

# Data Infrastructure Strategy Stakeholder Engagement

Shaping investment in social science data  
infrastructure for the next five years

July 2021

# Contents

<b>1. ABOUT THIS STAKEHOLDER ENGAGEMENT EXERCISE</b>	<b>3</b>
1.1. High-level summary	4
1.2. Questions	4
1.3. How to respond	4
1.4. Next steps	5
<b>2. POSITION STATEMENT</b>	<b>6</b>
<b>3. OBJECTIVES</b>	<b>8</b>
<b>4. ABOUT THE DATA INFRASTRUCTURE STRATEGY</b>	<b>10</b>
4.1. ESRC's role in the social science data infrastructure system	11
4.2. Why a Data Infrastructure Strategy is needed	12
4.3. Vision for the Data Infrastructure Strategy	13
4.4. ESRC Data Policy	13
4.5. What activities will support the Strategy's development and delivery?	14
<b>5. DATA INFRASTRUCTURE STRATEGY THEORY OF CHANGE     AND DRAFT OBJECTIVES</b>	<b>15</b>
Figure 1: Theory of Change	17
Figure 2: Objectives	18
Annex 1: ESRC managed investments in data infrastructure	19





# 1.

## About this stakeholder engagement exercise

## 1.1 High-level summary

ESRC is the largest UK investor in social science-led data infrastructure. A total investment of over £200m over five years supports world-leading longitudinal and cross-sectional studies, vital enabling data services for research, and projects which obtain and curate big data. This investment must be guided by a strategic and holistic approach.

This engagement exercise invites input from our key stakeholders to help shape ESRC's new Data Infrastructure Strategy. The strategy will guide our investment in data infrastructure from 2022.

The document below includes a vision for the Data Infrastructure Strategy (**section 4.3**), a theory of change (**Figure 1**) and draft goals and objectives (**Figure 2**).

We invite responses to our questions by **10 September**.

## 1.2 Questions

1. What do you think about the vision for the Data Infrastructure Strategy? For the vision, see **section 4.3**.
2. What do you think about the activities identified to support delivery of the strategy and do you think any others should be considered? For the activities, see **section 4.5**.
3. Do you have any comments on the draft theory of change (**Figure 1**), focusing on evidence of need and targeted impacts?
4. What do you think about the draft goals and objectives in **Figure 2**?
5. In your view, and after reviewing this document, what impact will the strategy make?
6. Is there anything else that you would like to comment on?
7. Would you be happy to be contacted to discuss your comments and would you like to be involved in any additional engagement activities?

## 1.3 How to respond

Responses to the above questions are invited by email to [datainfrastructure@esrc.ukri.org](mailto:datainfrastructure@esrc.ukri.org).

The deadline for comments is **10 September 2021**.

For convenience, a Word document compiling the questions in this document is available [here](#). ESRC is very happy to meet to receive verbal feedback (subject to availability). Please contact the above email address and we will arrange a meeting.

Accessible versions of this document's figures, suitable for screen-reading software, are available on the [ESRC website](#).

## UK DATA SERVICE

The UK Data Service (UKDS) plays a pivotal role in realising ESRC's investment strategy to support data infrastructure as a shared research resource. It provides a centralised archive and unified point of access to a very extensive range of high quality social and economic data, including UK census data, government funded surveys, longitudinal studies, international macrodata, qualitative data and business microdata, that would not otherwise be readily accessible. It is designed to provide seamless access and support to meet the current and future research demands of both academic and other users, and to help them maximise the impact of their work. It is highly effective in addressing challenges in data discovery and access for researchers and other data users, saving time and resources, and enabling a wider range of research questions to be addressed.

UKDS provides highly efficient and cost effective curation and provision of data for social science research ensuring delivery of public benefit from investment in data resources.



## 1.4 Next steps

2021

July - September

ESRC Data  
Infrastructure  
Strategy  
Stakeholder  
Engagement

September

ESRC Data  
Infrastructure  
Expert Advisory  
Group meeting to  
discuss stakeholder  
feedback

November

ESRC Data  
Infrastructure  
Strategy update  
provided to ESRC  
Council

2022

Spring

ESRC Data  
Infrastructure  
Strategy published  
/ Implementation  
begins

### LONGITUDINAL STUDIES

ESRC's suite of longitudinal studies provides an unparalleled shared research resource. By collecting data at intervals from the same people over their lives, our longitudinal studies are continuously building an ever-richer data resource that captures how life in the UK changes over time and how events and exposures have long-term impacts.

