ESRC Future Data Services Contextual positioning paper

Felix Ritchie, University of the West of England Bristol and Senior Strategic Fellow, FDS

Contents

Introduction

This paper presents a contextual framework to facilitate and promote access to data for research purposes. It provides relevant background information on theories and concepts which will help shape the thinking of ESRC's Future Data Services programme. The framework comprises two sections:

- **Core concepts** discusses basic concepts from the management, economics and data governance literature which form the background to the work of the FDS team
- **Models of decision-making** explore how decisions are made, what works, what creates efficiencies, how evaluation can (and can't) help, and sources of evidence

This paper draws heavily on the external literature, including the work of the authors, to provide a primer on the relevant topics. Much of the economic theory can be found in a standard textbook; for non-economic theory, particularly relating to data governance, the references will provide some additional insight or explanation. However, to keep the paper readable, we have tried to minimise references except where a particular point is being made that would benefit from an example. Some sections are based on a single paper or report; these are identified in the footnotes.

The views expressed in this paper are those of the author and do not necessarily reflect the views of ESRC or the Future Data Services team.

List of abbreviations

DEA	Digital Economy Act 2017
ESRC	Economic and Social Research Council
ESEe	Economic, social and environmental evaluation
HDR UK	Health Data Research UK
MRC	Medical Research Council
NSI	National statistical institute
ONS	Office for National Statistics
RDC	Research data centre
RJS	Remote job server
TRE	Trusted research environment
UKDA	UK Data Archive
UKRI	UK Research and Innovation
UKSA	UK Statistics Authority

Glossary

Term	Definition to be used in this document
Anonymous data	data which does not include sufficient detail to allow the data subject to be identified, under any reasonable conditions
Breach of confidentiality	the release of identified or de-identified data to an unauthorised system, environment or person; a breach of confidentiality may not mean a disclosure as it will depend on the circumstances
Breach of procedure	failure to follow appropriate operating procedures, irrespective of whether a breach of confidentiality occurs
Data collector	The body that acquires the data about a subject
Data depositor	The body that deposits collected data with third parties for re-use
Data holder	The body that has physical/technical possession of data
Data linker	The body responsible for linking data sources together
Data or Research user	The end user of the data for research purposes
Data subject	The person that the data refers to
De-identified data	data which includes sufficient detail to allow the data subject to be identified, but only with effort and with less certainty (for example, a combination of gender, age, type of employer, salary range and disability status)
Distributed access	restricting the physical location of the data, but allowing users in other locations to carry out analysis and extract statistical results (not microdata) through RJSs or vRDCs
Distributed data	sending microdata to users under licence, to analyse on their own machines
FAIR principles	Findability, Accessibility, Interoperability, and Reusability: principles for improving the use and accessibility of data
Five safes	Five Safes is a framework used to help decisions surrounding data access it considers five related but separate dimensions: safe projects, safe people, safe data, safe settings, safe outputs.
Identified data	Data directly related to an individual respondent for example name, employee number, address
Microdata	individual unit records about a person or organisation, such as information collected from surveys or administrative data

Output checker	The person checking research outputs for risk before release
Output SDC	the application of SDC methods to analysis
Principles-based	A regulatory regime or operating model where 'principles' (what you are trying to achieve) are the basis for planning rather than specific instructions, and compliance means checking to see if a solution is consistent with the principles
Public use file (PUF)	data file without restrictions on use or onward access
Raw data	the source data collected
Record-level data	synonym for microdata
Remote access	a system which allows users to 'see' and manipulate the source data from a physical location independent of the actual location of the data holder's facilities
Remote job server (RJS)	a system allowing a range of complex analyses to be carried out, not just tabulations, without seeing the source data
Research data centre (RDC)	a restricted access facility where users can manipulate the source data without restriction the environment is made secure so that users cannot bring information into or take data out of the facility without approval
Research data manager	The manager of data made available to researchers
Research-ready data	Data that has been processed and curated to be suitable for research use
Rules-based	A regulatory regime or operating model where explicit rules are the basis for planning
Scientific use file (SUF)	data file which retains some non-negligible confidentiality risk and so therefore has circulation restricted to authorised users for specific purposes
Secure Data Environment (SDE)	A secure facility that allows data to be used without leakage from the system; access, ingress and egress are strictly controlled. SDE and TRE are sometimes used interchangeably, but an SDE is concerned with secure *storage and management*, a TRE is about research/analytical *use*.
Secure use file (SecUF)	data file which contains non-negligible confidential information therefore circulation and use is restricted to authorised users in controlled facilities
Sensitive data	data where release to an unauthorised person is likely to cause nonnegligible harm or distress to the data subject
Statistical disclosure control (SDC)	applying statistical measures to make changes to the data or publication to reduce risk of reidentification
Synthetic data	generated data that can replace or augment sensitive source data
Table server	a system which allows users to generate their own tables from the data flexibly, but without seeing the source data; a specific form of RJS
Trusted research environment (TRE)	A secure facility that allows researchers to analyse data without leakage from the system; access, ingress and egress are strictly controlled. RJSs and RDCs are types of TRE
Virtual RDC or Remote RDC (vRDC)	an RDC where technology is used to provide equivalent security to a physical site and to separate the RDC from the actual location of the data

Part I Core concepts

This section introduces some basic concepts that provide the foundational framework of data access. The aim of this section is a neutral description of current perceptions of theoretical and empirical evidence.

1. The Five Safes[1]

1.1 Concept

The Five Safes was devised (initially with only four 'safes') at the beginning of 2003 for the Office for National Statistics (ONS). This is a framework that divides confidential data management into five dimensions: safe projects (the ethics, purpose and plan for access), safe people (who are the users, and what training do they have/need?), safe settings (how is the data stored and transferred, and with what safeguards?), safe data (is the detail in the data appropriate?) and safe output (could published results inadvertently breach the confidentiality of data subjects?). These are often characterised by five questions:

Element	Typical question	Example of problems being addressed
Safe projects	Is this appropriate use and management of the data?	• What is the purpose of the access request? • Is this an ethical and lawful use of the data? • What is the benefit to society or to the organisations sharing data? • Is there a data management plan in place? • What happens to the data at the end of the project?
Safe people	How much can I trust the data users to use it appropriately?	• Do the users have the necessary technical skills? • Do the users need training in handling confidential data? • Are users likely to follow procedures?
Safe settings	How much protection does the physical environment afford to the data?	• How is data stored? • Are there physical restrictions on the users? • Does the IT prevent unauthorised use? • Are mistakes by authorised users likely to be detected?
Safe outputs	How much risk is there in the outputs of the access breaching confidentiality?	• If the aim of access is to produce statistics, is there any residual risk by, for example, showing outliers? • If the aim of the access is to produce data for onward transmission, how do we make sure that the released data is appropriate for the next use?
Safe data	Is the level of detail in the data appropriate?	• Is there sufficient detail to allow the project to go ahead? • Is this excessive detail which is not necessary for the project?

Table 1 The Five Safes, represented as useful questions – Green and Ritchie (2023), based on Ritchie (2017)

The Five Safes considers each 'safe' as a separate but joint element of any solution (Ritchie, 2017). For each 'safe', there is a wealth of knowledge about how to achieve the goals (how to run ethics committees, how to anonymise data, how to design secure IT systems, and so on). Therefore, each

[1] This section is largely taken from Ritchie (2017) and Green and Ritchie (2023a)

safe can be considered separately, by assuming that the necessary controls in the other dimensions can be implemented as needed.

For example, a project lead considering the broad shape of the project (ethics, data collection methods, user groups, post-project use, and so on) can work on the assumption that user training, IT, data detail and output checks can be implemented to any relevant standard, once the shape of the project is clear. Alternatively, the project lead might start from considering the types of users, what they can be reliably trained in, and feeding this into the specification for system design knowing that it can be achieved.

The separateness makes for *efficiency* in decision-making: not everything has to be considered at once. The *security* comes from then considering all the 'safe' dimensions together: for example, do our assumptions about the staff training hold, given the IT system that we have designed? The joint-but-several approach illustrates that 'safe use' might have multiple solutions. For example, risks in outputs may be handled by training users, or by having an automatic output-checking system.

These five dimensions of control are scales, not targets; different solutions will have different levels of control in different dimensions. Treating controls as scales demonstrate one of the most useful aspects of the framework – the ability to dial up or dial down controls, based on context and how the safes are considered jointly. Each control should be effective and contribute to reducing risk, therefore limiting unnecessary risk control efforts. There may be non-existent controls in some dimensions, so long as the collective solution is appropriate.

Table 2 illustrates the application of controls to different versions of the same dataset: one released openly on the internet, one downloadable as an end-user licence, and one available through a TRE:

Situation	Controls				
	Project	People	Settings	Outputs	Data
Open data	None	None	None	None	Very high (full anonymization)
Licensed download	Some (online application)	Some (licence)	Some (online guidance)	None	High (eg little geographical detail)
Controlled use	High (application with human review)	High (compulsory training)	High (isolated environment)	High (all outputs manually reviewed)	Minimal (de-identification only)

Table 2 Different control levels potentially applied to the same source data (source: Green and Ritchie, 2023a)

For the open data, all the protection is contained in the data itself, as there are no feasible controls on use. In contrast, the secure facility needs only make minimal changes to the data because other controls provide an assurance of safe use. For release of the licensed download, the data holder has some confidence in users, but is aware that there is much scope for error beyond her control, and so reduces the data detail.

There is no inherent precedence in the Five Safes. Ritchie (2017) and others argue for the primacy of safe projects, with 'safe data' treated as the residual. However, there may be cases where, for example, the detail in the data is fixed in advance, and the other dimensions must adjust to it; for example, where data collection plans have been treated as separate from the research planning.

This also happens where data holders interpret the 'need to know' principle as "ensure that data availability is minimised in system design". The Five Safes places this principle in its proper context. Data access procedures with data as the residual control in the design allow for any data consistent with the other 'safes' to be released through the system; the actual detail in any dataset release can

be specific to the project. In other words, 'need to know' is a project characteristic, not a system feature.

The framework itself is no guarantee of good practice. Some critics refer to it as a box-ticking exercise; one UK government organisation unironically declares itself to be 'Five Safes compliant'. However, the problem is not the framework, but the implementation. It is not uncommon for infrastructures to focus on one safe without reference to others e.g. to consider if the data are safe or unsafe, without reference to whether the data are accessed in a safe setting, or if safe people accessing the data.

The lack of detail in the Five Safes is a limitation (although Green and Ritchie, 2023a, note that this is also a strength when it comes to flexibility), and several authors have tried to create practical guidelines. The most advanced is probably that of the Australian research group CADRE (Co-ordinated Access to Data Researchers and Environments); the initial "CADRE Five Safes Framework" (McEachern, 2021) is now available as a web resource. CADRE notes that Five Safes does not stipulate a quantitative evaluation of risk factors, and so it develops a checklist to provide consistency between data managers. In the UK, the SATRE project (https://satre-specification.readthedocs.io/en/stable/) has similarly tried to turn broad ideas into actionable tasks.

