

# Discipline and Research Data in Geography

By

Winnie Wan Ting Tam

Doctoral Thesis  
Submitted in partial fulfilment of the requirements  
for the award of  
Doctor of Philosophy of Loughborough University  
September 2016

© Winnie Wan Ting Tam 2016



# Abstract

Research data is essential to scholarship. The value of research data and its management has been increasingly recognized by policy makers and higher education institutions. A deep understanding of disciplinary practices is vital to develop culturally-sensitive policy, tools and services for successful data management. Previous research has shown that data practices vary across sub-fields and disciplines. However, much less is known about how disciplinary cultures shape data practices. There is a need to theorise research data practices based on empirical evidence in order to inform policy, tools and services.

The aim of the thesis is to examine the interrelation between data practices and disciplinary cultures within geography. Geography is well-established and multidisciplinary, consisting of elements from the sciences, social sciences and humanities. By examining a single discipline this thesis develops a theoretical understanding of research data practices at a finer level of granularity than would be achieved by looking at broad disciplinary groupings such as the physical and social sciences.

Data collection and analysis consisted of two phases. Phase one was exploratory, including an analysis of geography department websites and researcher web profiles and a bibliometric study of collaboration patterns based on co-authorship. Phase one aimed to understand the disciplinary characteristics of geography in preparation for Phase two. The second phase consisted of a series of 23 semi-structured interviews with researchers in geography, which aimed to understand researchers' data practices and their attitudes toward data sharing within the context of the sub-discipline(s) they inhabited.

The findings of the thesis show that there are contrasting intellectual, social and data differences between physical and human geography. For example, intellectually, these two branches of geography differ in terms of their research objects and methods; socially, they differ in terms of the scale of their collaborative activities and the motivations to collaborate; furthermore, the nature of data, how data is collected and data sharing practices are also different between physical and human geography.

The thesis concludes that differences in the notion of data and data sharing practices are grounded in disciplinary characteristics. The thesis develops a new three-dimensional framework to better understand the notion of data from a disciplinary perspective. The three dimensions are (1) physical form, (2) intellectual content and (3) social construction. Furthermore, Becher and Trowler's (2001) disciplinary taxonomy i.e. hard-soft/pure-applied, and the concepts urban-rural ways of life and convergent-divergent communities, is shown to be useful to explain the diverse data sharing practices of geographers. The thesis demonstrates the usefulness of applying disciplinary theories to the sphere of research data management.



## Acknowledgements

I would like to thank my supervisors Jenny Fry, Claire Creaser and Steve Probets for their support throughout my PhD journey, for their patience, guidance and knowledge.

I would like to thank all the interviewees who participated in the study, without whom this thesis would never have been possible.

I would also like to thank my friends for all the good moments spent together. Particular thanks to C.C. and H.C. for all the support and encouragement.

Thanks and love also to my parents and my brother for their constant support over many years.



## Table of Contents

<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1 Overview.....	1
1.2. Context .....	1
1.3 Overall aim and research questions .....	11
1.4 Methods .....	12
1.5 Expected contribution to current understanding .....	12
1.5 Definitions of key terms.....	12
1.6 Summary .....	13
<b>Chapter 2 Academic Disciplines and Research data .....</b>	<b>15</b>
2.1 Introduction.....	15
2.2 Academic disciplines and subjects.....	15
2.3 The characteristics of disciplines .....	16
2.4 Classification schemes for academic disciplines.....	20
2.5 Sub-fields and specialisms .....	28
2.6 Interdisciplinarity.....	30
2.7 Research data management and digital curation.....	32
2.8 Research and research data management lifecycle.....	33
2.9 Research Data Management surveys.....	39
2.10 The concept of data.....	40
2.11 Summary .....	43
<b>Chapter 3 Methodology .....</b>	<b>44</b>
3.1 Introduction.....	44
3.2 Philosophical assumptions .....	44
3.3 Overview of the research design .....	45
3.4 Research design of Phase 1 .....	47
3.5 Research design of Phase 2: Interviews.....	56
3.6 Ethical considerations.....	62
3.7 Data sharing arrangement.....	62
3.8 Summary .....	63
<b>Chapter 4 An Analysis of Web Content: Departments &amp; Research Groups .....</b>	<b>64</b>
4.1 Context .....	64
4.2 Data collection and analysis.....	65
4.3 The Organisation of Geography at the Institutional Level .....	66
4.4 Summary and conclusion.....	85
<b>Chapter 5 An Analysis of Web Content: Researchers' Profiles.....</b>	<b>88</b>
5.1 Introduction.....	88
5.2 Data collection and analysis.....	88
5.3 Demographic results.....	89
5.4 Findings from the analysis of web profile content.....	91
5.5 Intellectual characteristics.....	95
5.6 Methods in geography .....	99
5.7 Conclusion .....	100

<b>Chapter 6 Bibliometric analysis.....</b>	<b>101</b>
6.1 Introduction.....	101
6.2 Bibliometric study—approach taken and techniques used .....	101
6.3 Results of Part 1 .....	105
6.4 Discussion of Part 1 results.....	108
6.5 Results of Part 2 .....	110
6.6 Discussion of Part 2 .....	115
6.7 Summary and conclusion.....	117
<b>Chapter 7 Intellectual Characteristics in Geography.....</b>	<b>119</b>
7.1 Introduction.....	119
7.2 Selecting interview participants .....	119
7.3 Research topics and subject in geography .....	120
7.4 Research methods in geography .....	125
7.5 Research methods in physical geography .....	126
7.6 Research methods in human geography.....	142
7.7 Conclusion .....	163
<b>Chapter 8 Social Characteristics of Geography.....</b>	<b>165</b>
8.1 Introduction.....	165
8.2 Social activities in geography .....	165
8.3 Collaborations in geography .....	169
8.4 Collaboration in physical geography .....	170
8.5 Collaboration in human geography.....	175
8.6 Methods and issues of internal data sharing in geography .....	182
8.7 Summary .....	185
<b>Chapter 9 Data Sharing Practices in Geography .....</b>	<b>187</b>
9.1 Introduction.....	187
9.2 Diverse data sharing practices within one discipline.....	187
9.3 Data sharing practices in physical geography.....	187
9.4 Data sharing practices in human geography .....	199
9.5 Summary .....	203
<b>Chapter 10 Data Sharing Attitudes in Geography.....</b>	<b>204</b>
10.1 Introduction.....	204
10.2 Data sharing attitudes.....	204
10.3 Reasons behind supporting data sharing .....	205
10.4 Barriers and concerns about data sharing .....	208
10.5 Summary .....	217
<b>Chapter 11 Discussion.....</b>	<b>219</b>
11.1. Introduction.....	219
11.2 Characteristics of geography.....	219
11.3 A disciplinary conceptual framework for research data.....	222
11.4 Applying Becher’s taxonomy of disciplinary cultures .....	230
11.5 Conclusion .....	235



<b>Chapter 12 Conclusion .....</b>	<b>239</b>
<i>12.1 Research questions and main findings of the study .....</i>	<i>239</i>
<i>12.2 Contribution to knowledge.....</i>	<i>243</i>
<i>12.3 Practical Implications of the study .....</i>	<i>245</i>
<i>12.4 Future Research.....</i>	<i>247</i>
<i>12.5 Closing Comments.....</i>	<i>248</i>
<b>Bibliography .....</b>	<b>249</b>
<b>Appendix 1 Disciplinary coverage in studies of research data management .....</b>	<b>269</b>
<b>Appendix 2 Seed URLs used in the link analysis feasibility study .....</b>	<b>271</b>
<b>Appendix 3 Definitions of codes .....</b>	<b>272</b>
<b>Appendix 4 Revised code structure in NVivo.....</b>	<b>274</b>
<b>Appendix 5 Information Sheet .....</b>	<b>275</b>
<b>Appendix 6 Consent form .....</b>	<b>277</b>
<b>Appendix 7 Sub-fields of Geography .....</b>	<b>278</b>
<b>Appendix 8 Invitation Letter .....</b>	<b>279</b>
<b>Appendix 9 Profile of Interview Participants.....</b>	<b>280</b>
<b>Appendix 10 Interview guide .....</b>	<b>281</b>
<b>Appendix 11 Ethics Checklist.....</b>	<b>286</b>

## List of Figures

Figure 1.1 Becher's matrix of disciplinary cultures (1987) .....	8
Figure 1.2 Subfields in physical geography (adapted from Matthews & Herbert, 2008).....	9
Figure 1.3 Subfields in Human geography (adapted from Matthews & Herbert, 2008).....	10
Figure 1.4 Disciplines, subfields and specialisms.....	13
Figure 2.1 Fractal patterns (adapted from Abbott 2001) .....	29
Figure 2.2 The six data centric phases of the research cycle (Pryor, 2012:6) .....	34
Figure 2.3 The degradation of information content associated with data and metadata (Graphic from Michener et al., 1997: 332) .....	35
Figure 2.4 DCC Curation Lifecycle Model (DCC, 2013).....	36
Figure 2.5 Knowledge pyramid (adapted from Kitchin, 2014; 10).....	41
Figure. 3.1 The relationship between theory and research of this study .....	44
Figure 3.2 A screenshot of the results of the pilot study.....	51
Figure 4.1 A screenshot of the mapping exercise .....	66
Figure 4.2 A screenshot of the School of Environment, Education and Development, University of Manchester .....	69
Figure 4.3 A screenshot of the School of Geography, Politics and Sociology, Newcastle University .....	69
Figure 4.4 Numbers of geography research groups .....	73
Figure 4.5 A screenshot from the University of Exeter geography department .....	75
Figure 4.6 A screenshot from the University of Hull geography .....	76
Figure 4.7 Research themes in the University of Birmingham .....	76
Figure 4.8 Word frequency of the titles of human geography research groups .....	77
Figure 4.9 Numbers of subfields per research group .....	78
Figure 4.10 Frequency of occurrence of subfields in research groups .....	79
Figure 4.11 Relationships between subfields of human geography.....	80
Figure 4.12 Word frequency of research group titles in physical geography.....	81
Figure 4.13 Physical geography research groups in the UK .....	83
Figure 4.14 Sub-fields in physical geography (adapted from Matthews & Herbert, 2008) .....	83
Figure 5.1 Gender of research group heads.....	90
Figure 5.2 Research outputs which were submitted to ERA 2012 (Source: ERA 2012 report).....	92
Figure 5.3 Research in 60 seconds video .....	93
Figure 5.4 Twitter on institutional web profile.....	93
Figure 5.5 Personal research website .....	93
Figure 5.6 Subfields of the human geographers .....	96
Figure 5.7 Subfields of physical geography .....	98
Figure 6.1 A screenshot of the analysis of co-authors.....	104
Figure 6.2 Average number of publications in human and physical geography (2008-2013).....	105
Figure 6.3 Number of co-authored publications .....	107

Figure 6.4 The percentage of publication written by single author across fields. Data from Thomson Reuters Web of Science. Source: adapted from Nagurney, 2013 .....	<b>109</b>
Figure 6.5 Collaboration patterns in human and physical geography .....	<b>113</b>
Figure 6.6 Number of collaborators and publications in geography .....	<b>115</b>
Figure 6.7 Results comparison on a broader level .....	<b>117</b>
Figure 7.1 Field measured data .....	<b>130</b>
Figure 7.2 Multispectral infrared data (from left to right: visual, near infrared, short wave infrared, and thermal infrared) Source: Stevenson, J (2013).....	<b>133</b>
Figure 7.3 Workflow of producing a map using satellite imagery data (adapted from Storrar et al, 2013) .....	<b>135</b>
Figure 7.4 Supplementary material (Source: Storrar et al. 2013) .....	<b>137</b>
Figure 7.5 An example of an interview transcript .....	<b>147</b>
Figure 8.1 The social differences between physical and human geography .....	<b>167</b>
Figure 11.1 Conceptual Framework for Research Data .....	<b>223</b>
Figure 11.2 Different data aspects changes throughout research process.....	<b>229</b>
Figure 11.3 An example of interview data changes throughout research process .....	<b>229</b>
Figure 11.4 Degree of data sharing and hard /soft disciplinary characteristics .....	<b>231</b>
Figure 11.5 Degree of data sharing and urban/ rural disciplinary characteristics.....	<b>233</b>
Figure 11.6 Degree of data sharing and convergent/divergent disciplinary characteristics .....	<b>234</b>
Figure 11.7 Degree of data sharing and pure / applied disciplinary characteristics.....	<b>235</b>

## List of Tables

Table 2.1 Whitley's classification scheme (adapted from Whitley, 2000) .....	<b>24</b>
Table 2.2 Characteristics of Multidisciplinary, Interdisciplinary, and Transdisciplinary Research (Verbatim from Aboelela et al, 2007 : 340) .....	<b>31</b>
Table 3.1 Research design .....	<b>46</b>
Table 3.2 Research sub-questions and methods of data collection .....	<b>47</b>
Table 3.3 Summary of methods that have been considered in Phase 1, including both implemented and not implemented methods .....	<b>52</b>
Table 3.4 Initial codes .....	<b>61</b>
Table 4.1 Top 10 geography departments/ schools .....	<b>67</b>
Table 4.2 12 Institutions where geography is structured with closely related disciplines.....	<b>68</b>
Table 4.3 Institutions with broad disciplinary units.....	<b>71</b>
Table 4.4 Geography research groups of the University of Bristol.....	<b>77</b>
Table 4.5 Research centres in the top 5 UK geography departments .....	<b>85</b>
Table 5.1 15 institutions that were chosen for this part of the study .....	<b>89</b>
Table 5.2 Non-geography research group leaders .....	<b>90</b>
Table 5.3 Types of research output in physical and human geography.....	<b>91</b>
Table 5.4 Common research funders mentioned on web profiles of human and physical senior geographers ..	<b>94</b>
Table 6.1 Sampling strategy for the bibliometric study .....	<b>103</b>

Table 6.2 Categories for co-authors .....	<b>104</b>
Table 6.3 F-test.....	<b>106</b>
Table 6.4 T-test.....	<b>106</b>
Table 6.5 F-test.....	<b>107</b>
Table 6.6 T-test .....	<b>108</b>
Table 6.7 Results of part 1 .....	<b>108</b>
Table 6.8 Number of authors per publication in different research fields .....	<b>110</b>
Table 6.9 F-test.....	<b>111</b>
Table 6.10 T-test.....	<b>111</b>
Table 6.11 Results of collaboration pattern in geography .....	<b>112</b>
Table 6.12 Categories which have significance differences between human and physical geographers .....	<b>114</b>
Table 6.13 Results of part 2 .....	<b>115</b>
Table 7.1 Sample subfields .....	<b>120</b>
Table 7.2 Sample Institutions.....	<b>120</b>
Table 7.3 Research topics and methods of interviewees .....	<b>121</b>
Table 10.1 Reasons behind supporting data sharing.....	<b>205</b>
Table 10.2 Barriers to data sharing .....	<b>209</b>
Table 11.1 Different methods used by physical and human geography.....	<b>219</b>
Table 11.2 Differences in social organisation of physical and human geography .....	<b>220</b>
Table 11.3 Methods and data characteristics in physical and human geography.....	<b>222</b>

# Chapter 1 Introduction

---

## 1.1 Overview

This thesis investigates the mutual shaping relationship between the data practices of researchers and disciplinary cultures, focussing on the discipline of geography. In this chapter, the rationale for the research is explained. The chapter commences by establishing the context and importance of the topic. The research questions of this study are then given. Following that the expected contribution to current understanding is explained and key terms are defined.

## 1.2. Context

Research data are fundamental to scholarship and research. As Borgman (2007: 115) states, data are “outputs of research, inputs to scholarly publications, and inputs to subsequent research and learning”. Managing research data has become an important topic in information science (Corrall, 2012; Borgman, 2015). There is an increasing recognition of data as a primary research output and a need for data to be actively and properly managed (Borgman, 2007; Harley et al, 2010; Tenopir et al, 2013; Nielsen & Hjørland, 2014; RCUK, 2016). This focus on research data reflects a wider trend towards a concern with the management and sharing of data.

On 18 June 2013, the British government and other G8 leaders signed an Open Data Charter. The G8 Open Data Charter (2013) set out five strategic principles, including open government data by default, increasing the quality, quantity and re-use of the data and releasing high value data. Fourteen areas where data was particularly high value were identified, areas such as science and research, health, geospatial and earth observation. It is recognised that citizens of all nations will benefit from open data, as open data will unlock the economic potential of data, drive innovation and increase accountability (Open Data Charter, 2013, Open Data Strategy, 2014). This is just one of a number of developments reflecting an increasing move towards open access to, and the re-use of, data. Today many research councils and funders worldwide, such as the funding councils and the Wellcome Trust in the UK, the National Science Foundation and National Institutes of Health in the US, and the Australian Research Council require data to be managed according to best practices and standards. It is evident that policy-makers and research funders believe that research data are a crucial part of producing knowledge and that opening up data will benefit society. As a consequence, Higher Education Institutions (HEIs) are under pressure to develop policies and to provide infrastructure and supporting services related to research data management. Effective

research data management, however, is not simply about creating policy, infrastructure and putting support services in place. Without an understanding of the diverse processes, values, practices and needs of knowledge communities within disciplines, research data may not be shared and data repositories may not be used (Nelson, 2009; Tenopir et al., 2011). In order to understand the full potential of research data in relation to data sharing and management, a deeper understanding of actual research practices, cultural contexts and discipline specific needs are required (RIN,2010; Herold, 2015; Lämmerhirt, 2016).

