Industry access to public sector data:

Review of current operational practice

Research Data Scotland

23/05/2023

# Purpose and introduction

1. The then Scottish Government Minister for Business, Trade, Tourism and Enterprise, Ivan McKee, asked Research Data Scotland to lead a review of current operational practice in the use of data by the private sector in research as a contribution to the Unlocking the Value of Data programme.
2. The Unlocking the Value of Data programme was established by Scottish Ministers in early 2022. They established an Independent Expert Group to oversee the development of a high-level policy statement and with a view to developing a national Framework to guide ethical and confident decision-making by Scottish public bodies on personal data access by or with the private sector.
3. While a fully defined and specific Scottish Government policy position has yet to be developed, two current Scottish Government documents provide some direction, although both were produced a number of years ago.
	1. The [Data linkage guiding principles,](https://www.gov.scot/publications/joined-up-data-better-decisions-guiding-principles-data-linkage/) which were updated last in 2014 and include a relevant paragraph:

1.5 Where linkages resulting in commercial gain are envisaged, this should be clearly and publicly articulated and widely communicated.

* 1. The [2015 Safe Haven Charter](https://www.gov.scot/binaries/content/documents/govscot/publications/agreement/2015/11/charter-safe-havens-scotland-handling-unconsented-data-national-health-service-patient-records-support-research-statistics/documents/charter-safe-havens-scotland-handling-unconsented-data-national-health-service-patient-records-support-research-statistics/charter-safe-havens-scotland-handling-unconsented-data-national-health-service-patient-records-support-research-statistics/govscot%3Adocument/00489000.pdf) (which was produced in relation to health data accessible through the safe havens within the governance of the NHS). The key principle of this is laid out below.

“Safe Havens work in partnership with academia, public service providers and industry to undertake research using de-identified or anonymised data that is in the public interest. However, personal data cannot be sold by a Safe Haven or transferred to a commercial organisation. Nor can they be transferred, nor access provided, to a third party (i.e. researchers or others) unless specified explicitly by the Data Controller(s) holding the source records and unless the third party operates to, at minimum, equivalent standards and with equivalent safeguards.”

1. There is no policy statement covering the operational issues that underpin that principle that would guide consistent interpretation and implementation. However, this document attempts to address that by describing current practice.
2. The lack of clarity to direct industry ‘access-to’ data means that innovative approaches to improving people’s health and societal wellbeing are less likely to be researched and discovered in Scotland. This in turn means poorer outcomes for people in Scotland and fewer high paid Scottish based research jobs in both the public and private sector.
3. As such, the paper is a report to Scottish Government and is in two parts.
4. First, there is a description of the current arrangements experienced by industry in accessing public sector data for research in the five Trusted Research Environments (also known as Safe Havens) in Scotland. That is case level data relating to a specific person, place or business.
5. Second, there is a summary of a set of issues that need further work to reach a firm set of operational policies in this space that strike the right balance between greater research use and value coming from Scotland’s data with public view on appropriate uses of their data.
6. This document is also likely to be useful for those researchers from academia, industry or the Scottish public sector who are looking to use public sector data in their work, controllers of public sector data in Scotland and for those providing secure data access services.
7. The review was undertaken through a series of interviews with actual and potential users of data working in industry, academia and the public sector, with those working in Scotland’s Safe Havens, with data owners, and panels who support decision making about use of public sector data. In addition, a series of workshops were arranged by the Innovative Healthcare Delivery Programme (IHDP) for private sector organisations wanting to use data. This work has predominantly looked at the use of health data by the private sector. This is because health is the area with the most clearly articulated industry demand for case level data to enable research in the public good.
8. The findings and conclusions have been tested with these groups, and with a short life working group of the Scottish Government’s Health and Care Data Board. Many thanks to those at IHDP who set up and ran those workshops and to all those who fed into this review.

