Lessico Intellettuale Europeo

STAY TUNED TO THE FUTURE

IMPACT OF THE RESEARCH INFRASTRUCTURES FOR SOCIAL SCIENCES AND HUMANITIES

edited by
Bente Maegaard and Riccardo Pozzo
with Alberto Melloni and Matthew Woollard as co-editors

LEO S. OLSCHKI EDITORE
2019

Questo volume è stato pubblicato grazie al contributo della Fondazione per le Scienze Religiose Giovanni XXIII

ISBN 978 88 222 6643 9

TABLE OF CONTENTS

WELCOMING ADDRESSES

INTRODUCTION

CONFERENCE KEYNOTES

PART I

Conceptualisation of Impact

PART V

Increase of Impact through Interaction of Domains

Matthew Woollard, Victoria Moody

UK DATA SERVICE: IMPACT-DRIVEN APPROACH TO SERVICE DELIVERY

Abstract

This paper has two key aims. First, to explain, promote and communicate methods of measuring impact (defined below) in European Research Infrastructures. Second, to examine some of the commonalities of approach in using impact measures as a factor in funding, and refunding (sustainability) RIs. By implication we shall promote the maturity of the SCI community in this area and also by implication we want to try and turn the impact of a service into a mechanism which funders can use to continue their investments into data service infrastructure.

Introduction

This paper has two main aims. Firstly, we explain, promote and communicate methods of measuring impact (defined below) in European Research Infrastructures (RIs). Secondly, we examine some of the commonalities of approach in using impact measures as a factor in funding, managing the sustainability of RIs. By implication we want to promote the maturity of the Social and Cultural Innovation (SCI) community in this area. We've been capturing and using impact for our service delivery for some time and it's possible that others can learn. By implication we aim to try and turn the way in which we identify and understand the impact of our service into a mechanism by which our funders can use to continue their investment in the UK Data Service. This is not a one-sided approach: our funders have to justify to their funders and our host organisations have to see the value of our work. But the two principal rationales are for accountability – to demonstrate the wider value of our Research Infrastructure and secondly, for understanding – to understand the methods and routes by which our

---

This paper is an expanded and re-written version of a presentation given by Matthew Woollard at the workshop, «Stay tuned to the future, an international conference on the impact of research infrastructures for social sciences and humanities», Bologna, 24-25 January 2018.

RI leads to impact, and also develop better ways of delivering and communicating impact.[1]

From our long experience, we also believe that it is near to impossible to construct a broad set of impact measures which can be used across multiple RIs, since the specific missions and objectives obviate easy comparison. Below we show in detail some of the specific measures and outcomes which we can use to understand our impact, but we are clear that these are not all universally applicable.

The UK Data Service (the Service) acquires data from a wide range of data creators; national governments, researchers and international governmental and non-governmental organisations, third sector, local government and commercial organisations. We benefit researchers and data providers by curating their data and making them available to eligible researchers – broadly defined – not just in higher education. All that these organisations need to do is to provide the Service with data and sort out the legal side with us and we make it available for reuse – there's no real cost to the data owner. We ingest data into our preservation systems, ensuring that it can be used in the future. For some data we provide online interfaces for access, some data can just be downloaded, and some data must be used in a secure room or environment. We tailor the access mechanisms to suit the needs of the researchers and the owners of the data. Researchers access data freely – there's no cost to them and with more than 7,000 data collections we invest in metadata for discovery and harmonisation. Our question bank provides access to the full text of over 700,000 questions asked in surveys over the last 40 years or so. We version DOIs for all of our datasets, providing easy to use citations. Having more than 7,000 high quality data collections in one place, means serious researchers don't have to search for long. We're also one of the pioneers of secure access for data which has a risk of disclosure, data which can't be made open and which we hold in secure environments. We train researchers to use these systems and we manually check their outputs. We provide comprehensive user support and training for users of the data. All of our activities occur in order for researchers to do better research – thus our primary impact is to facilitate the impact of others.

The Service has been supporting researchers through its existence as an RI in its various forms for over 50 years and can be considered as 'impact-

---

[1] T. Penfield, M. J. Baker, R. Scoble, M. C. Wykes, *Assessment, evaluations, and definitions of research impact: A review*, «Research Evaluation», 23 (2014), n. 1, pp. 21-32, DOI: 10.1093/reseval/rvt021. See also: European Strategy Forum on Research Infrastructures Long-Term Sustainability Working Group, *Long-Term Sustainability of Research Infrastructures: ESFRI Scripta. Vol. 2*, Milan, Dipartimento di Fisica, Università degli Studi di Milano, 2017.

ful' throughout that period. We've had a formal focus on impact since receiving dedicated funding (in various forms) for this focus by the primary funder of the Service, the Economic and Social Research Council (ESRC), over the last six years.