### 1.2.1 Recognition of the importance of research data and its management

The value of research data and its management is increasingly recognised. One of the reasons is because of the advancement of technologies. In the era of big data, data has been increasing in volume, velocity and variety (Laney, 2001; Kitchin & Lauriault, 2015). Advanced information and computing technologies, including grid computing and digital technology, not only generate large scale scientific data-sets at much greater speed but also support the creation of large quantities of social sciences and humanities data such as images, texts, videos and audio files in digital form (Dutton & Jeffreys, 2010). The practice of scholarly work is shifting as a result. The proliferation of advanced digital technologies, networks and a vast growth in the availability of research data are supporting data exchange. Researchers are able to do research using new methods and increase research productivity. For example, ecologists are able to conduct large scale data analysis on a global-scale more efficiently by re-using and integrating datasets produced by other researchers, (Webb, 2013). Geographers discuss the opportunities which big data and new computational techniques can bring for geographical research, e.g. Graham & Shelton (2013), Miller (2014) and Shearmur (2015). Data-driven scholarship is increasing. It is evident that the “data deluge” (Hey and Trefethen, 2003) is transforming the landscape of worldwide scholarly communication and the information infrastructure for research (Lord et al, 2004; Borgman, 2007; Jankowski, 2009; Pryor, 2012; De Roure, 2014).

The awareness of the need to manage research data as an important institutional and national resource (both within and beyond academia) is growing internationally. Governments and research funders worldwide are increasingly seeing research data as valuable assets in institutions and the wider society. For example, the European Cloud Initiative by the European Commission (2016) states that research data plays an essential role in driving innovation and business. However, it has been realised that the new opportunities and new knowledge resulting from the data deluge can only emerge if data are effectively managed, accessible and usable (NSF, 2007; Royal Society, 2012; RCUK, 2016). Worldwide, a number of reports have emphasised research data as vital resources

whose value needs to be preserved for the research of the future, for example, the National Science Foundation's *Long-lived data collections* (2005) and *Cyberinfrastructure vision for 21st century discovery* (2007) in the US; the *From data to wisdom* report by the Prime Minister's Science, Engineering and Innovation Council (PMSEIC) (2006) and *Open research data report* (Houghton & Gruen, 2014) in Australia; the *eSciences data curation report* (Lord & Macdonald, 2003), the *Developing the UK e-infrastructure for science and innovation* (Poschen, 2007) and the *Science as an open enterprise* (Royal Society, 2012) in the UK; and *Making Science as a reality* (2016) from the OECD. It is evident that there has been an increasing strategic awareness of the need for data curation and data infrastructure on an international scale.

Open source, open access, open data and open science initiatives also have a strong influence on the political environment and policy making on data curation (Willinsky, 2005; Lyon, 2007; Open Data Institute, 2015). These international open science initiatives promote open access to information, data, ideas and research outputs (Berlin Declaration, 2003, Budapest Open Access Initiative, 2002, Open Source Initiative, 1998, Open Data Institute, 2015). They encourage open cultures in science and scholarship. Governments are increasing public access to their data for supporting transparency, driving innovation and stimulating social and economic growth (Open Data Strategy, 2014).

In the UK, the government is committed to supporting open public access, not only to government data, but also to published research and research data. The UK government is one of 34 countries that has signed the OECD Declaration on Access to Research Data from public funding in 2004 (OECD, 2004). In 2005, Research Councils UK (RCUK) published a position statement on access to research outputs. The RCUK position statement states that "publicly-funded research must be made available and accessible for public use" (RCUK, 2005). Given that data have been increasingly recognised as an important research output, research funders have applied the principles of openness to publicly funded research data. In April 2011, RCUK published a set of *Common Principles on Data Policy*. The principles advocate publicly funded research data as a public good, produced in the public interest, should be made openly available, discoverable and re-usable. The principles require that research data should be managed according to best practice and standards (RCUK, 2011). The *Guidance on best practice in the management of research data* (RCUK, 2015) provides further guidance on the interpretation of the common principles. Now all UK research funders, including Arts and Humanities Research Council, Biotechnology and Biological Sciences Research Council, Cancer Research UK, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Medical Research Council Science and Technology Facilities Council, Natural Environment Research Council and the Wellcome Trust, have data policies which

require researchers proposing a project to submit a data management and sharing plan with their grant application. In addition, the Engineering & Physical Sciences Research Council (EPSRC, 2011, 2015) required all institutions in receipt of its research funding to develop a roadmap and to achieve full compliance with EPSRC's research data management expectations on and after May 2015. The *Concordat on Open Research Data* (2016) developed by RCUK, JISC, the Wellcome Trust and Universities UK was also published. The 10 principals assume that data sharing should be the default, though also recognising the rights of researchers to first use of data and various exceptions. Furthermore, European Commission (2016) has declared that data from research projects under the Horizon 2020 research and innovation programme will be open up by default from 2017. All these initiatives provide evidence of a strategic push to improve research data management capability.

Universities are among the institutions under pressure to manage research data more effectively. The Data Curation Centre has suggested that formal policy for guidance is necessary for effective data curation (Whyte & Tedds, 2011). As a result many institutions have developed policies related to research data management (JISC, 2016; Brown et al, 2015; Digital Curation Centre, 2013; Whyte & Tedds, 2011). At least 37 universities now have a formal Research data management policy (Horton & DCC, 2016), such as the University of Hertfordshire's "Data Management Policy" (2011), the University of Oxford's "Policy on the Management of Research Data and Records" (2012), the University for the Creative Arts' "Research Governance" (2013), the University of Glasgow's "Research Data Management Policy" (2015) and the University of Salford's "Research Data Management Policy" (2016). Institutions are working on infrastructures and services to implement these policies.

### 1.2.2 Benefits of research data management

It is argued that the potential benefits of managing, sharing and re-using research data brings significant benefits for researchers and research communities and society more broadly (RDMRose, 2013; DCC, 2013; Beagrie & Pink, 2012; Fry et al, 2009). Managing research data is about "maintaining and adding value to a trusted body of digital information for current and future use" (Digital Curation Centre, 2013; Pennock, 2007). The point being made in the preceding quote is that for researchers having data management strategies should not only be about complying with requirements of funders. Effective research data management helps preserve research data and it minimises the risk of data loss and technological obsolescence. It has been suggested that digital data are very fragile (DCC, 2013). The increasing volume of data creates storage problems and makes data difficult to retrieve (Borgman, 2007). Digital data are at risk not only from the increasing volume of digital data but also from technological obsolescence (Lord et al, 2004; DCC 2013). For



example, data stored in a wide range of media formats, software and hardware can become obsolete in a few years. Some research data, such as observational and environmental data, are collected at a unique place and time, and cannot be recreated. It is vital to preserve such valuable data (Borgman, 2007; DCC, 2011). Thus, effective research data management protects researchers' data against potential loss and increases opportunities for researchers to reuse their data.

Research data management also enhances data security and provides access to research data. Good data management practices provide clear documentation, metadata, provenance and secure storage for data, enabling research data to be searchable, accessible and usable, not only for research but also for teaching (Beagrie & Pink, 2012).

Effective research data management helps improve the quality and trustworthiness of research data, as it emphasises the accurate, complete, authentic and reliable collection of data (University of Oxford, 2012). It not only preserves the data but also allows reliable verification of results externally, ensuring research integrity (Van den Eynden et al, 2011). Thus researchers can access, work and draw conclusions from reliable data.

Furthermore, it encourages re-use and sharing of research data. It is believed that new research can be built on pre-existing data (Veerle Van den Eynden et al, 2011). As mentioned earlier, research data management concerns the processes for creating and organising data, such as folder structures, versioning, naming conventions and documentation of data etc. It ensures research data remains accurate, valid, reliable and complete (Higgins, 2012). All these attributes increase researchers' confidence in open data and encourage data reuse and sharing which maximises the impact and global visibility of research as well as enhancing researchers' reputations (Beagrie & Pink, 2012). It allows datasets to be used in new and innovative ways, encouraging new collaboration between different disciplines, specialisms and transferring knowledge to industry. In addition, it can reduce the duplication of data collection costs (Fry et al., 2008). Borgman (2012) also states that there are four common rationales for sharing research data:

1. to reproduce or verify research
2. to make publicly funded research available to the public
3. to enable others to ask new questions of the data and
4. to advance the state of research and innovation.

Clearly, there are many benefits and incentives associated with effective research data management, although not all these benefits and rationales are relevant to all types of data and all disciplinary communities. However it is clear that improvements in research data management may

also bring about changes in the manner in which research is conducted, for example effective research data management may lead to changes to the research culture and practices of research communities in different ways, e.g. changes in research approach, the norms of data re-use, formal/informal sharing or increase in interdisciplinary research.

### 1.2.3 Disciplines and research data management

An understanding of disciplinary research practices and needs is essential to the development of policy, tools and services for successful research data management (RIN, 2008; Cragin et al, 2010; Cox et al, 2014). There are an increasing number of studies of research data management in the disciplinary context e.g. RIN (2008), Lyon et al. (2010), Van den Eynden & Bishop (2014), Borgman (2015). One of the key disciplinary focused projects in the UK was the SCARP<sup>1</sup> project by Research Information Network (RIN) and the Digital Curation Centre (DCC). The project investigated disciplinary attitudes and approaches to data curation (Lyon et al, 2010; RIN, 2008). In addition, discipline-focussed research data management training materials (RDMTrain) have been developed by JISC (2011). The body of literature in this field suggests that there are not only variations in types, file formats, value and long term viability of research data in different disciplines but also significant differences in data practices and culture of sharing data. For example, according to the RIN report (2008), re-use of data and data sharing is a norm in astronomy, while in climate science, the culture of data sharing is not strong.

Although some studies have been carried out on research data management in the disciplinary context, such as, RIN (2008), Lyon et al (2010), Whyte & Pryor (2011) and Aker & Doty (2013), much less is known about how these differences are shaped by the culture of disciplinarity. It is evident that practices of data creation, use and sharing vary widely across subject disciplines and their sub-disciplines. However, these differences are actually shaped by something more fundamental, the nature and the complexity of disciplines, such as heterogenous processes and practices of constructing knowledge within disciplines, informal and formal methods of sharing and collaboration patterns. Thus, this research will attempt to provide insights into the practices of research communities and explore how these shape data practices of researchers.

The nature and use of research data is strongly related to disciplines (Cragin et al. 2010; Williams & Pryor, 2009; Pryor, 2009). Effective support of research data management, therefore, is dependent on a detailed understanding of the characteristics of disciplines and data creation and use within

---

<sup>1</sup> DCC SCARP project <http://www.dcc.ac.uk/projects/scarp>

knowledge communities. As Corral et al (2013) state, there are different needs across disciplines. Terms such as, 'specialization', 'fragmentation', 'hybridity' and 'fluidity' have been used to describe the dynamic nature of disciplinarity (Klein, 1996; Dogan & Pahre, 1990). As knowledge becomes increasingly interdisciplinary, some research communities borrow concepts and methods from another discipline and become hybrid disciplines or sub-disciplines. Due to the dynamic nature of disciplines, data are highly specialized, fragmented and diverse. As knowledge production becomes more and more interdisciplinary in nature, the hybridity and fluidity of disciplines also create complexity in relation to research data management.

Disciplines provide basic structures for organising knowledge and research in Higher Education Institutions. Becher and Trowler (2001) show there are intellectual differences, e.g. research problems, research objects, and methods, and also social differences e.g. community culture, communication patterns, and funding systems, between disciplines and sub-disciplines. For instance, history typically uses interpretative methods and historians tend to work independently; while high energy physics typically favours quantitative methods, and large scale collaborations. It appears that history is more likely to generate small scale data, e.g. documents, manuscripts, photos, than the large computational datasets, digital born data, which are often generated in disciplines such as high energy physics. Furthermore, computational datasets are usually highly structured and anonymised, whereas interview data is more personalised and not easily anonymised (Borgman, 2012). This shows that the characteristics of disciplines are likely to affect approaches to managing data and attitudes towards data sharing.

It seems likely, therefore, that the nature of disciplines, e.g. research methodologies, research problems, processes and practices within and across disciplines and sub-disciplines, shape the nature of the data created, ways in which data are manipulated and stored, the possibility of data sharing and reuse, and what constitutes effective data management practices. It is essential, therefore, to understand how disciplinary cultures shape research data practices. In addition, there is a need to better theorise research data practices. Theoretical frameworks will provide a systematic way to understand research data and data sharing practices of researchers. Conceptualised understandings will help appropriate measures to be effectively developed to increase awareness, encourage best practices and further develop appropriate strategy, services and infrastructures for research data management.

#### 1.2.4 Developing a disciplinary perspective

Several classification schemes of academic disciplines have been developed to explain differences between disciplines (Braxton & Hargens, 1996). For instance, one of the well-known classification

schemes is Becher's (1987) matrix of disciplinary cultures, based on the hard-pure, hard-applied, soft-pure and soft-applied dimensions of knowledge (as shown in Figure 1.1 and explained more fully in Section 2.5.1). The terms hard/ soft, pure/ applied are used to refer to the knowledge structures of disciplines and their intellectual differences in disciplines (Becher & Trowler, 2001). In order to develop an understanding from a disciplinary perspective, one of the common approaches is to compare across a number of disciplines. For example, the DCC SCARP project examined disciplinary differences in research data sharing, and reuse by comparing sixteen case studies of disciplines (Key Perspectives, 2010).

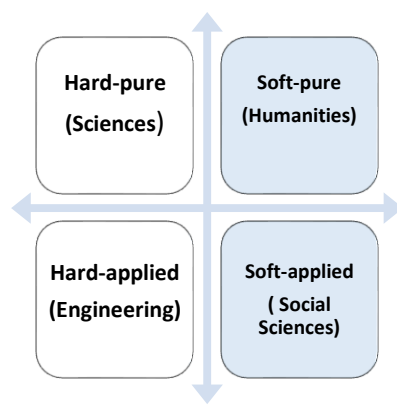


Figure 1.1 Becher's matrix of disciplinary cultures (1987)

However, in this research, comparisons within a single discipline are given more emphasis. Studies based on broad disciplinary categories may falsely represent disciplines as being unified. As mentioned earlier, there are an increasing number of studies focused on research data from a disciplinary perspective. e.g. RIN (2008) Lyon et al. (2010), Van den Eynden & Bishop (2014), Borgman (2015). A range of disciplines have been studied, for example, astronomy, systems biology, genomics, classics and sociology. Yet, overall, it appears that there have been relatively more studies focusing on the Sciences than on the Social Sciences (see appendix 1). Most of the results of these studies show that much diversity exists even within a single discipline e.g. at a sub-disciplinary level (Key Perspectives, 2010; RIN, 2008).

In order to provide an insight into how the nature of discipline shapes, and is shaped by research data practices, this research will focus on one single social science discipline, namely geography. Geography is chosen because despite being a 'data-intensive' discipline there is only limited research into its research data management and data practices, with a few exceptions, for example there are some studies of awareness/activities of curating geospatial data (McGarva; 2006; Bose & Reitsma, 2005).

Geography is the study of the surface of the Earth, which includes the 'Earth's landscapes, people, places and environments' (Royal Geographical Society, 2013; Matthews & Herbert, 2008:14).

Geography is a unique discipline. It bridges the social sciences (human geography) with the natural sciences (physical geography) (Royal Geographical Society, 2013). In fact, modern geography is not only connected to physical sciences and social sciences, but also the humanities (Matthews & Herbert, 2008; Holloway et al, 2003).

It is an interesting discipline to investigate because there is a clear division between the two main branches in geography: physical geography and human geography. Physical geography concerns the sciences of physical landscapes and environmental processes. It belongs to the 'hard' dimension of Becher's (1987) matrix of disciplinary cultures (Figure 1.1 and explained more fully in Section 2.4.1). In contrast, human geography belongs to the 'soft' side of Becher's (1987) matrix. It concerns human activities on the Earth, including cultures, societies and economies from a spatial perception.

Today, both physical geography and human geography consist of a wide range of sub-disciplines, reflecting diverse research problems in the discipline. Matthews & Herbert (2008) identified six core sub-disciplines of physical geography, i.e. geomorphology, geocryology, hydrology, soil geography, biogeography and climatology. Besides these six sub-disciplines, physical geography has had a close relation to other science disciplines and has developed some well-established inter-disciplines, such as Quaternary science and geo-archaeology as shown in Figure 1.2.

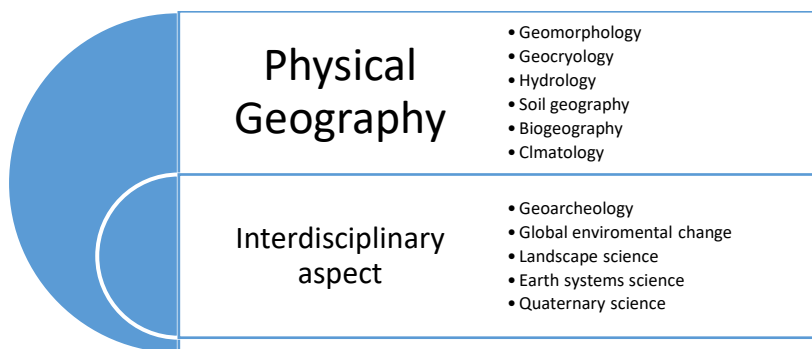
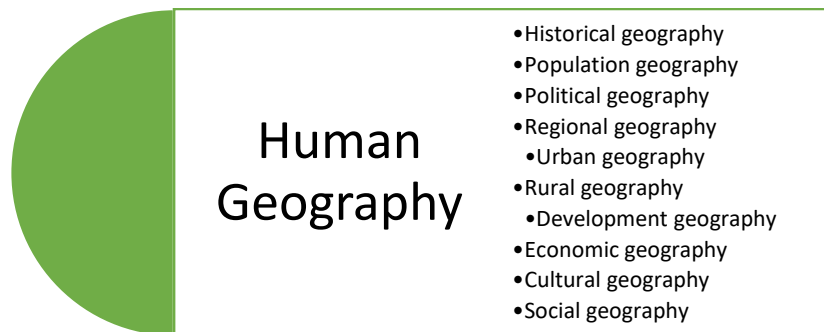


Figure 1.2 Subfields in physical geography (adapted from Matthews & Herbert, 2008)

Compared to physical geography, human geography has a wider range of subfields. Human geography is commonly divided into subfields, such as, historical geography, cultural geography, development geography, urban geography, economic geography, political geography, population geography and health geography (Matthews & Herbert, 2008; ESRC, 2013) (Figure 1.3). However, it should be noted that the nature of these labels of subfields are rather artificial. Matthews & Herbert (2008) suggest that the purpose of these divisions is mainly for supporting undergraduate teaching

and to bring academics with similar interests together. These subfields tend to be further subdivided into specialisms or form new fields which overlap between divisions, yet the labels of the subfields are always retained and used loosely (Matthews & Herbert, 2008).