The UK has longstanding investments in longitudinal data. These include:

- the flagship cohorts run by the **Centre for Longitudinal Studies (CLS)**, which includes a series of birth cohort studies that stretch from the **1958 National Child Development Study** to the recently commissioned **Early Life Cohort Feasibility Study**, plus other aligned cohorts.
- **Understanding Society, the UK Household Longitudinal Study**: one of the largest longitudinal household panel studies of its kind, with over a decade of detailed information on the life changes of its participants. The study has been a crucial resource for informing government policy, as illustrated recently by SAGE using the data to advise government on vaccine hesitancy amongst ethnic minorities.

If funding for these longitudinal studies were to cease, we would not have the benefit of rich, flexible, shared data resources relevant to the current era. Without the benefit of these continued investments, policy makers would not have been able to access high quality, robust, representative data across a range of studies in order to identify rapid societal changes and to inform critical policy decisions taken around the COVID-19 pandemic.





## 2.

# Position statement

The UK is facing unprecedented changes, challenges and opportunities. This creates a pressing need for economic and social science data infrastructures that are relevant to the current era, fit for purpose, agile and take advantage of innovative forms of data, technology and methodology. Such investments will enable high-quality, high-impact research to inform responses that maximise public benefits for the UK.

Social science is becoming more data intensive. New technology, data resources and linkages create new research opportunities and enable innovative approaches, cutting across traditional domain boundaries and pushing the frontiers of scientific enquiry for public good. Expectations are increasing regarding the integration and interoperability of infrastructures across disciplines. Public recognition of the importance of ethical, lawful, safe and secure handling of data is greater than ever.

By 2022 we will develop a Data Infrastructure Strategy that will shape ESRC's investment in research data infrastructure. It will establish guiding principles and priorities to inform our investment approach over the next five years (2022–7) and beyond, underpinned by priority research interests relevant to both UKRI and UK governments. ESRC cannot deliver this alone and collaboration between a range of organisations and communities will be essential. We are therefore seeking advice from stakeholders to help us shape our investment in social science data infrastructure for the years ahead.

## ELECTION STUDIES

The ESRC-funded **British Election Study (BES)** has carried out large-scale surveys of the British electorate immediately after every general election since 1964. It is one of the longest running election studies worldwide and the longest running social science survey in the UK – looking at why people vote and why they vote the way they do. ESRC also funds **aligned studies of devolved elections in Northern Ireland, Scotland and Wales**, as well as referendum studies.

This very long-running set of studies provides high-quality, comparable, time series data on representative samples to enable in-depth understanding of the way representative democracy functions in UK elections and provides insights into UK public opinion, participation, attitudes towards politics and explanations of electoral outcomes.





# 3.

## Objectives



The Data Infrastructure Strategy should enable ESRC to:

- transparently prioritise and balance needs between sustaining current investments, pursuing new opportunities and supporting innovation.
- establish a pipeline of ideas so we can respond to and access new funding opportunities.
- optimally manage our investment portfolio, including identifying, measuring and communicating impact.
- adapt or finish the funding of data infrastructures in a fair and transparent way when further investment would not provide value for the public compared to other opportunities.

The impact the Strategy is looking to achieve (as outlined in the Theory of Change in Figure 1) is as follows:

**The delivery of public services is improved through better policies, which are informed by the effective application of highly valued ESRC data investments and infrastructures by skilled researchers. These improvements, underpinned by public trust, deliver enhanced societal and economic outcomes, improved response to major challenges, and enable the advancement of science through innovation and from taking advantage of the opportunities that data offers. Embracing the value of data and effective data use enhances the UK's reputation as a front runner in innovation, which encourages collaboration and the international flow of data.**





# 4.

## About the Data Infrastructure Strategy





#### 4.1 ESRC's role in the social science data infrastructure system

ESRC is the main public funder of social science data infrastructure in the UK, investing over £200m over the last five years (not including the ESRC-funded Administrative Data Research UK (ADR UK) programme). These data infrastructures include data collections and the facilities and services which enable them to be fully, ethically and safely used to achieve excellence in research and innovation.

The UK has excellent social science data infrastructure by global standards and ESRC's approach and investments form an important part of the wider research and innovation landscape. Social science research is wide-ranging and is often essential to maximising public benefits from other research disciplines.

