



## **Research Data Scotland: 2022-23 business plan**

**Simplifying & speeding up data access:  
moving RDS from start-up to mature footing**

## **Purpose**

**This document summarises organisational priorities for 2022-23, how we'll use our income to deliver those priorities, and how this supports our wider strategy<sup>1</sup>.**

## **RDS Vision**

Our vision is to promote and advance health and social wellbeing in Scotland by enabling access to public sector data about people, places and businesses, and creating the conditions for collaborative data-driven research and innovation in the public good using that data.

## **RDS Mission**

To achieve our vision, RDS needs to coordinate access to the excellent data already collected about people, places, and businesses in Scotland, helping researchers and analysts to access that data in a timely and cost-effective way, enable working across boundaries to collaborate on common datasets, and access to the analytical tools they need to bring cutting edge insights that inform great decisions. We will enable research on different scales: from large international studies, to simple and quick analyses of individual datasets, and support people to use the latest data science and geospatial approaches.

RDS will build a portfolio of research data assets that are held securely with public support and enable these datasets to be up to date and accessed at pace. RDS will introduce new services that helps organisations to understand the impact of decisions they make and better hones their improvement of wellbeing and equalities.

## **Position at the start of 2022-23**

### **Partners**

RDS is owned equally by Scottish Government, Public Health Scotland and the Universities of Dundee, Edinburgh and Glasgow. We are working in partnership together with the National Records of Scotland to deliver on our shared vision. The **electronic Data Research & Innovation Service (eDRIS)** team in Public Health Scotland provides support for researchers in accessing data, and together with the Edinburgh Parallel Computing Centre (EPCC) at the University of Edinburgh manages the National Data Safe Haven – a Trusted Research

---

<sup>1</sup> The 2022 strategy can be downloaded alongside 2022-23 business plan [researchdata.scot/about-us](https://researchdata.scot/about-us)

Environment providing very secure computing to store and process data. National Records of Scotland provide an indexing service connecting data around a person. Scottish Government leads data acquisition as part of the Administrative Data Research (ADR) Scotland programme.

In February 2022, RDS signed a collaboration agreement with the four Regional Safe Havens (i.e. Glasgow, Tayside/Dundee and Grampian/Aberdeen and Lothian/Edinburgh) that bring together Universities and the local NHS to provide a range of services around detailed healthcare data. This outlines how we will work together to align data, data services, communications, and some processes. Universities associated with Regional Safe Havens in Scotland are taking steps to become company members. Linked to this, RDS is a member of the Health Data Research Alliance, which is a collaboration of data provisioning and research organisations united to establish best practice for the ethical use of UK health data for research at scale.

RDS is starting to work with the Office for National Statistics as a collaborator in the development of their Integrated Data Service, which had similar ambitions to RDS in connecting data from across UK Government and Devolved Administrations.

## **Services**

The eDRIS team provide three services for researchers

- a **researcher consultation and information service** that helps researchers at the start of projects understand what might be possible so they can make the case for funding. This is not currently revenue generating but provides a valuable service to researchers understanding of the processes and datasets available.
- a **research coordination service** that supports researchers in accessing data. This provides support for researchers in securing permissions to access data, the provision of the data and of a computing environment with analytical tools. They charge for this service, though the revenue received for providing this service does not currently cover their total costs.
- A **data analysis service** that supports people where they need some additional support on using the data. This is charged per day at cost.

Scottish Government provides a **data acquisition service** as part of delivery for the ADR-Scotland programme. This makes arrangements with data controllers for the use of data in research, and supports the process of bringing a copy of data into the National Safe Haven and making it “research ready”. As part of this service, National Records of Scotland (NRS)

index the data, which de-identifies the data whilst enabling it to be connected around person for an approved research project.

## **Users**

eDRIS is currently handling over 500 live projects and enquiries. These are predominantly for academic and NHS researchers, and are for access to NHS data, either from single or multiple datasets. They also support the academic teams as part of the ADR Scotland programme and have a few projects from Scottish Government analysts. These projects tend to be more complex and include linkage of data across different topic areas.