*Figure 1.3 Subfields in Human geography (adapted from Matthews & Herbert, 2008)*

Another reason that geography has been chosen for this study is because it is a well-established discipline in universities. Today, geography departments can be found in most parts of the world, for example, the US, Australia, South America, China, Japan, India (International Geographical Union, 2013). In the UK, 49 institutions, involving more than 1000 academic staff, were submitted to the Research Assessment Exercise, the Geography and Environmental Studies sub-panel, in 2008 (RAE, 2008). According to an international benchmarking review of UK human geography, UK human geography ranks first in the world (ESRC, 2008). Bibliometric data shows that UK human geography exceeds other countries in both volume and citation impact (ESRC, 2008). The review described research in human geography in the UK as being “radically interdisciplinary”, and of having “intellectual diversity” and “significant theoretical and methodological innovation” (ESRC 2008: 27). This means that its data practices in geography could be particularly dynamic and interesting.

Geographers employ a wide range of methods for data collection, such as, secondary data, questionnaire surveys, interviews, visual images, observations and measurements in the field, numerical modelling, laboratory methods, spatial modelling and remotely sensed images (Clifford et al, 2010). It is noted that some cultural and historical geographers collect written text and arts images as their data, whereas soil geographers collect remote sensing data (Clifford et al, 2010; de Paul Obade & Lal, 2013). Broadly one could claim that human geographers collect qualitative data, while physical geographers collect quantitative data. However, the nature of geography is dynamic and changing, with research practices in geography being reshaped by technological innovations. As Kitchin (2014) pointed out geographers are increasingly benefiting from open data, i.e. data which can be accessed and used by everyone (ODI, 2015). For example, geographers are able to get access to high value administrative and government datasets, e.g. the ONS Postcode Directory, because of

the open data movement. Researchers can potentially generate high social and commercial value data from open data (Wessels et al, 2017). Today, human geographers not only collect qualitative data but also quantitative data. For instance, among human geographers, economic geographers analyse quantitative spatial data, using open data as secondary sources, feminist geographers adopt GIS technology for qualitative studies (deLyser, 2014; Elwood & Cope, 2009; Kwan, 2002). It is evident that the nature of geography as a discipline and the very nature of the data are mutually shaping.

In general, geography has both hard and soft disciplinary characteristics (Becher & Trowler, 2001), it also has strong interdisciplinary links with other academic fields. Thus, by examining this single discipline, it will be possible to compare research practices at a finer level and gain understandings of the inter-relationship between cultural characteristics of a discipline and its research data.

### 1.3 Overall aim and research questions

On the basis of the need to conduct more research to understand disciplinary data practices, this research is underpinned by the following research question:

In what ways are the data practices of researchers in geography shaped by the nature of the discipline?

The aim of the research is to investigate the relationship between data practices of researchers and the culture of a discipline, focusing on the discipline of geography.

A possible outcome of developing this detailed understanding of the mutually shaping relationship between data and disciplinary cultures would be informed national and institutional research policy development based on cultural sensitivity. A deeper understanding of the applicability of disciplinary theories, such as Becher and Trowler's matrix, in explaining data practices within a single discipline will also be developed.

In order to answer the above question, the following sub-questions were developed:

1. What are the characteristics of geography as a discipline (i.e. intellectual, social and organisational aspects)?
2. What are the data practices (i.e. collecting, managing, using and especially sharing data) of researchers in geography and are there identifiable patterns within or across sub-disciplines?

3. Do patterns of data practices, especially data sharing, vary across different sub-disciplines geography and, if so, why?

## 1.4 Methods

As the nature of discipline is very complex, any single method or technique alone is unlikely to be sufficient to explore the dynamic and complex relationship between the intellectual and social structures of a discipline and research data practices. Thus this research uses multiple methods. The research consists of two phases. The first phase of this research was exploratory, relying on a range of desk-based research, including an analysis of geography department websites and researcher web profiles and a bibliometric study of collaboration patterns. This first phase of the research aimed to provide an understanding of geography as a discipline before beginning the interviews. The second phase of this research was a series of interviews with researchers in geography. The aim is to understand researchers' data practices, their attitudes toward data sharing, and also for them to reflect on the impact of current changes around the discipline and data environment.

## 1.5 Expected contribution to current understanding

This research will provide a better understanding of how data practices of researchers are grounded in the underlying culture of disciplines. The research will provide a theoretically valuable contribution to deepen understanding of the research data and data sharing practices of researchers. It will enhance the conceptual theory of disciplinarity and data through detailed analysis of research practices in geography. In addition, it will contribute to the practical implementation of effective research data management by deepening our understanding of the culture of disciplines.

## 1.5 Definitions of key terms

### *Disciplines, subfields and specialisms*

A discipline can be broadly defined as a field of knowledge. A discipline consists of sub-units and sub sub units. For example, in the discipline of geography, there are two main branches, i.e. human geography and physical geography. Human geography consists of social geography, cultural geography, economic geography, development geography etc. Within social geography, there are some smaller units, such as, geography of children, feminist geography.



In the literature terms such as segments, sub-disciplines, subfields, specialisms, sub-specialisms, specialities are commonly used to describe the sub units or sub-sub units of a discipline, e.g. Becher (2001) and Klein (1996). Sometimes these terms are used as synonyms. There is no standardisation of how to use them. This lack of standardisation reflects the reality that the internal structure of a discipline is complex. As Klein (1996) suggested the boundaries of these smaller units of a discipline are rarely easy to differentiate. The presence of these terminologies shows an attempt to capture the internal structure of a discipline, a field of knowledge, which is fundamentally abstract. Yet these terms are essential to help studies to conceptualise the theoretical aspect of the nature of disciplines.

In this study, the terms 'subfield' and 'specialism' are preferred. 'Subfield' is used to refer a subunit of a discipline. The term 'subfield' is preferred to be used than 'sub-discipline', because 'sub-discipline' implies a subdivision within one discipline. While in this research, 'subfield' is not only seen as a sub-division of one discipline, but it can also be a sub-division of other disciplines (Figure 1.4). For example, development geography is a subfield of geography, but it also can be as a subfield of the discipline of development studies. The term 'specialism' is used as a finer unit than subfield, such as, a sub-subunit of a discipline. For example, children's geographies is a specialism of geography.

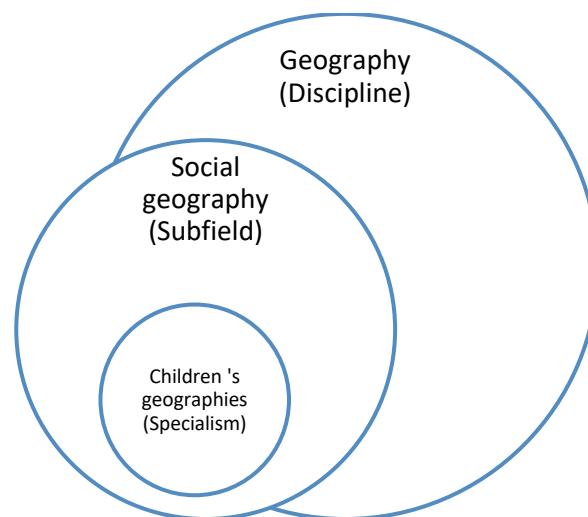


Figure 1.4 Disciplines, subfields and specialisms

## 1.6 Summary

The chapter has established the context and importance of the topic, outlining the increasing importance of managing research data effectively and the importance to understand disciplinarity. It has also explained the rationale behind focussing this research on one single discipline which is

geography. The research questions, aims and objectives have been presented. The main goal of this study is to explore the inter-shaping relationship between data practices in geography and disciplinary cultures using theoretical concepts of disciplines, which is a previously under-explored area. A deeper understanding of the applicability of disciplinary theories will help to achieve effective research data management.

# Chapter 2

## Academic Disciplines and Research data

---

### 2.1 Introduction

This chapter presents a review of previous literature on academic disciplines and research data. The first section focuses on academic disciplines. It begins by discussing the notion of academic disciplines, examining different classification schemes for them. It then discusses smaller units of discipline, i.e. subfields and specialisms, and interdisciplinarity. The second section focuses on research data. It begins by identifying differences between research data management and data curation. It then examines the concept of data. It further discusses the nature of the research data lifecycle, and institutional research data management surveys.

### Part 1 The notion of academic disciplines

Academic disciplines are very diverse. Each discipline and their sub-fields have their own character, including intellectual and social cultural characteristics. There are diverse methodologies and research instruments, different patterns of communication and collaboration, different research practices among disciplines. All of which imply that there are not only significant differences in data types but also in practices in managing and sharing data among researchers. In order to explore how disciplines shape researchers' data practices, first, it is essential to understand the notion of academic disciplines.

### 2.2 Academic disciplines and subjects

Academic disciplines are often identified as a branch of knowledge in higher education, or subjects taught at university. Disciplines are commonly seen in curricula (Berger, 1970). Yet, disciplinarity not only has a close connection with learning and teaching but also with research (Trowler et al, 2012). Disciplinary variations can be found in both teaching and learning practices and research practices in Higher Education (Lindblom-Ylännea et al., 2006; Becher & Trowler, 2001; Aditomo et al, 2013). Parker (2002: 374) argues there are subtle differences between "subject" and "discipline":

*“Subject’ is reassuringly concrete - a subject can be defined, has a knowledge base which can be easily constructed into a programme of knowledge acquisition and, perhaps most importantly, of quantitative assessment. Subjects are inclusive - anyone studying on a*

*subject programme belongs, whereas 'discipline' brings with it tricky questions about access and boundaries: about inter- and multi-disciplinarity, about who can be said practising the discipline. However, subjects are also passive—they are taught, learned, delivered."*

Trowler et al. (2012) maintain that curriculum structures are different from knowledge structures; subject is about disciplines as curriculum but not as research.

## 2.3 The characteristics of disciplines

The notion of academic discipline is very complex, it is not easy to identify a concise and conclusive definition of disciplines in the literature (Trowler et al, 2012, Krishnan, 2009). Several attempts have been made to define what disciplines are, e.g. Goodlad (1979), Donald (2002), Turner (2000), Weingart & Stehr (2000), Trowler (2012) and Jacobs (2014). In general, there are several approaches to understand the notion of disciplinarity.

### *Intellectual characteristics*

The first approach is concerned with the intellectual characteristic of a discipline, i.e. disciplinary epistemology, the form and focus of knowledge. Berger (1970) defines a discipline as

*"a specific body of teachable knowledge with its own background of education, training, methods and content areas".*

Donald (2002:8) refers to disciplines as

*"a body of knowledge with a reasonably logical taxonomy, a specialised vocabulary, an accepted body of theory, a systematic research strategy, and techniques for replication and validity".*

Both Berger's (1970) and Donald's (2002) definitions emphasise that a specific body of knowledge forms the essence of disciplines. They also show the significance of the intellectual character of disciplines, i.e. each has research problems, concepts, theories and methods of investigation which are specific to themselves (Becher & Trowler, 2001; Chettiparamb, 2007). However, the "body of knowledge" approach is not enough to fully understand what disciplines are. It has been noted that disciplines are always unfairly described as fixed, closed and limiting (Christie & Maton, 2011).

Trowler et al (2012) criticises the above definitions by Berger (1970) and Donald (2002), indicating that they only emphasise the form of knowledge as the essence of disciplines. The definitions have a static feel and fail to reflect the dynamic nature of disciplines which are continually changing and evolving (Becher & Trowler, 2001; Trowler et al, 2012).

### *Social characteristics*

The second approach to understand disciplines is concerned with the social and cultural aspects of a discipline. As Storer states (1966:3), disciplines act not only

*“as a body of knowledge or as a set of methods and techniques but as... the organised social activity of men and women who are concerned with extending man’s (sic) body of empirical knowledge through the uses of these techniques. The relationships among these people, guided by a set of shared norms, constitute the social characteristics”*

There is a very close relationship between knowledge forms and knowledge communities, as particular academic communities are organised in a way which is related to the specific body of knowledge in which they are engaged (Becher & Trowler, 2001). Each discipline not only has their distinctive academic ideas but also a distinctive culture, such as their own values, attitudes and behaviour. Becher and Trowler (2001:23) define culture

*“as sets of taken-for granted values, attitudes and ways of behaving which are articulated through and reinforced by recurrent practices among a group of people in a given context”.*

Becher and Trowler (2001) maintain that disciplinary epistemology and disciplinary culture are mutually dependent and inseparable. For instance, disciplinary epistemology acts as a structural factor which determines the intellectual tasks, shapes the groups and conditions, the practices and attitudes within academic communities. On the other hand, academic communities build, share and develop narratives about the nature of disciplinary knowledge; the beliefs, attitudes and behaviours within communities act as an actor shaping knowledge forms, for example, through the way they interpret the nature of knowledge in their disciplines ( Clark,1983 ; Becher & Trowler, 2001; Trowler et al., 2012).

Academic disciplines are dynamic in nature. One of the reasons is because to some extent knowledge and disciplines are socially constructed (Turner, 2000; Becher & Trowler, 2001; Young, 2008). As Weingart & Stehr (2000 : xi, xiv) states,

*“ Disciplines are not only intellectual but also social structures, organisations made up of human beings with vested interests based on time investments, acquired reputations, and established social networks that shape and bias their views on the relative importance of their knowledge...disciplines are diffuse types of social organization for the production of particular types of knowledge”.*

The social aspect of a discipline is significant. Disciplines can be seen as networks and social circles. Crane (1972) suggested that there is a vast interpersonal network and set of social circles behind the impersonal structure of knowledge. Individuals work as a group to discover and build new knowledge in academic disciplines. They work on similar research problems, share an interest in a common subject, communicate formally and informally and establish networks and form ‘invisible colleges’ (Price, 1963; Crane, 1972). However, these communities are subject to the influences of institutions. It is noted that researchers in a discipline would not be able to pass their knowledge to the next generation unless the discipline has been through the process of institutionalisation, which enables the discipline to transfer the knowledge by learning and teaching in universities (Goodlad, 1979; Krishnan, 2009).

### *Organisational characteristics*

The third approach to understanding the notion of disciplines is concerned with the organizational, institutional aspect of a discipline. Disciplines act as structures for knowledge and research (Del Favero, 2004). It appears that disciplines provide an organisational structural framework to Higher Education Institutions, such as faculty and departments. In this approach, disciplines are always represented by the existence of academic departments (Becher & Trowler, 2001; Turner, 2000). However, as Becher and Trowler (2012:41) state “it does not follow that every department represents a discipline”. For example, some universities establish departments in parapsychology and viticulture, however those departments’ academic validity and appropriateness of subject are challenged by the views of academics, and they tend to be regarded as fields of study instead of disciplines (Becher & Trowler, 2001). There are variations between departments in the same discipline. One of the reasons is because HEIs manage and organise knowledge and disciplines in different ways and with different styles (Lodahl & Gordon, 1972). As Becher and Trowler (2001) state, what counts as English, Mathematics or History can vary from institution to institution. For example, one institution might consider informatics as “the study of the structure, the behaviour, and the interactions of natural and engineered computational systems”( School of Informatics, University of Edinburgh, 2012) whereas another institution considers informatics studies as the study of “the application of information technology to practically any field, while considering its impact on individuals, organizations, and society” (Illinois Informatics Institute, University of Illinois, Urbana-Champaign, 2011).

Furthermore, departments in the same discipline can also be varied because of other factors, including historical, geographical and economic factors. There are different education systems, different educational ideologies and cultures, different stages of economic development in different

countries that affect the organisation of disciplines (Becher & Trowler, 2001). For instance, History is closely related to national culture. It cannot be a culture free discipline. The English Marxist historian, German Marxist historian and the Russian Marxist historian can all have very different interpretations of Marxism (Becher & Trowler, 2001). In Podgorecki's (1997) cross-national study of university culture, it is evident that different disciplinary practices and approaches differ in different countries, such as between China, United Kingdom, Poland and the United States. Although there are significant differences between departments in the same disciplines, they are still from the same branch of knowledge. By contrast, Turner (2000) argues that departments in different universities in the same discipline are interchangeable. Not only is the disciplinary epistemology similar to another department but also the business of a department, such as " hiring, tenuring, promoting, teaching, advising, granting degrees, placing students" are more or less the same (Turner, 2000: 46). Turner (2000:46) states

*"departments corresponding to disciplines have a long history in universities, and they represent, with some quite interesting national differences, a more or less standard worldwide organisational form; consequently, the disciplines they represent make for a quite different kind of collectively, with more or less correspond to the common intellectual interests and instructional tasks of a group of academics."*

This shows that in the diversity there is also unity in the institutional form of disciplines.

