

**'Barriers to uptake and use of data sharing systems at the
University of Auckland - Identification of differences in
researcher and academic librarian perceptions'**

by

Joanne Leigh Simons

Submitted to the School of Information Management,
Victoria University of Wellington
in partial fulfilment of the requirements for the degree of
Master of Information Studies

October 2016

Abstract

Research Problem: To identify the differences in researcher and academic librarian perception of barriers to the uptake of data sharing systems by researchers at the University of Auckland in order to address all possible barriers during implementation and improve researcher use of new systems.

Methodology: This study has a cross-sectional research design, using a mixed methods research strategy, in particular a sequential exploratory design where preliminary interviews with researchers and academic librarians informed the construction of an online survey tool distributed more widely to researchers and academic librarians within the University of Auckland. Statistical significance testing was performed using the Mann-Whitney U test.

Results: 103 survey responses were received from researchers, and 18 from academic librarians. There were observable differences in the relative impact assigned by researchers and librarians to the various factors that influence researcher decisions to share data. There are also significant differences in the perceptions of barriers to data-sharing between research disciplines.

Implications: There may be a need to improve communications between the library and researchers with regards to the tools and services that they can offer. Library staff may need additional training in support of University researchers, as a proportion did not feel confident answering questions about researcher data-sharing. The research discipline differences in perceptions of barriers to data-sharing mean that a “one-size fits all” strategy for education in and marketing of these services will not be the most effective strategy to address concerns and increase researcher engagement.

Table of Contents

Introduction	4
Topic Statement	5
Literature Review	6
Research Data Management	6
Academic librarian views on barriers to research data sharing	7
Researcher views on barriers to research data sharing	8
Research questions and hypotheses	11
Research Design	11
Methodology and methods	12
Phase 1 – Preliminary Interviews	12
Phase 2 – Survey	13
Population and sample	13
Data collection	14
Data analysis	15
Ethical considerations	16
Publication venues	16
Results	17
Interviews	17
Survey construction and pilot study	18
Comparison of Researcher and Librarian perceptions of barriers to data-sharing ..	19
Comparison of the perceived barriers to research data-sharing by researcher discipline	20
Discussion	29
Limitations	37
Findings and recommendations	38
Future work	39
Conclusion	40
References	41
Appendices	44

Introduction

Scientific data is the foundation of medical, engineering and scientific knowledge. Immense data sets are generated in scientific research, some parts used to produce research publications and some not. Published and unpublished data could reasonably be expected to be available for further study. Funders invest huge sums of money annually, and are increasingly requiring research data management (RDM) and data sharing as a condition of project funding (Kahn, Higgs, Davidson & Jones, 2014). A recent investigation of over 500 papers published in biology journals showed that the papers where data was extant dropped by 17% per year after publication (Vines et al., 2014), emphasising that researchers cannot be relied upon as sole preservers of research data.

Although many university libraries in the USA have integrated RDM into the research process (University of Central Florida Libraries Research Lifecycle Committee, 2012, Figure 1), RDM systems at New Zealand's tertiary education providers are mostly at the preliminary stages of development (Corral, Kennan & Azfal, 2013).

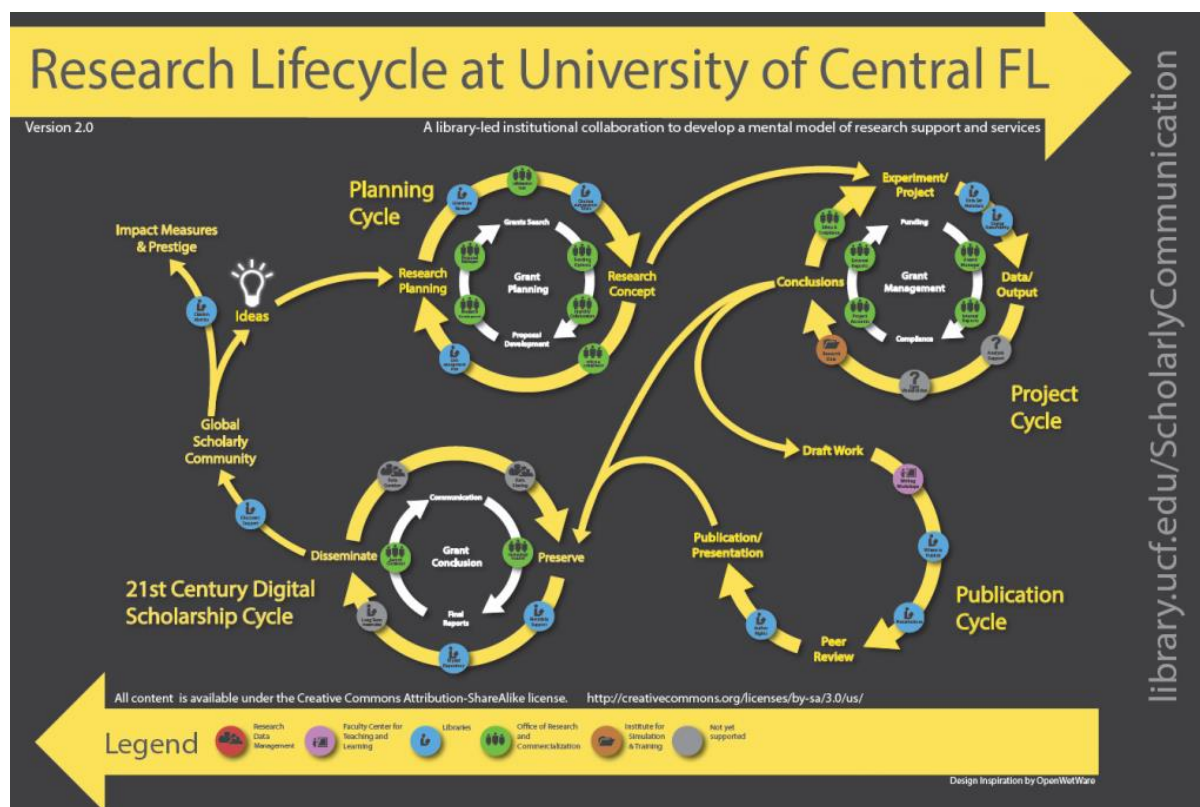


Figure 1. Research lifecycle at the University of Central Florida. The blue dots indicate library support services available at all parts of the research data cycle including data management planning, digital stewardship, metadata support and the development of data sharing services (University of Central Florida Libraries Research Lifecycle Committee, 2012).

Studies have indicated that researchers are often unable or unwilling to make their data accessible (Savage & Vickers, 2009; Vines et al. 2013; Nelson, 2009) and different fields have very different approaches to how they treat data (National Academy of Science, National Academy of Engineering, Institute of Medicine [NAS, NAE, IOM], 2009a). This could add levels of complication to implementation and uptake of RDM systems in universities.

Topic Statement

This research proposes to identify the differences in researcher and academic librarian perception of barriers to the uptake of data sharing systems by researchers at the University of Auckland in order to address all possible barriers during implementation and improve researcher use of new systems.

Although researchers are individually responsible for the management of their research data, library staff are heavily involved in the development and implementation of these systems, as well as promoting the systems to researchers and educating researchers in their use. Gathering the views from both sides will more effectively inform library and ICT department decisions on how systems should be developed or chosen, as well as policy around their implementation and use. Identification of any differences in perception, and knowledge of the barriers that have the strongest negative influence on researcher use of data sharing systems will enable these factors to be addressed during promotion and education activities. They may also mean that additional training of library staff will be required before promotion of the system to research staff is initiated. If barriers to use of a data sharing system are neutralised as much as possible during the early stages of implementation, it is expected that this would increase researcher uptake and use of the system. This, in turn, would increase the long term funding opportunities of research in the university.

Literature Review

Research Data Management

The development of the digital age has meant that researchers in the sciences and other fields have been presented with the opportunity to generate very large sets of data during their research (Boulton, 2012). This in turn has meant additional responsibilities for the management of such large sources of information. There are three main phases to research data management: maintenance of data integrity, data sharing and data stewardship (NAS, NAE, IOM, 2009b). Maintenance of data integrity is the responsibility of the researcher and their collaborators, and is carried out during the research project until publication of the data. Data stewardship is the long-term preservation of and maintenance of access to research data and is often carried out at an institutional or national level. National initiatives have been established in the UK and Australia to facilitate data curation and management (Simons & Richardson, 2013). Data sharing, the focus of this study, is the process of making available all of the raw data supporting publications that arise from a research project.

The reasons for making research data available through data sharing are two-fold. Firstly it makes research data part of the scholarly record that can be validated and tested by other researchers. Secondly, it makes the data available so that others can make use of it in new research, sometimes in unrelated fields of study (Swan & Brown, 2008).

Due to their significant financial investment into research on an annual basis, funding bodies are increasingly requiring a data management plan as part of a researcher's funding application (Kahn, Higgs, Davidson & Jones, 2014). This is particularly widespread in Europe and the United States, where there is a general acceptance that research carried out or sponsored by public agencies should be publically available (Uhlir & Schröder, 2007). In the UK, most public funders encourage data sharing and some such as the Economic and Social Research Council require it (Mauthner & Parry, 2013). Guidelines for managing open access to research data and information on best practice is readily available (Organisation for Economic Co-operation and Development, 2007; Murray-Rust, P., Neylon, C., Pollock, R., & Wilbanks, J., 2010).

Some science disciplines such as physics, mathematics and computer science have embraced data sharing, and the archive ArXiv hosted at Cornell University is widely used (Ginsparg, 2011). Although the main role of ArXiv is the sharing of publication preprints, researchers are increasingly using the system to share datasets. Data sharing is relatively rare in some other disciplines, particularly the social and natural sciences (Swan & Brown, 2008), although open access repositories such as Dryad are available for their use (Dryad, 2015). It is commonly accepted that a major barrier to data-sharing is that smaller institutions cannot afford the infrastructure to provide data sharing (Heidorn,

2008) and it has been suggested that technical barriers prevent scientists from uploading data into repositories (discussed in Mauthner & Parry, 2013).

Academic librarian views on barriers to research data sharing

Academic librarians have an important role in data sharing of research from their academic institutions and have thus been interested in identifying the barriers to use of data sharing systems by researchers. However, most studies led by librarians have involved surveying other academic libraries and librarians.

Tenopir, Sandusky, Allard and Birch (2014) used a large scale survey to investigate the involvement of North American academic research libraries in providing RDM services. Rather than directly addressing researcher uptake of data sharing, this study looked at barriers to provision of these services. It concluded that training and upskilling of librarians was a significant barrier to providing research data services. As academic librarians perform a role in promotion of data and bibliographic services to researchers, a lack of confidence on the part of librarians with regard to understanding data sharing could negatively impact researcher uptake of the service.

In a more geographically relevant study, the status of RDM services offered by universities in NZ, Australia, Ireland and the UK showed that less than 26% of the universities in these countries were currently offering any RDM services, although the Australian universities were leading the way (Corrall, Kennan & Afzal, 2013). An earlier study by Brown (2010) had focussed only on NZ libraries but had a very low response rate and showed little development of RDM. Corrall et al. (2013) surveyed academic librarians by questionnaire and had a relatively high response rate. The NZ respondents were most concerned by the different levels of demand for RDM due to low researcher uptake in some areas, and the technical issues of supporting different data needs across research disciplines.

A study of South African library professionals was carried out at an industry workshop to gauge the attitude of attendants to RDM policy and data sharing issues (Kahn et al., 2014). This paper was co-written and interpreted by members of the Digital Curation Council from the UK who have considerable experience in data curation at research institutions. Responses were measured by clicker devices issued to all workshop participants, and responses were viewed on a screen in real time, which could affect the way that participants answered. Regardless of the methodological issues, the data showed that the library participants had little knowledge of data management and did not believe it should be part of their jobs. If required to promote data sharing to researchers this would certainly affect their effectiveness and could negatively impact uptake of RDM systems.

Pinfield, Cox and Smith (2014) carried out a study that differed from previous work on RDM in libraries, in that it used 26 semi-structured interviews to gather information from staff at UK university libraries involved in RDM. This study aimed to analyse the library's contribution to data management and sharing within the institutional context, and to identify the main drivers of data sharing, and the factors that influence the development of RDM systems. Participants were selected from respondents to a previous survey on RDM, which may have skewed the results somewhat as all participants had already shown their engagement with data management systems. These respondents were also concerned with technical issues of storage, security and preservation, but the study showed that other possible barriers to use of data sharing were issues with compliance, jurisdiction and quality of research activity carried out with the data shared. The library staff also indicated a concern with different cultures and working practices across disciplines, as mentioned previously by Corral, Kennan and Afzal (2013). These factors all contributed to the development of a model by Pinfield et al. (2014) that attempts to address the multiple layers of stakeholders, activity and drivers of RDM and their interactions in a university environment.

A 2013 study by Williams (2013) was a library directed study that actually targeted researchers. They also used interviews and indicated that barriers to data sharing were the large amount of time and effort that it took to prepare data for submission, and then to submit it to a repository. The study specifically targeted members of the crop science faculty that already participate in data sharing, and so it did not address reasons why researchers might refrain from submitting data at all.

Researcher views on barriers to research data sharing

Although there is reluctance on the part of some researchers to share data, a study by Vines et al. (2013) has shown that sharing of research data is facilitated when journals mandate submission as part of their publishing policy rather than simply stating that they encourage data sharing. However, no journal had a data sharing rate of greater than 27% regardless of their policy. This was supported by earlier work by Savage and Vickers (2009) who found that nine out of ten researchers refused to supply research data when directly approached, despite having published in the journal PLoS which has a mandatory data sharing policy. Respondents indicated that barriers to sharing were the amount of time and effort required and institutional policies against data sharing. An international survey of scientists producing long-term datasets (Mills et al., 2015) has indicated that the type of data-sharing systems required for sharing is also an issue for some researchers, with institutional repositories preferred over mandatory submission to public databases.

Swan and Brown (2008) produced a report on data sharing commissioned by the Research Information Network in the UK. They carried out over 100 interviews with researchers and data managers in different academic fields but did not separate the responses of researchers and librarians. They found that differences between research fields were often more important than national differences in determining data-sharing practices. They identified researcher desire to retain exclusive use of their data until all publication value was extracted as the main barrier to data sharing. This was compounded by a belief that there were no career rewards for sharing data. Other barriers included lack of time and resources, lack of knowledge of processes, legal and ethical constraints, and concerns about inappropriate use of the data.

A mixed methods study, focussing more specifically on biodiversity data sharing, gathered data from more than 60 interviews and 700 survey responses (Enke, Thessen, Back, Bendix, Seeger & Gemeinholzer, 2012) and produced similar results to Swan and Brown (2008). Enke et al. (2012) found that loss of control over the data, time and effort invested, concern over incorrect use of the data and legal and ethical issues were the main barriers to data sharing cited by biodiversity researchers.

The Lyon, Rusbridge, Neilson and Whyte (2010) paper is an in-depth study of the variations between and within research disciplines of attitudes and approaches to data sharing, curation, access, reuse and preservation. The methodologies this study employed will not be applicable to the proposed research due to their immense scope. However, the data generated by the study has greatly informed understanding of the area and will assist in the development of questions for the project proposed. Lyon, et al. (2010) used a case study approach where five researchers were embedded in host research groups or curation centres, covering seven separate case studies for months at a time. Data was gathered using semi-structured interviews, by observation of research group meetings and laboratory work, demonstration of operational procedures by staff and by questionnaire.

Separate reports are available for each case study that focus on the differences in the way the disciplines address RDM, but common aspects included researchers protectiveness towards data they create and the tendency to reuse data as a means to advance their current research. The differences in RDM, even at sub-discipline level, led to a suggestion that requirements for curation needed to be established by engagement with specific research communities, rather than developing a general set of characteristics.

Mauthner and Parry (2013) conducted an analysis of data sharing policy in order to understand researcher reluctance to share data. They take an epistemological view and argue that the reluctance is not to do with data infrastructure, but that policies do not take into account data

sharing as relational practice. They suggest that researchers are reluctant to breach the moral contracts they have with respondents and collaborators in their research.

A recent study by Fecher, Friesiker and Hebing (2015) used a combination of methods to investigate data sharing by researchers. A systematic review of 98 papers published between 2001 and 2013 that addressed academic data sharing from the point of view of primary researchers was carried out. The results of this study were combined with those from a survey of secondary data users of a large open data set available in Germany. These users were likely to be highly familiar with data sharing, and have high affinity for its use, but the systematic review is more likely to include information from reluctant data users as well. The framework developed in this study identifies six broad categories of factors that influence data sharing by researchers; data donor, research organisation, research community, norms, data recipients and data infrastructure (Table 1).

Table 1. The categories of factors affecting data sharing by researchers according to the framework developed by Fecher, Friesiker and Hebing (2015).

Framework Categories	Sub-categories
Data donor	Sociodemographic factors Degree of control Resources needed Returns
Research organisation	Policies of data donor's organisation Policies of research funders
Research community	Data sharing culture Standards Scientific value Publications
Norms	Legal norms Ethical norms
Data recipients	Adverse use of data Recipients organisation
Data Infrastructure	Architecture Usability Management software

The identification of barriers to sharing research data will become increasingly important when trying to maximise use of a service that will be provided at considerable cost to an academic institution. Researchers who perform data sharing believe that libraries have a role in maintenance of data repositories, data education and training of researchers (Williams, 2013) but also have some doubts as to their current ability to do so (Frank & Pharo, 2015). Although librarians and researchers may have different views on what the most important barriers to data sharing are, there has not yet been a study that directly compares their perceptions.

Research questions

Research question

How do perceptions of the barriers to uptake and use of data sharing systems vary between science researchers and academic librarians at the University of Auckland?

Sub-questions

- How much use do University of Auckland researchers currently make of available data sharing systems?
- Which barriers to data sharing are considered by researchers and librarians to have the most impact on use of RDM systems?
- Is the relative importance of these barriers affected by the faculty or field a researcher or librarian works in?

Research Design

This study had a cross-sectional research design, using a mixed methods research strategy. Mixed methods research involves the collection of both qualitative and quantitative data, which are analysed together to form a more complete understanding of a research topic. In particular a sequential exploratory design was used where the initial qualitative data collection and analysis leads on to the quantitative data collection and analysis (Cresswell, 2014). The research was carried out in two phases, with a small qualitative investigation using interviews to gather data which were used solely to inform the construction of the phase two research tools. Phase two is a mainly quantitative research strategy based on surveys, but with some qualitative aspects where open text answers are allowed as part of the response to some survey questions. This set of data was the main focus of the research.

Methodology and methods

Phase 1 – Preliminary Interviews

Phase 1 was a preliminary study carried out to gather data that has only be used to inform the construction of the survey for Phase 2. Semi-structured interviews were carried out with four members of University of Auckland staff. Two were sourced from the academic research, and two from academic librarians who work with research staff at the University. In order to best inform the development of the survey, candidates were approached who work in research or with researchers from different disciplines. The two research staff selected worked in the Life Sciences and Medical research disciplines, and of the library staff chosen, one worked specifically with researchers from the Arts and Humanities, and one had worked with researchers in many disciplines.

Potential candidates for interviews were approached ahead of time, and were provided with interview guide covering the topics that the researcher expected to cover during the interview. The question sheet provided to participants and used in the semi-structured interviews is included in Appendix A. Given that the interviews will be carried out in person, anonymity was not possible, but confidentiality was maintained. Audio recordings of the interviews were made using a Sony stereo voice recorder ICD-UX71F, and were between 15 and 25 minutes in duration. The data were transcribed and coded using the NVivo (NVivo, 2012) system for identification of topics to be added to the survey. Recordings and transcripts have been kept in password-protected folders.