Impact is a fairly recent mode for describing and understanding benefit in the context of funded research. We acknowledge methodological uncertainties about impact, noting that impact does not offer a fixed methodology or universal accord as to what it can measure or claim for itself. Here we begin however, to understand how the Service as an RI might align with current discourses of impact and also how those discourses may be deployed to help understand the specific beneficial effect of the Service and the use of the data and resources it makes available.

An overriding basis for understanding our impact is through, the Service's relative disinterest in and independence from the research process, predicated upon our mission (of the curation and archiving of economic and/or social data, broadly defined, for research and teaching):

> To provide an easy to use, trusted and innovative 'one-stop-shop' for suppliers and users of the extraordinary economic and social data resources available in the UK, following the highest standards for data management, access and training and support, across the data life-cycle.

Such positioning perhaps makes impact *measurement* more difficult, including efforts to identify the direct link between an underpinning research output, and the impact, where – as we are focused on – the data we make available were used in that output. We possibly have a more difficult task: we are unlikely to be able to claim impact from simply assuring a particular data collection is available for a piece of research which uses the data and goes on to have that impact. Linear concepts of impact don't perhaps work here. We imply a more embedded effect of the RI where «knowledge provides the concepts, data and tools that underpin our knowledge of social and policy problems».[2] Research processes and RIs may in this definition, be perceived as mutually constitutive to some degree. RIs in this case are implicit in both supporting others in knowledge production (providing data for research) but also in the production of knowledge (curation of *these* data) and any assertions of impact in this context are potentially performative as well as constructivist; so we are careful what we claim as impact and how. We discuss here the opportunity to consider and contribute to methodologies of impact in a particular context, that of data impact and of being an RI which provides those

---

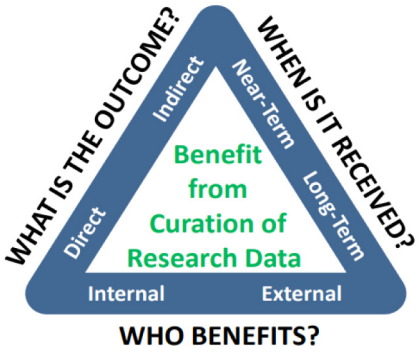
[2] C. Boswell and K. Smith, *Rethinking policy 'impact': Four models of research-policy relations*, «Palgrave Communications», 3 (2017), n. 44, DOI: 10.1057/s41599-017-0042-z.

data; we assess the possibility of identifying ways of understanding and claiming a specific beneficial effect from where and what we are as an RI.

We are starting with a reasonably conventional definition of the impact of the Service which makes sense from the point of view of the Service which we are involved in because it takes account of the dual benefits of costs avoided, and increased productivity and knowledge transfer in the research process. This leads to the construction of a definition of impact which focuses on the cumulative effect of the existence of RIs:

> Impact may be defined as a benefit accrued for the greater good (whether political, economic, socio-cultural, environmental, etc.), within any sphere, but most usually *beyond* the research community, in any way (direct and indirect) and at any time (the past, now, but mostly in the future).

The environment is essentially that which the well-known PESTLE analysis covers, with temporality and effect-type as additional factors.[3] This definition accords well with an earlier 'anatomy of a benefit' used in the Jisc[4] – funded Keeping Research Data Safe project.[5]



Anatomy of a benefit

Source: Introduction to the KRDS Benefits Analysis Toolkit[6]

---

[3] P for Political, E for Economic; S for Social; T for Technological; L for Legal and E for Environmental. The UK's Chartered Institute of Personnel and Development's website contains a helpful introduction to the PEST/PESTLE methodology. See: https://www.cipd.co.uk/knowledge/strategy/organisational-development/pestle-analysis-factsheet.

[4] Jisc is the UK's National Research and Education Network.

[5] N. Beagrie, B. Lavoie and M. Woollard, «Keeping Research Data Safe 2» (2010) [Available from http://repository.essex.ac.uk/2147/1/keepingresearchdatasafe2.pdf].

[6] Introduction to the KRDS Benefits Analysis Toolkit (2011). Available at: https://beagrie.com/static/resource/intro_benefits%20analysis%20toolkit_0711.pdf.

Thus we can construct a generic impact statement with these three elements as follows, adding in the various actors as well.

By acting as the licensee for data, researchers benefit from not having to negotiate licences with data owners and thus save time, cost and effort.

