ABERYSTWYTH UNIVERSITY

Department of Information Studies

Academics' perspectives on open research data

A pre-implementation study for the University of Malta

RAELENE GALEA

2020

A dissertation submitted in part fulfillment for the Master of Science Degree in Management of Library and Information Services.

Contents

List of Figures

List of Abbreviations

DMP - Data Management Plan

EU - European Union

HE - Higher Education

IPR - Intellectual Property Rights

IR - Institutional Repository

MCST - Malta Council for Science and Technology

OA - Open Access

ORD - Open Research Data

RDM - Research Data Management

REF - Research Excellence Framework

UKRI - UK Research and Innovation

UM - University of Malta

UML - University of Malta Library

Declaration and statement of originality

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed

Date 13th November 2020

This work is the result of my own investigations, except where otherwise stated. Where *correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).
Other sources are acknowledged (e.g. by footnotes giving explicit references).
A bibliography is appended.

Signed

Date 13th November 2020

I hereby give consent for my work, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed

Date 13th November 2020

Abstract

University institutional repositories around the world have increased the visibility of the intellectual property of academic institutions. Implementation of such repositories has given rise to discussions about the value of open science and open access to research. Such research includes peer-reviewed publications and the data collected for the purpose of these publications. Research data is an important component as it may be used to verify the results and to generate new insights. However, researchers may not always be willing to share this information for a number of reasons, including intellectual property rights, financial constraints and the time that needs to be invested to make sure that the data is re-usable and interoperable. Access to data may also allow other researchers and the general public to verify the results and challenge their interpretation. Data may also be used to produce different outputs. Those in support of the open data movement argue that transparency is fundamental to democracy and that people have the right to access data that has been gathered through public funding. On the contrary, data sharing raises a multitude of issues related to confidentiality, intellectual property rights, data protection and so forth. The aim of this dissertation is to explore the possibility of opening access to digital data for research funded by the University of Malta, which is a parastatal entity. This will essentially serve as a pre-implementation study for putting in place a policy that requires academic members of staff to share their data. Collaboration by academics is key to the success of such a venture, therefore their point of views and concerns will be explored.

Acknowledgements

Sincerest gratitude goes to my dissertation supervisor, Dr Sarah Higgins, for her valuable feedback and support. My appreciation also extends to all the lecturing and administrative staff at the Department of Information Studies and the distance learning team at Aberystwyth University. I am honoured to have completed my studies with such an esteemed university.

I am grateful to the University of Malta Library Management team for supporting and encouraging me through this endeavour. They have been instrumental in the choice of my research area.

I would like to thank all the academic members of staff at the University of Malta who contributed to this study by taking part in the questionnaires and the interviews, especially the ones who participated in the pilot study and gave their feedback.

Heartfelt thanks to my partner, soulmate and best friend, Marie-Clare for proof reading my work and for reminding me of my capabilities when I failed to see them.

Chapter 1: Introduction

If I have seen further, it is by standing on the shoulders of giants.

- Newton, 1675

Science has seen a momentous shift towards open science, an approach that is based on collaboration and open access in research (Elsayed and Saleh, 2018; European Commission, 2016). Open science has been defined as the sharing of "data, scientific opinions, questions, ideas, folk knowledge, workflows and everything else as it happens" (Nielsen, 2009, p.32). The open access (OA) movement aims to increase access to scholarly information by providing free web access to peer-reviewed publications and research data (Needham and Lambert, 2019; Suber, 2012). OA promotes wider dissemination of research and it has been made possible through the implementation of institutional repositories that serve to provide a platform for depositing scholarly works, making them available to users. Increasing access to research promotes learning and innovation, whilst enhancing scientific development (Enago Academy, 2018).

OA is not limited to research output, but it also includes open research data (ORD). Whereas restrictions on access to research data used to be present in the past, the situation is being challenged by the open data movement (Kitchin, 2014); a movement which is built on ideologies of openness, contribution and collaboration. In line with democratic values, the goal of the open data movement is to make data available for everyone and not only to individuals who can pay for access (Kitchin, 2014).

The European Union (EU) advocates in favour of open access (OA) to publications and research data. EU member states are encouraged to create policies on a national level in order to fulfil the initiatives set out by the European Research Area (ERA). As one of the member states, Malta has put OA on its research strategy for 2020 (OpenAIRE, n.d.). The University of Malta (UM) and the Malta Council for Science and Technology (MCST) are working in collaboration to create a national open access policy. This dissertation is essentially a pre-implementation study for compiling and executing such policy, specifically with regards to open access to research data. The aim of the research is to explore ways in which scholars will adhere to the set policies and procedures. For this reason, the barriers for the implementation of a policy that favours ORD will be explored and ways to overcome such barriers will be identified. The research took place among resident academic members of staff working at the University of Malta.

In its mission, the University of Malta Library (UML) claims to "steer and advocate in favour of Open Science." The aim of the open science department is to increase accessibility of all types of scientific and research material by making it available for free. Although open data is one of the main principles of the department, the concept of ORD remains yet in its early stages. The knowledge gap that will be filled to satisfy the primary objective of this research is to find ways in which the UM can encourage the sharing of primary data, whilst also identifying the role that the library has in filling this gap. The research questions have been designed around assessing academic staff readiness for research data deposit and identifying ways to encourage the sharing of raw data. Specifically, the objectives of the research include:

- Assessing academic staff willingness to share and reuse primary research data.
- Identifying motivators and demotivators for sharing and reusing data.
- Identifying concerns regarding ORD and outlining ways to overcome such concerns.
- Exploring the benefits and setbacks of implementing a research data policy at the University of Malta.
- Identifying the role of the library in research data management.

The above objectives lead to answering the research questions set for this study, namely:

- To what extent are University of Malta academic members of staff willing to make their research data digitally available in open access?
- What concerns do University of Malta academic members of staff have on opening access to their research data?
- How can the University of Malta foster a culture of research data sharing among its researchers?

The research methods used for this study include desktop research and a combination of quantitative and qualitative methodologies. Desktop research was carried out to understand the

topic of open science and open research data, whilst investigating how these are handled in other universities, especially in the UK. Data collection was done through questionnaires, where quantitative methodologies were employed to analyse survey data; and interviews, where qualitative methods were used to examine the topic.

The dissertation is organised in six chapters. Chapter 1 gives an overview of the study, outlining the purpose, research questions and knowledge gap that this research aims to fill. Chapter 2 gives an overview of the literature. The aim of the second chapter is to investigate the context within which the study takes place; that is, the stand of the European Union on open research data and how other countries and Universities are handling the issue. Chapter 3 outlines the research methods that were employed to carry out the research. A justification for each methodology is also provided. Chapter 4 presents the results of the research, whilst chapter 5 presents an analysis and discussion of such results. Chapter 6 concludes the study and describes the limitations encountered in the carrying out of the research.

Chapter 2: Literature Review

If you have knowledge, let others light their candles in it.

- Margaret Fuller

2.1. Introduction

Sharing of data gathered for scholarly research can be defined as the process of making data accessible through dissemination. In fact, IGI Global, an international academic publisher based in the United States, argues that data sharing is the process of making research data available for access to others (IGI Global, 2020). Increasing access to research is pivotal, with open research data (ORD) being at the core of the discussion about open access due to its complexities. The open data movement encourages free data sharing, thus making data available to everyone without subscription or access fees. This, in turn, promotes collaboration between researchers which results in publications of higher standards.

Sir Isaac Newton, an influential scientist, mathematician and philosopher, known for his crucial contributions in the scientific upheaval, proclaimed that most of his discoveries would not have been possible without the input of scholars that carried out research before him (Vernon, 2017; Williams, 2014). It is in fact very common for progress and innovation to be built on preceding discovery. Improving upon a successful experiment through further investigation, or taking a different approach to a failed test can lead to new findings and innovative discoveries.

At face value data sharing may seem as a giveaway of intellectual property that rightly belongs to the author, however there are counterarguments that encourage ORD. This chapter will examine existing literature on open research data with the aim of outlining the advantages and drawbacks, and subsequently identifying ways of overcoming the barriers to free access to raw data.

2.2. Open Science

Before understanding the notion of open research data, the concept of open science will be explored. Open science is the practice of making research available to all those who are interested; this includes persons coming from a low socio-economic background, to individuals enjoying wealth and status (UNESCO, 2017; Woelfle et al., 2011). At the heart of the open science

movement is an innovative method to the scientific process; a method which is founded on collaboration and cooperation among researchers, through integrating new approaches for disseminating knowledge using digital tools (European Commission, 2016).

Open science is an umbrella term that includes much more than open access to research outputs and research data. FOSTER, an EU-funded project which aims to serve as a portal for making open science the way forward in the research industry, identifies 8 branches of open science (FOSTER, n.d.), being:

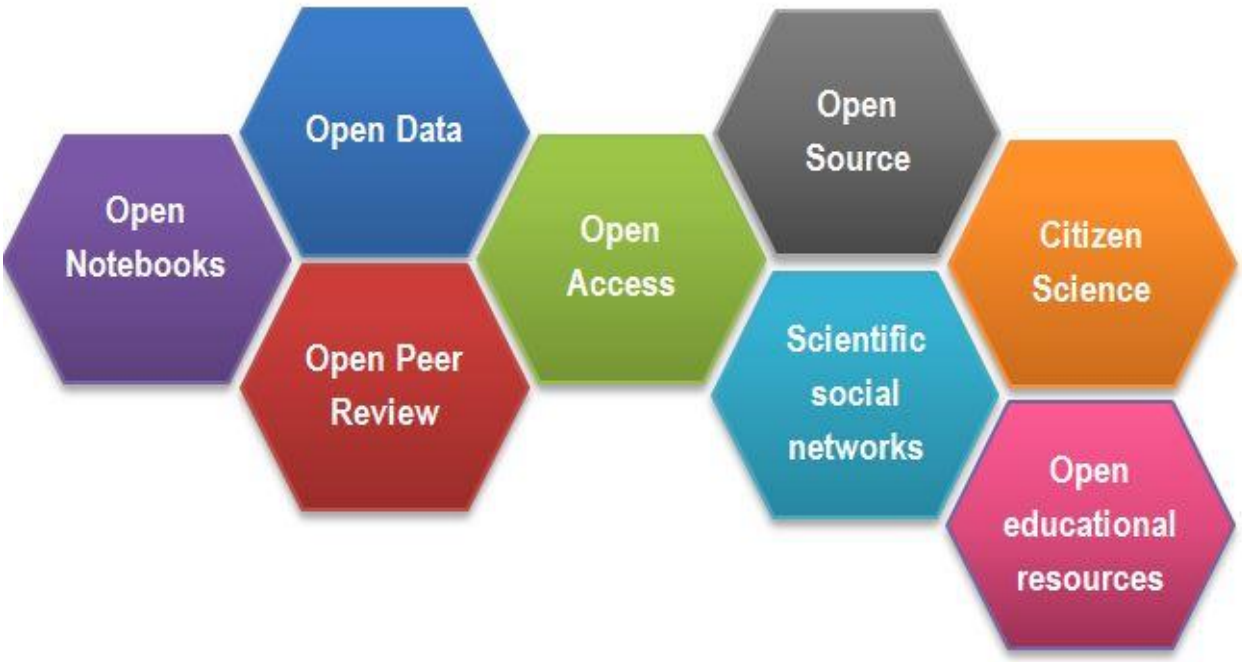

Figure 2.1: The 8 facets of open science. Source: https://www.fosteropenscience.eu/content/what-open-science-introduction

As the FOSTER project rightly notes, the attention of library and information centres is usually focused around open access to publications and open access to research data.

The leading objective of open science according to the Organisation for Economic Co-operation and Development (OECD) is to increase accessibility to publicly funded research, including publications and research data, by making such research available and publicly accessible with minimal constraint (OECD, 2015). In line with this, the Open Research Data Taskforce (2018)

argues that there are two main aims for open science and these should always be true for scientific research and research data:

- Open access: Research should be as open and accessible as possible. This implies that research should be in digital format, available online, free to use, and free of most copyright and licensing limitations.
- Research should be FAIR: **F**indable, **A**ccessible, **I**nteroperable and **R**e-useable.

2.2.1. Defining Open Research Data (ORD)

Research data is the raw, or primary data, collected by the researcher for the purpose of producing research. It is proof that the results obtained and the analysis carried out are reliable, thus they add to the credibility of the study (Emory University, 2020). Research data sharing may raise issues of validity and confidentiality, leading to reluctance from the researchers' end to openly share their work.

FOSTER (2017) defines open research data as:

> […] the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.
>
> (FOSTER, 2017)

The implications of data sharing vary according to discipline, since by nature some fields would have data that is more sensitive than others (Zuiderwijk and Spiers, 2019). In the science field, research and findings are improved through data sharing. In Nature Magazine (p. 435, vol. 534), Campbell (2016) reports that sharing of clinical information in medicine has helped scientists diagnose rare conditions. The example brought forward by Nature magazine is the discovery of the mandibulofacial dysostosis disease in Ottawa; the instances of this disease are so rare that a

single case could not identify it. Previous records of the disease together with their symptoms were recorded in a retrievable manner and this led to the identification of similar occurrences of the illness. This shows that availability of fundamental research data avoids repetition of the same experimentations and investigations, thus saving the researchers' time, energy and resources needed to carry out research that has already been done. In contrast, the same energy could be used to further build upon existing studies (Hahnel, 2013).

2.2.2. Benefits and setbacks of ORD

Open data provides a myriad of benefits for researchers, including an increase in visibility and citations (Piwowar et. al, 2007; Piwowar and Vision, 2013). In addition, access to research data increases the reliability of the research and allows for the verification of findings whilst providing insight for new research on its own or combined with other data sets (Emory University, 2020; Pampel and Dallmeier-Tiessen, 2019; Piwowar and Vision, 2013, Wouters and Haak, 2017, Zuiderwijk and Spiers, 2019). In the case of publicly funded research, policies and practices encouraging data sharing will secure an increase in research potential, thus ensuring greater returns from the public investment (OECD, 2007 in Pampel and Dallmeier-Tiessen, 2019).