Our role as an investor of public funds includes:

- Providing a backbone for research – investing in critical underpinning infrastructures which operate across the UK and constituent nations for wide use, with broad public benefits.
- Providing leadership – monitoring and shaping the social science and broader data infrastructure landscape, making connections, coordinating stakeholders to pool resources, driving efficiency through synergies, ensuring infrastructure can support evolving priority public policy needs.
- Fostering innovation – supporting exploration of and investment in emerging, innovative technologies or methodologies that are promising but unproven.
- Driving impact – ensuring that investments and ESRC leadership activities maximise and clearly articulate the impact and value for money for the public.

A list of ESRC's current investments in data infrastructure is provided at **Annex 1**.



## 4.2 Why a Data Infrastructure Strategy is needed

Social science is already in a strong position for data infrastructure due to ESRC's long-standing commitment to shared research infrastructure investment. ESRC's data sharing policy, supported by funding for resources and infrastructure (UK Data Service, data standards, training, culture) which enable broad use of data resources, means our data infrastructures are recognised globally for their contribution to research. ESRC's championing of representative population data ensures that our data infrastructures are well-placed to support rapid research responses to emergent public and policy challenges.

ESRC has invested in an evolving portfolio of data infrastructures since our inception over 50 years ago, to create public benefits and to support the Council's strategic objectives. Developments in recent years include new infrastructures to support use of administrative data (ADR UK) and the Urban Big Data Centre and Consumer Data Research Centre (UBDC and CDRC), as well as completing a major

review of longitudinal studies.<sup>1</sup> Alongside this, ESRC has continued to invest in developing long-standing infrastructures that accumulate in value over time.

Meanwhile, technological and methodological advances, new legal frameworks, and growth in the amounts and types of data being generated create a dramatically different context for ESRC's investment. While exciting, this means there are far more worthwhile investment opportunities than our financial allocation for data infrastructure can support.

All this creates a need for a Data Infrastructure Strategy that enables ESRC to make the best use of our resources, maximise the value of our investments and build flexibility into the data infrastructure portfolio to address long-term, emerging or changing needs.

<sup>1</sup> <https://esrc.ukri.org/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/>

### ADMINISTRATIVE DATA RESEARCH UK (ADR UK)

ADR UK's work plays an important role in bridging the gap between government and academia, helping enable government policy to be informed by the best evidence available and putting us on the path to a future in which the true potential of administrative data to improve society is realised.

ADR UK has continued to make significant progress and there are now several interconnected programmes where ADR UK infrastructure is a dependency:

- The Office for National Statistics (ONS, ADR UK's major data infrastructure partner) contribution to the UKRI National Core Studies (NCS) Data & Connectivity programme is funded through ADR UK, facilitating research access to a range of COVID-19 datasets.
- Similar contributions to COVID-19 data collection, access and research have been made by ADR UK Government and academic partners in the devolved administrations, directly leading to a suite of research outputs based on linked administrative data feeding through to SAGE and devolved administration equivalents.
- ADR UK funded a pilot collaboration between the Joint Biosecurity Centre (JBC), the Department of Health and Social Care, Ministry of Housing, Communities and Local Government and an academic team to help support local authorities in England to make use of the COVID-19 related datasets held within the ONS Secure Research Service.
- A suite of Research Fellowship calls has been launched, showcasing the new, linked datasets that ADR UK are making available to researchers. These include crime and justice data from the ADR UK-funded Ministry of Justice Data First programme, Police National Computer and education data. A future call is being designed to showcase 17 years of linked education and NHS Digital health data; the largest dataset NHS Digital have ever made available.
- In collaboration with the No. 10 Data Science team, ADR UK and ESRC are launching a Data Science Fellowship scheme, bringing academic expertise into the heart of government decision-making.
- The ADR UK programme is now gaining international recognition, with Co-Directors of the International Population Data Linkage Network (IPDLN) for 2020 and 2021. This involves running a major ADR UK/IPDLN international conference in Autumn 2022.

### 4.3 Vision for the Data Infrastructure Strategy

The Data Infrastructure Strategy should enable ESRC to deliver this vision:

*ESRC's data infrastructure investment ensures unique, high quality data will be used to address major societal challenges and advance the UK's reputation as a world leader in social science. Under this strategy, ESRC's investment will become more valuable, innovative, and integrated; ensuring the flexibility to respond to new opportunities and maximise the economic and societal contribution of data infrastructures and resources.*

To achieve the vision, the ESRC Data Infrastructure Strategy will have two components:

- A written strategy, for publication in 2022, which outlines our intended approach to investing strategically in data infrastructure over the next five years.
- A programme of activities and new approaches which supports the implementation of this strategy from 2022, building on our strengths and learning.

