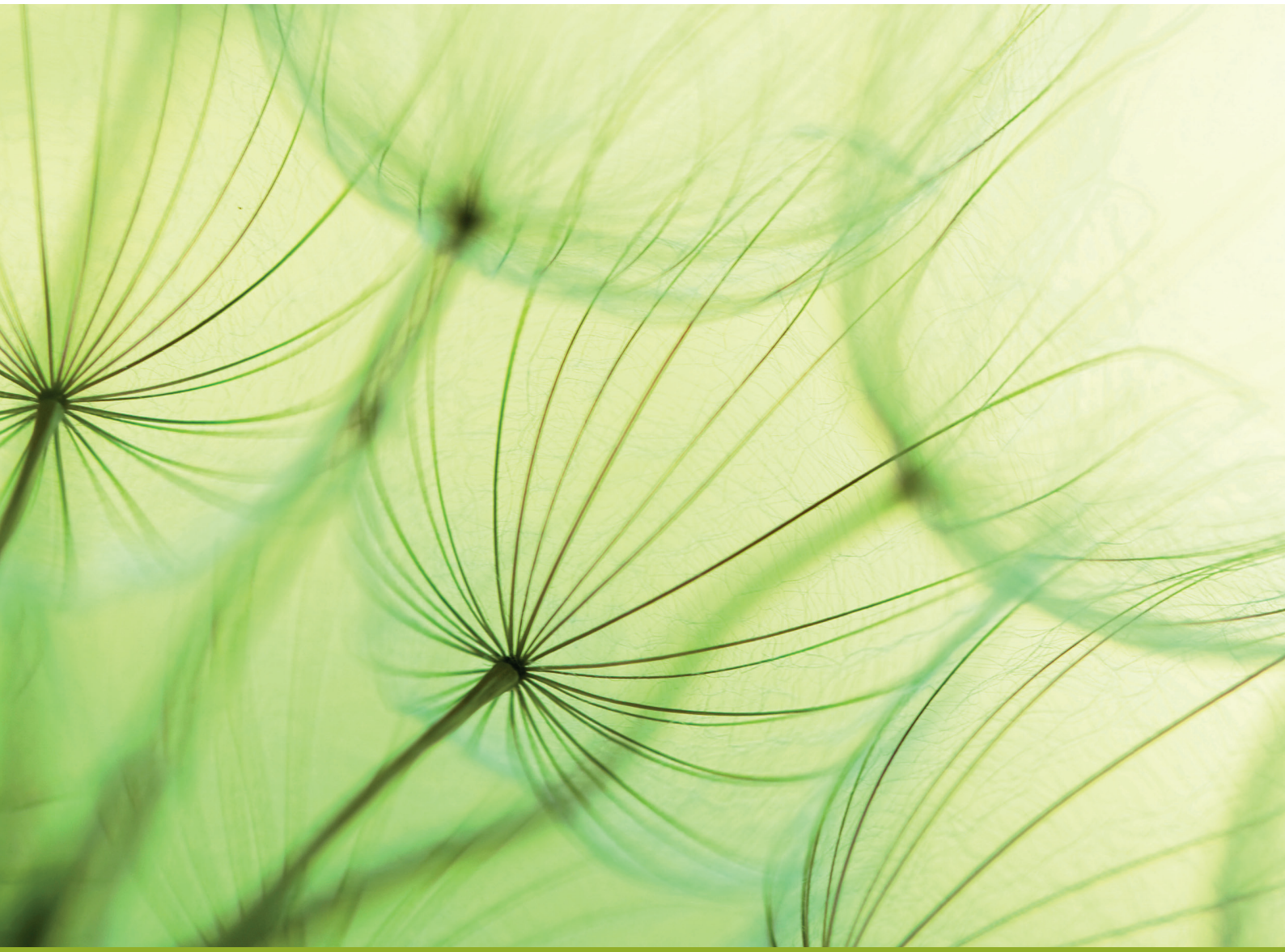
Knowledge Exchange



Sowing the seed: Incentives and motivations for sharing research data, a researcher's perspective

DFG Jisc CSC DEff SURF

“Incentives and motivations for sharing research data: researcher’s perspectives”

Veerle Van den Eynden and Libby Bishop
UK Data Archive, University of Essex

Please cite this document as:
Van den Eynden, V. and Bishop, L. (2014). Incentives and motivations for sharing research data, a researcher’s perspective.

A Knowledge Exchange Report, available from
knowledge-exchange.info/Default.aspx?ID=733

Acknowledgements

This project has been initiated and steered by a group of experts within the Knowledge Exchange research data domain. Knowledge Exchange wishes to thank the following experts for their contributions and comments, and the sharing of valuable insights:

- **Jens Nieschulze,** Göttingen University
- **Stefan Winkler-Nees**, DFG
- **Angela Holzer**, DFG
- **Christian Ertmann**, Royal Library of Denmark
- **Tanja Kantola**, CSC
- **Damien Lecarpentier**, CSC
- **Leena Jukka**, CSC
- **Irina Kupiainen**, CSC
- **Jeroen Rombouts**, 3TU Datacentrum
- **Ingrid Dillo**, DANS
- **Magchiel Bijsterbosch**, SURF
- **Rachel Bruce**, Jisc
- **Matthew Dovey**, Jisc
- **Ari Lukkarinen**, CSC
- **David Kernohan**, Jisc
- **Matthias Katerbow**, DFG

Published by Knowledge Exchange, November 2014

Executive Summary

This study, commissioned by Knowledge Exchange, has gathered evidence, examples and opinions on current and future incentives for research data sharing from the researcher's point of view, in order to provide recommendations for policy and practice development on how best to incentivise data access and reuse.

Whilst most researchers appreciate the benefits of sharing research data, on an individual basis they may be reluctant to share their own data. This study is based on qualitative interviews with 22 selected researchers of five research teams that have established data sharing cultures, in the partner countries of Knowledge Exchange: Denmark, Finland, Germany, the Netherlands and the United Kingdom. The five case studies span various academic disciplines: arts and humanities, social sciences, biomedicine, chemistry and biology. The research groups or projects are:

- **Retired men gathering in cities** (fsd.uta.fi/en/data/catalogue/FSD2842/meF2842e.html), Finland
- **LARM Audio Research Archive** (larm.blogs.ku.dk/), Denmark
- **Netherlands Bioinformatics Center** (nbic.nl/), Netherlands
- **Evolutionary Plant Solutions to Ecological Challenges** (ruhr-uni-bochum.de/dfg-spp1529/Seiten/index.html), Germany
- **Chemistry Department, University of Southampton** (southampton.ac.uk/chemistry/), United Kingdom

When researchers talk about 'data sharing' they frequently mean a variety of different ways in which research data are exchanged with other researchers. Six different modes of data sharing can be recognised: private management sharing, collaborative sharing, peer exchange, sharing for transparent governance, community sharing and public sharing. These modes vary by factors such as whether data are shared within a trusted group or with strangers and whether the primary purpose of sharing is transparency or further research.

Important motivations for researchers to share research data are (1) when data sharing is an essential part of the research process; (2) direct career benefits derived from sharing through greater visibility of one's work, reciprocal data exchanges, and the reassurance of having one's data recognised as valuable by others; (3) the norms that researchers are exposed to within their research circle or discipline; and (4) a framework of funder and publisher expectations, policies, infrastructure and data services as external drivers. While discipline norms were important, equally striking was the variation found within disciplines. Incentives to share also vary across a researcher's career trajectory. The strong influence of data sharing norms implies a key role for early training on research data sharing, e.g. as an integral part of research methods training.

Researchers' experiences, data sharing practices and motivations are shown to be heterogeneous across the studied research groups and disciplines. Patterns of data sharing and related incentives cross the disciplinary boundaries.

A strong emerging theme is the need to create a level playing field for all researchers to share data and change the collective attitude towards sharing. Data sharing training embedded in research methods training is crucial for data sharing to become standard research practice. There is equally a vital role to be played here by formal data policies as long as they do not interfere with informal sharing, and are sensitive to variations across disciplines. Policies operate through at least two channels. They provide a collective voice and help to clarify and change the norms of the research community. Policies also alleviate a mismatch of incentives: they create a positive motivation to share that benefits science in general, even in those cases where the direct benefits to individual researchers are weaker.

Based on the views and perceptions expressed by the interviewed researchers as found through the study, combined with the investigators' expertise, recommendations for incentivising increased data sharing are made for several stakeholder groups. Different stakeholders may opt to focus on incentivising different modes of data sharing. While all sharing modes are valid and serve particular purposes, transparency is needed to ensure fair access to data for all researchers.

Recommendations for research funders:

- All research funders to adopt a data sharing policy that clearly indicates expectations for data accessibility, in order to provide a level playing field with regard to data sharing for all funded researchers. Policies can consider measures such as requirements for data management planning and clear guidance on how a percentage of grants budgets can be allocated to data management for projects creating data with high potential reuse value

- Provide funding and support services to researchers where needed, e.g. for data documentation, annotation and data deposit. This should be similar to the funding of publication costs. Not all research disciplines have the same needs in this respect, as some data require more preparation than others to make them available for reuse. Also the type of data sharing influences what is needed, e.g. sharing raw vs. processed data; sharing data that supplements articles vs. sharing in repositories

- Focus data sharing funding towards two key intervention points:
  - early when research is being planned
  - upon completion of a research project, to prepare data and documentation for curation

- Continue to invest in data infrastructure that also provides rich context, detailed metadata and even a narrative account of the data creation. The kind of infrastructure researchers find most useful is where research data, papers and other outputs or resources are jointly available within a single data resource. Examples noted in this study are PubMed, TAIR, LARM, CCDC and ChemSpider

- Data sharing training embedded into research methods training for students and doctoral researchers, to help establish data sharing as standard research methodology and practice

- Promote reuse of existing data resources via specific funding streams for secondary analysis and by setting expectations for research grant applicants to justify the need to create new data in research (i.e. to demonstrate that existing data cannot address their research questions)

- Engage with publishers and commercial partners on IP and copyright of data that may limit data sharing by creating a working group to find ways to protect IP and share data, especially when research is intended for non-commercial use

- Provide guidance to peer reviewers to evaluate data sharing plans and strategies in research proposals

Recommendations for learned societies:

- Promote discussion of formal research recognition for data sharing and data publishing
- Set clear data sharing expectations for respective research disciplines, e.g. through code of conduct or best practice code
- Promote the development and uptake of data sharing agreements within specific research disciplines, that stipulate agreement over how research data can be shared in a timely and open manner (e.g. similar to the 1996 Bermuda Principles and 2003 Fort Lauderdale agreement over prepublication data release in genomics), to provide a level playing field to all researchers
- Promote the development of data sharing resources and standards for the research discipline

Recommendations for research institutions:

- Formally recognise and value data, alongside publications, as part of research assessment and career advancement
- Incorporate data impact into PhD career assessment, e.g. via a system of portfolio assessment where research data may be one element alongside other research outputs that provide evidence for research impact, or via a data CV
- Provide training in research data sharing to students, embedded into methods training, so data sharing becomes part of standard research practice
- Set expectations for data sharing for researchers within the institution
- Provide integrated support services to researchers, e.g. a one-stop-shop for all research data management and sharing guidance

Recommendations for publishers:

- Strengthen direct career benefits to researchers to share their data via data citation and data sharing metrics. This should provide what researchers in particular disciplines call for, e.g. micro-citation, micro-publications, data publishing with Digital Object Identifiers (DOIs), data citations to link with ORCIDs (Open Researcher and Contributor ID), and digital watermarking of data files to provide provenance of data
- Ensure that publishing terms and agreements of manuscripts do not create disincentives for micro-publishing of data, e.g. through overly restrictive requirements for manuscript content to be new
- Journals and innovative publishers to explore and actively encourage publication of negative findings, failed experiments, etc.
- Request that all data related to a published manuscript are made available, not only the data supporting the published results
- Set open or preservation standards for data formats, file formats, and supplemental documentation
- Ensure that data is fully and properly cited in all publications, and provide clear instructions to editorial staff and reviewers to check for correct data citation
- Make all supplementary data available openly (free of charge), even if the article is not

Recommendations for data centres and repositories:

- Develop and encourage pull factors for data sharing such as actively inviting researchers to share data by deposit or by other channels. Researchers feel valued and reassured about data quality when their research data are in demand
- Deliver specialist data sharing training for researchers on IP, copyright, technical standards and metadata
- Develop and provide flexible systems of providing access to data, allowing data owners to set controls where this is needed, e.g. embargo periods, defined access groups, etc.
- Provide data resources that combine data and related rich context such as publications and other outputs

Recommendations for Knowledge Exchange:

- Invest and engage in the development of data infrastructure with rich context or invest in infrastructure within the member's remit that underpins and allows these other infrastructures to flourish. The kind of infrastructure researchers find most useful is where research data, publications and other outputs or resources are jointly available within a single data resource. Examples noted in this study are PubMed, TAIR, LARM, CCDC and ChemSpider
- Explore and develop mechanisms for micro- and nano-citation of research data
- Lead the development of data sharing strategies and data sharing expectations at a national level, in collaboration with the various stakeholders
- Push for development of data infrastructures and data services at a European level
- Push for the recognition of data sharing in career progression at a national and European level
- Call upon relevant stakeholders to provide data sharing training to all undergraduate students
- Fund practical solutions for data sharing, possibly jointly with research funders [see recommendations for funders]
- Develop national registries of research data that link to journals, repositories, etc.