Albeit the various benefits of ORD, scholars also note the disadvantages this includes. Creating and maintaining research data involves costs attributed not only to the creation and maintenance of the data itself, but also documentation, creation of metadata, and formatting of the data by the researcher. In fact, the requirement for researchers to adhere to set standards and present their data in a readable and interoperable manner is a chief hindrance to open data sharing (National Academy of Sciences (US), National Academy of Engineering (US) and Institute of Medicine (US) Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, 2009). Other barriers to opening access to research data include intellectual property rights (IPR), ethical and legal concerns, technical issues concerning the compatibility of data types, insufficient time, competitiveness and financial issues (Ferguson, 2014; Puniewska, 2014; Verhulst, 2015).

Research has shown that although scholars recognise the significance and merits of ORD and at times make use of open data themselves, they may be disinclined to share their primary data (Van den Eynden and Bishop, 2014). The absence of a reward system, lack of training, or a combination of both may lead to this lack of sharing.

Research shows that there are a number of factors that motivate scholars to share their data and these include the nature of the research, intrinsic motivators, technical know-how, policies and regulations, and the possibility of collaboration (Ferguson, 2014; Zuiderwijk and Spiers, 2019). Zuiderwijk and Spiers (2019) also outline the factors that encourage researchers to reuse open data. These include contextual factors, effort required, experience and the possibility of generating new data based on existing data sets.

2.3. ORD in practice

2.3.1. The local situation

The National Open Access desk in Malta is the University of Malta (UM) (OpenAIRE, n.d.), which is the leading higher education (HE) institution on the island. The UM has a student population of 11,500, which includes a total of 1,000 international students and another 450 on Erasmus exchange programs (OpenAIRE, n.d.). As a publicly funded body, with the largest population compared to other HE institutions on the island, and the largest number of courses offered (European Commission, 2020), the UM produces research in all academic disciplines, including humanities, social sciences, natural sciences, formal sciences and applied sciences (OpenAIRE, 2020). The open science initiative in Malta is being financially supported by Horizon 2020.

The UM set up its Open Science Department in 2017 (OpenAIRE, n.d.). The Open Science Department aims to offer support in the "[…] dissemination of the intellectual output, research and data produced by the University of Malta" (University of Malta, 2020). In addition, its remit is to foster a culture of open science among academics and researchers working at the UM. Since

the UM's institutional repository (IR), OAR@UM, is the only IR in Malta (OpenAIRE, n.d.), it also serves as a national repository. For this reason, any publications of Maltese cultural heritage are deposited, irrelevant of whether any contributions by University academics were made (K Slonina 2020, personal communication, 3 February). Over the years, the UM's institutional repository, OAR@UM, has seen an exponential increase in the deposits of peer-reviewed articles (K Slonina 2020, personal communication, 3 February). Statistics for yearly deposits are provided in figure 2.2, below. The statistics are limited to deposits of articles and exclude books, book chapters and conference proceedings, which are also deposited on the IR.

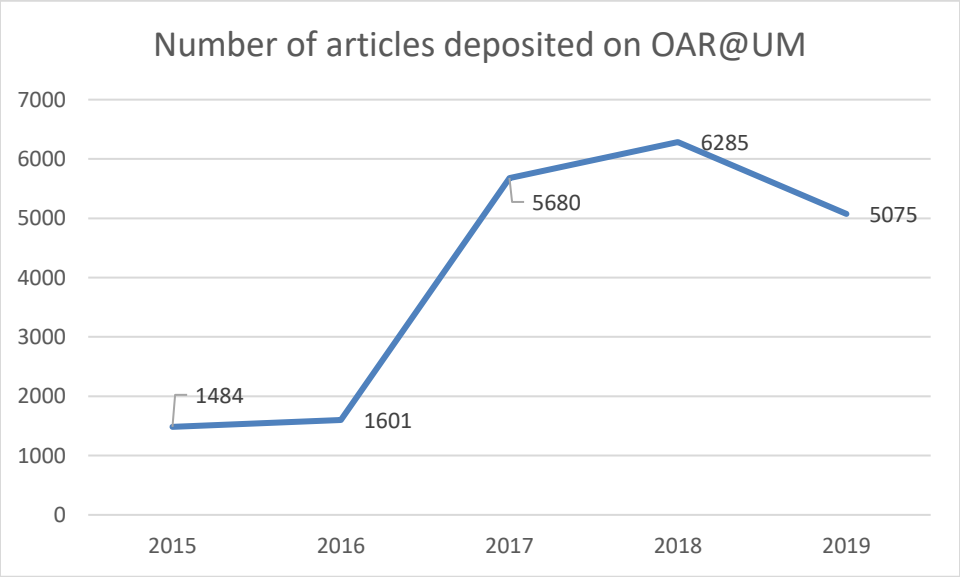

Figure 2.2: Number of articles deposited on OAR@UM per year

Academic members of staff were encouraged to submit their research outputs through constant urging, reinforcement, meetings and support given by the open science team. The open science office also showed their willingness to provide all the assistance needed to academics, including one-to-one training sessions. A good number of academics are now on board when it comes to sharing of research produced; however, this does not hold true for the data underlying such research. In fact, the UM's open access policy does not cater for the sharing of research data. The University of Malta Library was planning to introduce an ORD policy in 2020, however by the time this research was conducted the policy was not yet finalised.

Zhu (2020) notes that even though policies for ORD may be in place; in practice, persuading researchers to follow such policies may be a challenge. Along these lines, the University of Malta Library (UML) is aware that such policy may be met with resistance and academics may not be too willing to share their data, unless they believe that they can actually benefit from it. It can be argued that since the UM is funded by the Government through people's taxes, then the general public deserves access to the data gathered through research projects which have ultimately been paid by their tax contributions. In addition, sharing of data will reap the maximum benefit of such research.

2.3.2. ORD in the UK

The UK has been considered a world leader in research data management (Universities UK, n.d.). With open data policies in place, the UK mandates researchers to deposit data supporting their research in open data repositories. The national policy regarding open data is managed through a Research Excellence Framework (REF), which operates under the auspices of the UKRI (UK Research and Innovation). UKRI works with HE institutions and other research organisations (including the government) "to benefit everyone through knowledge, talent and ideas" (UKRI, n.d.). Their stand on open research stems from the belief that opening access to research increases its value. The UKRI Open Access principles state that:

> Open access policy should align and assist with the development of a research system that facilitates 'openness'. This includes opening up access to publications, research data, research methods and metrics.
>
> (UKRI, n.d., p.1)

Generally, the University of Malta follows the footsteps of UK Universities when it comes to the implementation and execution of policies and practices. In fact, over the past years, UML members of staff have carried out work placements and undergone in-service training in UK university libraries through EU funding. A desktop research exercise was carried out to find out

how many UK universities that use DSpace as the platform for publishing journal articles have already implemented an open research data policy. The directory of open access repositories (openDOAR) website (v2.sherpa.ac.uk/opendoar) was used to carry out this research. Out of the 31 institutions that satisfied the criteria below, a total of 14 institutions currently include datasets in their repository and 11 of these 14 institutions have a policy for open data.

- Repository type: Institutional
- Software name: DSpace
- Content types: Journal Articles
- Countries and regions: United Kingdom

Researchers conducting research in the UK are to follow the open science institutional and funder policy, available on the OpenAIRE website (openaire.eu). In addition, the majority of research funders require scholars to submit a data-sharing plan together with their request for funds (Zhu, 2020).

The recent political turmoil concerning Brexit and the UK's exit from the EU may impact research due to discontinuation of funding for research and science programmes (Nature, 2019), however the effects can only be evaluated in the future. Currently, the UK research and innovation (UKRI) is the leader in ensuring the continuation and development of research and innovation in the UK (UKRI, n.d.). This can be achieved through collaboration between all the institutions that produce research, such as universities, businesses, charity organisations and other research organisations.

The data policies of two UK universities, namely University of Cambridge and Loughborough University, were examined. Both universities have a Research Data Management (RDM) policy. Additionally, they have several policies and guideline documents related to data handling. For instance, the University of Cambridge have a set of guidelines on good research practice which includes information about collecting and handling primary data, record keeping, and disseminating and publishing results, all of which are crucial to data sharing (University of Cambridge, 2018). Loughborough University provide information regarding intellectual property, data protection, copyright and freedom of information (Loughborough University, 2020). Both

universities encourage researchers to publish their research data and opening access to it whenever possible. In its open research position statement, the University of Cambridge states:

> The University supports the academic freedom of researchers to pursue new knowledge, and to choose the means of dissemination, but within that free choice, the University encourages outputs of research, and where appropriate the accompanying data, to be 'as open as possible, as closed as necessary.'
>
> (University of Cambridge, Open Research Position Statement, 2019)

In its RDM policy, University of Cambridge proclaims that it is "committed to disseminating its research and scholarship as widely as possible" and in an effort to honour this obligation, it offers assistance to academics and students who need help to make their research data available. The management and curation of research data falls however under the responsibility of scholars who are duty-bound to abide by the University's policies including research procedures, ethics guidelines, records retention schedules, as well as funders' policies. Both the University of Cambridge and the Loughborough University assume responsibility for:

- disseminating information regarding policy requirements in connection with research data;
- providing infrastructure for data storage and other services needed for good research data management including back up and retention procedures;
- training and guidance on good practices in data management and data sharing, with Loughborough also providing templates for creating Data Management Programs (DMP);
- offering guidance and instruction related to ethical data management; this includes information regarding data protection, research integrity and Intellectual Property Rights (IPR).

(Research Data Management Policy Framework, University of Cambridge, 2019; Research Data Management Policy, Loughborough University, n.d.)

In addition to the above, the University of Cambridge has a dedicated website to provide guidance on research data management. (https://www.data.cam.ac.uk/). Whilst acknowledging their obligations, both Universities accentuate the researchers' liability for preparing a DMP, ascertaining compliance with lawful and moral restraints during the handling of data throughout its lifecycle, and planning the necessary time and finances needed for data management at the proposal stage. Researchers are also accountable for storing data appropriately and keeping backups, creating the metadata and providing information on how to access the research data in their research outputs.

2.3.3. The European Union on ORD

The European Union advocates in favour of open science and open access to research. Through the European Research Area (ERA) and their initiatives, member states are encouraged to create their own policies and mandates, in order to fulfil such initiatives. To date, Malta does not have a national open science policy, however the OpenAIRE website reports that open access (OA) is part of Malta's National Research and Innovation Strategy 2020. For this reason, OA is on the Maltese Government's agenda for 2020. Charge for the creation and completion of the national open science policy for Malta has been trusted to the UM and the Malta Council for Science and Technology (MCST). The latter is a public organisation whose aim is to develop national research policies and strategies.

2.4. The future of data sharing

Whereas the role of the librarian may have changed over the years, one thing has always remained constant; the librarian's fundamental duty is to provide access to information, whatever shape or form such information may come in (Pew Research Centre, 2013). This implies that, by nature, librarians are advocates for open access to research. Research does not have to be limited to the output result, but also unpublished data which can lead to further findings and

publications. Since data sharing increases access to research, it should be a priority for libraries and information centres worldwide.

A current concern for libraries is the management of research data (Pryor, 2012; Whyte and Tedds, 2011). This is mainly because digital data calls for measurements of preservation which are different and innovative when compared to the preservation of physical objects. In the case of digital items, preservation consists of processes that will guarantee continued usage and accessibility to the digital files (UNESCO, 2017). Dealing with the sharing of research data requires good management of datasets. This includes overcoming challenges of data handling such as storage, compatibility, technological obsolesce and security of sensitive data (Cox and Pinfield, 2014).

Data sharing is an inevitable progress that will develop new careers and change the role of libraries worldwide. Lupton (2019) explains the role of a data scientist, which is a career that specialises in the technical aspect of sharing research data. This includes specialisation in combining statistics, analytics and machine learning. In addition to the technical aspect, a change in the mind-set of researchers is essential for open research data to be successful. The accomplishment of this new philosophy is measured by the collaboration of the researchers. Getting scholars to adhere to policies without resistance is therefore within the remit of the librarian.

Pampel and Dallmeier-Tiessen (2019) suggest a number of factors that need to be employed in order to encourage academics to deposit their research data in open access. Primarily, it is imperative that institutions have a Research Data Management plan that supports researchers in data sharing. Data repositories need to be in place and support needs to be given by libraries or data centres in order for scholars to upload their research data in the correct manner. Universities should also have a number of incentives in place to encourage their academics to share their data. Lastly, data repository managers must ensure that the research data management plan is sustainable, especially in terms of human and financial resources. The next chapter considers the methodology employed for carrying out the research.

Chapter 3: Methodology

A research method is simply a technique for collecting data.

- Alan Bryman

3.1. Introduction

The research methods followed in the study, together with justifications for the selection of such methodologies, are outlined in this chapter. The chapter also gives information on the participants and the sampling methods that were used. Data collection practices and methods for analysis are outlined accordingly. The chapter concludes by giving an overview of the ethical procedures that were taken into consideration when carrying out the study.

3.2. Research methods and data collection

A desk research and mixed method approach, including quantitative and qualitative methodologies, were used for the purpose of this research. Data for this study was collected through questionnaires and interviews.

The idea for this study stemmed from the fact that the Maltese Government has included open access in its National Research and Innovation Strategy 2020 (Malta Ministry for Education and Employment, 2014). The University of Malta (UM) has been fostering the idea of open access amongst its academics in several ways, including the implementation of the first institutional repository in Malta, the inaugural of an open access department within the library and the introduction of an open access policy. The University of Malta Library (UML) has been the key stakeholder in fostering the idea of open science at the UM however, albeit the UML advocates in favour of open science, there still exists a research gap when it comes to open research data (ORD).

3.2.1. Desktop research

Desktop research, or desk research methodology, enables the researcher to gain knowledge of the field through secondary data (Travis, 2016). For the purpose of this study, desktop methodology was carried out to gain an understanding of open research data, which is the topic of investigation. Primo, by ExLibris, and DSpace were used to search for subscription material and

deposited electronic material. Both platforms were accessed through Aberystwyth University and the University of Malta portals for wider coverage. The initial stages of the research helped to better identify the research gap and formulate the research questions. With the aims and objectives in mind, further research was carried out to formulate the questions that were eventually asked in the questionnaire and subsequent interviews.