Much of the implementation will be incremental, embedded in our management of new funding opportunities and existing investments. In the implementation phase, we expect the following to be critically important:

- a transparent decision-making framework for funding.
- responsive funding mechanisms.
- aligned objectives for investments to maximise collaboration, integration, innovation and user-focused services.
- a clear impact strategy supported by effective communications and engagement.

### 4.4 ESRC Data Policy

The Strategy will be underpinned by ESRC's long-standing, evolving data policy, founded on the principle that the data resources ESRC funds are shared research infrastructure. ESRC will continue to support this principle through funding resources for data collection or creation, for cleaning and documentation to specified standards, for storage, curation, discovery and access via the UK Data Service and other platforms, for user support and training, and through measures to support this culture including data citation and use of appropriate metrics, and further development of the Secondary Data Analysis Initiative (SDAI). As part of the strategy development, the Data Policy is currently being updated and expanded.



## 4.5 What activities will support the Strategy's development and delivery?

The ESRC Data Infrastructure Programme is progressing work in a number of related areas to inform changes in the way that we work and the development of frameworks to support the implementation of the Data Infrastructure Strategy. This work interacts with the business as usual work of the team, which provides opportunities to test approaches and to learn. The Data Infrastructure Programme is engaging with other ESRC teams and other Research Councils to inform this work.

- **Impact:** to demonstrate more effectively to a wide range of stakeholders, the academic and wider economic and societal impact and value of investments, through our actions as a funder and through supporting our investments in order to recognise, value and magnify the meaningful outcomes and impacts arising from investments.
- **Finance:** to support both large, longer-term investments as appropriate and provide a more agile, responsive approach including enabling the development of new innovative ideas and potential new investments.
- **Landscape mapping:** to develop tools and processes enabling ESRC to create and regularly update a landscape map of the Data Infrastructure portfolio and other relevant social science research infrastructures.
- **Governance and assurance, monitoring and evaluation, reporting:** to outline how ESRC should govern, assure and manage investments.
- **Common assessment framework and proposals pipeline:** to create a framework to enable ESRC to invest in the most strategically important initiatives as funding opportunities arise.
- **Management information:** to set up the means to deliver information to provide an evidence-based approach to the development and delivery of ESRC's Data Infrastructure Strategy and ensure efficiency and effectiveness in how this information is acquired and curated.

- **Sustainability/realising long-term value:** to ensure that ESRC's mechanisms that seek to maximise the long-term value created through our investment in data infrastructure are fit for purpose and support the vision of the Data Infrastructure Strategy.
- **Communication and engagement:** to increase the recognition, by specific audiences, of social science data infrastructures' value and for engagement activities to enable well-informed investment decisions.

In addition, ESRC needs to consider how data services will be delivered in the future. ESRC has launched a programme of work called '**Future Data Services**' (FDS) to consider the optimum approach to making future data service investments. The aim of FDS is to ensure that ESRC gathers the evidence it needs to explore and assess the future needs of its communities (recognising that research is more cross-disciplinary and cross-sector in nature than ever before). As part of the FDS work, an options appraisal regarding future data services provision for the social sciences and beyond will enable ESRC to make timely evidence-based decisions. Work involved will include mapping, reviewing, consultation and scoping. New data services are expected to be commissioned from Autumn 2024 onwards.





# 5.






**Data Infrastructure  
Strategy theory of  
change and draft  
objectives**

To support delivering the vision, ESRC has produced a Theory of Change which shows how change can happen in the short and long term to achieve identified impacts. See **Figure 1**.

The Theory of Change method identifies a logical flow, defines what our work contributes to and indicates how each section is connected.

This method also identifies an evidence of need so that it is established with a clear understanding of the issues or opportunities we want to address. The overall desired impacts describe the broad or long term change we want to see happen.