# Definitions

1. **Private Sector company**: Private sector organisations are driven by profit, though many businesses deliver public benefit through their activities. In the context of research using public sector data, there are different types of interaction private sector firms currently have, such as:-
	1. Companies may be directly funded by the public sector to do work. Sometimes all their costs would be covered e.g. consultancies commissioned by public bodies to provide analytical/research services. Sometimes part of their costs may be covered e.g. an company wins an innovation grant that match funds their investment.
	2. Companies who have ongoing programmes of work/partnerships with academia or public sector researchers with established track records in public benefit from their use of public data. For example, pharmaceutical companies work closely with academic and NHS research teams on new medical treatments.
	3. Firms who don’t have established connections to public sector or academic researchers or history in using public data well, for example Start ups who are looking to developing new technologies using data analytics.
	4. Businesses who generate valuable data and are keen to derive further public value by connecting that data to other public sector data assets.
2. [**Trusted Research Environment**](https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens): not-for-profit highly secure computing spaces that act as data custodians and provide remote access to health data for approved researchers. Trusted Research Environments (TREs) can also be called Data Safe Havens and Secure Data Environments. In Scotland, they are independently accredited to [ISO 27001](https://www.iso.org/isoiec-27001-information-security.html) standards and the National Safe Haven is [accredited by the UK Statistics Authority](https://uksa.statisticsauthority.gov.uk/digitaleconomyact-research-statistics/better-access-to-data-for-research-information-for-processors/). In Scotland there are five Trusted Research Environments, namely the [National Safe Haven](https://www.isdscotland.org/Products-and-Services/eDRIS/), [West of Scotland Safe Haven](https://www.nhsggc.scot/hospitals-services/services-a-to-z/west-of-scotland-safe-haven/#:~:text=The%20Glasgow%20Safe%20Haven%20is,citizens%20in%20the%20Glasgow%20region.), [HIC -Health Informatics Centre](https://www.dundee.ac.uk/hic), [DataLoch](https://dataloch.org/) and [Grampian Data Safe Haven](https://www.abdn.ac.uk/iahs/facilities/grampian-data-safe-haven.php).
3. **Case level data**: data relating directly to an individual, place or business. This may be one record per person, place or business; or multiple records that are then linked to a person, place or business e.g. a record of medical prescriptions may be multiple entries for a person. Once the data is in the TRE it is de-identified and, in some cases, fully anonymised.
4. **Research**. This is investigative work in which the results apply to groups of people, rather than individuals being targeted. Thus for the purposes of this paper the definition of “research” excludes work to operationalise that research. For example where organisations (including firms) test data products on operational systems e.g. predictive algorithms that do lead to people receiving different services based on data.
5. Firms sometimes get access to operational systems containing public sector data when they are supporting public sector delivery or developing/maintaining the technology to run those operational IT systems. This is not a research use.

# Current operational practices

## Governance and data access

1. All Safe Havens operate processes that are aligned with the international best practice [“five safes” framework](http://fivesafes.org/) for helping make decisions about making effective use of case level sensitive data. The five safes are:-

Safe projects: Is this use of the data appropriate, lawful, ethical and in the public interest?

Safe people: Can the user be trusted to use it in an appropriate manner?

Safe data: Does the data itself contain sufficient information to allow confidentiality to be breached?

Safe settings: Does the Safe Haven limit unauthorised use or mistakes?

Safe outputs: Is the confidentiality maintained for research outputs coming out of the Safe Haven?