There is a question of terminology. Describing these as five 'elements' tends to work well for designers and implementers, as it implies a self-contained component which can be tackled as part of a project plan. However, in discussions with data holders and the general public, who tend to be more concerned that risk is being managed, 'dimensions of control' seems to be a more meaningful description; this carries the sense of "what you can't control, and what you can". Other writers have described the elements as themes, security controls, components, or standards. Some organisations describe their operations as a 'portfolio' model.

1.2 Use

The Five Safes framework has three main uses: description, including pedagogy; design and evaluation; and regulation.

The Five Safes is currently used to describe the governance arrangements for almost all the general-purpose UK government and academic trusted research environments (TREs) for health and social science research. Eurostat, the Bundesbank, the Dutch research infrastructure ODISSEI, and the NSIs of Canada, Australia, New Zealand, Mexico, France and Norway formally describe their RDCs in this way. As NSIs often have a strong influence over the data strategies of other parts of the public sector, the adoption of the Five Safes by NSIs has had significant spillover effects, particularly in Anglo-Saxon countries. It has also been used in confidential data management training since 2004. The ready-made structure appeals to the trainees, and the framework provide a context for the training itself as part of the 'safe people' element.

As awareness of the Five Safes has come to precede planning, it has become more common to use it for design. It has been used for designing data strategies in the public and private sector, scientific-use files, survey structure, risk assessment tools, and public consultations, amongst other uses. The US National Research Council incorporate it into their recommendations for anonymization of US health data for sharing (NRC, 2014). At the other end of the design process, the predefined structure can simplify evaluation. This has not been a major use so far.

Finally, the Five Safes is used in formal legislation, such as the UK Digital Economy Act 2017, state legislation in New South Wales, Victoria, and South Australia, and the Australian Federal Data and Transparency Act 2022. There is also a growth in principles-based regulation (such as the GDPR),

which needs an operational framework to enforce it; although not framed in the language of the Five Safes, this may be something that develops.

1.3 More or different safes?

As noted, the Five Safes is a generic framework. Some authors have tried to give more meaning to the dimensions, particularly safe projects and safe outputs. Several others (including HDR UK) have also suggested additional 'safes'. Finally, some authors have proposed ways to combine the safes more effectively. McEachern (2021) introduces 'safe groups'- combinations of different safe categories which would be automatically formulated as 'safe' for classes of use. Statistics Canada is exploring a similar issue. Green and Ritchie (2023a) provide a more detailed review of these proposed extensions, but most of them suffer from the problem.

1.4 Subjectivity vs objectivity

The Five Safes is an explicitly subjective framework. It does not attempt to quantify 'safe' in any dimension, let alone try to balance 'safety' in one dimension against 'safety' in another dimension. There are five reasons for this.

- There is no meaningful metric in any dimension, let alone a common metric; "risk of re-identification?" is often the proposed measure but is difficult to define or measure
- Discussion of 'trade-offs', 'risk-utility maps' or similar concepts implicitly assumes the independence of protection measures; in practice, this is unlikely to be the case
- In addition, 'trade off' models also do not allow for discontinuities: for example, a secure RDC presents a number of step changes in risks associated with different operating models
- Any data that does come out of modelling is subjective; there are no objective models of risk with external validity (Hafner et al, 2015); risk is affected by scenarios, motivations, timing, external information, the output publication scheme and so on; all of these need subjective assessment.
- The design and operation of the system affects the psychological perception and hence interaction of the participants in the system; in fact, much research into data governance is based on the assumption that behaviour *will* be changed by well-designed systems

In this light, the lack of quantitative measures for the Five Safes is an essential strength of the model. This focuses decision-makers' minds on the need to collect, evaluate and use subjective evidence, and to build consensus for decisions. This does not mean that no metrics are useful. Green et al (2024) for example suggest a range of metrics for examining the performance of output-checking processes. However, these are operational measures, and not risk measures.

2. Data Access Spectrum

The use of the Five Safes presents a representational issue: how do you represent five dimensions in a useful single image? Presentation of the five categories is straightforward, and there are lots of variants. These categorical graphs are less useful when trying to discuss degrees of control within the safes. The 'graphic equalizer' conceived by the Steve McEachern (Australian Data Archive) and the 'spider' model of Arne Wolters (Health Foundation) both allow visualisers to present some degree of the scale of each control.

One successful representation has been to recognise that four of the five dimensions tend to work together: projects, people, settings, and output are created as a coherent set of 'procedural' controls. Data acts as the residual dimension, being fitted into the appropriate procedural and

technical setting and adjusted to the required level of anonymity. This is also the approach recommended in Ritchie (2017) when designing data access systems.

With just one axis and a level of data associated with this, this allows a simple representation which can also, usefully, be identified with extant access solutions:

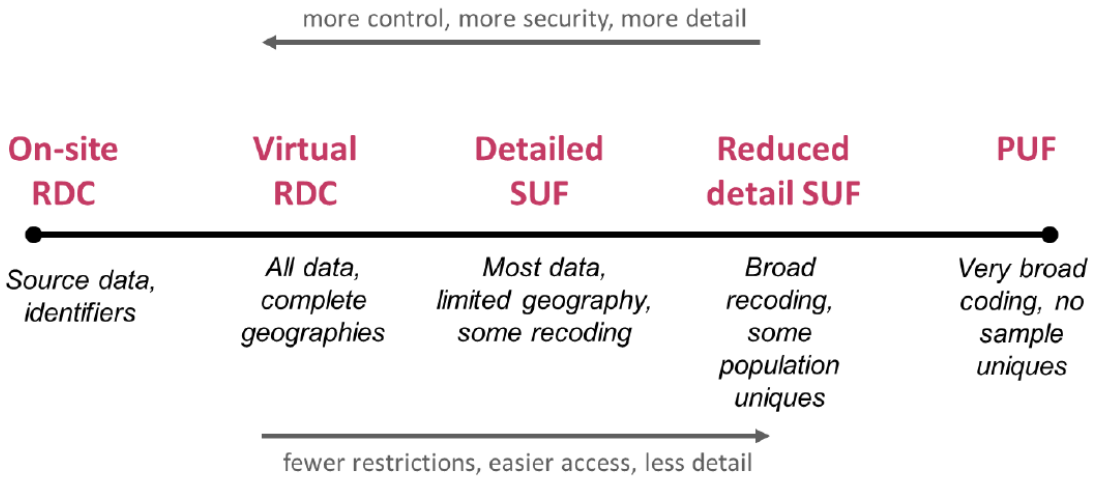


Figure 1 Data access spectrum (from Green and Ritchie, 2016)

For example, each version of the UK Labour Force Survey can be attached to a specific delivery mechanism. This representation was used in the ONS policy documents as a summary representation of ONS' data offerings, and informally by other UK organisations such as the UK Data Archive.

This model is called the 'data access spectrum' in the UK or 'continuum of access' in Canada where it was simultaneously and independently developed by Chuck Humphrey of the Canadian Research Data Centres Network. A similar version was adapted for government bodies in Greece and Australia. This model has also been promoted by other organisations such as the ODI, although whether as an innovation or as a development of the other models is not clear. It seems likely that such a simple but useful device would be re-invented multiple times.

The ODI representation is slightly different, and is a 'data spectrum' rather than 'access':

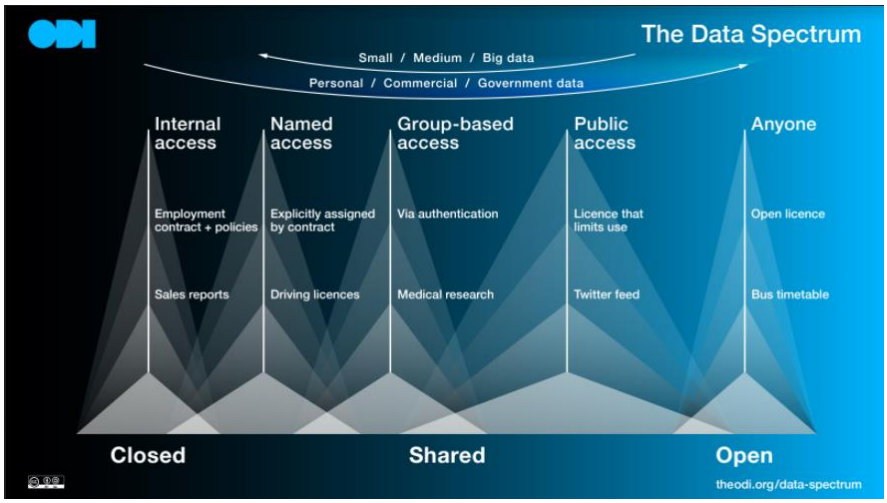


Figure 2 The ODI Data Spectrum (source: www.theodi.org)

This perhaps reflects the ODI's stronger focus in working with private sector organisations, where contractual agreements and commercial considerations play larger part than in public sector data models.

This representation is helpful in terms of the big distinction between "distributed data" and "distributed access". TREs offer the latter, but many ESRC data services also offer distributed data. For example, many thousands of researchers regularly download 'reduced detail SUF' by signing an End User Licence with the UK Data Service. This simply asks registered researchers to provide a short basic outline of why they wish to use the data. This is particularly useful for undergraduates or postgraduates who are using data for their research for the first time, but also for researchers in general who do not need to answer their research hypotheses with very detailed data.

The spectrum illustrates how access to data can be efficiently distributed according to the anonymised nature of the data, and the need by the researcher to use data of varying detail. Different services have different cost implications: distributed access implies significant initial IT investment, and ongoing service costs. In contrast, distributed data has high up-front costs (de-identification/anonymisation particularly) but few ongoing costs.

The DAS illustrates fairly well how data access works in the UK, and it can be used to classify different types of data service. It also provides a way of evaluating changes in data access policy. For example, the withdrawal of ONS 'Special Licence' data (detailed SUFs) from UKDA distribution has led to additional demand to access more sensitive data in TREs. Existing TREs have struggled to cope with this demand, but it is clear from the DSAS that an expected outcome of this would have been more demand for SecUFs.

These spectra are representations of data access where the researcher has full access to the data. For remote job servers such as OpenSafely or Microdata.no, these present slightly different issues as they increase the 'settings' controls and decrease the 'people' controls. Also, there are significant variations when data is being distributed eg licencing conditions, assumptions about compliance with IT specs, training for researchers. There might be value in exploring different ways to represent these.

3. Principles-based design and regulation

In recent years, the Five Safes model has become increasingly associated with the principles-based approach to regulation. There is an affinity between the two concepts. The Five Safes provides a framework for planning; the principles-based model provides a way of suggesting how goals should be specified. Neither is specific on the actual implementation; but both provide a way that the effectiveness of any implementation can be measured.