Disciplines provide a structural framework for institutions; on the other hand, the disciplinary framework within universities is highly influenced by institutions' decisions and the environment of Higher Education Institutions. It is noted that the UK Higher Education Institutions' environment has high levels of competition, especially, in the context of government spending cuts, the modifications of the funding structures for teaching and research and the rising cost of research (Becher & Trowler, 2001; Trowler et al., 2012). These changes have significant implications for disciplines and academic communities. As institutions need to compete for resources, funding and students etc, there are increasing pressures for institutions to adopt a "business strategy" in order to be more competitive in both education and research market. As a result, institutions shape their map of knowledge and organise the disciplinary structures according to market demands (Krishnan, 2009; Becher & Trowler, 2001). Universities see students, employers and governments as the key customers (Becher & Trowler, 2001). They try to increase reputation and incomes by marketising knowledge. For example, adding new attractive courses and programmes for their customers. In addition, some institutions close programmes or departments that are not successful, some combine departments to one department or research centre (Krishnan, 2009). By combining departments

into new interdisciplinary departments or research centres, cost can be reduced. As Davis (2007:49) states “You could get rid of that spindly comparative- literature department by combining it cleverly and “interdisciplinarily’ with the heftier English department, and then you’d have to pay only one secretary instead of two”. Disciplines are not only affected by institutions’ decisions but also affected by regulatory frameworks. As Trowler et al (2012) point out, the way that institutions organise their disciplinary departmental structures are usually associated with the Research Assessment Exercise (RAE), Research Excellence Framework (REF) and the subject centres of Higher Education Academy. Becher and Trowler (2001), Spruling (2012) and Trowler et al (2012) suggests that these homogeneity structures for assessing the quality of research and education have some effects on research practices in diverse disciplines.

There is a substantial amount of literature that has attempted to define what academic disciplines are. It has been suggested that various approaches can be used to understand the notion of disciplinarity, such as through the intellectual character, the social and cultural aspect and organisational aspect of a discipline. Disciplinary epistemology and disciplinary culture within knowledge communities is mutually dependent. Through the process of institutionalisation, disciplines become “visible”. However, the disciplinary structural form and disciplinary practices are always conditioned by institutions’ decisions and other regulatory frameworks such as REF. Trowler (2012: 9) concludes that disciplines are

*“reservoirs of knowledge resources shaping regularised behavioural practices, set of discourses, ways of thinking, procedures, emotional responses and motivations. These provide structured dispositions for disciplinary practitioners who reshape them in different practice clusters into localised repertoires. While alternative recurrent practices may be in competition within a single discipline, there is common background knowledge about key figures, conflicts and achievements. Disciplines take organisational form, have internal hierarchies and bestow power differentially, conferring advantage and disadvantage.”*

## 2.4 Classification schemes for academic disciplines

Different conceptual schemes have been developed for classifying disciplines and explaining the differences between them. Braxton & Hargens (1996) conducted an extensive review of theoretical frameworks for understanding disciplinary differences. Jones (2011) conducted a similar review, but focusing on classification frameworks that had been introduced after 1996. There are more than eleven classification schemes that have been identified (Jones, 2011), including earlier works such as, Conant’s (1951) distinction between theoretical fields and empirical fields; Kuhn’s (1962)



differentiation of groupings: paradigmatic vs. pre-paradigmatic fields and Pantin's (1968) study based on restricted vs. unrestricted fields.

Becher and Trowler (2001) particularly highlight the contribution of Pantin's (1968) and Kuhn's (1962) work to the understanding of forms of knowledge and knowledge communities. Pantin's classification focuses on the form of knowledge. Pantin (1968: cited in Becher & Trowler, 2001: 32) explained

*“there is one real, and graded, distinction between sciences like the biologies and the physical sciences. The former are unrestricted sciences and their investigator must be prepared to follow their problems into any other science whatsoever. The physical sciences, as they are understood, are restricted in the field of phenomena to which they are devoted. They do not require the investigator to traverse all other sciences. But while this restriction is the basis of their success, because of the introduction of this restricted simplicity of their field we cannot necessarily take them or their methods as typical of all the sciences.”*

Restricted fields are considered to have narrow, specific theoretical structures and clearer domain boundaries; they share the properties of Becher's (1989) description of hard pure knowledge. In contrast, unrestricted fields, such as social sciences, have more open-ended knowledge structures and boundaries that are more permeable; they share the properties of soft pure knowledge (Becher & Trowler, 2001).

In contrast, Kuhn's disciplinary groupings are about the levels of paradigm development. The notion of paradigm has been defined variously (Becher & Trowler, 2001). It is noted that paradigmatic development commonly relates to “the degree of consensus or sharing of beliefs within a scientific field about theory, methodology, techniques, and problems” (Kuhn, 1970; Lodahl and Gordon, 1972: 58; Biglan, 1973; Braxton & Hargens, 1996). Becher and Trowler (2001) suggest that Kuhn's groupings show a social aspect, as they are primarily defined in relation to the communities of knowledge. Fields with a high level of paradigm development (e.g. physics, chemistry) share paradigms. Researchers within these fields are mutually dependent, have a high degree of consensus on research topics and methodology and have clear ways to investigate the field of knowledge. In contrast, fields with a low level of paradigm development (e.g. non-sciences and social sciences) do not share paradigms, have less agreement about appropriate research problems to be investigated and methods to be used. They work more separately (Becher & Trowler, 2001).

### 2.4.1 Becher's classification

Broad disciplinary groupings, such as, hard pure, hard applied, soft pure, soft applied have been commonly used for investigating the differences among academic disciplines. This typology is derived by Becher (1989) from the earlier study of Biglan (1973) and Kolb (1981). Disciplines are divided into a four cell matrix. Hard disciplines are high paradigmatic. They have well-developed theory and universal law and generalisable results. In contrast, soft disciplines are low paradigmatic. They have less theoretical structure, and problems are loosely defined (Trowler et al, 2012). Pure disciplines are predominantly theoretical and self regulating. Applied disciplines are more applied to the professions and are regulated or influenced by external organizations.

Becher and Trowler (2001) argued that not only cognitive characteristics but also social and cultural characteristics of a discipline, such as, knowledge communities and networks, are significant to the differences among disciplines. Becher and Trowler (2001) developed a different set of dimensions when looking at communities of researchers, i.e. convergent versus divergent, urban versus rural. Convergent disciplinary communities have uniform standards and procedures and a stable elite. In contrast, divergent communities allow more intellectual deviance and experience of standards shifting. The Rural /urban dimension concerns the nature and the scale of problems and also communication patterns. Thus urban disciplines have a limited number of salient topics, high people to problem ratio, more collaboration and intense interaction. They tend to solve problems by breaking them down into smaller problems and solve them with quick, short-range solutions (e.g. publish short journal articles). In contrast, rural disciplines have a broader intellectual territory, more diverse research problems, with a smaller people to problem ratio and less interaction. Rural disciplines favour a holistic approach, solutions to problems take longer to solve, thus they have longer forms of research output, such as books and monograph. In general, by combining both cognitive and social characteristics, History can be described as soft, pure, divergent and rural. Biology can be described as hard, pure, convergent and urban.

### 2.4.2 Whitley's classification

Whitley (2000) developed another classification scheme for differentiating disciplines. He based his classification on the concepts of the degree of task uncertainty and the degree of mutual dependence (Whitley, 2000). Task uncertainty refers to the degree of predictability of research processes and task outcomes. Whitley (2000) differentiated two types of task uncertainty, i.e. technical task uncertainty and strategic task uncertainty. Technical task uncertainty relates to research techniques and methods. For example, fields with low technical task uncertainty, such as physics and chemistry, have a well-recognised set of methods or techniques, therefore results will be more "predictable, visible and replicable" (Whitley, 2000: 121). On the other hand, fields with high

technical task uncertainty, such as sociology, have a wide variety of methods to be used and different interpretations. Whitley (2000) argued that the question of what methods should be used and how can be highly personal in high technical task uncertainly fields. Thus the results are more likely to be uncertain and unpredictable.

The other type of task uncertainty, strategic task uncertainty, relates to the variability and stability of intellectual priorities, research topics and goals of research (Whitley, 2000). Fields with low strategic task uncertainty level, such as economics and physics, have restricted varieties of research problems (Whitley, 2000). As the importance of problems and intellectual priorities will be generally agreed by researchers and practitioners, their research and goals are relatively clearly ordered. Thus research strategies and activities are relatively more stable and uniform. In contrast, fields with high strategic task uncertainty, such as artificial intelligence, engineering and sociology, have varied research goals and problems. The importance and the value of those problems can be influenced not only by researchers, but also by others, such as the general public. Thus, the variability of research topics and priorities can be great.

Mutual dependence, Whitley's (2000) second principle of classification, refers to the degree of researchers' dependence upon colleagues in order to make contribution, to achieve research goals, status and access to resources. Whitley (2000) differentiated between two types of dependence: functional and strategic dependence. Functional dependence refers to the extent to which researchers need to depend on colleagues' research results and ideas in order to build knowledge. In addition, researchers need to demonstrate their results are useful and have influenced other's research in order to obtain higher reputation in the field. Thus, fields with a high degree of functional dependence, e.g. chemistry and mathematics, have extensive co-ordination of research topics and outcomes (Whitley, 2000). On the other hand, strategic dependence refers to the extent to which researchers have to convince their colleagues of their research problems and strategies are significant and fit in to their collective goals. As Whitley (2000: 89) stated co-ordination is

*“not just a technical matter of integrating specialist contributions to common goals.... It is a political activity which sets research agenda, determines the allocation of resources, and affects careers in reputational organisations and employment organisations.”*

Fields with a high degree of strategic dependence are, therefore, strongly bounded. They recognise the need of co-ordinating research strategies to achieve collective goals.

By combining the two dimensions, task uncertainty (technical and strategic) and mutual dependence (functional and strategic), sixteen possible types of fields are generated. However, Whitley (2000)

explained that only seven types of fields are stable and likely to create knowledge. Each type of field combines different degree of task uncertainty and of functional dependence (Table 2.1). The scheme can not only distinguish between fields, as mathematics and chemistry, but also between fields in different period and fields in different countries with various traditions (Whitley, 2000). As can be seen from Table 2.1, there is a close relationship between technical task uncertainty and functional dependence. It is an inverse relationship, i.e. a high degree of technical task uncertainty is associated with a low degree of functional dependence and vice versa. High degree of technical task uncertainty implies there is a low degree of standardised methods and techniques, which restrict the possibility that research results can be compared or evaluated across research centres. Research results are not easily compared and integrated. Thus, researchers tend to work independently. They are not necessarily dependent on colleagues' research results in order to build knowledge. On the other hand, a low degree of technical task uncertainty implies that methods and analytical techniques are standardised. As research results would be able to be compared and co-ordinated, collaboration becomes more significant. Thus, researchers depend on others' research results and ideas to build knowledge.

16 possible combinations of degree of mutual dependence and degree of task uncertainty.  Shaded boxes indicate 'unlikely' combinations:				Degree of Functional Dependence			
				Low		High	
				Degree of Strategic Dependence		Degree of Strategic Dependence	
				Low	High	Low	High
Degree of Technical Task Uncertainty	Low	Degree of Strategic Task Uncertainty	Low	1	2	3 Technically integrated bureaucracy	4 Conceptually integrated bureaucracy
			High	5	6	7 Professional adhocracy	8 Polycentric profession
	High	Degree of Strategic Task Uncertainty	Low	9	10 Partitioned bureaucracy	11	12
			High	13 Fragmented adhocracy	14 Polycentric oligarchy	15	16

Table 2.1 Whitley's classification scheme (adapted from Whitley, 2000)

Whitley's scheme identifies seven types of fields, as elaborated in the following sections.

### 1. Fragmented adhocracies

Fragmented adhocracies feature high technical task uncertainty, high strategic task uncertainty levels, with low degrees of functional and strategic dependence.

The main feature of these fields is their intellectual diversity and fluidity. These fields have a variety of research problems and methods. Researchers are encouraged to pursue a variety of goals with different approaches. Thus these fields usually do not have a stable configuration of specialised problem areas. Research in these fields, e.g. Management studies, political studies, English literature, is divergent and relatively speaking highly personal, as researchers do not need to depend on the research results of a particular group of researchers or make specific contributions to fit in collective research goals. In general, they have little need for internal co-ordination of research strategies and outcomes.

Audiences of these fields are also varied. As boundaries are relatively permeable and varied, audiences are not limited to a small group of specialists in institutions. For example, in management studies, there are not only researchers from university, but also people from consultancy firms and large organisations. While in sociology, researchers can not only build a reputation through publishing articles for small groups of specialists, but can also publish books for the general public and students.

### 2. Polycentric oligarchies

Polycentric oligarchies are similar to fragmented adhocracy, and are characterised by high technical, high strategic task uncertainty, low degrees of functional dependence but with high degree of strategic dependence. This means researchers depend more on other members of the fields for reputations. And theoretical co-ordination in these fields becomes more significant. Polycentric oligarchies, e.g. British social anthropology, German philosophy, are often organised by a small group of intellectual leaders or universities who control the production and certification of research skills. Different research schools and centres have better co-ordination of research strategies to achieve collective goals. However, polycentric oligarchies still have high levels of uncertainty of results and interpretations. This means the degree of co-ordination of research across research centres is restricted by the low degree of standardised research skills and interpretative procedures. Thus, research results of distinct research centres cannot be easily compared and evaluated.

In addition, polycentric oligarchies feature common sense intellectual concepts which are similar to fragmented adhocracies, but polycentric oligarchies use more technical terms and theoretical structures to describe general concepts.

### 3. Partitioned bureaucracies

Partitioned bureaucracies combine high technical task uncertainty, low strategic task uncertainty, low functional dependence and high strategic dependence. Partitioned bureaucracies, for example, economics and finance business, partition the central theoretical work away from empirical applications. Analytical and theoretical work is valued higher than empirical exploration in these fields. Empirical research only develops as sub-units around the central theoretical core. It is because the control of empirical exploration in these fields is limited. For example, economics feature theoretical uniformity and technical uncertainty. Empirical findings in economics are not always consistent with market hypothesis and economic theories. Yet research skills, training programmes and analytical procedures in these fields are highly standardised, which enable the uniformity and stability of research strategies and problems. In general, central theoretical work dominates the control of intellectual boundaries.

### 4. Professional adhocracies

Professional adhocracies, e.g. bio-medical science, artificial intelligence, combine low technical task uncertainty, high strategic task uncertainty, low strategic dependence and high functional dependence. Different from the previous three types, a professional adhocracy has low technical task uncertainty. It has extensive standardisation of technical skills and procedures which enable co-ordination and control of task outcomes. Thus, research outcomes can be easily compared and evaluated. However, research problems and goals are diverse. Theoretical co-ordination is not very important in these fields. No single group dominates the development of research strategies. For example, in bio-medical sciences, research goals and problems of universities and non-academic organisations are often not identical. Knowledge in these fields is highly specialised and empirically focussed with varied problems.

### 5. Polycentric professions

Polycentric professions, e.g. experimental physiology, continental mathematics, have low technical task uncertainty, high strategic task uncertainty, high strategic dependence and high functional dependence.

Polycentric professions feature segmentation of research problems which centre on a few dominant research schools. They have relatively standardised skills and procedures, results can be compared and evaluated across schools. High strategic dependence implies there is a greater degree of competition between different research schools. Theoretical co-ordination and research strategies in these fields are emphasised.

## 6. Technologically integrated bureaucracies

Technologically integrated bureaucracies combine low technical task uncertainty, low strategic task uncertainty, with high functional dependence and low strategic dependence. These fields, such as modern chemistry, have standardized methods and tools. Low degree of task uncertainty means research outcomes are relatively predictable. Thus, research can be organised and co-ordinated systemically. Extensive division of tasks and skills becomes possible in technologically integrated bureaucracies. Research can be carried out across research centres and countries. Low degree of strategic dependence means theoretical co-ordination of problems is not very important.

Researchers and practitioners study different topics and use different analytical methods in order to distinguish themselves from others. Generally, knowledge in these fields is highly specialised and distinguished.

## 7. Conceptually integrated bureaucracies

Conceptually integrated bureaucracies, e.g. physics, combine low technical task uncertainty and low strategic task uncertainty, with high functional dependence and high strategic dependence. A conceptually integrated bureaucracy and a technologically integrated bureaucracy are very similar. They both have standardized methods and tools, formal standardised reporting system and stable theoretical framework. However, a high degree of strategic dependence in conceptually integrated bureaucracies means these fields are more competitive and strongly bounded. Problem areas and sub-fields are not as separate as in technologically integrated bureaucracies. As theoretical co-ordination is highly valued in these fields, researchers need to demonstrate the importance of their work and their research can fit into collective goals to their colleagues. As Whitley (2000: 203) stated, "the search for unity and coherence is a major characteristic" of conceptually integrated bureaucracies, such as physics.

### 2.4.3 Limitations of classification schemes

As outlined above, there are many different classification schemes for distinguishing differences between disciplines. Different conceptual schemes try to explain disciplinary differences from different perspectives. For example, from an epistemic perspective, such as, high / low paradigmatic fields (Kuhn, 1970), restricted / unrestricted fields (Pantin, 1968), hard/pure, soft/applied (Becher & Trowler, 2001); or from a social perspective, such as, convergent/ divergent, urban/ rural (Becher & Trowler, 2001). Each classification of disciplines has their own advantages and limitations. For example, Whitley's scheme can identify more subtle distinctions between disciplines than single dimensional schemes, such as Kuhn's and Pantin's. Becher and Trowler (2001) argue, however, that Whitley's scheme can be too complex and can become un-functional.

There is one common limitation of all these classification schemes for disciplines. Both Whitley (2000) and Becher and Trowler (2001) have pointed out that the weakness of their own schemes is that no discipline can perfectly fit into one single category. The nature of disciplines is very complex and dynamic. Disciplinary boundaries and the divisions of disciplines are not fixed. As Ostreng (2007) has stated, disciplinary boundaries can be illusory. For example, social sciences and humanities disciplines usually are classified as “soft” knowledge in Becher’s scheme. However, one may argue that geography can be classified as “soft” or “hard” knowledge. As geography can be divided into two core branches, physical geography and human geography, which have very different research methods and concepts between them. Human geography needs to be placed on the “soft” side while physical geography needs to be placed on the “hard” side. Becher and Trowler (2001) pointed out that broad disciplinary groupings sometimes can be a troublesome unit of analysis. Nevertheless, although disciplinary classification schemes “cannot do justice to the complexity of inquiry processes and knowledge structures in various disciplines”, they provide “useful dimensions for describing disciplinary variations” (Kolb, 1981:244 cited in Becher & Trowler, 2001:39).

