of data sharing communities, but there is room for further improvement, particularly with respect to the sharing of data which combines interests from different disciplinary groups, for sharing of microdata across international boundaries (both from official statistical sources and from academic and commercial-led surveys), and for establishing links between personal records held by public and private sector organisations to facilitate research under appropriately controlled conditions of access. It is essential that the UK Data Forum pays close attention to ensuring adequate follow-up on these initiatives.

2.2 Cross-sectional data on individuals and households

Collaboration between the Office for National Statistics and the Economic and Social Research Council's Economic and Social Data Service, now the UK Data Service, has fostered progress in providing more rapid and enhanced access to a wide variety of cross-sectional microdata sources: some 500 new datasets were deposited with the Economic and Social Data Service between 2009 and 2012. The Integrated Household Survey, Labour Force Survey and other major datasets are now available to browse or download from the ESDS, often within months after the completion of fieldwork. In addition, some important new household surveys have been deposited at the ESDS by the Office for National Statistics, including the Wealth and Assets Survey and the Life Opportunities Survey. The ONS programme of work on Measuring National Well-being has produced a broad span of outputs on both subjective and other measures of well-being. While this progress is encouraging, there is a need to ensure that a clear protocol is agreed to provide quality assured data and metadata on new government surveys and that the long-term access and preservation of government surveys is achieved at an early date in order to maximise their research value.

Plans for flexible access to data collected within the 2011 Census of Population are well advanced. Work has also commenced to provide historical samples of anonymised records from the UK Population Censuses for 1961, 1966, 1971 and 1981. More information on issues relating to the future of the Census of Population is included in section 3.7. Given that the Census is a key source of information underpinning many social scientific studies, the Office for National Statistics, the ESRC and the research community

need to continue to work closely together to ensure that future data needs, particularly those relating to detailed spatial areas, continue to be met.

2.3 Data on and from organisations

Progress in this area includes completion of fieldwork for the 6th Workplace Employee Relations Survey and the launch of a pilot scheme ('Retail Data Navigators') to test research access to a variety of customer databases held by private sector organisations. Access to a wide range of official data on organisations is now available at both the ONS Virtual Microdata Laboratory and the secure access facilities of the UK Data Service. Research access to HMRC survey data from organisations is currently being piloted. Other major surveys of organisations completed in this period and now becoming available include the 2011 National Employer Skills Survey and the 2011 English Business Survey.

Major changes are taking place in the ways in which government and other public bodies make data available for research and other uses. The Open Data Institute (ODI) has been set up as a result of the UK Government Open Data Policy and is working to promote and improve access to a wide range of government datasets, especially for geo-coded 'micro aggregated' datasets. It is necessary to ensure that this development is harnessed by funders to enable social scientific research.

Administrative data, routinely collected by public sector organisations and relating to individuals, have enormous research potential either to enhance existing surveys or census data, or in their own right. However, access and linkage to such data for research purposes has been problematic. The Administrative Data Taskforce has investigated these problems and has made important recommendations covering the need for new infrastructure and legislation (see section 2.1). The implementation of these recommendations will be a key strategic aim over the next two years.

2.4 Longitudinal data on people and households

The last three years have seen the continuing development of Understanding Society as the UK's major longitudinal household study; the commissioning of a new national birth cohort study (now named the Life Study); and establishing the Cohorts and Longitudinal Studies Enhancement Resource (CLOSER) – a world-leading collaboration across the major UK longitudinal and cohort studies to provide an enhanced research environment for improved access to these valuable data resources. In addition to these major developments, the

ESRC, MRC and Wellcome Trust have collaborated to establish an Expert Advisory Group on Data Access (EAGDA), to provide strategic advice on the emerging scientific, legal and ethical issues associated with access to data, especially from cohort and other studies which combine biomedical with social and economic data. On-going attention is required to ensure these developments come to fruition.

Joint ESRC and MRC investment launched in October 2012 which brings together nine major UK longitudinal studies. It aims to maximise their use, value, visibility and impact and to stimulate interdisciplinary research across the studies. It will provide harmonised and shared resources for research and will assist with training and development for researchers. See: closerprogramme.co.uk/wordpress

This major new UK-wide birth cohort study with a strong interdisciplinary emphasis is a joint ESRC and MRC investment.

Recruitment of pregnant women for pilot begins in 2013. Up to 100,000 babies will be included in the study. Data collection will take place at two or three points through to 2016.
See: lifestudy.ac.uk/homepage

2.5   Developing international data resources

In the international arena, some major initiatives have taken place or are taking shape. These include:

- The Department of Business, Innovation and Skills has agreed that the UK should host the European Social Survey and, through the ESRC, will assist in promoting this research infrastructure to the status of a European Research Infrastructure Consortium[7].

- The ESRC is committed to co-funding the Consortium for European Social Science Data Archives, now seeking to become a European Research Infrastructure Consortium.

- Seventeen funding agencies are currently participating in a global collaboration to improve discovery, access and sharing of public health data. This collaboration, led by the Wellcome Trust, has already undertaken reviews and

is establishing pilot projects to improve access to public health data both for comparative research and to assist poorer countries with the development of these important data resources.

- More than twenty research projects have been completed, or are nearing completion, under the Pathfinders Programme, a cross-national initiative designed to promote comparative research using existing data across a range of disciplines in the social and economic sciences[8].

These developments indicate the strong interest there is in promoting an international social science research agenda. To promote this interest there is a need for researchers and research funding organisations often to act in concert. The proposed amendments to the *Regulation on Data Protection*[9] and its potential impact on data sharing is an example of the need for the voice of the UK research community to be heard. Equally, social scientists must develop a mechanism to coordinate their requirements for future multinational data infrastructure developments within the European Strategy Forum for Research Infrastructure[10].

The UK Data Service is a comprehensive resource funded by the ESRC to support researchers, teachers and policymakers who depend on high-quality social and economic data. The new service integrates several long-established data services into one central point of access, including the Economic and Social Data Service (ESDS), the Census Programme and the Secure Data Service. Its key aims are to:
- Support innovative, policy-relevant research
- Provide access to high-quality data
- Manage change in attitudes to data access
- Promote secondary data analysis
- Support and improve data skills
See: ukdataservice.ac.uk

2.6   Reviewing and promoting existing data services

The ESRC has completed reviews of the Census Programme (resources designed to provide a variety of access procedures to detailed Census of Population data), the Economic and Social Data Service (ESDS) and the Secure Data Service.