3.1 Definitions, pros and cons

Ritchie and Green (2022) provide the following definition of rules-based and principles-based regulation:

> Rules-based regulation aims to specify in a binary manner what is allowed or not allowed; the primary source of direction is the regulation itself. The strength of rules-based regulation lies in simplicity and a common understanding of the regulation. It works best when the terms can be unambiguously defined; for example, the UK Data Protection Act 2018 makes attempted re-identification of de-identified or anonymised data a criminal act, except in very specific, easily understood cases […]

> Principles-based regulation focuses on what any system is trying to achieve, and then questions whether the system actually achieves those objectives. In a principles-based system, implementation decisions are primarily under the control of the implementor; regulation is there to specify the goals, and to identify what evidence should be presented that the goals have been achieved.
>
> (Ritchie and Green, 2022)

Principles-based regulation is likely to be supported by accreditation. Rather than specifying actions, the actors, systems and procedures in a system undergo ex ante validation to ensure that they are fit for the purpose specified in the regulations.

The tax system provides an example. Under a rules-based tax system, allowable tax exemptions are specified either in law or in the statements of the regulator. Any exemption not specified in the regulations should not be allowable. The advantage of this approach is the apparent certainty. However, in practice the difficulty of specifying exemptions with no ambiguity tends to provide incentives for high-tax individuals to explore ways round those exemptions, which can then be the basis for a war of attrition between tax authorities and prospective tax avoiders. This can be a considerable waste of resources.

For this reason, some propose a principles-based tax system, whereby the regulation makes reference to concepts such as 'reasonable' allowances. This reduces the chance of gaming the system by high-tax payers, as the tax authority can always decide what counts as 'reasonable' and the context in which it is assessed. However, this provides less certainty over tax liabilities, and it cannot eliminate the prospects of a war of attrition between taxpayers and tax authorities over what is 'reasonable'. Overall, tax systems are more likely to be rules based.

A contrast is provided by the regulation of financial markets. The same arguments as above can be made about the merits of certainty versus flexibility and the opportunity for gaming. However, there is more variation: European regulation of financial services markets tends to be rules-based, whereas in Anglo-Saxon countries this is more likely to be principles-based (Keenan, 2020).

Ritchie and Green (2020) identify the advantages of regulation regimes as

Rules-based regulation	Principles-based plus accreditation
• Clarity • Transparency • Advance rather than post-hoc knowledge of liabilities or duties	• efficiency, as solutions can be adjusted to circumstance rather than a legislative context • flexibility, as multiple accreditation pathways can be set up • adaptability to circumstances, rather than requiring legislative/formal change • adaptability to collective learning as processes develop • cultural change through positive reinforcement and engagement with goals • engagement with stakeholders

Table 3 Rules- vs principles-based regulation (from Ritchie and Green, 2020)

An effective principles-based system therefore requires a much greater level of engagement, but it seems to address some of the flaws of older legislation which struggled to provide adequate guidance.

3.2 Regulation regimes and the Five Safes

In the context of the Five Safes, rules- and principles-based regulation may be contrasted using the following examples, which all exist somewhere:

Safe…	Rules-based standard	Principles-based standard
Projects	Identify list of valid uses	Specify benefits that must be demonstrated, and risks to be considered
People	Require specific accreditation eg meet Civil Service appointments criteria	Require 'appropriate' training
Setting	Follow government IT standard	Follow ISO27001 practices (ie choose system and be able to demonstrate integrity of it) and review
Output	Apply threshold rule tabular statistics	Apply threshold 'rules of thumb' but allow appeals if important and demonstrably non-disclosive
Data	Clear boundary between anonymised and other data	Data must be 'appropriate' to the environment

Table 4 Regulation regimes and the Five Safes

The Five Safes therefore can accommodate either approach, or a mixture of the two, but moves towards principles-based models are more likely to be done in the context of the Five Safes. This is explicit in legislation in the UK and Australia but even the European GDPR follows this line. The GDPR is not explicitly principles-based and does not cite the Five Safes; nevertheless, its balancing of data detail against 'procedural and technical measures' and the avoidance of specific technical or statistical standards in favour of solutions 'having regard to' outcomes places it in the same camp.

Data legislation increasingly is framed as principles-based. This includes the UK Digital Economy Act 2017, the European General Data Protection Regulation 2016 and the Australia Data Access and Transparency Act 2021, although only the latter explicitly references principles. However this does not necessarily follow through in practice. The implementation of the DEA research access provisions, for example, has been heavily rules-based and criticised for its inflexibility or reference to the public benefit goals.

However, in general specification of rules in data governance is difficult. Regulating digital activities has severely tested regulations which rely upon clear statements; for example, no legislation provides an unambiguous definition of 'anonymous' data because it relies on context except in very trivial cases.

3.3 Moving beyond regulation into strategic design

In the context of regulation, it is straightforward to contrast principles- and rules-based approaches. However, in the wider context the difference is not between principles and rules but between principles-based and operational planning. We can characterise as

Approach	First stage	Second stage
Principles-based	Where do we want to be/what outcomes/goals have we identified?	How do we achieve this?
Operational	What have we got/what can we do?	Where will this take us?

Table 5 Principles-based vs operational planning models

At the micro level there are advantages in the operational approach: it limits discussion to what is practical and achievable, and keeps the conversation focused. However, in general the principles-based approach is to be preferred: if you don't know where you want to be, how do you if you are using your resources effectively?

Starting from the goal and working backwards provides four key advantages (note these are also relevant to regulation but we are expanding the discussion here):

- There may be multiple ways to achieve the end goal, and so there is an immediate structure for deciding between competing methods
- Principles are less dependent on specific circumstances (eg legal frameworks) and so can (1) absorb change in technology law etc without needing to be updated, and (2) be applied consistently across very different environments eg training researchers in the UK, Eurostat, Australia, and Nepal
- Principles are not dependent on any specific technology or environment and so can be devised without reference to feasibility; conversely, they can be changed to reflect eg changing cultural preferences, without the need to root that change in a specific delivery mechanism
- Principles are more likely to be expressed in simple language, therefore adding communication with and buy-in from non-specialist groups

Principles can be nested in planning: very high level ("there should be a range of data delivery mechanism that meet different user needs"); next level ("TREs should be the default mechanism for delivering access to data which cannot securely be distributed"); next level ("To be considered a 'TRE', a TRE should show…."); and so on.

3.4 The role of accreditation

The way to deal with the downsides of principles-based regulation (lack of clarity or transparency, ex ante agreement on what is needed) is accreditation. If adherence to a principle is met by accreditation to a specific standard, then this can provide the necessary clarity, certainty and transparency. At a simple level, consider the requirements for being an ONS 'accredited researcher':

- Successful completion of the Safe Researcher Training (SRT) course and associated test
- Evidence of having the necessary research education and experience

The relevant *principle* is that users of TREs holding ONS data should have the relevant skills to be able to use that data *efficiently* and *securely*. The SRT is designed to ensure that the principle is met. Once a researcher has passed the test (which can be done separately from putting in a project application), they have an accreditation which can be used as evidence across all the main UK social and health science TREs, for a period up to five years.

3.5 Relevance to FDS

Principles-based thinking should be at the core of FDS, both for strategic thinking and implementation:

- On strategic thinking, the goal of FDS is to look forward; starting from very high-level principles without reference to implementation allows the raison d'etre for ESRC funding to be considered, without focusing on value for money in specific cases
- On implementation, detailed principles derived from high-level principles as the outcome allows for both innovation in delivery and more effective evaluation (Alves et al, 2021)

4. Monopolies and regulation

Some data services would seem to be 'natural monopolies'; that is, economies of scale suggest that the market for good or services is best served by one, or one dominant, supplier. The economic theory of monopolies is straightforward, if simplistic. Monopolies have good and bad points, and regulation is proposed as the usual way of overcoming the negatives.

The issue of monopoly is closely related to standards and efficiency. A monopoly may arise from standards, be the basis for enforcing standards, or be both or neither. A monopoly can be efficient (in the sense of having lower costs than competing interests) but may not be fully efficient in terms of society's benefit. Finally, standards have many of the same characteristics of monopolies: they can also lead to short-term efficiencies but can stifle innovation in the long run. Moreover, the efficiency of standards is related to the question of who defines the standard: is it for the benefit of the monopoly, or society?

This section and the next explores these issues.

4.1 Standard economic theory of monopolies

The dominant (neoclassical) economic theory of monopolies has been largely unchanged for a century. Most economic analysis focuses now focuses on (a) sources of monopolies and (b) regulation. To keep this analysis simple, we use 'prices' to mean the explicit or implicit market value of whatever the monopoly supplies, and 'profit' the net benefit to the managers or owners of the monopoly. We do not distinguish between managers and owners.

In economics, a monopoly is a supplier of goods and services that has sufficient market power to allow it to set its own prices without regard to what its competitors might be doing. This does not mean that the monopolist can change any price that it likes; it is still limited by the law of demand, that as price goes up demand tends to fall. But in general the monopolist can behave with a significant amount of freedom.

A monopoly can arise from several factors. It may be a statutory monopoly, such as the police force in a particular area. It may be the result of high fixed costs to enter the market, such as an internet search engine. It may be because of historical specialisation, such as a defence equipment maker. It may be the product is unique and capability to supply the goods or services is in limited supply, such as a famous actor. It could be that the monopolist has access to deeper capital markets than competitors and so can indulge in predatory pricing to keep out competitors, although of course no examples of this exist in real life.

The one most relevant for data services is the 'natural monopoly'. This is where economies of scale mean that larger business can operate at lower costs than smaller ones to provide the same service. The implication of this is that the largest firm will always be able to beat its smaller competitors on price. Natural monopolies particularly occur in businesses that have network effects: electricity grids, train services, water supply. This is why, UK excepted, these tend to be run by local or national monopolies. Significant fixed costs requiring a minimum efficient scale of production can also lead to natural monopolies.

For the rest of this chapter we assume that monopolies do not arise from unfair manipulation of markets, but from the natural characteristics of the business: economies of scale, uniqueness, limited supply of input. For our purposes, it will be sufficient to only consider economics of scale in this section, as the others can be seen as 'solvable' given sufficient time and money.

A monopoly which operates at scale to lower costs is clearly good for society: one large steelworks consumes less resources to make the same amount of steel as hundreds of mini foundries. However, this is not the optimum for society. The monopolist can combine resources efficiently to produce goods at a lower cost, but while some of that translates into lower prices, it is also likely to lead to higher profits for the monopolist. A monopoly may also lead to less of the good being produced than under a competitive system, because the monopolist can use its market power to restrict supply.

Neoclassical economists agree that a natural monopoly does use resources efficiently, and that the way to extract more of the benefit for society is to regulate it: fixing maximum prices, limiting profit, requiring a minimum level of service and so on. These all have limitations; in particular, they tend to encourage the wastage of resources on non-productive activities for internal reasons while still meeting regulatory goals.

The perspective on innovation is less clear. Early economists argued that monopolists stifle innovation because they have no incentive to innovate: they maximise profit in the current state of the world and innovations might offer opportunities for potential competitors. More recent economists argue that monopolists do innovate to keep their competitive advantage; an arms race between the monopoly and potential competitors, but designed to keep the monopolists position secure, not necessarily to increase the societal benefit of goods and services.