Phase 2 – Survey

A survey was selected as the main data collection tool due to several factors. It has the advantage of enabling collection of a relatively large amount of information in a short period of time, and enables the identification of information that can be representative of a larger population from a smaller subset of participants (Fowler, 2009).

Population and sample

The population surveyed for Phase 2 of the proposed research has two components. The first sub-population is comprised of researchers based at the University of Auckland. This population included academics, post-doctoral researchers, research assistants and PhD candidates. The second sub-population is the library staff who work with researchers at the University of Auckland, and included subject specialist librarians, and Research Support librarians. The academic librarian sub-population is much smaller than that of the researcher population at the university, and so all relevant staff identified will be invited to participate in the research. Permission was received from the acting University Librarian to use the subject librarian mailing list, to directly contact the Research Support Librarian team, and to post the information about the survey on the in-house social media tool Yammer in the librarian group as a reminder. As there is no sampling of the librarian population, the invitation was not sent to a sample of the researcher population, but rather an open invitation to participate in the survey was sent to the research mailing lists of the faculties and departments.

As part of the ethical considerations for this study, permission to contact University staff was obtained from the heads of faculties or departments. Permission was received to contact researchers at the Faculty of Education and Social Work, and the Faculty of Medical and Health Sciences. No response was received from the Faculties of Arts and Science, but permission was received from the Heads of School for Environment, Physics and Business; as well as the Classics and Ancient History department.

The different faculties, schools and departments all had slightly different ways that they were willing to have the survey invitation and information disseminated. The Schools of Environment and Physics were willing for the survey invitation to be sent to their academic staff, post-doctoral researcher and PhD student email lists, and so it is likely that it reached all of their researchers. The Faculty of Medical and Health Sciences were willing for the invitation to be sent to the whole Research mailing list for the faculty, however subscription to this list is voluntary and it is possible that not all researchers in the faculty received the invitation. The Faculty of Education and Social Work were not willing to give me access to their mailing lists, but were willing for the information to be distributed via the lists through the mediation of a faculty administrator. The School of Business were not willing

for all of their research staff to be contacted about the survey, but selected a sample of 20 research staff who were provided the survey information and invitation to participate, through the mediation of one of their research administrators.

Survey Design and Data collection

Data was collected by an online survey constructed on the Qualtrics (Qualtrics, 2016) survey tool; this is an extremely flexible tool providing convenient access to the data collected. The questions and options included in the survey were informed by reviews of the data sharing literature, and by additional topics and issues raised in the interviews from Phase 1. The factors identified in all of the studies described in the current literature review as barriers to data sharing seemed to fit well within the categories described by Fecher et al. (2015, Table 1). This set of categories and sub-categories provided the basic framework for the questions that were included in the survey.

The types of data collected were both categorical and continuous in nature. The majority of questions for describing interactions with data-sharing had multiple choice answers for rating the importance of various factors on a symmetric ordinal scale, with the option for a free-text response in some cases. Consideration was made as to the structure of questions and the order of questions within the survey as these are known to affect the way that participants make their responses (Saris & Gallhofer, 2014). In particular, in Question 9 participants were asked to rate the impact of each of 20 factors on a five point scale. Such a question is susceptible to fatigue bias, where less focus or consideration is put into the questions asked last (Egleston, Miller & Meropol, 2011). In order to reduce this bias, the order in which the factors were presented to the participant were randomised. The exception to this was the "Other" factor, with its associated free-text box, which was always presented to the participant last.

A pilot study was performed that involved sending the preliminary version of the survey to six researchers who had previously worked at the University of Auckland, but now worked for other employers, and so would not be part of the final population being sampled for the study. Feedback from those participants was used to improve the survey structure and questions for clarity and usability. The final set of survey questions can be found in Appendix B.

The invitation to participate in the study included the participant information sheet, a link to the survey website and information around participant consent. A reminder to participate in the survey was sent out to the researcher mailing lists several weeks after the original contact where possible, and was directly sent to the individual library staff. An additional reminder to librarians was made by

posting the survey information and invitation to participate on the librarians group on the in-house social media system Yammer.

Data analysis

Results of the data collected was compared between the two main population sub-groups of researchers and librarians, but comparisons were also made between groups of researchers from different disciplines. Analyses were carried out using the in-program analysis tools and by exporting data into Microsoft Excel for description and summary and in order to calculate the statistical significance of any differences observed in the quantitative data.

Data has been described using standard measures of the central tendency such as mode, median and mean, depending on the type of data recorded at each question. Student's t-tests could not be used to determine the statistical significance of differences between the populations as the data is likely to be non-normal in distribution and ordinal in nature. The Mann-Whitney U test has been used to determine the statistical significance of observed differences in the data, as they did not meet the requirements for parametric tests (Hole, 2011). The Mann-Whitney test measures significant differences in the central tendencies of two groups, in this case using medians rather than means as the data is ordinal in nature. Mann-Whitney also has the benefit that it does not require that the data be normally distributed in order to be valid (Vogt, 2014).

Data from the rating scale questions has been plotted using diverging stacked bar-graphs, as these make it easier to compare the attitudes of respondents in different demographic categories. Other forms of graphical presentation can be misleading and do not facilitate comparisons well (Robbins & Heiberger, 2011).

Response rate of the librarian population has been calculated from the direct mailing of individual staff with known involvement in research. Information about the total number of researchers on the mailing lists and still currently working for the University could not be obtained as the faculties and departments participating in the study were unwilling to disclose this information. It has been taken into account that as responding to the survey is voluntary, it is likely that the data responses have been skewed due to respondent interest in or engagement with data sharing practices and may not be representative of the whole community.

Ethical considerations

Given that the population that has been sampled for this research are employed by the same institute that the researcher works for, it was made clear that participation is not compulsory and will in no way affect an individual's employment or status at the institution. It was not expected that any of the questions presented in the survey were likely to cause significant emotional or psychological distress to participants.

All responses have remained confidential and will only be viewed by the researcher and their supervisor. The survey part of the study may be anonymous, unless identifying information is provided during the survey by participants. Any identifying data has been removed from the responses before analysis and publication. Although the respondents may choose to provide an email address with which to be contacted with the final results of the study, this information has been kept separately from the survey responses for that individual. Information sheets were provided with the invitation to complete the survey, and participants were notified that participation in the survey indicates informed consent. Data has been kept in password secured folders and has only been accessible to the researcher and their research supervisor.

Human Ethics approval was sought and received from the School of Information Management Human Ethics Committee (SIM HEC) at Victoria University of Wellington (see Appendix C).

Publication venues

The results of this research will be made available on request to any of the responding interview and survey participants, who will be emailed with a summary of the results. It is expected that the data will be published in a peer reviewed journal, and that the final project report produced will be submitted to the institutional repository at Victoria University of Wellington. The results will be presented in a "Library Update" seminar addressed to the Libraries and Learning services staff of the University of Auckland and it is hoped to also present the data at a wider professional conference in New Zealand or Australia.

Results

Part 1 - Interviews

Four interviews were undertaken in order to inform the construction of a more effective survey tool to better gather data and understand the barriers to data-sharing by researchers at the University of Auckland. These interviews took place over the course of a month, with the interviews with Researcher 1 and 2 being conducted before those with Librarian 1 and 2. The interviews were transcribed and analysed using the NVivo system, and many of the themes that emerged after the coding process closely reflected the framework outlined by Fecher et al. (2015). Themes of particular interest were those relating to the data donor category, and the research community category, with issues relating to ethical and legal norms also arising. Both of the researchers spoke of times when a lack of appropriate data-sharing by other researchers had impeded their work, with one stating

“Data-sharing issues hit at the end of your PhD, the start of your post-doc and become more evident as you continue working”.

This references the socio-demographic category affecting data-sharing by researchers, and is also reflected in the comment

“As I have become more established in my career I have become more confident about sharing data”.

Research organisation themes also arose, particularly regarding the data-sharing policies of international funders, but both researchers indicated that they had not seen much evidence of interest in data-sharing from New Zealand researchers or funders, with most data-sharing requests coming from overseas researchers, and only funders in Europe and America requiring that the data from their projects be made openly available.

Although the researchers had recently seen evidence that the institutions where research takes place were interested in encouraging open access in both data and publication one stated

“This has little effect, data-sharing needs to be mandated in order to work”.

The policies of the National Institute of Health (NIH) in the USA was given as an example, where data-sharing was mandated but not initially policed, to give the research community time to adjust to the concept of mandated data-sharing before the funder started to enforce the sharing of data from the projects it funds.

The interviews with the library staff also high-lighted themes from the Fecher et al. (2015) framework, particularly those around the resources needed, the effect of sharing-data on the ability of the researcher to publish, and legal norms such as copyright and intellectual property.

One factor that seemed to fall outside this framework was identified by Librarian 1 who worked with researchers in the Arts and Humanities fields. Particularly with regards to researchers working in fields such as literature or history

“There just isn’t any unpublished data apart from their initial notes which contain their personal thought-processes. Everything else, all of the arguments and evidence goes into the publication”.

Apart from the lack of time and resources that frustrated university researchers when the task of sharing data was added to an already busy schedule of teaching, research and administration, Librarian 2 said that in their experience researchers often were simply not aware of the facilities, tools and support services available to them at the university to help in this process.

These points that did not seem to fall within the Fecher et al. (2015) framework of factors affecting data sharing by researchers were also integrated into the construction of the survey tool for the second part of this study.

Survey construction and pilot study

Although socio-demographic factors are a major category in the Fecher et al. (2015) framework, the survey was designed with minimal demographic factors. The main focus in this area was on the length of time a respondent had worked in or with researchers, and the research disciplines the respondent had worked in or with. Although there may be an impact on data-sharing caused by factors such as gender or age, and some disciplines may have increased proportions of researchers in particular demographic sub-groups, it is not believed that including questions in these areas would have added value to the survey. Additionally, despite the survey being anonymous, it is drawing respondents from a restricted population, and fewer demographic questions mean less likelihood of individual respondents being identifiable. Many researchers and library staff, the target populations, are likely to know this as many of them design surveys as part of their own work.

The next set of questions addressed how the respondent viewed the culture of data-sharing in their discipline, how much they shared data, and the methods that the respondent used to share data. The part of the survey addressing the respective levels of impact that various factors had on the decision to share data was presented as a single question with multiple factors to be rated on a

continual scale. The survey closed with a multiple choice question addressing the respondent's awareness of data-sharing systems and support at the university, and a free text box for comments.

A pilot study of this survey was carried out, using a sample of six researchers who had previously worked at the University of Auckland, but were currently working at other institutions and so would not be part of the population targeted by the main survey. They worked in research disciplines including Engineering, Mathematics, Business (Marketing research) and the Life Sciences. All of the targeted participants completed the survey and then provided feedback on the content and structure by email. The pilot study identified areas where the wording of the questions made the meaning unclear, and technical issues with the way that the survey was presented to participants in Qualtrics. These improvements were integrated into the version of the survey that was presented to study participants. An example is the suggestion to use sliders rather than buttons on the continuous scale questions to make the survey more compatible with mobile devices, and places where the free-text box presented was too small.

A major suggestion by the Marketing research pilot study participant, was that there can be issues with grid-style questions where respondents are expected to rate many factors along a scale and respondents select the same value for each factor. Such data is effectively meaningless and may need to be excluded from the analysis. It was recommended that I add an additional question where the most important of these factors be selected by the participant, ensuring both some usable data from the response, and a different way of evaluating the relative impact of each factor.

Part 2 – The Online Survey

The survey was made available via an anonymous URL for participants to link through to the survey on the Qualtrics website, and was open to responses for a month. After the survey was closed there had been 102 submissions by researchers, and 18 by librarians. The response rate for the librarians can be calculated, as the survey invitation was sent to 96 librarians within Libraries and Learning Services at the University of Auckland. This is a response rate of 19%, however, when you consider that only 9 of those surveys from librarians were completed, this response rate drops to 9.3%. Of the research submissions, 83 out of 102 were completed, but the variable nature in which the surveys were distributed to researchers meant that it is not possible to calculate an overall response rate for research staff. There were 12 responses from researchers who selected Business and Economics as a research discipline which is a response rate of 60% from the 20 researchers targeted, but it is

unlikely that a response rate of this level was achieved across all of the research disciplines. Only one respondent identified themselves as both a researcher and a librarian.

Comparison of Researcher and Librarian perceptions of barriers to data-sharing

Responses to the demographic questions in the survey showed that the greatest proportion of researcher respondents had been working in research for over fifteen years, with the next highest proportion belonging to those who had been working in research 1-5 years, the majority of whom are probably PhD students. A similar spread of time spent working with researchers is seen for the librarian respondents, with an increased proportion having worked for 5-10 years (Figure 2). The category of <1 year has unsurprisingly fewer respondents as it is a smaller temporal period than the other categories.



Figure 2. Demographic data for Researcher (n=102) and Librarian (n=18) responses, describing the number of years they have worked in or with researchers.

Research is increasingly becoming a multidisciplinary endeavour, and a large proportion of the researcher respondents selected more than one research discipline. There were researcher respondents from all of the categories of research discipline, although significantly more in Social Science and Education, Physical Chemical and Earth Sciences, and Medicine and Health (Figure 3).

The nature of working in an academic library is reflected in the fact that 80% of the librarian respondents indicated that they had worked with researchers in more than one discipline. The “Other” category was selected by librarians who wished to indicate that they worked with researchers from all disciplines.

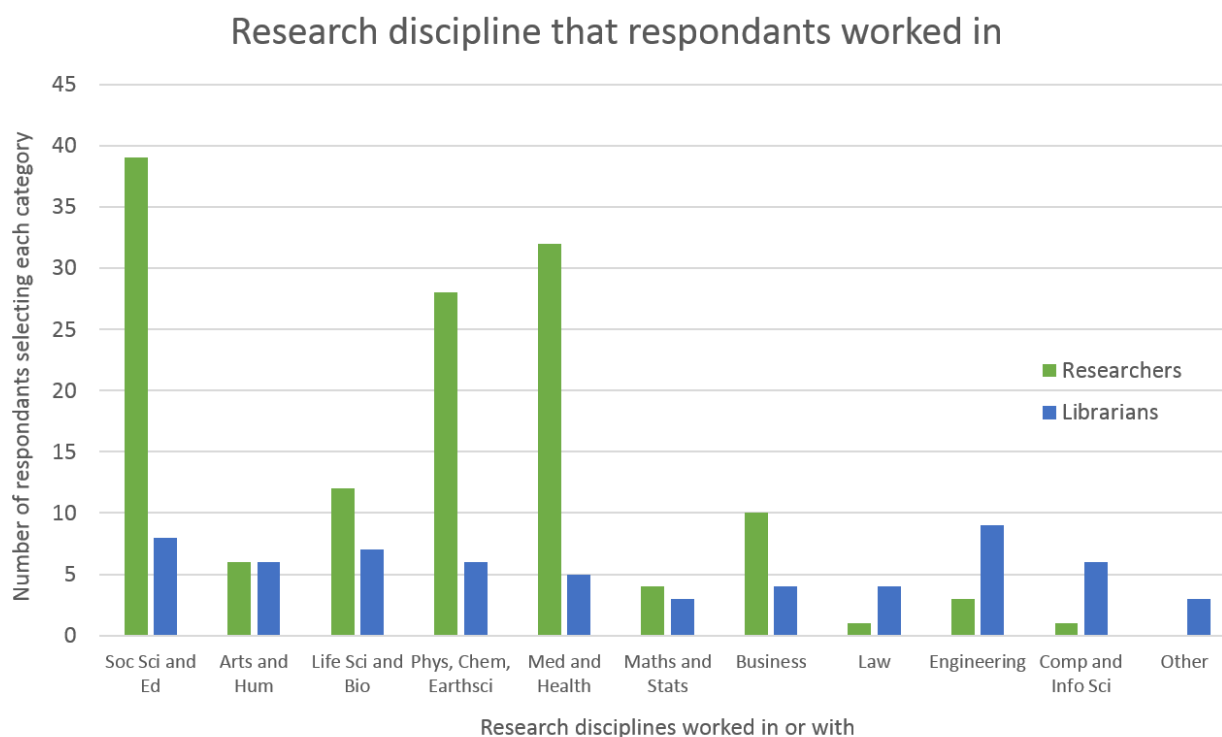


Figure 3. Demographic data showing the research disciplines that respondents worked in or with for Researchers (n=102) and Librarians (n=18).

There were very similar patterns in the ways that researchers and librarians thought that research data was being shared, but librarians were much more likely to think that researchers were using institutional and community repositories, and also to indicate that they were not sure how data was being shared (Figure 4). The most common ways that researchers indicated they shared data were by email (59%) and secure sharing systems such as drop-box or Google drive (69%). Funder repositories such as the one managed by the Wellcome Trust are relatively new as a way to share data, so it is perhaps not surprising that this was not selected by any respondents as a way to share data.

Researchers were more likely to think that data-sharing was part of the research culture, with a greater proportion of researchers (26%) indicating that data sharing was a common or essential part of the research culture compared to only 4% of librarians (Figure 5A). Researchers were also twice as likely to say that they regularly or frequently shared data (30%) while only 15% of librarians thought

that researchers regularly shared data and none of them thought that researchers frequently shared data (Figure 5B).

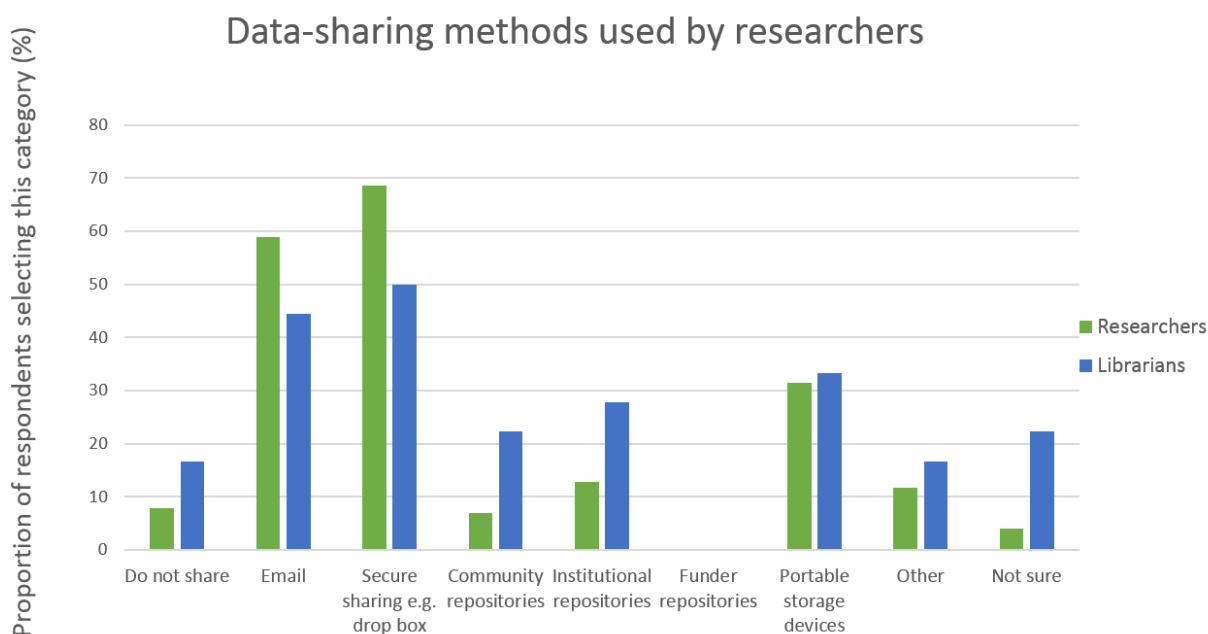


Figure 4. Data-sharing methods used by researchers, and methods librarians think researchers use to share data. Graphical representation of percentage of respondents selecting each category.