The colours show specific pathways of change that will interact to deliver the identified impacts:

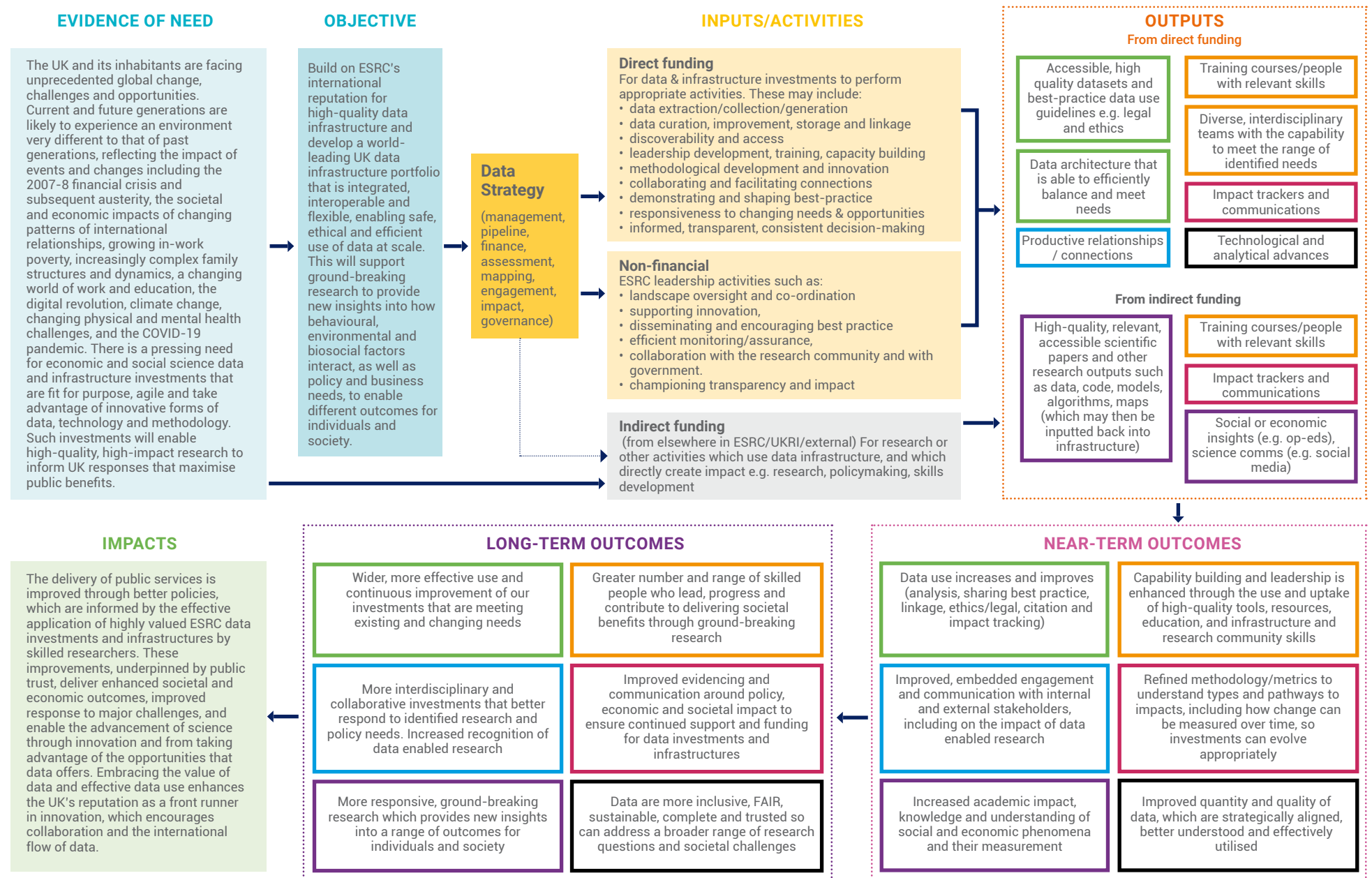
-  **Best-practice data use**
-  **Communication and engagement**
-  **Capacity building**
-  **Research impact**
-  **Improved data quality**

The Theory of Change shows how embedding the activities set out in the Strategy will enable ESRC to fulfil its role in a more strategic and streamlined way, which will maximise data infrastructure investments' contribution to real-world impacts.

In addition to the Theory of Change, an objectives diagram (see **Figure 2**) has been drafted which will also support delivering the vision and will be used as guiding principles for ESRC, its investments and investment decisions. This figure outlines areas of focus, goals and objectives of the Strategy.