## **Data**

There are around 70 datasets listed in the RDS catalogue. These cover many aspects of NHS activity (GP, hospital care, and prescribing), the 2011 census, vital events, and datasets about public services for children. Many of these datasets have been brought together through the Administrative Data Research Scotland and Health Data Research UK National Core Studies programmes. Some datasets are very up to date and have regular feeds that keep it so. And some of the data have significant restrictions upon their use. There are also the Scottish Medical Imaging datasets that have been brought together through the PICTURES programme.

The datasets have often been used in production of official statistics and have been enhanced on their ingestion.

There are also some standing linked datasets in development, such as a dataset of protected characteristics. And through the ADR-Scotland programme a number of datasets are being prepared (particularly around justice, veterans, and household surveys). We also can create household level datasets from data about individuals.

## **Priorities for 2022-23**

Our focus will be on improving the speed and quality of the research data access service. We have developed a detailed service design blueprint that establishes user journeys that are quick with reduced effort whilst maintaining scrutiny of public value for all uses of data. We will have a detailed delivery programme underway to create the components of that service by June and will publish this later in the summer.

We will develop and roll out streamlined approaches to information governance for the use of data for research. Our intention is to put all datasets current available for research in the Covid and ADR-S (Administrative Data Research Scotland) data holdings in the National Data Safe Haven based on this innovative approach – agreeing the extent of use with data controllers. We will also start to engage with a wider range of data controllers for the next wave of datasets we will look to make available for research in the public good.

And we will move Research Data Scotland from a start up to a more mature footing, regularising our budgeting, people, policy, and governance functions, filling key positions and skills gaps on a permanent basis where this is needed, and using the flexibilities RDS possesses to support our partners.

We will continue to take user informed approaches that simplify and standardise where possible across Scotland, the UK and internationally. Public engagement will be at our heart, establishing approaches to do this systematically around Scotland. In our work, we will take every opportunity to promote inclusion and demonstrate trustworthiness. And we will be a learning organisation – looking at best international practice and working in iterative ways. As such, we will establish a small set of demonstrator projects to test new data and innovative approaches, taking time to learn and adjust what we do.

With initial management data about the length of time projects take and from hearing user views about what will make the biggest difference, we are confident that our plans are focussed in the right areas. However, a lot of change is required, and this will not all be achieved this year. So, it is important that we phase our work well and take actions that will set RDS up to move forward at pace going forward.

## **During 2022-23 we plan to do the following**

### Partnerships and communications

- Develop and publish a RDS public engagement strategy, establishing a Scottish public engagement for research data network, using this to establish ongoing approaches to public engagement, testing out changes we make with the public
- Work with UK partners through HDR-UK (Health Data Research UK) to develop robust and collaborative approaches to public engagement.
- Develop formal agreements with RDS founding partners: Public Health Scotland, National Records of Scotland, Edinburgh Parallel Computing Centre to clarify roles and expectations
- Start driving forward approaches to federate regional healthcare data through the Regional Safe Havens Technical Working Group, and with UK data through the Office for National Statistics (ONS) Integrated Data Programme around business data
- Agree a memorandum of understanding with the ADR-S programme to ensure we are making the most of our collective investment
- Develop strategic partnerships with organisations: in the health innovation space, around services to support machine learning and other Artificial Intelligence approaches, and with other research data services in Scotland.
- Support the work on “unlocking the value from public sector data” and consider the impact of its findings for how best to work with industry

### Services

- Develop a system blueprint for the Researcher Access Service that simplifies the researcher user journey
- Establish a user engagement panel to ensure changes meet expectations of researchers.
- Scope and start delivery of digitisation of the researcher application to access data
- Further refine the Scottish metadata catalogue for datasets available in the National Safe Haven
- Put together comprehensive guidance and support materials, and offer training for new researchers in the data access process, setting clearer expectations and roles based on the new user journey.
- Support development of the eDRIS model, building skills in the new user journey and newly available datasets.

- Consider whether we need to develop an approach to cohort creation, based upon user demand.