The researcher has affiliations with UK universities, as a student and ex-intern, as well as the UM, being an employee. For this reason, and in virtue of the fact that the UK is considered a world leader in research data (Universities UK, n.d.), the secondary research focused primarily on the way open research data is handled in the UK. A desktop search was performed with the aim of finding out which UK universities have implemented an ORD policy. This research was done through the OpenDOAR website which provides an exhaustive index of open access repositories and their respective policies (https://v2.sherpa.ac.uk/opendoar/). Additionally, Google searches were performed to gain insight on the barriers that universities encounter when implementing an ORD policy. The following quests were carried out and several crawled websites and peer-reviewed articles were consulted:

- struggles+universities+opendata
- researchers+views+open+research+data

Other academic material, including recently published Journal articles, were consulted for the literature review. Since ORD is a considerably recent and on-going topic for libraries, a good number of recent material was found on the subject.

3.2.2. Data collection techniques

Survey data was gathered through the use of a questionnaire mainly consisting of pre-determined close-ended questions. The questionnaire was designed using Aberystwyth University's subscription to the Jisc Online surveys, previously known as Bristol Online Surveys (BOS). The reason for selecting this survey tool was two-fold; it is user-friendly and it is GDPR

compliant (Jisc, 2020). Further in-depth data was collected through interviews held with participants who willingly agreed to further discuss the subject by participating in an online interview.

3.2.3. Mixed method approach: quantitative and qualitative research methodologies

The terms “convergent methodology”, “triangulation” and “mixed method approach” have all been used to describe a methodological approach whereby quantitative and qualitative methods are merged to produce fruitful outcomes (Bryman, 2006; Jick, 2006). The mixed method approach draws on the strengths of each method, that in turn “offset” the weaknesses of the other (Bryman, 2016). Bryman (2016) claims that the combination of both methodologies offers “completeness” in that it gives a holistic perspective of the topic being researched. The research method employed for a study may affect the findings (Powell and Connaway, 2004); therefore merging two methodologies to investigate a topic increases the reliability of the outcomes.

3.2.4. Quantitative research methodology

The questionnaire was divided into three major sections:

- data management
- data sharing
- policies and procedures

The questionnaire was distributed amongst 950 researchers working at the University of Malta (UM); comprising the total population of full-time resident academics at the UM. The aim of the questionnaire was to evaluate the current situation vis-a-vis open research data at the UM, whilst assessing the point of view of researchers from various backgrounds. The questionnaire consisted of 12 questions; some of which had further dependent sub-questions. The majority of the questions were dichotomous, asking for a simple “yes” or “no” answer; easing data analysis.

Some questions were trichotomous, having an additional neutral answer. Some questions consisted of multiple-choice answers where the participant was asked to choose one or more answers. The questionnaire also included a five-point scale question and an open-ended question; here participants could add feedback. The purpose of the latter was partly to help formulate the interviews and to possibly gather further understandings that were not thought of by the researcher.

The questionnaire was administered by email through the 'Marketing, Communications & Alumni Office' of the UM; this department has the contacts of all UM employees and it is a standard UM procedure to send such communication through it. The questionnaire gathered a broad opinion of UM academics on open research data, whilst serving as a sampling tool for discussing the topic further during interviews.

3.2.4.1. Pilot study

The questionnaire was piloted with 11 academics prior to being sent to the entire population. These academics were selected because the researcher was well-acquainted with them both on a professional and on a personal basis, hence a mutual level of trust existed. The researcher believed that the academics chosen for the pilot study would provide constructive feedback.

Bryman (2016) highly suggests the piloting of self-administered questionnaires, stating that it not only ensures that the research tool performs well, but it is also crucial for making sure that the questions are clear since the interviewer would not be present when the participants are answering them. In addition, Sincero (2012) claims that pilot testing increases efficiency by increasing the reliability of results. This is because through pre-testing the researcher can ensure that the instructions and questions are clear to the respondents. In fact, 6 of the academics who took part in the pilot study communicated valuable feedback which led to the questionnaire being updated. The other 5 participants did not suggest any changes.

3.2.5. Qualitative research methodology

Purposive sampling was employed to select participants who would contribute to the study by answering a set of questions during a semi-structured interview. Purposive sampling is explained by Bryman (2016) as a technique whereby the researcher selects participants in a strategic manner that ensures that the data gathered is relevant to the research questions.

The participants who took part in the questionnaire were asked to provide their contact details if they were willing to be interviewed. Academics who agreed to be interviewed were stratified according to the discipline in which they teach and perform research. In this way, the researcher ensured representation from different backgrounds and research interests. This, in turn, heightened the reliability of the findings since studies have shown that levels of data sharing vary according to discipline (Houtkoop et al, 2018; Severin et al, 2020; Wouters and Haak, 2017; Zuiderwijk et al, 2019). A total of 3 interviews were carried out, thus representation from every discipline was not possible, however out of the academics who agreed to be interviewed, the ones chosen were from faculties that enrol a large number of students every year and with academic backgrounds that could presumably have strong opinions on open research data; both positively and negatively. The reason why 3 interviews were held is two-fold. Firstly, only a small number of academics agreed to be interviewed, out of whom not each participant got back to the researcher when actually contacted for the interview. Secondly, a purposive sampling technique was preferred over random sampling owing to the fact that research has shown that a higher level of efficiency is achieved when employing the former (van Rijnsoeve, 2017; Vasileiou et al., 2018).

Initially 4 academics from different disciplines were selected to be interviewed. The disciplines selected included:

- Information and Communication Technology (ICT), which hypothetically as a discipline would be more knowledgeable and interested in the technical aspect of the data sharing process.

- Health Sciences was chosen because research in this field may contain patient data which could raise data protection issues. Conversely, sharing clinical datasets could improve scientific research (Kalkman et al, 2019).
- Education was selected because a lot of research in this discipline would involve human participants that are below consenting age, therefore gaining permissions and ethics clearance may be a struggle.
- The social sciences were chosen because they also involve human participants, however the person selected to represent the social sciences is an advocate for open data, therefore the researcher hoped that this participant could provide a better perception regarding the situation on open research data in Malta.

The academic from the Faculty of Health Sciences did not get back to the researcher when contacted for the interview and given the information sheet, so the latter contacted another academic from the same faculty, who also did not get back to the researcher. Therefore no representation was present from the Health Sciences discipline.

Due to the pandemic caused by the Covid-19 coronavirus, the interviews were held via video-conferencing using Zoom and Google Meet. The video-conferencing tools were being recommended by the University of Malta during the rise of the pandemic, thus both the researcher and the participants were familiar with the tools since they were using them on a frequent basis for meetings and lectures.

3.3. Validity of the methodology

Even though the mixed method approach involves additional resources for collecting and analysing two different types of data (Creswell & Plano Clark, 2011, in McKim, 2017), it is advantageous as it increases the validity of the research findings, making the research more truthful and comprehensive (Daniel, 2011). Additionally, mixed methods research draws on the strengths of both qualitative and quantitative methods.

The questionnaire asked the same questions to all participants, with only a small number of questions allowing participants to add their opinions without choosing from pre-set answers. Since the questionnaire was distributed to a large number of persons, having close-ended questions allowed for easy analysis of the data. On the other hand, the aim of the interviews was to delve deeper into the subject by allowing interviewees to sound their opinions through open-ended questions. In fact, interviews are used when informative, in-depth data is being sought from the individual (Pickard, 2007). Hence, the combination of both methods increased the validity of the results. That being said, the results would have been even more valid if the sample size was larger since this would have reduced the margin of error. This is further discussed in section 5.2 of this dissertation.

The idea of having focus groups instead of interviews was considered for this research. Having people from different disciplines discussing the benefits and drawbacks of ORD would have substantiated the hypothesis that the researchers' background affects their opinions on ORD. Since the focus group method gives the researcher the chance to examine multiple interactions on a single topic in a relatively short period of time (Powell & Connaway, 2004), the researcher would have presumably been able to grasp further concerns during the discussions. Several reasons contributed to the dismissal of the focus groups idea. Firstly, more participants would have been needed to organise the focus groups and that would have been an issue. Secondly, from previous personal experience, it is quite difficult to find a common timeslot which would be convenient for all the participants, especially when the subjects in question are academics. Additionally, and most importantly, as explained further on in section 4.2, this study was conducted during a partial lockdown on the Maltese islands following the outbreak of a world pandemic. Conducting interviews via video conferencing worked out well, however organising online focus groups would not have been practical.

3.4. Methods of Analysis

The analysis component of the Jisc online survey tool was used to examine data generated by the questionnaire. Jisc produces graphical representations of the data gathered, which eases the analysis. A thematic analysis approach was employed to present the results and analysis categorically according to the research questions and objectives of the study.

For the qualitative component of this study, analysis was carried out by coding of the answers according to the content. This is referred to as content analysis, which has been explained by Medelyan (2019) as the categorisation of topics in qualitative data. The starting codes were set according to the research questions, and new codes were created for new information that developed during the interviews.

Manual analysis was used to code the data since the size of the dataset could be handled easily. In addition, although automatic analysis is unbiased, manual analysis is more accurate (Medelyan, 2019; Squires, 2020)

3.5. Ethical Considerations

The questionnaires and interviews used for data collection were held with consenting adults and no sensitive data was collected. Ethics clearance was obtained from Aberystwyth University and from the University of Malta. The former was needed since it is the institution certifying this dissertation, whilst the latter was needed since data subjects were employed with this institution. The objectives of the research were communicated with participants in writing (see Appendix I). The only personal data that was collected for this study were the names and email addresses of participants who voluntarily wished to take part in an interview. An information sheet and consent form (see Appendix II) were sent to the participants who were then selected to be interviewed. The participants were asked to sign a consent form agreeing to the data usage conditions outlined in the information sheet. Such personal information was held and safeguarded in line with GDPR requirements. The identity of the data subjects was not disclosed

or transferred to third parties. Participants were made aware that they could withdraw from their involvement in the research at any point, without providing any reason for doing so. All efforts were made to present data in the most objective manner possible so as to avoid biases from the researcher's end.

Chapter 4: Results

The common facts of today are the products of yesterday's research.

- Duncan MacDonald

4.1. Introduction

This chapter starts off by reporting the number of respondents that completed the questionnaire and the number of participants that contributed to this study by sitting for an interview. The chapter also presents a contextual description of the results obtained through the questionnaire and the interviews. The results are presented as four main themes that are directly related to the research questions and objectives set out at the beginning of this research. Themes are further divided into narrower topics. The areas presented in this chapter constitute ideas that emerged from the questionnaire and the interviews carried out. The four main themes include:

- Researchers' opinions on open research data (ORD)
- Data management and storage practices
- Researchers' concerns on data sharing
- The role of the library in ORD

4.2. Online questionnaire (Appendix III)

A total of 88 responses were received for the questionnaire, which was available between the 22nd of March until the 12th of April 2020, both dates included. 45 responses were received in the first 15 days, after which a reminder was sent and another 43 responses were gathered. The questionnaire was distributed among the entire population of resident academics who work at the University of Malta, which amount to 950 potential respondents. The distribution medium used was email and it was sent through the UM's Marketing, Communications & Alumni Office on behalf of the researcher. The response rate was of 9.26% (88/950).

A number of factors could have contributed to the low response rate. Firstly, since several questionnaires are distributed through the UM's 'Marketing, Communications & Alumni Office', it may be the case that some individuals disregard such emails. Secondly, this study was carried out during the outbreak of a virus that became a world pandemic; the COVID-19 coronavirus. A few days before the distribution of this questionnaire, the Maltese government had closed all

education institutions, including the University, in a measure to minimise the quick spread of the infection. This measure meant that academics had to find alternative ways to deliver lectures through technologies such as video conferencing. Therefore, the dissemination of the questionnaire happened in a time where the subjects under study were busy focusing on performing their jobs' duties at home during an anomalous situation, whilst making sure that they kept themselves and their families safe during a viral epidemic; finding the time to answer a questionnaire on open data was definitely not a priority.

4.3. Interview (Appendix IV)

Three interviews were held with researchers from different research backgrounds:

- Information and Communications Technology;
- Education;
- Social Sciences.

The same questions were asked to all three participants.

4.4. Themes

4.4.1. Researchers' opinions on open research data

This dissertation aimed to understand how the University of Malta can encourage data sharing by their academic staff. To understand this, the attitude of the University's academics towards both sharing their own data and using others' data were examined.

4.4.1.1. Sharing own data

Participants were asked whether they had ever shared research data generated for their research outputs; the majority of the respondents (54 out of 88 participants, 61.4%) stated that they had shared their research data, however only 11 (20.4%) had shared it in open access. Other participants had shared data with members of their research team, an external individual known to them, an external individual not known to them, or a smaller group of external people. Participants were allowed to choose more than one answer for this question. A graphical representation of the data collected is provided as figures 4.1 and 4.2, below.

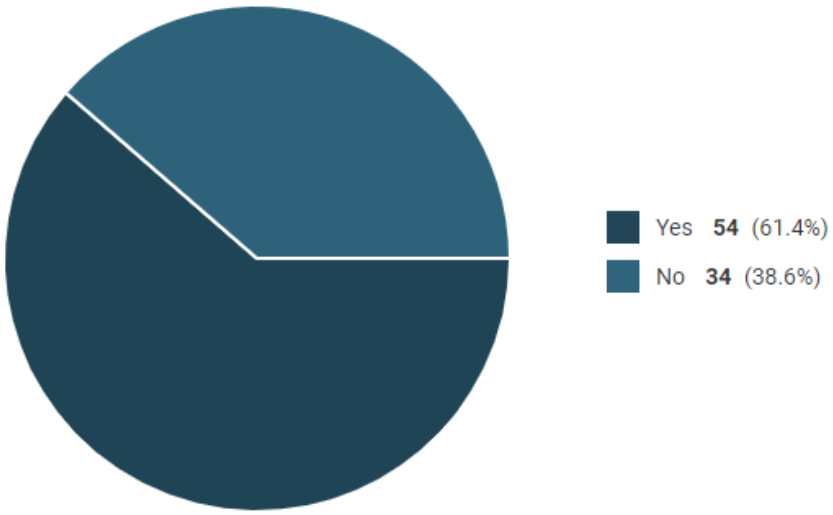

Figure 4.1: Question: Have you ever shared research data that you gathered for your research and publications?