The general view of economists now is that this second position is closer to the truth, using as evidence household chemical or cosmetics, where a small number of dominant firms flood the market with different brands to close off space for new firms. On this view, monopolies do change and innovate but for their own benefit.

4.2 Flaws in the theory

The usual complaints about neoclassical economics can be made in respect of the above discussion: no business or regulator has the cost information to carry out this analysis; no businesses use marginal or even average cost pricing; few companies make just one product with no variation; businesses vary enormously in their resources, capabilities and operating environment.

A second criticism more pertinent to FDS is the assumption that senior executives are wholly focused on short-term profit (or a similar univariate internal goal such as executive pay, or seniority). This is debatable even with private sector firms – individuals have complex motivations – but it becomes very problematic when considering public sector operations. Here, the goals are more diffuse, and individuals working in the public sector are likely to be more focused on the public good.

However, the broad principles stand up well to this criticism: monopolies can be good for the public but without government intervention are likely to benefit themselves more than the public; monopolies are innovative to maintain their position, but the innovation may be focused on protecting their position rather than improving public outcomes; poorly-designed regulation may lead to wasting resources on unproductive compliance activities.

4.3 Relevance to FDS

There are significant opportunities for monopoly in data services. Investment in data infrastructure or metadata systems can lead to natural monopolies due to economies of scale. Product-based monopolies can arise because multiple parties negotiating for access to the same data is likely to lead to no access: data suppliers usually look to build a trust relationship with a single delivery partner. UKRI funding processes can encourage predatory bidding, and grant funding can create a monopoly through exclusion during the period of the grant. Finally, UKRI data service investments require some input from specialist resource (people) in short supply, and where new capability

cannot be easily or quickly developed; this has been seen particularly in the DARE investment programme, which have been concentrated in a small pool of specialist organisations. Beyond UKRI, a clear source of monopoly is legislation and regulation, with ONS having a unique statutory position to acquire and share data for research.

5. Standards, innovation and efficiency[2]

5.1 Types of standard

Standards can encourage innovation by providing a framework for new ideas, or restrict it by stifling new developments. Standards can be sources of efficiency by encouraging economics of scale, or can be a drag on productivity by requiring adherence to inappropriate or outmoded methods. Whether standards bring benefits is therefore an empirical question.

There are two types of standard:

- **process** standard: a common terminology and security model – the way of identifying goals, principles, or a framework
- **outcome** (or **design**) standard: a set of specific 'levels of achievement' for those goals

These can be developed as either open or closed processes:

- **open**: developed by co-operation in a network, with input into the standard definition open to a wide group
- **closed**: defined and/or imposed by a dominant party

This can be further developed as the difference between 'systems' and 'networks'. A system implies a monolithic solution, perhaps designed and/or implemented by a single authority, and embodying outcome standard; this in turn suggests a closed standard with ownership and a central architecture. In contrast, a network focuses on gateways and communication protocols; what goes on behind those gateways is of no concern to the network.

The focus on the network rather than system characteristics decentralises the decision making process and is designed to encourage innovation in solutions. For example, the world-wide web was developed without the need to change the basic operation of the internet which predated it by two decades. Not specifying particular solutions can encourage alternatives to be explored.

The downside of setting standards is that they themselves can become blocks to progress, however well-intentioned the original plan. This is particularly the case where network standards cannot be changed without massive expenditure or disruption. Adoption of standards is enhanced by being able to adapt to the needs of those who were not part of the original consultation.

Technological lock-in similarly is a risk to be managed, and decentralisation can help by providing the incentives for competition. The internet is a positive example: the 1970s protocols, designed for a much simpler world, have been steadily augmented by new ones, driven by the enormous success of the World-wide Web, in ways which the 1970s designers never envisaged.

There is a time-dependency in the development of successful standards: early movers may generate 'co-evolutionary' responses in other participants. For example, the Five Safes was developed for ONS but has since spread widely. The apparent conformity of terminology has encouraged the wider adoption of the model into the data security literature.

[2] This section is largely adapted from Ritchie (2013)

5.2 Effectiveness of standards delivery and implementation

There is a wide management literature on the development of standards and the roles that dynamics and the attitude of affected parties have on the outcome. Brunsson et al. (2012) provide a summary, identifying four elements of particular importance for the adoption and efficiency of a standard:

- the number of participants in the standard-setting process: more participants can increase acceptance, but more participants can also make agreement harder or water down the proposed standard to the lowest common denominator
- the tension between flexibility and adoption or diffusion: responsiveness to events or the interests of stakeholders may lead to a widespread adoption, but too much flexibility can make the standard too woolly, and limit adoption
- if standards are voluntary, 'stick' is more effective at encouraging adoption than 'carrot', but this is more likely to be driven by those with market power who have reached the critical mass to be able to impose their interests on others
- an 'expert group' setting standards creates a division between 'insiders' and 'outsiders', potentially creating unwillingness to comply irrespective of the value of the standard; the most effective way to kill any development may be to insist on it being adhered to

There has been less interest in standards in the economics literature, as the topic is theoretically straightforward: if a standard produces more benefits than costs, a rational actor should adopt it; the effort expended by an agent in developing a standard should outweigh the cost, as should the choice of whether to pursue an open or closed standard. Implicitly then the economics literature argues for the 'carrot' rather than the 'stick': standards are more likely to be a negative influence if imposed as a directive without relating compliance to operational benefits. Compliance is more likely to be achieved if it saves time or effort, or improves credibility. However, this is largely a conceptual argument.

The economics perspective does throw light on the value of a standard as a pure information device in a world of uncertainty. Standards may be the cost-effective way for an agent to acquire useful knowledge, as well as acting as a signal of intention for other agents. The interest here is when the agent has little incentive to develop a standard (or form a club to develop it), but there is a benefit to society generally in developing such standards. This provides a case for government intervention.

5.3 Communities of interest or practice

One approach to the development of common understanding may be the development of 'communities of interest'[3]. These are usually described as groups of interested/expert parties getting together to share experience and develop practices or protocols. As such they have a direct relevance to the development of standards by providing either review groups for both open and closed standards. In the case of open standards, there is the opportunity for the community of interest to form the core of the development group.

Several expert groups in the UK have grown up organically to address issues of research data access:

- Safe Data Access Professionals (SDAP) – co-ordinated by UKDA
- Safe Researcher Training (SRT) Expert Group – co-ordinated by ONS
- International Secure Data Facility Professionals Network – co-ordinated by UKDA
- SDC-REBOOT – co-ordinated by UWE Bristol

[3] In the computing literature this has a specific meaning in network theory. This is not the meaning here.

- UK-TRE – co-ordinated by Universities of Dundee/Swansea

Most of these are associated with social science trusted research environments (TREs) such as the Secure Data Service or Scottish Safe Havens, as these comprise a relatively compact community who interact regularly. SDC-REBOOT and UK-TRE were both funded by MRC under the DARE programme for several months in December 2023 - March 2024 (UK-TRE as a pre-existing group, SDC-REBOOT as a development of the DARE SACRO project) to support their establishment, and are likely to continue.

The aim of these groups varies. The SRT expert group is primarily about sharing information on SRT and related training developments within the group. ISDFPN is fairly new, and so far is mainly bringing interested parties to present material of value to the group. SDAP is more focused on operations, and has developed a widely-used SDC manual (Greci et al, 2019) and staff progression profiles in the past. UK-TRE is a discussion group for general TRE matters, mainly led by technical teams at TREs. SDC-REBOOT focuses on output checking (traditional and AI), and has been used to promote and get feedback on the SDC manual developed under SACRO (SACRO, 2023), and on an operational guide to SDC developed under the FDS aegis (Green et al, 2024).

The groups also vary in terms of membership. SDAP and SRT Expert Group are fairly fixed. ISDFPN is new and so seems to be acquiring members. SDC-REBOOT and UK-TRE are large networks, with mailing lists and regular events open to all interested parties.

While these provide effective engagement opportunities, it is not clear (certainly to outsiders, possibly to members as well) what the functions of these groups are: are they for the members, or for the community? If the latter, how do others engage? Are they the source of standards and protocols, or the review body for things defined elsewhere? Should they have authority?

The main concern for communities of interest is how to manage the balance between being a talking-shop (nice for the members, but achieving little) and being a closed shop (trying to impose a particular view of the participants on the world).

5.4 Relevance to FDS

Both process and outcome standards potentially can have a role to play in data services. At present there are only three extant standards commonly accepted across a wide range of data services:

- Safe Researcher Training: a pre-requisite for access to most social science TREs, with the material taught by multiple organisation and with accreditation details shared between organisations
- MRC training: this has widespread acceptance amongst the health researchers as a minimum standard of training
- Digital Economy Act accreditation: this is required for TREs to hold data made accessible under the DEA

SRT is widely seen as an effective tool for TRE researchers, and the presence of an Expert Group (including, but not limited to, all the delivery partners) to review it has ensured that it continue to meet TRE needs. MRC training is often seen as a baseline for public health researchers, but it is generally not accepted by itself for TRE access. DEA accreditation has proved problematic, and provides a good example of where the standard appears to have stifled innovation. It is worthwhile comparing SRT and DEA accreditation, both of which are ‘process standards’. SRT was developed by a small team, but with significant and ongoing input from delivery partners, with the aim of creating an open standard (Green and Ritchie, 2023b). In contrast, DEA accreditation is a closed standard,

developed and imposed by the UK Statistics Authority with very limited input from non-governmental bodies.

Other standards currently under development include common data access agreements being developed by HDR UK, and a Researcher Registry being funded by MRC. Both of these are system standards but are open, having been developed with a wide range of stakeholders; both will also aim to have adoption by 'carrot' (ie because it solves problems) rather than the stick (there is limited opportunity to enforce adoption). Both are also the product of a small team taking charge of development whilst consulting widely on the path of development. This may be the model that FDS needs to adopt as providing an effective operating model, neither monolithic directive nor unrestricted free-for-all. FDS is sponsoring the development of common specifications for job roles, training needs and vacancy advertisements for data services staff; this is also a system-open standard, and being developed on the small-team-big-network model.

6. Incentives and disincentives of alternative funding models

Funding streams will have an impact on what data services can be provided, and how. In this section we briefly discuss the conceptual issues.

6.1 Uncertainty, measurability and flows of activity

UKRI provides long-term funding for data services. Uncertainty about the future means that a funder needs to choose where risks and incentives should lie. The two extreme cases can be characterised as follows:

	Option 1: Fixed-price	Option 2: cost recovery
What it means	• The funder and contractor agree a payment in advance of any work	• The contractor carries out the work, and charges the funder
Financial risk	• Borne by contractor	• Borne by funder
Advantages to funder	• Known cost • Contractors incentivised to minimise costs of tenders	• Ability to respond to changes in circumstance (positive and negative)
Disadvantages to funder	• Supplier may go bankrupt or otherwise fail to deliver • Supplier has incentive to minimise work done	• Risk of unexpected costs • Supplier has incentive to maximise profitable work charged for
Works best when	• Cost of providing services is relatively well-known and predictable	• Contractor and funder have shared incentives

The usual answer is to have risk-sharing models, where there is a mix of fixed funding and flexibility; the funding may be staggered over time to allow regular evaluation, and it may be subject to maximum and minimum payments. This is generally how large government projects are funded. The economic theory behind this is solid, based around principal-agent theory: the less measurable the target, the more scope for cheating by the supplier, the more the funder should try to be specific.