### Data

- Establish the data acquisition team.
- Rework the agreements we have with data controllers of existing ADR-Scotland and Covid data holdings to broaden use under the SECURE (**S**afe, **E**nhanced, **CUR**ated **E** dataholding) data access approach
- Bring together data from key administrative and census data to create a pseudonymised protected characteristics dataset available for research
- Complete the pilot project with the ONS Integrated Data Programme to improve the range and interconnectivity of data on Scottish businesses and use learning from this to better connect key UK wide and Scottish datasets.
- Further simplify and standardise the process for data acquisition, in particular our offer to data owners to work with them, and support data quality improvements as part of the data preparation process.
- Establish a community of practice around synthetic datasets, agree common policies and approaches, and pilot and agreed approach to offering synthetic datasets to researchers
- Start working with further data controllers on children, post-16 learning and healthcare to scope the potential for making this data available during 2023.
- Work with EPCC (Edinburgh Parallel Computing Centre), NRS (National Records of Scotland) and eDRIS teams to refine the way that data is securely stored in the National Safe Haven, and simplify the ways that data flows.

### Information Governance

- Agree a new SECURE Information Governance approach. Develop, and deliver a plan for its introduction, and establish a group of data controllers to oversee its operation.
- Start to build clear, consistent approaches to Information Governance to data access for research across the Scottish public sector working to understand the range of approaches used, the differences between these and opportunities for simplification.

## Organisational Development and Governance

- Complete the establishment of RDS the organisation: establishing the board and the other governance and advice needed,
- Agree and publish an organisational strategy, finish initial recruitment of staff, and establish appropriate financial/performance management and procurement approaches.
- Finish establishing relevant organisational policies
- Refine the funding model for the researcher access service so it better reflects how services are consumed, and establish RDS role in managing finances for 2023-24 onwards
- Aim to secure Independent Research Organisation status with UKRI (UK Research and Innovation) and start to lead collaborative research bids for funding relevant to RDS' mission.
- Develop thinking about approaches to diversify funding

## **Medium-term (from April 2023)**

In the remainder of 2023, we will focus on widening the range of service users and datasets we offer whilst further improving the speed and quality of data access services to researchers. During 2023 and 2024 our aim is to develop new services and new valuable research data resources. We are planning to do the following...

### Partnerships and communications

- Ramp up our marketing and communications about the Researcher Access Service offer to attract academic and Scottish public sector analysts as service users.
- Continue to work as part of a UK public engagement network to support a public communications campaign about trustworthy research using administrative and clinical data.
- Work with other Scottish organisations to take the message of our services across the UK and internationally, and attract investment into Scotland
- Develop strategic partnerships with organisations across Scotland, UK and internationally to allow greater federation of data and collaborative research.

### Researcher Access Services

- Complete work to digitise the application for research data
- Continue to improve guidance, support materials, and training for new researchers in the data access process, and on the datasets we have made available for reuse.
- Make the production of synthetic datasets to researchers a standard approach.
- Implement new approaches to cohort creation and federated analysis that help researchers refine their asks for data and speed up data access.
- Broaden the range of analytical software and improve compute power in response to user priorities
- Enable researchers to bring together key open datasets alongside case level data
- Develop approaches to automation of published dashboards driven by data held through RDS
- Examine approaches to automate privacy enhancement, disclosure control and other parts of the data access and linkage process
- Develop a “connect your data” service to enable researchers to bring their own data in ways that integrate with existing RDS curated datasets

- Work up a “measuring impact service” that allows organisations to share details of people they are working with and receive a report on the impact of that service

### Data Sourcing

- Move forward with data acquisition and construction of standing cohort datasets based upon user feedback
- Further simplify and standardise the process for data acquisition, in particular our offer to data owners to work with them, and support data quality improvements as part of the data preparation process.
- Further automate the process of data ingestion with a focus on improving the timeliness and regularity of the data held through RDS.
- Complete the pilot project with the ONS Integrated Data Programme to improve the range and interconnectivity of data on Scottish businesses, and use learning from this to better connect key UK wide and Scottish datasets.
- Establish further research cohorts, building on existing longitudinal studies and in response to research needs
- Ensure data flowing the ADR Scotland programme are available to researchers and there are approaches to updating these datasets.