1. All projects (including those involving firms) need a clear statement of the benefit to the public.
2. In most projects, approval of applications is based upon firms being part of a research partnership that is led by an academic or someone from a public sector organisation. In the West of Scotland TRE, there are a small number of projects that are directed by companies without partners. These are for addressing short term specific research questions.
3. Applications for research data access are initially considered by staff at the TRE. Any application where staff at a private sector firm are part of wider research team are triaged into “simple” and “complex” cases. While the processes in each TRE are slightly different, the principles remain the same. In complex cases a full scrutiny of the application must take place and include
	1. Agreement of a Data Protection Impact Assessment
	2. Approval by an independent “Public Benefit and Privacy Panel” or NHS Caldicott Guardian
	3. A consideration of the ethics of the use of data proposed
	4. An assessment of whether a view from a public panel needs to be sought
4. For projects in the National Safe Haven, there needs to also be an external independent scientific peer review.
5. The boundary between simple and complex cases differs between the five TREs. In most cases, complex cases are where the applicant is asking for new data about individuals to be brought together, there are many data items or particularly sensitive data, the project is using data across multiple data controllers, or where there are any doubts about strong public benefit of the research. Simple data access therefore is where there is clear precedent for the research i.e. that there has already been an independent approval of privacy and ethical risks, no particularly sensitive data is being used, and there is a clear public benefit case for the research. In this situation, approval for data access is made by Safe Haven staff. In one safe haven there is no distinction between complex and simple cases with every research project going through the complex governance route.
6. Virtually all data analysis of case level data happens within one of the five TREs and the firms do not become controllers of the data i.e. in this situation the data themselves are not released to the firms as they are only able to access but not remove data. All the data accessed by researchers is deidentified and is limited to the data approved for use in that specific research project.
7. Direct access to the data within a safe haven is allowed for staff working in private sector firms, though data access is only possible by staff that are accredited as “Safe Researchers” i.e. that each researcher has completed governance and information security training and passed an exam on safe data handling. This is managed in some cases by industry researchers being given an honorary contract with a UK based public sector organisation for the duration of the project. DataLoch have developed a training course bespoke to researchers working in firms that needs to be completed in addition to standard training before they can access data directly.
8. Each researcher must sign a user agreement that stipulates the terms on which data access is given and sanctions for breaking the terms of use. With any serious breach of the terms of the user agreement, the researcher is reported to the Information Commissioners Office. Depending on the severity and implications of that breach, action including a fine or imprisonment can be taken against the perpetrator.
9. Those researchers need to work in “Safe organisations” i.e. the organisation signs up contractually for each project to safe handling of data, and of being transparent about the purpose of the work and any outputs. An independent assessment of the organisation is conducted by most safe havens to check for a public benefit motive.
10. Any outputs from the TRE i.e. the analytical findings must be checked for disclosure risk (i.e. whether it might be possible to identify individuals) by a member of staff of the Safe Haven. This is complex for research projects involving AI algorithms. A recent project [[GRAIMatter](https://gbr01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdareuk.org.uk%2Fgraimatter-guidelines-for-safe-deployment-artificial-intelligence-models-in-trusted-research-environments%2F&data=05%7C01%7Croger.halliday%40researchdata.scot%7C4d0b24a21b934d0511a308db220cf415%7C039076c6a0fc453a99c4a9626e422172%7C0%7C0%7C638141211717069283%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=iAKNKZzL91TV1OXM9e8G8%2B6MJl3G716UW4XUbXIdIG4%3D&reserved=0), led by HIC at the University of Dundee] provided a review and recommendations for developing practice on this.
11. A small number of research projects require data from a range of data controllers, for example studies of rare health conditions where data from multiple NHS Boards come together. In some cases, data is moved into a secure computing facility run by a firm. This is based upon having data sharing agreements, a Data Protection Impact Assessment and a full System Security Policy Assessment. In these cases staff in TREs fully anonymise the data and securely transfer it.
12. There are some projects that are directly commissioned by public sector organisations and carried out by private sector staff. In these situations, public bodies will procure services of firms to fill gaps in skills or short-term capacity. These projects with either take place in TREs using all the standard safeguards as for projects delivered by public sector or academic researchers, or they will take place on the premises/systems of the public sector organisation.
13. Navigating data access processes is challenging for many firms. This is perhaps because of the lack of a clear engagement mechanism on defining the public benefit from industry involvement, and lack of consistency across Scotland.

## Finance

1. There are direct and indirect costs for data access borne by staff at Safe Havens. This includes staff time for curating and preparing data, scrutinising project applications, and supporting ethics assessments, and public engagement. It also includes non-staff costs such as computing, data curation and accommodation.
2. Access to public sector data for projects involving a private sector organisation is currently based upon paying the associated direct costs in full together with a contribution towards indirect costs. Each Safe Haven publishes a pricing structure for this, though charges are often made on a case-by-case basis, informed by NHS Research and Innovation teams. The national Health and Care Public Benefit and Privacy Panel (PBPP) expects benefits to the NHS / wider Scottish economy to be explicitly detailed and for the intellectual property of the outcomes to be clear within the application.

## Transparency

1. Currently staff in safe havens collect data on who accesses data, what datasets/data items are accessed, the purpose for this, and any published outputs. This is held by the Safe Havens. Publication of this information is done consistently by the Lothian, HIC and Grampian Safe Havens. In addition, elements of this are made public, for example through the IRAS annual report, research impact reports, or research funders. The PBPP publishes a list of studies that have been approved and in due course their end of project reports.

## International data access

1. Arrangements for access to the data by researchers from outwith the UK (whether from the private sector or not) have an additional level of security. These are considered on a case-by-case basis and data controllers have different risk appetites on whether they allow international access. This often means it takes time and effort on the part of the researcher to set this up. Access is typically quite restrictive e.g. access is only given for a specified and short period of time. The Grampian Safe Haven have developed an assessment process for this.