While a useful theoretical framework, the risk-sharing model depends on lots of assumptions about what information is available. The key problem is the measurability of necessary cost, outcomes and funding targets. There is a substantial information asymmetry, with funders unlikely to have either the detailed knowledge of the costs of running a data service, or the resources to interrogate accounts with sufficient detail. There is however a power symmetry too: the requirement on public funders to achieve value for public money, and the limited options for data services to find alternative sources of income, means that the needs of the funders may dominate the outcome.

It also matters whether:

- the project is small (for example, developing SRT cost under £50,000), which limits financial risk for the funder
- the impact is small (for example, building the ARDx data set) which limits non-delivery risk
- the activity is repeated and measurable at different times (for example, UKDA service provision) which gives multiple opportunities for feedback

If the cost is high and/or the impact is large (and hard to undo), and if this is a one-off investment, managing uncertainty is going to be a key element of the plan.

There is one major problem with the theoretical framework: the assumption of competing goals (outcomes for the funder; profit for the contractor). This may not be valid when the data service shares the same, or similar, mission as the funder. Without conflicting incentives, it seems feasible to treat the funder-contractor partnership as a single goal. Unfortunately, this has rarely been examined in the economics literature as, from a mathematical perspective, the outcomes are all trivial and not worthwhile publishing.

6.2 Funding for large infrastructure projects

Large infrastructure projects face additional issues, compared to funding for services/small projects:

- Need for significant upfront commitment to enable planning and development
- May be few (or only one) potential suppliers
- Long time scale which may outlast contracting parties
- May be difficulty specifying the project goals
- Once funding is committed, there may be a strong incentive to carry down that road

The third issue is a perennial problem in UK contracting, with large infrastructure projects agreed on unrealistic terms by people in govt who then move on. At least, that this the perception; it is not clear how much this is a stereotype, and how much it is a genuine problem.

The Integrated Data Service provides a strong example of the last two problems. The IDS programme was started with a very limited understanding of user needs or research data provision[4], but with a tightly-defined timetable to spend significant amounts of public money. The IDS regularly passes the gateway reviews which compare actual performance against target performance, but widespread criticism of the unsuitability of the programme for the UK data landscape reflects concerns that the benchmarks for the IDS are poorly aligned with societal outcomes.

6.3 Relevance to FDS

There is lots of economic theorising around this, and some applied work. However, it is not clear whether this is the big problem suggested by the literature on government funding. The principal-agent assumption of conflicting goals may not be relevant for data services, where the suppliers are also heavily invested in the intrinsic value of the project (desire for recognition, shared ambition between funders and suppliers, personal commitment of individuals to the field). Moreover, although data services within a country may be unique, internationally, they are rarely so. There is therefore at least some opportunity to do comparative cost analysis, as well as to have experts judging proposals.

[4] Source: regular discussion by the author with IDS team at various levels

Part II Models of effective decision-making

In this section we consider the evidence of what works in data services: what is effective, what is efficient, what makes decisions happen, how decisions should be made. As this section seeks to identify good practice models, we move away from the neutral study of concepts of the previous section, to normative evidence-based models (ie what we should be doing).

7. Why attitudes matter

7.1 Framing

All decisions in data services are subjective, whether directly (choosing to follow one or another course of action) or indirectly (choosing to be bound by an external set of rules or conditions). Hence, all decisions are influenced by the attitudes of the decision-makers.

Consider the basic question of acting lawfully. Is this an objective for any data service? When the authors conduct shows-of-hands with audiences familiar with data governance, the answer is overwhelmingly 'yes': why would you not want to act lawfully? In fact, as Ritchie (2014a) shows, the answer is 'no'; not because acting unlawfully is encouraged, but because acting lawfully is not an *objective*; it is a *constraint* on one's action. Acting ethically is a legitimate objective, but acting lawfully is a condition of the service.

Once lawfulness is seen as a constraint rather than an objective, other changes follow. Challenging established practices becomes easier, as the challenge is to the *objective* of the data service, not the constraint of lawfulness. Formal advice becomes more useful: rather asking legal advisors what one's options are, advice becomes "this is what I *want* to do; *how* do I make sure this is lawful?" which is a much better question from the perspective of both advisor and advisee. It breaks the link between what *is* done and what *can be* done.

7.2 The psychology of default values

The importance of framing is well-founded in psychology. Equally important is the choice of 'reference point' – what is seen as the default. Losses are seen as more important than gains, as so where you start from matters. Economists, in particular, have had significant success in demonstrating, both theoretically and empirically, the importance of defaults. Consider a survey of village residents being asked to consider whether developers should be allowed to build near to a popular viewpoint. Asking the question "how much are you willing to accept to allow this to happen?" (ie the default is no change) always returns a higher valuation than the question "how much would you pay to stop this happening?" (ie default is that the change happens). This is so well established by economists and psychologists that is merits no discussion outside classrooms.

This is directly relevant to data services. Consider making a decision about whether to share data for research. The data holder could take the position "no release unless I am satisfied all meaningful risks have been dealt with" (default-closed), or alternatively "release unless I am presented with a reasonable risk I cannot deal with". In theory, these two positions should give the same outcome; in practice, they differ substantially. This is not surprising when we consider what each position entails giving up:

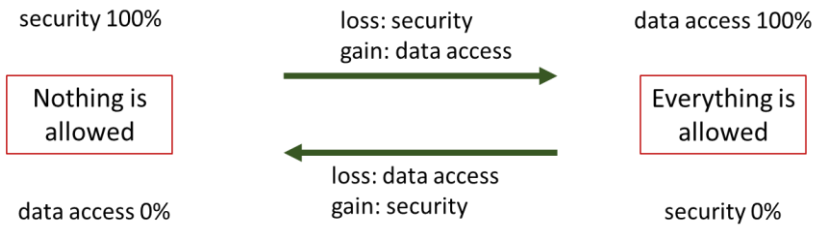


Figure 3 Impact of different default positions (from Ritchie, 2014a)

The default-closed perspective ("nothing is allowed") starts with complete risk protection and zero data access; any change for that position entails increasing access but removing risk protection. In contrast, the default-open position ("everything is allowed") starts from zero risk protection but full data access; any change entails more risk protection but a loss of access. We expect each decision-maker to choose a position closer to their starting point than the other would choose, and therefore the default matters.

Decision-makers sometimes argue that their default position is the 'correct' one. There is no foundation for this. It is clear in the above example that the 'allowable' criterion is the same in both cases (all reasonable risks have been managed); this is the point.

Consider the argument that data should only be shared on a need-to-know basis. This seems to imply a default-closed position: data can only be shared once the use of it has been established. But it can also be applied to a default-open position: data can be shared unless it can be demonstrated there is no use for it. The burden of proof shifts: in the default-closed world, the user must demonstrate need and the value of that use is being judged; in the default-open model, the data holder must demonstrate that there is no value in allowing the user that data, and the risk to privacy is being judged. In a world of perfect information it is possible that the two sides would agree. However, in an uncertain world such as research use, where the activities and outcomes by definition are unpredictable, differences are more likely; and as the uncertainty grows, so do the differences.

Two common arguments for defensive (default-closed) models are (1) the 'precautionary principle' (2) social justice demands that privacy is the most important criterion. We consider each in turn.

The precautionary principle argues that, where the outcome of taking (or not taking) an action is uncertain, greater weight should be given to negative outcomes, particularly those which have permanent effects. It is a variation on the 'do no harm' idea. Superficially, this seems sensible. In practice, this gives enormous weight to the status quo. It also encourages over-regulation, as constraints almost always pass the 'do no harm' test, whereas relaxation of rules needs to be demonstrably harmless. This can be seen in the sclerosis of many UK government processes, where protective rules are applied without effective review (see case studies below). The precautionary principle can also be abused to block change indefinitely, as hypothetical cases of downsides can never be exhausted.

The argument that privacy is the key criterion rests on the confusion between objectives and constraints noted at the start of this section. If the 'need to maintain privacy' is an objective then the default-closed option is sensible; moreover, the objective is achieved easily, by refusing all use of data. However, once it is acknowledged that maintaining privacy is a constraint, not an objective, then the default-open perspective is a better starting point. A proposed solution can be tested against the constraint: does it meet the privacy criterion or not? If not, then the solution can be

revised until the constraint is met. The alternative, of starting with constraints and thinking "what can we do now?" is both inefficient and ineffective.

Ultimately, the choice of a default reflects power dynamics. A default-closed position gives power to those who hold data; a default-open position gives power to those who want to use it. Outcomes also depend on personal factors: a data holder who is not prepared to countenance any release, or a user who has no understanding of confidentiality concerns, may be able to exert undue influence on the outcome, and the starting point will affect that individual's success. However, it is essential to note that, even with all parties well-briefed and well-disposed to reach agreement, the default position will affect outcomes.

But default positions are not fixed: the framing of the topic affects them. In our experience, alternative framings can have substantially different results; see, for example, Hafner et al (2019) where a new framing of risk dramatically affects the de-identification of a dataset. More broadly, the entire early history of the Virtual Microdata Laboratory at ONS (2003-2011) is a continual reframing of problems in data access.

Finally, we note that a default-open position aligns better with principles-based design and regulation, which is focused on outcomes rather than constraints.

7.3 Risk, risk aversion and incentives[5]

Ritchie (2014b) argues that public sector decision-making is generally default-closed, and that this is largely due to incentive structures in government which emphasise decision-avoidance and penalise innovation, rather to any inherent risk-aversion on the part of public sector employees. There is no evidence that public sector employees are more risk-averse than is beneficial to society (being more risk-averse than the private sector is not sufficient criterion in itself), but there is evidence that such employees are more willing to accept or follow the decisions of others.

Ritchie (2014b) argues that, if this indeed the case, the solution is not simply to reduce bureaucracy; this might in fact increase decision-avoidance by reducing the 'comfort blanket' allowing innovations to be supported. Instead the solution is change attitudes through (1) education and (2) changing incentives to innovate. While the latter is a problem of the public sector, data services can play a significant role in helping to educate decision-makers in government about relative risks. In particular, actively promoting good practice can make a significant difference to the attitudes of decision-makers by reducing the zone of uncertainty.

8. The EDRU model of data governance

8.1 The 'traditional' model

Models of data governance have been fundamentally defensive in nature; the focus is on the costs and risk to the data owner, and models assume that the primary aim of any data access strategy is to prevent malicious misuse. There are strong incentives in the public sector to take a risk-averse approach to decisions around confidential data (Ritchie, 2014b). This approach (which we refer to as 'traditional') is *default-closed*; that is, it assumes that no access will be granted unless it can be proven to be safe. Natural risk-aversion in decision-makers is encouraged by a literature that emphasises worst-case scenarios and protection against hypothetical possibilities.