## 2.5 Sub-fields and specialisms

Most modern disciplines consist of a wide range of sub-fields which can be further divided into specialisms. Generally, specialisms can be divided into theory-based, methods-based and subject matter specialisms (Becher & Trowler, 2001). Specialisms are not easy to identify. They usually are seen as ‘networks’, ‘social circles’ or ‘invisible colleges’ (Becher and Trowler, 2001). Becher and Trowler (2001) stated that specialisms are less officially recognisable than disciplines, which can be identified in professional groupings, journals or take organisational shape in university departments.

Becher and Trowler (2001) pointed out that there are two main features when disciplines are looked at closely. Firstly, is complexity. Secondly, there is little consistency and stability. Disciplines look fragmented under the microscope (Becher and Trowler, 2001). It is difficult to find a single set of features to characterise one particular discipline, especially at a detailed level of analysis.

While Becher and Trowler (2001) suggested that the complexity of disciplines and sub-fields like biological cells continuously subdivide and recombine. Abbott (2001) suggested disciplines have a fractal pattern of division, i.e. disciplines and sub-fields develop through a pattern of split, conflict and ingestion. Abbott (2001) argues that “fractal distinctions”, such as the object of study, methods within a discipline, repeats a pattern as geometric fractal within themselves (Figure 2.1).



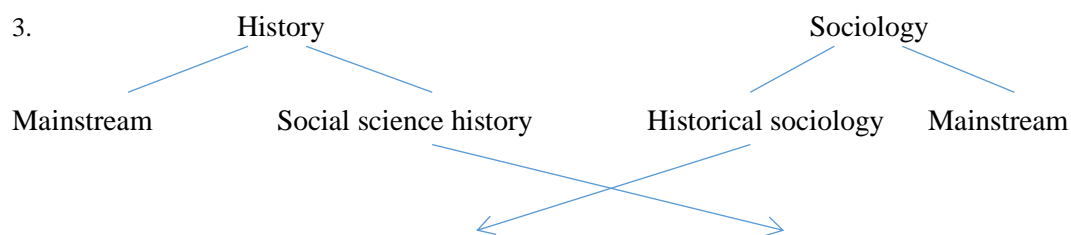
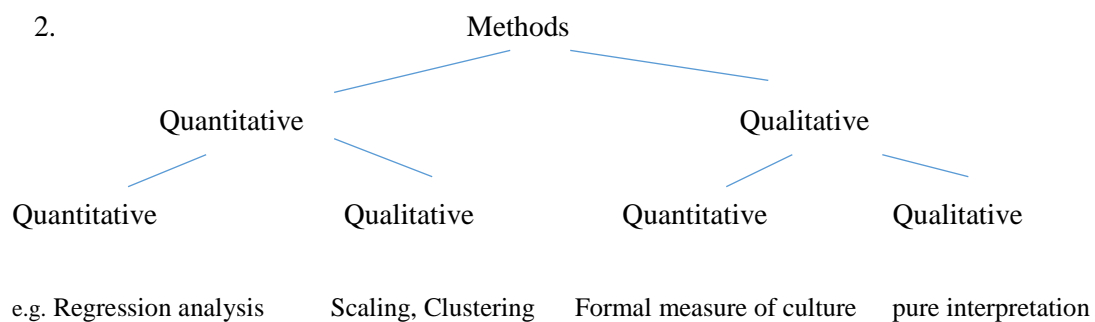
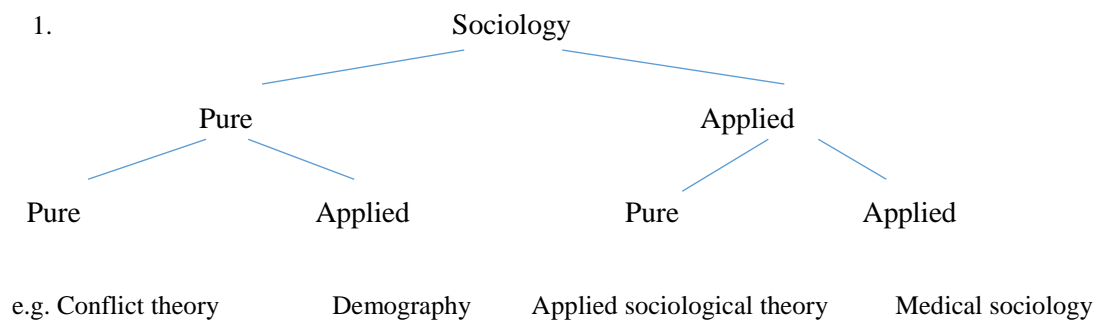


Figure 2.1 Fractal patterns (adapted from Abbott 2001)

These patterns illustrate how sub-fields and methods within a discipline are organised and develop. For example, generally, there are two broad methodological aspects in sociology, quantitative and qualitative. Each aspect can be further divided into quantitative and qualitative. On the quantitative side, there are statistical methods e.g. regression analysis on one hand and more descriptive methods such as scaling technique and cluster analysis on the other. While on the qualitative side, there are some formal models and procedures to measure culture contrast with pure interpretative approaches.

Abbott (2001) also used the patterns to show how a discipline can be split into specialisms. For instance, in Figure 2.1 the second example, sociology is generally divided into “pure” sociology and “applied” sociology. Within “pure” sociology, there is not only theoretical research such as, conflict theory and structural functionalism, but also empirical research such as stratification and

demography. While within “applied” sociology there is also a more theoretical perspective, e.g. applied sociological theory in contrast with real application in medical sociology. Thus it shows that it is quite common that a discipline contains sub-fields and specialisms with opposing properties, such as hard versus soft, pure versus applied (Klein, 1993; Becher & Trowler, 2001; Abbott, 2001).

Abbott (2001) also uses the pattern to show the relationship between two sub-fields from two different disciplines, history and sociology (Figure 2.1). Abbott (2001) argues that it is possible for social science history and historical sociology to have more similarities with each other than with their parent disciplines. It is noted that there are intellectual interests that overlap between specialisms within neighbouring disciplines.

## 2.6 Interdisciplinarity

Interdisciplinary research has been increasing in the last three decades. In 1972, the Organisation for Economic Co-operation and Development (OECD, 1972) published a volume *Interdisciplinarity* to promote interdisciplinarity in research and teaching in Higher Education Institutions. As knowledge advances, there is an increasing recognition of the need to link different disciplines together in order to solve problems outside traditional disciplines (Aboelela et al, 2007). It has been argued that a new mode of knowledge production has evolved, which has been referred to as Mode 2 (Gibbons et al, 1994). Gibbons et al (1994) used the term Mode 1 to describe the traditional form of knowledge which comes from academic disciplines, primarily located in institutions; Mode 2 refers to knowledge production by solving real world problems. Mode 2 knowledge is not limited to universities. It is about interdisciplinarity, crossing boundaries and innovation. It has close interaction with scientific, technological and industrial form of knowledge, it brings different expertise to solve problems (Klein, 2000). Mode 2 knowledge has been favoured by research funders (Frodeman et al, 2012).

Although the term interdisciplinarity is frequently used, it is often loosely used and defined. Krishnan (2009) stated that terms such as multidisciplinary, crossdisciplinary, interdisciplinary and transdisciplinary which have different meanings have been used interchangeably. Aboelela et al (2007) reviewed fourteen definitions of interdisciplinarity and concluded that multidisciplinary, interdisciplinary and transdisciplinary all are associated with two or more disciplines (Table 2.2). However, there are differences between them. A multidisciplinary team includes researchers from different disciplines. They work on the same question but without much interaction. They do not usually share publication. Interdisciplinary researchers interact with each other. They combine

different methodologies to solve problems together. Transdisciplinary researchers fully synthesize theories and methods; develop some new understanding and language to solve complex problems.

	<b>Participants/ Discipline</b>	<b>Problem Definition</b>	<b>Research Style</b>	<b>Presentation of Findings</b>	<b>Examples from Infectious Disease Literature</b>
<b>Multidisciplinary</b>	<i>Two or more disciplines</i>	<i>Same question but different paradigm OR different but related questions</i>	<i>“Parallel play”</i>	<i>Separate publications by participants from each discipline</i>	<i>Medicaid cost containment and access to prescription drugs, <a href="#">Cunningham (2005)</a>, <a href="#">Lichtenberg (2005)</a>: The effect of access restrictions on the vintage of drugs used by Medicaid enrollees</i>
<b>Interdisciplinary</b>	<i>Two or more distinct academic fields</i>	<i>Described/defined in language of at least two fields, using multiple models or intersecting models</i>	<i>Drawn from more than one, with multiple data sources and varying analysis of same data</i>	<i>Shared publications, with language intelligible to all involved fields</i>	<i>The “Minimizing Antibiotic Resistance in Colorado” Project: Impact of patient education in improving antibiotic use in private office practices, <a href="#">Gonzales et al. (2005)</a></i>
<b>Transdisciplinary</b>	<i>Two or more distinct academic fields</i>	<i>Stated in new language or theory that is broader than any one discipline</i>	<i>Fully synthesized methods, may result in new field</i>	<i>Shared publications, probably using at least some new language developed for translation across traditional lines</i>	<i>Assessing the implementation of the Chronic Care Model in quality improvement collaboratives, <a href="#">Pearson et al. (2005)</a></i>

Table 2.2 Characteristics of Multidisciplinary, Interdisciplinary, and Transdisciplinary Research (Verbatim from Aboelela et al, 2007: 340)

The first part of the chapter established some core understandings of the notion of academic disciplines and disciplinary theories. In general, academic disciplines can be understood from three perspectives: intellectual, social and organisational. The section looked at some key theoretical frameworks for understanding disciplinary differences, such as, Becher’s (1989), Whitley’s (2000) and Abott’s (2001). Among the frameworks reviewed, Becher’s framework is the best framework to be the conceptual basis for this research. The framework not only can be applied at broad disciplinary levels but also sub-field levels. Thus, it is suitable for this research, particularly this study focuses on one single discipline, i.e. geography. In addition, Becher’s framework is not too complex to be operational in practice, e.g. compared to Whitley’s framework. This research will look at the applicability of Becher’s disciplinary theories to explain data practices within geography.

## Part 2 Research data

The second part of this chapter focuses on research data. Managing research data has become increasingly important. As mentioned in the first chapter, research data management is high on the agenda of research policy makers, national funding agencies and institutions. Thus, this section examines what research data management is and the concept of research data through literature. It also looks at some research data management surveys which have been conducted by UK institutions, which shed some light on data sharing practices of researchers and disciplinary patterns.

### 2.7 Research data management and digital curation

In literature, a variety of terms were found, such as, research data management (RDM), data curation, digital curation and data preservation, which are associated with managing research data. Research data management, data curation, digital curation and data preservation have similar meanings. The term “research data management” has become widely used by universities in the UK, as it has become an important national agenda with its emphasis on good research practices in HEIs. Interestingly, a detailed definition of research data management is lacking in the literature. Nevertheless, according to Whyte & Tedds (2011), research data management refers to the management and organisation of data throughout the entire research cycle. Research data management includes processes such as creating, documenting, maintaining, archiving, preserving and sharing data (EDINA & University of Edinburgh, 2013). Effective research data management is an active and continuing process to ensure data can remain discoverable, accessible and usable over the long term (Pryor, 2012).

Another commonly used term associated with managing research data is data curation/ digital curation. Comparing to ‘research data management’, data curation/ digital curation has been defined in more detail. A number of definitions were found. For example, the Digital Curation Centre state that

*“digital curation is concerned with actively managing data for as long as it continues to be of scholarly, scientific, research and/or administrative interest, with the aim of supporting reproducibility of results, reuse of and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its long-term accessibility and preservation, authenticity and integrity.” (DCC, 2008:3).*

While the University of Illinois' Graduate School of Library and Information Science (2012) has also defined data curation:

*“data curation is the active and ongoing management of data through its life cycle of interest and usefulness to scholarship, science, and education. Data curation activities enable data discovery and retrieval, maintain its quality, add value, and provide for reuse over time, and this new field includes authentication, archiving, management, preservation, retrieval, and representation”.*

Both definitions have not only illustrated the aims of data curation, but also described the advantages of data curation activities, e.g. enabling data reuse, supporting reproducibility in research results.

It appears that there are no clear differences between data curation, digital curation and research data management. All of these terms are concerned with the active management of research data throughout the research cycle and ensuring data can be accessible and useable over the long term. Data curation was used in early reports of e-science in the UK, for example, Lord & MacDonald (2003), whereas the term research data management is relatively new. As Lewis (2010) commented, research data management is still at its developing stage, terminologies are still evolving. Nevertheless, research data management has become a more widely used term across Higher Education Institutions in the UK. One of the possible reasons is data curation gives a stronger impression of an archival and curatorial perspective. For instance, 'curation' has connotations of the activities of archives or museums. However, managing research data is not solely the curators' responsibilities. Data creators, i.e. researchers, are also responsible for managing their own data (Pryor & Donnelly, 2009; Lyon, 2007). It is possible that universities would like to increase researchers' awareness of data management and so use a more engaging term. Freiman et al (2010) reported that some researchers do not understand the term digital curation. It is suggested that specialised language should be avoided in data management resources, and clear, simple and jargon free language should be used in order to communicate with researchers more effectively (Freiman et al, 2010). As this research is focused on the academic research community and researchers' data practices, research data management is used as the preferred term throughout this thesis.

## 2.8 Research and research data management lifecycle

A lifecycle approach has been commonly used for research data management. Research processes and data are closely related. Models of research process and data management could be used in this research as the basis for the understanding of differences between research data practices in

different disciplines/sub-fields. This section, therefore, looks at the lifecycle approach to the management of research data, some of the available models and concludes with a longer description of what is probably the most comprehensive and widely used model in the field in the UK, namely the DCC curation life cycle which provide a deeper understanding of research data management.

### 2.8.1 The research lifecycle

Research data management is about managing and organising data throughout the research cycle, from its entry to the research cycle through to publishing and beyond (Ingram, 2016; Whyte & Tedds, 2011). Pryor (2012) illustrated a research lifecycle with six phases (Figure 2.2), i.e. hypothesise, research, interpret, synthesise, publish and reuse.

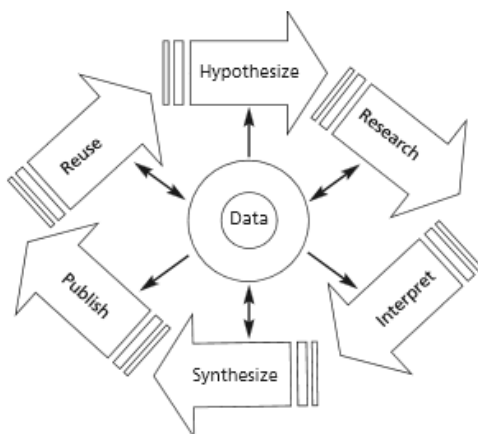


Figure 2.2 The six data centric phases of the research cycle (Pryor, 2012:6)

In each phase, researchers are involved in some aspects of data activities directly or indirectly. In phase 1, researchers gather and review existing data to design research and structure research questions and hypotheses. During phase 2, new raw data are collected and generated during the research phase. In phase 3, researchers process, analyse and interpret the data. In phase 4, the data are synthesized, e.g. using modelling or statistical methods. In phase 5, a new set of data is produced. Finally, in phase 6, researchers disseminate and publish the data which can be reused for the foundation of new research. As Pryor (2012) state the lifecycle is 'datacentric', it represents researchers as data creators and data users, and it illustrates how data can be related to research and researchers in each stage of the research process. However, one of the limitations of the 'datacentric' model is it fails to present researchers as data managers (Pryor, 2012). Nevertheless, lifecycle models which are focused on data management have been developed.

## 2.8.2 Data management lifecycle approach

A lifecycle approach has been widely adopted to managing data in research data management. It is noted that many institutions have adapted the lifecycle approach to the management of research data, e.g. University of Leicester (2013) and the University of Oxford (Uribe, 2008). Institutions have used the lifecycle model to develop resources and services for researchers, such as data management planning checklists, to ensure all the essential steps in the lifecycle are covered (University of Leicester, 2012).

As Ball (2012:3) states, one of the advantages of using the research data management lifecycle model is it provides a useful framework for considering different data related activities that need to be performed on data throughout its lifecycle. Pennock (2007) emphasises a life-cycle approach to managing digital materials is necessary, pointing out three reasons for the necessity of adopting the lifecycle approach. Firstly, digital data are very fragile, Michener et al. (1997) have shown the normal degradation of data and metadata over time (Figure 2.3). Without managing research data throughout its lifecycle, the content of the data can be quickly lost after the time of publication.