Contents

1. Background

This study was commissioned and funded by Knowledge Exchange, a cooperative effort that supports the use and development of Information and Communications Technologies infrastructure for higher education and research.

Knowledge Exchange (KE) is formed of five partners in five European countries:

- **CSC** (csc.fi/english), IT Center for Science in Finland
- **Denmark's Electronic Research Library (DEFF)** (deff.dk) in Denmark
- **German Research Foundation (DFG)** (dfg.de) in Germany
- **Jisc** (jisc.ac.uk) in the United Kingdom
- **SURF** (surf.nl) in the Netherlands.

The central benefit of research data sharing is well-recognised amongst researchers and society in general: to enhance and accelerate scientific progress for the benefit of science and society (RECODE, 2013; Tenopir et al., 2011). The last decade has seen rapid growth in the policy drivers of data sharing in the form of data sharing mandates by research funders, data policies by publishers and government transparency initiatives. Human and material capability has also developed exponentially, through the creation of increasing numbers of generic, community and institutional data repositories, and an increased focus on training and capacity building for researchers and supporting services.

The importance of increased data sharing and (open) access to research data has been heavily promoted by international and national entities. The OECD, in its Principles and Guidelines for Access to Research Data from Public Funding (OECD, 2007), declares that publicly funded research data are a public good that should be made openly available with as few restrictions as possible in a timely and responsible manner without harming intellectual property. The Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Berlin Declaration, 2003) calls for the promotion of knowledge dissemination via the worldwide web as a sustainable, interactive, transparent and openly accessible global information network. The report of the High Level Expert Group on Scientific Data notes the need for a European scientific e-infrastructure that supports seamless access, use, reuse and trust in data in order for open infrastructure, open culture and open content to go hand-in-hand (European Commission, 2010). The Royal Society's report, Science as an Open Enterprise, states that publishing data in a reusable form to support research findings should be mandatory; stronger still, that not sharing research data should be considered bad science (The Royal Society, 2012). The adoption of common data standards is seen as crucial, as are mandates for data publishing. A recent G8 science ministers statement calls for open scientific research data that are easily discoverable, accessible, assessable, intelligible, useable and wherever possible interoperable to specific quality standards (G8 UK, 2014).

1.1. Barriers and enablers for sharing research data

Whilst funder policies and high profile statements and recommendations may accelerate data sharing, individual researchers, though recognising the wider benefits of data sharing for scientific progress, often remain reluctant to share their own research data with other researchers.

Both detailed qualitative studies and wider surveys assessing data sharing practices, barriers and enablers amongst researchers (Borgman, 2012; Piwowar, 2011; RECODE, 2013; Savage and Vickers, 2009; Sayogo and Pardo, 2013; Tenopir et al., 2011; Wicherts et al., 2006; Youngseek and Stanton, 2012) - some of which focus on specific research disciplines, others look across a range of disciplines – identify numerous perceived or real barriers to data sharing:

- fear of competition, of being scooped and therefore reduced publication opportunities
- cost in both time and money to prepare data and documentation for sharing and absence of funding to do so
- absence of professional rewards for data sharing
- lack of standards and data infrastructure
- ethical and legal constraints

Enablers of data sharing generally reported in literature are:

- data sharing expectations of funders and journals
- peer expectations and sharing practices in the research community
- availability of data repositories and standards
- desire to showcase data quality
- researchers' data management skills
- organizational support
- acknowledgement received for data sharing
- data publication and metrics

Research also reveals strong disciplinary differences in data sharing practices, as well as more granular differences between research groups, because in different research disciplines, data may be perceived as different things or play a different role in the research, and come in varying formats and sizes. Youngseek and Stanton (2012) found data sharing being seen as critical for new science by biological, chemical and ecological scientists, whilst computer scientists, engineers, mathematicians and physical oncologists stated the opposite. RECODE (2013) in exploring attitudes towards open data found data sharing to be limited in particle (astro)physics (despite the collaborative nature of research), in health and clinical research (due to ethical constraints), and in archaeology; and prevalent in bioengineering and in environmental research. Significant variations in attitudes and practices were found within each discipline as well (RECODE, 2014). The Digital Curation Centre's SCARP project, investigating researchers' attitudes and approaches towards data deposit, sharing, reuse, curation and preservation over a range of research fields in different disciplines, found that the diversity of data types, working methods, curation practices and skills found within domains means that data sharing requirements should be defined at the finer-grained level, such as the research group (Lyon, et al. 2010).

1.2. Motivations and incentives for sharing research data

Fewer studies have focussed specifically on motivations and incentives for researchers to share data. A recent report by the Expert Advisory Group on Data Access in the UK on incentives for data sharing, based on interviews with a small number of key stakeholders (research funders, senior academic managers, postdoctoral researchers, a chair of a Research Excellence Framework panel and a senior data manager) and a web survey with researchers and data managers, finds similar barriers and restraints as those known. It recommends as essential incentives, that research funders should: strengthen and finance data management and sharing planning requirements; continue funding and development of infrastructure and support services; recognise high quality datasets as valued research outputs in the Research Excellence Framework; and establish career paths and progression for data managers as members of research teams. In addition, they recommend that research institutions should develop clear policies on data sharing and preservation and provide training and support for researchers to manage data effectively; and for journals to establish clear policies on data sharing and processes, with datasets underlying published papers readily accessible, and with appropriate data citation and acknowledgement (EAGDA, 2014).

The European RECODE project, investigating values, motivations and barriers towards open access to data in five case studies, found as motivations for researchers to share their data openly: easier access to data for comparison, error testing and to avoid duplication; faster advancement of science; more reliable research results; combined work; encouraging industrial uptake of data for commercialisation, and cumulative knowledge (RECODE, 2013). Overall, the project reported that incentives for providing open access to data were quite weak, and also found a lack of incentives for researchers to participate in data review processes (RECODE, 2014).

Boosting data citation metrics and impact factors that reflect data reuse, and the weight data sharing and publishing may therefore carry in career progression, (similar to the importance of paper citation indices and the impact factors of journal articles), are often flagged up as potential incentives for increased data sharing by researchers (Costas et al, 2013; Force 11, 2013).

Many studies reinforce a common message: data and the sharing of information is what research is truly about. Research findings are shared in publications and expertise is shared through networking.

1.3. Policy background

In the Netherlands, the Netherlands Organisation for Scientific Research (NWO), which is the main scientific funding organisation, expects publicly funded research data to be made available for reuse by researchers wherever possible. The NWO regulation on granting stipulates that data that emerge from research funded by NWO is co-owned by NWO, with NWO having a say in making data accessible (Knowledge Exchange, 2013). NWO encourages open access to both publications and research data. The Royal Netherlands Academy of Arts and Sciences (KNAW), via its policy on open access and digital preservation (KNAW, 2011) requests its researchers to digitally preserve research data, ideally via deposit in recognised repositories, to make them openly accessible as much as possible; and to include a data section in every research plan stating how the data produced or collected during the project will be dealt with. ZonMw, a funder of health research in the Netherlands, requires funded researchers to make collected research data available for reuse (ZonMw, 2014).

In Germany, the German Research Foundation (DFG), the main funder of basic research, encourages open access publishing and sharing of research data. It introduced in 2010 a requirement for grant applicants to state what they plan to do with their research data during and after a proposed research project. Such statements are reviewed as part of the application review process. The requirement aimed to

raise awareness of data sharing and data management amongst applicants. Rather than developing single shared infrastructure, the focus in Germany is currently on bottom-up initiatives and discipline-specific or institutional policies, such as the Max-Planck-Society Data Policy, or the Grundsätze zu Forschungsdaten an der Universität Bielefeld (Knowledge Exchange, 2013). Recently in July 2014, the universities of Göttingen, Humboldt-Universität zu Berlin and Heidelberg have also adopted research data policies, expecting sustainable data management practices from researchers, as well as timely access to research data, especially those underlying scholarly publications. The Alliance of German Science Organisations, a consortium of nine academic organisations, adopted in 2010 principles for the handling of research data, supporting long-term preservation and open access to research data for the benefit of science (Alliance of German Science Organisations, 2010).

In Denmark, the funding entities Council for Independent Research and Council for Strategic Research have currently no requirements on data sharing or research management for research grants. However, the Danish e-Infrastructure Cooperation (DeIC) and the Danish KE partner DEFF are developing a national strategy for data management for all research institutions. Also individual institutions are in the process of developing data strategies, e.g. the University of Copenhagen is developing recommendations for the management of research data (Knowledge Exchange, 2013).

In Finland, the government published a resolution on improving the availability and promoting the reuse of public sector digital data in 2011. A national roadmap for utilization of electronic data in research was published in 2011. It was followed by two initiatives of the Finnish Ministry of Education and Culture, Open Science and Research (2014-2017), and the National Research Data Initiative (2011-2013), which are key actions in developing data infrastructures and services, standards, university and funder policies and recommendations (Knowledge Exchange, 2013). Research funder Academy of Finland requires data management plans as part of funding applications and offers the possibility to apply for funding to cover the costs of open access publishing. Selected funders such as Tekes (Finnish Funding Agency for Innovation) require data to be disseminated. For institutions of higher education, standardised data policies are still being defined and much variation remains in methods and practices amongst the institutions (TTA, 2014).

In the United Kingdom, the seven research councils that are the public funders of research - Arts and Humanities Research Council (AHRC), Biotechnology and Biological Sciences Research Council (BBSRC), Economic and Social Research Council (ESRC), Engineering and Physical Sciences Research Council (EPSRC), Medical Research Council (MRC), Natural Environment Research Council (NERC) and Science and Technology Facilities Council (STFC) – each have a data policy and jointly have adopted Common Principles on Data Policy (RCUK, 2011) and a Policy on Open Access (RCUK, 2012). Between them, these require that publicly funded research data should be made openly available with as few restrictions as possible in a timely and responsible manner that does not harm intellectual property; and for peer reviewed research papers to be published in journals that are compliant with the Policy on Open Access; and to include a statement on how the underlying research materials such as data, samples or models can be accessed. ESRC and NERC had adopted data sharing policies as early as the mid-1990s (Corti, et al., 2014); they both mandate data sharing as a condition of funding, whilst other councils encourage data sharing. Most councils require data management and sharing plans to be submitted with grant applications. Councils place responsibility for managing and sharing research data with research grant holders, except EPSRC who place the responsibility with host research institutions (EPSRC, 2011). Research organizations receiving EPSRC funding will from May 2015 be expected to have appropriate policies, processes and infrastructure in place to preserve research data, to publish metadata for their research data holdings, and to provide access to research data securely for 10 years

beyond the last data request. As a result, there is now an increase in UK universities and research organizations having institutional data policies and developing research data repositories and data support services for their researchers, as well as training programmes.

Following on from public funders, some private research funders, charities and governmental departments that also fund research have adopted data sharing policies: Cancer Research UK, Department for International Development, Department of Health, Nuffield Foundation and the Wellcome Trust.

UK Research Councils also fund a variety of data sharing support services and data infrastructure, such as the UK Data Service (ESRC-funded) and NERC data centres. BBSRC, Cancer Research UK, MRC and the Wellcome Trust - the main funders of medical and public health research - are partners in UK PubMed Central.

At a European level, the European Commission (EC) published in 2012 a communication and recommendations on access to, and preservation of, scientific information (European Commission, 2012a; 2012b), calling for coordinated actions across all member states to drive forward open access, long-term preservation and capacity building to promote open science, not just for EC-funded research, but also for national research funding. For the current framework funding programme, Horizon 2020, the commission has published requirements for data management planning, and has included a pilot on open access to research data in seven thematic areas, in particular for data underpinning open access papers.

In line with data sharing policies amongst research funders, journal publishers also increasingly have data policies that require research data underpinning the findings published in peer-reviewed articles to be available for readers and reviewers, for scrutiny and verification of published findings, as well as future reuse of such data. This is driven both by prominent research funders urging publishers to adopt stronger data policies to help drive the open data movement, and by the publishing community itself to counteract fraudulent research. The open access publisher BioMed Central has strongly pushed the adoption of journal data policies since 2010, through its open data statement, its cross-publisher working group on open data publishing and its ongoing efforts on the copyright and licensing of data underpinning publications (Hrynaszkiewicz and Cockerill, 2012). The Joint Data Archiving Policy, adopted by many leading journals in the field of evolution in 2011, and since adopted by numerous other journals across various disciplines, strongly influences the increase in journal data policies. This requires data that supports the results described in a paper to be archived in an appropriate public archive as a condition for publication, often recommending Dryad as an appropriate repository.