# Case study: Industrial Centre for Artificial Intelligence Research in Digital Diagnostics (iCAIRD)

## Summary

iCAIRD is one the UK’s largest healthcare artificial intelligence research programmes, with over 50 ground-breaking research projects spanning technology platforms, pathology and radiology. It is a 4-year programme funded by Innovate UK and industrial partners until March 2023.

Using the quadruple helix approach, iCAIRD is delivered from Scotland by a collaboration of 40 partners including health boards, universities, multinational industrial partners, SMEs, technology providers, healthcare professionals and the public.

## Public data

iCAIRD uses a large volume of diagnostic images including x-rays, CTs, mammograms, MRIs and digital images of histopathology tissue slides from patients in Grampian, Lothian and Greater Glasgow & Clyde health boards.

The programme has digitised over 1.75m pathology tissue slides whilst modernising the pathology laboratory at NHS Greater Glasgow & Clyde. For the first time, histopathology images are available at scale to Scotland’s research community.

In support of the imaging data, iCAIRD has used around 20 linked datasets that include everything from GP referral letters to radiology reports and laboratory test results.

In total over 1bn individual data points have been processed, relating to almost 2m patients.

## How public data is used

iCAIRD projects typically use a category of artificial intelligence called ‘machine learning’, where computer algorithms are taught by clinicians to recognise patterns in historical patient data.

The resulting computer software is then able to recognise these same patterns in the data used by clinicians to investigate, diagnose and treat patients.

Artificial intelligence needs to use a large volume of varied and representative data from across Scotland, to be accurate and usable within patient care.

A small selection of the iCAIRD outcomes that wouldn’t otherwise be possible without access to patient data, include:

* A tool that recognises the features of COVID-19 on a chest x-ray, within minutes of the scan being taken
* A clinical ‘cockpit’ that searches a patient’s full medical record and shows clinicians the most relevant data for treating a suspected stroke
* Software for use by the Scottish Breast Screening Service, as a ‘second pair of eyes’ to detect signs of potential breast cancer
* Tools that quickly identify wrist and ankle fractures on x-rays
* Pathology software that can detect, classify, measure, report and visualise different cancers (breast, prostate, kidney, liver, gastric and gynaecological).

## Data protection

iCAIRD takes the confidentiality and security of patient data very seriously. It uses over 40 different governance checkpoints and a combination of contractual, technical and behavioural controls to protect patient data, including:

* Data is deidentified before researchers are given access.
* For most types of data, a researcher can only access it using a secure data environment provided by one of Scotland’s Data Safe Havens. Researchers cannot download this data.
* Before any artificial intelligence models or research outputs are released, they must first pass through a disclosure control process that ensures no patient-level data is present.
* Individual researchers must pass a basic disclosure vetting process, follow strict procedures for data access and be suitably trained in using sensitive medical data.
* Every research organisation is governed by a collaboration agreement, that includes a clear statement of how they can use the data, and penalties for mismanagement.
* Before researchers are granted access to patient data, they must have the approval of an ethics committee, which examines the purpose of their research and ensures that the data to which they are given access is proportionate and appropriate.
* Each research project is independently governed by a suite of quality-assured processes.

## Project benefits

Artificial intelligence is particularly sensitive to local data and the environment in which it is deployed. By using Scottish patient data, the AI software produced by iCAIRD researchers will be more accurate and usable across Scotland’s health boards.

Having access to realistic data reduces the cost of developing AI software and increases the likelihood that it will operate correctly in the clinical setting, where it’s deployed. This makes iCAIRD an attractive choice for industrial partners,

Some of the problems that iCAIRD tries to solve, such as summarising key information for clinicians from an extensive patient history, are highly complex. It’s impossible to solve these difficult problems without access to a wide variety of linked patient data at scale.

## Challenges

Obtaining the data and governance approvals necessary for artificial intelligence research in healthcare, is extremely challenging. It is particularly difficult when operating a large research programme with interdependent teams and projects. Some examples of where improvements are necessary include:

* Process duplication - some or all the processes (data request, research ethics, information governance, security compliance, order placement and invoicing, etc.) must be repeated for each health board, depending on the governance route. This takes an enormous amount of time, effort and cost.
* Research governance (which is inextricably linked to the data being used by researchers) isn’t consistent between boards and between safe havens. There are differing attitudes to industrial innovation, different processes and different documentation.
* Within the national ethics process, there were concerns about providing data to commercial organisations. In at least one site, local governance has been streamlined for academic research and commercial research must follow a different process. There appears to be an unhelpful bias against commercial research in some areas.
* There is no common architecture or approach to systems integration and data extraction across health boards. Data is stored in a myriad of different IT systems, and this differs between boards. Data extraction is often a resource-intensive, manual process with limited automation.