7   See ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric
8   The countries involved were Brazil, China, India, South Africa and the United Kingdom.
    See www.esrc.ac.uk/funding-and-guidance/funding-opportunities/international-funding/pathfinder.aspx
9   For further information see ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm
10 See ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri

As a result of these reviews, a new integrated service has been launched, the UK Data Service, which combines these services in a flexible and efficient manner. This new service will develop outreach procedures to promote access to the data currently held and will extend its holdings to include new forms of data with research value. It will encompass most of the current activities of the ESDS, the Secure Data Service, and the Survey Question Bank, and the data access and data support activities of the Census Programme. The UK Data Forum will need to work with the new service to ensure that it meets the needs of the social science community.

2.7 Summary and recommendations

The role of the UK Data Forum is to help identify, promote and monitor developments in the data infrastructure for the social sciences. With these goals in mind, the UK Data Forum endorses the plans and recommendations made in the reports mentioned in section 2.1 and will work to promote and monitor their implementation. Moreover it is crucial that the other recent developments outlined above are kept under review and progressed during the period 2013 to 2018. In particular, the UK Data Forum will:

- monitor progress towards the implementation of the recommendations of the Royal Society report *Science as an Open Enterprise*, particularly those relating to the sharing of personal information for research purposes;

- monitor and, where possible, assist with the implementation of the recommendations of the Administrative Data Taskforce. These recommendations have the potential to open up new areas of research, without reliance upon survey data collection, and to enhance existing survey data through linkage to administrative records;

- encourage government departments and other commissioners of data to work with the UK Data Service to ensure that all survey data with potential research value are transferred in a timely manner for long-term preservation and access, together with appropriate metadata;

- work with ESRC to ensure that social science research benefits from the Open Data Initiative;

- monitor the work undertaken by EAGDA, the Cohorts and Longitudinal Studies Enhancement Resource (CLOSER) and the UK Data Service to promote harmonised and consistent data access policies across agencies, government departments, disciplinary and national boundaries and, where possible, the private sector;

- in line with the recommendations made in the OECD Report *New Data for Understanding the Human Condition*, establish further links with data forums in other countries, particularly the countries of the European Union, seeking to exchange information on data access, sharing and data security and working to help identify cross-national data infrastructure needs;

- continue to support a strategic approach to the development of data and related resources for research in the social sciences and at the boundaries with other scientific disciplines.

Significant progress has been made over the past three years towards achieving the ambitious work plan outlined in the previous version of the *UK Strategy for Data Resources for Social and Economic Research*. However, there are certain areas where progress has not been as rapid as we would have wished. These include the development of access to data about organisations; and the need to promote the interest of UK social and economic researchers at the international level and within the European Union. Further work to progress these activities is outlined in the next section.

3 MEETING THE CHALLENGE AHEAD

DATA INFRASTRUCTURE NEEDS

3.1 The challenge

A thriving and vital research community is one that is engaged with major new developments in data resources, ensuring that they meet current research needs and provide the opportunities for future research. The challenge is to ensure that these needs are met and that research opportunities are grasped.

3.2 Enhancing longitudinal data about people and families

Longitudinal data are the essential resource for understanding life-course pathways. Past experiences and influences play a key role in shaping outcomes and behaviours, which in turn affect the ability of individuals to respond to life events and to policy interventions. The UK investment in these studies for over 65 years places our longitudinal data resources at the forefront internationally. Retaining world leadership in research that uses these important scientific and policy resources is a priority.

The UK comparative advantage is most evident for the wide portfolio of birth cohort studies. There is a need to ensure greater comparability and complementarity among these and other longitudinal studies, whilst encouraging innovation and greater depth. Since most key longitudinal studies are funded by the ESRC and the MRC, with the Wellcome Trust and Cancer Research UK also playing an important role, the establishment of CLOSER and of EAGDA provides a vehicle for funders to pursue such developments and work towards an appropriate balance of socioeconomic, behavioural, health and biological elements in coverage.

Several longitudinal studies cover a fairly wide age-range and quintessential among these is Understanding Society, the largest household panel study in the world. This study was launched following the first strategic plan[11] and covers a broad range of socioeconomic and health outcomes. It has recently been enriched by the addition of cognitive measures and the collection of biomarkers. It is essential that the linkages of blood analyses and genomic markers, including epigenetic signatures, to the rich socioeconomic information are enabled to help facilitate the emerging biosocial science research agenda.

Longitudinal interview-based survey data can also be enhanced via further linkage to administrative records, which enable additions to the content at relatively low cost. Small additional investments can create data which have significant research potential. A specific example is the planned continuation of the Longitudinal Study of Young People in England (LSYPE)[12] through administrative data linkage following the decision of government to discontinue annual face-to-face interviews with cohort members due to resource constraints.

Innovation in data collection is required if the UK is to retain world leadership in these longitudinal studies. The emerging efforts to explore the potential for mixed mode and web-based collection of information need consolidation and much further development. Such approaches may help retain quality, maintain response rates, enable more frequent contacts and enhance tracing, as well as help cost containment. Innovation in content is also of critical importance to enable exploration of emerging scientific and policy agendas. New ways of collecting information on attitudes and motivations should be sought, linking these to behavioural outcomes, as the human/computer interface becomes more embedded within our everyday lives.

Longitudinal data sources are expensive to maintain and develop. However, unlike physical infrastructure that depreciates in value through time, there are major returns to each additional wave that add value to all of the previous investments. These studies play a critical role in enabling understanding of long run processes and mechanisms involved in shaping behaviours and outcomes. While the Research Councils have made significant progress over the last five years in bringing a more strategic approach to their funding, further efforts must be made to secure a sustainable funding base for those studies which have a long run perspective.