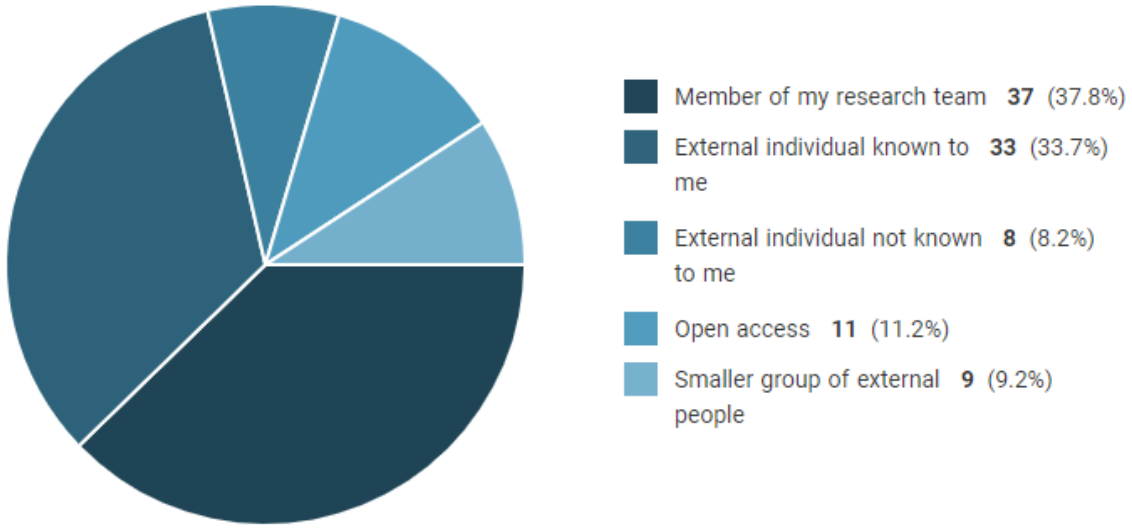

Figure 4.2: Question: Who did you share your data with?

Researchers who had previously shared their research data were asked to select the reasons behind the sharing of data. 37 respondents (68.5%) shared primary data due to the belief that in so doing they would contribute to the research field by allowing other researchers to use their data in order to generate new insights. 26 participants (48.1%) believe that sharing their raw data increases the credibility of their research. 20 (37%) have shared their research data in order to increase the visibility of their research. A smaller percentage of the respondents shared their data because it was required by the publisher or by the funding body; 13% (7) and 9.3% (5) respectively. 6 participants (11%) chose "other" and their replies are quoted below:

- "To support students' research. To generate funds to our institute"
- "To generate more knowledge with my colleagues"
- "We agreed initially to pool data as a team"
- "To write a new collaboration research proposal"
- "Shared with UOM colleague who asked for help"
- "Collaborations"

The influence of the faculty's position on data sharing hypothetically affects scholars' openness to sharing. For this reason, participants were asked whether or not their faculty encourages data sharing. The answers given for this question were quite balanced, with 18 (20.5%) positive and 18 (20.5%) negative answers. The other 52 (59.1%) participants stated that the faculty's position in this regard is neutral.

4.4.1.2. Using shared data

To test the readiness of scholars to use data generated by other researchers, the participants were asked if they ever made use of research data generated by others. 47 (53.4%) participants answered in the affirmative, whilst 41 (46.6%) answered adversely. Out of the 47 respondents who had used raw data produced by others, 24 (51.1%) said that such data was freely available in open access whilst 23 (48.9%) said that it was not available in open access. 33 (56.9%) of the respondents used such data to generate new findings whilst 17 (29.3%) of the respondents used it to verify the research output produced by that data. Other participants gave different reasons for using raw data generated by other scholars, such as collaborations, training purposes, data analysis and the generating of new data.

The results surrounding researchers' opinions on open research data show that academic staff at the UM are most inclined to share their data with members of their own research team or other individuals known to them. Thus, collaborations among small groups of researchers seems to be popular, however the idea of sharing primary data with the general public is less favoured.

4.4.2. Data management and storage practices

Managing data is crucial for data sharing, with the storage of data being a critical part of the process. This may hypothetically affect the researchers' willingness to engage in data sharing. This dissertation set out to pinpoint the concerns that academics have towards making research data available in open access, thus examining scholars' readiness to engage in the additional administrative tasks that ORD entails, was pivotal.

Scholars taking part in this study were asked about their data storage practices with a total of 66 (75%) respondents affirming that they do preserve the data generated for their research output, 4 (4.5%) participants claiming that they do not preserve such data and 18 (20.5%) reporting that they sometimes preserve research data. A graphical representation is provided as figure 4.3, below.

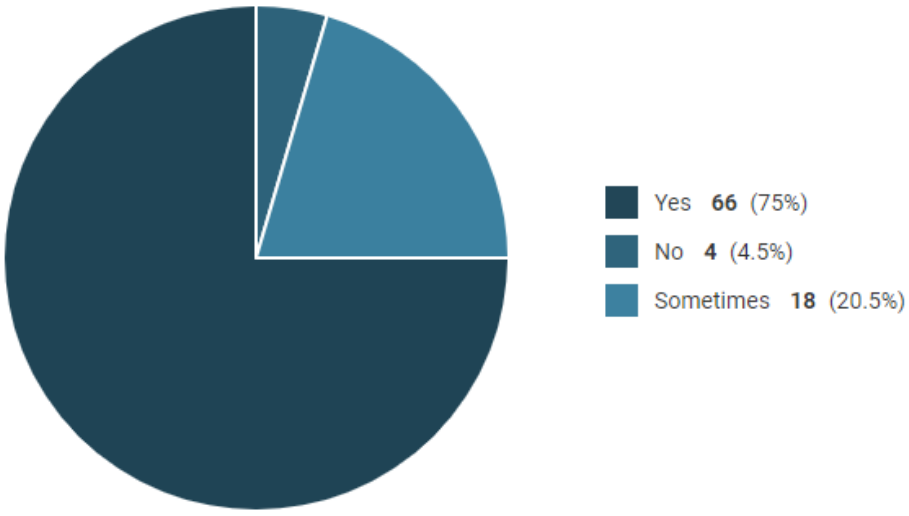

Figure 4.3: Question: Do you preserve the research data that you gather for your research?

Respondents were asked where they store their data by means of a multiple-choice question, for which more than one answer could be selected. 71 (84.5%) participants use physical storage devices, 33 (39.3%) participants use cloud-based storage, 7 (8.3%) participants use research data repositories and 11 (13.1%) participants chose “other”. The latter were asked to specify the storage means that they use; results showed that some researchers preserve paper-based copies of the raw data they collect. These may be in various forms including interview transcripts, questionnaires and lab reports. Such hard copies are kept in box files or under lock and key conditions. One of the respondents stated that primary data is kept in special dedicated servers until required and as per ethical guidelines.

Participants were also asked whether they document the data they gather for research. This question was only presented to participants who stated that they always, or sometimes, preserve their research data. Out of the 84 scholars who were eligible to answer this question, 83 responded. A total of 38 (45.8%) researchers said that they always document their data and 6 (7.2%) researchers said that they never do. 39 (47%) researchers stated that they sometimes document their research data.

The concept of data management was brought up during one of the interviews when one of the participants sustained that Maltese people still need to learn how to handle data effectively and unless they learn how to manage data they cannot recognize the value of sharing the same.

The results gathered in the questions related to data management show that a number of UM academics store data haphazardly, if they do at all. Documenting data and storing it in standardised formats does not seem to be the norm.

4.4.3. Researchers' concerns on data sharing

Recognising researchers' concerns on data sharing was pivotal in reaching the aims of this dissertation. By doing so, the UM can ensure that such worries are attended to when preparing the policy document. Overcoming such barriers will eventually increase the level of acceptance of the policy.

4.4.3.1. Sharing own data

Participants who stated that they never shared their primary data were asked to select reasons why they did not. More than one reason could be selected for this question. Ethical and legal concerns were the most common motives, being selected 19 (55.9%) times, followed by safeguarding the privacy of the research objects, with 15 (44.1%) respondents selecting this option. 12 (35.3%) participants believe that sharing their research data would lead to loss of intellectual property rights on the same. 9 (26.5%) participants feel that they need assistance or

training on the technical aspect of sharing raw data. 8 (23.5%) respondents believe that they would not be properly accredited. 6 (17.6%) researchers who took part in this study do not share their data because they believe that they do not gain anything by doing so. A total of 6 (17.6%) researchers also believe that research data is private to them and therefore they do not share their data. 3 (8.8%) survey participants do not share their primary research data because according to them the process is too time-consuming. 4 (11.8%) participants selected "other" for this question and their replies included that they never received any request to share their research data and that the faculty does not support data sharing.

The concerns and barriers mentioned by the interview participants included the following:

- As a population Maltese people are not used to sharing data, with one of the interviewees declaring "I don't think that we [Maltese people] have a sharing culture" and another interviewee stating that "some people are very territorial about their data";
- The type of data being collected may be concerning in its nature, for example one of the participants uses a lot of videos as her primary data;
- It is not always possible to acquire informed consent from vulnerable people;
- Misuse of data by third parties;
- It would not be possible to inform data subjects how exactly their data would be used by other parties once it is shared publicly;
- Data out of the original context may not make sense;
- The processes for sharing data and presenting it in an acceptable manner are time-consuming; one of the interviewees claimed that these practices are an "additional administrative burden on academics";
- Anonymising data is not always possible;
- Open data sharing may have a negative impact on participation, with individuals and companies being discouraged from participation as they would not want information about them to be accessible by the public;
- Loss of financial opportunities;
- Privacy, sensitivity, confidentiality.

4.4.3.2. Using shared data

Researchers who never used research data generated by others were also asked to select the reasons why they refrained from doing so. More than one answer could be selected for this question. 16 (39%) respondents prefer generating their own data, 3 (7.3%) respondents do not believe that research data available in open access is reliable, 22 (53.7%) respondents have not had the need to use data generated by others so far and 13 (31.7%) respondents claimed that the data they needed was not available in open access.

The questionnaire included an optional open-ended question, where participants could add feedback. The feedback given mainly mentioned concerns regarding open research data, with data protection, ethical issues, copyright and intellectual property rights being the topmost cited. Costs and time were also listed as concerns. Some of the participants referred to the nature of the data as being a factor that affects whether data should be in open access or not. Two respondents mentioned that presenting raw data out of its original context may be misleading. Another two respondents said that data should be shared out of free will and no policy enforcement should be placed in this respect. Other researchers mentioned the possible setbacks of premature sharing of data including loss of patent rights. Two researchers commented on the importance of a professionally structured policy, with one of them emphasising the importance of first enforcing an intellectual property rights policy and then implementing an ORD policy. One of the replies for the open-ended question mentioned the involvement of data scientists, stating that: “Data science and data analysis is becoming more complex. Researchers cannot be experts in everything and also HR is limited. It is important to invest in data scientists to support researchers.”

The main concerns surrounding ORD are related primarily to legal and ethical matters, and intellectual property rights. The additional work embedded in the process of data sharing is also an issue as some academics may struggle with the technicalities and find this too time-consuming.

4.4.4. The role of the library in ORD

Through this study, the researcher aimed to identify ways in which the UM can nurture a society of researchers who are open to data sharing. Subsequently, understanding the role of the library in ORD was a critical component of this dissertation.

4.4.4.1. Fostering a culture of research data sharing among UM academics

The library is hypothetically a strong influencer in fostering a culture of ORD among researchers working at the UM. A question with a trichotomous scale answer asked participants about their stand on data librarians; specifically, if they think that the UML should employ data librarians. The question had 54 positive replies (61.4%), 9 negative replies (10.2%) and 25 neutral replies (28.4%). Participants were also asked if they would use the service of a data librarian if this was available, for which 66 (75%) responded positively and 22 (25%) responded negatively. Another question sought to understand whether researchers believe that collaboration between librarians and researchers will improve the quality of research data management in an institution. 72 respondents (81.8%) chose “Yes”, while 16 respondents (18.2%) chose “No”.

During the interview, participants were asked about their expectations of the role of a data librarian. All 3 interviewees mentioned that they expect a data librarian to have technical expertise; this includes knowledge on how to upload data to a repository, formatting data and presenting it in a reusable manner, and indexing. The data librarian is expected to be the go-to person and a specialist for data management, with the ability of offering training, guidance and consultancy. A participant also mentioned that the data librarian should be conversant with creating data collection tools in a way that would then allow researchers to present them adequately.

4.4.4.2. Research data in Malta

Most of the researchers who took part in this study were, at the point of taking the questionnaire, unaware that Malta is currently working on establishing a nationwide policy on open science, which will include procedures on ORD. In fact, only 21 participants (23.9%) were aware of this. Researchers were also asked to rate their willingness to follow a policy that requires them to share their primary data in open access. The choices were presented as a five-point Likert scale from "very willing" to "very unwilling". The responses are presented in figure 4.4, below.

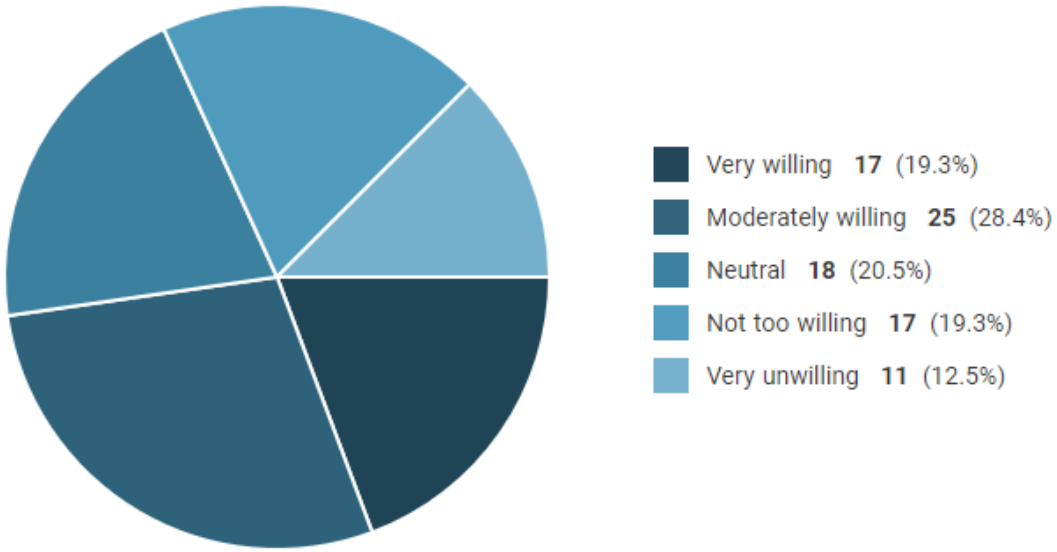

Figure 4.4: Researchers' willingness to adhere to a national policy on ORD

Through the interview, the researcher sought to identify reasons as to why Malta, as a country, has not kept abreast with other EU countries when it comes to the openness and sharing of research data. They were asked whether they think that researchers' resistance to share data is impeding the country from moving forward in this regard. Two of the respondents stated that in their opinion it is more lack of knowledge rather than conscious resistance to the concept. One of these respondents admitted that she had not heard of open research data prior to taking part in this study. The other participant said that there is a general lack of knowledge on the subject, as well as scepticism, especially from unions, who also showed resistance when open access to

research was introduced at the University. All three respondents pointed out that the population's culture plays a significant part in this, asserting that Maltese scholars do not have a sharing ethos. One of the respondents claimed that shifting the "mentality" of a society is very difficult, however, a culture change in this regard is necessary in order to move forward.