As an example of how defensive planning can reduce the effectiveness of data service planning, consider 'safe data'. Statistical disclosure control (SDC) applied to microdata reduces data detail and

[5] This topic is covered in considerably more depth in Ritchie (2014b) and Ritchie and Welpton (2014)

so risk; this is called 'input SDC' (to distinguish it from the prevention of breaches in outputs, or 'output SDC'). Input SDC is a well-established field of research stretching back to the 1970s. It has a coherent methodological framework, a large range of statistical techniques, open source software to support data holders, and general agreement on problems and the advantages and disadvantages of solutions. However, Hafner et al (2015) and Hafner et al (2019) argue that the application of input SDC in real situations over-protects the data and undermines public benefit due to

- A data-centred perspective which assesses inherent data risk in isolation from its use
- The use of hypotheticals and mathematically tractable 'worst case' models rather than evidence in the assessment of risk
- A 'default-closed' position, placing the onus on the data holder to consider whether all possible risks have been addressed
- Use of arbitrary statistical models to generate spurious objectivity

Similar concerns can be raised about historical attitudes to the other five safes, often stemming from an assumption that data users are inherently untrustworthy and require active policing. It has been argued that the lack of breaches of confidentiality is evidence that the traditional model works well. As the traditional model has, until recently been the universal ethos for all data access solutions, this statement cannot be challenged. However, an increasing number of data access solutions adopting the EDRU ethos (see next sub-section) provides evidence that the traditional model has been over-cautious and missed opportunities to create synergy with the research community; as such it has not served society well.

8.2 The evidence-based, default-open, risk-managed, user-centred (EDRU) approach

The EDRU model (Green and Ritchie, 2016) proposes an alternative conceptual framework, reversing many of the implicit assumptions used in decision-making. The elements are:

- Evidence-based: use of empirical (rather than modelled) evidence and plausible scenarios for risk assessment, and acknowledging uncertainty
- Default-open: intending to supply or share data unless negative factors cannot be overcome, in contrast to the default-closed models where nothing happens unless every risk has been managed (see section 8 above)
- Risk managed: considering risk as a spectrum, including benefits foregone by not sharing; acknowledging trade-offs made, rather than trying to 'minimise' risk as traditional models are wont to claim
- User-centred: identifying user needs as the primary objective and working backwards, rather than the data-centred approach of the traditional model

This approach emphasises decision-making focused on achieving the goal of data access: "how do we make X happen?" This contrasts with the historical, defensive approach which is perhaps characterised by "Should we make X happen?"

The EDRU approach conceptualises and integrates coherently a number of underlying themes in this area, such as 'safe statistics', 'active researcher management', 'circles of trust' or user-centred training. As such, it is difficult to say whether organisations have adopted this approach or are merely selecting useful elements from it. Green and Ritchie (2016) is an exception where the organisational data strategy was explicitly designed using EDRU motifs.

Ultimately, the EDRU ethos provides a more sustainable world-view as it starts from first principles rather than historical precedent. On the limited evidence available, EDRU-based operations are

more likely to provide a secure and useful data access solution; they also seem better suited to exploit the gains from increased data access by engaging with researchers more.

9. Determining performance standards

For data services, performance standards can be hard: much of their business might involve intangible inputs, activities or outputs. However, some idea of relative or absolute performance is essential for measuring the effectiveness of operational change. In this section we consider alternative models for defining performance.

9.1 Cost-effectiveness analysis

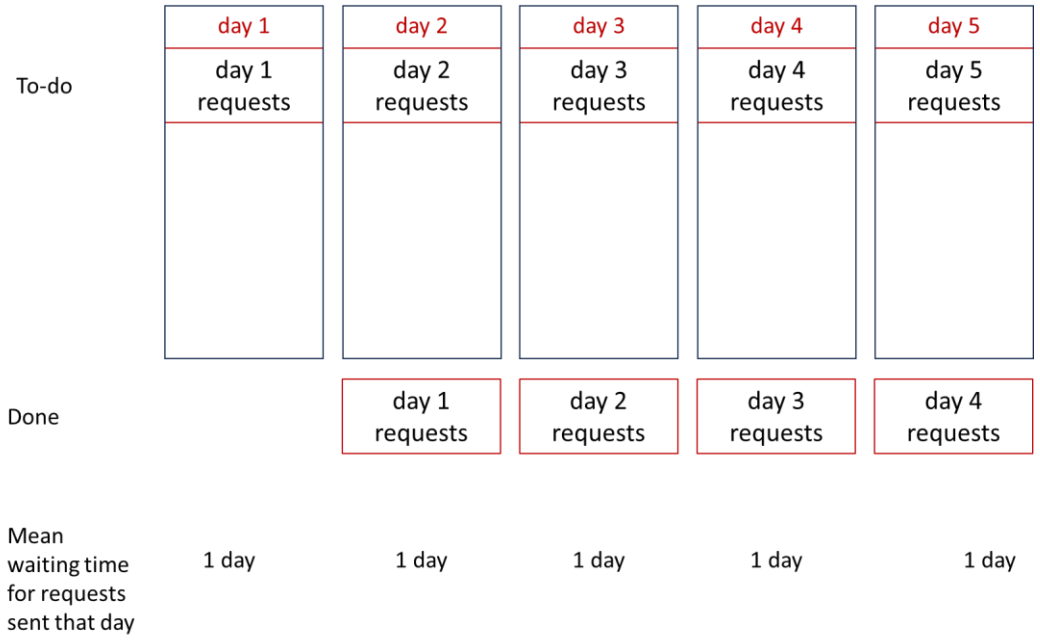
Cost-effectiveness analysis (CEA) seeks to answer a relatively simple question: Is activity, output or outcome X being achieved with the minimum amount of resources? This can be an effective question to ask, as it does not require the outcome to be evaluated or changed; the focus is on delivery mechanisms.

The limitation of CEA is that it focuses on known ways of doing things, and less on challenging accepted assumption. It can be transmuted into "is the way we currently do things as good as we can get, given our current staff, skills, infrastructure and resources?" or, even worse, "are we wasting any resources?" The answers to these are likely to be 'yes' and 'no'; contradiction can be seen as criticising staff for failing to spot potential improvements.

9.2 The operational potential model

Many data management processes have time allocated to them far in excess of the actual amount needed. For example, output checking response times are often in terms of days; access approval processes can take weeks or month; and yet in both cases the actual time spent on the approval of the output or application is significantly lower. The bulk of the time is spent with activities in queues.

The common response to this is 'we don't have the staff/resources'. This argument makes no sense if the response time is constant. Suppose it takes 1 day to clear the output requests in a typical day (we use a day as the minimum response time in all these examples, to allow for administrative delays and other demands on both parties). Then the response time is a constant 1 day:



Supposing the data provider clears no outputs on day 2. Then the response time increases but remains constant at a higher level:

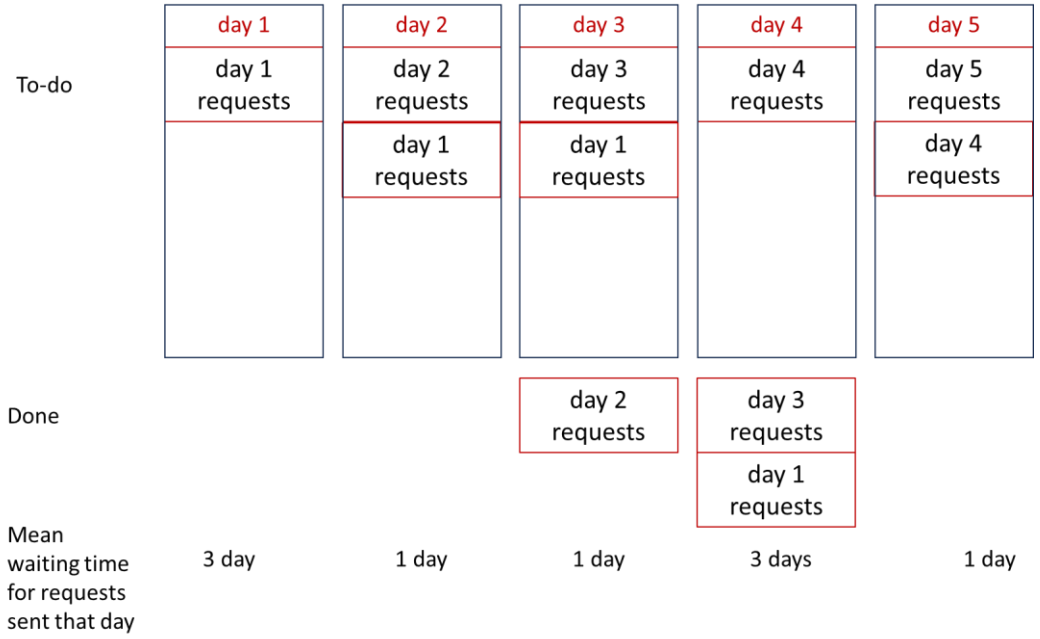
	day 1	day 2	day 3	day 4	day 5
To-do	day 1 requests	day 2 requests	day 3 requests	day 4 requests	day 5 requests
		day 1 requests	day 2 requests	day 3 requests	day 4 requests
Done			day 1 requests	day 2 requests	day 3 requests
Mean waiting time for requests sent that day	2 days	2 days	2 days	2 days	2 days

And if another day has no responses:

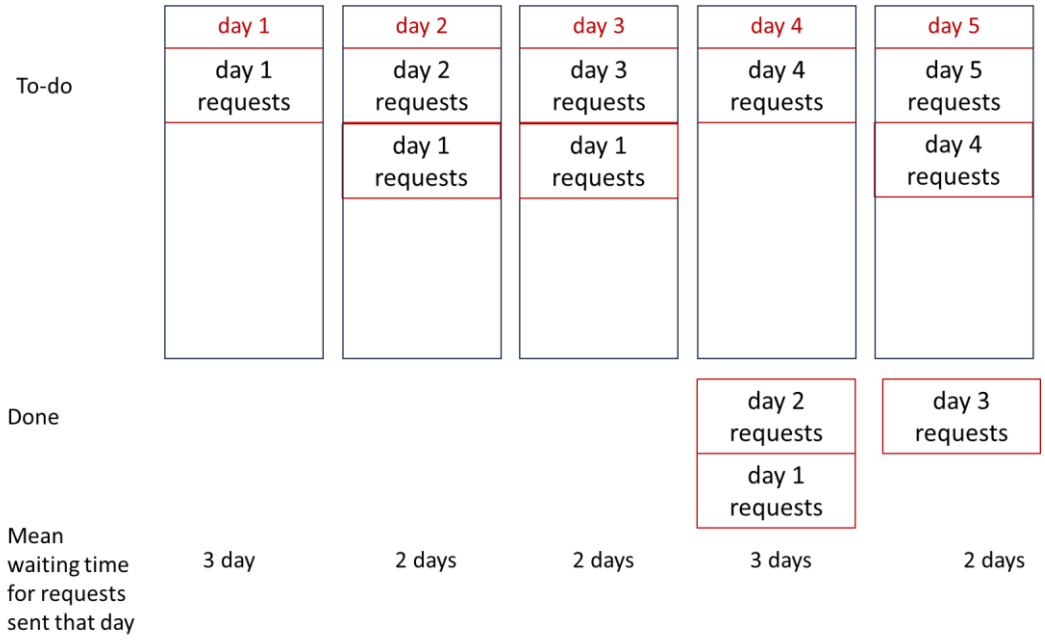
	day 5	day 6	day 7	day 8	day 9
To-do	day 5 requests	day 6 requests	day 7 requests	day 8 requests	day 9 requests
	day 4 requests	day 5 requests	day 6 requests	day 7 requests	day 8 requests
		day 4 requests	day 5 requests	day 6 requests	day 7 requests
Done	day 3 requests		day 4 requests	day 5 requests	day 6 requests
Mean waiting time for requests sent that day	2 days	3 days	3 days	3 days	3 days

Thus the wait time increases, *despite there being enough staff to handle the daily throughput*. So, a constant response time implies enough staff to meet the needs of the service. A response time greater than the minimum implies a backlog has been allowed to build up. An increasing response time implies insufficient resources, and a shortening response time implies more than adequate resources to deal with the regular traffic.