3.3 Addressing multidisciplinary challenges

Many of the most interesting and innovative challenges in today's complex world require multidisciplinary approaches both within the social sciences, but particularly engagement with disciplines beyond the social sciences. An example of the former is the collaboration that has developed between economists and psychologists in the field of behavioural economics that underlies the 'Nudge' unit in the Cabinet Office[13]. Virtually all government policies require a strong social scientific input, whether on energy and climate change, the built environment, food security and military recruitment, or other areas including more obvious concerns with poverty reduction, early child development and social mobility. In the

---

[11] *The National Strategy for Data Resources for Research in the Social Sciences, 2006-2009.* Swindon: ESRC.
[12] The LSYPE is a major panel study of young people which brings together data from several sources, including annual interviews with young people and their parents, and administrative sources. 15,500 respondents were first interviewed in the spring of 2004 (at age 13) and were interviewed annually until 2010, resulting in a total of seven 'waves'. For the first four waves of LSYPE, the parents or guardians of the respondents were also interviewed.
[13] See cabinetoffice.gov.uk/news/governments-nudge-unit-goes-global

same context a broader understanding of behaviour requires engagement of the social and behavioural sciences with neurobiology and epigenetics[14].

A key area of development will continue to be growing engagement of the social sciences with health sciences and biological sciences. Increasingly data resources that have their origins in the social sciences or in the biomedical sciences are rightly dismantling these barriers and moving towards a biosocial science. There is evidence to suggest that lifetime experiences get 'biologically embedded' among the pathways or mechanisms involved in lasting effects. These biological changes induced by the external environment can include changes in brain architecture or epigenetic effects, as well as socioeconomic or personality trait legacies. Ensuring that data resources are rich enough to enable such processes to be studied requires repeated collection of biomeasures including DNA. We thus need data resources which have suitably rich information to permit elucidation of pathways and mechanisms that may shape or be shaped by behaviours and subsequent experiences.

Many other multidisciplinary challenges can be better addressed by linking longitudinal data resources to administrative and other records. For example links to environmental data to enable study of the mutual interplays between areal contexts, the built environment, pollution and people's behaviours and experiences; links to customer transactions data enable better understanding of consumer choices and are important in relation to research on diet and obesity.

A multi-disciplinary research agenda requires coordination of relevant interests, ensuring that social scientists engage with scientists from other disciplines and that other disciplines recognise the key role of the social sciences. For this to happen, higher education institutions need to recognise and promote the integration of different disciplinary approaches to research and analysis, and research funders must be proactive in promoting programmes of research that draw on data that span disciplinary domains and require the pooling of skills, knowledge and analytical methods from a range of disciplinary areas.

3.4   Data about organisations

Organisations – social units of people systematically structured and managed to meet a need or pursue collective goals – form an essential part of our societies. It is through organisational structures that economic growth, education, health and social wellbeing are pursued. They underlie paid employment, generate profits for owners and shareholders, facilitate investments in infrastructure across the range of human activities and provide for the health and wellbeing of populations.

There are many areas where an improved understanding about the role of organisations would have major benefits for policy. These include:

- organisational structure and economic growth – the roles played by R&D, networks, innovation activities, skills development;

---

[14] Neurobiology is the study of cells of the nervous system and the organisation of these cells into functional circuits that process information and mediate behaviour. Epigenetics is the study of heritable changes in gene expression caused by mechanisms other than changes in the underlying DNA sequence.

- organisational efficiency, productivity and the link with industrial relations;

- organisational behaviour and financial structure; investment and tax policy;

- employee health and wellbeing at the workplace and organisational policies;

- the social and economic contributions of voluntary organisations;

- 'not-for-profit', local and national government organisations and the measurement of value added.

In contrast to the amount of data that is available for research on individuals and households, information on organisations may be difficult to locate or access for research and is fragmentary. Currently there are research groups or networks that have formed around specific types of data on organisations and/or have particular disciplinary affiliations. These include the study of employment relations, innovation, financial performance, organisational productivity, training and skills development, managerial behaviour, marketing and organisational location. These related research interests tend to focus upon organisational data of a specific type, say from a survey, statutory reporting requirements or customer databases. A mechanism is required which provides for the cross-fertilisation of ideas and expertise to enable new resources to be developed and for existing resources to be integrated in ways that inform a multi-disciplinary approach to understanding organisations. A first step in this direction should include a stock-take of available data and a review of the range of interests in the research use of data on organisations, with consideration of the benefits that could emerge from a more coordinated, cross-disciplinary approach to the development of data about organisations. Dependent upon the outcome of this review, further steps could be taken to provide improved research access to the wide range of data on organisations. In tandem with these steps research activity should be built up in ways that encourage and promote the development of analytical methods and data linkage techniques to help gain a better understanding of the role organisations play in the economy and society.

3.5. Customer databases

As society becomes ever more differentiated in terms of lifestyles and consumption habits, retailers and many other private sector organisations have come to rely upon customer databases covering purchases or services provided in order to gear their product and service offerings to the needs of consumers. The most obvious example of this data revolution in retailing is the use of shopping cards by customers to record their purchases in exchange for incentives offered by retailers. There are many other examples, including the deployment of smart meter technology in the retail energy industry, deep data mining of mobile phone data and the variety of research uses that can arise from records of internet-based activities.

These developments have potential research and policy interest well beyond the organisations themselves. Energy company data are of value in understanding fuel poverty for example, while pharmaceutical purchases are of interest to researchers investigating the development and maintenance of healthy lifestyles.

Preliminary work from a 'Retail Data Navigators'[15] scheme launched in March 2012 has indicated interest from several commercial organisations in the use of their customer databases for wider social and economic research. A strategic goal for the period 2013-2018 will be to build on this interest, exploring the potential research value of such data and establishing data access protocols.

3.6 Social media as research data

Through social media, millions of human interactions occur and are recorded daily, creating massive data resources which have the potential to aid our understanding of patterns of social behaviour (e.g. participation in activities; political attitudes; risk-taking behaviours; group motivations; religious beliefs; lifestyle choices, etc.). Social media analytics represent an opportunity to invest in extensive, large scale social research that is within a temporal frame that cannot be achieved through either snapshot surveys or interviews or via longitudinal panel research.

Social media are formed of complex networks that cross national boundaries, meaning a coordinated international social science response may be the most appropriate approach to the development of social media as research data. The social sciences should provide the impetus for such research, and the development of knowledge about social media data as research resources must be a key element of this strategic plan.