2. Objectives and conceptual framework

The overall objective of this study was to provide evidence and examples of useful incentives for data sharing from the researchers' point of view to inform scientists and policy makers.

The specific aims were to:

- Identify - known but especially as yet unknown or unofficial - incentives of researchers for making data available with a focus on the values and intrinsic motivations of the individual, as well as on the interactions within research teams and in the larger research community
- Analyse existing and possible future benefits for researchers sharing their data
- Investigate the influence of existing policies on the practice of data sharing throughout the whole life-cycle of the research process as well as the influence of existing institutions and infrastructures offering support services for data sharing
- Consider the whole research lifecycle and determine the most opportune moments in the research process for incentivising data sharing
- Provide recommendations for policy development regarding the incentivising of data access and reuse

In this report the following concepts are used:

- **Data sharing** can be defined as the release of research data for use by others (Borgman, 2012). Diverse modes of data sharing exist, and are described in chapter 5, in line with degrees of openness (Whyte & Pryor, 2011). Data sharing requires the systematic collection, curation and dissemination of data
- **Research data** is information relevant to or of interest to researchers, and is not limited to data which is the output of research. It is any research materials resulting from primary data collection or generation, or derived from existing sources intended to be analysed in the course of a research project. This can be numerical data, textual data, images, multimedia recordings, scripts, code, etc.
- **Data repository** is a centre or facility that holds data, and where that entity assumes some responsibility for holding, storing and disseminating data, but does not necessarily provide long-term curation or preservation
- **Supplementary data** is the data submitted to journals with manuscripts in advance of publication. This data provides the underlying evidence for outputs and findings in the article (e.g. tables, graphs, statistical results), and does not necessarily include all data used to produce the article

3. Methodology

The study is based on interviews with 22 selected researchers within five research groups or projects with established data sharing practices as case studies; each case study representing a different academic discipline.

By studying researchers who actively share their data, we gathered knowledge about what happens on the ground, rather than only the perceptions researchers may have about data sharing.

3.1. Sample selection

The sampling strategy employed was purposeful, with a small number of cases (five) chosen to meet specific criteria. All five cases were selected based on the target phenomenon of interest - data sharing. In all cases, data sharing was known, *a priori*, to be taking place, albeit in different forms. Maximum variation (Sandelowski, 1995) was sought in factors such as scope, size and funding sources of the research groups (national vs. international; small vs. large; national vs. international funding, localized vs. cross-institutional), the career stage of researchers (from postgraduate researchers through to research programme leaders), and the types of data generated in the research (crystal structures, genetic sequence data, biosamples, audio recordings, etc.). The research explored variations within and across cases to better understand the incentives that motivate such data sharing. Selected case studies and researchers are thus not representative for the entire research community.

The five research groups or projects (and their disciplines) were selected by Knowledge Exchange (KE) in partner countries, based on their existing data sharing practices. Each of the case studies was classified according to the specified criteria: (inter)national, maturity of data sharing, etc., and available information on research and data sharing practices was collected for each case. KE also identified and selected one or more local researchers in each country to conduct the interviews.

KE identified a primary contact for each site, and that person in turn suggested additional researchers to interview, based on defined criteria to ensure variability: career stage, specialist knowledge and data sharing practices.

3.2. Interviews

Information was collected through semi-structured qualitative interviews with the selected researchers, in English or in the local language. Interview questions were developed, based on the investigators' expertise and literature review, in collaboration with KE and the local interviewers. An initial list of questions was discussed with KE representatives and local interviewers. Feedback and comments were integrated to develop the final list of interview questions [Annexe 1].

Interviews were carried out by local interviewers in each country, in consultation with the investigators. The investigators conducted the interviews in the UK. Interviews were audio-recorded and transcribed. All interviews were summarised using a standard template to provide comparability of topics [Annexe 2] and, where necessary, translated into English. Time did not permit conducting pilot interviews, but substantive and logistical issues were reviewed and resolved after the first interview in each case. Basecamp was used as file sharing space for interviewers to share and exchange relevant interview information and documents.

The investigators proposed four to five interviews for each case. The final numbers were: Denmark (4), Germany (6), Netherlands (1), Finland (1), and the UK (10). In all cases, primary decisions regarding the sample were made by the local interviewer(s). In both cases where a single person was interviewed, the investigators discussed options for boosting the sample with KE and the local teams. Due to various factors, such as interviewer schedules, risks of delay to overall project, and the total sample achieved, it was agreed that no additional interviews would be done for those cases. The final sample of 22 fell within the planned range of 20-25 interviews, albeit with more variation in sample size per case than was originally intended.

All interviewees have given permission for their names to be used in outputs, for the use of exact quotations, and photographs taken during interviews. They equally consented to the interview recordings and transcripts being archived and made available for future use via the UK Data Service and Knowledge Exchange. The archived data from this study are held in UK Data Service ReShare (Van den Eynden and Bishop, 2014).

Informing the study was a comprehensive literature review of existing information on barriers, enablers, incentives and benefits of data sharing, including the influence of funder and publisher policies, IT capability, data infrastructure and support services provided by funders and institutions.

3.3. Methods of analysis

A comparative analysis across two dimensions was carried out. Firstly, the cases were selected explicitly to explore known factors that influence data sharing, such as project scale and discipline. The semi-structured questions allowed the collection of reasonably consistent information across all the cases. Secondly, comparisons were developed within cases, in particular for the three cases with multiple interviewees: Denmark, Germany and the UK. Each of these cases proved to show significant variation in sharing attitudes and practices, making it clear that broad discipline and research group alone do not determine all aspects of data sharing. This method is based on the Framework approach, in use at the National Centre for Social Research in the UK (NatCen, 2012).

Both investigators analysed the full set of interviews, starting from the comparable summaries, through synthesis and comparison of the topics discussed across all interviews. This was followed by comparison and discussion of emerging themes and findings. A draft report of findings and recommendations was circulated to KE members and interviewees for feedback. Modifications included a revised category scheme for the types of incentives identified.

Finally, it should be noted that for the final recommendations and action points, the investigators have drawn on all available sources: literature, research findings, and their own expertise.

4. Case studies

The five case studies span various academic disciplines: arts and humanities, social sciences, biomedicine, chemistry and biology.

Each case has been given an acronym, e.g. [RetiredMen], and these are used to reference the cases in further chapters throughout the report.

4.1. Retired men gathering in cities, Finland [RetiredMen]

This case was selected as an exemplar of small-scale, qualitative sociological research conducted by a single researcher for a master thesis. The methodology combined ethnographic observations and interviews with elderly men on a town square, as well as a field diary kept by the researcher. The intent was to study behaviour and social organisation in the tradition of urban ethnography. A distinguishing feature of the case was the archiving of the field diary, alongside the interviews. Anthropologists and ethnography researchers debate as to whether field diaries should be shared, many arguing that the information may be too sensitive or personal to share; yet some responding that sharing field diaries is a way of being open and transparent about observational research methods. The researcher in this case believed strongly in the value of sharing diaries because of the benefits he could realise were he able to learn from diaries of experienced researchers.

However, the researcher did not intend to archive or share the data at the start of the research. The project won a national award, bringing it to the attention of the Finnish Social Sciences Data Archive (FSD). They requested that the data be deposited and the researcher agreed. The Archive provided extensive support in processing the data, including cleaning, helping with anonymisation, and developing necessary documentation.

The main form of data sharing in this case was through a formal data repository (Wallin, 2011).

This case study consisted of a single interview with the researcher who conducted the study.

4.2. LARM Audio Research Archive, Denmark [LARM]

The LARM Audio Research Archive is an interdisciplinary project with a goal of producing a digital infrastructure to facilitate researchers' access to the Danish radiophonic cultural heritage. It comprised a consortium of ten Danish Universities and cultural institutions. This project is based in the humanities. It focused on the infrastructure of a digital audio archive that makes audio files of radio broadcasts available to researchers and allows annotation of such audio files. The content of radio programs is diverse, and included print materials of radio programs. Distinctive features of this case include its large size across institutions, national scope, multi-disciplinary research teams (musicology, socio-linguistics, media studies), and the emphasis on both enhancing content and expanding infrastructure.

LARM allows data to be shared through a Creative Commons licence. One LARM project, LARM.fm, developed the digital platform to enable searching, sorting and annotating of audio files. Researchers access data, adding their own annotations to them, based on their various research goals. LARM is available to researchers and students at Danish universities, but not the general public.

Data are shared both through formal infrastructures (LARM.fm) and through personal networks. For some researchers, both inadequate infrastructure and confidentiality constraints made personal networks the preferred mode for sharing their data.

The former project manager of LARM was interviewed, as well as a musicology professor who participated in a LARM work package on radio and music and uses LARM.fm, a PhD student in sociolinguistics who used LARM to study pronunciation, and a PhD student in media studies who used it study how Danish radio has used media to develop adult education.

4.3. Netherlands Bioinformatics Center [NBIC]

This research centre was selected as an example of a national centre of expertise, with strong international links, where research is funded by both academic funders and private enterprise. Bioinformatics research applies computerised tools and methods to generate meaningful biological knowledge by processing and interpreting complex and large datasets. Researchers analyse sequence-based information, genotype-phenotype relations, proteomics and metabolomics expressions, etc.

The research case studied focuses on biosemantics research, mining existing literature and data resources for associations and patterns between genetic sequences and their expressions (meaning) in the field of biomedicine. Data sharing and access to information (literature) in the form of computer-readable expressions, as well as being able to attribute (cite) the provenance of information, is essential for this kind of research.

A senior investigator who plays a prominent role in making the case for increased data sharing for the benefits of bioinformatics research (Mons et al, 2011), was interviewed. He advocates rapid and public sharing of research data, as well as nano-publication or nano-attribution of meaningful assertions, as smallest elements in a database and linked with a publication.

4.4. Evolutionary Plant Solutions to Ecological Challenges, Germany [Adaptomics]

This DFG Priority Programme, also named Adaptomics, is funded by the German Research Foundation (DFG) and currently has 22 ongoing research projects across universities in Germany and beyond. It was selected as an example of a large international consortium, where collaboration, sharing expertise and sharing research data within and across projects is crucial for the research. The core research is funded by DFG, but individual researchers are also funded by the private sector and venture capital for other research activities, and operate within international informal research networks.

The central focus of this cross-disciplinary programme is gaining insight into the genetic and molecular traits that Brassicaceae plant species develop to match local environmental demands, the relevance of specific sequence variation for plant performance in the natural environment and its evolutionary role. Concentrating on the ecologically diverse Brassicaceae family capitalizes on the in-depth molecular understanding of the Brassicaceae model plant *Arabidopsis thaliana*.

The main research disciplines within the programme are evolutionary biology, population biology, taxonomy and biodiversity research, and the research data created and used are gene sequence data, microarray data, taxonomy records and ecology data. Research data are typically shared early in the research process for collaborative purposes, with collaborators and trusted networks. Processed data are shared as supplementary data with published papers (e.g. Koch et al, 2013), with most journals by default expecting supplementary data to be submitted as proof of results, either to the journal itself, or lodged with established international repositories such as GenBank or ENA for gene sequence data (e.g. ENA accession ERP000102 for Sharbel et al, 2010). DFG also directly funds data sharing through the development of database resources within the programme, such as BrassiBase, an online knowledge and database system on Brassicaceae taxonomy, systematics and evolution, including chromosome numbers, traits and characters and germplasm resources (Koch et al, 2012).

Six researchers of four projects were interviewed: the coordinator of the entire programme, two PhD students, two principal investigators of individual projects and the data manager for BrassiBase.

4.5. Chemistry Department, University of Southampton, United Kingdom [Chem]

The Chemistry Department at Southampton has a relatively long history of engagement with complex data sharing. The university hosts the National Crystallography Service and eCrystals archive, directed by a staff member of the department, and also has links with the Cambridge Crystallographic Data Centre and the Dial-a-Molecule Network, a network for rapid development of synthetic molecules. Staff members have been involved in many initiatives, such as promoting the adoption and use of electronic lab notebooks (ELNs) to enable the sharing of descriptive information about reactions and experiments, data citation and the development of metadata standards for crystallography. This case provided an opportunity to closely examine a domain (crystallography) often identified as an exemplar of long-established and well-functioning data sharing practices, side-by-side with other domains such as organic chemistry, synthetic chemistry and physics where sharing is much less widespread. Uptake of ELNs is particularly promoted for the Dial-a-Molecule Network, as sharing reaction protocols from notebooks helps design rapid pathways for developing new molecules. Researchers in these groups also span diverse disciplines - besides organic and synthetic chemistry, computer sciences, physics and social science research into data sharing practices, uptake of ELNs and academic publishing take place within the research group.

Another distinctive feature of this case is the department's multiple connections with the private sector, for example in the form of Knowledge Transfer Partnership for supporting and testing ELNs in private company research labs. The competitive nature of some genres of research in the group also raises challenges of intellectual property, patents, and commercial impacts of sharing data. Finally, the department received funding from the EPSRC, which - uniquely amongst UK funders - places primary responsibility for data sharing at the institutional level, rather than with the individual researcher.