Data services may argue that the minimal response time is too tight, as there are legitimate delays: staff may be temporarily absent, or take time to bring up to speed, or specific cases might take more time than the minimum. For example, consider the case where one output in 3 takes 2 days to clear:



The average response time, measured over three days is 5/3=1.67 days. So, a target response time set at 2 days would cover this. A problem occurs if the data service, rather than seeing this as a margin of error over the minimal response time, interprets this as a target response time:



By using the full allocated time (designed to allow for variation) as the expected response time, the data service ends up with being unable to meet the target. In the above example, 3-day average response time are now 7/3=2.33 days.

Hence the service needs to have two target times: the **internal** target, reflecting the **median** time to clear the output; and the **external** target communicated to users, reflecting the **expected maximum clearance** time for **most** outputs. In the above example, this gives an internal target of 1 day, and an external target of 3 days; one would expect 100% success in meeting the target and 67% success in over-achieving the target.

The data service could consider a third time, the **exceptional** time: this covers extreme events. In the output-checking example, this may require getting third-party opinion. Suppose this second opinion requires three weeks. Clearly including this in the mean calculations for the external target misrepresents the expected response time (it is in the data service's interest to provide a target which it regularly beats, but it does provide temptation not to meet the internal target). A more honest solution might be to triage outcomes: at the end of the internal target time, it should be clear whether the output can either be

1. Cleared
2. Cleared within the external target time
3. Referred elsewhere for a longer consideration

In the third case, the data service needs to manage this by communication with the user; extending the target time to cover extreme events is not acceptable.

Some data services will argue that this does not reflect forces beyond their control; for example, an access process which requires getting external approval from a data owner. This is nonsense: the same arguments as above still hold. Suppose approval requires the agreement of Ministry X, which takes one day once someone in the Ministry looks at it. If the Ministry is given one day to respond, it should take one day. If it is given five days to respond, there is a very large amount of organisational behaviour research which says that it will typically take five days. In the mean time, another four days of requests have accumulated, so from day 5 onwards the Ministry is still dealing with one requests a day but it is also facing justified criticism for taking five days to approve a request. Putting off the request benefits no-one.

Again, it could be argued that a response time above the minimal time is needed to allow for, for example, key individuals to be away or have higher-priority duties. This can all be in the service-level agreement between the data service and the Ministry, and it makes clear that the reason for the delay is not the time taken to assess it, but the delay caused by the Ministry's operational needs.

The value of this approach, which we call 'operational potential' is threefold. First, it clarifies where and why delays are *allowed* to happen; and 'allow' is important, as it is makes clear that these are organisational choices. Second, it reinforces that procrastination does not reduce the problem; on the contrary, the problem remains but the blame is shifted onto the non-responding party. Third, this provides an incentive for organisations to improve their performance: in a well-designed system, being better than the mean and no worse than the target maximum should be achievable.

9.3 Comparative (international) analyses

Data services are often seen as a special type of organisational unit: they are part of a small group of specialist service providers; they provide a service which is unusual; and they embody a lot of expertise which is not available – or required - elsewhere. This allows them to claim that what they do is justified by their own specific needs. Daa services can be natural monopolies within their countries, particularly those associated with NSIs.

However, there is no data services organisation which operates in such a way as to be unique in human experience. Some models – such as remote job servers like Microdata.no, OpenSafely or Lissy – are relatively rare, but there are sufficient examples to learn from each other. Other models, such as traditional Remote Data Centres, are very familiar. In terms of specific services (such as researcher certification or output checking) these are extremely well understood, with a very large number of examples.

International comparators can also be helpful to allay fears of not-me-first. Ritchie (2014b) gives this example:

> *in one country I observed a new, potentially high risk, process being proposed. The relevant committee was unwilling to commit to an innovative new system, despite the numerous advantages. However, when it was pointed out that a similar process had already been used in another country for a similar purpose, the committee changed its mind. Note that the committee did not directly question the appropriateness of the comparison; the key point was that this country was not the first to try this system for this general purpose. The committee's decision was validated by the precedent set by the other country.*

(Ritchie, 2014)

The major concern with comparative studies is ensuring that the comparison is meaningful. All comparisons will differ in some respects; the question is whether those differences matter. This is akin to considering bias in a statistical study: no study will be perfectly bias free, but is the potential for bias likely to affect the results of the analysis?

For example, consider comparing the UK remote job system OpenSafely to the Norwegian remote job system Microdata.no. Comparisons about the way that code is shared, output checks, or the use of standard or restricted coding languages are all valid. On the other hand, comparing the access route for Microdata.no to OpenSafely is invalid, as the former is coupled to Norway's personal registration system, which is not available to UK facilities. However, the Norwegian access route can still provide useful examples of how an automatic approval process can facilitate secure access.

10. Reducing research bureaucracy – the Tickell report[6]

Finally, we consider lessons learned from another in-depth study into how the research landscape can be improved. In 2023 the Tickell report on research bureaucracy[7] was published. Bureaucracy is described as "the accumulation of everyone's good ideas". The report explores the ways that bureaucracy can weigh down processes. What can be seen is that many of the concerns (and solutions) above are reflected in this overall study.

10.1 Principles and key themes

The Review developed the following seven principles for cutting unnecessary bureaucracy:

- **Harmonisation** - Reducing the volume of administration through the use of common processes between different funders to make essential work easier.
- **Simplification** - Reducing the complexity of individual processes.
- **Proportionality** - Ensuring that the obligations placed on researchers and institutions are commensurate with the size of the risk or reward.
- **Flexibility** - Supporting excellence wherever it is found and not excluding research that does not fit within narrowly defined parameters.

[6] This summary was written by Cara Kendal of the DRAGoN team at UWE.
[7] https://www.gov.uk/government/publications/review-of-research-bureaucracy

- **Transparency** - Communicating the rationale for systems and processes which have a bureaucratic burden.
- **Fairness** - Developing approaches to systems and processes that support fairness, rather than erode it.
- **Sustainability** - Cutting bureaucracy in ways that avoid destabilising the system to deliver a more efficient system over the long term.

10.2 Key findings and recommendations

Assurance

Findings:

- The research sector faces numerous complex and duplicative assurance requirements, leading to risk aversion and over-compliance. Lack of trust, coordination, and knowledge exchange hinders effective assurance processes. The bureaucracy grows incrementally, introducing new requirements but few attempts to remove or reduce redundant ones.
- Recommendations:
  Government departments funding research should work together to improve alignment of assurance approaches, removing duplication. UKRI should take action to achieve coordination across councils. Government should facilitate collaboration with other funders. Funders and research organisations should develop collective approaches to support institutions in managing assurance processes and explore self-certification or earned autonomy for institutions with robust assurance records.

Applying for funding

Findings:

- Researchers are concerned about the length and complexity of application processes, with low success rates (around 20%) for research grant applications. Single-stage processes, which require applicants to provide all information at the outset, lead to unused information and wasted time for most applicants. Two-stage approaches may improve the system but may present resourcing challenges or take more time. The review discusses managing the prospect of streamlined application processes that could lead to higher application numbers. Funders are already addressing these issues, but there is potential for further improvement.

Recommendations:

- Funders should experiment with application processes to reduce applicant burdens, including two-stage processes with information requirements scaling with likelihood of receiving funding. They should work together to standardise language and questions, with UKRI facilitating this across Research Councils. They should review adaptations to assessment processes, including information for national security assessments and innovative approaches like peer reviewer triage to experimenting with new models. They should ensure application processes support equality, diversity, and inclusion, and remove letters of support requirements from most circumstances.

Grant Implementation and In-Grant Management

Findings:

- The short award letter issue to project commencement period can hinder procurement, recruitment, and financial administrations. Conversely, it can take too long to get agreement

from funders on project changes. Project monitoring information requested by funders is often unclear in purpose to recipients, and contracting and collaboration agreements can cause delays due to research organizations using their own versions instead of standard formats like Brunswick or Lambert Agreements.

Recommendations:

- Funders and recipients should allocate sufficient time for tasks prior to project commencement, build in flexibility, and prioritize ethical and regulatory approvals. Universities and research organizations should use standard templates for contracts and collaboration agreements allowing faster and easier collaborations. Funders should also consider no-cost extensions to reduce delays and queries. Lead partners should be responsible for ethical and regulatory approvals, and counterparties should not require duplicative approvals.

Digital Platforms

Findings:

- Digital platforms in the UK need to support institutional diversity and adapt to change without being too complex. Greater harmonisation of digital platforms is possible, but it may be limited by individual funder objectives. Inter-operability and data sharing could reduce bureaucracy. Key funders like UKRI, NIHR, and Wellcome are moving away from older platforms and may be able to deliver improved services. Funders are driving forward programs to reduce bureaucracy, such as the Simpler and Better Funding programme, and the UKRI Funding Service piloted by UKRI, providing end-to-end functionality for all Research Council grant applications starting 2024.

Recommendations:

- Jisc should lead creation of sector-wide groups to oversee research information ecosystem development and integration in higher education. Funders, universities, and regulators should prioritize interoperability and improved data flows in any new digital systems. Improvement of existing systems using tools such as application programming interfaces, point-to-point integration, and machine learning should be explored.

Institutional Bureaucracy

Findings:

- Institutional bureaucracy is a significant source of unnecessary bureaucracy, with a culture of risk aversion affecting decision-making processes. This can lead to unnecessary approval hierarchies, delays, and operational difficulties. The use of generalist professional services departments for research support can cause longer delays due to lack of familiarity or confidence in handling grant agreements or contracts.

Recommendations:

- Research organizations should consider delegating research-related approvals to closer research managers and officers. Universities UK should collaborate on research management issues, such as increasing risk appetite and streamlining burdens through standardization. If not already in place, research organizations should establish "Trusted Fund" policies to enable risk-averse projects within specific parameters.