The main focus of the survey was on the impact that different factors have on whether researchers will share data sets associated with a publication of a piece of funded research. The responses to all of the factors measured in Question 9 on the survey were collated for Researchers (Figure 6) and for Librarians (Figure 7). For the group of all researchers, the most important factors in whether to share a data set seem to be concerns over inappropriate or adverse use of the data by others, ethical concerns, whether sharing data will affect the ability of the researcher to publish and how much control the researcher has over how the data is shared. The patterns of relative impact in Figures 6 and 7 indicate that there are considerable differences in the factors that researchers and librarians think are influencing whether researchers will share data. The count data for the responses to Question 9 for researchers and librarians is available in Appendix D.

Some of the factors where the librarian and research views agree on the relative impact are shown in Figure 8. Librarians have similar perceptions to researchers of the relative importance of researcher control over how the data is shared, the quality of the research being shared, and of the impact of concerns of inappropriate or adverse use of data by others after sharing.

A Data-sharing is part of the research culture

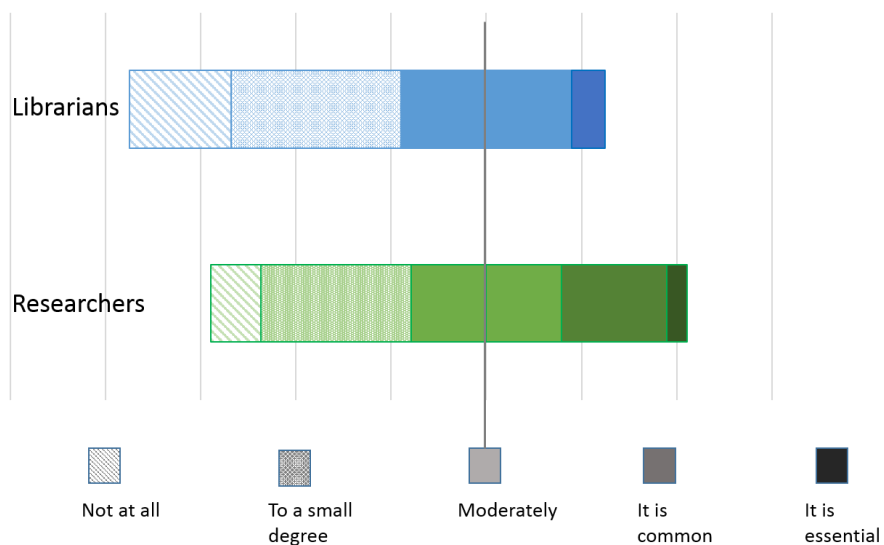
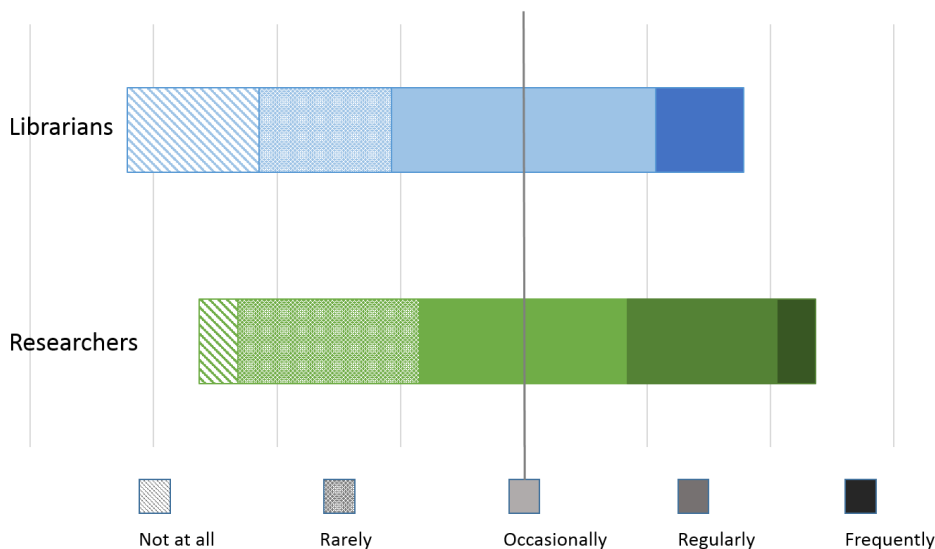


Figure 5. Comparison of researcher and librarian views of whether data-sharing is part of the research culture (A), and whether researchers regularly share data (B). The data shown are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Researchers n=95, Librarians n=14)

B Researchers regularly share data



Although Figure 8 shows some of the factors where librarians and researchers agreed, there were several areas where they did not agree on the relative impact of the factors involved. Librarians and researchers have a different perspective on the impact of data-sharing policy on the decision by a researcher to share a data set (Figure 9). Librarians are more likely to think that the policies of the institute where a researcher works (56%), the policies of research funders (78%) have “definite” or “significant” impact when compared to researchers (34% and 40% respectively).

All researchers - relative importance of factors influencing data-sharing

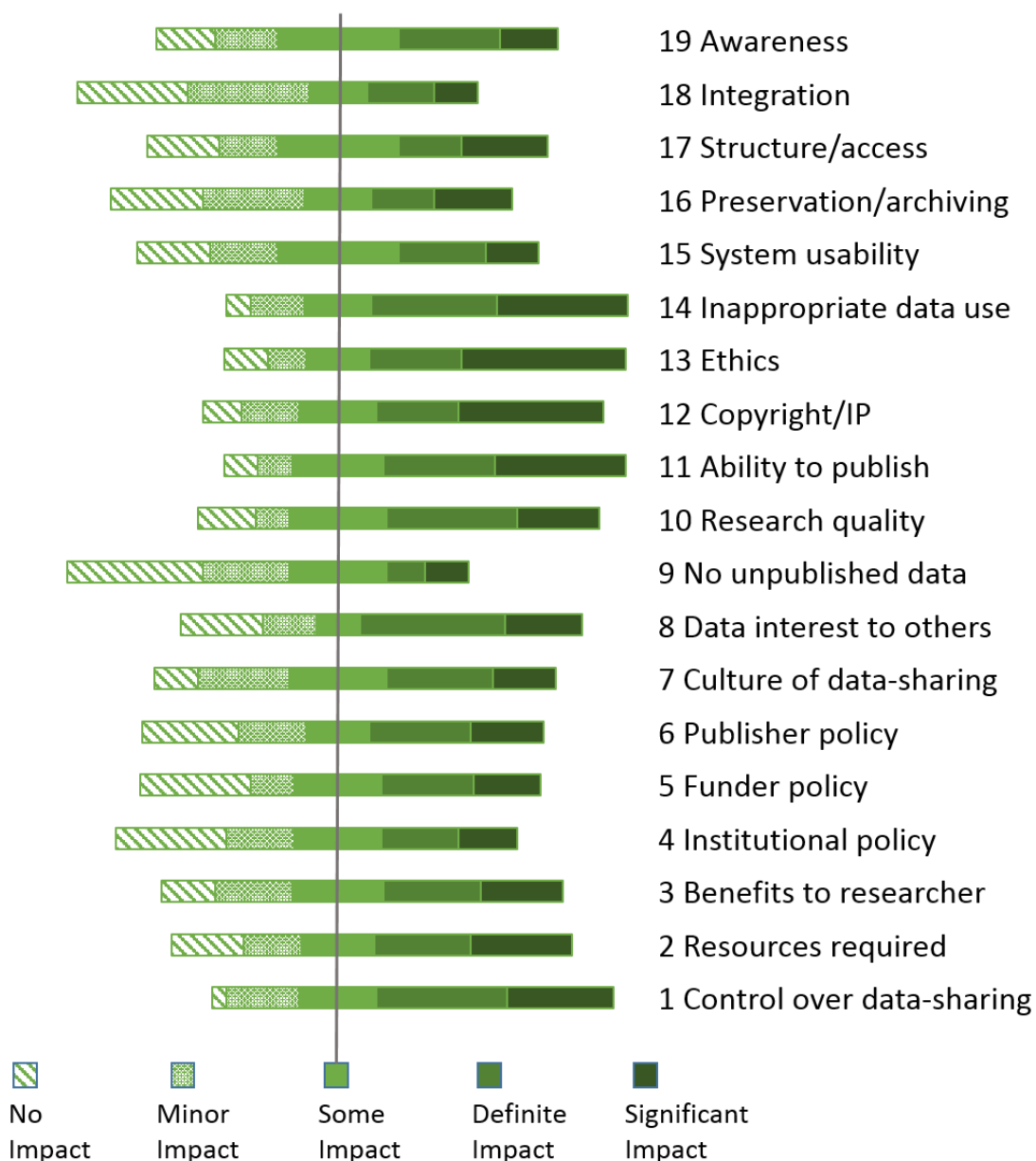


Figure 6. Relative impact of the factors affecting the decision by researchers to share data, as described by respondents who are researchers. The data shown are the percentage of responses in each category. The grey line indicates the median response of the middle value on the continuous scale (n=83)

Librarians - relative importance of factors influencing data-sharing

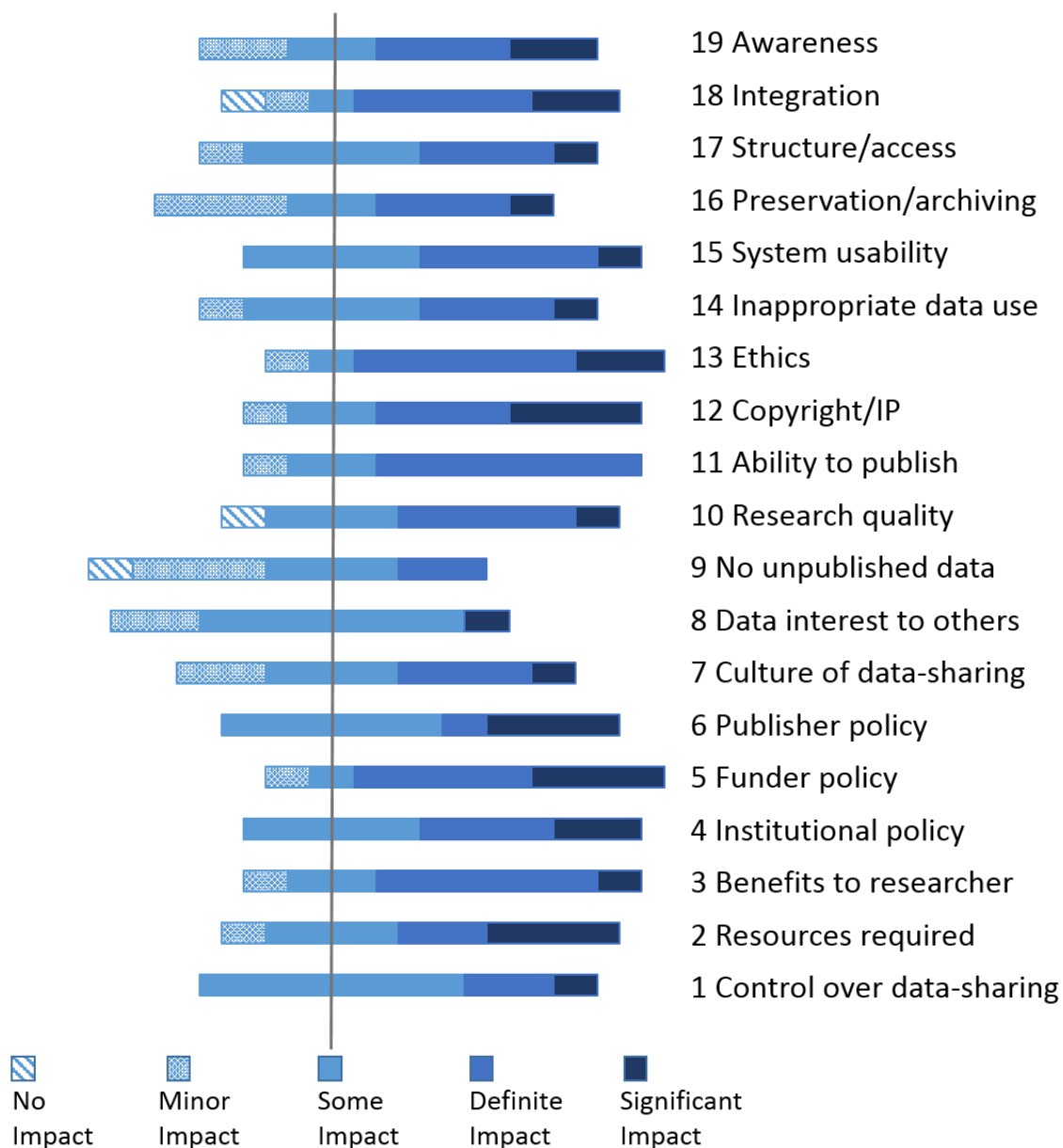


Figure 7. How librarians perceive the relative impact of the factors affecting the decision by researchers to share data. The data shown are the percentage of responses by librarians in each category. The grey line indicates the median response of the middle value on the continuous scale (n=9).

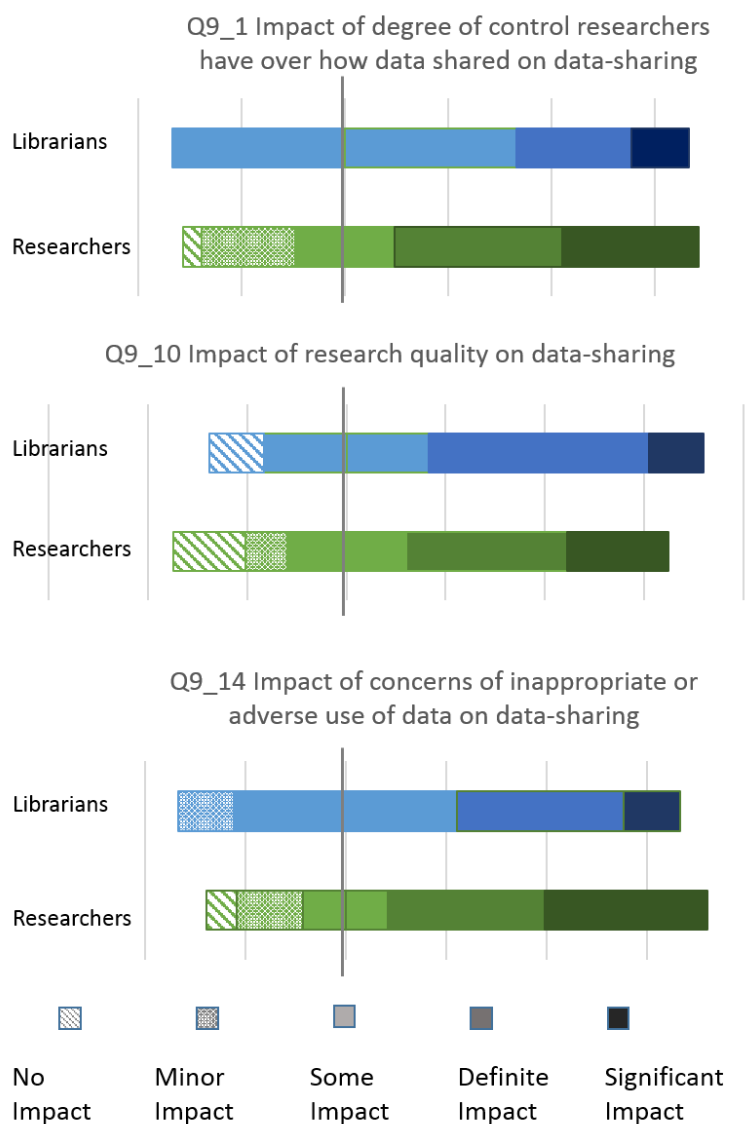


Figure 8. Agreement between researchers and librarians. Comparison of researcher and librarian views of the relative impact of factors affecting researcher decisions to share data.

Three factors were selected where the relative impact was assessed similarly by researchers and librarians: Q9_1 researcher control over data sharing, Q9_10 quality of the research, Q9_14 concerns over inappropriate use of the data.

The data displayed are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Researchers n=83, Librarians n=9).

While the impact of publisher policy also looks different (Figure 9), this is due more to the proportion of researchers indicating that it has “no impact” or “minor impact” (41%) when the proportion for librarians was 0% in those categories.

Another area that librarians thought had more impact than researchers was around the data-sharing systems area (Figure 10). Librarians were considerably more likely to think that whether the data-sharing system was integrated into other university research systems had a “definite” or “significant” impact on data-sharing (67%) than researchers (28%). More librarians also thought that the usability of the data-sharing system had “definite” or “significant impact (56%) than researchers (39%).

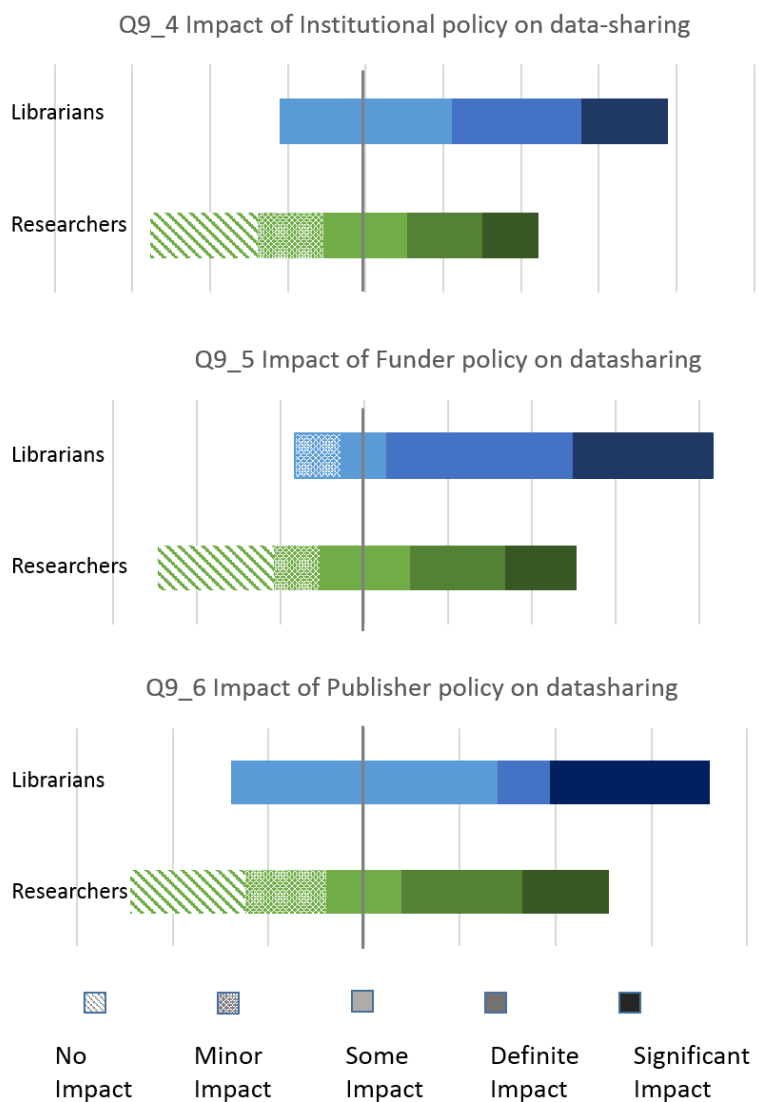


Figure 9. Disagreement between researchers and librarians on the relative impact of data-sharing policy factors affecting researcher decisions to share data.

Three factors discussed the impact of data-sharing policy Q9_4 Institutional policy, Q9_5 funder policy, and Q9_6 publisher policy.