A very wide array of systems were used for formal sharing, such as the National Crystallography Service, eCrystals and the Cambridge Crystallographic Data Centre for sharing crystal structures (e.g. Tizzard et al, 2011). CCDC extract crystal structures from publications. Other formal channels such as publications and ePrints Soton (the publications and data repository of the University of Southampton) were also mentioned, as was GitHub for sharing code. Personal sharing took place too, both through hierarchical relationships (student data shared with supervisor) and peer-to-peer. Some disciplinary differences in data sharing practices can be noted. Researchers applying physics (e.g. in the development of lasers and x-ray sources for x-ray microscopes) indicate that data sharing is not customary and journals do not expect data as supplementary files, since the development of experiments (rather than the output data) is the crucial part of the research, with little reuse potential for the output data.

In chemistry, the Royal Society for Chemistry (RSC) has played a leading role in the development of community resources and databases such as ChemSpider and ChemSpider Synthetic Pages.

Interviews were conducted with two professors of chemistry, a reader in optoelectronics research, the coordinator of the Dial-a-Molecule Network, the director of the National Crystallography Service, a trainer in a Knowledge Exchange programme and four PhD students.

Characteristics of the five case studies

Country	Finland	Denmark	Netherlands	Germany	UK
Case study	Retired men gathering in cities	LARM Audio Research Archive	Netherlands Bioinformatics Center	Evolutionary Plant Solutions to Ecological Challenges	Chemistry Department, University of Southampton
Acronym used in report	RetiredMen	LARM	NBIC	Adaptomics	Chem
Research discipline	Sociology	Humanities	Biomedicine	Plant genetics, taxonomy, ecology	Chemistry
Key contact	Antti Wallin	Bente Larsen	Barend Mons	Ute Krämer	Jeremy Frey
Research group type	National project	National project	National team	International consortium (research programme)	Research group with international links
	Small project	Large project, multi-disciplinary	Medium-sized project	Large programme; 22 projects; cross-disciplinary in ecological and biological sciences	Cross-disciplinary
Research	Ethnography of behaviour and social organisation	Linguistics, media sociology, cultural studies, media studies	Bioinformatics and biomedicine	Brassicaceae research: evolutionary biology, population biology, biodiversity research	Organic chemistry, synthetic chemistry, electronic lab notebook development; also studying data management and sharing attitudes and practices, as well as promoting good data practices.
Research funding	Ministry	National Programme for Research Infrastructure.	NWO	DFG	EPSRC, RCUK, EU
Research data	Interviews, observations, field diaries	Radio broadcasts, radio programme info	Using third party data and literature for data mining	Microarray data, taxonomy, ecology data, gene sequence data, phenotype measurements	Crystal structures, chemical fractions, chemical processes, diffractive images,
Data policy background		Funding of archive development to provide access to digital content (radio programmes) to researchers	NWO expects research data to be made available for reuse by researchers wherever possible	DFG encourage open access publishing and sharing of research data	EPSRC expect institutions to disseminate data from 2015
Data sharing features	The project was for his master thesis. It won a national award. Then the Finnish Social Sciences Data Archive asked him to deposit. They worked on the data, cleaning and documentation, to make it suitable for sharing.	Digital archive to make sound files available to researchers (e.g. radio broadcasts). Audio files of diversified content, as well as the printed radio programs. LARM was conceived to support a variety of research topics related to this material.	Mons wrote Nature commentary on data sharing (the value of data); data mining is essential for his research (biosemantics).	Data sharing between programme projects is essential; also sharing through collaborative research and as supplementary files with publications; data sharing via genetic databanks, e.g. NCBI GEO, Genbank; online BrassiBase.	Sharing electronic lab notebooks (ELNs); Soton had Jisc MRD project on data citation from ELNs; Crystal structures shared via Crystallographic Data Centre.
Interviewers	Irina Kupiainen, Damien Lecarpentier	Anders Sparre Conrad	Joeri Nortier	Jens Nieschulze, Juliane Steckel	Libby Bishop, Veerle Van den Eynden
Interviewees	Antti Wallin	Bente Larsen Jacob Thøgersen Morten Michelsen Janne Nielsen	Barend Mons	Ute Krämer Nora Hohmann Marcus Koch Stefan Wötzel Timothy Sharbel Markus Kiefer	Jeremy Frey Simon Coles Susanne Coles Richard Whitby Will Fryson Oliver Brand Cerys Willoughby Bill Brocklesbury Xiaoping Tang Phil Adler

5. Data sharing and data reuse practices amongst researchers

5.1. Diverse modes of sharing

When interviewed researchers talk about sharing data, different forms of sharing can be distinguished, similar to the six degrees of openness described by Whyte and Pryor (2011):

- **Private management:** sharing data with colleagues within a research group
- **Collaborative sharing:** using data within a consortium
- **Peer exchange:** sharing data with trusted peers in informal networks
- **Transparent governance:** sharing data with external parties such as funders and institutions for accountability, research assessment, scrutiny or inspection
- **Community sharing:** with members of a research community
- **Public sharing:** making data available to any member of the public

Besides these different forms, data sharing can also be reciprocal between different parties (mutual benefit); or can be in the form of donation from one party to another (possibly unknown) party, separating data producers from data users.

Sharing data is seen by most researchers interviewed, across all case studies, as part of the normal scientific process. This usually refers to cases where data are shared within the research group, in collaborative research or with trusted researchers (peer exchange); or where data that underpin findings published in papers are deposited as evidence to allow replication or validation of results. Data that a researcher feels could still be exploited for future publications are usually not shared.

In the [Chem] and [Adaptomics] case studies, research data were routinely shared within the research group or with collaborating researchers in a project. Within [LARM], researchers had access to each other's data and annotations. Especially for confidential data (e.g. interviews), such sharing within the research group is seen as the safest option for sharing data [LARM]. When data are shared in collaborative research or via peer exchange, this is usually early in the research cycle, whereby raw data or slightly processed data are shared, usually shortly after data have been created or captured. There is usually an informal agreement of understanding about the ownership of the data and how other researchers can use them. In some research domains, this is seen as a strategy to optimise the analysis of research data [Chem; Adaptomics], with different researchers contributing different expertise to data analysis. Such sharing typically provides mutual benefits, such as co-authorship, the sharing of expertise, or the possibility that by sharing data with peers, one might in future equally be able to request those peers' data. Sharing data within a collaborative or peer network also provides insight into the kind of research taking place within the network and the resources available to researchers for future use [LARM]. Where shared data are essential for someone else's research, co-authorship would be expected [Adaptomics; Chem]; where data would provide supporting information for research, citation is expected. When data are shared for collaborative research, the data are usually similar, with a familiar file format, so sharing is easier, requiring little data preparation besides annotation, and is by some researchers considered as more meaningful. Where research is competitive [Adaptomics] data sharing through collaboration provides control.

Researchers frequently share their data via publications, as published findings (directly included in the article) and via supplementary data files that may be deposited with the journal itself or in a repository [Adaptomics; Chem]. Data shared in this way provide evidence for published findings (transparent governance); but can also be seen as community sharing or public sharing if the supplementary data are detailed enough and of high quality to enable reuse in future research. Data shared as evidence for published findings tend to be processed data rather than raw data [Adaptomics; Chem], allowing verification or validation of the findings; but may be less versatile for reuse. Also, only those data that directly relate to the published

findings are usually included in supplementary files, rather than entire datasets. Much data are therefore never shared. Although there are exceptions, with some researchers submitting very detailed and raw data files as supplementary materials [Adaptomics]. According to the interviewed researchers, journals usually do not set standards or expectations for supplementary files. Most interviewees state that journals expect supplementary data to be deposited with the journal itself, rather than with a dedicated data repository. This provides fewer guarantees that such data files will remain accessible in the long term. In some research areas domain repositories are well established so that journals do recommend such repositories as place of deposit. In crystallography, journals require crystal structures to be submitted to a central database such as the Cambridge Crystallographic Data Centre or the Crystallography Open Database [Chem]. In biology, gene sequence data are normally expected to be deposited at the time of publication in community repositories such as GenBank and the European Nucleotide Archive (ENA) at the European Molecular Biology Laboratory. Other types of biology data underpinning publications are submitted to Dryad or to the journals themselves [Adaptomics]. Some chemistry researchers comment that supplementary files are often just tables of processed and synthesised data, often in PDF format, not the raw reusable data [Chem]. In biosemantics, research depends strongly on data being available alongside publications in a usable format, since data provide machine-readable information (which written articles may not do), and data being available to other researchers rapidly [NBIC].

Sharing for transparent governance is indicated in chemistry research, where research data are typically kept for at least ten years as a safety record and as a funder requirement [Chem]. Where applicable, lab notebooks are kept to substantiate patent applications [Chem].

Where researchers create community or topical databases as direct outputs from research, this is because they are directly funded to do so, e.g. BrassiBase [Adaptomics] and the LARM radio archive [LARM].

Only a few of the interviewed researchers have placed entire datasets into a public or community repository, either through invitation by the archive [RetiredMen] or for data that underpin a publication [Chem]. Crystal structures may be submitted to eCrystals (as prepublication / micro-publication) or to CCDC (extracted from publications), as reference for chemical compounds, but for most compounds the structure is never described.

Researchers recognise that the longevity of data resources can vary according to the type of data or research. For example, DNA sequences may be outdated after a decade, whilst taxonomy data may be relevant forever [Adaptomics].

5.2. Optimal data sharing moments in the research lifecycle

Research data tend to be shared at two main phases in the research cycle :

- early in the research process
- at the time of publication

Early in the research process, shortly after data have been created, raw data are shared through collaborations and informal contacts, sometimes with some descriptions, annotation or interpretation [Chem, Adaptomics, LARM]. For some researchers trust is an essential condition for sharing data in this way. Raw data may be easier to share compared to processed, as sharing the latter would give away methods prior to publishing and requires effort for data preparation.

At the time of publication, data are either provided as evidence for published results and submitted as supplementary files with the manuscript to a journal or a recommended repository [Chem, Adaptomics]; or are deposited just after publication into a repository [Chem, Adaptomics, RetiredMen]. Such data are quality checked and usually processed, although some researchers submit raw data with scripts for traceability, or simply raw

data such as raw sequence data reads submitted to GenBank [Adaptomics]. This tends to be a well-established practice in many of the natural sciences, but not so in the social sciences and humanities.

Researchers made a few additional remarks about the timing of data sharing. In some research domains data sharing before publication is difficult since journals require that the information contained in a manuscript has not yet been published or made public elsewhere [Chem]. One researcher who submitted data to a data archive after publication stated that in future earlier sharing would be easier, or at least early planning and preparation [RetiredMen]. Data management and sharing support is generally considered to be needed by interviewees during the planning stages of research.

5.3. Using data

Whilst all interviewed researchers share their own research data and many use data from other researchers, not many access data from public data repositories or community repositories. Data are most frequently sourced via literature, as an integral part of literature review, - either data/information that is in the article itself or from supplementary files to the article [Chem; Adaptomics] - or via collaborations, networking and personal contacts [Adaptomics, LARM]. Data may be needed as a primary resource for research, or to support the research, e.g. as contextual information. Interviewees indicate that supplementary data to published articles can vary widely in quality and usability: these data are typically behind subscription walls, like the articles themselves, and may be in formats not readily suitable for reuse (e.g. PDF of a data table). This makes the reuse of such data difficult.

Some researchers source research data from closed information collections such as museums, radio archives and web archives [RetiredMen; LARM]. It often requires inside knowledge to know which data may be available, as well as negotiating cumbersome procedures to gain access.

Research data resources that interviewees appreciate most are those where data, articles and other information are available from a central resource, so the research data are accompanied by documentation and context. Examples mentioned by interviewed researchers are PubMed, The Arabidopsis Information Resource (TAIR), LARM, ChemSpider and the Cambridge Crystallographic Data Centre. These are public resources or community resources. Some researchers indicate that even more detailed narratives of the process of generating data would be beneficial [Chem].

5.4. Tracking reuse

Overall, none of the interviewed researchers are overly interested in tracking use of their shared data, although they routinely expect citation or co-authorship when data are used by other researchers. This is probably influenced by the fact that most reuse is via collaborative research or as peer exchange, so the reuser is known to the original researcher.

One researcher [Chem] uses Google scholar citations to track citation of publications and is starting to see data citations appear in such metrics.

5.5. Sharing culture

Norms within research groups, departments, projects or entire disciplines strongly influence data sharing in these case studies, either in favour of sharing [Adaptomics; NBIC; Chem; LARM]; or against sharing [Retired Men]. Many interviewed researchers refer to the fact that data sharing is just something they do within their research group, or something they've been used to from the start of their research career [Chem; Adaptomics]. In some disciplines data sharing is indeed well established and has become part of standard research practice, e.g. biology and genetics [Adaptomics], biomedicine [NBIC] and crystallography [Chem]. In other research communities data sharing is not yet that common, e.g. humanities [LARM], sociology [RetiredMen], organic chemistry and physics [Chem]. Researchers may see data as being private and too revealing about individual research ideas to be shared.