Impact as expressed here is direct and persistent. One form, perhaps the overriding form of impact is economic because it means researchers are not spending time and effort in negotiation, nor in data curation, management and infrastructure for these data when it could be more profitably spent in research: a longer term cultural benefit, since it places researchers in a position whereby they become reliant on a service/archive doing this for them. (Such reliance may have a corresponding long-term impact which is not positive.) If the Service were to close tomorrow, then each and every use made of data in our holdings would have to be renegotiated between the researcher and the data owner. Moreover, data owners would incur costs in separately managing not only data access infrastructure for researchers but also user support, training and output checking for data security. The strength of this negative impact could be reduced if all researchers were better aware of what their (now thankfully distant) predecessors had to do in order separately to access the data from the whole range of data creators: companies, government departments, other academics, etc.

Direct impact of the archive/service stems from our actions which *offer this benefit* to researchers and an indirect impact of our funders. They have the foresight to fund us, and make the decision that it is better to fund us to do *this*, rather than fund the researcher.

The impact statement is defined in such a way therefore, as it could act as a template for RIs to use more widely.

- The first part is the action carried out by the RI: 'by acting as the licensee for data'.
- The second is the beneficiary: 'researchers'.
- The third is the benefit: from not having to negotiate…
- And the fourth is the *impact type* in this case economic.

This is all very straightforward. But, it doesn't help us quantify any of this impact. We could say that we spend £x thousand per year on our negotiation activities, and that that is the amount saved; but this is of course the lowest end of any estimation. In this case let us imagine that our negotiation for *all researchers* has the same cost as a single researcher negotiating for their use of the data. If the data are used 200 times by researchers, then having us doing the negotiation is 200 times cheaper. Unfortunately, we

cannot test in any real way whether this is anything close to the real benefit. We can't assume that demand would remain the same if we hadn't negotiated for the data in the first place. Perhaps half of the 200 users would not have bothered to negotiate for the data; or put another way perhaps half of the 200 users are only inquisitive and don't actually want to use the data.

In this case, our ability to measure economic impact is hugely hampered, because the counter-factual makes the real demand for the data uncertain. Are researchers taking the data from us because we exist or because they really need to use it in their research? The big challenge from this simple question is different. It makes us wonder about the value of one of our Key Performance Indicators which is the number of registered users. Our funders assume that we are doing better if the number of our registered users increases. Unfortunately, this takes no account of data which has been downloaded, and it also takes no account of the use of the data. (And, by the way, we have something in the region of 26k registered users of the UK Data Service.) The theory is that if we increase the number of registered users, there is a likelihood that we are increasing the amount of good science which is being done, and (more speculatively) that there must be an increase (not concomitant) of the benefit of that research. However, because we have thought this through, the UK Data Service, no longer actively promotes itself to increase the number of registered users. All our promotion is focussing on keeping existing users registered, and maintaining and increasing our data holdings. At present, we are reasonably complacent in terms of numbers of registered users. There is a small upward trend over the last decade, but net increase is less than 10%. What our funders see every six months are two numbers, representing at two fixed points in time, the number of registered users. If the difference between these two numbers is an increase, then there's a pat on the back; if the difference is negative, then there's a reproachful comment – but that is it.

Part of our income is allocated to demonstrating the impact of our funder's investment, both in [applied] research (which we do not carry out) but supply part of the raw materials for and in our infrastructure itself. We recognise first and foremost that ascribing direct benefits of a data service is complicated by the primarily indirect nature of its impact. In the current scene and specifically for the economic and social sciences, the impact of research on policy is generally considered the 'purest' form of impact. And of course, economic and social science data archives can only *indirectly* affect the work of researchers.

Approach

Accepting that the Service operates to bring data to researchers, saves them cost, time and supports them to use the data, we now consider how to structure a framework for data impact which aims to understand that usage from its appearance in research, to methodologies for demonstrating (and 'claiming') this impact. As we are interested in the impact of the data underpinning research (as opposed to the research itself); in mechanisms for understanding the effect, rather than simply that data were used in an output, that is – as it *appears* in research, policy, debate or the evidential process (although important), our focus is on where (ideally) cited data can be tracked through the specific beneficial outcome and on to an evidenced effect, corroborated by the end user. This is not easy.

In ideal terms, if a social or economic benefit is realised or if one person's life is changed for the better as a result of the use of Service data, resources, expertise, or from the policies and activities of the Service as a data infrastructure, it is important for the Service to aim to understand how, jointly with partners to share that benefit and use it to strengthen and expand its impact.