---

15 See esrc.ac.uk/funding-and-guidance/collaboration/working-with-business/index.aspx

3.7 Cross-sectional data on people and families

Cross-sectional data on people and families arises mainly from surveys and the decennial Census of Population. While much progress has been noted for household surveys, the future of the UK Census of Population is now being considered carefully[16].

The Censuses of Population are strategic national assets. Their flexibility to meet future needs in areas relating to area population projections, service provision, and physical infrastructure planning is vital to the wellbeing of the nation. The decision to review the traditional census-taking method was agreed across the United Kingdom. National Records Scotland is working closely with colleagues in the Office for National Statistics (ONS) and the Northern Ireland Statistics and Research Agency (NISRA) to ensure a consistent and coordinated approach to the future production of UK population statistics. By 2014 they will have made recommendations to Government on the next Census and how it should be conducted. Options range from an on-line census, to some combination of linked administrative data plus sample survey methods. Whatever shape it takes, it is vital that the key users of census data, especially those with an interest in the detailed geography of issues relating to ethnicity, religion and social inequalities, make the case for the retention of provision for their data needs. Equally, researchers who have made use of the Samples of Anonymised Records from the censuses and the three Census Longitudinal Studies must be assured

that the research potential and ongoing linkage of these resources is not diminished by any changes to census plans.

Following the 2010 Comprehensive Spending Review, resource constraints have had a significant impact upon the ability of various government departments and the Office for National Statistics to continue with existing cross-sectional surveys or to launch new surveys, and created financial pressures on devolved administrations too. This is an area where the UK Data Forum can help monitor the impact of reduced data collection on research activity, identifying where information needs are most affected and reviewing options for alternative strategies to develop relevant data sources.

3.8 International data access

Access to social and economic data from other countries, and to UK data by researchers from other countries, is vital for the support not just of cross-national comparative research, but also to facilitate a growing research agenda where national boundaries are of less importance. Issues such as poverty and migration, the spread of infectious diseases, understanding and responding to environmental catastrophes are examples of areas where social scientists needs to work together regardless of their national locations. This, in turn, drives the need for better access for knowledge about data in different countries, for data and metadata to be harmonised between countries and for improved cross-national access to survey, census and administrative data at the microdata level.

With such issues in mind, the OECD Global Science Forum established an Expert Group which published its findings in 2013[17]. Foremost among its recommendations are the following:

- National statistical organisations and international organisations should ensure that all data they collect and process are documented to agreed and common standards. Such documentation should be easily discoverable on their websites.

- National statistical agencies should establish mechanisms to improve access by the research community to social science microdata in their possession.

- National and international statistical agencies should strengthen the efforts they are making to harmonise social and economic data at the international level, seeking to prioritise these activities in specific areas.

---

16 See ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/index.html
17 www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.htm

- National funding agencies should ensure that new research awards have accompanying data management plans and should assign resources for this purpose. They should cooperate at both national and international levels to share this information.

- National funding agencies should collaborate with national libraries at the international level to ensure that a common system is adopted for referencing datasets in research publications.

To take these recommendations forward, national funding agencies and national statistical agencies should work together to identify lead responsibilities and to coordinate their actions. This could be achieved through the creation of a body located within a permanent facility and tasked to promote and monitor activities that support these recommendations.

3.9   Public engagement and communication

Many of these proposals for meeting the challenges ahead could raise concerns with the public about the potential for misuse of data that they have provided, or which relates to them, their families, their communities and their workplaces. The development of procedures to facilitate linkage between administrative datasets is an area of particular concern, as is the possible increased use for research of data about or from organisations. An important element of the UK Strategy for Data Resources for Social and Economic Research will be to ensure that there is wide and effective consultation and discussion on issues that relates to the use of personal and organisational data for research purposes. This will require the research community, data holders and research funders to cooperate in the formulation a Public Engagement Strategy – a plan which specifies how such engagement should take place, with whom, for what types of data development and how the resulting information should be acted upon. A key element of the strategy must be the specification of the communication

channels that should be established to ensure that public engagement is a two-way process, through which researchers can describe the potential public benefits arising from their proposed use for research of particular types of data and the measures that they will be taking to protect such data, and members of the public have the opportunity to voice any concerns they may have about such use of their data.

3.10   Recommendations

There are seven areas where the development of data resources needs to be prioritised over the next few years.

Enhancing longitudinal data about people and families

- Using new and proposed facilities for data linkage (e.g. e-Health Centres, Administrative Data Research Centres), to improve linkage between longitudinal survey data, biomedical data (genomic, health records, etc.) and administrative records. The use of these facilities to be encouraged by research funders collaborating to develop a stronger interdisciplinary basis for research programmes (both within differing social sciences and to the biological, medical and environmental sciences) which will benefit from linked data.

- For research funders to ensure that the range of longitudinal studies achieves a balanced portfolio to meet research needs. This to include an audit of longitudinal data resources, identifying overlaps and gaps in content and coverage required to address major interdisciplinary research themes.

- For research funders to promote continued innovation in social scientific understanding (behavioural economics, psychology of decision making and choice, etc.) taking advantage of developments in the human/computer interface where appropriate and working to promote greater integration with the biological and health sciences.

- For the governing bodies of the major longitudinal studies and related resources to encourage exploration of linking to new forms of data from social networking and other online sources and the use of mixed mode and web-based approaches to collecting data.

- For government departments and research funders to consolidate a stable, substantial and sustainable funding basis for key longitudinal studies.

- For research funders and policy users to cooperate in the promotion of seminars, workshops etc. designed to

raise awareness of the potential value for policy-relevant research among policy users.

Addressing multidisciplinary challenges
- The Economic and Social Research Council collaborate with other funders to develop a strategic approach to biosocial research, including working to ensure the establishment of a wide range of appropriate biosocial data sources and encouraging research that links information on life experiences and individual behaviours to neuroscience, genomics and health sciences.

- For higher education institutions to foster a more collaborative approach to the development of cross-disciplinary research skills and data analysis.

Data about organisations
- The Economic and Social Research Council to commission a review of the range of interests in organisational data and propose ways to encourage cross-fertilisation of research and resources across these interests.

- In the light of this review, the Office for National Statistics, Department of Business, Innovation and Skills, HMRC, the UK Commission on Employment and Skills, the ESRC and the statistical authorities in the devolved administrations to collaborate with the private sector, seeking ways to develop and promote access to and linkage between data about and from organisations.