Besides research data being valuable resources for research, datasets are also considered necessary for teaching and training young researchers in analysis [Adaptomics; Chem]; or to be able to test algorithms or scripts during research processes [Chem].

5.6. Conditions for data sharing

Researchers identify a range of conditions that they perceive as conducive to data sharing. Some of these conditions involve control or access limitations placed on the data, controlling who can access and use the data, for which purpose, or when data will become accessible. Confidentiality of data or competition are typical factors motivating the need for regulating data access [Chem, Adaptomics, LARM]. The ability to limit data sharing to a trusted group is crucial to researchers in some cases to share data in a competitive research environment [Adaptomics]; or data may be available specifically for research and learning purposes [LARM]. In other cases, timing was important, being able to embargo data until key publications could be written and published [Chem, Adaptomics].

Other conditions researchers highlight as important for sharing data are appropriate acknowledgement via citation or co-authorship. The degree of acknowledgement researchers expect depends on whether the data that are reused provide supporting information in the new research (citation), or a crucial element in the new research (co-authorship).

Good data infrastructure was also cited as a necessary condition for greater sharing [NBIC, Adaptomics].

In general, there was a sense that sharing data should not impose an undue (and unfunded) burden on the data creator. If data sharing imposes additional costs beyond the primary research, such costs should be met with extra funding [Chem, Adaptomics, LARM].

In one case, relatively tight access controls were assigned to archived data, requiring permission from the data creator before the archive can release the data. A few years on, the researcher felt this condition was too restrictive and could be relaxed [RetiredMen].

5.7. Barriers to sharing

Interviewees reported various barriers to sharing data, similar to those described from literature in the introduction. We did find a wide range of barriers identified, especially considering the modest size of this sample. Formal research assessment, with its emphasis on the impact of articles, was seen as doing untold damage to motivations to share data. Since sharing data is not rewarded, any time or effort taken to share data is wasted for career progression; with a risk that data users would benefit from publishing based on the data. In areas where commercial entities or content are prominent, intellectual property (IP) and copyright restrictions were also seen as major inhibitors of sharing. Finally, insufficient infrastructure was cited by some, though again, there was wide variation in responses on this topic.

Researchers described in general terms conditions or requirements that did, or could, help to promote sharing. Researchers find it easier when sharing activities build on normal and routine work practices. The general atmosphere and relationship of trust was also highlighted, most important in peer exchange and collaborative sharing. A number of researchers flagged up the need to retain some levels of control over the kind of data they share, and with whom. A rich and diverse infrastructure that supports different modes of sharing, e.g. within a team, discipline, with vetted strangers or the public, was seen as useful too.

5.8. Sharing negative findings and failed experiments

Many researchers in the [Chem] and [Adaptomics] cases regret the fact that currently failed experiments and negative findings are typically not published and therefore the data or information related to these is not shared. This is seen as particularly important to avoid numerous repetitions of the same experiments by different researchers, since the knowledge that a particular experiment fails is never published or shared.

5.9. Variation across stage of career

Some researchers in every research group connected incentives to share data with phases of research careers. Generally, three career phases were relevant: early, middle and established.

Researchers early in their careers, typically bachelor, master and doctoral students up until completion of their degrees, experience two competing pressures. First, they confront fears of exposing their work, both out of concern about getting scooped, and also because of potential embarrassment for showing immature, naive, and possibly wrong data or procedures [RetiredMen; Chem]. On the other hand, it is these same early career researchers who are highly motivated to make a name for themselves by getting credit for new methods, procedures, and increasingly, by sharing data.

For researchers in the middle of their careers, there is less concern about showing immature work, although fears of being scooped can still persist. These researchers are focused on building their careers, and typically are responding to the major incentives in their disciplines, usually publication in key (high impact) journals. They may be less eager to share, making sure to maximise the publications they can get from their valuable data resources, and keeping data for future research [Adaptomics].

As researchers enter the later stages of their careers, interest in sharing data can grow quickly in some cases [Chem; LARM]. In these cases, the contribution of a major corpus, such as data, is important in establishing a reputation as a major scholar in a domain, leaving a legacy for others to build upon.

In terms of incentives, researchers point out that different incentives apply at different career stages. For example, for early careers researchers, invitations to share from recognised sources (e.g. conferences) can help to overcome fears [Chem], as can the ability to control the scale of sharing [Chem]. The idea of embargoes until after the publication of theses is relevant for several cases [Adaptomics; Chem]. In contrast, those in latter phases of their careers are motivated by the desire to leave a legacy, but often need support and assistance with tasks such as file format conversion and documentation in order to share their work.

Significantly, given the variety of this sample, most researchers in all the cases stressed the importance of early career training in data management and sharing, suggesting the importance of establishing expectations about and practices of sharing as essential elements of researcher education.

5.10. Funding models for data sharing

Two issues were identified by researchers regarding models for funding data sharing. The first concerns sustainability. Even when short-term project funds are available specifically for data sharing, this does not necessarily produce a predictable funding stream to finance core data infrastructure [LARM, NBIC]. Some researchers noted that standards, such as for metadata and file formats, were a bigger challenge than funding [Adaptomics].

The second issue concerns the appropriate allocation of costs to prepare data to be shared. Data requiring extensive anonymisation or documentation can place a burden on the data creator [LARM]. Where these tasks went beyond what data creators needed for their own use, it was generally felt that such costs should be more centrally funded [Adaptomics]. Research data shared as supplementary material with papers usually require no or little cost (but may be of low reuse value). The upgrading and curation of full data collections for deposit in repositories often requires extra time and effort, and therefore funding [Adaptomics].

Charging for data use was not common, but seen as justified in cases where the data repository provided extensive value-added services, as well as making the data available [Chem].

6. What motivates researchers to share data?

The incentives that currently motivate interviewed researchers to share their research data, fall within three main categories:

- Direct benefits
  - for the research itself (more robust)
  - for the career of the researcher (recognition)
  - for discipline (get wiser)
  - for science (better science)
- Norms of the project, research group, and/or discipline
- External drivers: policies and expectations from research funders and publishers

6.1. Direct benefits

6.1.1. Data sharing is essential for the research

In certain research fields or for certain researchers, data sharing is an essential part of the research process, or the research depends entirely on the sharing of data. This is the case where various researchers need to be engaged in analysing complex data and the data are shared through collaboration or peer exchange, e.g. in genomics research [Adaptomics]; or where lots of information and data accessed via community sharing or public sharing are needed to find results, e.g. in biosemantics research [NBIC].

In the case of collaborative data sharing or peer data exchange in research, various researchers often contribute different expertise when analysing particular datasets [Adaptomics]. The collaboration and data sharing means that different interpretations of the data can be done. Exchange of ideas results in increased knowledge and overall leads to better research. In these cases, there are mutual benefits for the researchers that share data with one another, often in an informal collaborative way. Publications are typically co-authored, so there is a balance in that all researchers engaged in the exchange and sharing of data, benefit equally from this sharing. Where data are shared through such collaborative research, the main drivers are exchange of ideas and knowledge that will lead to better analysis and findings.

In other research domains, research depends on the availability of vast amounts of information and data that can be mined to find patterns and relationships that the human brain cannot pick up [NBIC]. Here, the researcher depends on the willingness for other researchers to make their data and information accessible, in a more altruistic manner. The primary researcher does not necessarily benefit from the sharing, other than via citation of data reuse. For this kind of data sharing, the researcher who needs the data is dependent on other researchers to share their data, without necessarily seeing immediate benefits. Data are typically accessed via community or public sharing.

Some researchers see the ultimate way to ensure that science is unbiased and rational, and stands up to criticism, is by publishing the raw data as direct evidence of the interpretations made by the researcher, i.e. the findings described in published articles [Adaptomics; Chem; LARM]. Best research practice is seen to be openness about the data that form the base material on which publications are based. Sharing the raw data then leads to the most efficient way to communicate science. This is in line with requirements of journals to submit supplementary data with publications.

Where researchers share their data in community repositories or via databases they develop as part of the research, they see the ability for other researchers to interpret their data in different ways as beneficial [Chem; LARM].

Some interviewees emphasise the importance for researchers - especially early career researchers - to learn and gain expertise from using (or seeing) other researchers' data [RetiredMen; LARM].

Some interviewed researchers were funded specifically to compile information resources into public databases or to digitise information into accessible formats, in order to create data and information resources for the research

community. Here creating community and public data resources **is** the research. Examples are BrassiBase [Adaptomics] and the LARM audio archive [LARM]. Overall, only a very small percentage of research funding is used directly to fund data sharing but it indicates that direct funding increases data sharing.

6.1.2. Data sharing to enhance the research career

Clearly, researchers expect that doing better research by sharing data, as described above, will also help their own careers. In addition, three more specific incentives were identified that could be considered as career-enhancing: visibility, reciprocity and reassurance.

Visibility

Sharing research data, either via publications, in community or public repositories, or through collaborations with trusted researchers, is seen by various researchers interviewed as being an important strategy to make their research and the entire research group more visible. It helps them to stand out from the crowd and to therefore attract more research funding and enhance career opportunities [Chem; Adaptomics]. It is part of building a reputation and getting recognised for new contributions to the field.

Indirectly, data sharing within the context of collaborative research is essential for building networks and collaborations, and therefore again to increase visibility and standing within the research discipline and improve career and funding opportunities.

In genomics research, a large-scale analysis of data sharing shows that studies that made data available in repositories received 9% more citations, when controlling for other variables; and that whilst self-reuse citation declines steeply after two years, reuse by third parties increases even after six years (Piwowar and Vision, 2013).

Reciprocity

Reciprocity featured in many descriptions of data sharing. In some fields, data is shared via direct exchanges between known and trusted researchers through peer exchange and collaborative sharing [Adaptomics]. Data are either traded for mutual benefit, or data are shared with the view that therefore in future data could be received in return, i.e. the data receiver is in debt to the data sharer. Even when data are shared through more formal repositories there can still be a strong sense of a responsibility to share, especially if one has benefitted from others sharing their data in the past.

Importantly, this experience of reciprocity was woven into career trajectories. Students and early career researchers described feelings of gratitude for data that had been shared with them. This contributed to their sense of responsibility, to return the favour and make their materials available to later generations [Chem; Adaptomics].

Reassurance

Many early career researchers express a reluctance to share data, not through disinclination, but from modesty. They believe their data cannot possibly be valuable to others. In some cases, research data were shared because they were specifically requested (in demand), either by other researchers, or by a data repository [Chem; RetiredMen]. This experience of being asked was powerful in making people overcome the feelings of inadequacy of their data. The reassurance that their data were good enough to be desirable to other researchers contributed to their sense of being recognised as legitimate members of their disciplines. This can be an important incentive for less confident early career researchers.

6.2. Cultural norms of the research group or discipline

The research culture within which an individual researcher operates has a very strong influence on his/her attitude towards data sharing. In some research disciplines such as genomics or crystallography, data sharing is so well established that it is taken for granted. It has become part of standard research practices. This applies to all degrees of data sharing. Many researchers state "sharing is just what we do / have always done"

[Chem; Adaptomics]. This is often the case where research handles such complex issues and datasets that collaboration across researchers is essential for being able to analyse the data [Adaptomics].

Equally there is a strong influence if junior researchers start their career within a research group where data sharing is standard practice [Chem; Adaptomics]. This sets the scene of how the research is done and emphasises the role of training and culture change for data sharing.

Data sharing as standard research practice also results in data practices that makes data sharing easier and more straightforward. The [Adaptomics] and [Chem] cases show that researchers share data in a hierarchical way from the start of their career. Bachelor and master students share data with doctoral students; doctoral students with postdoctoral researchers and supervisors, etc. This means that from the start of their research careers, researchers create and process their data in such way that colleagues - and therefore also other researchers - can easily understand the data.

Data are seen by some interviewees as having no research value if they cannot be fully exploited, which sharing facilitates [Adaptomics].

Sometimes a non-sharing culture can also motivate data sharing. In research disciplines where data sharing may not be well established, not be customary, or not be done at all, some researchers want to challenge these assumptions that data can or should not be shared [RetiredMen; Chem]. This may be the situation for junior researchers, who - influenced by a general open attitude towards information sharing through social media - feel that non-sharing attitudes are too conservative and unfounded. They apply their open attitudes towards research. Here non-sharing norms stimulate the sharing of research data.

6.3. External factors as extrinsic incentives

As explained in the background section, a range of policies, mandates, expectations and requirements from research funders and publishers may exist to share data. While these are not incentives per se, researchers are motivated to follow or comply with such demands, as a means to other objectives, such as receiving funding or getting papers published.

Many researchers cited journal requirements as currently having an influence on data sharing [Chem; Adaptomics]; fewer mentioned funder or institutional policies or data planning requirements as being strong incentives [Chem].