The Service takes an approach of 'with, not for' to supporting data impact, ensuring impact activity is focused on data user, partner, funder and policy-maker defined concepts of impact.

In addition to the definition of impact of particular relevance to the Service and RIs discussed above (p. 45), the Service's focus on impact is aligned with that of our primary funder, the ESRC (Economic and Social Research Council), which defines impact as:

> the demonstrable contribution that excellent research makes to society and the economy and can involve academic impact, economic and societal impact or both:
>
> – Economic and societal impact is the demonstrable contribution that excellent social and economic research makes to society and the economy, and its benefits to individuals, organisations and/or nations;
>
> – Academic impact is the demonstrable contribution that excellent social and economic research makes in expanding understanding and advancing scientific, method, theory and application across and within disciplines.[7]

The ESRC considers that academic, economic and social research impact can form:

---

[7] What is impact? <https://esrc.ukri.org/research/impact-toolkit/what-is-impact/> accessed 21 June 2018.

Instrumental impact	Influencing the development of policy, practice or service provision, shaping legislation, altering behaviour
Conceptual impact	Contributing to the understanding of policy issues, reframing debates
Capacity building	Through technical and personal skill development

For the Service impact is conceptualised as supporting the impact of others using Service-curated and hosted data as well as the impact of the Service as an RI. Our focus is on coordinating data impact activity through the concept of data impact by:

- Expanding methodologies of data impact
- Capitalising on the role of the Service as a critical part of the UK's research infrastructure, internationally
- Contributing to processes of developing social benefit through supporting the re-use of Service data, where (ideally) cited data can be tracked through the specific beneficial outcome and on to an evidenced effect, corroborated by the end user; and
- Demonstrating data impact leadership.

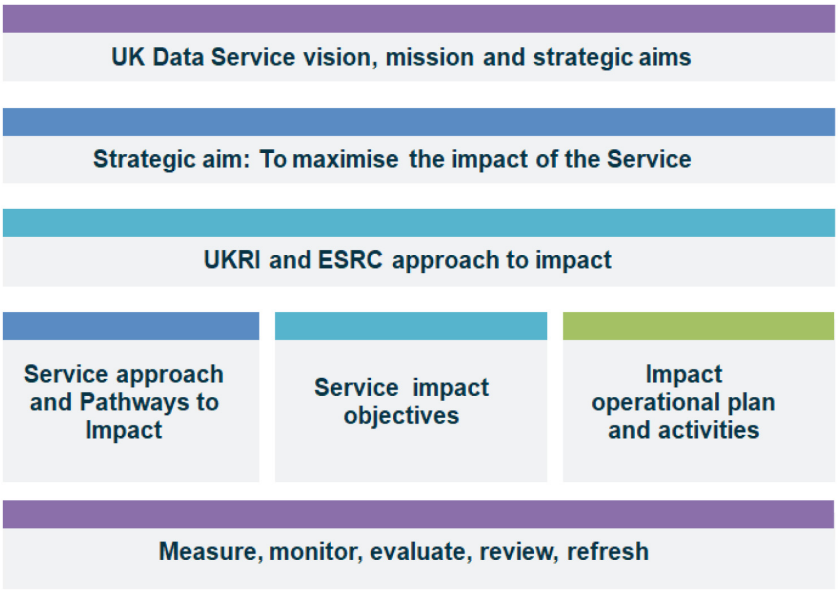
Our focus is therefore on the demonstrable contribution the Service and its data and resources make to the economy, society, culture, public policy and services, health, the environment and quality of life. Our emphasis is on drawing together evidence about the reach and significance of the impact of the use of the data and resources, of the Service as a whole. We frame impact in terms of the Service's strategic approach so that it can be understood in terms of the Service's strategic aims and support the Service in the achievement of its vision:

To support high quality social and economic research, teaching and learning through assuring long-term access to quality economic and social data, supporting and promoting their use, value and impact.

What this means in effect it that our impact measurements are mostly non-quantitative and almost all directly related to our overall strategy. The figure below shows how we frame impact in terms of our overall strategy.

Our Pathways to Impact, a requirement of UK research funding, were developed as part of the impact strategy in the funding period 2012-2017. The Pathways are the prospective approach to impact, designed at the bidding stage, the impact objectives are designed to structure how it is proposed to achieve and demonstrate impact at various stages throughout the grant, in fulfilment of the Pathways. The objectives present detail and context to guide

activity and may change over the phase of the grant as they are achieved; as the impact develops; or as impact priorities and opportunities change.