These issues represent a significant barrier to AI research and innovation. Scotland must adopt a fast, simple and consistent national approach to healthcare data.

For more information, please contact info@icaird.com. *iCAIRD (project number 104690) is funded by Innovate UK on behalf of UK Research and Innovation [UKRI].*

# Future work: Areas to clarify and explore

## Aim of work

1. Research Data Scotland has heard from many stakeholders who have suggested potential ways in which the way in which staff in firms access public sector data can be improved, and there are areas where current arrangements could be developed.

## Context

1. While there are no arrangements where case level data is “sold”, there is a long history of selling access to non-personal public data for commercial use. In Scotland this happens (for example) through Registers of Scotland and at a UK level through Ordnance Survey and the Met Office. Significant public benefit comes through services the data insights these arrangements allow.
2. Industry access to data needs particular consideration as public views on this are different from public sector, academia or third sector uses. While there is widespread conditional acceptance of private sector use of public sector data especially among the informed public, many people’s gut reaction is of mistrust of the motives of industry.
3. If there were a loss of trust in use of public sector data, this would set back the ability to use data to improve wellbeing through research.
4. We also know that many people want to see the development of equitable benefit sharing models for collaborations or partnerships between private and public sector organisations, as they expect a share of benefits (including profits) accruing directly from access to public sector data to be returned to publics and reinvested into the public sector. There is also sometimes a blurred line between research and development in commercial and academic environments with a number of spin out companies coming from university-led innovation.
5. The public expect to see stringent oversight, governance, and safeguard arrangements around private sector use of public sector data, especially concerning an oversight or governance body, transparency and accountability processes, and arrangements for data security and safety, consent, and confidentiality. However, the precise nature of what the safeguards should be is contested, and it may be that the nature of the safeguards is less important than the fact that effective safeguards exist.
6. The public want there to be public involvement or engagement processes and activities around private sector use of public sector data, but the precise nature of what this should look like, who should be involved and in what ways is contested – while some want to be actively involved in decision making, others prefer more passive forms of communication and information distribution, and proportionality matters.
7. A strong message from industry is that alignment across Scotland and the UK is beneficial to give clarity and certainty to researchers.
8. There were a number of areas for further exploration and these have been brigaded under the following themes:-

## Purpose

1. Use of public sector data by firms for research only happens for work in the public good and that will always continue. There is, however, some contention about what work in the public good means, and what to do when the public benefit at the project outset is dependent upon the findings of the research. Therefore, clarifying where the boundary lies between what we're comfortable doing versus what we're less comfortable doing will be helpful.
2. There are occasions where a project starts with a clear public benefit by a trusted private sector organisation, but that the ownership of that organisation changes during the project potentially affecting the motivation and way of working on the project. This is a consideration of any clarification of arrangements.

## Transparency

1. There are plans to make information on who accesses data, what datasets/data items are accessed, the purpose for this, and any published outputs publicly available via a “data use register”. To make this happen will require coordination with the Scottish Safe Havens to agree standards and ensure roll out of these registers.
2. In addition, ensuring the is a clear and straightforward narrative about the use of people’s data for research is vital for maintaining people’s trust. This should be a collective challenge for those operating Safe Haven environments.

## Governance

1. While the overarching principles are clear and consistent across Scotland at a high level, practice differs somewhat in different Safe Havens. Aligning approach for industry to best practice would simplify and speed up cycles of innovation whilst and continuing to provide robust scrutiny of applications for data. Indeed, it is important to recognise the role that the secure network of Scottish Safe Havens deliver through a best practice governance pathway.
2. There is the potential to work with the Safe Havens around Scotland, NHS Caldicotts, other public sector data controllers, and the Public Benefit and Privacy Panels to identify steps to alignment around best practice approaches. This would require consultation with public panels to advise data controllers and develop a policy they support.

## International

1. Scotland has a policy choice about ease of access by international researchers or firms. Being part of international research consortia can bring significant benefit and new research to Scotland, though allowing easier access to Scottish data by international researchers may mean reduced investment in new research jobs in Scotland. There are also significant Information Governance challenges here. So, clarifying a policy and operational processes in support of that position would be very helpful.