Customer databases
- The Economic and Social Research Council to continue to develop relationships with retailers and other private sector organisations (including internet based businesses) that have built large customer databases, with the aim initiating some exploratory analytical projects of mutual interest.

- For the ESRC and the ONS to cooperate in investigating the potential use of customer databases and other large scale datasets for the 'Beyond 2011' project.

- In collaboration with private sector organisations, the Economic and Social Research Council to carry out a review of future skill needs and training requirements to meet the demand for analysis of large and complex customer databases generated by organisations.

Social media as research data
- The Economic and Social Research Council to explore with other national funding agencies the scope for an

international approach to ascertain the value of various forms of social media data as research resources.

Cross-sectional data on people and families
- As an important part of the 'Beyond 2011' consultation, the ONS to gather input from the research and policy communities regarding their needs for detailed geographical information on people and families.

- For the ONS to give careful consideration to the impact that any proposed changes will have on the nature and availability of census microdata records for research, particularly for research based on the Samples of Anonymised Records and the Census Longitudinal Studies.

International data access
- The ESRC to support the establishment of the ESS and CESSDA as European Research Infrastructure Consortia and assist with their activities to promote access to harmonised data for comparative research.

- The ESRC, working with research funding agencies in other countries, to assist with efforts to explore the research potential of new forms of data.

- For the major national social science funding agencies to cooperate in finding ways to support the recommendations of the OECD Global Science Forum Expert Group.

  Additionally, it is important to ensure that there is strong public support for those activities which will develop data by re-using and linking between data provided by people for purposes other than research. There is also a need to ensure that the long-term future of many of these resources is assured.

Public engagement and communication
- For researchers, data holders, research funders and other relevant stakeholders to develop a Public Engagement Strategy covering engagement and communication with the public on issues relating to the reuse of and linkage between data which have the potential to identify individuals and/or organisations.

Sustainable funding for key data services
- For research funders to explore mechanisms that provide for stable and predictable funding of those data services which, by their very nature, must be sustained over the long term.
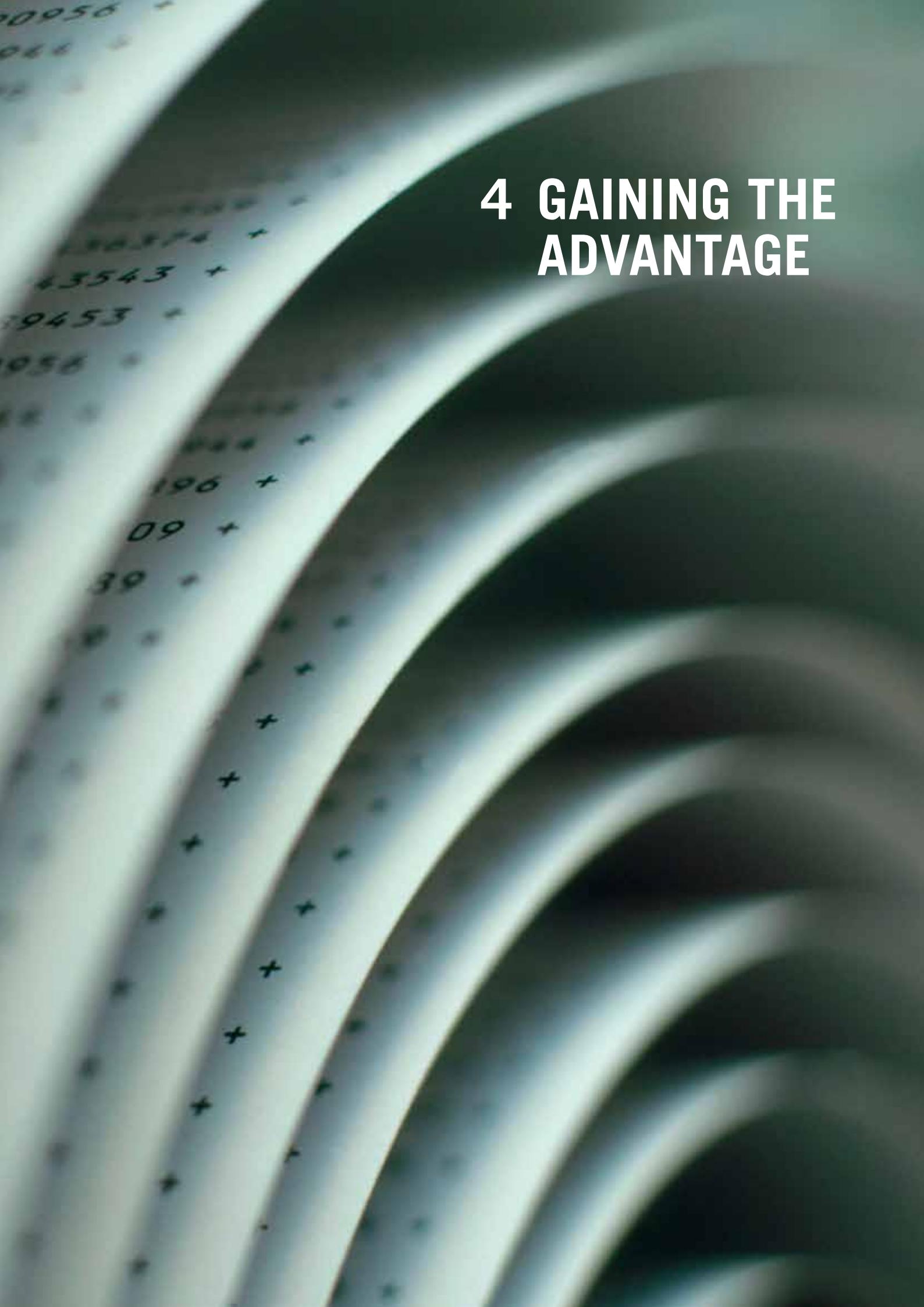
4 GAINING THE ADVANTAGE

ENABLING RESEARCHERS TO MAKE FULLER USE OF NEW AND EXISTING DATA SOURCES

4.1 The challenge

Effective research combines three elements: physical research infrastructure (data and tools for analysis); the technology and processes to facilitate access (data curation, computational resources, communication networks and relevant security measures) and researchers with the appropriate skills and knowledge to undertake leading edge research which takes advantage of the data infrastructure and technological developments. The challenge is to ensure their effective combination – that users of data are equipped to exploit the full potential of new and enhanced data resources and have the knowledge and technical skills to derive maximum advantage from these resources.