Thus our Pathways were to:

- Support development of impact derived from research which uses our data
- Engage more non-academic organisations and communities in using our data and services
- Articulate the role of the Service in terms of its contribution to societal benefit, developing the impact of the Service as a whole in the context of that benefit and having an impact on other data services and data infrastructure internationally
- Establish methods for expanding innovation and collaboration in using our data, resources and expertise
- Formalise the evidencing and corroboration of our impact
- Promote our impact through targeted communications activities in an engagement, collaboration and co-creation framework
- Promote impact capacity building

We have developed impact objectives to structure how we achieve and demonstrate the achievement of the Pathways to Impact, and against which we plan annual activities and initiatives. The objectives are to:

1. Derive impact from the Service as a research infrastructure, its assets, resources and expertise, and embed impact across activity

2. Demonstrate data impact leadership and innovation, expanding methodologies of data impact
3. Maximise the use of Service data, resources and expertise by new and non-academic users
4. Increase collaboration between the research community and non-academic organisations and communities, acting as a 'data facilitator'
5. Align Service data and resources to supporting partner priorities focused on addressing societal challenges
6. Engage researchers in ensuring that and understanding how their research using Service data and resources has impact
7. Expand upon and remain creative in how the Service develops, celebrates and promotes its own and others' impact
8. Understand and identify the potential for inter/multidisciplinary approaches to contributing to data impact
9. Increase data impact through increased data citation
10. Develop systems and processes for tracking and evidencing the impact of the Service as a whole
11. Generate more, and more authentic, corroboration
12. Maximise income to the Service from its impact activities.

How to collect information efficiently

Our focus is on understanding and leveraging data sources which already exist and can be repurposed for impact evaluation. We collect information about our impact from the perspective of the role of the Service as a critical part of the UK's research infrastructure, structuring how we conceptualise that impact, through researching and collating external (and internal) activity we are able to demonstrate the impact the Service has through both usage of the data by others and through the categories which define our area of operation as follows:

RI impact
Offering trusted digital repository status
Setting standards in data infrastructure
Providing a unified collection once, for all
Offering efficiency and value for money
Promoting data reuse
Enhancing research capacity
Promoting research data ethics and integrity

Leading data management/policy consultancy, nationally and internationally
Assuring long-range access
Offering support and training in data use, a skills focus
Offering expertise in data management
Stable, enduring and innovative, a Service approach

Examples include:

– *Leading data management/policy consultancy, nationally and internationally*: Matthew Woollard, Director and Louise Corti, Functional Director for Collections Development and Producer Relations at the UK Data Service provide an update on how the UK Data Service has supported the development of the new Indian Council for Social Science Research (ICSSR) Data Service:
http://blog.ukdataservice.ac.uk/welcoming-progress-on-the-new-indian-councl-for-social-science-research-iccsr-data-service/

– *Enhancing research capacity*: Rob Dymond-Green, Technical Manager for the UK Data Service Census Support Service, describes the creation and processing of the 2011 UK Census dataset:
http://blog.ukdataservice.ac.uk/creating-a-unified-2011-census-dataset-for-the-four-nations-of-the-uk/

We plan to establish Service impact champions from among our staff where we 'triage' activities identified by Service colleagues and the external organisations they engage offering potential impact through the categories. Some activities identified by the impact champions will assist in our communication and marketing activities, some will indicate the need for a more impact-focused approach.

We have created a range of dedicated impact channels to develop and enhance our impact activity and its profile. These channels also enable us to promote the Service's and others' impact in a range of ways, helping achieve our impact objectives in the areas of communicating, promoting and building impact capacity through sharing its outputs widely. The channels themselves could also be considered as constitutive of impact because they provide a series of outlets for the development and consideration of methodologies for it:

Impact web pages	We have developed dedicated impact web pages where we bring together the elements of the impact activity which endure over the period of the strategy, such as Impact Case Studies, the #DataImpactFellows, Blog, Lab and #DataImpact events.
Impact case studies	Our *c.* 200 case studies focus on users of the data, including an opportunity for them to describe their impact or findings for policy. The case studies also cover use of the data in teaching. At present we are also focussing on developing case studies with a focus on early career researchers and the Ph.D journey.
Data Impact Blog	The blog is a hub for researchers, students, communities, policy-makers, government and anyone interested in maximising the impact of economic, social and population and data in teaching, research and policy. The blog is where we encourage debate about data impact, • share best practice in data impact; and • keep the data impact community up to date with news, events and the latest data-driven impactful research and policy making.
Impact and Innovation Lab	Through the development of the lab we enhance our impact through working more closely with innovators developing inspiring data solutions to social challenges. The lab focuses on methods and technology.
#DataImpactFellows	The #DataImpactFellows programmes aims to establish additional ways to support the long-range use of its data and resources by new generations of scholars, extending this usage through the research partnerships they develop and by the students they teach – from the earliest stages of, and throughout their career. The programme aims to provide career development opportunities for scholars at a relatively early stage of their academic careers with a proven record of research which has a dedicated focus on impact and includes engagement beyond academia.
#DataImpact events	Annual/biennial event where panels comprising leading data innovators explore data re-use in policy and research, sharing their experiences of demonstrating data enhanced impact and focusing on defining collectively what the data solvable policy challenges are across the higher education, public, commercial and civil society sectors.