Whether or not journals and peer reviewers request data files that underpin published findings as supplementary material, to be submitted to the journal or to a repository, seems very discipline-specific [Chem; Adaptomics]. Such supplementary data are needed for verification and as quality control, both at the time of publication - when peer reviewers can assess how the findings were achieved, based on the raw or processed data - and in future for researchers to replicate or validate research. IT and data infrastructure plays an important role here too. In recent decades it has become much easier to store such supplementary files and make them available to journal audiences.

The availability of data support services to help researchers with preparing data for sharing, where they exist, are recognised as making data sharing easier for researchers [RetiredMen; Chem].

In some disciplines, learned societies play an active role in pushing data sharing forward, through development of infrastructure and resources, for example, the Royal Society of Chemistry actively developing community data resources such as ChemSpider and ChemSpider Synthetic Pages [Chem].

7. Future incentives for data sharing

Incentives that researchers feel would motivate more researchers to share their data, or would motivate researchers to share more of their data were described as follows, and are addressed further in the recommendations.

A strong theme that emerged in many interviews is a role for funders and other formal institutions such as learned societies to provide a level playing field for all researchers to share data and change the collective attitude towards sharing. Individual researchers, whilst recognising the overall benefits of data sharing for science, may not see the benefit of data sharing for themselves. Moreover, an individual researcher sharing data may be at a disadvantage, in spending time doing things that are not rewarded, or taking the risk of being outcompeted when sharing ideas freely. However, the field or discipline as a whole could benefit if everyone would share data. Funder and publisher policies that mandate or promote data sharing, or that set expectations for data sharing are seen as being able to provide such level playing field for increased data sharing and establish an important standard of fair and equal access to data. This also includes promoting requirements for data management planning.

Some researchers see the need for direct funding for research data management support, e.g. to facilitate data documentation and annotation. Such data management and sharing support is especially needed early in the research process.

Alongside this, the training of students in essential data sharing and management practices is seen as an important factor in increasing the incidence of data sharing. Researchers themselves state that textbooks or their studies never teach them about data sharing.

Data infrastructure and standards are indicated to be needed at various levels. This can be a research community agreeing on standards for data formats, documentation, etc.; and publishers setting standards and requirements for supplementary data files, to increase their usability, their quality and their longevity. Data infrastructure is needed in the form of repositories and publishers' infrastructure for supplementary data.

In biology, the need was felt for making all shared data openly accessible in the case of supplementary data files held by publishers. The data are often behind subscription walls alongside the articles, and therefore inaccessible to many researchers.

The sharing of failed experiments is mentioned as being of paramount importance in different research fields [Chem; Adaptomics].

The [Chem] and [NBIC] cases indicated that progress is needed in finding ways to share and cite research data in smaller segments or chunks, i.e. smaller than the supplementary files that are typically submitted with a refereed journal article or smaller than datasets typically deposited in a repository. Especially when sharing is linked with journal publications, this level of output represents a very high standard and much work to publish a paper. A theme in the interviews was the need for new forms of micro-sharing and micro-citation to be developed. The key element is to be able to generate, share and cite smaller outputs (micro- or even nano-bundles) or being able to cite subsets in a data collection and having such citations being taken into account in career progression and impact factors (see also Mons et al, 2011). Micro-shared data can be statements that are shared and published early in the research process to make initial ideas visible and accessible. These can also be factual extracts from larger publications or datasets. It is then also essential that such micro-publications do not jeopardise future acceptance of papers by journals.

Researchers who recognise the strong influence norms can have, advocate the need to create research environments where it is routine practice to share data.

8. Influence of policies and support services on data sharing

For most researchers interviewed, funder and institutional data policies are not currently very important.

None of the interviewed researchers was mandated by a funder or institution to share their research data, except where data sharing was directly funded: development of the LARM radio archive and BrassiBase. Overall, however, funding for database projects or data sharing is perceived to be difficult to obtain. Whilst researchers recognise an interest amongst funders and institutions for increased data sharing, they are perceived as unwilling to fund this.

Many researchers actively sharing data or needing data for their research want to see stronger leadership from research funders, with funders requiring data stewardship plans, data management plans and specifying data sharing requirements. Researchers indicate this as essential to provide a level playing field for data sharing or to change the mind-set of researchers. Other researchers feel differently, indicating that science should drive sharing [Adaptomics]. UK funder data policies and open access policies for scholarly publications are seen as inspiring examples by researchers in other countries. It is well recognised that funder data sharing policies need to go hand in hand with investment in data infrastructure and data services for researchers; and that together these can contribute strongly to establishing a culture of data sharing. The [LARM] and [Adaptomics] cases are examples of research groups with a strong sharing culture, where the funding of data infrastructure development by the research funder – the LARM audio archive and BrassiBase respectively – results in valuable data resources being made available to the wider research community; thereby boosting the sharing culture even more. A similar example is the Economic and Social Research Council in the UK; for years it has mandated the archiving of research data from all funded research projects. This policy goes hand in hand with the funding of supporting data infrastructure and services: the UK Data Service (previously the Economic and Social Data Service) provides the data infrastructure to curate, preserve and disseminate such research data, and provides data management training and support to researchers.

Research funding by industry and private companies typically limits the sharing of research data to protect commercial interests, but is also seen as essential by researchers to advance research [Adaptomics; Chem].

In some domains, learned societies are proactive in pushing data sharing through the development of community data infrastructure and standards, e.g. the Royal Society of Chemistry [Chem].

Researchers have mixed experiences with data policies or data requirements of journals, even in similar research domains. For some researchers, all journals they publish in require supplementary data to be provided as evidence; for other researchers that is not the case. Overall, supplementary data are not considered to be curated and not usually held in a repository. Some researchers feel strongly that all supplementary data should be open access, even if articles are not.

Patent applications require the release of supporting data and metadata, e.g. lab notebooks [Chem].

All these aspects show how key stakeholders in the research environment can positively influence data sharing, both through policies and requirements, as via infrastructure and services investments.

9. Discussion

9.1. Experiences of data sharing

This study, based on five research groups as case studies in five European countries, has gathered evidence, examples and opinions from selected researchers on what motivates them to share their research data; and what in their view would incentivise more data sharing in the future. Although the study was designed to provide maximum variation in research disciplines (across various academic disciplines: arts and humanities, social sciences, biomedicine, chemistry and biology), the study is not representative across all research disciplines or the entire academic community. Instead the aim was to gain in-depth insight into the motivations of individual researchers through detailed interviews.

Whilst the phrase 'data sharing' is frequently used to describe making one's research data available to other researchers, the study shows clearly that different researchers may mean very different things when they talk about data sharing. Different forms of data sharing are referred to, ranging from a very controlled sharing with trusted colleagues or peers, to the sharing of data via public repositories, where such data may be used by unknown researchers for unknown purposes. The different types of data sharing meant by interviewees are comparable to the six degrees of openness reported by Whyte and Pryor (2012): private management, collaborative sharing, peer exchange, transparent governance, community sharing and public sharing.

A significant finding in the study is that researchers have heterogeneous experiences with and views of data sharing that cut across disciplinary boundaries and case study circles. Rather than finding disciplinary or group patterns in data sharing practices, the interviews show how individual researchers across the case studies have similar data sharing practices and recognise similar incentives and motivations.

9.2. Different incentives for individuals vs. science

Whilst it is well known that most researchers appreciate the benefits of sharing research data for science in general, the benefits for an individual researcher can be weak or mixed. Researchers are often acting in ways that make sense for them, that is, not to share data. For example, they don't waste precious research time on making data sharable, and they don't risk getting scooped by exposing valuable data too early. Although everyone is acting sensibly according to their own interests, the collective result is that far less data is shared than everyone agrees would be optimal. Change can be induced through outside interventions such as policies, agreements and regulations; whereby all researchers are bound by the same rules. By setting standards, expectations and requirements for data sharing that apply to all researchers, a level playing field can be created, where no researchers are penalised for acting in ways that enhance the common good.

9.3. Incentives for sharing data

The study shows that researchers are motivated by three main types of incentives to share research data: direct benefits to the research or to their career; the strong influence of the sharing norms within the research circle; and external drivers such as funder policies, publisher requirements and support services provision.

One could argue that the best incentive for data sharing is where such sharing is essential for the research. The research itself then drives the need to share. This is seen clearly in the [Adaptomics] case study, where data are frequently shared in collaborative research, as such collaboration is needed to be able to analyse often complex data. Research becomes more efficient, and publications more frequent, the more data are shared with trusted researchers. Since in this case the need to share data is based on mutual benefits to all researchers that produce and use data, there is an equilibrium of incentives: data are both 'produced' by all researchers in the research circle and 'used' in analysis and publications for mutual benefit.

On the other hand, the [NBIC] case study shows that other kinds of research may depend equally heavily on the availability of large amounts of (structured) data and information, yet here, not all data users are also data producers and vice versa. This biosemantics research depends on researchers making data available via community sharing or public sharing, when they may not derive direct benefits themselves from reuse of such data by biosemantics researchers. Citation of data use alone may not provide a strong enough direct benefit to the researchers producing data, to share such data with the research community. Such data reuse cannot only depend on reciprocity. If data producers see no immediate direct benefit to share their data with unknown future users, other strategies or drivers are needed to encourage the sharing of data for the benefit of science. This will be in the form of policies or expectations set by research funders, learned societies and publishers. Therefore, relying on science alone to drive data sharing can never apply to all research disciplines. Other incentives are needed.

Career progression is a second form of direct benefit that has an important influence on data sharing. Interviewees state that increased visibility of their research and increased publications - especially following collaborative sharing - are important motivations in data sharing. At the same time, the current career progression which is entirely based on metrics for articles publications, fails to motivate, and in many cases penalises, data sharing. Altruistic data sharing can be perceived by researchers to harm career progression if time is spent on sharing data rather than on writing publications, or if shared data can be used by competing researchers for their research and publications. Changes are needed, with a demand by researchers keen to push data sharing further forward for alternative metrics on which career progression can be based (besides journal impact factors). Data citation metrics and social media citation metrics can also reflect impact of research, besides the traditional journal article citation metrics.

It is clear that the norms within a researcher's research circle have a very strong influence on whether or not data are shared. Many interviewees state that they share their research data because that is just what they do and have always done within their research group or community. When junior researchers develop their research practices within a group where data are routinely shared - within the group or with collaborating groups or peers - then that is the practice they adopt. When they reach the stage of developing their own research projects - winning and managing their own research budgets - their attitude may change slightly, as the need to protect their ideas and to secure publications and future funding may increase their incentives for holding onto valuable data themselves. But overall, the cultural norm is there and remains there, and they progress on to continuing to share research data. Changing norms within all research disciplines and within all research groups is thus an important factor. In some cases this will be driven by the direct benefits to the research. In other cases this needs external drivers in the form of expectations (and policies) set by the research community, the funders, the publishers, etc.

The finding that academics are not overly interested in the analytics of data reuse corresponds with similar findings with regards the sharing and reuse of open educational resources (McGill et al, 2013); instead, what is important is high quality reuse such as citations, personal feedback, finding collaborators, and additional questions being directed towards the original researcher.

9.4. Accounting for how incentives work

In trying to determine which incentives might be most effective, it is helpful to understand how and why they operate. Basically, we are trying to find levers to motivate data sharing. In all but a very few individuals, data sharing occurs in response to extrinsic motivation. It is not something done for its own sake, but to achieve other ends. Those ends can range from immediate career reward to altruistic goals of better science, but they are extrinsic. However, it is established that extrinsic

motivations are stronger when they are internalised, i.e. personally endorsed. Moreover, research suggests that this kind of endorsement is stronger when at least one of three factors is present: relatedness, competence, and autonomy (Ryan and Deci, 2000).

This study did not intend to explain why various incentives might be effective. Nonetheless, it is reassuring that the findings do seem to be consistent with existing accounts. Researchers' competence and autonomy are clearly promoted when data sharing enhances the research itself and the researchers' career. Relatedness, the sense of being a respected and valued member of a community, is supported when that community has data sharing as one of its norms. Researchers become recognised as valued members of the community by joining the practice of sharing.

9.5. Concluding remarks and areas for further research

All degrees or forms of data sharing are beneficial in research, but may be favoured by different actors or stakeholders. For example, funders are often keen on open sharing of all research data to maximise the benefits from their investments. But this is often not relevant for researchers themselves, who may be more in favour of collaborative sharing or peer exchange.

Leadership is needed from funders, institutions, learned societies and publishers, to jointly advance data sharing, to stimulate the direct benefits that researchers get from sharing their research data, and to ensure that data sharing becomes part of standard research practices - simply something every researcher does, just as they publish findings.

There will always need to be a 'mixed economy' of incentives that consider the different forms (degrees) of data sharing that exist. And within those different degrees of sharing, there should be no value judgement about which forms of data sharing may be best or preferred.

Collaborative sharing or peer exchange may directly benefit the research and motivate researchers via co-authorship. Community and public sharing may be essential to advance research beyond the realm of the primary research and to enable innovation. All kinds of sharing serve their purposes and fit within a particular context. What needs to be avoided, however, is the favouritism that can affect collaborative sharing or peer exchange, therefore excluding groups of researchers from fair access to data. Transparency is needed about all forms of data sharing, so any researcher knows under which conditions data may be accessible. Different stakeholder groups may be interested in promoting particular sharing modes.