One of the participants claimed that Malta has lagged behind because "we haven't had a champion"; claiming that now that the University library has taken the lead in this, his prospects for scholars within the University to cooperate are high. The interviewee also hopes that this will inspire other institutions in the country, leading to a national cultural revolution.

4.4.4.3. Mandating ORD

All three interviewees agreed on the idea of a phasing-in period prior to mandating ORD, claiming that it is important to educate scholars and making sure they recognise the benefits for sharing their data in open access. Training, communication, and feedback were also mentioned during the interviews.

Prior to formulating policies, the University would need to decide on whether it would provide a repository to host the data, as well as statistical tools to monitor the usage of data; this was mentioned by two of the interviewees.

The feedback received through both the questionnaire and interviews showed that the concept of a data librarian as the 'go-to' person for any issues related to data management and data sharing would be well-received. The fact that there would be a person who academics can refer to if they find any difficulties seems to put their mind at rest and possibly encourage them to further contribute to the research community by sharing their data.

Chapter 5: Analysis and discussion

…research data achieve significance in sociology when viewed in relation to theoretical concerns.

- Alan Bryman

5.1. Introduction

This chapter aims to analyse the findings in light of the literature review and research questions posed at the beginning of the study. These are then discussed in relation to the themes brought out through the research. It also discusses the reliability and validity of the data collected. The desired outcome of the research was to discover ways to encourage Open Research Data (ORD) at the University of Malta (UM) and identify the role of the library in the process of data sharing. This was the starting point that led to the research questions and objectives, that set out to identify:

- the willingness of scholars to make their research freely available in open access, whilst using data generated by others that is also available in open access;
- concerns perceived by scholars regarding ORD;
- ways how the UM can nurture a culture of open data sharing among scholars.

A thematic approach was adopted to analyse the results; consequently, this chapter is organised in topics that tackle the different yet interrelated issues that come into play when introducing an institutional policy for ORD. The same themes used for the Results section (Chapter 4) are also used for this chapter as it is believed that this will increase consistency and encourage readers' apprehension. Nonetheless, the arguments were further evolved and linked together where deemed appropriate. A fundamental goal is that the policy is adhered to because the motives behind it are understood, which in turn achieves increased compliance.

5.2. Reliability of data collection

With a total population of 950 resident academics (all of whom received the questionnaire and were eligible to participate), a response rate of 9.26% and a confidence level of 95%, the margin of error for this study was of 10%. The initial response rate was as low as 4.74%, however an additional email reminding academics to participate in the survey was sent. This increased the representation of the total population, thus lowering the margin of error. Nonetheless a margin

of error of 10% is still high, hence the sample may not be representative of the whole population. The possible factors that may have contributed to the low response rate are discussed in section 4.2. In addition to the determinants outlined in 4.2, it has been noted that this response rate is common for other studies on related topics that have targeted the same cluster of participants. Saliba, K. (2020) and Borg, J.J. (2018) have both conducted research about open access amongst University of Malta academics, using the same distribution medium that was used in this study; the former achieved a response rate of 8.5% with a margin of error of 11% whilst the latter achieved a response rate of 10.6% with a margin of error of 9%; both margins of error were calculated at a confidence level of 95%.

Research was also carried out through interviews with the aim of further discussing certain issues at a deeper level, with researchers of different backgrounds. Three interviews were carried out and the replies were used to further substantiate the questionnaire's findings. Carrying out more interviews proved to be hard given the outburst of a global pandemic, however the data gathered was enough for the purpose of this research. Additionally, the aim set out for the interviews was reached. In order to safeguard the privacy of the interviewees, and for ease of reference, they were named interviewee 1, 2 and 3 respectively.

5.3. Themes

5.3.1. Researchers' opinions and concerns regarding ORD

The amenability of researchers to share their data is closely linked to the concerns perceived by the same regarding ORD. Several preoccupations feature when sharing own data and also when using data generated by others, which issues need to be dealt with if universities want to move towards notions of open access and ORD. Hence, the research questions are interrelated by nature as one links to another in scope.

The research carried out in this study showed that the notion of data sharing is still at its inception phase in Malta. This could be related to the fact that there are currently no country mandates nor institutional policies that encourage research data sharing habits. Even though the percentage of academics that had shared their data was higher than the ones who had never shared their data, there was only a small number of people who shared such data in open access.

Total number of respondents	Have never shared data	Have shared data	Have shared data in open access
88 (100%)	34 (38.6%)	54 (61.4%)	11 (12.5%)

Figure 5.1: Participants who have shared data in open access

The virtues of ORD may increase researchers' motivation levels to honour policies regarding sharing data in open access; especially when such virtues are outlined by their peers. Researchers who took part in this study shared their raw data for various extrinsic, altruistic and also intrinsic reasons. Some researchers simply shared data because it was required by the publisher or funding body, others shared data because they believed that they would be contributing to the research field, and others shared data because they believed that by sharing the raw data they would be increasing the credibility and visibility of their research. The latter is in line with a study carried out by Piwowar and Vision (2013) that concludes that the availability of open datasets in accompaniment with research output is a strong variable that favours citation rates. Some Maltese academics may be unaware of such intrinsic benefits; in fact a particular interviewee specifically asked the interviewer "really and truly what's the point of making it available for everybody?"

It has been noted that the scholars' background and area of research affects the extent to which scholars are willing to share their data and reuse data shared by others. Even though the hypothesis could be tested further, it is safe to say that this research has evidenced that the FICS (Faculty / Institute / Centre / School) position may also have an impact on the academics' views on data sharing. The majority of the questionnaire respondents stated that the FICS's position regarding ORD was neutral, whilst an equal amount of questionnaire respondents answered this question with a 'yes' or 'no'.

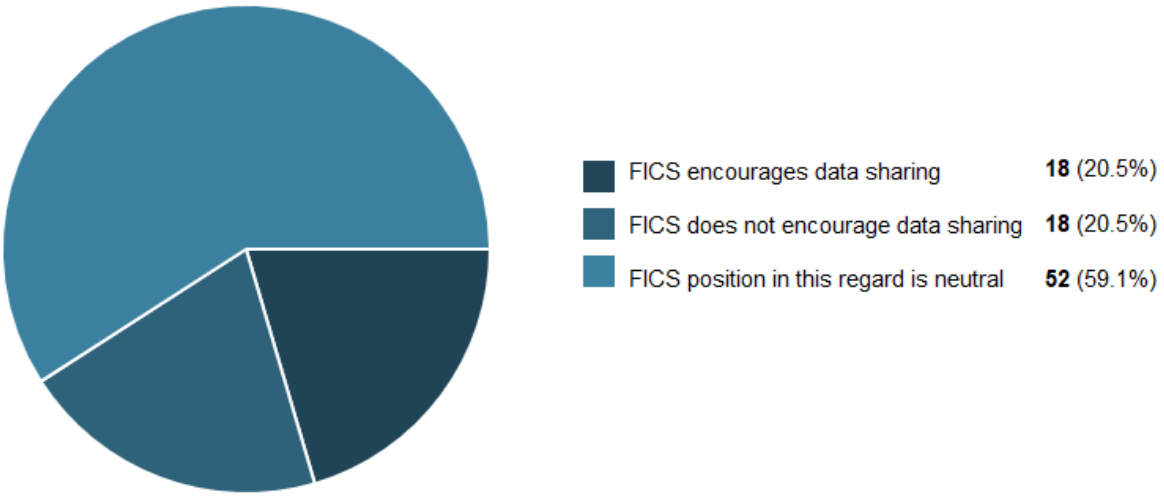

Figure 5.2: FICS position on data sharing

The areas of study and the FICS's position on data sharing as determinants of researchers' views on ORD were also brought up during the interviews. Interviewee 1, who works in education, claimed that an issue that she finds with ORD might be related to the type of data that is being gathered. She mentioned that the type of data that she collects sometimes consists of videos of children of a non-consenting age; adding that even if she sought parents' approval, the fact that such data will be in open access may discourage participation. The same participant emphasised her concern about people being reluctant to participate albeit being consenting adults, stating that even she would not be eager to take part in a study where data is free-for-all with possibilities for people using it for motives, which differ from the original study. She also said that sometimes people are easily identifiable, especially in a small country like Malta, where it is hard to completely anonymise the participants.

Interviewee 2, who has a background in Information and Communications Technology (ICT) stated that in his area of research it is expected that codes and traces, which are his raw data, are made available for reviewers via a repository. He also said that the faculty culture is one that encourages sharing and collaboration. This has shown contrasting views when compared to interviewee 1. Nonetheless, interviewee 2 voiced his concerns on private companies who might not be willing to share certain information openly; this is because the data might end up being

used for commercial purposes by their competitors. He added that removing certain components of the raw data may solve the problem.

As a social scientist, interviewee 3 claimed that the data he collects concerns people and therefore some of it may be private or sensitive. Publicising such data may require extra attention from him as a researcher to make sure that the data is then accessed and used as per established ethical protocols. He adds that "the responsibility for respecting data and what is behind it isn't just with the data producer but also with the data consumer."

Although other factors may affect researchers' views on ORD, this study has shown that the type of data being collected as well as the scholarly background of the researcher are strong influencers on the researchers' disposition to collaborate with policy makers when it comes to open data sharing. This has also been confirmed in the literature as Zuiderwijk and Spiers (2018) state that "the prevalence of data sharing and re-use differs between disciplines [...and] although the unique challenges and characteristics of each discipline need to be carefully considered [...] disciplines with low sharing rates may learn from disciplines where sharing is common" (p. 228).

5.3.2. Research data management: a culture change

> "Digital data preservation should be a key aspect of all research projects. Some research data are unique and cannot be replaced if destroyed or lost, yet only by referring to verifiable data can your research be judged as sound."
>
> (Digital Curation Centre, 2020)

Preservation goes hand in hand with documentation, which involves the organisation of data in a way that makes it discoverable by the researcher and also by others, especially if the primary data will be published with the intention of being used by others. This entails what interviewee 3 referred to as "an additional administrative burden." In order for researchers to be able to share their data in open access, they would need to organise and document their data, create metadata, ensure that the data is presented in an appropriate format that guarantees

compatibility and longevity of the datasets, whilst also ensuring that the necessary backups are detained. Subsequently, data management is at the heart of data sharing and ORD. In fact, the desktop research held in the initial stages of this research has shown that rather than focusing solely on ORD, the two UK Universities that have been consulted for the purposes of this study have a data management policy, which includes procedures for ORD.

The results obtained in the questionnaire showed that amongst the cluster of researchers that took part in this study, not all researchers preserve and document their data. Although the majority of researchers do preserve the raw data, the storage medium used is sporadic rather than consistent throughout the UM. The majority of researchers, in fact, store data on a physical storage device, with only 7 (8.3%) out of the total 88 respondents stating that they store data on research data repositories. Additionally, not all researchers document their data, with 5 respondents not even answering this question. The reason behind this inconsistency may be related to a national unawareness about the concept of data sharing, which became more evident during the interviews.

> "I don't think there's any resistance as such, it's more lack of knowledge, it's more ignorance, it's more, perhaps, some scepticism. It's lack of information. This is a cultural revolution that has yet to happen."
>
> (Interviewee 3)

All three interviewees referred to the Maltese people's mentality and culture at some point in the interview, claiming that this could be a major factor affecting scholars' standpoint on ORD. Interviewee 1 believes that Maltese researchers do not have a sharing culture and agreed with the idea of adopting the European Commission's recommendation for Malta to employ a "phase-in" approach, stating that this would help researchers to "get used to the idea" whilst helping policy-makers to identify the "teething problems". Interviewee 2 emphasised the need for increased education about data sharing and open access, stating that older generations may find it harder to conform to such ideologies. The idea of showing scholars the benefits of ORD and

educating them instead of obliging them was mentioned by all three interviewees, with interviewee 3 suggesting the adoption of “the persuasive route rather than the coercive one.”

Interviewee 2 associates the lack of awareness with the cultural background of the Maltese society, a society which was described by interviewee 3 as “territorial about their data”. The former claimed that “before even considering open access data, [...Maltese scholars need to...] learn how to deal with data.” The next section will further discuss education and the role of the library in data sharing and open research data.

5.3.3. The role of the library in successful ORD

As previously discussed in this chapter, having scholars collaborate in the ORD endeavour is the ideal way forward as opposed to pressuring them to follow policies blindly. One of the open-ended questions of the questionnaire asked for additional feedback of the participants and a respondent replied “I share data if I want. If I am forced to, I will do everything not to share it.” Notwithstanding that not all scholars share the same view as this particular researcher, the feedback received through both the questionnaire and the interviews has shown that education and awareness are crucial for having researchers on board. The readiness of academics employed at the UM to share their data is displayed as figure 5.3, overleaf.