The data displayed are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Researchers n=83, Librarians n=9).

There was one area, however, that researchers thought had more impact on data-sharing than librarians. 55% of researchers thought that the question of whether the data would be interesting or relevant to other researchers had “definite” or “significant” impact, compared to only 11% of librarians.

Question 10 in the survey asked respondents to select up to 3 of the most important factors that impact the decision of whether to share a research data set. Although many people selected three factors, some selected two or only one factor. Concerns around ethical issues was the most important factor for researchers with 42% of the respondents selecting this factor, while it was selected by only 22% of librarians (Figure 11). The next most important issues to researchers were whether sharing data would affect the researcher’s ability to publish, concerns around inappropriate or adverse use of the data by others, and issues around copyright and intellectual property.

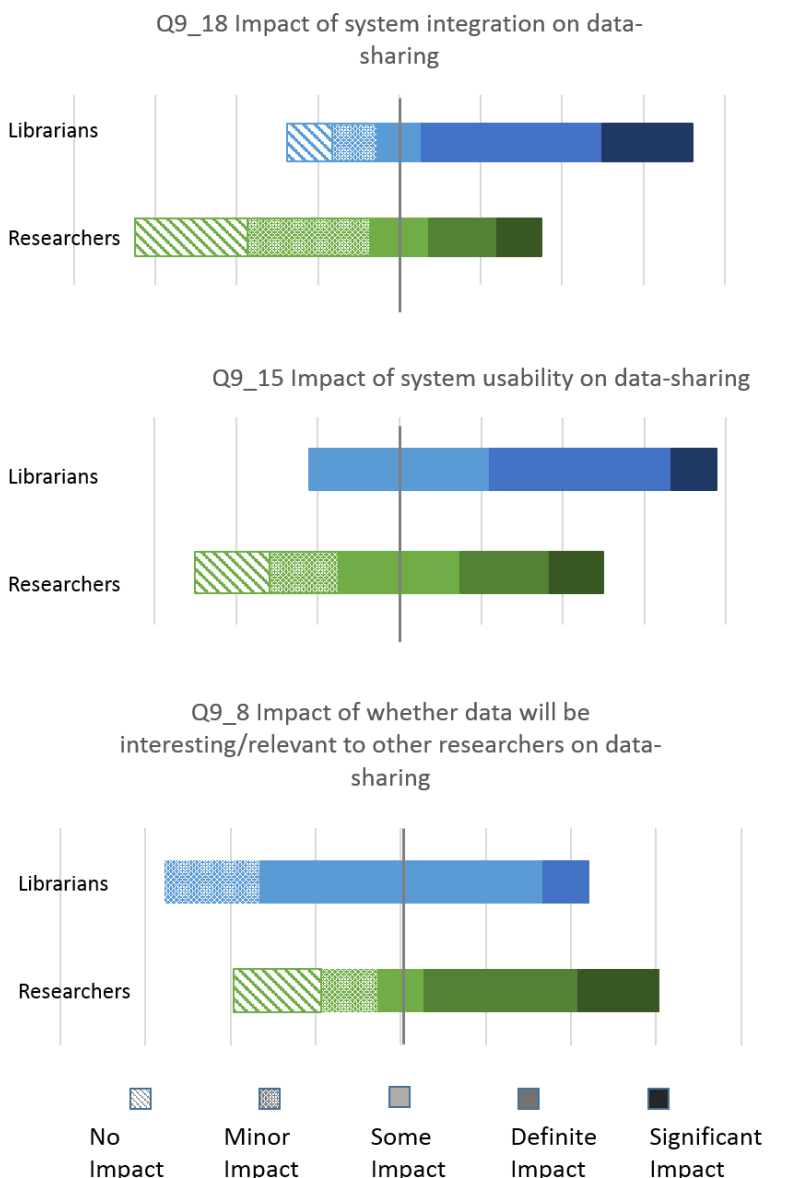


Figure 10. Disagreement between researchers and librarians on the relative impact of data-sharing policy factors affecting researcher decisions to share data.

Two factors discussed the impact of data-sharing systems Q9_18 System integration and Q9_15 System usability. A factor more important to researchers was Q9_8 Interest/relevance of data to other researchers.

The data displayed are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Researchers n=83, Librarians n=9).

All of these factors were also selected by librarians at a similar proportion to the researchers (Figure 11). The culture of research sharing, and the quality of the research being shared were both selected as most important factors by over 10% of researchers, but were not selected as being the most important by any of the librarians. This difference could also be due to the small population size of the librarian respondents.

Librarians again selected institutional policy as being more important (22%) than researchers (5%), but the difference in selecting funder policy was not as great selected by only 11% of librarians and 5% of researchers. Librarians also selected the integration of data-sharing systems into research as being one of the most important factors at a much higher rate (22%) than researchers (1%), reflecting the results observed in the answers given to Question 9 of the survey.

Q10 - most important factors affecting data-sharing

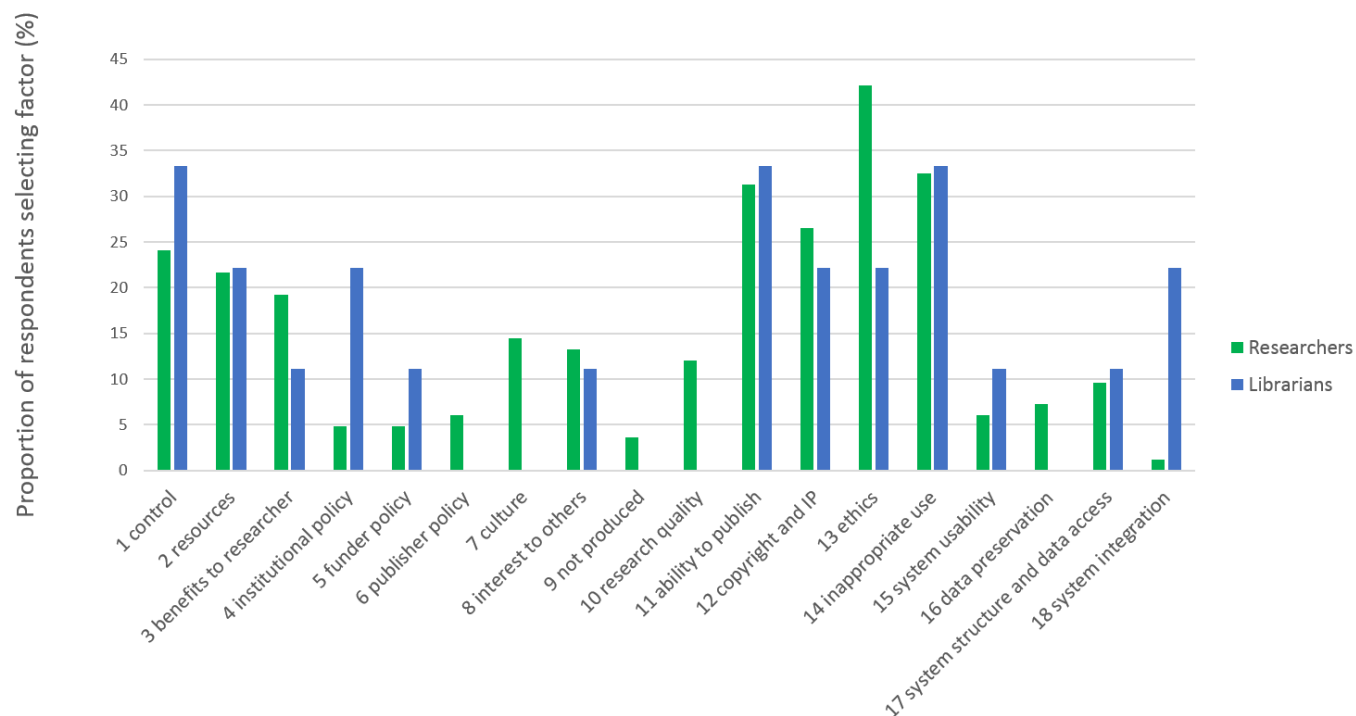


Figure 11. Comparison of the factors with the most impact on data-sharing by researchers, as selected by researchers and librarians. Data presented as proportions of respondents selecting a factor. (Researchers n=83, Librarians n=9).

Comparison of the perceived barriers to research data sharing by researcher discipline

As some of the research disciplines had very few respondents, comparisons by discipline were only made between research disciplines with greater than 10 respondents, in order to allow both fair comparisons and to enable statistical testing. The five research disciplines selected were Social Science and Education, Life Science and Biology, Physical Chemical and Earth Sciences, and Business and Economics.

At a demographic level, all disciplines had representatives in each of the respondent categories for the length of time spent working in research, with points of difference being a reduced proportion of Life Science and Biology respondents in the 1-5 year category and an decreased proportion of Physics, Chemistry and Earth Science respondents in the >15 years category when compared to the other disciplines (Figure 12).

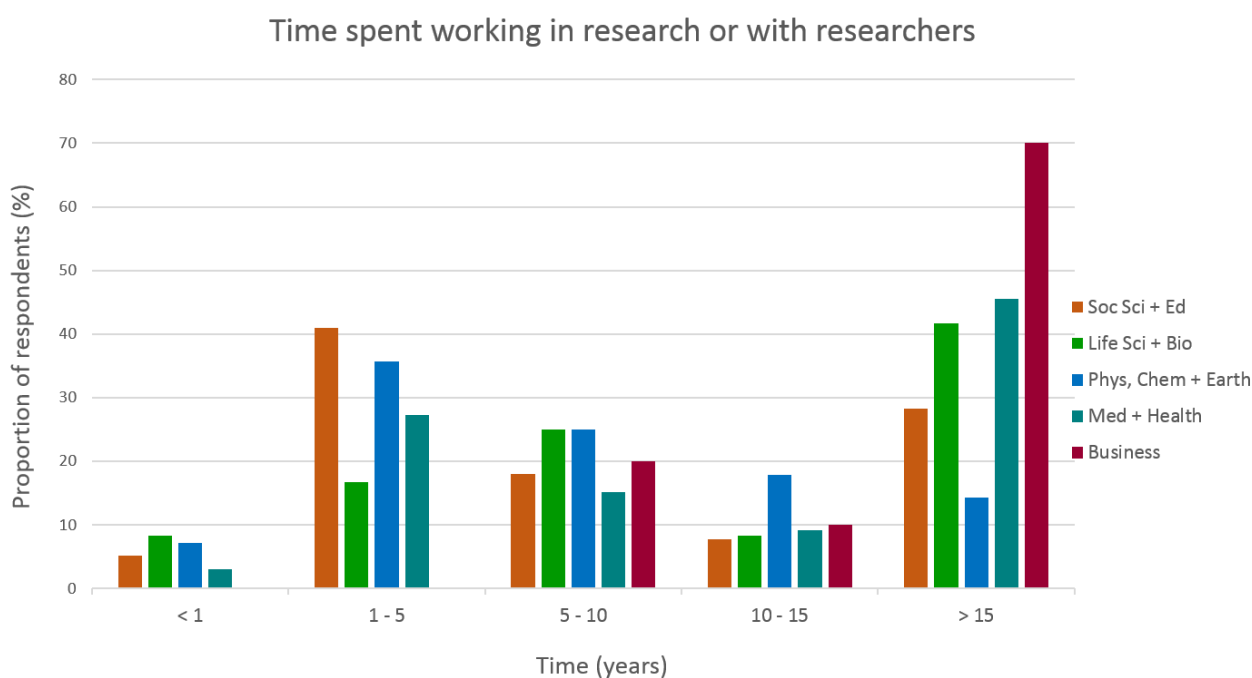
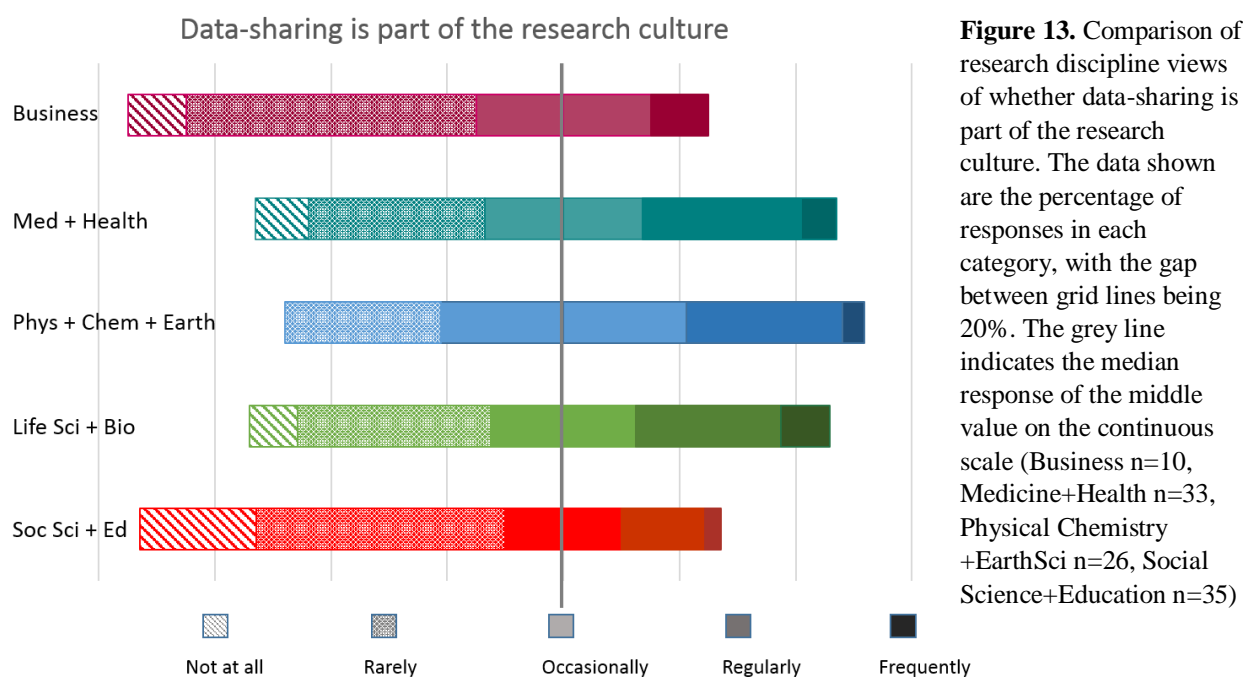


Figure 12. Demographic data for Social Science and Education (n=35), Life Science and Biology (n=12), Physical Chemical and Earth Sciences (n=26), Medicine and Health (n=33), and Business and Economics (n=10) researcher responses, describing the number of years they have worked in research.

Figure 13 shows that a greater proportion of both Social Science and Education (63%), and Business and Economics (60%) researchers believe that data-sharing is “not at all” or “rarely” a part of the research culture in their discipline when compared to Life Science and Biology (41%), Physical Chemical and Earth Sciences (27%), and Medicine and Health (39%).



Mann-Whitney U tests were performed and confirmed that researchers from Social Science and Education (Mdn=2) ranked data-sharing as a significantly smaller part of the research culture than researchers in the Physics, Chemistry and Earth Sciences (Mdn=3, $U=291.5$, $z=-2.38$, $p=0.008$, $\alpha=0.05$, $r=0.31$) and a significantly smaller part of the research culture than those from Medicine and Health (Mdn=3, $U=426$, $z=-1.86$, $p=0.031$, $\alpha=0.05$, $r=0.23$).

The same test showed that there was no significant difference in the way that Business and Economics researchers (Mdn=2.3) and Life Science and Biology researchers (Mdn=3, $U=46.5$, $z=-0.89$, $p=0.19$, $\alpha=0.05$, $r=0.31$) ranked data-sharing as part of the research culture, although they appear different in Figure 13. This may reflect the smaller sizes of these data sets.

A much greater proportion of Social Science and Education researchers (57%) indicated that they shared data to only “a small degree” or “not at all” when compared with Medicine and Health (36%), Physical Chemistry and Earth Science (19%) and Life Sciences and Biology (33%) researchers. Mann-Whitney U tests confirmed that Social Sciences and Education researchers (Mdn=2) share data significantly less than both Physics Chemistry and Earth Science researchers (Mdn=3, $U=253.5$, $z=-2.94$, $p=0.0016$, $\alpha=0.05$, $r=0.38$) and Medicine and Health researchers (Mdn=3, $U=391$, $z=-2.29$, $p=0.011$, $\alpha=0.05$, $r=0.28$).

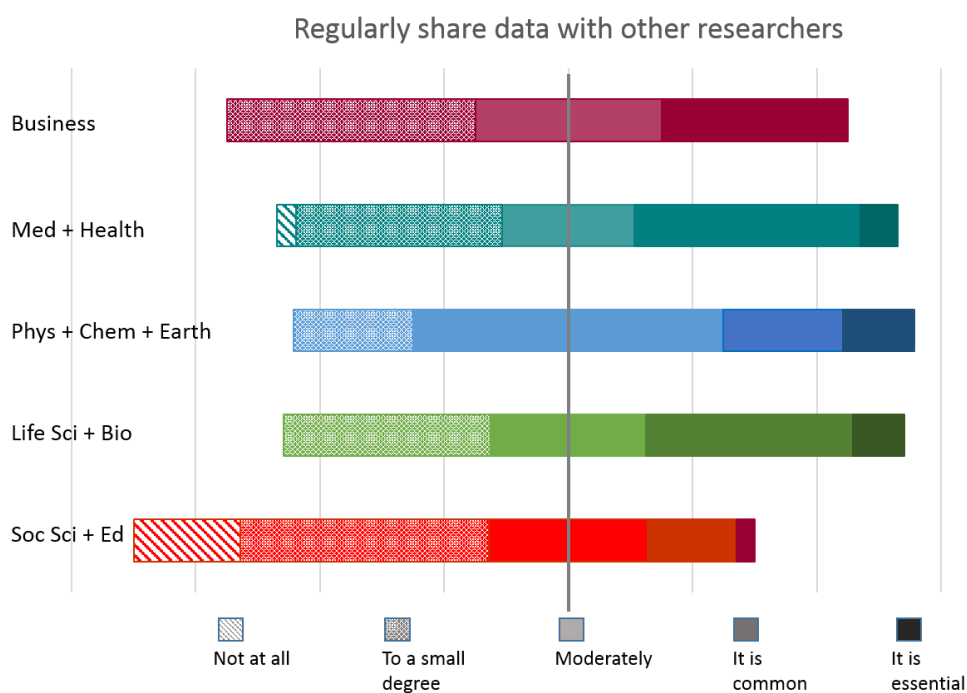


Figure 14. Comparison of research discipline views of whether data-sharing is part of the research culture. The data shown are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Business n=10, Medicine+Health n=33, Physical Chemistry +EarthSci n=26, Social Science+Education n=35)

Graphs that show the comparisons within the research disciplines of all of the factors ranked on their impact on data-sharing can be found in Appendix E.

There were several factors that were ranked consistently on impact across the research disciplines compared, including highly ranked factors such as whether the data will be of interest or use to other researchers, and low ranked factors such as whether research from a discipline will not produce unpublished data (Figure 15).

Although they were overall ranked as less important by researchers than librarians, there is still considerable variation in the ways that individual disciplines view the importance of data-sharing policies on their decision whether to share data. Life Sciences and Biology researchers (50%) and Medicine and Health researchers (57%) rank funder policy the most highly with larger proportions of researchers ranking it of “definite” and “significant” impact. In comparison, Business and Economics (60%) have the highest proportion of researchers selecting publisher policy as of “definite” or “significant” impact. The high impact given to funder policy by Life Sciences and Biology researchers may reflect the presence in this discipline of funders such as the NIH and Wellcome Trust who require data-sharing from projects they fund.