While there is always a need for more research (we are researchers ourselves), on balance this work suggests that the best use of resources would be toward practical support for researchers already eager to share data. More general investigations of barriers, or even incentives, does not seem warranted. That said, within the domain of practical support, more knowledge about how to develop and implement certain tools and processes for data sharing could be beneficial.

This study indicates that further research is needed to explore practical options for micro- and nano-citation of research data in ways that benefits research. A recent article by Pröll et al. (2014) evaluates a model based on time-stamped queries for different data types and application domains. Another area of focus would be practical options for integrating research data into research assessment, e.g. via data portfolios. Finally, more investigation is still needed regarding ways to resolve tensions between Intellectual Property and confidentiality restrictions and data sharing.

10. Recommendations

Leadership is needed from various actors for research data sharing to become the norm in a wider range of disciplines, institutions and research groups.

For research data sharing to become more prominent and to become the norm within a wider range of research disciplines, institutions and research groups, and for there to be a level playing field so researchers can share their data with confidence and without the fear of being outcompeted, leadership is needed from various actors: research funders, learned societies, research institutions, publishers, data centres and others. The following recommendations for incentivising the sharing of research data result from the findings of this study, for those different actors. This categorisation of the recommendations does not imply that these actors should not collaborate in their efforts, indeed, collaboration is essential in many cases. Furthermore, the different actors may choose to focus on incentivising the mode(s) of data sharing that they want to see enhanced or that best suits their area of expertise. Different research disciplines and groups will always have different forms of data sharing, be it through informal collaborations, through publications, or by deposit into public repositories, that may need different approaches to incentivise.

10.1. For research funders

- All research funders to adopt a data sharing policy that clearly indicates expectations for data accessibility, in order to provide a level playing field with regards to data sharing for all funded researchers. Policies can consider measures such as requirements for data management planning and clear guidance on how a percentage of grants budgets can be allocated to data management for projects creating data with high potential reuse value
- Provide funding and support services to researchers where needed, e.g. for data documentation, annotation and data deposit. This should be similar to the funding of publication costs. Not all research disciplines have the same needs in this respect, as some data require more preparation than other to make them available for reuse. Also the type of data sharing influences what is needed, e.g. sharing raw vs. processed data; sharing data that supplements articles vs. sharing in repositories
- Focus data sharing funding towards two key intervention points:
  - early when research is being planned
  - upon completion of a research project, to prepare data and documentation for curation
- Continue to invest in data infrastructure that also provides rich context, detailed metadata and even a narrative account of the data creation. The kind of infrastructure researchers find most useful is where research data, papers and other outputs or resources are jointly available within a single data resource. Examples noted in this study are PubMed, TAIR, LARM, CCDC and ChemSpider
- Invest in the provision of data sharing training embedded into research methods training for students and doctoral researchers, to help establish data sharing as standard research methodology and practice

- Promote reuse of existing data resources via specific funding streams for secondary analysis and by setting expectations for research grant applicants to justify the need to create new data in research (i.e. to demonstrate that existing data cannot address their research questions)

- Engage with publishers and commercial partners on IP and copyright of data that may limit data sharing by creating a working group to find ways to protect IP and share data, especially when research is intended for non-commercial use

- Provide guidance to peer reviewers to evaluate data sharing plans and strategies in research proposals

10.2. For learned societies

- Promote discussion of formal research recognition for data sharing and data publishing

- Set clear data sharing expectations for respective research disciplines, e.g. through codes of conduct or best practice codes

- Promote the development and uptake of data sharing agreements within specific research disciplines, that stipulate agreement over how research data can be shared in a timely and open manner (e.g. similar to the 1996 Bermuda Principles and 2003 Fort Lauderdale agreement over prepublication data release in genomics), to provide a level playing field to all researchers

- Promote the development of data sharing resources and standards for the research discipline

10.3. For research institutions

- Formally recognise and value data, alongside publications, as part of research assessment and career advancement

- Incorporate data impact into PhD career assessment, e.g. via a system of portfolio assessment where research data may be one element alongside other research outputs that provide evidence for research impact, or via a data CV

- Provide training in research data sharing to students, embedded into methods training, so data sharing becomes part of standard research practice

- Set expectations for data sharing for researchers within the institution

- Provide integrated support services to researchers, e.g. a one-stop-shop for all research data management and sharing guidance

10.4. For publishers

- Strengthen direct career benefits to researchers to share their data via data citation and data sharing metrics. This should provide what researchers in particular disciplines call for, e.g. micro-citation, micro-publications, data publishing with Digital Object Identifiers (DOIs), data citations to link with ORCIDs (Open Researcher and Contributor ID), and digital watermarking of data files to provide provenance of data

- Ensure that publishing terms and agreements of manuscripts do not create disincentives for micro-publishing of data, e.g. through overly restrictive requirements for manuscript content to be new

- Journals and innovative publishers to explore and actively encourage publication of negative findings, failed experiments, etc.

- Request that all data related to a published manuscript are made available, not only the data supporting the published results
- Set open or preservation standards for data formats, file formats, and supplemental documentation
- Ensure that data is fully and properly cited in all publications, and provide clear instructions to editorial staff and reviewers to check for correct data citation
- Make all supplementary data available openly (free of charge), even if the article is not

10.5. For data centres and repositories

- Develop and encourage pull factors for data sharing such as actively inviting researchers to share data by deposit or by other channels. Researchers feel valued and reassured about data quality when their research data are in demand
- Deliver specialist data sharing training for researchers on IP, copyright, technical standards and metadata
- Develop and provide flexible systems of providing access to data, allowing data owners to set controls where this is needed, e.g. embargo periods, defined access groups, etc.
- Provide data resources that combine data and related rich context such as publications and other outputs

10.6. For Knowledge Exchange

- Invest and engage in the development of data infrastructure with rich context or invest in infrastructure within the member's remit that underpins and allows these other infrastructures to flourish. The kind of infrastructure researchers find most useful is where research data, publications and other outputs or resources are jointly available within a single data resource. Examples noted in this study are PubMed, TAIR, LARM, CCDC and ChemSpider
- Explore and develop mechanisms for micro- and nano-citation of research data
- Lead the development of data sharing strategies and data sharing expectations at a national level, in collaboration with the various stakeholders
- Push for development of data infrastructures and data services at a European level
- Push for the recognition of data sharing in career progression at a national and European level
- Call upon relevant stakeholders to provide data sharing training to all undergraduate students
- Fund practical solutions for data sharing, possibly jointly with research funders [recommendations for funders]
- Develop national registries of research data that link to journals, repositories, etc.

Literature references

(all accessed 9/10/2014)

- Alliance of German Science Organisations (2010), **Principles for the Handling of Research Data**. allianzinitiative.de/en/core-activities/research-data/principles.html
- Berlin Declaration (2003), **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities**. Max Planck Society. oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung
- Borgman, C. L. (2012), **The conundrum of sharing research data.** Journal of the American Society for Information Science and Technology 63: 1059-1078. dx.doi.org/10.1002/asi.22634
- Corti, L., Van den Eynden, V., Bishop, L and Woollard, M. (2014), **Managing and Sharing Research Data: a Guide to Good Practice.** Sage Publications Ltd. uk.sagepub.com/books/9781446267264
- Costas, R., Meijer, I., Zahedi, Z. and Wouters, P. (2013), **The Value of Research Data - Metrics for datasets from a cultural and technical point of view**. A Knowledge Exchange Report. knowledge-exchange.info/datametrics
- de Cock Buning, M., van Dinther, B., Jeppersen de Boer, G.G. and Ringnalda, A. (2011), **The legal status of research data in the Knowledge Exchange partner countries**. Centre for Intellectual Property Law (CIER), Molengraaff Institute for Private Law, Utrecht University. knowledge-exchange.info/Default.aspx?ID=461
- EPSRC (2011) **EPSRC Policy Framework on Research Data,** Engineering and Physical Sciences Research Council. epsrc.ac.uk/about/standards/researchdata
- Expert Advisory Group on Data Access (2014), **Establishing incentives and changing cultures to support data access**, wellcome.ac.uk/stellent/groups/corporatesite/@mshpeda/documents/webdocument/wtp056495.pdf
- European Commission (2010), **Riding the wave: How Europe can gain from the rising tide of scientific data.** Final report of the High Level Expert Group on Scientific Data. European Commission. cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf
- European Commission (2012a), **Communication: Towards Better Access to Scientific Information: Boosting the Benefits of Public Investments in Research.** European Commission. ec.europa.eu/research/science-society/document_library/pdf_06/ era-communication-towards-better-access-to-scientific-information_en.pdf
- European Commission (2012b), **Recommendation on Access to and Preservation of Scientific Information**, European Commission. ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf
- Force11 (2013), **Joint Declaration of Data Citation Principles**. force11.org/datacitation
- G8 UK (2013) **G8 Sciences Ministers Statement, 12 June 2013**. gov.uk/government/uploads/system/uploads/attachment_data/file/206801/G8_Science_Meeting_Statement_12_June_2013.pdf

» Hrynaszkiewicz, I. and Cockerill, M.J. (2012), **Open by default: a proposed copyright license and waiver agreement for open access research and data in peer-reviewed journals**, BMC Research Notes, 5: 494. doi:10.1186/1756-0500-5-494

» KNAW (2011), **KNAW policy on open access and digital preservation**. knaw.nl/en/topics/open-access-and-digital-preservation/open-access/policy

» Knowledge Exchange (2013), **Data sharing and reuse policies - a (brief) comparative overview from the Knowledge Exchange countries.** knowledge-exchange.info/Files/Filer/downloads/Documents/Presentations/130925%20KE%20DARUP%20comparative%20overview%20final.pdf

» Koch, M.A., Kiefer M., German, D.A., Al-Shehbaz, I.A., Franzke, A., Mummenhoff, K. and Schmickl, R. (2012), **BrassiBase: Tools and biological resources to study characters and traits in the Brassicaceae – version 1.1,** Taxon 61(5): 1001-1009.

» Koch, M.A., Scheriau, C., Betzin, A., Hohmann, N. and Sharbel, T.F. (2013), **Evolution of cryptic gene pools in *Hypericum perforatum*: the influence of reproductive system and gene flow.** Ann Bot (2013) 111 (6): 1083-1094. doi:10.1093/aob/mct065

» Lyon, L., Rusbridge, C., Neilson, C. and Whyte, A. (2010), **Disciplinary Approaches to Sharing, Curation, Reuse and Preservation:** DCC SCARP Final Report to JISC. Edinburgh: Digital Curation Centre. dcc.ac.uk/sites/default/files/documents/scarp/SCARP-FinalReport-Final-SENT.pdf

» McGill, L., Falconer, I., Dempster, J.A., Littlejohn, A. and Beetham, H. (2013), **Journeys to Open Educational Practice: Impact.** UK OER/SCORE Review Final Report. oersynth.pbworks.com/w/file/67270745/OER%20Review%20impact.pdf

» Mons, B., van Haagen, H., Chichester, C., 't Hoen, P., den Dunnen, J.T., van Ommen, G., van Mulligen, E., Singh, B., Hooft, R., Roos, M., Hammond, J., Kiesel, B., Giardine, B., Velterop, J., Groth, P. and Schultes, E. (2011), **The value of data.** Nature Genetics 43(4): 281-283. doi:10.1038/ng0411-281

» NatCen (2012), **The Framework approach to qualitative data analysis.** surrey.ac.uk/sociology/research/ researchcentres/caqdas/files/Session%201%20Introduction%20to%20Framework.pdf.

» OECD (2007), **OECD Principles and Guidelines for Access to Research Data from Public Funding.** Organization for Economic Co-operation and Development. oecd.org/dataoecd/9/61/38500813.pdf

» Open Data Exchange (2011), **Ten Tales of Drivers and Barriers in Data Sharing.** Opportunities for Data Exchange Project. alliancepermanentaccess.org/wp-content/uploads/downloads/2011/10/7836_ODE_brochure_final.pdf

» Piwowar, H.A. (2011), **Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data**. PLoS ONE 6. plosone.org/article/info:doi/10.1371/journal.pone.0018657

» Piwowar, H.A. and Vision, T.J. (2013), **Data reuse and the open data citation advantage**. PeerJ 1:e175 http://dx.doi.org/10.7717/peerj.175

» Pröll, S., Rauber, A., Jones, C., Corti, L., Watkins, J. and Barateiro, J. (2014), **Providing Concise Data Citations to Enable Reproducibility Across eScience Disciplines.** IEEE Transactions on Emerging Topics in Computing (forthcoming).