ScoopIt	Our ScoopIt enables us to curate data use as it appears across the media. The ScoopIt is an important mechanism for gaining broad insight into the 'appearance' of the collection in terms of its indication of the concerns of the time as both research and media focus, as well as demonstrating evidence of reach and significance through the data's visibility in media outlets. Our analysis of ScoopIt shows over 3,000 articles about research which uses the data in the collection, at around 1,000 per year (and is only indicative).
Impact focused Twitter	Here we promote elements of the impact programme and involve ourselves in impact as an emerging empirical approach.
Our youtube channel	Here we include videos of the data impact events and initiatives such as depositor stories where data creators corroborate the impact of the Service from their perspective.
Wakelet	Our Wakelet is where we curate social media from conferences and events from across the Service, promoting our impact.
Google Analytics	Which we use for demonstrating reach in terms of international coverage and also for giving a sense of how impact channels are performing.

How do we support impact development more directly?

Objectives 4 and 5 of our impact strategy offer us the opportunity to facilitate impact development in partnership with others:

- Increase collaboration between the research community and non-academic organisations and communities, acting as a 'data facilitator'
- Align Service data and resources to supporting partner priorities focused on addressing societal challenges

Our dual role of acting as a data facilitator and supporting partner priorities focused on addressing societal challenges has enabled us to co-ordinate specifically impact focused activities.

Examples include:

- We have held two open data dives, attended by academic, government, and commercial research professionals, where we built impactful applications from open data in the collection together with external open data. The data dives took the form of competitions with the prises being 3-D prints of the winning dives. One was of UK house prices using the Service's Census data and the other was themed as a SWOT analysis of Greater Manchester and the Region

– internationally as a gift for the new mayor of Greater Manchester with the winning teams creating maps of deprivation and community resilience, prescriptions and income inequality.
– We have proposed an additional emphasis on moving the impact continuum forwards, to include a greater focus on orienting our impact practice towards addressing intractable social challenges from the perspective of those organisations and agencies charged with addressing them. We have identified a need for ensuring that impact practice is targeted towards removing organisational barriers to developing social benefit. To that end we have established an impact ambassadors programme where we support public, commercial, voluntary and community sector representatives and data users from among the academic community jointly to understand and support their data needs, and coordinate a programme of activity focused on deriving impact through identifying and 'data solvable' social challenges defined by partners and supported by a range of academic data experts.

How we understand impact where data in the collection are used by others

We are in a good position to collect and demonstrate information about the impact of research which uses the data we make available. We understand usage of much of the data in the collection through our user registrations and are able to trace publications post usage. Moreover, we have recently required citation as a qualification for releasing outputs after disclosure checking and can follow up to ascertain the impact which we then can work with the researcher to create a case study.

The metric tide a report on the independent review of the role of metrics in research assessment and management, chaired by Professor James Wilsdon recommends that:

> The use of digital object identifiers (DOIs) should be extended to cover all research outputs. This should include all outputs submitted to a future REF[8] for which DOIs are suitable, and DOIs should also be more widely adopted in internal HEI and research funder processes. DOIs already predominate in the journal publishing sphere – they should be extended to cover other outputs where no identifier system exists, such as book chapters and datasets.[9]

---

[8] The REF is the UK's Research Excellence Framework which is the mechanism where the UK government assesses the quality and impact of research doing by Higher Education Institutions and distributes funding to those institutions. See http:www.ref.ac.uk.

[9] J. Wilsdon *et al.*, *The metric tide: Report of the independent review of the role of metrics in research assessment and measurement* (2015), DOI: 10.13140/RG.2.1.4929.1363.

The citation of research data (and metadata) can support the understanding and promotion of research impact through the tracking of the use of data in research and on into policy and product development, influencing decisions about public and commercial spending and service provision.

Citing research data isn't new; the Service and other data repositories around the world have been requiring it as part of their standard user agreement for many years. Citing data using persistent identifiers (such as DOIs) supporting verification and attribution or research, helps people to understand the impact of the research and offers the realisation and demonstration of efficiencies through re-use. A DOI is automatically assigned to any data collection deposited into the UK Data Service.