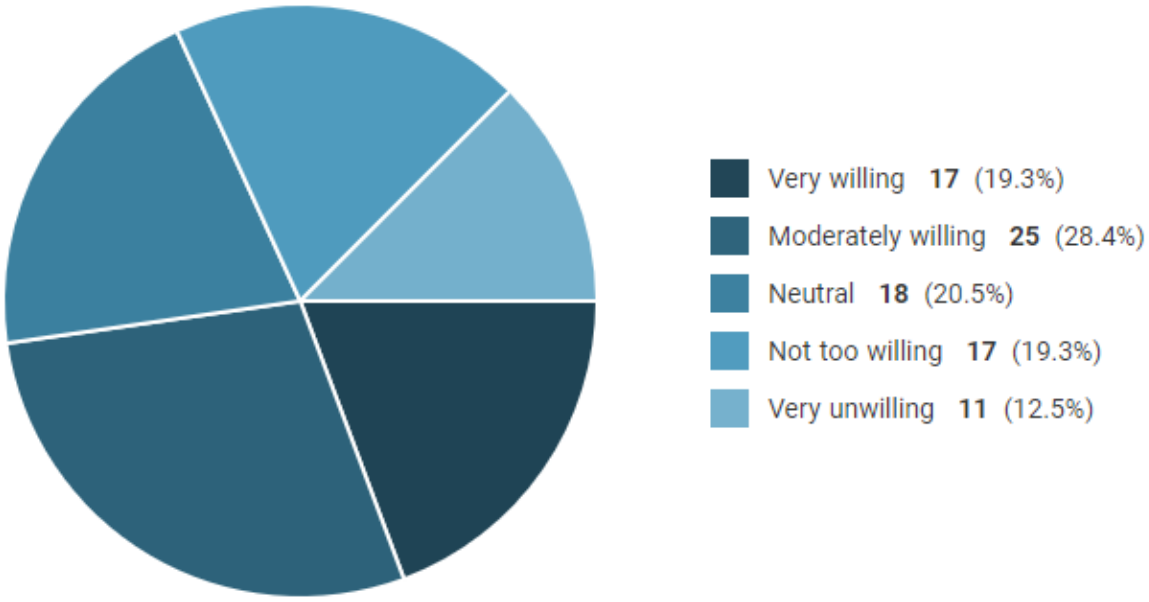

Figure 5.3: Willingness of full-time resident academics at the UM to follow an open data policy.

The administrative work that needs to be done in order to share raw data seems to be an impediment for researchers who agree with the general idea of ORD but have never actually shared their data, because doing so is too time-consuming, or because they are unaware of the benefits they might reap out of sharing their own data. The majority of the participants who engaged in this study welcomed the idea of having a data librarian, with 72 participants (81.8% of the total population) believing that collaboration between librarians and researchers will enhance research data management. Through the interviews, the researcher sought to identify the responsibilities that academics expect librarians to undertake. All three interviewees highlighted that the data librarian would be the go to person for any queries regarding data sharing and data management, with two of the interviewees emphasizing the fact that the librarian should have technical expertise. Interviewee 1 also mentioned that it would be a good idea to have a dedicated website maintained by the library, consisting of frequently asked questions (FAQs). In fact, the University of Cambridge has a very insightful website dedicated exclusively to research data; covering information about data management, the university's data repository, data policies, events such as conferences related to research data, and FAQs[1]. As the

[1] https://www.data.cam.ac.uk/

entity steering the open data project in Malta, the UM library is therefore expected to provide training and assist researchers in technicalities such as indexing of data, saving and presenting it in the right format and so forth.

The type of training that should be organised for academics was also discussed during the interviews. One of the salient points that emerged out of the interviews was awareness, as a lot of scholars are not cognizant of what open data entails and this was also proven through certain answers obtained in the questionnaire. For instance, it was clear that some participants were oblivious to the fact that the primary data could be shared after publication of the research output, thinking that by sharing raw data prematurely they are risking theft of their intellectual property, with third parties publishing ahead of them. It was also clear that some concerns could be eliminated if the researchers knew that they could embargo the research data for a specific period of time. Having academics who have previously shared their data in open access divulge their experience with their colleagues is also beneficial in raising awareness about the advantages of ORD. This lack of awareness about data sharing in the Maltese research population was also evident in a report by the European Commission, which states that "there is significant lack of awareness on the topic, with researchers expressing a keen interest in using Open Data, but often not showing a corresponding willingness to share their own data" (European Commission, 2020, p.41).

Training on the legal and ethical aspects of ORD is vital in order to ensure that data will be shared within lawful and mindful confinements. Another important point mentioned during the interviews links to section 5.3 of this dissertation, where the hypothesis that discipline affects the level of sharing was discussed; two of the three interviewees recommended that discipline-specific or data type-specific training should be organised, with one of the interviewees stating that it is important for instructors to tailor the training according to the audience's research area. Providing researchers who deposit research data with statistical analysis tools to help them understand how their data is being used may also be an incentive.

Librarians are therefore enablers of ORD and key players in helping researchers present their data according to FAIR principles. As Ohaji, Chawner and Yoong (2019) put it, the data librarian

role is a position that entails responsibility for the research data management (RDM) of the institution within which they are employed. FOSTER (n.d.) identify several responsibilities that should be encompassed in the role of a data librarian, including cultivating awareness, governing and supporting the infrastructures, developing RDM policies, and supporting and instructing researchers.

5.4. Limitations of the study

The sample size used for the questionnaire is not necessarily representative of the whole academic population of the UM. Although there was no bias from the researcher's end, since the questionnaire was sent to all full-time resident academics, it is possible that the ones who actually answered the questionnaire were scholars who are either in favour of ORD, or are at least curious about the topic. Scholars who are strongly against the concept might have also been interested in participating in order to voice their viewpoint. It would be interesting to know the position on ORD of the 862 academics who did not take part in the research. Some of them may have simply ignored the email, yet a cluster of these non-participants may be the ones who will not even be interested in attending training on ORD, simply because they don't have the time or aptitude for it. Although the field within which academics perform research was identified as a strong determinant of the researchers' disposition for data sharing and ORD, the sample size used for the interviews was quite small and other factors could have affected the disparity of opinions. For instance, age is another factor that could have been tested; the correlation between this demographic and the disposition towards ORD could have been an important component in this study. An additional limitation of this dissertation was the period when the research took place. As discussed in section 4.2, this research was conducted during the outbreak of a world pandemic; implying that the subjects were focused on engaging in alternative ways to deliver lectures given that the University had closed down its doors. This meant that academics who were not very technologically-savvy were going through a learning curve in their career. The possibility of these academics finding time to answer a questionnaire was minimal. The

shortcomings outlined above could have hindered the study in one way or another by lowering the reliability of the results attained.

The word limit of this thesis proved to be a challenge and a hindrance. The challenge was to keep to the point and deliver a large amount of information in a small amount of words; the hindrance was that certain ideas had to be discarded so as to keep within the word limit. For instance, it would have been interesting to hold focus groups in addition to interviews. I would have organised focus groups with people from different disciplines in order to study the interaction amongst them and examine the effect that one academic's opinion may have on another. If the constraints outlined in section 3.3 had to be overcome, focus groups could also be organised as a follow up to this research. A different perspective to this study would have been to gather viewpoints from Master's and Doctoral students for instance, rather than just from academics. The next chapter concludes the study and findings, whilst outlining recommendations for how the University of Malta should go about implementing an open data policy. The concluding chapter also suggests further research in the area.

Chapter 6: Conclusion and Recommendations

We are dealing with academics; they are a very strange bunch. My suggestion is always to adopt the persuasive route rather than the coercive one.

- Interviewee 3 (personal communication, 2020)

This research aimed to identify ways in which academics working at the University of Malta (UM) could get on board with the implementation and execution of a policy that favours open research data (ORD). This was done by establishing the reasons why a lot of researchers in Malta do not have a sharing culture. A general unfamiliarity with the concept of data sharing, together with the fact that Malta was one of the last EU countries to have a national open data portal (Times of Malta, 2016) and the reality that Malta is still in the process of having a national mandate on ORD, contributed to a nation of researchers who are reluctant to share their intellectual property with the general public. Lack of technical know-how, including familiarity with the technological interface, knowledge of the steps needed for good documentation of data, cognisance of the ideal formats, storage, and longevity of the data itself have also been identified as barriers to ORD.

As attested through the qualitative and quantitative research methods employed for this study, it can be argued that the majority of academics who are not willing to share their raw data in open access are basing their opinions on misinformed beliefs. Lack of education and awareness have been identified as primary factors contributing to a non-sharing culture in the Maltese population. The study also showed that presently, scholars are not incentivised enough to share their data in open access; this was evident as several academics said that documenting and sharing their research data is too time-consuming. This process is seen by some as an 'administrative burden' rather than a measure for increasing the reliability and credibility of their research output. It was also concluded that the research area of the academics impacts the disposition of the same to disclose their raw data, primarily due to subject sensitivity.

Universities practicing ORD have adopted the concept of research data sharing being 'as open as possible, as closed as necessary' and are requiring researchers to submit a data management plan (DMP) at the research proposal phase. The DMP would ensure that data acquired during the course of research is handled appropriately throughout its lifecycle, whilst correct analysis and storage protocols are practiced. The DMP would also include a plan for how and which data will be shared. The researcher believes that implementing this approach at the UM would be

beneficial. This is because one cannot see ORD in isolation; rather, it must be part of a process towards better management of data.

Fostering a culture of data sharing in Malta may be a challenge that will change librarians' roles. Teaching researchers about open data, data management, documentation, digital preservation, indexing, creation of metadata and everything in between, is not a one-size-fits-all programme. Librarians should acquire the skills enabling them to provide subject-specific or data-type specific training, whilst providing professional assistance with the technicalities of the process, and also being able to offer legal advice about which data could be shared, how to anonymise data, and so forth. In the Maltese scenario, the University could appoint a team of people with the responsibility of assisting lecturing staff in data management. Ideally the team is made up of librarians, as experts in data management, and members of the legal office, who could offer guidance on issues regarding data privacy and sensitivity as per Maltese laws and regulations.

In conclusion, this study recommends that once the national policy for ORD in Malta is available, procedures for the UM are drawn up. Such procedures should be phased-in gradually and library representatives should organise meetings with each faculty with the objective of raising awareness, educating the academics, and attending to their concerns. Prior to organising such meetings, a study needs to be undertaken to identify the type of data produced by each faculty. Follow-up meetings should be organised so that scholars who have shared raw data in open access may share their experience, making this a learning curve for both librarians and other scholars who wish to follow suit. The researcher recommends that a DMP is submitted with every research proposal that requires funding from the UM. This will cultivate a mentality where irrespective of whether the data is being shared or not, it is well documented and stored for future reference. It will eventually become part of the research process and will not be regarded as an added burden. Once this is mastered, sharing will become a simple extra step that does not require a lot of additional effort. Finally, as outlined by the participants who took part in this research, academics need to know and believe that sharing primary data in open access will be beneficial for them and their research, as this would be the best intrinsic motivator that drives them to adhere to the set policies and procedures; the key is 'education over coercion'.

This study could be further developed by including the opinion of researchers from other higher institutions in Malta, such as MCAST (Malta College of Arts, Science and Technology) and the American University of Malta. A comparative study would be interesting, especially in the near future when Malta introduces a national mandate on ORD. It would be interesting to find out what policies and procedures have been implemented by each institution, what worked and what did not, and how such policies could be enhanced to further encourage a culture of open data in the Maltese islands.

Bibliography

Astell, M., 2017. *Ask not what you can do for open data; ask what open data can do for you.* [Online]
Available at: http://blogs.nature.com/naturejobs/2017/06/19/ask-not-what-you-can-do-for-open-data-ask-what-open-data-can-do-for-you/
[Accessed 19 September 2019].

Borg, J.J. (2018). *The academics' perceptions of open access: a case study at the University of Malta. Dip.LIS Dissertation.* University of Malta. Available at: https://www.um.edu.mt/library/oar/handle/123456789/39541 (Accessed: 10 June 2020).

Brochu, L. & Burns, J., 2019. Librarians and research data management - a literature review: Commentary from a senior professional and a new professional librarian. *New Review of Academic Librarianship,* 25(1), pp. 49-58.

Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done?. *Qualitative Research,* 6(1), pp. 97-113.

Bryman, A., 2016. Social research methods. 5th ed. Oxford: Oxford University Press.

Campbell, P., 2016. Data sharing. *Nature Magazine*, pp. 435-436. [Online]
Available at: https://www.nature.com/news/the-ups-and-downs-of-data-sharing-in-science-1.20124
[Accessed 31 January 2020]

Coughlan, S., 2018. The challenges facing universities in 2018. [Online]
Available at: https://www.bbc.com/news/education-42413636
[Accessed 20 September 2019].

Cox, A. & Pinfield, S., 2014. Research data management and libraries: current activities and future priorities. *Journal of Librarianship and Information Science,* 46(4), pp. 299-316.

Daniel, H., 2011. *Benefits of mixed methods research.* [Online]
Available at: http://benefitof.net/benefits-of-mixed-methods-research/#:~:text=Benefits%20Of%20Mixed%20Methods%20Research%201%20More%20comprehensive.,time%20and%20cost%20savings%20benefits.%20More%20items...%20
[Accessed 06 September 2020].

Data Science Degree Programs Guide, 2020. *What is a data science generalist?.* [Online]
Available at: https://www.datasciencedegreeprograms.net/faq/what-is-data-generalist/
[Accessed 27 January 2020].

Digital Curation Centre (DCC), 2020. *What is digital curation?.* [Online]
Available at: https://www.dcc.ac.uk/about/digital-curation
[Accessed 24 August 2020].

Elsayed, A. & Saleh, E., 2018. Research data management and sharing among researchers in Arab universities: An exploratory study. *IFLA Journal,* 44(4), pp. 281-299.

Else, H., 2019. *What Boris Johnson's leadership could mean for science.* [Online]
Available at: https://www.nature.com/articles/d41586-019-02279-z
[Accessed 14 December 2019].

Emory University, 2020. *Benefits of open data to researchers.* [Online]
Available at: https://sco.library.emory.edu/research-data-management/open-data/benefits-research.html
[Accessed 4 February 2020].

Enago Academy, 2018. *How open science is helping researchers succeed.* [Online]
Available at: https://www.enago.com/academy/how-open-science-is-helping-researchers-succeed/
[Accessed 25 January 2020].

European Commission, 2016. *Open innovation, open science, open to the world - a vision for Europe.* [Online]
Available at: https://publications.europa.eu/resource/cellar/3213b335-1cbc-11e6-ba9a-01aa75ed71a1.0001.02/DOC_2
[Accessed 29 January 2020].