Four factors that were ranked significantly differently across disciplines include: the ability of the researcher to publish, which is ranked quite highly by researchers overall; Ethical considerations, also ranked very highly overall; preservation and archiving of data within the system, with a

moderately low overall impact ranking; and integration of the data-sharing system into other university systems, which received a low overall impact rating from researchers (Figure 16).

Factors where disciplines were similar in their estimation of the impact on data-sharing

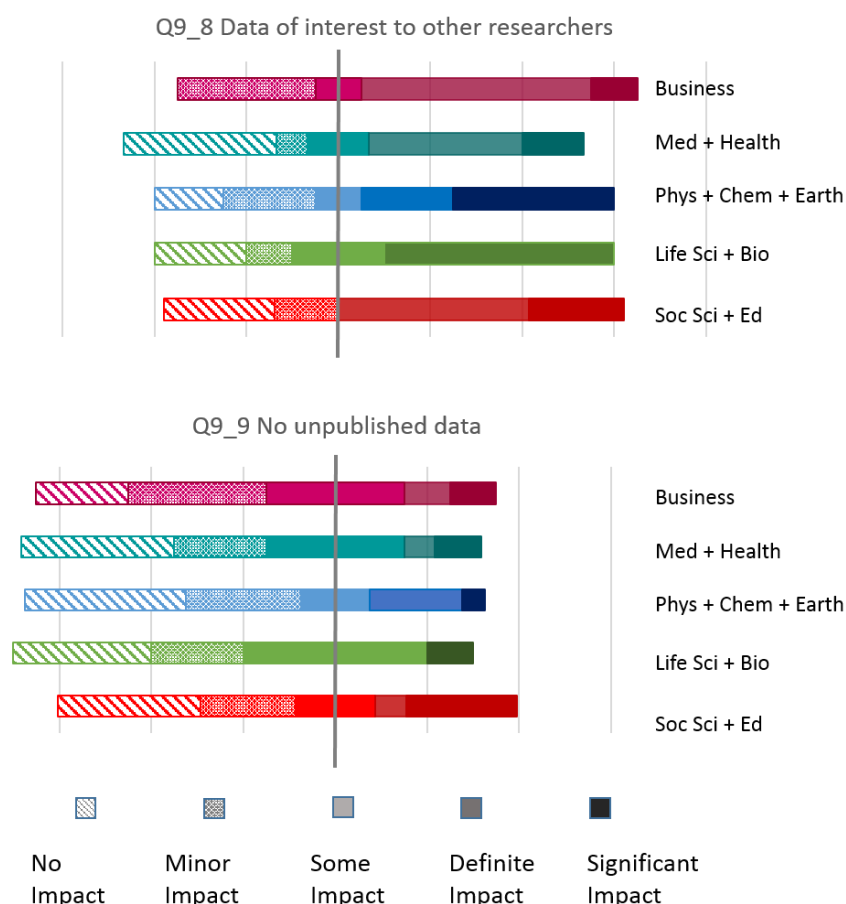


Figure 15. Agreement between researchers across disciplines. Comparison of researcher views of the relative impact of factors affecting researcher decisions to share data.

Two factors where the relative impact was assessed similarly across disciplines: Q9_8 whether data is of interest to other researchers, and Q9_9 that there may be no unpublished data from a study.

The data displayed are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Business n=10, Medicine+Health n=30, Physical Chemistry +EarthSci n=20, Social Science+Education n=29)

Figure 16 shows the differences between disciplines in the way they rank these impact factors, and the significance of the differences was investigated with the Mann-Whitney U test. Business and Economics researchers (Mdn=4.5) rank the ability of the researcher to publish as having significantly more impact than do researchers in Social Science and Education (Mdn=4, $U=88$, $z=-1.83$, $p=0.033$, $\alpha=0.05$, $r=0.29$). Physics Chemistry and Earth Science researchers (Mdn=3) rank Ethical concerns as having significantly lower impact on data-sharing than Medicine and Health researchers (Mdn=5, $U=115$, $z=-3.66$, $p=0.00012$, $\alpha=0.05$, $r=0.52$). This result makes sense, as Physics Chemistry and Earth Science researchers are much less likely to be working on research projects that involve either people or animals.

Factors where disciplines differed in their estimation of the impact on data-sharing

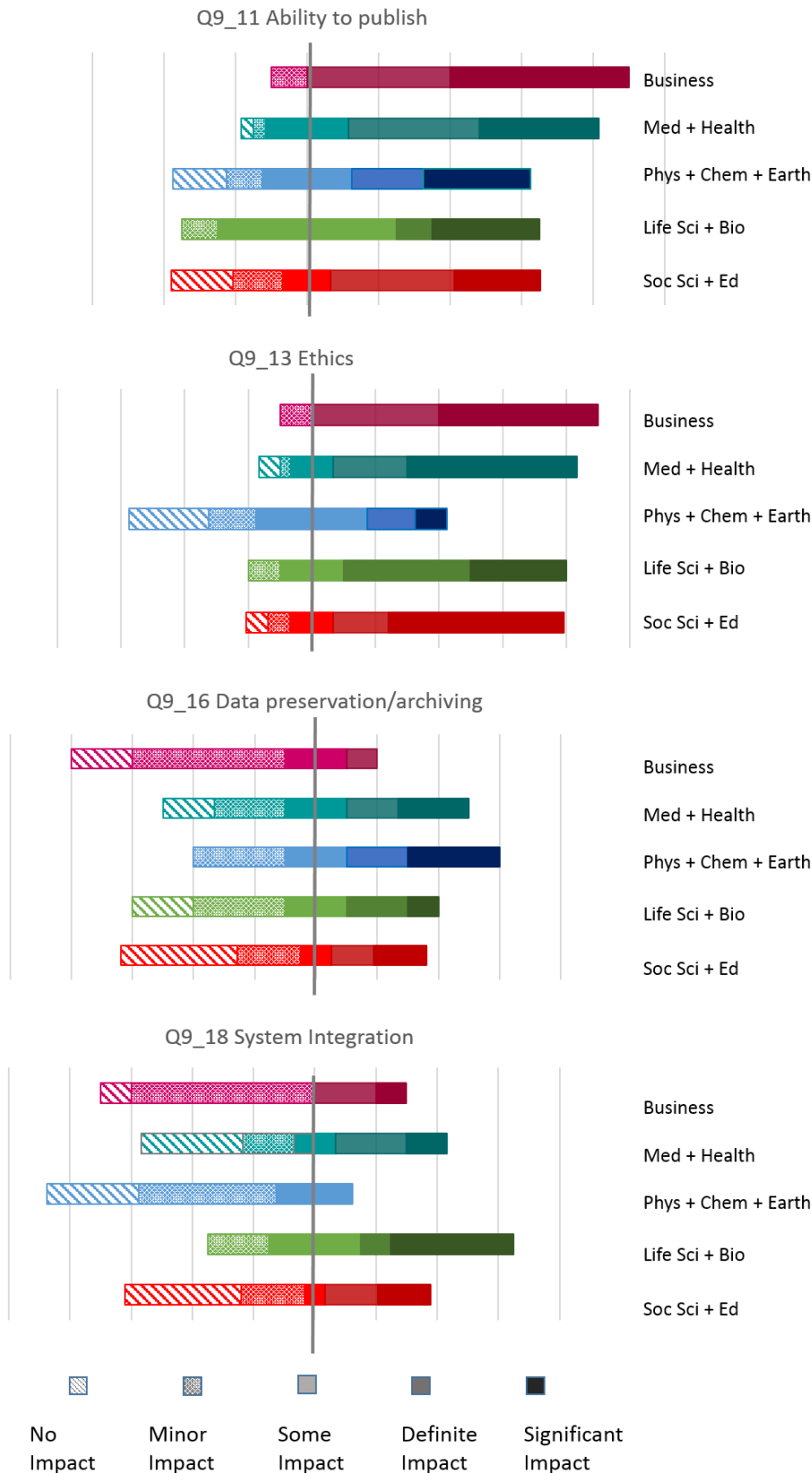


Figure 16. Difference between researchers across disciplines. Comparison of researcher views of the relative impact of factors affecting researcher decisions to share data.

Four factors where the relative impact was assessed differently across disciplines: Q9_11 the ability of researchers to publish, Q9_13 Ethical concerns, Q9_16 data preservation and archiving, and Q9_18 system integration.

The data displayed are the percentage of responses in each category, with the gap between grid lines being 20%. The grey line indicates the median response of the middle value on the continuous scale (Business n=10, Medicine+Health n=30, Physical Chemistry +EarthSci n=20, Social Science+Education n=29)

Physics Chemistry and Earth Science researchers (Mdn=4) rank data preservation and history as a significantly more important factor in data-sharing than researchers in Business and Economics do (Mdn=2, $U=12$, $z=-3.87$, $p=5.4 \times 10^{-5}$, $\alpha=0.05$, $r=0.71$). Life Science and Biology researchers (Mdn=3.5) rank system integration as a significantly higher impact factor than both Physical, Chemistry and Earth Science researchers (Mdn=2, $U=24.5$, $z=-3.32$, $p=0.0004$, $\alpha=0.05$, $r=0.61$) and Business and Economics researchers (Mdn=2, $U=25$, $z=-1.89$, $p=0.029$, $\alpha=0.05$, $r=0.42$).

The difference in perception of research barriers between research disciplines is also evident when investigating the data given as a response to Question 10: which of these factors have the most significant effect on whether you will share data sets associated with a publication or piece of funded research? Figure 17 shows this data separated by research discipline, and the responses vary greatly between the disciplines.

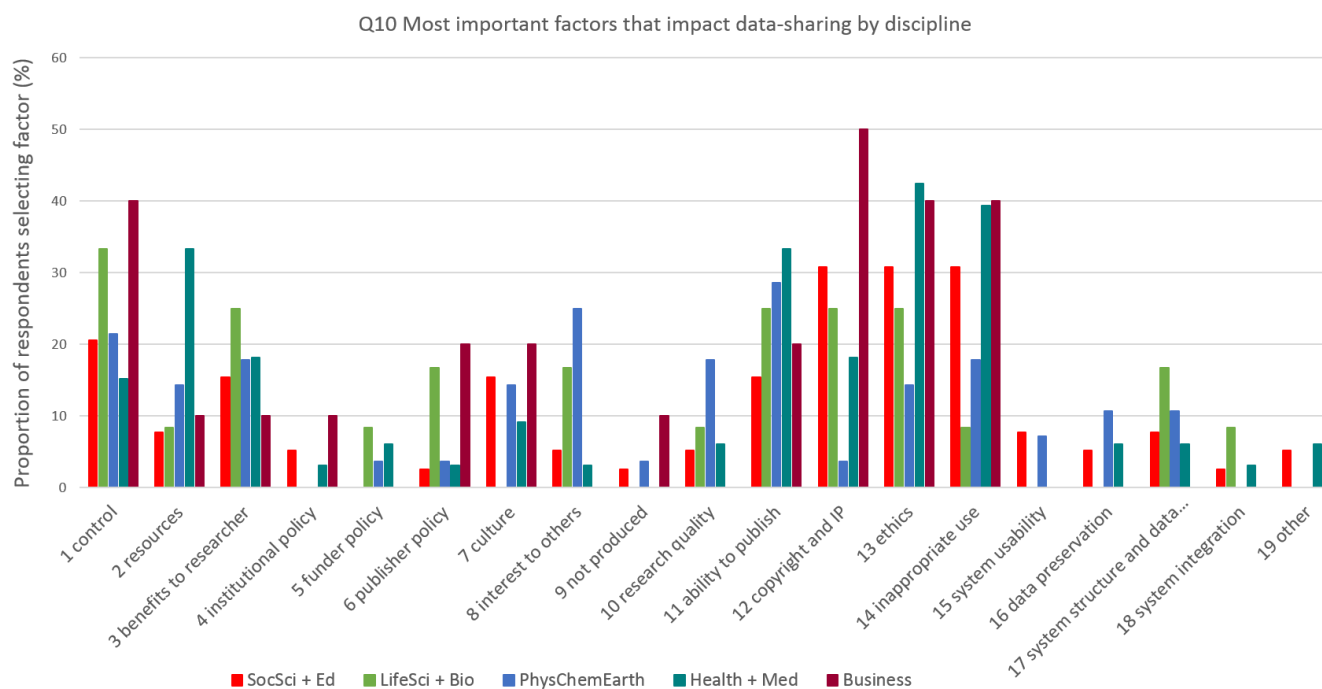


Figure 17. Comparison of the factors with the most impact on data-sharing by researchers, separated by research discipline of the respondent. Data presented as proportions of respondents selecting a factor. (Business $n=10$, Medicine+Health $n=30$, Physical Chemistry +EarthSci $n=20$, Social Science+Education $n=29$)

While 50% of researchers from Business and Economics selected copyright and intellectual property issues as the most important factor, it was chosen by only 3.6% of Physics Chemistry and Earth Science researchers. 33% of Medicine and Health researchers have selected resources in time and money as one of their most important factors in deciding whether to share data, but the other disciplines are all hovering around the 10% mark. Although Business and Economics researchers did not rank either “institutional policy” or “data not produced” as very high impact factors overall in the comparative scale, 10% of Business and Economic researchers selected each of these categories as one of their most important factors. Some factors are only selected by a small proportion of particular disciplines, and not at all by other disciplines (Figure 17).

It is also worth noting that although they were not included in the direct comparison, that every researcher respondent who worked within the Arts and Humanities discipline selected “concerns over inappropriate or adverse use of the data by others” as a factor with the most impact on whether they shared data in response to Question 10.

With regards to the category “Other”, the responses included the more generic “data policies” from one respondent, and “the value of the research question being asked” which links to the idea of control of data access by the researcher who produced it. The final factor in the “Other” category was “Maori world views understood?” This factor is of particular relevance to research in a New Zealand context.

One of the last questions on the survey was whether the respondent was aware of the data-sharing systems and support available through the University of Auckland. 47% of researchers answered “No” and a further 30% answered “Maybe”.

The responses to the free-text box for additional comments at the end of the survey fell into several themes. Some people commented that they were reluctant to try to share data using University of Auckland systems due to previous poor experiences, several researchers emphasised the need for researchers to have control over who gains access to the data they wish to share, and others commented about the restrictions around sharing data that were required by the Human Ethics committee.

Discussion

Overall the Fecher et al. (2015) framework seems to be an effective model for the factors affecting data-sharing by researchers. The factor “awareness of the availability of data-sharing systems” could fit under the sub-category of “Data-sharing culture” in the Research community category. A small proportion of researchers (3.6%) selected “Research from my discipline may not produce data” as one of their most important factors in question 10 of the survey. Although only a small proportion of respondents selected this category, there were only a small proportion of respondents from the Arts and Humanities discipline, whom it is most likely to affect. This indicates that there may be a place for an additional sub-community under the “Research community” category of “Unpublished data not produced”. This factor does not really fall under the auspices of the category “Data-sharing culture”, as it is not that they do not wish to share data, but that there is no data to share. The new factor raised by a research respondent in a free-text box of “consideration/understanding of Maori world views” is a factor of particular relevance to researchers in a New Zealand context. This could also fall under the auspices of the sub-category “data-sharing culture” within the Fecher et al. (2015) framework.

Limitations of the study

One limitation of this study is that the different ways that the survey was disseminated to the different groups of researchers made it very difficult to calculate the response rate to the survey. Time restraints around this piece of research meant that I could not commit additional time to negotiating with research departments with regards to how their researchers could be approached, and it is likely that the different dissemination methods have affected the rates of response and the resulting non-response bias. It is generally accepted that good surveys focus on smaller samples with high response rates (Evans, 1991), however Holbrook, Krosnick and Pfent (2007), found that although surveys with lower response rates showed a decrease in population representativeness, that it was not by a large amount. Although reminder notices can be sent to email lists, the nature of an anonymous survey means that you cannot specifically target non-respondents to increase the response rate.

Unfortunately the extreme differences in the population sizes of the researcher and librarian respondents meant that the data do not meet the requirements for most tests of statistical significance and so the validity of the observations of difference between researcher and librarian perceptions of barriers to data-sharing have not yet been confirmed.

Of the 19 researchers who did not complete the survey, 12 had been working as a researcher for less than five years, indicating that they are probably PhD students. This may mean that they were not yet invested in the idea of data-sharing being an issue for researchers.

The high proportion of librarians who did not complete the survey is of some concern, but I was individually approached by five librarians who wished to apologise for not completing the survey but felt that they were not sufficiently knowledgeable about the way that researchers worked with data in their discipline to be comfortable in answering the rest of the survey. This, combined with the high proportion of librarians answering “not sure” to the question on how data is shared, indicate that there may be a place for further training of academic librarians at the University of Auckland Libraries and Learning Services, in order to provide ongoing research data support to research staff.

Findings and Recommendations

Academic librarians and researchers perceived the relative impact of various factors on data-sharing by researchers in different ways. Librarians are very aware of factors such as system usability and integration having an effect on whether users will continue to use a system, but researchers did not consider that these factors had much impact when deciding whether to share data or not.

Librarians also believed that the data-sharing policies of research institutions and funders had much more impact on research decisions, while researchers ranked these factors quite low in impact. This could be linked to a point raised by one of the researchers interviewed in Phase 1 of the study, who indicated that data-sharing needed to be mandated by funders and institutions, that simply recommending that researchers share their data would not be sufficient to change entrenched behaviours.

Significant variation was observed in how researchers from different disciplines perceive the factors that influence the decisions on whether to share data associated with a publication or funded research project. This indicates that researchers cannot be treated as a homogenous population, particularly with regards to education and engagement around the use of data-sharing systems at the university. A more targeted approach is recommended for dealing with the different faculties and Schools within the University, as they will each have a unique focus and particular concerns that need to be addressed in order to increase their engagement.

The lack of awareness by researchers of University systems and support for sharing data may indicate a requirement for increased levels of communication from the library to researchers, and

also better collaboration across the services that exist within the University to support research, such as the Research Office, and UniServices who deal with legal issues and intellectual property.

Concerns around ethical uses of data, ranked very highly as a factor influencing the decision to share data for many researchers. Some researchers were concerned that ethics approvals might preclude the sharing of data. Researchers may need additional training in how to write ethics approvals. They may be making promises to destroy all data, and maintain complete confidentiality, which might not be necessary in order to receive ethics approval for their particular study (Corti, Van den Eynden, Bishop, & Woollard, 2014).

Future Work

Given the low response rate of academic librarians to this survey, it would be interesting to perform a qualitative study investigating what researchers and librarians thought the roles of the library and librarians were with regards to supporting research and sharing of research data. It would also be interesting extend this study to other New Zealand universities to determine whether the patterns of discipline specific perception of the factors affecting data-sharing are consistent within disciplines but across institutions. Additionally to extend the study to Crown Research Institutes to investigate the different ways that researchers with a more commercial focus share data, and how the smaller special libraries operating within Crown Research Institutes interact with researchers and research data.

It was outside the scope of this research project, but reanalysing the results of this survey focusing on the demographic of time spent working in research could provide some interesting insight. Are researchers who are new to research less likely to share data because they are concerned with establishing a foothold in the discipline and do not want to lose publishing opportunities, or are they more likely to share data because they have been immersed in digital culture, the open access and open source movements throughout their research careers and it is an integral part of their research culture? Are researchers who have been working for more than 15 years less likely to share data because they are reluctant to change their behaviour, or are they more likely to share data, as indicated by Researcher 1, because they are established in their career and have less to lose by sharing data?