We have developed a #CiteTheData campaign which we are proposing to widen through engagement with data providers and journal publishers. We also have the opportunity to support data citation through researcher profile platforms such as ORCID.[10]

Mining publicly-available case-studies

We pilot data-mined the API (Application Programming Interface) that the Higher Education Funding Council of England (Research England) made available on its Research Excellence Frameworks (REF) Impact Case Study website.[11] Citation of data, even the appearance of data, was not mandated in impact case studies for the 2014 REF, so to understand the usage of data in the collection in REF impact case studies we needed to pre-define a set of scripts for mining; we started with high usage terms such as «Labour Force Survey», «Crime Survey for England and Wales». The pilot, focusing on data named in the case study database found 60 or so impact case studies which clearly used data in the UK Data Service collection (from the search terms we used) to support their development (or the underpinning research).

As previously considered however, we are careful what we claim as impact and how. The table below shows some of the ways in which data were a feature in research across a range of institutions and Units of Assessment in the REF. Further work includes reviewing specifically how the data were used in the case studies. Initial indications are that use of the data provid-

---

[10] http://blog.ukdataservice.ac.uk/citethedata-impact-tracking-and-the-metrics-debate/.

[11] http://impact.ref.ac.uk/CaseStudies/.

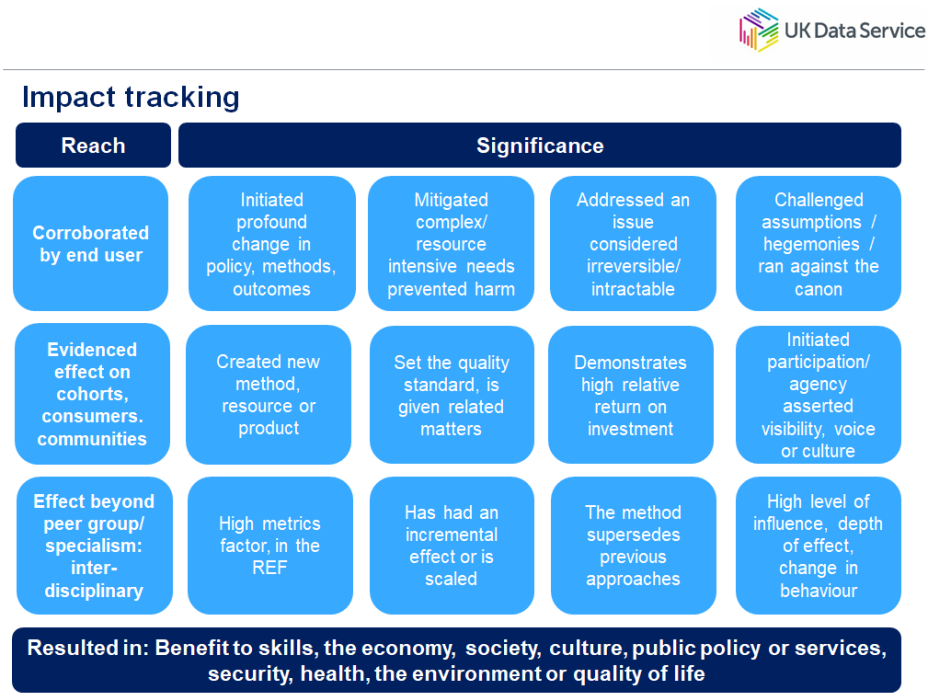
ed evidence of improvement or detriment to people's lives, supporting research processes with the evidence of change.

Assessment Unit	Institution	Case Study Title	Data
Business/ Management	University of York	Employee ownership plans: individual behaviour and company outcomes	Wealth and Assets Survey
Business/ Management	City University London	A fairer approach to compensation for personal injury and fatal accident cases	Labour Force Survey
Social Work/ Social Policy	Nottingham Trent University	Crime drop, security and victimisation	Crime Survey, Census
Economics/ Econometrics	University of Kent	Improving the Economic Role of State Education in Britain: Lessons from the Independent Education Sector	Labour Force Survey
Sociology	University of Surrey	Improving the quality of life for citizens in the UK through shaping the organisation and practice of policing	Crime Survey for England and Wales
Economics/ Econometrics	University College London	Setting national minimum wages	Labour Force Survey
Public Health, Health Services and Primary Care	The University of Oxford	Shaping international and UK tobacco policy and practice	General Lifestyle Survey
Geography, Environmental Studies and Archaeology	University of Portsmouth	The Smoking Epidemic in England and Scotland: Shaping Public Health Policy and Planning	Census & Health Survey for England
Law	Birkbeck, University of London	Trust in justice: mapping public attitudes towards the police and other legal institutions and how these findings have changed EU and UK policy	Crime Survey for England and Wales (European Social Survey)
Education	Institute of Education	University fees and social mobility: a difficult balancing act	Labour Force Survey

Mandating data citation in future REF impact case studies – as part of a broader programme to raise the value imperative of quality, cited data as output – in the wider REF would be a further step towards understanding the impact of the Service and of value for research assessment frameworks for other RIs.