4.2 The issues

There is clear evidence of a shortfall between the supply of and the demand for researchers with the requisite data analysis and data manipulation skills in the academic, government and private sectors. Some key developments are under way to ensure greater quantitative skills training. At the undergraduate level the £15.5 million Quantitative Methods Programme investment by the Nuffield Foundation, the ESRC and HEFCE aims to promote a step-change in quantitative methods training for UK social science undergraduates. At the postgraduate level and beyond the needs and skills become deeper and more varied. The ESRC's £11.5 million Secondary Data Analysis Initiative is encouraging more researchers to engage with data resources. Enabling capacity for high quality research using 'Big Data' in the academic, government and private sectors is increasingly being regarded as important for achieving goals of economic growth and improved wellbeing.

To realise the potential of the wide range of developments in research infrastructure, researchers need to be equipped to make full use of new and existing data sources. Many of the recommendations in earlier sections of this Strategy are intended to make it simpler to discover and access a growing range of data. However, the increasing complexity and coverage of such resources poses demanding challenges for greater cross-disciplinary collaboration, for more complex analytic tools, and for access to specialist knowledge and support.

The sheer volume and scale of many of the new or developing sources of research data will require researchers to have access to methodological support, particularly in the areas termed 'informatics' or 'data science'. In the biological, medical and behavioural sciences it is usual for any sizeable research unit to employ such specialist support as part of their infrastructure. There is a need to ensure that all Doctoral Training Centres and many major research centres in the social sciences can provide such support. There is also a role for the National Centre for Research Methods (NCRM) in encouraging related activities. However, complex data analysis usually requires more expertise than that which can be gained via short training courses. Mechanisms have to be provided to enable on-going engagement with specialist statisticians, informaticians and data managers.

Research on the interplays between life experiences and biology is inherently multidisciplinary, requiring collaboration among social scientists, neuroscientists, and biomedical and biological scientists. A minimum requirement for doctoral training should be that all students are made aware of these developments and given a basic knowledge of key topics, such as neuroscience or epigenetics, emphasising the links to the social and behavioural sciences. At a more advanced level, researchers will need more extensive training through multidisciplinary Doctoral Training Centres and jointly funded postdoctoral fellowships. There is also a need for jointly funded research programmes that enable experienced researchers to begin to realise the huge potential from the biosocial data emerging from longitudinal studies.

There are parallel requirements to enable research using administrative and organisational data. The skills required here relate more to methods of data linkage, data cleaning and data mining. The skills required for understanding and working with collaborators from government or the private sector and knowledge of resources and opportunities to enable this should be part of basic doctoral training. Researchers who wish to pursue such opportunities need to have access to specialised training and support.

At a more basic level there is evidence of a growing shortfall in the recruitment of 'data scientists': individuals who can bring skills in data manipulation and data management to the research process. This is a need to develop training and collaborative arrangements between the academic, government and private sectors to enable skilled and timely analysis of large-scale data sources, with improved methods of data visualisation and spatial analysis, for rapid provision of evidence for policy makers.

4.3   Recommendations

In comparison with some other countries, social and economic researchers in the UK have less incentive or may have been less well prepared to obtain the methodological skills and knowledge that are needed for what are rapidly becoming increasingly complex fields of research. It is vital that steps should be taken to address this shortfall. It is recommended that the research councils and the education funding councils should take action as follows:

- In its review of the operation of Doctoral Training Centres, the ESRC should have regard to the delivery of leading edge courses on data manipulation, management and analysis. This should also include knowledge of links to biological data and links to health sciences, the role of longitudinal studies in establishing links between experience and behaviour and of the potential use of administrative data as research resources.

- For Research Councils to develop collaborative interdisciplinary doctoral and postdoctoral training that spans Research Council remits, coupled to initiatives to demonstrate and enhance multidisciplinary data analysis linking the social sciences to biomedical, biological and environmental sciences.

- For government, academic and private sectors to identify the shortfall in key data skills and to cooperate in the development of training programmes designed to meet specific needs.
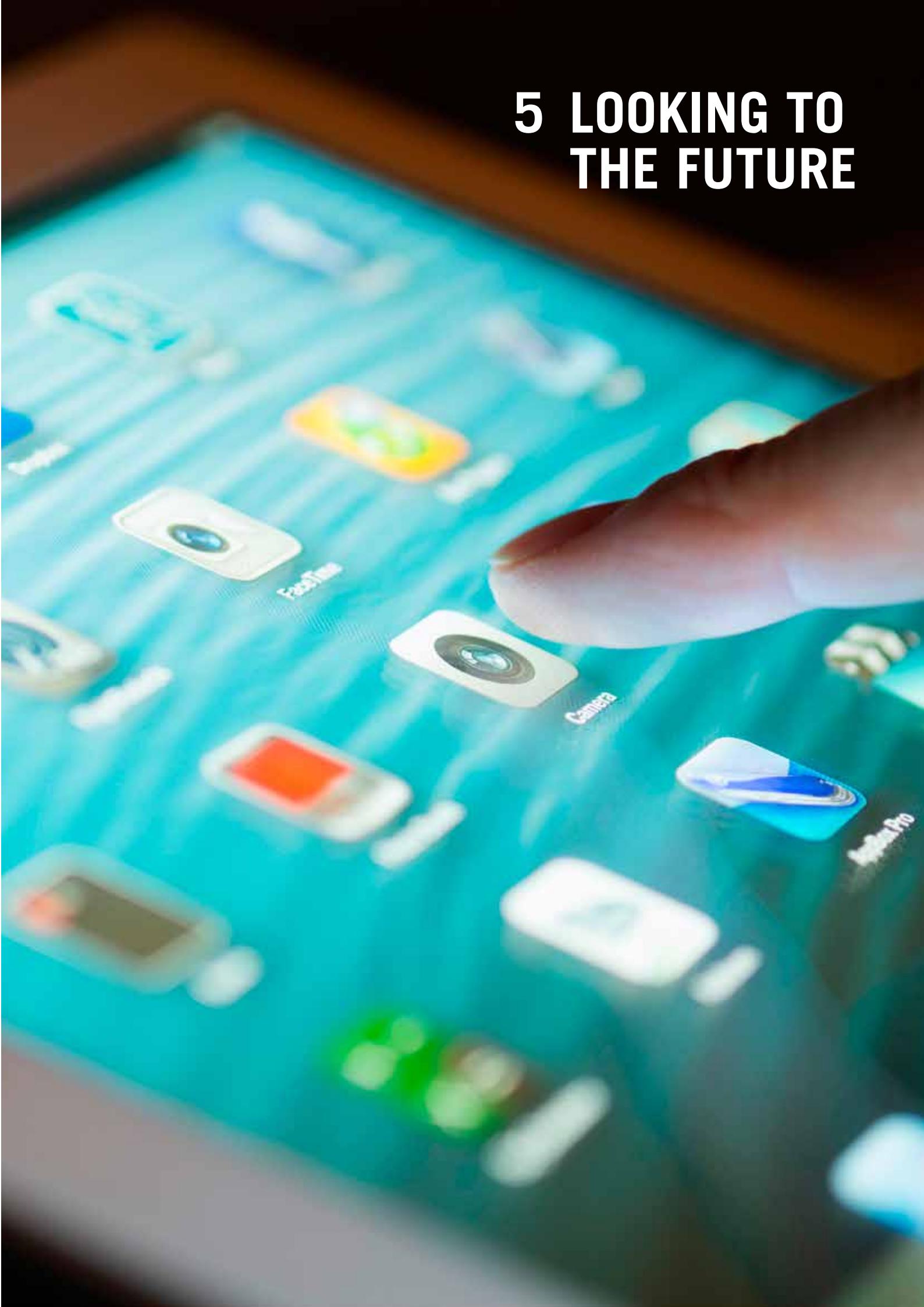
5.1   The challenges