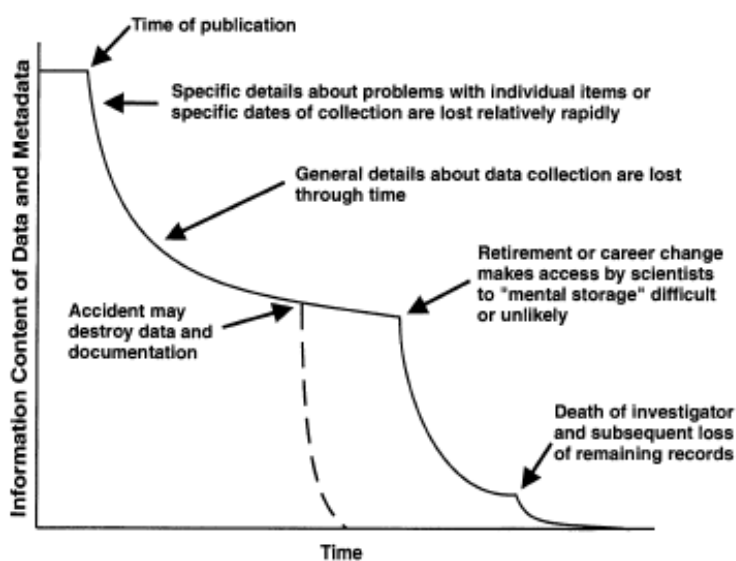


Figure 2.3 The degradation of information content associated with data and metadata (Graphic from Michener et al., 1997: 332)

In addition, data are vulnerable to change from technological obsolescence throughout the lifecycle (Lord et al, 2004; DCC 2013). Technological advances can put the integrity of data at risk. For example, data stored in a wide range of media formats, software and hardware can become obsolete in a few years. Secondly, activities at all stages in the lifecycle are vital. All activities or lack of activities, in one stage can greatly influence the capability to preserve the data in later stages (Pennock, 2007). For example, if metadata is not added at the relevant stages in the lifecycle, the

data may not be able to be identified and re-used in later stages. Thirdly, Pennock (2007) believes that the lifecycle approach is the only possible way to ensure the authenticity and integrity of data can be retained which enable reliable reuse.

### 2.8.3 DCC Curation Lifecycle Model

Various data management lifecycle models have been reported in the literature, e.g. DCC curation lifecycle model (Figure 2.4), UK Data Archive data lifecycle (2013), Data Documentation Initiative conceptual model (2004) and DataONE Data lifecycle (2003). All models show the important stages for effective research data management.

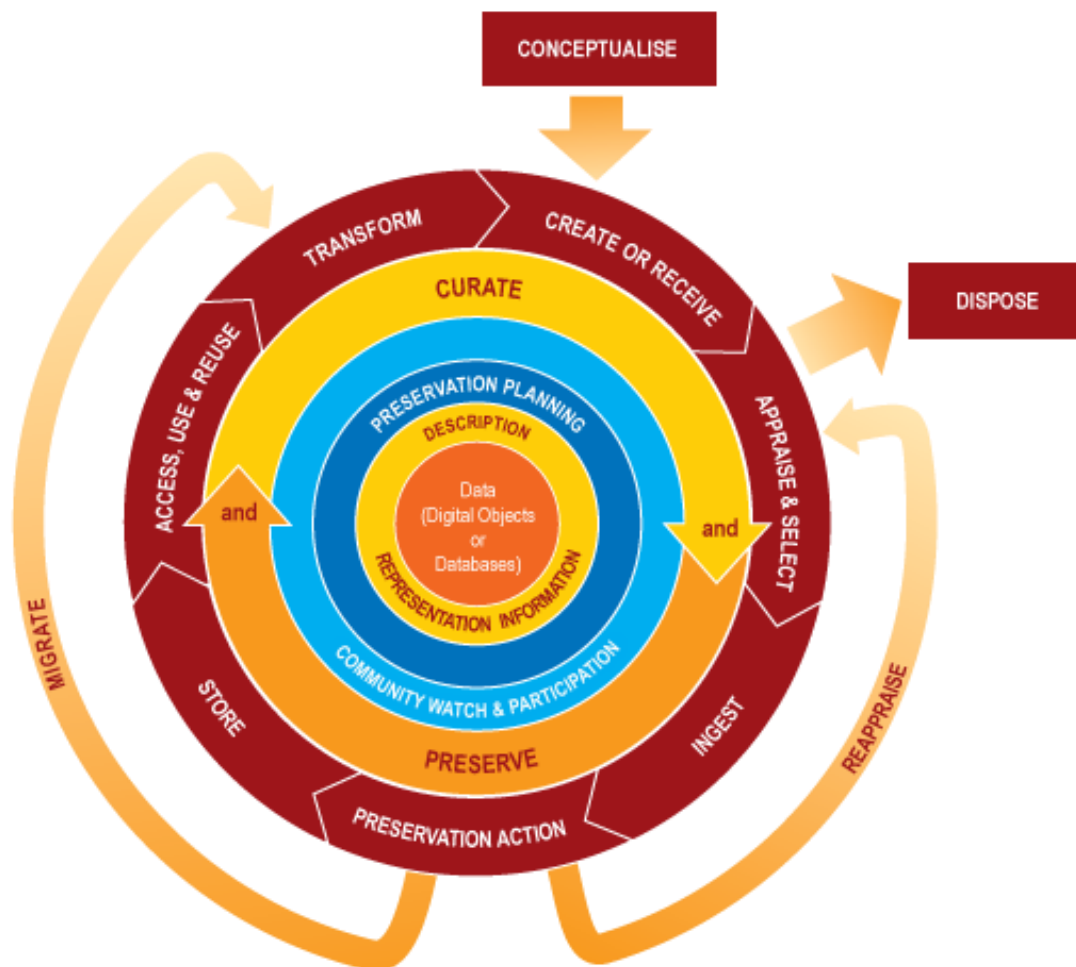


Figure 2.4 DCC Curation Lifecycle Model (DCC, 2013)

Compared to other models, the DCC curation model provides more details of the essential activities in research data management. For example, the DCC lifecycle differentiates three types of actions for effective RDM, i.e. full lifecycle actions, sequential actions and occasional actions. The DCC life cycle model is probably the most comprehensive and provides a suitable framework for this research. The DCC lifecycle is an organisational tool not only for data curators but also for data



creators and users (Higgins, 2008). As Digital Curation Centre (2013) stated, it can be used for planning activities within institutions, ensuring all necessary stages are performed, and in correct sequence.

### Data

At the centre of the model is digital data, which includes databases and digital objects (such as text files, image files, sound files, websites along with metadata Simulation data)

### Full lifecycle actions

Full lifecycle actions include 'Description and Representation Information', 'Preservation Planning', 'Community Watch and Participation' and 'Curate and Preserve'. These actions are at the inner rings of the model. The Digital Curation Centre (2013) suggest that all these four actions should be considered at all stages throughout the data management lifecycle.

- 'Description and Representation Information' concerns activities such as assigning sufficient metadata and representation information and using appropriate standards. These are essential to retain the value of the data over the long term.
- 'Preservation Planning' is required throughout the data management lifecycle, for instance, produce plans for procedures and policies, revise strategic plans of different stages.
- 'Community Watch and Participation' are also essential. Stakeholders, such as researchers, technical staff, need to maintain a watch on in community activities in order to keep up to date with developments of shared standards, tools and software.
- In addition, 'Curate and Preserve' are needed throughout the lifecycle, i.e. execute data management actions, and promote curation and preservation.

### Sequential actions

Sequential actions are presented in the outermost ring of the model. The model shows the stages at which actions need to be undertaken. It begins with 'Conceptualise' which is the planning stage of research data management, for example, activities such as: addressing practical issues, identifying roles and responsibilities and considering storage options for data. It is noted that the planning stage should begin even before the data is created (Higgins, 2012). Then the lifecycle begins properly with 'Create or Receive', i.e. when research data are created or pre-existing research data are received, for example, secondary data from data centres or data repositories. Adequate administrative and descriptive metadata should be assigned at this stage. The next stage is "Appraise and Select". This is when the data are evaluated and selected according to legal, funder or institutions' requirements. Then the data are transferred to an archive, repository or data centre, i.e. the 'Ingest' stage. The

next stage is 'Preservation Action' which includes data cleaning, validation, assigning preservation metadata, and ensuring standard data structures or file formats. The purpose of these actions is to ensure data remains authentic, reliable and usable while maintaining its integrity and the authoritative nature of data. After that, the data are stored securely, adhering to relevant standards. Actions such as backing up copies, maintaining storage hardware can be associated with this "store" stage (Ball, 2012). The 'Access, Use and Reuse' is the following stage. At this stage, data are accessible to users and reusers. Actions such as, making information available publicly, implementing access controls and authentication procedures can be included in this stage. The final stage of the cycle is 'Transform', i.e. new data are created from the original data, for instance, by migration into a new format, by selection, by creating a subset of the data. So that the new data begins its lifecycle again.

#### Occasional actions

The three occasional actions are 'dispose', 'reappraise' and 'migrate'. They are key actions for long-term management of data. Data that has not been selected for long-term curation will be disposed according to policies or legal requirements. It is possible that data may be transferred to other data repository. Reappraisal is also essential. As Higgins (2012) states, in order to ensure the data are relevant to collections, it needs to be reappraised over time. Finally, in order to prevent software obsolescence, migration of data is also necessary, for example, migrate data to a different format.

This section has described what is regarded as good research data management practices using the DCC lifecycle model. Generally, the lifecycle approach helps institutions and researchers to plan data related activities to ensure all necessary actions are carried out for effective research data management. However, in reality, complete control of the data lifecycle is always not possible (Pennock, 2007). Pennock (2007) pointed out research and research data lifecycle are generic frameworks. They are idealised situations. It is noted that, at a deeper level, research cycles can be very different in different disciplines and sub-fields. Thus the way to manage research data can be different between disciplines. Nonetheless, the lifecycle approach can be a good way to understand the data practices of researchers. For example, to understand how researchers create data, manage, store and share their data through the process of research. Thus, the lifecycle approach is not only useful to evaluate research data management services but can also be useful for investigating data practices of researchers.

## 2.9 Research Data Management surveys

A number of research data management surveys have been conducted by UK institutions, for example, University of St Andrews (2016), University of Sheffield (Cox and Williamson, 2015), London School of Hygiene & Tropical Medicine (Knight, 2013), University of East London (Grace, 2013), University of Nottingham (Parsons, 2013), University of Exeter (Open Exeter Project Team, 2012), University of Hertfordshire (Nassiri & Worthington, 2012), University of Lincoln (Winn, 2012), University of Oxford (Wilson et al, 2012), Newcastle University (Kometa, 2011) and University of Northampton (Alexogiannopoulos, 2010). All of these surveys aim to understand current researcher data practices, identify needs and requirements for research data management, and to use findings to inform and develop policy framework, services and infrastructures.

All of these surveys have examined three main areas, i.e. types of data that researchers collect, how data are stored and backed up, and data sharing practices. The results across these surveys are fairly consistent. For example, most of the results show that researchers create a diverse range of data types, such as, audio files, images, websites, geospatial data, raw data etc., in which text documents and Excel spread sheets are the most common (e.g. Cox & Williamson 2015 and University of St Andrews, 2016). A couple of surveys also point out that the most common form of data collection is survey, interview, focus group and observation (Open Exeter Project Team, 2012; Alexogiannopoulos, 2010).

Results showed researchers store and back up data regularly and have multiple copies. The most popular places to store their data are private laptop, office PC, external hard drive, university shared drive and USB drive. However, answers in the follow up interviews of the Northampton survey revealed that researchers use different storage methods at different stages in the research life cycle (Alexogiannopoulos, 2010). Furthermore, most of the results indicated that most researchers intended to keep their data 5-10 years. However, some of the surveys suggested that data are usually neglected by researchers once the research project is finished (Alexogiannopoulos, 2010; Open Exeter Project Team). There was limited action to archive or to preserve the data properly.

Results of the surveys, e.g. Cox & Williamson (2015), also show sharing research data openly is not common. It is found that researchers usually only share their data with collaborators, research groups members and colleagues at the moment (Parsons, 2013; Nassiri & Worthington, 2012; Winn, 2012). And email attachment is the most common methods of data sharing.

In general, these surveys give an overall view of current researchers' practices in relation to research data management. However, most of these surveys are predominately quantitative and lack a

qualitative dimension. In addition, none of the survey results are presented with analysis at disciplinary level, although some of the results are presented according to faculty, such as, Open Exeter Project Team (2012), Kometa (2011). One of the surveys (Open Exeter Project Team, 2012) shows that humanities researchers prefer using privately owned computers to store data, while science researchers prefer using office computers and university network drives to store data. The results also show that Medical science has the greatest proportion of researchers that do not back up data because they are automatically backed up or backed up by IT staff (Open Exeter Project Team, 2012). In addition, researchers in the life sciences, physical sciences, mathematics and engineering are the most willing to share data while medical researchers are the least willing to share it (Open Exeter Project Team, 2012; Wilson et al, 2012). Evidently, there are significant differences in research practices between faculties. Thus in order to develop effective data infrastructure, services and training for research communities, cultural characteristics of disciplines have to be understood. Yet the surveys to date have not analysed social sciences in depth, let alone specific disciplines like Geography and their sub-fields.

## 2.10 The concept of data

In order to manage data effectively, it is not only essential to have a deeper understanding of what research data management involves but also what constitutes data. While research data management is a relatively new agenda for institutions, the notion of data is not new. The term 'data' has existed for a long time. According to the Oxford English dictionary (OED), the etymological origin of the word is from Latin "datum", a plural form of data, meaning 'something given'. It was suggested that the term data was first used in English in 1646 (Rosenberg, 2013). Rosenberg (2013) conducted an historical analysis of the usage of 'data' suggested that the term "data" was used in a wide range of contexts in the 18<sup>th</sup> century, for example, in theology and mathematics. The term "data" could refer to facts from biblical scripture, or evidences prior of an argument (Kitchin, 2014; Borgman, 2015). The usage of the term had been steady increasing over the centuries.

Today, data has become a diversely understood concept (Kitchin, 2014). For example, while data can be seen as a representation of facts, it also can be seen as binary data processed in computers (Floridi, 2008). It can be seen that the exact meaning of data varies from context and perspective. There has been attempts to conceptualise the notion of data. A common approach to conceptualise data in information sciences is through the knowledge pyramid (Ackoff, 1989, Fricke, 2009; Jennex & Diego 2013) (Figure. 2.5).

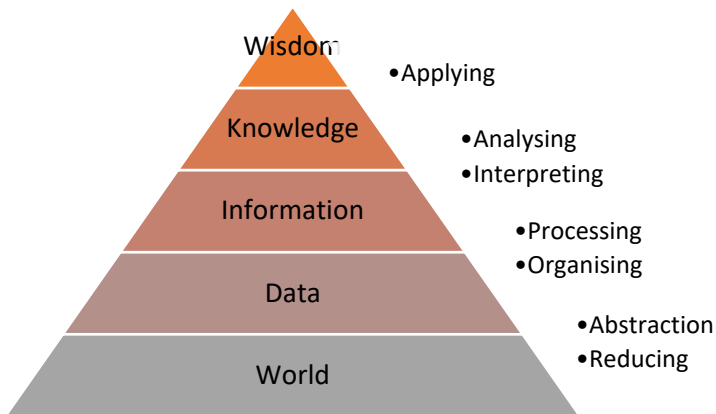


Figure 2.5 Knowledge pyramid (adapted from Kitchin, 2014; 10)

In the hierarchy, data refers as symbols, a representation of objects or activities. Information involves data that are processed with meaning and purpose. Knowledge answers how-to questions (Ackoff, 1989). Kitchin presents an interpretation of the differences between each layer of the pyramid (2014: 9)

*“Each layer of the pyramid is distinguished by a process of distillation (reducing, abstracting, processing, organising, analysing, interpreting, applying) that adds organisation, meaning and value by revealing relationships and truths about the world.”*

Although, the knowledge pyramid is commonly use, yet it has been criticised on the grounds that it oversimplifies reality and fails to represent the complex constructs of these relationships (Borgman (2015).

### 2.10.1 Research data

Numerous definitions of data in the context of academic scholarship can be found, as research data management is increasingly important. Pryor (2012) refers to data as

*“The primary building block of information, comprising the lowest level of abstraction ..., where it is identifiable as collections of numbers, characters, images or other symbols ...”*

This definition is similar to the view from the knowledge hierarchy mentioned earlier, which sees data as a fundamental block of information. Definitions of research data not only can be found in the literature, but also in research funders’ and institutional policies. For instance, EPSRC (2016) defines research data as:

*“recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is*

*created in digital format, all research data is included irrespective of the format in which it is created.”*

Thus the Research Council views research data as a collection of facts. Another example of definitions of research data is from University of Southampton’s (2015) research data management policy. It states:

*“1.3 “Research Data” means information in digital, computer-readable format or paper-based that:*

*1.3.1 is contained or presented in various ways including notes, facts, figures, tables, images (still and moving), audio or visual recordings.....”*

*1.3.2 which is collected, generated or obtained during the course of or as a result of undertaking research (which includes but is not limited to conducting field or laboratory experiments, conducting trials, surveys, interviews, focus groups or analysis of data); and*

*1.3.3 which is subsequently used by the Researcher as a basis for making calculations or drawing conclusions to develop, support or revise theories, practices and findings.”*

In this definition research data is viewed as information gathered in the process of research. It also shows the diverse nature of research, as it is identified data can be digital and non-digital. It also can be notes, images, or visual recordings. Research data can be categorised in different ways for operational and general purposes (Borgman,2015). Kitchin (2014) identified some categories of data, for example, quantitative data vs qualitative data, structured data, primary, secondary and tertiary data. Research data can also be categorised by degree of processing. One of the recognized examples is NASA’s Earth observing system data information system (2010), which divided data products to different levels, ranging level 0 to level 4, from raw, high resolution data to model output, results from analyses of lower-level data.

Furthermore, research data can be classified by its origin. For example, National Science Board (2005) categorised research data into three types, i.e. observational data, computational data and experimental data. The implication of these categories was to show the different requirements for managing and curating in different types of data (NSB,2005; Borgman,2015). It was suggested that these categories can not only be applied to sciences, technologies, social sciences but also to health sciences, arts and humanities. However, these categories seem to fail to recognize disciplinary differences. For example, observational data in sciences can be very different from observational data in qualitative social research. Requirements of for curating observational data in sciences, such

as, remote sensing data can be hugely different from data from observations in social sciences, e.g. interview data. Thus, these categories cannot truly differentiate the different requirements for curation in different disciplines.