Communications

Findings:

- Frustration with bureaucratic requirements may stem from unclear communication of R&D funding systems and processes. Scope to increase awareness of existing tools that can reduce bureaucratic burdens. Proactive communication and engagement by funders and regulators can address uncertainty about implementing new requirements. Government and funders should engage with the sector on new requirements to identify the most efficient approach. Concerns about communication approach with the sector such as jargon and inconsistent language, ensuring communications are received by the right audiences, and ensuring timely submission deadlines.

Recommendations:

- Government, funders, and regulators should consult with research organizations before introducing new regulations. Proactive communication from government and funders on emerging regulatory issues is crucial, using the (RCAT)i model is good practice. Important research messages should be sent from funders to research office contacts as well as Vice Chancellor/Pro-Vice Chancellor Research.

11.Linking theories of change, planning and evaluation

FDS uses a standard Theory of Change (inputs, activities, outputs, outcomes, impacts) to identify potential areas of improvement. In this section we briefly consider how theories of change like this can be linked to broad principles and objectives, adapted from Alves et al, (2021, s4.3.1). For our purposes, 3 stages of the Alves et al (2021) method are relevant.

Stage 1 (programme planning): identify programmatic elements

The expected impact of the FDS programme should directly embody UKRI objectives

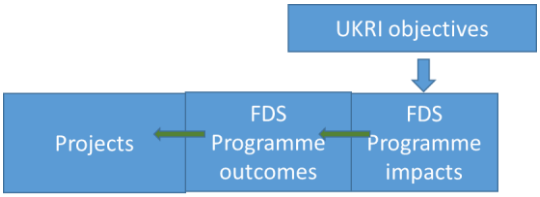


Figure 4 Linking mission to programme impacts (after Alves et al, 2021)

This is a strategic view to ensure that funded projects are linked to the overall goals of UKRI via the programme objectives. Creating a more practical, structured and detailed programmatic theory of change focuses on the impact of the programme as a whole, before working backward to identify potential projects through which the change is delivered.

Stage 2 (project planning): identify project theory of change

The programmatic theory of change documents the path from programme inception to projects. For each project, outcomes and impact can be defined, along with their relationship to programme outcomes and impacts:

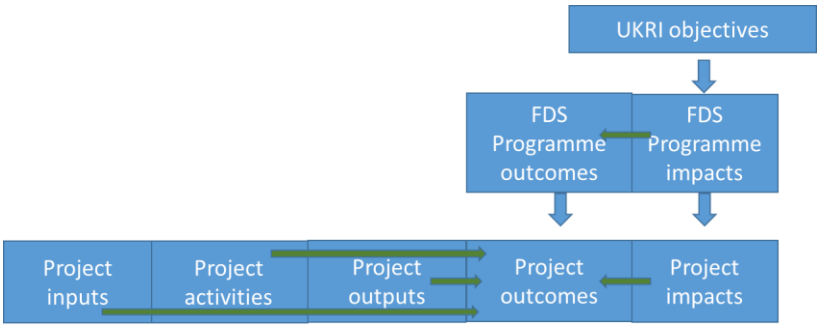


Figure 5 Linking programmes and projects (after Alves et al, 2021)

This ability to map the overall programme and then focus down on key projects is particularly important in this context given the wide range of recommendations likely to arise from the FDS programme. Linking programme goals to project goals means that gaps in the former can be more easily identified.

Stage 3 (project planning): identify what is measurable, and why

This approach also provides the basis for evaluating outcomes. FDS is likely to be required to demonstrate that any expenditure of public funds brings public benefit, ideally through a formal Economic, Social and Environmental evaluation (ESEe). Best practices for ESEe are defined in great detail in the Green Book (HMT, 2018). In practice, ESEe suffers from two significant problems.

First, ESEe is often carried out post hoc, asking "What did this project achieve and was it worth it?" Being aware of project goals changes the question to "What did this project achieve compared to what was expected?" This is a much more useful question. Stating the goals in advance limits the chance of post-hoc rationalisation of failure as a success on more limited terms, which is why it is the gold standard for clinical trials.

Second, and more important, ESEe assumes that everything is measurable to some extent. In practice, measurability tends to drop as one moves from inputs to impacts:

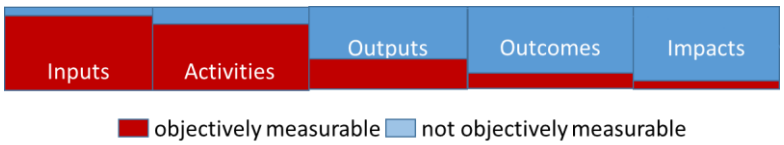


Figure 6 How measurability of outputs varies (after Alves et al, 2021)

Alves et al (2021) note that metrics often focus on what is measurable, rather than what is meaningful. This is because the economic approach to economic evaluation seeks to generate estimates of costs and benefits in comparable metrics, so that estimates of 'net benefit' can be generated. This is a largely pointless exercise: valuations are very subjective, and most of the outcomes or impacts have no useful valuations. Even an objectively measurable metric such as "% reduction in output clearance times" cannot be translated into impact metrics without making heroic assumptions about the time value of research. Hence, Alves et al (2021) and Whittard and Ritchie (2022) argue that the ESEe should take the form of trying to usefully assess: "*This* is the resource cost; *this* is the outcome; do we believe the latter is worth the former?"[8]

[8] Whittard and Ritchie (2022) explore the problem of trying to measure unmeasurable values in detail using a specific example.

12.The use of evidence and theories

The FDS project uses four sorts of evidence: primary, secondary, comparative and experiential.

As part of the FDS project, a large amount of evidence has been gathered from stakeholders such as

- research users
- research managers
- funders
- data service managers
- specialist professionals (eg output checkers or metadata specialists)
- data owners or those with input into the data release process
- regulators
- representatives of the general public

Evidence-gathering processes have included one-to-one interviews, reverse science cafes, and online and face-to-face workshops. This is **primary** evidence. An example is the evidence suggesting that data services are struggling to provide career profiles for specialist staff in non-academic roles.

Secondary evidence is evidence cited from existing studies, as in Part I of this document. This may be empirical, but it also includes theoretical or conceptual studies (such as, say, Ritchie, 2018, discussing the fallacy of 'spontaneous recognition'). An example is well-founded findings from organisational psychology that extended deadlines tend to result in procrastination rather than more flexible production.

Comparative evidence is primarily evidence from other countries, such as the way that different remote job systems are organised. For example, the Australian DataPlace system is an example of an effective application review process that is minimises user steps. Comparative evidence can also include historical evidence, such as information on the way the Secure Research Service has evolved from its predecessors; for example, the operation of the Microdata Release Panel provides an example of an efficient principles/exception-based approval process incorporating precedence and triaging.

Finally, **experiential** evidence is based on our own extended experience of what does or does not work in data services. Unlike secondary data, this may not have been formally written up, but it provides an important corollary to other evidence, and to theoretical models. This particularly true for assessing evidence around human behaviour. For example, theoretical models show substantial advantages for both researcher and data services to allowing screen sharing for remote users, and strong primary evidence in support. Concerns about screen sharing are focused on the theoretical argument of whether researchers can be trusted to use screen sharing appropriately. Experiential evidence (of this particular example, and of other examples where researchers have been given responsibility for self-regulation) supports the ability of researchers to self-regulate, and highlights the lack of evidence for the general 'risky researcher' hypothesis.

References

Alves, K., Tava, F., Whittard, D., Green, E., Beata Kreft, M., & Ritchie, F. (2021). Process and economic evaluation of the ODI R&D programme: Final report. London: Open Data Institute

Brunsson N., Rasche A., and Seidl D. (2012) The dynamics of standardization: three perspectives on standards in organizational studies, *Organization Studies* 33 613-623 doi:10.1177/0170840612450120

Green E and Ritchie F (2016) Data Access Project: Final Report. Australian Government Department of Social Services. June.

Green E. and Ritchie F. (2023a). "The Present and Future of the Five Safes Framework". *Journal of Privacy and Confidentiality* 13 (2). https://doi.org/10.29012/jpc.831.

Green, E. and Ritchie F. (2023b) "Using Pedagogical and Psychological Insights to Train Analysts Using Confidential Data". *Journal of Privacy and Confidentiality* 13 (2). https://doi.org/10.29012/jpc.842.

Green E., Kendal C., Ritchie F. and Alves (2024) Guide to output checking processes. UWE DRAGoN. https://uwe-repository.worktribe.com/output/13324670

Hafner, H., Lenz, R. & Ritchie F. (2019). User-focused threat identification for anonymised microdata. Statistical Journal of the IAOS, 35(4), 703-713. https://doi.org/10.3233/SJI-190506.

Hafner H-P., Lenz R., Ritchie F., and Welpton R. (2015) "Evidence-based, context-sensitive, user-centred, risk-managed SDC planning: designing data access solutions for scientific use", in UNECE/Eurostat Worksession on Statistical Data Confidentiality 2015, Helsinki

HMT (2018). The Green Book: Central Government Guidance on Appraisal and Evaluation. London: HM Treasury.

Keenan P. (2020) "Dictum meum pactum": UK regulation: Rules or Principles. Keenan Regulatory Consulting.

McEachern S. (2021) CADRE Five Safes Framework - Conceptualisation and Operationalisation of the Five Safes Framework. Co-ordinated Access for Data, Research and Environments. doi:10.5281/zenodo.5748610

NRC (2014) Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences. National Research Council doi:10.17226/18614. ISBN 9780309298063. PMID 25032406.

Ritchie F. (2013) International Access to Restricted Data: A Principles-based Standards Approach. *Statistical Journal of the IAOS* 1 Jan. 2013 : 289 – 300. https://doi.org/10.3233/SJI-130780

Ritchie F. (2014a) Access to Sensitive Data: Satisfying Objectives Rather than Constraints. *Journal of Official Statistics*, Vol.30 (Issue 3), pp. 533-545. https://doi.org/10.2478/jos-2014-0033

Ritchie F. (2014b) Resistance to change in governments: risk inertia and incentives. UWE Department of Economics working paper no.1412. December

Ritchie F. (2017) The 'Five Safes': a framework for planning, designing and evaluating data access solutions. Data For Policy Conference 2017. September

Ritchie and Green (2020) Frameworks, principles and accreditation in modern data management. UWE DRAGoN working paper. https://uwe-repository.worktribe.com/output/6790882

Ritchie F. and Welpton R. (2014) Addressing the human factor in data access: Incentive compatibility, legitimacy and cost-effectiveness in public data resources. UWE Economics working paper 1413 https://uwe-repository.worktribe.com/output/807316

SACRO (2023) *The SACRO Guide to Statistical Output Checking*.
https://zenodo.org/records/10054629

Tickell A (2022) Independent Review of Research Bureaucracy.
https://www.gov.uk/government/publications/review-of-research-bureaucracy

Whittard D., Ritchie F., Musker R. and Rose M. (2022) Measuring the value of data governance in agricultural investments: A case study. *Experimental Agriculture* 58:8
https://doi.org/10.1017/S0014479721000314