» RCUK (2011), **Common Principles on Data Policy**, Research Councils UK. rcuk.ac.uk/research/datapolicy

» RCUK (2012), **Research Councils UK Policy on Open Access,** Research Councils UK. rcuk.ac.uk/research/openaccess/policy

» RECODE (2013), **Policy Recommendations for Open access to Research Data in Europe - Stakeholder values and ecosystems**. http://recodeproject.eu/wp-content/uploads/2013/10/RECODE_D1-Stakeholder-values-and-ecosystems_Sept2013.pdf

» RECODE (2014), **Policy Recommendations for Open access to research Data in Europe - Institutional barriers and good practice solutions**. http://recodeproject.eu/wp-content/uploads/2014/09/RECODE-D4.1-Institutional-barriers-FINAL.pdf

» Ryan, R. and Deci, E. (2000), **Intrinsic and extrinsic motivations: classic definitions and new directions.** Contemporary Educational Psychology 25:54-67. doi:10.1006/ceps.1999.1020

» Sandelowski, M. (1995), **Focus on qualitative methods: sample size in qualitative research.** Research in Nursing and Health, 18: 179-183.

» Savage, C.J., and Vickers, A.J. (2009), **Empirical study of data sharing by authors publishing in PLoS journals.** PloSOne, 4(9): e7078.

» Sayogo, D.S. and Pardo, T.A. (2013), **Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data.** Government Information Quarterly, 30(1): 19-31. doi 10.1016/j.giq.2012.06.011

» Sharbel T.F., Voigt M.L., Corral J.M., Galla G., Kumlehn J., Klukas C., Schreiber, F., Vogel, H. and Rotter, B. (2010), **Apomictic and sexual ovules of *Boechera* display heterochronic global gene expression patterns**. Plant Cell 22, 655–671 doi:10.1105/tpc.109.072223

» Tizzard, G., Threlfall, T., Ling, S and Coles, S (2011), **(E)-2,4,5-trifluoro-N-(4-fluorobenzylidene)aniline.** University of Southampton, Crystal Structure Report Archive. http://dx.doi.org/10.5258/ecrystals/1424

» TTA (2014). **Policies of higher education institutions and funding agencies.** tdata.fi/en/korkeakoulujen-ja-rahoittajien-politiika?param=param.

» Tenopir, C., Allard, S., Douglass, K., Aydinoglu ,A.U,, Wu, L., Read, E., Manoff, M., and Frame, M. (2011), **Data Sharing by Scientists: Practices and Perceptions**. PLoS ONE 6. doi:10.1371/journal.pone.0021101 plosone.org/article/info:doi/10.1371/journal.pone.0021101

» The Royal Society (2012), **Science as an Open Enterprise**. The Royal Society Science Policy Centre Report 02/12. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf

» UK Data Archive (2014), **UK Data Archive model consent form.** data-archive.ac.uk/media/210661/ukdamodelconsent.doc

» Van den Eynden, V. and Bishop, L. (2014), **Interviews of incentives for sharing research data. [Data Collection]**. Colchester, Essex: UK Data Archive. http://reshare.ukdataservice.ac.uk/id/eprint/851540

» Wallin, A. (2011), **Gatherings of Finnish Male Pensioners on a Market Square 2011 [computer file]**. FSD2842, version 1.0. Tampere: Finnish Social Science Data Archive [distributor]. https://services.fsd.uta.fi/catalogue/FSD2842?tab=summary&lang=en&study_language=en

» White, D. and Manton, M. (2011), **Open Educational Resources: The value of reuse in higher education**. University of Oxford. http://www.jisc.ac.uk/media/documents/programmes/elearning/oer/OERTheValueOfReuseInHigherEducation.pdf

» Whyte, A. and Pryor, G (2011), **Open Science in Practice: Researcher Perspectives and Participation.** International Journal of Digital Curation 1(6): 199-213. http://ijdc.net/index.php/ijdc/article/viewFile/173/241

» Wicherts, J. M., Borsboom, D., Kats, J., and Molenaar, D. (2006), **The poor availability of psychological research data for reanalysis**. American Psychologist, 61(7): 726.

» Youngseek, K. and Stanton, J.M. (2012), **Institutional and Individual Influences on Scientists' Data Sharing Practices.** Journal of Computational Science Education, 3(1): 47-56.

» ZonMw (2014), **Toegang to data**. zonmw.nl/nl/programmas/programma-detail/toegang-tot-data-ttdata/algemeen/

Annexe 1. Interview question topics

Interviews will be carried out as semi-structured interviews. The questions below are a guiding list of questions for the interviews.

1. The kind of research the interviewee does [and his/her team] ; which entities typically fund this research (national, international funder); collaborative research with other institutions/countries?
2. What types of research data the interviewee creates / generates during research? Explore also data formats.
3. Does interviewee use any data produced by other researchers? This can be data that result from other research, but also from heritage collections, such as broadcasting data, radio, text, video, etc.
4. If so, how does interviewee access such data / information, i.e. what are the data sharing circles ?
5. Does interviewee share any of the research data he/she creates? If so, which types/formats of data? How are data shared ? Within which context ? And with whom ? What is the data sharing circle?
6. Why does interviewee share research data? [What motivates him/her to share...]
7. Explore data sharing policies in the country / institution, and how policies relate to practices for data sharing; also explore policies of research funders that may apply (incl. international funders); explore whether international collaboration (with different policies in different countries) influences data sharing...
8. Is data sharing and data reuse being tracked by the interviewee?
9. How long has interviewee been sharing data (e.g. influence of expertise, training,...)?
10. When in research cycle are data being shared?
11. Explore funding models for data sharing and how data sharing is funded
12. Would interviewee prefer to share more/less data ?
13. What are the barriers or obstacles to data sharing for interviewee / for institution / for research domain?
14. What conditions are placed (or would like to place) on sharing of research data?
15. Which incentives currently apply to sharing research data and data sharing rewarded in research assessment / career promotion?
16. Which incentives (see above; and others) would encourage you to share more research data? Explore policy and practice developments and incentives
17. Would support in any of these areas make any difference?
18. Are there better or worse times during your research when support for data sharing would most help? (grant allocation, acceptance of publication by journal, ...)
19. Are there any specific policy or other changes that would help?
20. Anything else we haven't touched on so far?

Annexe 2. Interview summary sheet

	Topic	(insert interviewee name)
1.	Kind of research (funder, collaborations)	
2.	Types of research data	
3.	Use of existing data	
4.	If so, how access such data / information	
5.	Sharing research data (types, formats, how, with whom)	
6.	Why share research data (motivations)	
7.	Data sharing policies in country / institution, and practices for data sharing	
8.	Data sharing and data reuse tracking	
9.	How long sharing data	
10.	When in research cycle are data shared	
11.	Funding models for data sharing	
12.	Prefer to share more/less data	
13.	Barriers or obstacles to data sharing	
14.	Conditions placed on sharing of research data	
15.	Current incentives for sharing research data and data sharing rewarded in research assessment / career promotion	
16.	Which incentives would encourage data sharing	
17.	Would support make any difference	
18.	Optimal times during your research for support for data sharing	
19.	Policy or other changes that would help	
20.	Other	

Annexe 3. Data repositories and resources mentioned or used by interviewed researchers

- BrassiBase: online knowledge and database system on Brassicaceae taxonomy, systematics and evolution, including chromosome numbers, traits and characters, germplasm resources
- Cambridge Crystallographic Data Centre (CCDC): repository of experimentally determined organic and metal-organic crystal structures and provider of related applications software
- ChemSpider: chemical structure database
- ChemSpider Synthetic Pages: database of practical procedures for research workers in synthetic chemistry
- Danish Radio Archive
- Data Fairport, an initiative for data publishing to ensure that valuable scientific data is FAIR in the sense of being Findable, Accessible, Interoperable and Re-usable
- Dryad: repository for data files linked to peer-reviewed publications for any research discipline
- eCrystals: archive for crystal structures generated by the Southampton Chemical Crystallography Group and the UK National Crystallography Service
- ePrints Soton: University of Southampton institutional research repository, for data and publications
- European Nucleotide Archive (ENA): database of nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation, at the European Molecular Biology Laboratory (EMBL)
- Figshare: cloud-based online repository where researchers can preserve and share their research outputs, including datasets
- Finnish Social Sciences Data Archive (FSD)
- Genbank: genetic sequence database at the National Center for Biotechnology Information (NCBI)
- GitHub: web-based repository for sharing source code that supports distributed revision control
- LARM.fm: web interface of the LARM Audio Research Archive, providing access to Danish radio broadcasts
- Netarkivet.dk: archive preserving the Danish internet (websites)
- PubMed: citations for biomedical literature
- The Arabidopsis Information Resource (TAIR): a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*
- UK National Crystallography Service: data collection and crystal structure analysis for the UK chemistry community, including the use of the UK synchrotron Diamond

Annexe 4. Consent form

Interview consent form

The use of incentives-case studies on current and future incentives for research data sharing

This study is carried out by the UK Data Archive, University of Essex, and funded by Knowledge Exchange, a co-operative effort that supports the use and development of information and communications technologies infrastructure for higher education and research.

- ☐ I have read and understood the project information and agree to take part in the study.
- ☐ I have had the opportunity to ask questions about the study.
- ☐ I understand that my taking part is voluntary and I can withdraw from the project whilst it is ongoing.
- ☐ I agree to be interviewed and my contributed information can be used in research outputs and publications by the UK Data Archive, by Knowledge Exchange and partners, whereby I may be quoted by name (I can indicate off-the-record information during the interview).

The interviews will be archived at the UK Data Archive and Knowledge Exchange and disseminated so other researchers can reuse this information for research and learning purposes:

- ☐ I agree for the audio recording of my interview to be archived and disseminated for reuse
- ☐ I agree for the transcript of my interview to be archived and disseminated for reuse
- ☐ I agree for any photographs of me taken during interview to be archived and disseminated for reuse

- ☐ I agree to be contacted in future by Knowledge Exchange and partners to participate in data sharing promotion events.

Name of participant [printed]	Signature	Date

Project contact details for further information:
Veerle Van den Eynden & Libby Bishop
UK Data Archive
University of Essex

Annexe 5. Information sheet

The use of incentives-case studies on current and future incentives for research data sharing

What is the project about?
This project was commissioned by **Knowledge Exchange** (knowledge-exchange.info) to investigate current incentives for research data sharing. KE is a co-operative effort that supports the use and development of information and communications technologies infrastructure for higher education. The research will involve qualitative interviews with five research teams in partner countries of Knowledge Exchange (Finland, Denmark, Germany, United Kingdom, and the Netherlands). The aim of this pilot study is to provide evidence and examples of useful incentives for data sharing from the researchers' point of view to inform scientists and policy makers.

How will you be involved?
You will be interviewed by a member of our research team about your attitudes, practices and experiences with sharing research data. We will provide you with a list of topics in advance. We intend to conduct all interviews in person, but some may need to be by phone. You may choose to have the interview in English or in your nation's language. The interview will probably take 30-60 minutes and will be recorded and then transcribed.

How will we address confidentiality and security?
We hope you will be comfortable using your real name for this research. You can indicate during the interview if you prefer certain information to be off-the-record and therefore anonymous. During the project, only the research team will access your contributed information. Any others, such as our transcriber, will be required to sign a non-disclosure agreement.

We expect to use your contributed information in various outputs, including a report and content for a website. Extracts of interviews and some photographs may both be used. We will get your permission before using a quote from you or a photograph of you.

After the project has ended, we intend to archive the interviews at the **UK Data Archive** (data-archive.ac.uk/) at the University of Essex in the United Kingdom and with Knowledge Exchange. Then the interview data can be disseminated for reuse by other researchers, for research and learning purposes.

Can I withdraw from the project?
Your participation is completely voluntary and you may withdraw from the research whilst the project is ongoing. If you should decide to withdraw after any materials have been published, we can remove your information from any copies we control, but we cannot ensure that copies will not persist in other locations.

What are the benefits and risks?
The project benefits include better understanding of researchers' motivations for sharing data. Such understanding will inform data policies and the design of infrastructures for sharing. We do not foresee any risks of participating in this project.

Where can I get more information?
The two researchers running the project are Libby Bishop and Veerle Van den Eynden, both at the UK Data Archive, University of Essex. You may contact Libby at xxx and Veerle at xxx. The contact for the project funder, Knowledge Exchange, is Angela Holzer.

Acknowledgements

This research project was carried out in dialogue with Knowledge Exchange and in collaboration with local interviewers in the partner countries: Anders Conrad (DK), Damien Lecarpentier and Irina Kupiainen (FIN), Jens Nieschulze and Juliane Steckel (GE), and Joeri Nortier (NL).

The authors thank Knowledge Exchange for funding the study, with special thanks to Angela Holzer, Irina Kupiainen and Bas Cordewener.

Thanks also to all researchers interviewed: Antti Wallin, Barend Mons, Bente Larsen, Bill Brocklesbury, Cerys Willoughby, Jacob Thøgersen, Janne Nielsen, Jeremy Frey, Marcus Koch, Markus Kiefer, Morten Michelsen, Nora Hohmann, Oliver Brand, Phil Adler, Richard Whitby, Simon Coles, Stefan Wötzel, Susanne Coles, Timothy Sharbel, Ute Krämer, Will Fryson and Xiaoping Tang.


UK • DATA
ARCHIVE