**Figure 1: Theory of change**





**Figure 2: Objectives**

<b>VISION</b>	ESRC's data infrastructure investment ensures unique, high quality data will be used to address current and future major societal challenges, and advance the UK's reputation as a world leader in social science. Under this strategy, ESRC's investment will become more valuable, innovative and integrated; ensuring the flexibility to respond to new opportunities and maximise the economic and societal contribution of data infrastructures and resources			
<b>AREAS OF FOCUS</b>	<b>Building and sustaining a foundation</b>	<b>Leadership and connectedness</b>	<b>Responsiveness, flexibility and innovation</b>	<b>Impact, engagement and public benefit</b>
<b>GOALS</b>	ESRC's portfolio of data investments is a foundational pillar for UK social science's role in research and policy-making	ESRC and investments works collaboratively and makes connections between people, organisations and infrastructures	ESRC and investments take decisions that support innovation and changing research and policy needs	ESRC and investments facilitate public and community engagement to ensure its investments deliver benefits for all UK communities
<b>OBJECTIVES</b>	<p>To create evidence to inform investment in and lay foundations for a sustainable mainstream research resource, complementing other UKRI infrastructures</p> <p>To ensure research users across all sectors and localities can discover and access data and data-driven research in order to deliver societal benefits</p> <p>To build on the UK's position as a highly collaborative player that makes a significant contribution to national and international challenges</p>	<p>To demonstrate, empower, develop and facilitate collaborative leadership in the research data landscape</p> <p>To support and facilitate effective integration and working together</p> <p>To enhance connections with other research councils to explore linkages with other disciplines</p>	<p>To invest proactively, transparently, and agilely to support innovation and capacity building of both people and infrastructure potential</p> <p>To ensure a holistic, comprehensive, ambitious and sustainable approach over the longer term with due focus on funding considerations</p>	<p>To ensure the performance and impact of both individual investments and the whole portfolio form part of a comprehensive reporting, monitoring and evaluation strategy</p> <p>To develop greater understanding of different outcomes for individuals and groups, through data and data-enabled research that delivers social and economic benefits</p> <p>To engage and secure recognition from data owners, policy-makers, other users, and the general public that ethical data-enabled research has delivered social and economic benefits</p>

# Annex 1: ESRC managed investments in data infrastructure

The following ESRC investments are currently managed by the ESRC Data Infrastructure Programme (ADR UK is managed by the ESRC's Strategic Hub team). These are categorised as Data Collections or Data Services, which can be defined as follows:

- **Data Collections** investments collect new data for research purposes.
- **Data Services** investments make available existing and new data for research purposes and facilitate the use of data through (among other activities) the provision of support, training and curation.

In practice, many investments include elements of both collections and services. There are also many more active ESRC research grants, not listed here, which generate research data which is deposited for re-use.

## DATA COLLECTIONS

### Centre for Longitudinal Studies (CLS) cohorts

1958 National Child Development Study

1970 British Cohort Study

Next Steps

Millennium Cohort Study

COVID Social Mobility and Opportunities Study

Early Life Cohort Feasibility Study

### Other cohorts

Born in Bradford (as co-funder)

English Longitudinal Study of Ageing (as co-funder)

Generations and Gender (UK lead funder)

Healthy Ageing in Scotland Survey (as co-funder)

### Household panel studies

Understanding Society

(UK Household Longitudinal Study)

### Election studies

British Election Study

Northern Ireland Election Study

Scottish Election Study

Welsh Election Study

### Other surveys

European Social Survey (UK lead funder)

International Social Survey Programme  
(UK lead funder)

National Survey of Sexual Attitudes and Lifestyles  
(as co-funder)

Skills and Employment Survey

Workplace Employment Relations Survey  
(as co-funder)

## DATA SERVICES

UK Data Service

### Census longitudinal studies

Centre for Longitudinal Study Information and User  
Support (CELSIUS)

Northern Ireland Longitudinal Study Research  
Support Unit

Scottish Longitudinal Study Development and  
Support Unit

### Big data

Consumer Data Research Centre

HateLab

Human Rights and Big Data Technology

Social Data Science Lab

Urban Big Data Centre

### Other services

Cohort and Longitudinal Studies Enhancement  
Resource (CLOSER)

Consortium of European Social Science Data  
Archives (CESSDA) – (UK lead funder)

Health Data Research UK (as co-funder)

Integrated Census Microdata

Software Sustainability Institute (as co-funder)



Economic  
and Social  
Research Council