### Information Governance

- Continue to take steps to standardise approaches to access to data for research across the Scottish public sector, aligning where possible with UK and international best practice.
- Work to introduce streamlined processes for securing access to data where data ownership is currently fragmented.
- Establish a group of expert users across each of individual and linked datasets, and therefore approaches to building a wider understanding in the use of these datasets

## Organisational Development and Governance

- Establish further tangible examples of investment into Scotland because of the RDS offer. This requires RDS to work with Scottish Government, Scottish Trade International and several other partners to establish a data driven innovation offer to potential investors that is rooted in use of data for public benefit.
- Continue to scan for international best practice in keeping data safe, acting in trustworthy ways, and speeding up access to data.
- Take steps to diversify funding sources and put RDS on a sustainable financial footing.

## Budgets

RDSS was given a budget of £6.0 million for 2022-23. This included using £2.4 million underspend from 2021-22.

	Budget (£m)
Partnerships and communications	1.0
Of which support for projects with Regional Safe Havens	0.7
Researcher Access Services	3.0
Of which support for projects with eDRIS	1.0
Data sourcing	1.1
Of which support for demonstrator projects	0.6
Information Governance	0.5
Organisational Development and Governance	0.4

## Measuring success

A set of initial KPIs were agreed by the Research Data Scotland Board. Their intention is to help measure and manage progress against priorities for 2022-23 and the overall outcomes RDS is aiming for.

### RESEARCH DATA SCOTLAND KEY PERFORMANCE INDICATORS,

		Recent performance	Target performance		
		Apr 21-Mar 22	2022-23	2023-24	2024-25
<b>1</b>	<b>Outcomes</b>				
1a	Number of jobs created	No data	Need baseline data to set target levels		
1b	Published impact case studies	No data	Need baseline data to set target levels		
1c	Number of citations to research / analysis derived from RDS data	No data	Need baseline data to set target levels		
1d	Number of publications using RDS data	No data	Need baseline data to set target levels		
<b>2</b>	<b>Use of Researcher Access service</b>				
2a	Number of new enquiries	259	259	311	373
2b	Projects supplied with data	49	55	93	187
2c	Conversion rate: projects supplied with	19%	21%	30%	50%

	data as proportion of all enquiries				
2d	Enquiries from Non-NHS or academic users	Government Oth pub sector 3 <sup>rd</sup> sector other	Need baseline data to set target levels		
<b>3</b>	<b>Speed of Researcher Access Service</b>				
3a	Average project duration  - Initial inquiry to researcher access (median days). Sum of 2a - 2c	207 days (13 projects)	179	114	74
3b	Initial inquiry to first Data Access panel submission (median days)	87 days (36 projects)	78	55	38
3c	First Data Access panel submission to approval (median days)	50 days (26 projects)	45	32	22
3d	Latest Data Access panel approval to researcher access (median days)	70 days (1 project)	56	28	14
<b>4</b>	<b>Data offering</b>				
4a	Number of datasets available via eDRIS	72 (May 22)	72	To be agreed	To be agreed
4b	Number of synthetic datasets available	0 (May 22)	1	To be agreed	To be agreed
<b>5</b>	<b>Marketing / Communications</b>				
5a	Website Traffic	page views <sup>2</sup> = 2,424	To be agreed	To be agreed	To be agreed

<sup>2</sup> Tracking was only activated on 20 April 2022

		Users <sup>1</sup> = 475			
5b	Social Media Likes	Engagement rate <sup>3</sup> = 3.3%  Likes = 142  Link clicks=212  Retweets=99	To be agreed	To be agreed	To be agreed
5c	Social Media Post Reach	Impressions/day [Twitter] = 206  Page views [LinkedIn] = 643  Unique visitors [LinkedIn] = 228	To be agreed	To be agreed	To be agreed
5d	Customer feedback – user forum attendance	Sep '21 = 116  Nov '21 = 91  Feb '22 = 102  May '22 = 82	To be agreed	To be agreed	To be agreed

---

<sup>3</sup> the proportion of people who engaged with our content