The UK has a leading position in both the quality and quantity of social and economic data that is available for research. This has been achieved over a long period through innovative approaches to investment in data infrastructure and through an ability to capture the impact of these investments in terms of the benefits derived from the research they support. Section 3 has outlined the steps that must be taken over the next few years to consolidate this position. This section looks further ahead, to identify the areas where the social and economic sciences will be positioned five years from now and to highlight the most significant changes in the nature of the ways in which research data will be identified, captured and analysed.

5.2   Exploring the research potential of new forms of data

Traditional forms of large-scale data in the social and economic sciences rely on survey techniques associated with representative samples of populations (of households, business organisations, etc.) or census methods. New forms of data arise from the growth of electronic communication, from internet usage, the growth of web-based services and from the expansion of the digital economy. Such data have huge potential research value and will become ever more pervasive. Plans to explore and capitalise on this potential will benefit from cross-national coordination and collaboration.

New forms of data bring considerable opportunities to have more frequent reporting, serial recording and rich geographic identification. The very much greater volume and velocity of information will need to be complemented by gains in methodology to assess and analyse these data and by new information management practices, including visualisation methods. It is expected that there will be a much faster rate of change in access to public and commercial records in the future than in previous decades.

A strategic approach is required in order to reap the potential benefits arising from such data sources. While they are not specifically designed for research, there is already sufficient evidence[18] of their research value to warrant a major programme of research and development, possibly undertaken in collaboration with funding partners from outside the UK. This programme of research will have to address the issues raised by the selective coverage of many of these data sources and the implications for inference to the general population.

5.3   Recommendations

To stimulate the use of new forms of data for research purposes, a number of recommendations are made:

- For national social science research funding agencies from a number of countries to collaborate in ways which will increase the range and depth of international research infrastructures, particularly those which harness new forms of data from social networks, the web, internet providers and the digital economy and enable such usage through appropriate methodological developments.

- For research funders to develop programmes of research that address the extent to which new forms of data are representative of the population and that exploit major recent developments in the handling of 'Big Data'.

- For national social science research funding agencies and data protection authorities to develop ethical guidelines for the research use of new forms of data, to complement existing guidelines.

---

18   For an example, see google.org/flutrends/

6 ACTION PLAN

Lead responsibilities and timetable

The following table lists the recommendations from each of the preceding sections of the UK Strategy for Data Resources for Social and Economic Research. In most cases, these recommendations will require collaboration between groups of bodies (government departments, research funding agencies, statistical authorities and the devolved administrations). The table below indicates where the lead responsibility might lie for initiation and monitoring of the activities associated with each recommendation and gives an estimate of the timetable for these activities.

Lead responsibility	Nature of the recommendation	Action by	Completed by
UK Data Forum	Monitor and assist with the implementation of the recommendations of the Administrative Data Taskforce and the Royal Society report 'Science as an Open Enterprise', particularly those relating to the sharing of personal information for research purposes.	June 2013	December 2014
	Encourage government departments to work with the UK Data Service to ensure that all survey data with potential research value are transferred in a timely manner for long term preservation and access, together with appropriate metadata.	Agree protocols by mid 2013	Ongoing, subject to annual review
	Work with ESRC to ensure that social science research benefits from the Open Data Initiative.	Ongoing	
	Monitor the work undertaken by EAGDA and the Cohorts and Longitudinal Studies Enhancement Resource (CLOSER) to promote harmonised and consistent data access policies across agencies, government departments, disciplinary boundaries and, where possible, the private sector.	Ongoing	Review annually
	Establish further links with data forums in other countries particularly the countries of the European Union, seeking to exchange information on data access, sharing and data security and working to help identify cross-national data infrastructure needs.	Ongoing	Review twice yearly
	Continue to support a strategic approach to the development of data and related resources for research in the social sciences and at the boundaries with other scientific disciplines.	Ongoing	Next version of strategy by October 2017
	Promote research using new and proposed facilities for data linkage (e-Health Centres, Administrative Data Research Centres), improve linkage between longitudinal survey data, biomedical data (genomic, health records, etc.) and administrative records.	Ongoing	