How do we turn a measurement on service activity into an impact description?

Our aim is to structure a framework for data impact which aims to develop methodologies for demonstrating (and 'claiming') impact for RIs. And given as discussed, that we are interested mechanisms for understanding the effect, rather than simply that data were used in an output and where (ideally) cited data can be tracked through the specific beneficial outcome and on to an evidenced effect, corroborated by the end user, we have developed a mechanism for understanding the specific beneficial outcome. From analysis of impact case studies developed as part of the UK's REF exercise in 2014 the following figure offers ways of understanding impact by type – that is through that specific beneficial effect:



Once it is possible to be fairly satisfied that the data were instrumental in the research process from which the impact is identified, we assess typologies of impact as described in order better to understand the reach and significance of the data in the impact. (The figure distils elements of what were concluded to be in that exercise 4* impact from the REF pilot impact case studies exercise designed by (then) HEFCE to structure impact in the exercise.)

We review the impact of the use of data and map it against the elements of the table. Similarly; not easy. Indeed, corroboration; the 'gold standard' of impact evidence and the most elusive, may be understood as further removed from linking data to impact. Other methodologies are proposed here, including a focus on elucidating the importance of data in supporting the understanding of lived experience over time with a focus on the pressures and changes in people's lives. Of benefit, may be the coordination of the development of aggregated, anonymised personas in association with partners, so that it is possible better to understand data impact from the perspective of the data subject. Such an approach may offer a more 'Uber' demonstration of the impact of the data as a consolidated asset of some significance to the social landscape and what it offers as a whole to policy and economics focused initiatives focused on social benefit. Indeed, impact from that perspective offers a focus on methods for understanding the way in which the available data frames the available hypotheses or even structures them, and that without which, much of the research which depends on the data may not be possible.

In this paper we have presented the ways in which we measure our impact both as an RI and through the use of the data and resources we make available. We hope that the paper makes a contribution to emerging frameworks for data impact and will continue to develop and enhance the activity towards understanding the impact of RIs as described here through our broad methodology of:

- Demonstrating data impact leadership
- Expanding methodologies of data impact
- Capitalising on the role of the Service as a critical part of the UK's research infrastructure, internationally; and
- Contributing to processes of developing societal benefit through supporting the re-use of Service data, where (ideally) cited data can be tracked through the specific beneficial outcome and on to an evidenced effect, corroborated by the end user.

What we have not presented here are easy to digest indicators for impact. The reason is that we do not believe that many proposed impact measures

are relevant to RIs like ours and we do not believe that there is a broad set of impact measures which can be used across multiple RIs, since the specific missions and objectives obviate easy comparison. If RIs are asked to report on impact measures which are meaningless or irrelevant, then funders must be expected to receive irrelevant or meaningless information for the subject. Consequently, we have concentrated on trying to identify impact and tracking back to understand our involvement. This is not a perfect methodology, but it is one from which we can get some idea of the real benefits the UK Data Service provides, and it is not based on spurious and unreliable metrics which are constructed from pertinent and reliable performance indicators.

We believe, like the authors of the recent draft *Reference framework for assessing the socio-economic impact of research infrastructures* that it is harder for RIs to assess impact because much of that impact is indirect. We also believe that the uniqueness of RIs means that comparing the socio-economic impact of different RIs is not recommended.[12] On the other hand we know that funding organisations need to make decisions on continued funding for RIs, and that impact is a major feature of this assessment process. Therefore, the Service has opted to maximise the opportunity for our funders to see for themselves through reporting as well as a variety of communication channels how we have had an impact.

[12] OECD Global Science Forum. Expert Group meeting on: «Reference framework for assessing the socio-economic impact of research infrastructures». *Draft Assessment Framework. Initial Consensus Indicators*. Paper dated 13 March 2018. [Available at: https://www.innovationpolicyplatform.org/system/files/Draft%20OECD%20GSF%20SEIRI%20framework%20v16.pdf].