In general, it is not easy to define and categorise research data. The nature of data is not only diverse but also abstract. Rosenberg (2013: 18) concluded, the nature of the term 'data' is still rhetorical in nature over centuries, as he writes,

“...data is rhetorical. A datum may also be a fact, just as a fact may be evidence [...] When a fact is proven false, it ceases to be a fact. False data is data nonetheless.”

This nature of data can pose challenges for researchers and information professionals to manage research data effectively.

## 2.11 Summary

In this chapter, two key topics of the study, disciplines and research data, were reviewed. The first part of the chapter provided a critical review of key concepts and theories of academic discipline. The notion of discipline is complex. In general, academic disciplines can be understood from three perspectives: intellectual, social and organisational. The chapter showed that there are different theoretical frameworks for understanding disciplinary differences, such as, Becher's (1989), Whitley's (2000) and Abbott's (2001). In order to investigate the applicability of disciplinary theories to explain data practices within geography, it is essential to have a deep understanding of these conceptual frameworks. The second part of the review looked at topics relating to research data. The research data management agenda has largely been driven by policies rather than researchers. This section looked at what research data management is and examined the research data management lifecycle from a curation perspective. These shed some light not only on what constitute good data practices for researchers but also introduce a lifecycle approach to investigate researchers' data practices, e.g. how researchers create, store and share their data in the research cycle. Furthermore, the concept of data was also reviewed in the chapter. Defining what is data is not straight forward. The concept and meanings given to data are not only diverse but also abstract. The section showed that the definition and meaning of data are change over time.

# Chapter 3 Methodology

---

## 3.1 Introduction

This chapter presents the methodology and research design of the study. It introduces the philosophical assumptions underpinning the research design and describes the research design itself. The research design consisted of two phases and used multiple methods. The chosen methods are evaluated in the context of alternate possible methods that after careful consideration were not included in the research design. Having described and justified the methodology in this chapter a detailed explanation of how each method was implemented is provided at the beginning of each relevant empirical chapter (chapters 4, 5, 6 and 7).

## 3.2 Philosophical assumptions

This section discusses the philosophical assumptions and rationale underlying the research, including, the relationship between theory and research and epistemological and ontological considerations.

### 3.2.1 The relation between theory and research

The following diagram (Figure 3.1) illustrates the approach taken to linking theory and research. In this study, an inductive strategy is adopted. As opposed to a deductive strategy, which is concerned with deducing a hypothesis and testing a theory (Bryman, 2016), new theories in this study are generated from the data rather than formed before the data collection. Nevertheless, existing theories are used to guide the empirical inquiry of the study. The study uses relevant theoretical literature, e.g. Becher's and Whitley's frameworks on disciplinary cultures to guide empirical data collection and analysis. New theories are then generated from the findings, which are fed back into the knowledge of relevant IS domains, such as, scholarly communication and research data management.

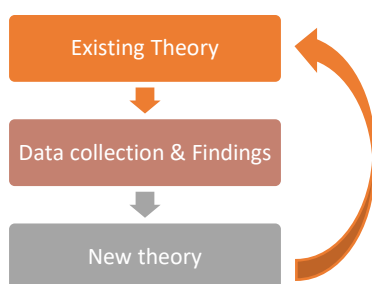


Figure. 3.1 The relationship between theory and research of this study



### 3.2.2 Ontological and epistemological considerations

In general, ontological issues are concerned with what is the nature of reality (Crotty, 2003; Creswell, 2013). Epistemological assumptions concern how reality is known and what is regarded as valid knowledge in a field (Creswell, 2013; Bryman, 2016). As Creswell (2013) has stated, a researcher's ontological assumptions and epistemological beliefs shape methodological perspectives, such as decisions about quantitative/ qualitative approaches, data collection methods and analysis.

This research adopts an interpretivist stance. Interpretivism is also commonly labelled social constructivism (Mertens, 2010; Denzin & Lincoln, 2011; Creswell, 2013). It is associated with an idealist ontology (Blaikie, 2007). Interpretivists believe that social reality, truth and meaning are constructed by human beings. The existence of the external world is not independent from our thoughts and knowledge. Thus, social reality cannot be simply one absolute objective reality (Bryman, 2016). As Creswell (2013) states, the social constructivist believes that there are multiple realities which are constructed through individuals' experiences and interactions. This social constructivist belief shapes the study, such as in the nature of research questions, research design and interpretation of the findings. This research seeks to explore the relationship between research data and discipline. Instead of seeing data as value free, and disciplines as objective identities, this study regards the notion of disciplines and data as being constructed by the interactions of social actors and institutions. The research, therefore, seeks to understand the complexity of disciplines and data from multiple dimensions, such as intellectual and social perspectives. Furthermore, the interpretivist position of this study acknowledges that subjective meanings are important. This research seeks to be epistemologically consistent. Consistent with the interpretivist stance, it adopts an inductive strategy and primarily qualitative methods. For example, the semi -structured interviews, focuses on an interpretive understanding of the view of research participants rather than measurements and objectivity.

### 3.3 Overview of the research design

The research approach of this study is predominately qualitative, although it does use multiple methods. As Jupp states (2006: 249), qualitative research "focus[es] on the meanings and interpretation of social phenomena and social processes in the particular contexts in which they occur". Given that the research is concerned with developing a deeper understanding of researchers' perceptions that underly disciplinary research data practices, a predominately qualitative approach is appropriate.

A range of data collection methods are used in this research. As the notion of discipline and research data are very complex, any single method or technique alone is unlikely to be sufficient to explore the dynamic and complex relationship between research data practices and the culture of a discipline, such as, intellectual and social structures of a discipline. Thus, this research comprises two sequential phases using multiple methods (Table 3.1). The first phase was exploratory, relying on a range of desk-based research, including an analysis of geography department websites and researcher web profiles and a bibliometric study of collaboration patterns. This first phase of the research aimed at providing an understanding of geography as a discipline and its research data before beginning the interviews. The second phase of the study was a series of interviews with researchers in geography. The aim was to understand researchers' data practices, their attitudes toward data sharing, and also for them to reflect on the impact of current changes around the discipline and data environment.



Table 3.1 Research design

### 3.3.1 Purposes of methods

The purpose of each method in Phase 1 and 2 are outlined as follows:

#### **Phase 1**

##### *An analysis of web content*

a. Departments & research groups

To examine how geography and its subfields are structured and organised within institutions (i.e. the organisational aspect of geography).

b. Researchers' web profiles

To look at key research topics and methods in geography on an individual researcher level (i.e. the intellectual aspect of geography)

### A Bibliometric study

To investigate collaboration patterns in geography through co-authorship (i.e. the social aspect of geography).

#### Phase 2

#### Interviews

To understand the impact of current changes around data environments and the discipline on researchers' data practices, especially their attitudes toward data sharing.

### 3.3.2 Research questions and methods of collecting data

Having outlined the aims of the methods used in this research, it is also important to relate these methods to the research sub-questions. Thus, the table below shows how these methods are linked to the research sub-questions (Table 3.2).

Research sub-questions	Methods
1. What are the characteristics of geography as a discipline (i.e. intellectual, social and organisational aspects)?	Web based content analysis Bibliometric study Interviews
2. What are the data practices of researchers in geography and are there identifiable patterns within or across sub-disciplines?	Interviews
3. Do patterns of data practices, especially data sharing vary across in different sub-disciplines geography and, if so, why?	Web based content analysis, Bibliometric study Interviews

Table 3.2 Research sub-questions and methods of data collection

### 3.4 Research design of Phase 1

This section describes the details of the research design of phase 1. Phase 1 was exploratory in nature. According to Blaikie (2007), there are two research stances that can be adopted related to the relationship between the researcher and the researched, that is being an expert or a learner. In Phase 1, the research problem was approached as a learner. As Blaikie (2007: 11) states, adopting a learner research stance,

*“the researcher aims to set aside existing social scientific knowledge [...] the answers to the RQs emerge from this learning process, rather than from a body of social scientific knowledge.”*

### 3.4.1 Purpose of Phase 1

Phase 1 aimed to understand the nature of geography and its-sub disciplines prior the second phase, e.g. to identify specific intellectual and social characteristics of geography, such as research problems and collaboration patterns. Furthermore, the purpose of phase 1 was to identify suitable participants for interviews in phase 2. A multi-method strategy was adopted in Phase 1. As mentioned earlier, Phase 1 consisted of a Web-based content analysis and a bibliometric study. This multi-method approach was used to extend the breadth and range of inquiry (Clark et al., 2008). As the nature of a discipline is complex and dynamic, using different methods in this way allowed different disciplinary aspects of geography to be explored, e.g. research methods, objects of study and collaboration patterns in Geography. Results of these preliminary studies not only helped sampling in phase 2, but also informed the development of interview questions.

### 3.4.2 Methods considered for phase 1

A variety of approaches were considered, including:

- 1) A questionnaire survey of research practices of geography researchers;
- 2) A small number of interviews with UK research data management experts;
- 3) A link analysis study of disciplinary interpersonal networks (e.g. based on links between researchers and researcher groups);
- 4) An analysis of data citations in geography
- 5) A web-based content analysis of geography departmental websites;
- 6) A small scale bibliometric study of geography.

Only 5 and 6, i.e. web-based content analysis and the bibliometric study, were selected for phase 1. The following explains why other methods, such as, a survey, interviews and link analysis, were not selected. A quantitative questionnaire survey could have potentially been used to cross-validate findings of the qualitative Phase 2 (Creswell, 2003), yet it was not chosen to be conducted in this research. One of the main reasons was because response rates of surveys are typically low. The researcher would have to put a lot of effort into gaining a response across institutions. There are no convenient forums through which the survey could be distributed to researchers and the kinds and numbers of questions that can be asked are limited, e.g. only small numbers of open questions can be asked (Bryman, 2016).

A small number of interviews with UK research data management experts were considered for phase 1, as it could be useful for identifying best practices for effective data management in geography. However, it was considered that interviews can be time consuming (Blaxter et al,2010).

Furthermore, research data management experts would not necessarily be able to provide a strong geography perspective which is essential to this research, as this research focuses on the understanding of data practices in geography. Thus, this method was not chosen for phase 1.

A link analysis was also considered. A pilot feasibility study of link analysis was conducted. It was found that the method was not feasible after the pilot study. The details of the pilot link analysis are reported below. Since it is a somewhat less familiar method, a fuller explanation of link analysis follows.

Link analysis is used to obtain useful information through analysing hyperlinks between websites. Typically, link analysis software uses a web crawler to obtain link data. A web crawler is an automated computer program, that starts crawling web pages from a starting URL(s) list, downloads pages and follows new links. The visited hyperlinks are automatically collected, calculated and processed. Link analysis reports and network diagrams are then generated by the link analysis software.

Link analysis has been commonly used in scholarly communication research (Thewall, 2004). Analysing hyperlinks between academic webpages can reveal useful information about disciplines. For example, Harries et al. (2004) used link analysis to examine disciplinary structures and networks between three disciplines, i.e. maths, physics and sociology. The study shows that there are differences in hyperlinks between these three subjects, such as, mathematicians tended to link to other mathematicians, while physicists tended to link to departmental home pages but not fellow academics. Harries et al (2004) suggested that this may be due to the high degree of institutional research in physics. It would therefore appear that link analysis can be a useful method. However, results sometimes can be quite limited; for example, it can be unclear what a link means. It may not necessarily indicate a real connection between disciplines. In addition, there could be connections that are not expressed as links. Harries et al (2004) point out that link analysis can be used as the starting point of an exploration rather than as the end result. Therefore, the results of link analysis in this research were planned to be integrated with qualitative explorations in order to be fully understood.

In the pilot study, link analysis was used to investigate the relationship between research groups, departments and interdisciplinary interpersonal networks in geography. It is to be noted that the nature of this part of the study was very exploratory. Different sets of seed URLs, such as, a list of geography department home pages URLs or a list of individual researcher webpages URLs, were used to address different objectives. For instance, in order to explore the relationship between

research groups, a list of research group webpage URLs will be used as the seed URLs. The link analysis study followed Thelwall's (2004: 3) approach, which is summarised below:

- Formulate research questions
- Conduct a pilot study
- Identify webpages that are to address research questions
- Collect data from a web crawler
- Apply data cleansing techniques
- Correlate link count data with other data sources, such as literature review
- Validate results through a link classification exercise
- Report results of the study
- Report the limitations of the study

A pilot feasibility study was carried in August 2013. Sixteen Geography departments were selected and the URLs for the webpages listing all the academic staff in these departments were used as seed URLs in the link analysis. These initial URLs were manually identified through web searching (Appendix 2). The link analysis was carried out using SocSciBot, a free web crawler software for link analysis research (<http://socscibot.wlv.ac.uk/index.html>). Based on the link analysis, network diagrams were produced. The network diagram below shows that there are only very few connections between geographers in the UK. Connections were mostly between authors and co-authors, and previous PhD students and supervisors. The results of the pilot study were unable to reveal the relationship between research groups and departments in geography. The results indicated there were very limited hyperlinks on geographers' website (Figure 3.2). Therefore, the method of link analysis was considered not suitable for further examining social characteristics of geography as a discipline for this research. Rather web-based content analysis of geography departmental websites and bibliometrics were selected to investigate the characteristics of geography in phase 1.



Table 3.3 is a summary of the methods that were considered for phase 1. It details the focus, advantages and limitations of each method.

Methods	Focus	Related to sub-research questions (sRQ)	Advantages	Issues/ Limitations
<b>1. Questionnaire Survey</b>	Collect basic overview of geographers' views of their discipline and data practices	sRQ 1: the nature of geography  sRQ 2: Researcher's understanding and attitudes towards RDM	Wide participation	How to distribute questionnaire  Response rates likely to be low
<b>2. Interviews with data management experts</b>	Data management experts' views on key issues in research data management in geography	sRQ 2: Data practices of researchers	Able to identify best practices for effective data management	Time consuming  Data management experts do not necessarily have extensive knowledge of data practices of geographers
<b>3. Link analysis</b>  site inlinks/ outlinks for research groups/ departments	Relationship between research groups/ departments	sRQ 1: the nature of geography, focus on academic networks	May be able to reveal interdisciplinary interpersonal networks	Labour-intensive method  Results can be quite limited. It can only answer a small part of a research question
<b>4. Citation analysis</b>	Data citation in geography	sRQ 1: the nature of geography - social aspect  SRQ2: Data practices of researchers	Able to investigate the impact of data and potentially to identify researchers' network	The practices of data citation have not been well established in geography
<b>5. Web-based content analysis of geography departmental websites</b>  <i>examine departmental websites and researchers' profiles</i>	Institutional departments and research groups of geography in the UK  Current key research subjects in geography  Methods and sub-disciplines in geography	sRQ 1: the nature of geography as a discipline (esp. organisational and intellectual characteristics)	Easy access to information  Expand and deepen disciplinary knowledge of geography for interviews	Sampling issues: How to sample departments, researchers and publications  Time consuming
<b>6. Bibliometrics</b>  <i>Co-author analysis</i>	Collaborations In geography	sRQ 2: the nature of geography as a discipline- the social aspect	Able to identify collaboration pattern in geography	Researcher's knowledge of method  Results can be limited.  It can only answer part of the research question

Table 3.3 Summary of methods that have been considered in Phase 1, including both implemented and not implemented methods



### 3.4.3 Web-based content analysis

The first study conducted in Phase 1 was a web based qualitative content analysis. It was carried out between December 2013 and March 2014. The coding categories used in Phase 1 were piloted and modified. Details of the development of the coding categories are presented in Chapters 4.2 and 6.2.2.

The Internet and the Web has become part of the everyday life of academics, for example, most academics have a webpage which contains information such as research interests, a list of publications, professional involvement etc. There is vast amount of information about scholars on the Web. The growth of the Internet has provided new research opportunities for researchers. Social science researchers have been exploring potential internet research methods, for example, using webpages as objects of analysis (Bryman, 2016; Fielding et al, 2008). Bryman (2016) points out that websites and web pages can be useful sources of data for both quantitative and qualitative analysis. He considered websites and webpages as 'virtual documents', as a form of document to analyse (Bryman, 2016).

Increasingly, academic departments showcase their research through departmental websites. Departmental websites provide information not only about disciplinary research but also interdisciplinary research within departments. Examining the websites of geography departments, can increase our knowledge of current research trends within geography and its sub-disciplines in the UK. There are some advantages of using this approach. Firstly, such information on the internet is freely available and has 24-hour accessibility. Second, the audiences of departmental websites are generally mixed (e.g. the general public), thus their content is usually clear and comprehensible. In addition, Webb et al (1966) identified that data collection that is nonreactive and unobtrusive is beneficial. Data collection from web-sites offers these benefits. As research participants are not aware that they are being studied, therefore the data gathered are not affected by the presence of researchers (Bryman, 2016; Berg, 2009). Yet since it is public it can be considered as published, so does not require ethics permission to use. This approach requires no effort on subjects' behalf. Janetzko (2008) points out that this unobtrusive method allows researchers to study large numbers in a sample.

There were also some potential issues to be aware of in this study, however. Bryman (2016) has pointed out the need to be aware of possible distortions and the credibility of website materials. As sources of organisational information can be about promoting the organisation themselves and websites do not necessarily represent what the organisation actually do in the real world (Gauntlett & Horsley, 2004). In addition, departmental websites are not designed for this particular research,

thus information on them will be non-standardised and can be incomplete. Furthermore, some websites can be out of date while some websites can be changing rapidly. The websites that were used in data analysis can change frequently or no longer be available in the future (Bryman, 2016). Since the focus of this part of the study was not about finding out the latest information on geography departments, but to understand geography as a discipline, the web-based study still can provide useful information for the research even if some websites are out of date. The next section describes details of the web based content studies.