Conclusions

This study has shown that although there were some similarities, there were also observable differences in academic librarian and researcher perceptions of the barriers to research data-sharing. There were also statistically significant differences by research discipline in the level of impact that researchers thought particular factors had on the decision whether to share data sets. There may be a need to improve communications between the library and researchers with regards to the tools and services that they can offer. Library staff may also need additional training in support of University researchers, as a proportion did not feel confident answering questions about researcher data-sharing. The research discipline differences in perceptions of barriers to data-sharing mean that a “one-size fits all” strategy for education in and marketing of these services will not be the most effective strategy to address concerns and increase researcher engagement. Data-sharing and open access are gaining momentum as a movement and there is a great deal of interesting research still to be done in this area.

References

- Boulton, G. (2012). Open your minds and share your results. *Nature*, 486(7404), 441 (world view).
- Brown, E. (2010). "I know what you researched last summer": How academic librarians are supporting researchers in the management of data curation. *The New Zealand Library & Information Management Journal*, 52(1), 55–69. Retrieved from http://ndhadeliver.natlib.govt.nz/delivery/DeliveryManagerServlet?dps_pid=IE3006750
- Corrall, S., Kennan, M. & Azfal, W. (2013). Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends* 61(3), 636-674. Accessed from http://muse.jhu.edu/journals/library_trends/v061/61.3.corrall02.html on 14/8/2015
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and sharing research data: A guide to good practice*. London, UK: Sage
- Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA.: Sage Publications.
- Cullen, R. & Chawner, B. (2009) Institutional repositories and the role of academic libraries in scholarly communication. *Proceedings of the Asia-Pacific Conference on Library & Information Education & Practice, 2009*, 268-277.
- Dryad (2015). *About the organization*. Accessed at <http://www.datadryad.org/pages/organization> on 21/3/2016
- Egleston, B., Miller, S. & Meropol, N. (2011). The impact of misclassification due to survey response fatigue on estimation and identifiability of treatment effects. *Stat Med* 30(30), 3560–3572 doi: 10.1002/sim.4377
- Enke, N., Thessen, A., Bach, K., Bendix, J., Seeger, B. & Gemeinholzer, B. (2012). The user's view on biodiversity data sharing – Investigating facts of acceptance and requirements to realize a sustainable use of research data. *Ecological Informatics*, 11, 25-33. doi:10.1016/j.ecoinf.2012.03.004.
- Fecher, B., Friesike, S. & Hebing, M. (2015). What Drives Academic Data Sharing? *PLoS ONE* 10(2), e0118053. doi:10.1371/journal.pone.0118053
- Fowler, F. (2009). *Survey Research Methods* (4th Edn.). Thousand Oaks, CA: Sage Publications.
- Frank, E. & Pharo, N. (2015). Academic Librarians in Data Information Literacy Instruction: A Case Study in Meteorology. *College & Research Libraries*, preprint. Accessed from <http://crl.acrl.org/content/early/2015/09/24/crl15-828.full.pdf> on 26/4/2016.
- Ginsparg, P. (2011). ArXiv at 20. *Nature*, 476(7359), 145-147. doi:10.1038/476145a.
- Heidorn, P. (2008). Shedding light on the dark data in the long tail of science. *Library Trends*, 57(2), 280-299.
- Holbrook, A., Krosnick, J., & Pfent, A. (2007). The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In *Advances in Telephone Survey Methodology*. Hoboken, NJ: Wiley.

- Hole, G. (2011). Mann-Whitney test handout version 1.0. *Teaching Resources for Research Skills One*. Accessed from <http://users.sussex.ac.uk/~grahamh/RM1web/teaching08-RS.html> on 30/9/2016.
- Kahn, M., Higgs, R., Davidson, J. & Jones, S. (2014). Research Data Management in South Africa: How we shape up. *Australian Academic and Research Libraries* 45(4), 26-308.
- Lyon, L., Rusbridge, C., Neilson, C. & Whyte, A. (2010). *Disciplinary Approaches to Sharing, Curation, Reuse and Preservation: DCC SCARP Final Report*. Digital Curation Centre; 2010. Accessed from <http://www.dcc.ac.uk/sites/default/files/documents/scarp/SCARP-FinalReport-Final-SENT.pdf> on 15/9/2015
- Mauthner, N. & Parry, O. (2013) Open Access Digital Data Sharing: Principles, Policies and Practices. *Social Epistemology*, 27(1), 47-67. doi:10.1080/02691728.2012.760663.
- Mills, J., Teplitsky, C., Arroyo, B., Charmantier, A., Becker, P., Birkhead, T., ... Zedrosser, A. (2015). Archiving Primary Data: Solutions for Long-Term Studies. *Trends in Ecology & Evolution* 30(10), 581-589. doi:10.1016/j.tree.2015.07.006
- Murray-Rust, P., Neylon, C., Pollock, R., & Wilbanks, J. (2010). *Panton Principles, Principles for open data in science*. Retrieved from <http://pantonprinciples.org/> on 31/3/2016.
- National Academy of Science, National Academy of Engineering & Institute of Medicine (2009a). Ensuring Access to Research Data. In *Ensuring the integrity, accessibility and stewardship of research data in the digital age* (59-93). Washington, D.C.: National Academies Press.
- National Academy of Science, National Academy of Engineering & Institute of Medicine (2009b). Research Data in the Digital Age. In *Ensuring the integrity, accessibility and stewardship of research data in the digital age* (11-31). Washington, D.C.: National Academies Press.
- Nelson, B. (2009) Data sharing: Empty archives. *Nature*, 461(7261), 160-163. doi:10.1038/461160a.
- NVIVO (2012). *NVivo Qualitative Data Analysis Software Version 10*; QSR International Pty Ltd.
- Organisation for Economic Co-operation and Development (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding*. OECD Publishing.
- Pinfield, S., Cox, A. M. & Smith, J. (2014). Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. *PLoS ONE* 9(12), e114734. doi:10.1371/journal.pone.0114734
- Qualtrics (2016). *Qualtrics Survey Software*; Qualtrics Labs Inc. <http://www.qualtrics.com>
- Robbins, N., & Heiberger, R. (2011). Plotting Likert and Other Rating Scales. Proceedings of the 2011 Joint Statistical Meeting. Accessed from http://www.amstat.org/sections/srms/proceedings/y2011/files/300784_64164.pdf on 13/9/2016
- Saris, W., & Gallhofer, I. (2014). *Design, evaluation, and analysis of questionnaires for survey research* (Second ed., Wiley series in survey methodology). Oxford, England: Wiley-Blackwell.
- Savage, C. J. & Vickers, A. J. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS ONE* 4, e7078. doi: 10.1371/journal.pone.0007078

- Simons, N., & Richardson, J. (2013). Research data: the new gold. In *New content in digital repositories: the changing research landscape*. (115-146) Oxford, UK: Chandos.
- Swan, A. & Brown, S. (2008). *To share or not to share: Publication and quality assurance of research data outputs*. Report commissioned by the Research Information Network (RIN). Accessed from <http://www.rin.ac.uk/system/files/attachments/To-share-data-outputs-report.pdf> on 10/10/2015
- Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. (2014) Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research* 36, 84-90. doi:10.1016/j.lisr.2013.11.003
- Uhler, P. & Schröder, P. (2007). Open data for global science. *Data Science Journal* 6, OD36-OD53, doi: 10.2481/dsj.6.OD36.
- University of Central Florida Libraries Research Lifecycle Committee. (2012). The research lifecycle at UCF [Online Graphic]. Retrieved from library.ucf.edu/ScholarlyCommunication/ResearchLifecycleUCF.php on 22/4/2016.
- Vines, T. H., Andrew, R. L., Bock, D. G., Franklin, M. T., Gilbert, K. J., Kane, N. C., ... Yeaman, S. (2013). Mandated data archiving greatly improves access to research data. *FASEB Journal* 27(4), 1304-1308. doi: 10.1096/fj.12-218164
- Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., ... Rennison, D. J. (2014). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* 24 (1), 94-97. doi:10.1016/j.cub.2013.11.014
- Vogt, W. (2014). *Selecting the right analyses for your data quantitative, qualitative, and mixed methods*. New York: Guilford Publications.
- Whyte, A. & Tedds, J. (2011). *Making the case for research data management*. Edinburgh: Digital Curation Centre. Accessed from http://www.dcc.ac.uk/webfm_send/487 on 14/8/2015.
- Williams, S. (2013). Data sharing interviews with Crop Sciences faculty: Why they share data and how the library can help. *Issues in Science and Technology Librarianship, Spring* doi:10.5062/F4T151M8

Appendix A: Interview questions

1. How long have you been working in research/with researchers?
2. What is the main discipline that you/your researchers work in?
3. How familiar would you say you were with research data management and data sharing?
4. Is research data management and data sharing an important part of your/their research?
5. Why is/isn't it important to you/them?
6. Do you think that the importance of RDM and data sharing is changing?
7. How prevalent is RDM system use in your field? In UoA? Nationally? Globally?
8. What sorts of data do you/your researchers share?
9. What systems do you/they use?
10. Why do you/they use these systems in particular?
11. Do you/they use RDM and data sharing systems for internal/external collaboration?
12. How do you think researchers at UoA feel about using RDM and data sharing?
13. What do you think are the main barriers to researchers using RDM and data sharing systems?

(follow up points to possibly suggest) Do you think researchers have issues around ... ?

- a. Legal –
 - i. privacy
 - ii. copyright
- b. Time –
 - i. to convert and upload
 - ii. to learn system
- c. Economic –
 - i. cost of time
 - ii. implications for funding
- d. Technical –
 - i. System knowledge
 - ii. Lack of standards and infrastructure
- e. Ethical –
 - i. Misuse of Data
 - ii. Confidentiality
- f. Professional –
 - i. Competition
 - ii. Absence of professional rewards

Appendix B: Qualtrics survey questions

University of Auckland Datasharing Survey

Q3 Thank you for your interest in this survey. Today you will be asked about sharing of research data sets associated with published works and funded research, and about any barriers to use of data sharing systems at the University of Auckland.

Q6 My position at the University of Auckland includes work as

- a researcher (1)
- a librarian or information professional (2)

Q1 How long have you been working as a researcher, or with researchers?

- less than 1 year (1)
- 1-5 years (2)
- 5-10 years (3)
- 10-15 years (4)
- More than 15 years (5)

Q2 Under which research areas does your work (or the work of the researchers you support) fall? Select as many as are appropriate.

- Social Science and Education (1)
- Arts and Humanities (2)
- Life Science and Biology (3)
- Physical, Chemical and Earth Sciences (4)
- Health and Medicine (5)
- Mathematics and Statistics (6)
- Business and Economics (7)
- Law (8)
- Engineering (9)
- Computational and Information Sciences (10)
- Other (please specify) (11) _____

Q5 Is data-sharing part of the culture in your research community?

_____ Data-sharing is a part of the culture (1)

Q7 Do you (your researchers) regularly share your data with other researchers?

_____ Regularly share data (1)

Q8 Which of the following methods for data-sharing do you (they) use?

- Do not share data (1)
- Email (2)
- Secure sharing systems such as drop-box, GoogleDocs or equivalents (3)
- Community repositories such as Figshare, ArXiv, or Datadryad (4)
- Institutional repositories such as ResearchSpace the online repository at the University of Auckland (5)
- Funder repositories such as that operated by the Wellcome Trust (6)
- Portable storage devices such as USB drives (8)
- Other methods (please specify) (7) _____
- Not Sure (9)

Q9 How much impact do the following factors have on whether you (your researchers) will share data sets associated with a publication or piece of funded research?

- _____ Concerns around the degree of control researchers have over how data is shared (1)
- _____ The resources needed in time or money to share the data (2)
- _____ A lack of acknowledgement or benefits to the researcher for sharing data (3)
- _____ The data-sharing policy of the institution where research is carried out (4)
- _____ The data-sharing policies of the funding institutions (5)
- _____ The data-sharing policies of the publishers of research works (6)
- _____ The culture of data-sharing in the research community (7)
- _____ Whether the data will be of interest or value to other researchers (8)
- _____ Whether research from my discipline produces unpublished data (9)
- _____ The quality of the research that is being shared (10)
- _____ Whether sharing the data will affect the ability of the researcher to publish (11)
- _____ Issues around copyright and intellectual property (12)
- _____ Ethical issues such as confidentiality (13)
- _____ Concerns around inappropriate or adverse use of the data by others (14)
- _____ Usability of the data-sharing systems (15)
- _____ Long-term preservation and archiving of data within the data-sharing system (16)
- _____ How well the data-sharing system is structured and how accessible the data is (17)
- _____ How well the data-sharing system is integrated into other research systems at the institution (18)
- _____ Awareness of the availability of data-sharing systems (19)

Q10 Which of these factors have the most significant effect on whether you (your researchers) will share data sets associated with a publication or piece of funded research? (select up to three answers)

- Concerns around the degree of control researchers have over how data is shared (1)
- The resources needed in time or money to share the data (2)
- A lack of acknowledgement or benefits to the researcher for sharing data (3)
- The data-sharing policy of the institution where research is carried out (4)
- The data-sharing policies of the funding institutions (5)
- The data-sharing policies of the publishers of my works (6)
- The culture of data-sharing in the research community (7)
- Whether the data will be of interest or value to other researchers (8)
- Research from my discipline may not produce unpublished data (9)
- The quality of the research that is being shared (10)
- Whether sharing the data will affect the ability of the researcher to publish (11)
- Issues around copyright and intellectual property (12)
- Ethical issues such as privacy and confidentiality (13)
- Concerns around inappropriate or adverse use of the data by others (14)
- Usability of the data-sharing systems (15)
- Long-term preservation and archiving of data within the data-sharing system (16)
- How well the data-sharing system is structured and how accessible the data is (17)
- How well the data-sharing system is integrated into other research systems at the institution (18)
- Other (please specify) (19) _____

Q12 Are you aware of the data-sharing systems and support available through the University of Auckland?

- Yes (1)
- Maybe (2)
- No (3)

Q13 Which of the following tools and support teams for data-sharing at the University of Auckland are known to you?

- Figshare at University of Auckland (1)
- ResearchSpace, the University of Auckland repository (2)
- Research Support Services at the Library (3)
- The Center for eResearch (4)

Q11 Do you have any further comments or concerns about sharing of research data at the University of Auckland?

Q12 Thank you for your participation. This has been an anonymous survey, but if you wish to be contacted with a summary of the results of this study, you are welcome to supply an email address to which the results can be sent. This address will be kept separate to any responses to the survey and will only be used by the researcher to provide you with the study summary.

Appendix C: Human Ethics Application and Participant Information Sheet

ResearchMaster Human Ethics Application

Application ID : 0000023305

Application Title : Identification of Differences in Researcher and Information Professional Perception of Barriers to Uptake and Use of Data Sharing

Systems at the University of Auckland

Date of Submission : N/A

Primary Investigator : Joanne L Simons

Other Investigators : Christine King

Dr ChernLi Liew

27/07/2016 Page 1 / 9

Research Form

Application Type

What type of application is this?

Research Only

Information

Welcome to the Human Ethics Application Form

The following advice will assist you in completing this process:

Useful information

For information about Human Ethics, go to the [Human Ethics web page](#). A sample Participant Information Sheet and Consent Form are available on this page.

For help, please email the [Ethics Administrator](#).

Policy

You must read the [Human Ethics Policy](#) before beginning your application. Appendix B of the Policy contains a sample consent form, information sheet, and transcribing

confidentiality form which may be useful (see last page).

Health research may require HDEC approval. To find out if your research requires this, read the [HDEC Guidelines](#) or contact the chairperson of your committee for clarification. If

your research does require HDEC approval, a copy of the application must be sent to the Administrator of the HEC. Evidence of approval should also be sent to the HEC

administrator

Student research

If you are a student, check with your supervisor before filling in this form. You may need to complete School requirements before applying for ethical approval.

Student applications are automatically forwarded to supervisor(s) for approval when the form is submitted. Once the supervisor has approved it, the form is automatically forwarded

for committee review. Staff applications go directly to the Committee on approval.

Technical

This online system works best on Internet Explorer and Safari. It may not work on your iPad or tablet.

A guide to using this online form, which includes a workflow showing how the approval process works, can be downloaded [here](#).

If your application involves other researchers, you can use the Comments function of this form to communicate about the application with each other. Click on the Application

Comments or Page Comments icon on the top right of the screen to view and add comments. Comments left on the form once it is submitted will be visible to your Head of

School and committee reviewers, so **remember to delete any private comments before submitting the form.**

Process

You will normally receive an outcome of the review of your application within three weeks, unless you apply during an advertised closedown period (i.e. December and January).

NO part of the research requiring ethical approval may commence prior to approval being given.

To apply for an amendment or extension to an approved application, open the approved form and click on Apply for amendment/extension. You will then be able to complete the

Amendment/Extension page and resubmit the form.

Application Details

1. Ethics category code*

Human

Application ID

0000023305

2. Please select the appropriate committee below. Please note that:

Education applications are now handled by the Human Ethics Committee.

27/07/2016 Page 2 / 9

School of Information Management

4. Title of project*

Identification of Differences in Researcher and Information Professional Perception of Barriers to Uptake and Use of Data Sharing Systems at the University of Auckland

School or research centre*

Information Management

5. Please list all personnel involved in this project. Ensure that all are listed with the correct role. **If you are a student, do not add your supervisor here:**

you will be asked

to add this information on the next page.

Please ensure that only one person is listed as Principal Investigator.

To add a person, search for their Victoria ID if known, otherwise either their first or last name (whichever is the most unusual). Click on the magnifying glass to search for results.

Press the **green tick** at the bottom right corner to save the person record.

Add anybody who is involved in this project as:

Associate Investigator

Other Researcher

PhD Student

Masters Student

Research Assistant

Click on the help button if you are having difficulty adding people to the list.*

1 Given Name Joanne

Surname Simons

Full Name Joanne L Simons

AOU SCIM

Position Principal Investigator

Primary? Yes

6. Are any of the researchers from outside Victoria?*

Yes

No

7. Is the principal investigator a student?*

Yes

No

Next time you save this form or move to a new page, a Student Research page will appear after this one. Please complete the two questions on the Student

Research page.

Student Research

7a. What is your course code (e.g. ANTH 690)?*

INFO 580

7b. Please add your primary supervisor (the supervisor who should review this application).

If your supervisor is also the Head of School or the school ethics officer, you will need to discuss with your School who should approve this application as Head of School or

delegate. The supervisor and Head of School or delegate **must not be the same person.**

To add your supervisor, search for their Victoria ID if known, otherwise either their first or last name (whichever is the most unusual).

Press the **green tick** at the bottom right corner to save the person record. *

1 Given Name Chern

Surname Liew

Full Name Dr ChernLi Liew

AOU SCIM

Position Supervisor

If your supervisor is also the Head of School, you will need to assign a different person to the Head of School or Delegate role on the Signoff page.

7c. What is your email address? (this is needed in case the committee needs to contact you about this application)*

simonsjoan@myvuw.ac.nz; j.simons@auckland.ac.nz

27/07/2016 Page 3 / 9

Note that systemgenerated

emails (eg approval notifications) will not necessarily come to this address. Systemgenerated

emails will come to the email address stored for you in

Student Records. To change the record in Student Records, log into My Victoria, and click on Student Records. You will be able to update your email address from there.

Project Details

9. Describe the objectives of the project*

This research proposes to identify the differences in researcher and information professional (IP) perception of barriers to the uptake of data sharing systems by researchers at the University of Auckland in order to address all possible barriers during implementation and improve researcher use of new systems.

Research question

In what ways are the perceptions of the barriers to uptake and use of data sharing systems vary between science researchers and information professionals at the University of Auckland?

Subquestions

How

much use do University of Auckland researchers currently make of available data sharing systems?

Which

barriers to data sharing are considered by researchers to have the most impact on use of RDM systems?

Which

barriers to data sharing are considered by IPs to have the most impact on use of RDM systems?