European Commission, 2018. *EACEA national policies platform: Eurydice.* [Online]
Available at: https://eacea.ec.europa.eu/national-policies/eurydice/malta/types-higher-education-institutions_en
[Accessed 31 January 2020].

European Commission, 2020. *Specific support to Malta: "Open Access: an opportunity for Malta": Recommendations for the Development of a National Policy for Open Access to publications, research data and related issues,* Luxembourg: Publications Office of the European Union.

European Commission, n.d. *Facts and figures for open research data.* [Online]
Available at: https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/facts-and-figures-open-research-data_en
[Accessed 11 November 2019].

European Data Portal, 2020. *Open Data maturity in Europe.* [Online]
Available at:
https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n3_2017.pdf
[Accessed 14 May 2020].

Ferguson, L., 2014. *How and why researchers share data (and why they don't).* [Online]
Available at: https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont
[Accessed 29 July 2020].

FOSTER, 2017. *What is open science? Introduction.* [Online]
Available at: https://www.fosteropenscience.eu/node/1420
[Accessed 29 January 2020].

FOSTER, n.d. *Libraries: roles and opportunities on Open Science.* [Online]
Available at:
https://www.fosteropenscience.eu/node/1426#:~:text=Librarians%20can%20offer%20guidance%2C%20training,analysis%3B%20outputs%20sharing%20and%20publication%3B
[Accessed 28 August 2020].

Fowler, F., 2002. *Survey research methods.* 3rd ed. London: Sage Publications Ltd..

Fuller, M., n.d. *Quotespeak.* [Online]
Available at: https://www.quotespeak.com/professional-quotes/business-quotes/best-sharing-knowledge-quotes-inspirational-motivational/2/
[Accessed 27 January 2020].

Gell, T., 2019. *Desk research: What it is and how you can use it.* [Online]
Available at: https://www.driveresearch.com/market-research-company-blog/desk-research-what-it-is-and-how-you-can-use-it/
[Accessed 19 February 2020].

Government of Malta, 2020. *Malta data portal (beta version).* [Online]
Available at: https://open.data.gov.mt/
[Accessed 20 July 2020].

Hahnel, M., 2013. *Open data: we need to share research results, even when they are wrong.* [Online]
Available at: https://www.theguardian.com/higher-education-network/blog/2013/oct/08/open-research-data-science-papers
[Accessed 19 September 2019].

Hayes, A., 2020. *Stratified random sampling.* [Online]
Available at: https://www.investopedia.com/terms/stratified_random_sampling.asp
[Accessed 9 April 2020].

Holstein, J. & Gubrium, J., 1995. *The active interview*. London: SAGE Publications, Inc.

Houtkoop, B. et al., 2018. Data sharing in psychology: A survey on barriers and preconditions. *Advances in methods and practices in psychological science*, 1(1), pp. 70-85.

IGI Global, 2020. *What is data sharing?* [Online]
Available at: https://www.igi-global.com/dictionary/bioinformatics-clouds-for-high-throughput-technologies/6815
[Accessed 29 January 2020].

Jick, T., 2006. Mixing qualitative and quantitative methods: triangulation in action. *Administrative Science Quarterly,* 24(4), pp. 602-611.

Jisc, 2020. *Online surveys (formerly BOS): Powerful, flexible online surveys.* [Online]
Available at: https://www.onlinesurveys.ac.uk/
[Accessed 3 January 2020].

Kalkman, S. et al., 2019. Responsible data sharing in international health research: A systematic review of principles and norms. *BMC Med Ethics*, 20(1).

Kitchin, R., 2013. *Four critiques of open data initiatives.* [Online]
Available at: http://progcity.maynoothuniversity.ie/2013/11/four-critiques-of-open-data-initiatives/
[Accessed 12 February 2020].

Kitchin, R., 2014. *The data revolution: big data, open data, data infrastructures and their consequences.* California: Sage Publications.

Kroes, N., 2012. *Opening science through e-infrastructures.* [Online]
Available at: europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/12/258
[Accessed 11 October 2019].

Lamb, D., 2019. *Better science through better data 2019 conference collection.* [Online]
Available at: https://researchdata.springernature.com/channels/1411-events/posts/56507-now-live-conference-collection-of-better-science-through-better-data-2019
[Accessed 27 January 2020].

Loughborough University, 2020. *Data Management.* [Online]
Available at: https://www.lboro.ac.uk/services/it/research/datamanagement/rdm/
[Accessed 19 August 2020].

Loughborough University, n.d. *Research data management policy.* [Online]
Available at:
https://www.lboro.ac.uk/media/wwwlboroacuk/content/library/downloads/researchsupport/2016_09_21_ResearchDataManagementPolicy.pdf
[Accessed 19 August 2020].

Lucey, T., 2002. *Quantitative techniques.* 6th ed. Hampshire: Cengage Learning EMEA.

Lupton, E., 2019. *Looking to the future of data sharing.* [Online]
Available at: https://researchdata.springernature.com/users/208739-emily-lupton/posts/43809-looking-to-the-future-of-data-sharing
[Accessed 27 January 2020].

Malta Ministry for Education and Employment, 2014. *National research and innovation strategy 2020,* Kalkara: The Malta Council for Science & Technology (MCST).

Maltatoday, 2020. *Malta takes another step towards a national open access policy.* [Online]
Available at:
https://www.maltatoday.com.mt/sponsored/sponsored/101506/malta_takes_another_step_towards_a_national_open_access_policy#.XsKGyWgzbIV
[Accessed 10 May 2020].

McKim, C., 2017. The value of mixed methods. *Journal of Mixed Methods Research,* 11(2), pp. 202-222.

Medelyan, A., 2019. *Coding qualitative data: how to code qualitative research.* [Online]
Available at: https://getthematic.com/insights/coding-qualitative-data/
[Accessed 9 May 2020].

Mims Adrian, A., 2018. *Challenges in creating open data policies for universities.* [Online]
Available at: https://www.openaccessgovernment.org/challenges-in-creating-open-data-policies-for-universities/54922/
[Accessed 12 February 2020].

Mims Adrian, A., Emison, B., Musker, R. & Hopkinson, J., 2018. *Open access & open data at PUSH universities.* [Online]
Available at: http://wp.auburn.edu/push/wp-content/uploads/2018/06/GODAN-report.pdf
[Accessed 12 February 2020].

Mortensen, D., 2020. *How to do a thematic analysis of user interviews.* [Online]
Available at: https://www.interaction-design.org/literature/article/how-to-do-a-thematic-analysis-of-user-interviews
[Accessed 18 April 2020].

National Academy of Sciences (US), National Academy of Engineering (US) and Institute of Medicine (US) Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age., 2009. Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age. Washington (DC): National Academies Press (US); 2009. 3, Ensuring Access to Research Data. Available at: https://www.ncbi.nlm.nih.gov/books/NBK215271/
[Accessed 12 February 2020].

Needham, P. & Lambert, J., 2019. Institutional repositories and the item and research data metrics landscape. *Insights,* 32(1), p. 26.

Newton, I. & Hooke, R., 1675. *Isaac Newton Letter to Robert Hooke.* [Online]
Available at: https://discover.hsp.org/Record/dc-9792/Description#tabnav
[Accessed 8 November 2019].

Nielsen, M., 2009. Doing science in the open. *Physics World,* 22(5), p. 30.

Ohaji, I., Chawner, B. & Yoong, P., 2019. *The role of a data librarian in academic and research libraries.* [Online]
Available at: http://informationr.net/ir/24-4/paper844.html
[Accessed 17 March 2020].

OpenAIRE, n.d. *National open access desks.* [Online]
Available at: https://www.openaire.eu/contact-noads
[Accessed 31 January 2020].

OpenAIRE, 2020. *Malta: national open access desk.* [Online]
Available at: https://www.openaire.eu/os-malta
[Accessed 30 June 2020].

The Open Research Data Task Force, 2018. *Realising the potential: final report.* [Online]
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/775006/Realising-the-potential-ORDTF-July-2018.pdf
[Accessed 11 October 2019].

OECD, 2015. Open government data. [Online]
Available at: oecd.org/gov/digital-government/open-government-data.htm
[Accessed 15 March 2020].

Orion, n.d. What is open science?. [Online]
Available at: https://www.orion-openscience.eu/resources/open-science
[Accessed 28 January 2020].

Orszag, P., 2009. *Memorandum for the heads of executive departments and agencies.* [Online]
Available at:
https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf
[Accessed 27 January 2020].

Pampel, H. & Dallmeier-Tiessen, S., 2019. *Open research data: from vision to practice.* [Online]
Available at: http://book.openingscience.org.s3-website-eu-west-1.amazonaws.com/vision/open_research_data.html
[Accessed 11 October 2019].

Pew Research Center, 2013. *Library services in the digital age: the present and future of libraries.* [Online]
Available at: https://www.pewresearch.org/internet/2013/01/22/part-5-the-present-and-future-of-libraries/
[Accessed 1 February 2020].

Pickard, A., 2007. *Research methods in information.* 1st ed. Cornwall: Facet Publishing.

Piwowar, H., Day, R. & Fridsma, D., 2007. Sharing detailed research data is associated with increased citation rate. *Plos One,* 2(3), p. 308.

Piwowar, H. & Vision, T., 2013. Data reuse and the open data citation advantage. *PeerJ*, Volume 1, p. 175.

Popkin, G., 2019. *Data sharing and how it can benefit your scientific career.* [Online]
Available at: https://www.nature.com/articles/d41586-019-01506-x
[Accessed 20 September 2019].

Powell, R. & Connaway, L., 2004. *Basic research methods for librarians.* 4th ed. Westport: Libraries Unlimited.

Pryor, G., 2012. *Managing research data.* London: Facet.

Puniewska, M., 2014. *The Atlantic: Scientists have a sharing problem.* [Online]
Available at: https://www.theatlantic.com/health/archive/2014/12/scientists-have-a-sharing-problem/383061/#:~:text=One%20could%20pin%20the%20problem,is%20the%20lack%20of%2

0funding.
[Accessed 23 July 2020].

Research Excellence Framework, 2021, n.d. *What is the REF?.* [Online]
Available at: https://www.ref.ac.uk/about/what-is-the-ref/
[Accessed 14 December 2019].

Saliba, K. (2020). *Academics' perspective of open access and institutional repositories, University of Malta: A case study*. M.A. Thesis. Aberystwyth University.

Severin, A., Egger, M., Eve, M. & Hürlimann, D., 2020. Discipline-specific open access publishing practices and barriers to change: an evidence-based review. *F1000Research*, 7(1925).

Silverman, D., 2006. *Interpreting qualitative data.* 3rd ed. London: SAGE Publications Ltd.

Sincero, S., 2012. *Pilot survey.* [Online]
Available at: https://explorable.com/pilot-survey
[Accessed 9 May 2020].

Slonina, K. (2020). *Some queries*. [email]

Squires, M., 2020. *Watermelon research.* [Online]
Available at: https://www.watermelonresearch.com/blog/post/automated-vs-manual-coding-whats-right-for-your-voc-programme
[Accessed 12 May 2020].

Stern, N., Spichtinger, D., Manola, N. & van der Vooren, R., 2020. *Open Access: an opportunity for Malta: recommendations for the development of a national policy for open access to publications, research data and related issues.* [Online]
Available at: https://mcst.gov.mt/wp-content/uploads/2020/04/OPEN-ACCESS-MALTA-REPORT.pdf
[Accessed 18 May 2020].

Suber, P., 2012. *Open access.* Cambridge: MIT Press Ltd.

Swan, A. & Brown, S., 2018. *The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs (Report to the JISC),* Truro, UK: Key Perspectives Ltd.

Times of Malta, 2016. *Malta among EU's worst when it comes to Open Data, report finds.* [Online]
Available at: https://timesofmalta.com/articles/view/malta-among-eus-worst-when-it-comes-to-open-data-report-finds.627967
[Accessed 12 May 2020].

Travis, D., 2016. *Desk research: the what, why and how.* [Online]
Available at: https://www.userfocus.co.uk/articles/desk-research-the-what-why-and-how.html
[Accessed 19 February 2020].

Ünal, Y., Chowdhury, G., Kurbanoğlu, K., Boustany J. & Walton G., 2019. Research data management and data sharing behaviour of university researchers. *Proceedings of ISIC, The*

Information Behaviour Conference, Krakow, Poland, 9-11 October: Part 2. Information Research, 24(1).

UNESCO, 2017. *Concept of digital preservation.* [Online]
Available at: https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-preservation
[Accessed 3 February 2020].

Universities UK, n.d. *The open research data task force.* [Online]
Available at: https://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Pages/open-research-data-task-force.aspx
[Accessed 11 December 2019].

University of Cambridge, 2020. *Research data.* [Online]
Available at: https://www.data.cam.ac.uk/
[Accessed 16 August 2020].

University of Cambridge, 2018. *Good research practice guidelines.* [Online]
Available at: https://www.research-integrity.admin.cam.ac.uk/files/good_research_practice_guidelines_2018.pdf
[Accessed 21 August 2020].

University of Cambridge, 2019. *Open research position statement.* [Online]
Available at: osc.cam.ac.uk/open-research/open-research-position-statement
[Accessed 28 March 2020].

University of Cambridge, 2019. *Research data management policy framework.* [Online]
Available at: https://www.data.cam.ac.uk/university-policy
[Accessed 28 March 2020].

University of Malta, 2020. *L-Universita' ta' Malta.* [Online]
Available at: https://www.um.edu.mt/
[Accessed 1 December 2020].

UK Research and Innovation (UKRI), n.d. *RCUK policy on open access and supporting guidance.* [Online]
Available at: https://www.ukri.org/files/legacy/documents/rcukopenaccesspolicy-pdf/
[Accessed 14 December 2019].

UK Research and Innovation (UKRI), n.d. *UKRI Open Access Principles and High Level Policy.* [Online]
Available at: https://www.ukri.org/files/legacy/oadocs/ukri-open-access-principles-and-high-level-policy-pdf/
[Accessed 14 December 2019].

Van Rijnsoever, F., 2017. (I can't get no) saturation: A simulation and guidelines for sample sizes in qualitative research. *Plos One,* 12(7).