Lead responsibility	Nature of the recommendation	Action by	Completed by
ESRC with other research funding agencies	Conduct an audit of longitudinal data resources to identify overlaps and gaps in content and coverage.	June 2011	December 2014
	Develop collaborative interdisciplinary postdoctoral training that spans Research Council remits, coupled to initiatives to demonstrate and enhance multidisciplinary data analysis linking the social sciences to biomedical, biological and environmental sciences.	April 2014	Ongoing
	Develop a strategic approach to biosocial research, including working to ensure the establishment of a wide range of appropriate biosocial data sources and encouraging research that links information on life experiences and individual behaviours to neuroscience, genomics and health sciences.	Ongoing	
	Develop programmes of research that address the extent to which new forms of data are representative of the population and that exploit major recent developments in the handling of 'Big Data'.	Ongoing	
Cross Research Council and higher education institutions	Foster a more collaborative approach to the development of cross-disciplinary research skills and data analysis.	Ongoing	
ESRC with UK Government, departments, UK research funding agencies, Governing Boards and CLOSER	Consolidate a stable, substantial and broad funding basis for key longitudinal studies.	September 2014	December 2017
	Encourage exploration of linking to new forms of data from social networking and other online sources and use of mixed mode and web-based approaches to collecting data in longitudinal studies.	June 2014	December 2015
	Co-operate in the promotion of seminars, workshops etc. designed to raise awareness of the potential value for policy-relevant research among policy users.	Ongoing	
	Develop a Public Engagement Strategy covering engagement and communication with the public on issues relating to the reuse of and linkage between data which have the potential to identify individuals and/or organisations.	March 2014	December 2014
ESRC	Review of the range of interests in organisational data and propose ways to encourage cross-fertilisation of research and resources across these interests.	March 2014	June 2014

Lead responsibility	Nature of the recommendation	Action by	Completed by
ESRC continued	Develop relationships with retailers and other private sector organisations (including internet based businesses) that have built large customer databases.	Ongoing	
	Support the establishment of the ESS and CESSDA as European Research Infrastructure Consortia.	Ongoing	
	In its review of the operation of Doctoral Training Centres, the ESRC should have regard to the delivery of leading edge courses on data access and analysis, including coverage of the potential for use of longitudinal data to establish links between experience and behaviour, neuroscience, genetics and the health sciences and the potential research value of administrative data.	January 2014	December 2014
ESRC with UK Government departments and private sector organisations	Collaborate with the private sector over leadership and funding to develop and promote access to data resources about and from organisations.	June 2013	Ongoing
	Identify the shortfall in key data skills and to cooperate in the development of training programmes designed to meet specific needs		
ONS with ESRC	Cooperate in investigating the potential use of customer databases and other large scale datasets for the 'Beyond 2011' project.	2013	Report early 2014
	Gather input from the research and policy communities regarding their needs for detailed geographical information on people and families.	2013	Report early 2014
	Give consideration to the impact that any proposed changes from 'Beyond 2011' will have on the nature and availability of census microdata records for research.	2013	Report early 2014
ESRC with research funding agencies in other countries	Assist with efforts to explore the research potential of new forms of data.	April 2014	Ongoing
	Cooperate with other national funding agencies finding ways to support the recommendations of the OECD Global Science Forum Expert Group.	October 2013	Ongoing

Appendix

Membership of the UK Data Forum

Chair:

Professor Tim Holt

Members (in alphabetical order)

Mr Richard Bartholomew	Department for Education
Ms Rachel Bruce	Joint Information System Committee
Mr Dave Carr	Wellcome Trust
Dr Norman Caven	Northern Ireland Statistics and Research Agency
Ms Vanessa Cuthill	Economic and Social Research Council
Mrs Pam Davies	National Statistician's Office
Ms Janet Dougharty	Department for Communities and Local Government
Mr Keith Dugmore	Demographic Decisions Ltd
Dr Peter Dukes	Medical Research Council
Mr Peter Fullerton	Office for National Statistics
Mr Rod Harrison	Scottish Executive
Mr Denis Huschka	German Data Forum
Mr John Kaye	British Library
Mr David Frazer	Department for Work and Pensions
Ms Linda Stewart	The National Archives
Dr Andy Sutherland	Health and Social Care Information Centre
Mr Mark Thorley	Natural Environment Research Council
Professor Gert Wagner	German Data Forum
Ms Amanda White	Home Office
Professor Chris Whitty	Department for International Development
Ms Sharon Witherspoon MBE	Nuffield Foundation
Dr Matthew Woollard	UK Data Service/UK Data Archive, University of Essex

Executive

Ms Lynne Conaghan	University of Warwick
Professor Peter Elias	University of Warwick
Professor John Hobcraft	University of York

Terms of Reference for the UK Data Forum

The terms of reference of the UK Data Forum are:

- to develop a clear view of the key national and international data resources that are required to meet the future needs of social and economic research and research in related fields of interest;

- to identify gaps in the current data infrastructure, and determine how those gaps might be closed;

- to prioritise future data provision, including the balance between:
  - enhancing access to and sharing of existing datasets
  - maintaining ongoing support for major datasets of national importance
  - new data collection
  - withdrawal of support and archiving of existing datasets that have outlived their usefulness;

- to agree principles and strategies for the co-ordinated development and co-funding of prioritised future data resources, as appropriate;

- to agree training and capacity building initiatives to promote the effective analysis of data resources;

- to identify and seek to resolve the practical and ethical issues associated with linking between data sources.

The UK Data Forum recognises that various concordats between the ESRC and government departments are already in place or are being established and that efforts at data coordination are being made across government departments and the devolved bodies. These represent a strong foundation for future collaboration. The Forum builds on and strengthens such work to provide a cross-council, cross-government mechanism for the promotion of collaborative working.

List of acronyms

CESSDA	Council for European Social Science Data Archives
CLOSER	Cohorts and Longitudinal Studies Enhancement Resource
EAGDA	Expert Advisory Group on Data Access
ESDS	Economic and Social Data Service
ESRC	Economic and Social Research Council
ESS	European Social Survey
HEFCE	Higher Education Funding Council for England
HMRC	Her Majesty's Revenue and Customs
LSYPE	Longitudinal Study of Young People in England
MRC	Medical Research Council
NCRM	National Centre for Research Methods
ODI	Open Data Institute
OECD	Organisation for Economic Cooperation Development
ONS	Office for National Statistics
RCUK	Research Councils UK
UKDS	UK Data Service