Is

the relative importance of these barriers affected by the faculty or field a researcher or IP works in?

10. Describe the benefits and scholarly value of the project*

Although researchers are individually responsible for the management of their research data, IP staff are heavily involved in the development and implementation of these systems, as well as promoting the systems to researchers and educating researchers in their use. Gathering the views from both sides will more effectively inform library and ICT department decisions on how systems should be developed or chosen, as well as policy around their implementation and use. Identification of any differences in perception, and knowledge of the barriers that have the strongest negative influence on researcher use of data sharing systems will enable these factors to be addressed during promotion and education activities. They may also mean that additional training of IP staff will be required before promotion of the system to research staff is initiated. If barriers to use of a data sharing system are neutralised as much as possible during the early stages of implementation, it is expected that this would increase researcher uptake and use of the system. This, in turn, would increase the long term funding opportunities of research in the university.

11. Describe the method of data collection. Note that later in this form, in the Documents section, you will need to upload any relevant documentation such as interview schedule,

survey, questionnaires, focus group rules, observation protocols etc. Delays are likely if the interview questions are missing from the Documents section. *

Data will be collected by an online survey constructed on the Qualtrics survey tool. The questions and options included in the survey will be informed by reviews of the data sharing literature, and by additional topics and issues raised during consultative discussions with researchers and information professionals. The majority of questions will have multiple choice answers for rating the importance of various factors on a symmetric ordinal scale, with the option for a freetext

response in some cases. Consideration will be made as to the structure of questions and the order of questions within the survey as these are known to affect the way that participants make their responses.

1. Q11: While the majority of this application indicates data collection will be in the form of a survey, the approval letter from Hester Mountfield indicates interviews may be involved as well. Please confirm that in spite of this, data collection will solely be based around the survey.

I can confirm that the data collection will be solely based around the survey. I had discussed with Hester that at some point in the future it might be interesting to take these questions further with a study including interviews, which is probably why she covered both in the approval letter, but with regards to the study proposed in the Human Ethics Application, all data collection will be by the survey.

12. Does your research have more than one phase that requires HEC approval?*

Yes

No

Key Dates

If approved, this application will cover this research project from the date of approval

13. Proposed end date for data collection*

10/09/2016

14. Proposed end date for research project as a whole*

14/10/2016

Proposed source of funding and other ethical considerations

15. Indicate any sources of funding, including selffunding (selffunding means that you are paying for research costs such as travel, postage etc. from your own funds) (tick all that apply)

Internally funded

Externally funded

Selffunded

16. Is any professional code of ethics to be followed?*

Yes

No

16a. Name the professional code(s) of ethics *

27/07/2016 Page 4 / 9

LIANZA code of professional conduct

17. Is ethical approval required from any other body?*

Yes

No

18 Depending on the characteristics of your participants or location of the research, you may need to arrange permission from another body or group before proceeding. If this is

the case, explain and describe how you are addressing this*

Permission to use email lists at the University of Auckland to contact possible participants and invite them to complete my survey has been obtained from the Acting Director of Libraries and Learning Services, the Dean of the Faculty of Education, the Dean of the Faculty of Medical and Health Sciences, the Head of Department Physics, and the Head of School Environment.

Permission has also been requested to contact researchers in the Faculties of Arts and Humanities; Creative Arts and Enterprises; Business and Commerce; and Law. When permission is received from these faculties, the survey invitation would then be extended to researchers working in these areas at the University of Auckland.

Treaty of Waitangi

19. How does your research conform to the University's Treaty of Waitangi Statute? (you can access the statute from Victoria's [Treaty of Waitangi page](#))*

Māori research and information professional staff will be included in the groups invited to participate in this research, and it is expected that their input will have a significant effect on the relevance of this work. It is hoped that this work will contribute to consideration of Māori research interests in the way data is managed and handled.

Information about participants

20. How many participants will be involved in your research? If you are using records (e.g. historical), please estimate the number of records*

The number of participants will depend on the response rate to the email inviting them to participate in the online survey.

The pool of possible participants being targeted is approximately 3000, including Library staff, research staff and PhD students.

21. What are the characteristics of the people you will be recruiting?*

The population being surveyed for the proposed research has two components. The first subpopulation

is comprised of researchers based at the

University of Auckland working in the STEM, Medical research, Social Science or Arts fields. This population will include academics, postdoctoral researchers, research assistants and PhD candidates. The second subpopulation

is the IP staff who work with researchers at the University of

Auckland, and will include librarians, library assistants, library managers, and ICT staff such as developers and systems support.

22. Are you specifically recruiting any of the following groups?

Māori

Pasifika

Children/youth

Students

People who are offenders and/or victims of crime

People with disabilities

People in residential care

People who are refugees

Please indicate below.*

Yes

No

23. Have you undertaken any consultation with the groups from which you will be recruiting?*

Yes

24. Provide details of consultation you have undertaken or are planning*

I have discussed my project and areas of interest with two researchers and two information professionals working in different disciplines at the University of Auckland, points of interest raised in these discussion have been used to inform the construction of the survey questions.

25. Outline the method(s) of recruitment you will use for participants in your study*

The IP subpopulation

is much smaller than that of the researcher population at the university, and so all relevant staff identified will be invited by email to participate in the research. As there is no sampling of the IP population, it is not planned to send the invitation to a sample of the researcher population, but rather to send an open invitation to participate in the survey to the research mailing lists of the faculties and departments. The email will include a participant information sheet, and the email will include the information that completion of the survey will be taken as evidence of participant consent to participate.

26. Will your participants receive any gifts/koha in return for participating?*

Yes

No

27. Will your participants receive any other assistance (for instance, meals, transport, release time or reimbursements)?*

Yes

No

27/07/2016 Page 5 / 9

28. Will your participants experience any special hazard/risk including deception and/or inconvenience as a result of the research?*

Yes

No

29. Is any other party likely to experience any special hazard/risk including breach of privacy or release of commercially sensitive information?*

Yes

No

30. Do you have any professional, personal, or financial relationship with prospective research participants?*

Yes

No

30a. Give details and indicate how you will manage this*

I work in collaboration with many of the information professionals at Libraries and Learning Services at the University of Auckland, and the researchers at the School of Environment are part of the community I serve in my position as Subject Librarian, Environmental Science, Earth

Sciences, and Geography.

It will be emphasised that participation is voluntary, and will in no way affect an individual's employment or status at the institution, and as the responses are anonymous, that I will not know who has participated. It is not expected that any of the questions presented in the survey are likely to cause significant emotional or psychological distress to participants.

31. What opportunity will participants have to review the information they provide? (tick all that apply)*

They will be given a transcript of their interview

They will be given a summary of their interview

Other

They will not have an opportunity to review the information they provide

31a. Explain how participants will be able to review the information they provide*

The data will be collected by online survey, and the participants will be able to change the answers they make on each page, before moving on to the next part of the survey; until submission of the completed questionnaire.

31b. Please give details about why the participants will not be able to review the information they provide*

The data will be collected via online survey on an anonymous basis. Once the questionnaire is submitted, participants will not be able to return to review their responses.

Informed consent

32. Will participation be anonymous? **'Anonymous' means that the identity of the research participant is not known to anyone involved in the research, including**

researchers themselves. It is not possible for the researchers to identify whether the person took part in the research, or to subsequently identify people who took part

(e.g., by recognising them in different settings by their appearance, or being able to identify them retrospectively by their appearance, or because of the distinctiveness of the

information they were asked to provide).*

Yes

No

32a. How will anonymity be assured in terms of access to the research data?*

Names and personal identifiers of participants will not be collected as part of the survey process. Any identifying information given in freetext boxes will be removed from the data before analysis. Demographic information will not be collected at a level that would allow individual participants to be identified.

32b. How will anonymity be maintained in terms of reporting of the data?*

No identifying data will be collected or reported and data will be aggregated for analysis and reporting so that no individual participant responses are reported at any stage.

33. Will contributions of participants be confidential? Confidential means that those involved in the research are able to identify the participants but will not reveal their identity to

anyone outside the research team. Researchers will also take reasonable precautions to ensure that participants' identities cannot be linked to their responses in the future.*

Yes

No

33a. How will confidentiality be maintained in terms of access to the research data? (tick all that apply)*

Access to the research will be restricted to the investigator

Access to the research will be restricted to the investigator and their supervisor (student research

Focus groups will have confidentiality ground rules

Transcribers will sign confidentiality forms

Other

33b. How will confidentiality be maintained in terms of reporting of the data? (tick all that apply)*

27/07/2016 Page 6 / 9

Pseudonyms will be used

Participants will be named only in a list of interviewees

Data will be aggregated and so not reported at an individual level

Participants will be referred to by role or association with an organisation rather than by name

Names will be confidential, but other identifying characteristics may be published with consent

Other

34. How will informed consent be obtained? (tick all that apply to all phases of the research you are describing in this application)*

Informed consent will be implied through voluntary participation (anonymous research only)

Informed consent will be obtained through a signed consent form

Informed consent will be obtained by some other method

Access, storage, use, and disposal of data

35. What procedures will be in place for the storage of, access to and disposal of data, both during and at the conclusion of the research? (tick all that apply)*

All written material will be kept in a locked file; access restricted to investigator(s)

All electronic information will be passwordprotected;

access restricted to the investigator(s)

Any files stored on a USB will be encrypted or password protected*

All questionnaires, interview notes and similar materials will be destroyed

Any audio or video recording will be returned to participants and/or electronically wiped

Other procedures

*Storage of data on a USB or similar device should be avoided if possible.

35b. Will the data be destroyed immediately after the conclusion of the research?*

Yes

No

35c. How many years after the conclusion of the research will the materials be destroyed?

3.00

36. If data and material are not to be destroyed, indicate why and the procedures envisaged for ongoing storage and security

Data will be kept for three years after the research project has concluded in order that the research might be prepared for publication in peerreviewed journals. It is possible that during the review process, the PI on the project could be asked to reinterpret

or analyse the data in a

different manner. The data will be managed by the PI who will keep the electronic data in a passwordsecured

folder and the data will be

destroyed three years after the completion of the project.

Dissemination

37. How will you provide feedback to participants?*

Should participants wish to receive a summary of the research results and notification of their dissemination (publication, presentation at a

conference), they will be invited to leave a contact email at the conclusion of the survey. This data will be collected independently to their

survey responses (via a separate survey link), and will not be able to be connected to the other data they may have provided.

38. How will results be reported and published? Indicate which of the following are appropriate. The proposed form of publications should be indicated on the information sheet

and/or consent form*

Publication in academic or professional journals

Dissemination at academic or professional conferences
Availability of the research paper or thesis in the University Library and Institutional Repository
Other

39. Is it likely that this research will generate commercialisable intellectual property? (check the help text for more information about IP)*

Yes

No

Documents

40. Please upload any documents relating to this application. A sample Participant Information Sheet and Consent Form are available on the [Human Ethics web page](#).

Please ensure that your files are small enough to upload easily, and in formats which reviewers can easily download and review. To replace a document, click the tick in the column to the right of the document title. A green arrow will appear click this arrow to upload a new document. To add a new document click on 'Add New Document', at top right of the documents window. Then enter the document name in the box that appears and click the green tick. A green arrow will appear to the right of the file name which allows you to upload the new file. *

27/07/2016 Page 7 / 9

Description	Reference	Soft copy	Hard copy
Participant information sheet(s)	InformationSheetSurvey_		
Simons_2016_V2.docx			
Participant consent form(s)	Project_approval_letters_UoA_June2016.pdf		
Questionnaire or survey	Survey questions.docx		

Getting feedback on your application

You can seek feedback on your draft application, for instance from a mentor or a school Ethics representative **before** submitting it for review.

There are two ways of doing this:

1. Emailing your application to someone

You can email your application and any associated documents to another person at Victoria. To do this:

1. Click on the Action tab (on the left of the screen)
2. Click on Email application
3. Search for the person using **either** their first name **or** their last name (whichever is the most unusual)
4. Select the documents to include from the Document list (eg the Application PDF)
5. Click on Send or Zip and send

If you wish to send your application to someone outside Victoria, one option is emailing the application to yourself and then forwarding it.

2. Assigning a peer reviewer

You can add someone to the form as 'peer reviewer'. This means that they will be able to access your form by logging onto ResearchMaster. They will also be able to comment on

your form online. **If you are a student, don't add your supervisor to the form as a peer reviewer to get supervisor feedback, submit the form. Your supervisor may then make comments on it and ask you to review it further before it goes to the committee for review.** To do this:

1. Click on the Review tab on the left of the screen
2. Click on 'Peer reviewers'
3. Search for the person using their person code if known, or **either** their first name **or** their last name (whichever is the most unusual)
4. Click on the person's name
5. You may then also want to send the peer reviewer a notification, by clicking on Notify Peer Reviewer on the Actions tab

Checklist

Please check the information below and tick the box at the bottom of the page. Then follow the instructions to submit.

Have you read the Human Ethics Policy?

Have you included an information sheet for participants which explains:

the nature and purpose of your research;
the proposed use of the material collected
who will have access to the material collected
whether the data will be kept confidential to you
how anonymity or confidentiality is to be guaranteed?

Does your information sheet also include:

a statement about participants' right to withdraw and the final date for doing so (and is this also referred to in the consent form)?
a statement confirming that the research has been approved by Victoria University of Wellington Human Ethics Committee?
a statement about the destruction of the data at the end of the project?

(for students) your supervisor's name and email address?

Contact details for the HEC Convener should participants have ethics queries? (AProf Susan Corbett, email susan.corbett@vuw.ac.nz, telephone +644463 5480)

Have you used your VUW email address?

Have you included a written consent form?

If not, have you explained on the application form why you do not need to get written consent?

Are your information sheets and consent forms on VUW letterhead?

Have you included a copy of any questionnaire or interview schedule you propose using?

I have gone through the checklist and completed all the relevant tasks. *

Yes

Signoff

27/07/2016 Page 8 / 9

41. This section records signoff by all other researchers involved in the project (the other team members listed at Q.6). Principal investigators do not need to complete this section you signoff by submitting the application.

If coresearchers

are external to Victoria University they may be unable to access this site. In this instance, the Principal Investigator may sign off on their behalf. Please upload

evidence of the coresearchers'

signoff (e.g., a scanned email) to the Documents page.

To sign off, do ALL 5 of the following 5 steps::

1. Click on the pencil icon on the far right of the line with your name on it
2. Click on I Accept
3. Add the date
4. Click on the green tick icon on the bottom of the signoff window
5. Go to the Actions tab and click on 'Notify lead researcher that signoff is complete'

This question is not answered.

Please add the Head of School or delegate the

person in your School who is responsible for Human Ethics. This person will be notified when your application is approved, and will have online access to the form.

1 Given Name Christine
Surname King
Full Name Christine King
AOU SCIM

Position Head of School (or delegate)

Please ensure that you **save your application before submitting it**. Once you have saved your application, to submit it, click on 'Actions' on the left hand side of the screen

and then 'Submit for review'.

If you are a student, your application will go to your supervisor once you submit it. If you are a staff member, your application will go straight to the

committee for approval once you submit it.

If you have any feedback about this online form, please email it to ethicsadmin@vuw.ac.nz

Amendment or extension request (available only for approved applications)

43. Are you applying for an extension, an amendment, or both?*

Extension

Amendment

Both an extension and an amendment

This question is not answered.

Please check that you have answered all mandatory questions and have saved the application before submitting your form. Upload any amended documents (e.g. Participant

Information Sheet) at Question 40 on Documents page. To submit your form, click on the Action tab and then click on Submit for review

Subsequent Amendments (further requests after initial amendment request has been approved)

If you have already had an extension or amendment in the past, please answer the questions below:

44a. Do you have a second amendment request to make?

Yes

No

This question is not answered.

44b. Do you have a third amendment request to make?

Yes

No

This question is not answered.

44c. Do you have a fourth amendment request to make?

Yes

No

This question is not answered.

27/07/2016 Page 9 / 9

Participant Information Sheet

Research Project Title: Identification of differences in researcher and information professional perception of barriers to uptake and use of data sharing systems at the University of Auckland

Researcher: Dr Joanne Simons, School of Information Management, Victoria University of Wellington

As part of the completion of my Master of Information Studies, this study is designed to identify differences in the way that researchers and information professionals view barriers to sharing of research data and use of systems with which to share data at the University of Auckland.

Sharing of research data is increasingly becoming a pre-requisite for obtaining research funding from a range of providers, and tertiary institutions are investing in providing systems to develop this capability for their researchers. Information professionals at tertiary institutions are involved in development, maintenance of, promotion and training of staff in the use of data-sharing systems. However, researchers may have different views of the barriers to use and uptake of data-sharing systems when compared to information professionals.

This research aims to identify the factors that are considered to have the most impact on data-sharing by researchers and information professionals, and whether the importance of these factors is affected by the research discipline within which the data is produced.

Victoria University requires, and has granted, approval from the School's Human Ethics Committee.

I am inviting researchers, and information professionals who work with researchers, at the University of Auckland to participate in this research. Participants will be asked to fill in an online survey that should take no longer than 15 minutes to complete.

Participation is voluntary, and as the survey is anonymous, you will not be identified in any written report produced as a result of this research or identifiable to the others within the University of Auckland, including possible publication of the data in academic conferences and journals. All material collected will be kept confidential, and will be viewed only by myself and my supervisor Dr Chern Li Liew. The research report will be submitted for marking to the School of Information Management, and subsequently deposited in the University Library. All data collected from participants will be destroyed within 3 years after the completion of the project.

If you have any questions or would like to receive further information about the project, please contact me at simonsjoan@myvuw.ac.nz or telephone 021 344 521, or you may contact my supervisor Dr Chern Li Liew at ChernLi.Liew@vuw.ac.nz or telephone 04 463 5213.

If you have any queries about the human ethics approval for the project, the Convenor of the School of Information Management Human Ethics Committee is Dr. David Johnstone, who can be contacted by email at david.johnstone@vuw.ac.nz, or by telephone on 04 463 5877.