## Finance

1. It is unclear currently whether the financial and wider benefit gained by private sector firms is close to the cost they pay for accessing data. A significant amount of work has explored more sophisticated approaches to benefit sharing and a [literature review](https://blogs.gov.scot/digital/2022/12/07/exploring-public-sector-personal-data-benefit-sharing/) was done as part of the unlocking the value of data work.
2. To move this forwards requires an assessment of the costs/benefits and a better understanding public views of different approaches and existing frameworks. It would then need to establish ways of actually rolling out preferred approaches, recognising that frameworks are likely to need to vary for different types of business and sector.
3. This also needs to explore the related issue of intellectual property of products derived from the use of publicly available data.

## New technology

1. Some companies are keen to work with specific technology. It is important to clarify what the offer to industry is technology wise from the public sector.
2. A significant number of private sector organisations bring new insights by using public sector data to do discovery research (i.e. what are the topics on which public benefit may be possible) and to develop Artificial Intelligence models.
3. Using AI approaches brings new challenges (for example disclosure risk of data embedded in the models themselves). There are a number of evolving issues around deploying the models in public sector setting, like ethics, intellectual property and regulation. There are a number of organisations who are collaborating on arrangements for resolving those issues.

## Data availability and quality

1. A number of firms mentioned that key to them working in Scotland was having data available for use on a consistent basis across the country and of a quality and regularity that would enable useful research to be done. This includes improving the published metadata for the datasets available for research, and the ability to allow firms to combine and analyse data they have collected with public sector datasets.

## Stages of research and development

1. Data is needed at both the research phase, and at a testing/deployment phase. The Safe Havens offer a route for supporting the research phase, though support is needed to secure access to data for the testing/deployment phase.

## Opt-out and consent

1. While the vast majority of the public are happy for their data to be used for research in the public good, some people’s gut reaction is to express uncertainty, particular when private sector organisations are involved. For some enabling opt out is desirable, however this risks introducing bias into the research and reaching false and harmful conclusions. Resolution of this issue aligns with the Research Data Scotland principle of “Aligned with the Scottish data strategy, we will support people’s appropriate choice over the use of their data in research”.
2. A number of organisations mentioned the need to be able to bring together consented data, in particular as part of a clinical trial, with unconsented data from medical records.

# Conclusion and recommendations

1. There are some issues on which a relatively swift policy decision could be made on the use of data in research by the private sector. These are
	1. the need for firms to work in partnership with academia/public sector,
	2. clarifying the boundary of research classed as in the public benefit,
	3. whether all research should be undertaken within a Safe Haven,
	4. arrangements for transparency about use of data and the outcomes stemming from that use
2. For these issues, a possible approach would be to undertake public engagement to test the optimal policy position on these issues, as well as engagement with data controllers and the Scottish Safe Havens. It would then be important to formalise these in an agreed written policy statement that is rooted in a firm legal position, public acceptability/social licence, and one that would maximise the public benefit from industry using public data.
3. There are other issues around data access that whose resolution is fundamental for industry to innovate with data. However, these are also improvements needed irrespective of whether this is research involving a private sector organisation. These are
	1. The alignment and simplification of Information Governance across Scotland
	2. The arrangements for access to Scottish data by researchers outwith Scotland and the UK
	3. The collection and curation of high-quality high value datasets
	4. Arrangements for public to choose that their data would not be used in all or specific research projects
4. Work is already underway to address these challenges, such as through the National Information Governance Programme. Some of these issues also align with the mission of Research Data Scotland who will work with others to reach a consensus amongst stakeholders and garner public support. They will continue to report back to Scottish Government on those matters.
5. There are also issues directly related to industry access to data where more in depth consideration is needed. Two of these came out strongly in discussions, namely:-
	1. How firms pay for access to Scottish data for research and share benefits coming from that access,
	2. How new technologies used by many firms would be handled.
6. Research Data Scotland is keen to support Scottish Government with these challenges. This will require engaging widely with industry, data controllers, the public, the Safe Haven network, and others across Scotland and the UK given the need placed by industry on alignment of process.

# [Ends]

#### About Research Data Scotland

Research Data Scotland is unlocking the power of public sector data.

We help researchers find and make use of health, social care and administrative data to improve the lives of people in Scotland. Working in partnership, we do this by safely and securely widening the range of data available, creating new data assets and providing a single point of contact for effective access to public sector data.

We are a not-for-profit charitable organisation created and funded by Scottish Government. We are a partnership between Scottish Government, leading universities and public bodies, such as Public Health Scotland (PHS) and National Records Scotland (NRS).

#### www.researchdata.scot

#### @RDS\_Scotland