Van den Eynden, V. & Bishop, L., 2014. *Incentives and motivations for sharing research data, a researcher's perspective..* [Online]

Available at: https://arkivum.com/wp-content/uploads/2015/02/KE_report-incentives-for-sharing-researchdata.pdf
[Accessed 23 July 2020].

Vasileiou, K., Barnett, J., Thorpe, S. & Young, T., 2018. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology,* 18(148), pp. 1-18.

Verhulst, S., 2015. *The downside of open data.* [Online]
Available at: https://thelivinglib.org/the-downside-of-open-data/
[Accessed 4 April 2020].

Vernon, J., 2017. *American scientist: On the shoulders of giants.* [Online]
Available at: https://www.americanscientist.org/article/on-the-shoulders-of-giants
[Accessed 29 January 2020].

Vernon, J., 2017. On the shoulders of giants. *American Scientist,* 105(4), p. 194.

Whyte, A. & Tedds, J., 2011. *DCC Briefing papers: Making the case for research data management,* Edinburgh: Digital Curation Centre.

Wilkinson, M., Dumontier, M., Aalbersberg, I. & al., e., 2016. *The FAIR guiding principles for scientific data management and stewardship.* [Online]
Available at: https://rdcu.be/b4fnp
[Accessed 10 May 2020].

Williams, M., 2014. *What Did Isaac Newton Discover?.* [Online]
Available at: https://www.universetoday.com/38643/what-did-isaac-newton-discover/
[Accessed 29 January 2020].

Woelfle, M., Olliaro, P. & Todd, M. H., 2011. Open science is a research accelerator. *Nature Chemistry,* 3(10), pp. 745-748.

Wouters, P. & Haak, W., 2017. *Open data: the researcher perspective.* [Online]
Available at: https://www.elsevier.com/__data/assets/pdf_file/0004/281920/Open-data-report.pdf
[Accessed 11 October 2019].

Zhu, Y., 2020. Open-access policy and data-sharing practice in UK academia. *Journal of Information Science,* 46(1), pp. 41-52.

Zuiderwijk, A. & Spiers, H., 2019. Sharing and re-using open data: A case study of motivations in astrophysics. *International Journal of Information Management,* 49(December 2019), pp. 228-241.

Appendices

Appendix I: Questionnaire consent

Dear Participant,

My name is Raelene Galea and I am currently reading for a Master of Science degree in Management of Library and Information Services with Aberystwyth University, Wales. For my dissertation I am investigating academics' perspectives on open research data. Research data is the primary data collected by a researcher for the purpose of producing research. Such data can then be utilised by other scholars to produce new research outputs. This is the way forward in information science.

My dissertation will serve as a pre-implementation study for the creation and enforcement of an open research data policy at the University of Malta (UM). My primary aims are: to assess researchers' willingness to share and reuse primary data; to identify concerns regarding raw data sharing, and ways to overcome such concerns; to identify the benefits and drawbacks of implementing a research data policy at the UM; and to define the role of the library in research data management.

The following questionnaire can be completed online and will take approximately 5 minutes to complete. Some questions allow participants to tick more than one answer. Some questions may contain technical words, which have been marked with an asterisk. A definition for these technical words can be accessed by clicking on "More info".

No personal data will be collected so that participants will not be identified, unless they willingly agree to contribute further by participating in an interview; in which case any data collected will be handled in accordance with GDPR requirements. Ethics clearance has been attained from both the University of Aberystwyth and the University of Malta. Participation in this questionnaire is voluntary and may be withdrawn at any time. By filling out this questionnaire, you are agreeing that the data gathered can be used and analysed for the purpose of this research.

Further information regarding my research can be found hereunder:

Student's Name: Raelene Galea

Student's email address: [redacted]

Sponsor: University of Malta

University: Aberystwyth University

Department: Information Studies

Course: MSc Management of Library and Information Services

Supervisor's Name: Dr Sarah Higgins

Supervisor's email address: [redacted]

Dissertation title: Academics' perspectives on open research data: A pre-implementation study for the University of Malta

Appendix II: Information sheet and consent form

(pto)

Information sheet

Dissertation title: Academics' perspectives on open research data: A pre-implementation study for the University of Malta

Purpose of the research: To assess researchers' willingness to share and reuse primary data; to identify concerns regarding raw data sharing, and ways to overcome such concerns; to identify the benefits and drawbacks of implementing a research data policy at the UM; and to define the role of the library in research data management.

Type of research intervention: Interview

Nature of participation: Voluntary

Procedures of withdrawal: Withdrawal from the study may be done during the interview or within 24 hours after the interview. This should be done by writing an email to the interviewer asking for participation to be withdrawn.

Usage of data: Data gathered and results produced will be used for the sole purpose of this study. No personal information will be shared in the study other than the faculty with which the interviewee is affiliated. This is considered necessary since the participant's area of specialisation may affect the level of data sharing. Other information such as names and contact information will not be shared. Names of participants and signatures will only be collected for consent purposes, and safely stored (indefinitely) by the interviewer for legal purposes, in accordance with the General Data Protection Regulation (GDPR). Interviews will be recorded, and transcripts will be produced. The interviewees will be given a copy of the transcript upon request. The research output may be shared in open access.

Researcher's details:

Name: Raelene Galea

Email address: [redacted]

Sponsor: University of Malta

University supervising research: Aberystwyth University, Wales, UK

Department: Information Studies

Course: MSc Management of Library and Information Services

Supervisor's name: Dr Sarah Higgins

Supervisor's email address: [redacted]

Consent

By signing this form the interviewee is agreeing to the following:

- Due to the current pandemic situation caused by the Covid-19 Coronavirus, the interview will take place online via Google hangouts, Zoom or Skype (this will be agreed beforehand between the interviewer and interviewee on an individual basis).
- Audio-recording of the interview will be captured.
- A transcript will be produced.
- The transcript will be sent to the interviewee upon request; any errors may be amended accordingly.
- Any interview content, such as direct quotations, taken from the interview will be anonymised.

By signing this form the interviewee is also agreeing that:

- Contribution to this study by means of interview is voluntary.
- The transcribed interview may be used as described above.
- The interviewee has read and understood the information sheet attached to this consent form.

Participant's full name: _____

Participant's Signature: _____

Date: _____

Researcher's full name: _____

Researcher's Signature: _____

Date: _____

Contact Information

This research has clearance from the Aberystwyth University Research Ethics Board and the Research Ethics Committee of the University of Malta (UREC).

For any concerns regarding this research, you may send an email to Raelene Galea (researcher) on [redacted] or Dr Sarah Higgins (dissertation supervisor) on [redacted]

Appendix III: Online questionnaire

(pto)

MSc_Questionnaire_OpenData

Note: This questionnaire was distributed through Jisc. Answers could be easily selected on-screen, and some questions depended on others.

Background

1. **Choose your faculty / centre / institute / school:***

[drop-down menu with the following options:]

- Faculty of Arts
- Faculty for the Built Environment
- Faculty of Dental Surgery
- Faculty of Economics, Management and Accountancy
- Faculty of Education
- Faculty of Engineering
- Faculty of Health Sciences
- Faculty of Information, Communication and Technology
- Faculty of Laws
- Faculty of Media and Knowledge Sciences
- Faculty of Medicine and Surgery
- Faculty of Science
- Faculty for Social Wellbeing
- Faculty of Theology
- Centre for Biomedical Cybernetics
- Centre for Distributed Ledger Technologies
- Centre for English Language Proficiency
- Centre for Educational Research
- Centre for Entrepreneurship and Business Incubation
- Centre for Environmental Education and Research
- Centre for Labour Studies
- Centre for Liberal Arts and Sciences
- Centre for Literacy
- Centre for Molecular Medicine and Biobanking
- Centre for Resilience and Socio-Emotional Health
- Centre for the Study and Practice of Conflict Resolution
- Centre for Traditional Chinese Medicine
- Institute of Aerospace Technologies
- Institute of Anglo-Italian Studies
- Institute of Baroque Studies
- Institute of Climate Change and Sustainable Development
- Confucius Institute
- Institute of Digital Games
- Institute of Diplomatic Studies
- Institute of Earth Systems
- Edward de Bono Institute

- Institute of European Studies
- Islands and Small States Institute
- Institute of Linguistics and Language Technology
- Institute of Maltese Studies
- Mediterranean Institute
- Institute of Physical Education and Sport
- Institute of Space Sciences and Astronomy
- Institute for Sustainable Energy
- Institute for Tourism, Travel and Culture
- Doctoral School
- International School for Foundation Studies
- School of Performing Arts

Data Management

2. Do you preserve the research data that you gather for your research?*

Yes / No / Sometimes

a. Where do you store your data? (Dependency: Yes / Sometimes in Q.2)

- ☐ Physical storage devices (e.g. computer, laptops, external hard drive, pen-drive)
- ☐ Cloud-based storage (e.g. Google drive)
- ☐ Research data repositories
- ☐ Other

i. If you selected Other please specify:*

b. Do you document* your data? (Dependency: Yes / Sometimes in Q.2)

(Documenting your data means providing tags or descriptive information about your data, in a way that makes it easily findable and usable by yourself or by others.)

Always / Never / Sometimes

3. Do you think that the University of Malta Library should employ data librarians*?*

(Data librarians are "people originating from the library community, trained and specialising in the curation, preservation and archiving of data" (Swan and Brown, 2008, p.1).

Yes / No / I have no opinion in this regard

4. Would you use the service of a data librarian if this was available?*

Yes / No

5. Do you think that collaboration between librarians and researchers will improve the level of research data management in an institution?*

Yes / No

Data sharing

6. **Does your faculty / centre / institute / school encourage research data sharing?***

Yes / No / Faculty position is neutral in this regard

7. **Have you ever shared research data that you gathered for your research and publications?***

Yes / No

a. **Who did you share your data with? (Dependency: Yes in Q.7)**

- ☐ Member of my research team
- ☐ External individual known to me
- ☐ External individual not known to me
- ☐ Open access
- ☐ Smaller group of external people

b. **Identify the reasons why you shared your research data. You can tick more than one option.** (Dependency: Yes in Q.7)

- ☐ To increase the visibility of my research
- ☐ It was required by the funding body
- ☐ It was required by the publisher
- ☐ I believe that sharing research data increases the credibility of my research
- ☐ To contribute to the research field, allowing other researchers to use my raw data to generate new insights
- ☐ Other

i. **If you selected Other please specify:***

c. **Why haven't you ever shared your research data? (Dependency: No in Q.2)**

- ☐ It is too time-consuming
- ☐ Ethical reasons or legal concerns
- ☐ To safeguard the privacy of the research objects
- ☐ I need assistance / training in the technical aspect of sharing research data
- ☐ I do not think that I will be given the appropriate credits
- ☐ I believe that by sharing my research data I will be losing the intellectual property rights that I have on that data
- ☐ I do not believe that I will gain anything by sharing my research data
- ☐ It is private to me
- ☐ Other

i. **If you selected Other please specify:***

8. **Have you ever made use of research data generated by other researchers?***

Yes / No

a. **Was it freely available in open access? (Dependency: Yes in Q.8)**

Yes / No

b. **Why did you use research data generated by other scholars? (Dependency: Yes in Q.8)**

- ☐ To generate new findings
- ☐ To verify the research output produced by that data
- ☐ Other

i. **If you selected Other please specify:***

c. **Why haven't you ever made use of research data generated by others? (Dependency: No in Q.8)**

- ☐ I prefer to generate my own research data
- ☐ I do not think that research data that can be found in open access is reliable
- ☐ I have not had the need to use data generated by others so far
- ☐ The research data I needed was not available in open access
- ☐ Other

i. **If you selected Other please specify:***

Polies and Procedures

9. **The European Union (EU) advocates in favour of open science and open access to research. EU member states are encouraged to have their own policies in place in order to fulfil such initiatives. Are you aware that Malta is currently working on establishing a nationwide policy on open science, which will include procedures on open access to research data?**

Yes / No

10. **How willing will you be to follow such policy if it asks you to share your research data in open access?**

Very willing / Moderately willing / Neutral / Not too willing / Very unwilling

11. **Is there any additional feedback that you would like to share?**

12. **Would you like to contribute to this study by possibly participating in an interview?**

Yes / No

a. **If you answered yes, please provide your name and email address:**

Appendix IV: Interview

Academics' perspectives on open research data: A pre-implementation study for the University of Malta

Interview Questions.

1. What is your opinion on open research data?

2. In its recommendations for the development of a national open access policy in Malta, the European Commission states that open research data should be published following the FAIR principles (FAIR is an acronym for Findable, Accessible, Interoperable and Reusable) (Stern et al., 2020; Wilkinson et al., 2016). This may put an extra burden on the researcher who will need to create the metadata, present research data in adequate formats, and so forth. On the other hand, this process increases the credibility and reliability of the research output whilst possibly increasing citations. What is your opinion on this and how do you see this working out in Malta?

3. The European Commission recommend a "phase-in" approach for open access in Malta, stating that ORD should be published under the principle of "as open as possible as closed as necessary" (Stern et al., 2020). Raising awareness and investing in training and support are also recommended.

    a. What do you think of these recommendations?

    b. Is there any other recommendation that you would add for the development of a national OA policy?

4. Think of your last research project.

    a. Would you have been able to share any research data under the principle of "as open as possible as closed as necessary"?

b. Was there anything that would have held you back from sharing your research data?

5. If sharing research data had to be mandatory, what kind of training do you think should be given to Maltese academics?

   a. How do you suggest this training is offered?

   b. Should it be compulsory or optional?

6. 75% of the questionnaire respondents replied in the affirmative when asked if they would use the services of a data librarian if this was available. What would you expect of the role of a data librarian?

7. Think of one or two research projects that you did in the past or are doing right now. In what ways could a data librarian have helped you if the service was available?

8. Malta has been mentioned as one of the worst EU countries when it comes to open research data (Times of Malta, 2016), being one of the last countries to have a national Open Data Portal, not having a systematic process to gather metadata from other portals in the country, and being one of the last two countries to organise events relating to open data (European Data Portal, 2020).

   a. Do you think that researchers' resistance to share data is impeding the country from moving forward in this regard?
   b. Do you think that there are any other reasons why Malta has not kept abreast with other countries when it comes to open data sharing?

9. Is there anything else you would like to add?