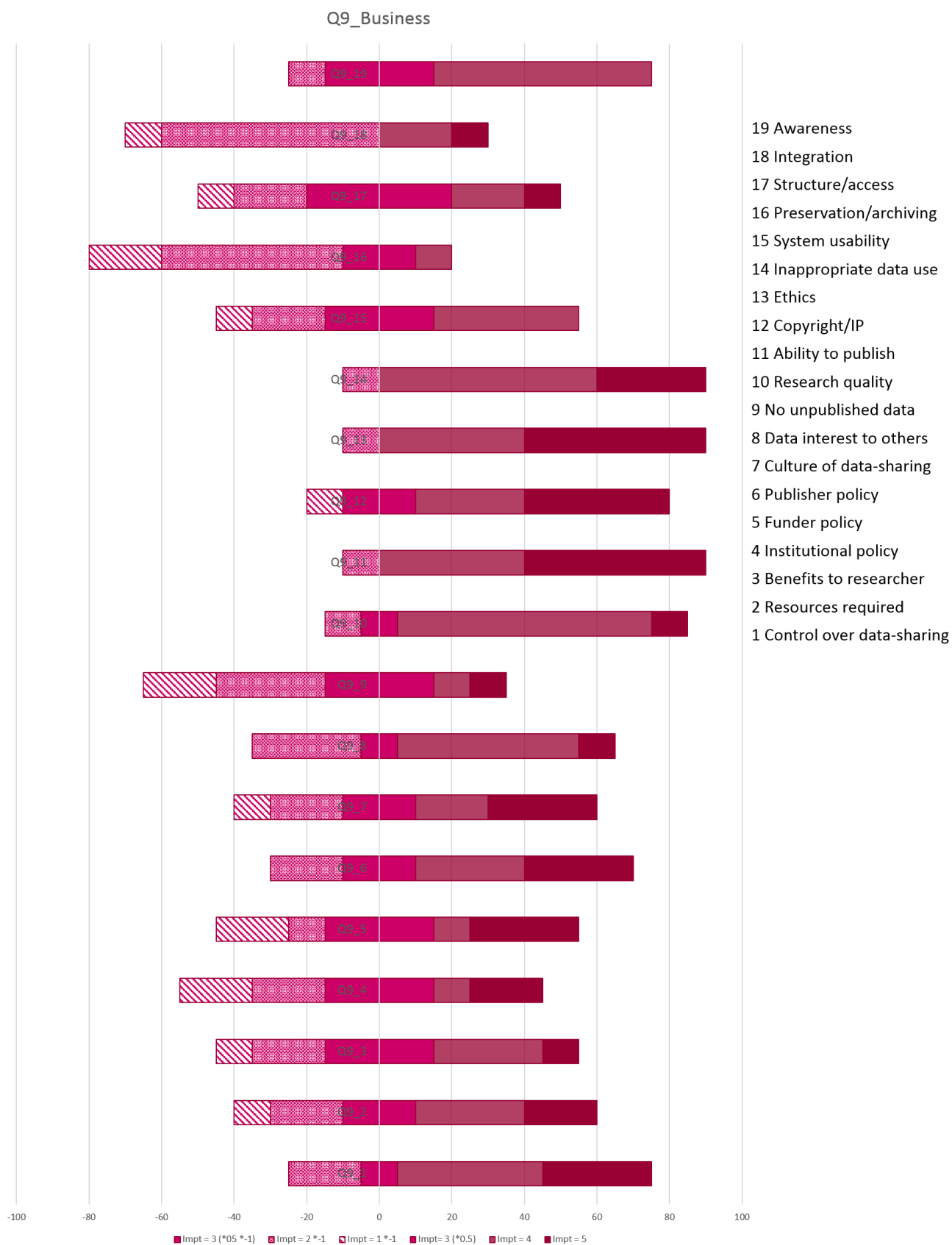
Dr Joanne Simons

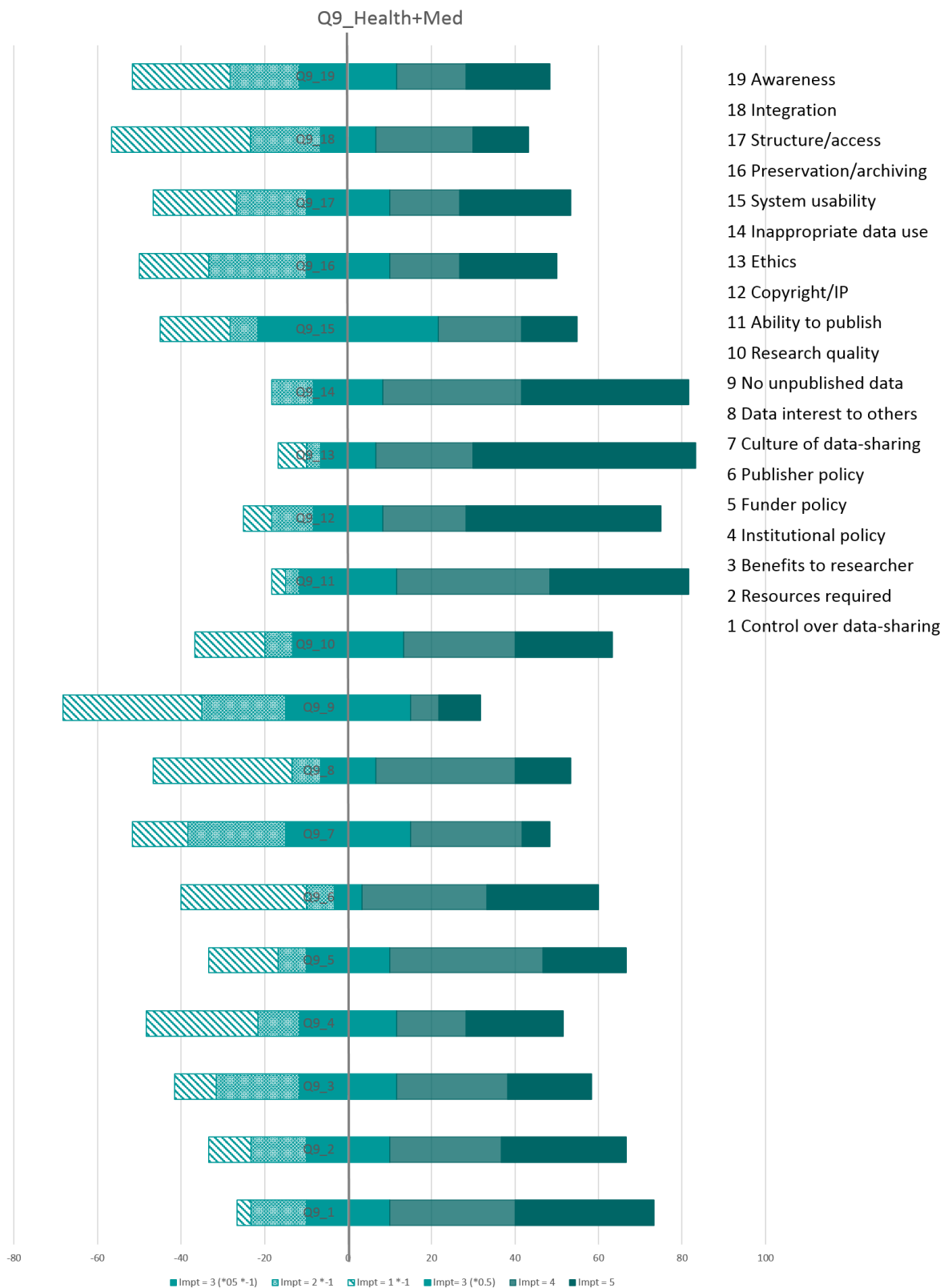
Appendix D Count data for Survey Question 9 – Relative impact of 19 factors on data-sharing

	Q9_1	Q9_2	Q9_3	Q9_4	Q9_5	Q9_6	Q9_7	Q9_8	Q9_9	Q9_10	Q9_11	Q9_12	Q9_13	Q9_14	Q9_15	Q9_16	Q9_17	Q9_18	Q9_19
Researchers																			
none	3	15	11	23	23	20	9	17	28	12	7	8	9	5	15	19	15	23	12
minor	15	12	16	14	9	14	19	11	18	7	7	12	8	11	14	21	12	25	13
some	16	15	19	18	18	13	20	9	20	20	19	16	13	14	25	14	25	12	25
definite	27	20	20	16	19	21	22	30	8	27	23	17	19	26	18	13	13	14	21
significant	22	21	17	12	14	15	13	16	9	17	27	30	34	27	11	16	18	9	12
total	83	83	83	83	83	83	83	83	83	83	83	83	83	83	83	83	83	83	83
Librarians																			
none	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0
minor	0	1	1	0	1	0	2	2	3	0	1	1	1	1	0	3	1	1	2
some	6	3	2	4	1	5	3	6	3	3	2	2	1	4	4	2	4	1	2
definite	2	2	5	3	4	1	3	0	2	4	6	3	5	3	4	3	3	4	3
significant	1	3	1	2	3	3	1	1	0	1	0	3	2	1	1	1	1	2	2
total	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

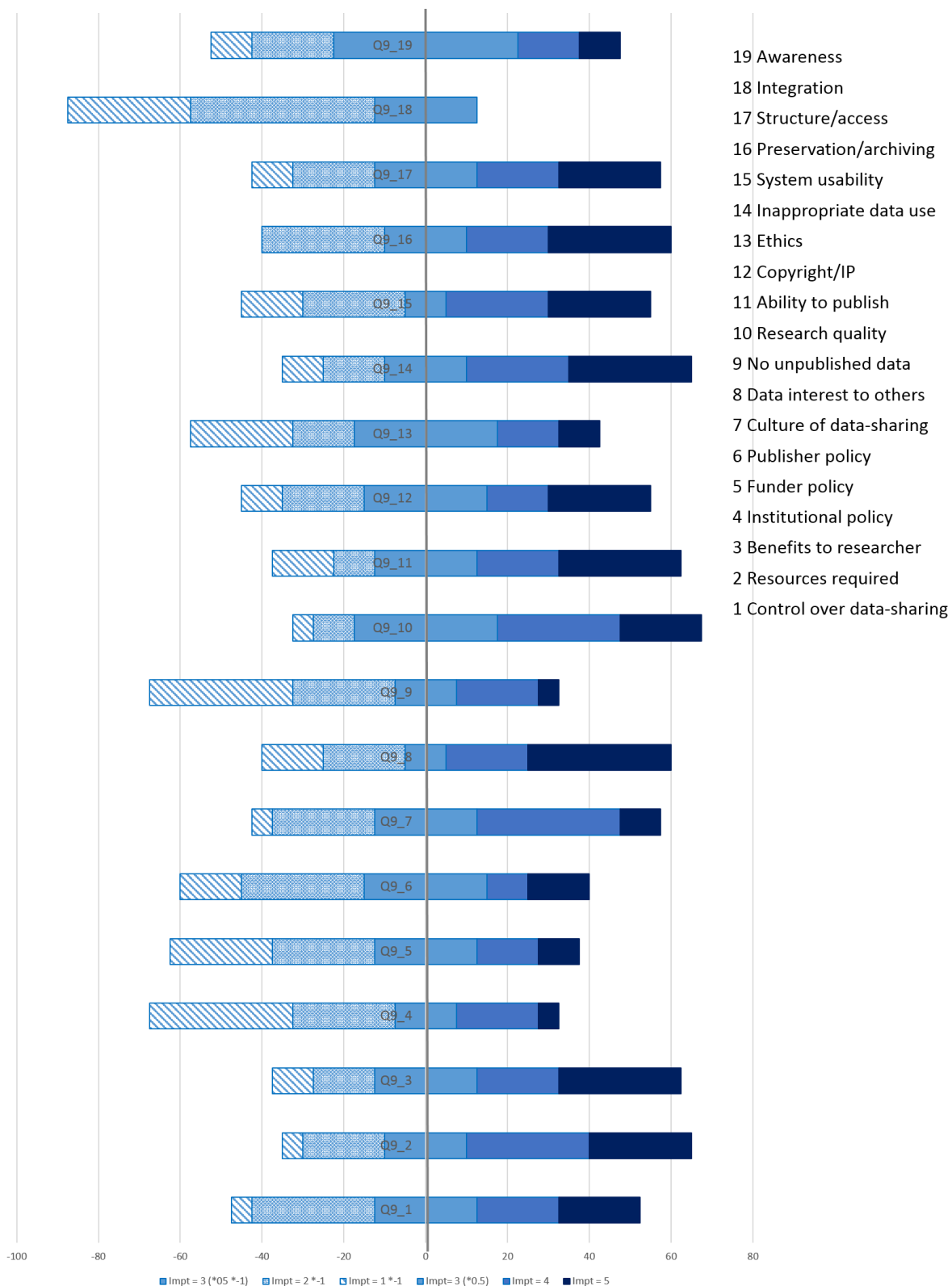
Question number	Factor name	Question number	Factor name
Q9_1	Control over data-sharing	Q9_11	Ability to publish
Q9_2	Resources required	Q9_12	Copyright and IP
Q9_3	Lack of benefits to researcher	Q9_13	Ethics
Q9_4	Institutional policy	Q9_14	Inappropriate use of data concerns
Q9_5	Funder policy	Q9_15	System usability
Q9_6	Publisher policy	Q9_16	Preservation/archiving of data
Q9_7	Culture of data-sharing	Q9_17	Structure of system/access to data
Q9_8	Data interest to others	Q9_18	Integration of systems
Q9_9	No unpublished data	Q9_19	Awareness of systems available
Q9_10	Research quality		

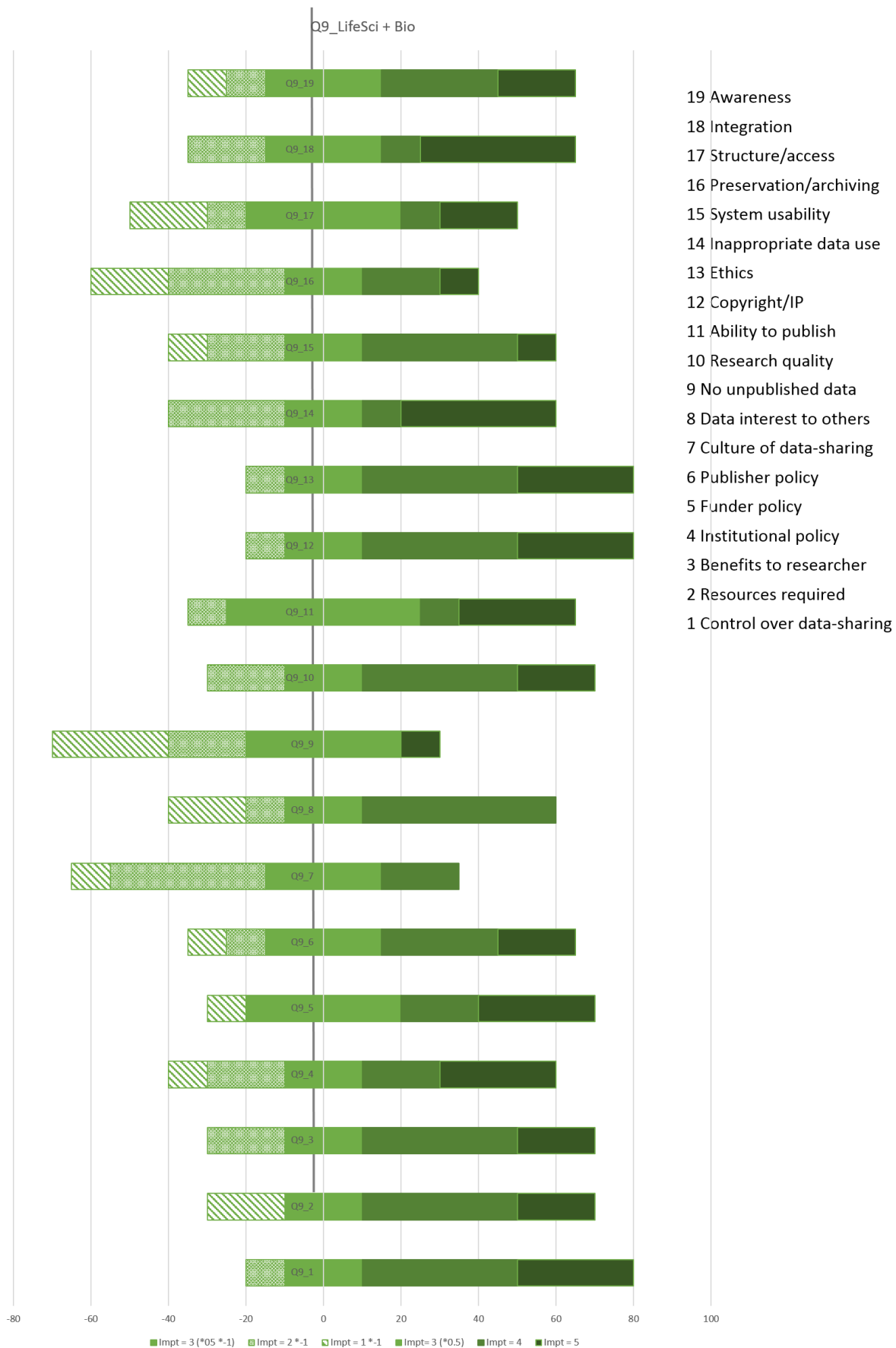
Appendix E: Discipline Graphs of all scaled impact factors from Question 9



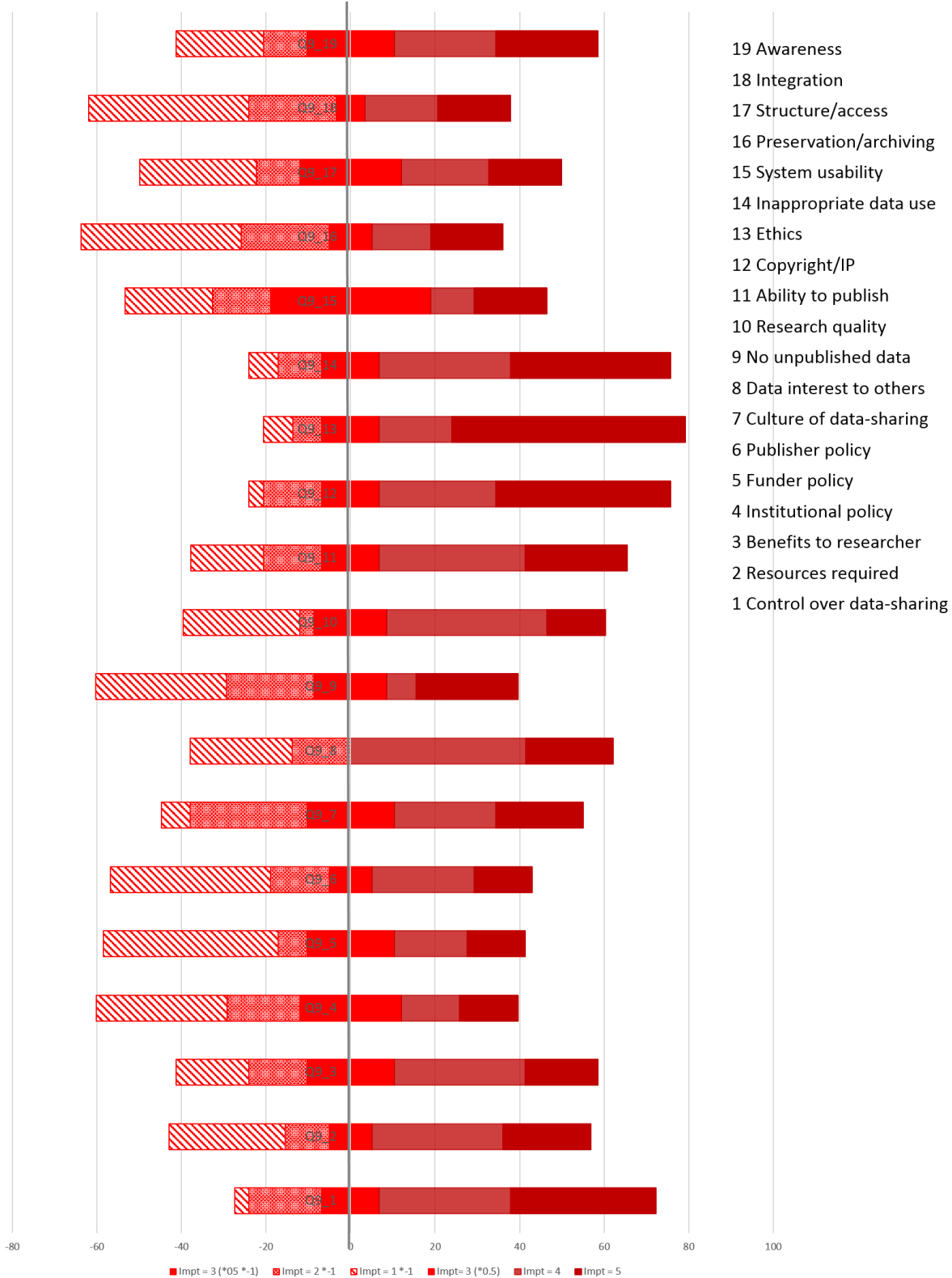


Q9_PhysChemEarth





Q9 – SocSci + Education



INFO 580

Assignment 2

Research Report

Joanne L Simons

ID- 300327897

Word count -