### Part 1: Departments & research groups

The first part of the web based content study examined UK geography departments and research groups in order to understand how geography and its subfields are structured and organised within institutions. In the UK, research communities in departments are usually organised into research groups and clusters. In order to explore how geography and its sub-disciplines are structured in an institutional context, a mapping exercise of UK geography research groups was carried out. A systematic approach was used to identify geography research groups in the UK. The sample of this part of the study consisted of geography departments from 49 institutions that took part in the Research Assessment Exercise 2008 for HEIs in the UK. The data was stored and analysed in an Excel spread sheet.

### Part 2: Researchers' profiles

The second part of the study examined research group leaders' profiles on departmental websites. This part of the study not only provided an opportunity to explore the intellectual characteristics of geography but also to acquire background knowledge for in-depth interviews for Phase 2. The study examined some disciplinary characteristics of geography through an analysis of researchers' web profiles, such as:

- research topics
- research methods
- types of publication (e.g. books, journals)
- research funders in geography

The sample of this part of the study consisted of research group leaders from 15 geography departments, i.e. five top-ranking departments in the 'geography and environmental studies' unit of assessment in the Research Assessment Exercise 2008, five middle-ranking departments, and five low-ranking departments. Research group leaders were chosen to be examined because they usually have relevant research experience to the focus of the research groups. In addition, the sampling approach can cover a range of universities with diverse research strength, such as,

research intensive universities vs. teaching universities (further details on the implementation of the web content study are presented in chapters 4 and 5).

### 3.4.4 Bibliometrics

A small bibliometric study was also used to explore the social characteristics of geography in geography. Bibliometrics has been defined as the quantitative study of published literature (Hood & Wilson, 2001). Bibliometrics has been used extensively in the study of scholarly communications, such as to explore emerging research topics of disciplines (Glanzel, 2012; Van den Besselaar & Heimeriks, 2006). This part of the study focused on identifying co-authorship patterns in geography, e.g. the number of co-authors and who are the co-authors of geographers in order to understand the social disciplinary characteristics in geography.

Analysis of co-authored papers have long been used for measuring research collaboration (Subramanyam, 1983, Katz & Martin, 1997). For example, Glanzel (2014), Velden et al (2010), Melin & Persson (1996) and de Solla Price (1963) explored research communication and collaboration using co-authorship. The nature of research collaboration is complex. Subramanyam (1983) suggested, due to the fact that research collaborations often happen over a period of time, it is difficult to use traditional methods such as, interviews, observations or questionnaire to capture the complex nature and the degree of research collaboration. It is identified that there are a few advantages of using bibliometrics to examine research collaboration in this research, such as that it is quantifiable, inexpensive, nonreactive and unobtrusive (Subramanyam ,1983). Yet, like any other research methods, it has its own limitations. Generally, researchers make assumptions when using bibliometrics to measure collaborations (Katz & Martin, 1997; Subramanyam, 1983). It is assumed that:

1. collaborative effort results in research papers;
2. all collaborators are stated as co-authorship in research; and
3. all collaborators stated in the papers have actually collaborated

(Katz & Martin, 1997; Subramanyam, 1983)

None of these assumptions is necessarily completely correct. Thus, Katz & Marti (1997) pointed out that co-authorship can only be a proxy, a partial indicator of research collaboration. Nevertheless, bibliometrics is a useful way to examine research collaborations in geography. In this research, *Scopus* was used as the source of data. Data then was processed and analysed using excel (further details see chapter 6).

### 3.5 Research design of Phase 2: Interviews

After gaining some understandings of the disciplinary characteristics of geography through web-based content analysis and bibliometrics, data collection of Phase 2 was carried out between September 2014 and November 2014 using semi-structured interviews. The purpose of phase 2 was to investigate researchers' research practices, understandings and attitudes of data sharing and thus was related to research sub-questions 1, 2, and 3. This phase was a key part of this research.

A research interview has been described as “a conversation with a purpose” (Berg, 2009; Kvale & Brinkmann, 2009; Marshall & Rossman, 2006; Kahn & Cannel, 1957:149). In a research interview, the interviewer seeks to elicit information from the interviewee (Bryman, 2016). The interview is one of the most popular methods in qualitative research (Bryman, 2016; Packer, 2011). Indeed, the interview has become the method of choice for data collection in a diverse range of qualitative research approaches, including ethnography, phenomenology, case study and grounded theory (Creswell, 2013; Packer, 2011). It can be used as an overall strategy or as one of the methods employed in a research project (Marshall & Rossman, 2006). Many qualitative research projects, in fact, use only interviews to collect empirical data (Packer, 2011).

In this research, interviews were used as a core data collection method. According to May (2011) and Kvale & Brinkmann (2009), qualitative interviews seek to understand the world from interviewees' perspective. Using interviews in this phase is in line with the qualitative interpretative philosophical approach of the research. As the purpose of Phase 2 was to understand researchers' research practices and their attitudes towards research data management, using interviews was an appropriate method which can provide rich insights into interviewees' experience, views and attitudes related to research data.

#### 3.5.1 Types of interviews

Semi-structured interviews are employed to obtain qualitative data in this research. There are three main types of interviews: the structured interview, the semi-structured interview and the unstructured interview. Typically, the structured interview is associated with quantitative research, while the un-structured interview and semi-structured interview are associated with qualitative research (Bryman, 2016). The differences between these interview types are not only their degree of flexibility in connection with interview structure, i.e. the wording and sequence of the interview questions but also their underlying assumptions (Berg, 2009; Gillham, 2000).

### Structured interviews

In a structured interview, the researcher uses a formally structured interview schedule. Questions in this type of interview are usually very specific and close-ended (Bryman, 2016). Each interviewee is asked identical interview questions in the same wording and order. Thus, answers can be aggregated, standardised and compared (Berg, 2009; Bryman, 2016; Denscombe, 2010). Research using structured interviews is usually based on a few assumptions. First, researchers using this type of interview have solid ideas of what things they want to reveal and what kinds of answer they need in the interviews (Gillham, 2000). It is assumed questions used in interviews are sufficient to gather all information for their study (Berg, 2009). Second, it is assumed that there are sufficient common vocabularies between the interviewer and interviewees, so interviewees can clearly understand what they are being asked (Berg, 2009). Third, it is assumed that the meaning of each question is the same for every interviewee (Berg, 2009; Frankfort-Nachmias & Nachmias, 2000). A structured interview approach is not suitable for this research, in particular because it cannot be assumed that there are enough common vocabularies to formulate the questions so that they are understood in the same way. It is expected that terms such as research data, research data management need to be clarified during interviews, for instance, research data can mean different things to different researchers. Overall, this type of interview does not allow enough flexibility to gain insights into researchers' attitudes and practices towards research data management.

### Un-structured interviews

In contrast to the structured interview, the un-structured interview does not have an interview schedule. It is like a free-flowing conversation which is started by a theme or a topic (Bryman, 2016). Unlike the structured interview, the underlying assumption here is the researcher does not know all the necessary questions prior to the interview (Berg, 2009). Thus the researcher does not have a full list of interview questions. Interviewers need to develop and generate appropriate questions and follow-up probes in the interviews (Hakim, 2000). Researchers using structured interviews are interested in the interviewees' views. Interviewees are encouraged to use their own words to describe their own experiences, beliefs or issues etc. (Frankfort-Nachmias & Nachmias, 2000).

### Semi-structured interviews

Similar to un-structured interviews, semi-structured interviews also place emphasis on the interviewees' point of view (Bryman, 2016). Semi-structured interviews involve using an interview guide. The interviewer will have prepared a list of questions. However, unlike the structured interview, interviewers do not necessarily follow the exact wording or the order of the questions. In addition, follow-up questions and probes are used to get more details. In semi-structured interviews,

researchers assume that in order to standardise questions, questions need to be expressed in words that are familiar to the interviewees (Berg, 2009). This assumption shows that researchers are actually aware of different interviewees understand the world differently (Gubrium & Holstein, 2003). Therefore, researchers understand and interpret the world from interviewees' viewpoint.

In this research, the semi-structured approach was chosen because this approach provides flexibility to gain insights into researchers' research practices, data management practices and their disciplinary cultures. Yet, the semi-structured approach also provides some structure to allow comparability, for instance, to compare differences between different sub-disciplines, which unstructured interviews cannot offer.

### 3.5.2 Sampling

A total of 23 interviews with academic researchers in geography were conducted during the period 23 June – 7 Oct 2014. The first interview served as pilot. The pilot resulted in the minor modification of one or two questions, but did not necessitate any major revision to the interview protocol, therefore the first interview was included as part of the dataset. In a review paper, Baker and Edwards (2012) investigated the question of 'how many qualitative interviews is enough'. The answers suggested by experts in the paper range from 1 to 60 interviews, with 30 being the mean. As Baker and Edwards (2012) and Pratt (2009) point out, there is no right answer and it depends, for example, on the research questions, objectives and resources. More interviews does not necessary mean better. Qualitative research is more focussed on 'richness, complexity and detail' than quantity (Baker and Edwards, 2012:5). Tracey Jensen advised that the quality of analysing interviews is actually more significant than the quantity of the interviews (Baker and Edwards, 2012). In this research, the sample was 23 researchers because the number is not large enough to enable a comparison across human geography and physical geography.

Purposive sampling was used to select participants, which was consistent with the overall qualitative approach of the research. Sampling in qualitative research is quite different from in quantitative research (Kumar, 2011). Quantitative research usually involves probability sampling in which a sample is selected randomly to ensure each unit in the population has an equal chance to be selected (Bryman, 2016). In contrast, most sampling in qualitative research involves purposive sampling. Purposive sampling is a form of non-probability sample. Researchers do not select a sample using a random selection approach but rather select sample participants or cases strategically in order to ensure the sample is relevant to the research goals (Bryman, 2016).

Typically, researchers select a wide range selection of types of individual to represent factors that might affect responses to the research questions. A major problem with purposive sampling is that

researchers are unable to generalise findings to a population (Bryman, 2016). Nevertheless, as Polit & Beck (2010) point out qualitative researchers are more interested in gaining rich, contextualised understandings about issues, individuals or a situation rather than generalising findings, which is the goal of most quantitative research.

As this qualitative research is exploratory by nature, suitable research fields and interviewees were identified after Phase 1. The criteria for selecting suitable participants were the research area of researchers, the volumes and types of research data they produced and the ranking of the university.

### 3.5.3 Interview guide

An interview guide was developed (Appendix 10). The interview guide not only developed detailed interview questions but also determined the purpose of each question. Interview questions were divided into few topic areas, such as, research areas, data practices, data sharing and data re-use. All interviews were recorded and notes were taken for important things during interviews. Data were then transcribed and imported to Nvivo for categorising and coding.

### 3.5.4 Data analysis

Thematic analysis was employed to analyse the qualitative interview data. Braun and Clarke (2006:79) describe thematic analysis as:

*“a method for identifying, analysing, and reporting patterns (themes) within data. It minimally organises and describes the data set in (rich) detail. However, frequently it goes further than this, and interprets various aspects of the research topic.”*

Although thematic analysis is widely used, there is no clear standardised agreement about the definition of thematic analysis or the way to produce it (Bryman, 2016; Braun & Clarke, 2006; Boyatzis, 1998). It is sometimes argued that this approach is not an identifiable method of analysis, as searching and categorising themes featured in thematic analysis are essential in a lot of approaches of qualitative data analysis, such as grounded theory, qualitative content analysis, and narrative analysis (Bryman, 2016). Nevertheless, this approach was chosen for this research because of its flexibility. As Braun & Clarke (2006) point out many other analytic approaches are theoretically bounded. For instance, interpretative phenomenological analysis is bounded to phenomenological epistemology, while grounded theory requires detailed technological knowledge of the approach in order to develop theory which is grounded in data (Bryman, 2016). Thus, thematic analysis provides a theoretically-flexible approach. It is an alternative to researchers who do not want to commit totally to produce a fully grounded theory analysis (Braun & Clarke, 2006).

In this research, data was analysed in six phases. The analytical approach was adopted from Braun and Clarke (2006). It is noted that the process of analysis is not a linear set of steps but a recursive process (Braun & Clarke, 2006): It involves moving back and forward between phases.

#### Familiarisation with data

The analysis started with familiarisation with the content of interview data, including, reading and re-reading whole transcripts and field notes, taking notes for ideas, annotating and highlighting significant quotes. Previous disciplinary knowledge gained from web-based content analysis and bibliometric work during the phase one of the study were also used to be sensitive to the context of the interview data. This stage of the analysis was used to gain a holistic understanding of the whole set of empirical data. It was an initial exploration of the data before coding (Saldana, 2013; Bazeley, 2013). This preliminary stage of data analysis was carried out throughout and after the data collection process.

#### Generating initial codes

Second, a first cycle coding was carried out (Salanda, 2013). Codes were applied to a short section of data. According to Salanda (2013:3), a code in qualitative research is

*“a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of [...] data”.*

Coding can be seen as a process of indexing. As Bryman (2016) suggested, coding creates an index of terms which help interpret and theorise data. Thus, this stage of analysis of this research was about organising data in a meaningful way. Interesting aspects of the data were coded systematically to generate initial codes from the data. Table 3.4 below shows some initial codes generated from interviews of the research. As this stage of analysis was exploratory in nature, it can be seen that codes were not organised and some codes in some cases overlapped.



Scale of the field	Impression of other fields
Academic audiences	Informal communication
Anonymise	Institutional support
Authorship	Interdisciplinary, multidisciplinary
Changes in geography	Issues in archiving data
Collaboration	Journal policy
Consensus	Key research audiences
Conferences	Main area of research
Control	Methods- carefully planned
Culture of data sharing in different communities	Methods- preference
Data after the completion of research project	Methods- questionnaire
Data citation	Methods used
Data collection	Motivations to share data
Data misuse	Open culture
Data policies awareness	Physical geography model
Data requests by others	Policy interests
Department	Process data
Differences between sub-fields	Published
Difficulties of collecting data	Pure and applied research
Ethics	Quantitative
Fieldwork	Raw data
Geography as a discipline	Research assistant
Images	Research data produced by researchers

Table 3.4 Initial codes

### Refining codes and theorising concepts

The third stage of the analysis was to review and refine identified codes and develop themes or concepts. Second cycle coding was conducted. As Saldana (2009) stated, the purpose of the second cycle coding is to further develop a conceptual, thematic and theoretical organisation from the first cycle codes. Annotations and analytical memos were to help identify conceptual connections between codes and reflect how codes can be integrated into themes. Codes were then reviewed and evaluated to ensure the codes work with the coded data extracts as well as the entire data set. Definitions of codes and a reviewed code structure of this stage are shown in Appendix 3 and 4. At the end of phase five, a thematic map of the analysis and an initial plan of the write up of the interview findings were produced.

### Writing up the analysis

The final stage was to conduct the final analysis and write up the report. At this final stage, work was focused on linking the analysis to the literature and research questions in order to produce a coherent and logical report of the analysis (Braun & Clarke, 2006). The findings of the interview are reported in chapters 7-10.

### 3.6 Ethical considerations

The research was conducted ethically according to the ethics policies of Loughborough University and in conformity to wider ethical norms in the social sciences. As human subjects were involved in the research, ethical issues, such as informed consent and privacy had to be addressed throughout the research. Ethical issues in social science research are usually divided into four main areas: protection from harm, informed consent, privacy and honesty (Bryman, 2016; Leedy & Ormrod, 2010). All participation in this research was strictly voluntary and based on complete information about the purpose of the project (See appendix 5 and 6 for information sheet and consent form). Participants' welfare and rights were respected. In addition, all the personal information of participants was handled with anonymity and confidentiality. Using the internet to collect data raises specific ethical issues, particularly, the boundary between public and private space can be unclear (Association of Internet Researchers, 2012; Bryman, 2016). However, in this research, data that was investigated in the web-based study are in the public domain, i.e. university departmental websites, and do not involve sensitive personal messages or vulnerable communities. As Thewall (2010) and Hewson et al. (2003) suggest, published information that is voluntarily available on the public web can be used by researchers without informed consent.

### 3.7 Data sharing arrangement

The possibility of sharing data from this study was considered. Data collected from this research which can be potentially shared were 68 research group leaders' profiles from departmental websites (Phase 1), 68 researchers' publication lists (Phase 1) and 23 interview transcripts (Phase 2). In this research, the sampling strategy of the Phase 1 and Phase 2 research are interrelated, e.g. the initial sampling choice of Phase 2 interviewees was chosen from the 68 researchers of the 15 institutions from Phase 1. Sharing data from Phase 1, such as webpages and publication list of the researchers is likely to expose interviewee identifiable. Furthermore, given the public nature of the web and publication material and the uniqueness of some research topics of researchers, it is very difficult to anonymise the data from the interviews without reducing the research value of data. Additionally, the participant information sheet for interviewees guaranteed confidentiality. Thus, a decision was made not to publish the data in order to protect confidentiality of interview participants.

### 3.8 Summary

This chapter presented the methodology and research design of the study. The approach to this research is interpretivist, and this influences the research design and choice of methods. The study adopts a predominately qualitative approach. As one single method is not sufficient to explore the complex interrelationship between discipline and the data practices of researchers, multiple methods are used. The study consists of two phases. Phase 1 aims to provide a disciplinary understanding of geography prior to Phase 2. Phase 1 methods include an analysis of geography department websites and researcher web profiles and a bibliometric study of collaboration patterns. Phase 2 is the key part of this research, i.e. a series of interviews with researchers in geography. The aim is to understand researchers' data practices, their attitudes toward data sharing. The current chapter provides a justification and rationale for the choice of methods. Furthermore, the chapter presents an account of the procedures adopted for the interview, including, the analysis of data and ethical considerations. The following chapter introduces the findings of the first phase of the study.

# Chapter 4 An Analysis of Web Content: Departments & Research Groups

---

## 4.1 Context

Data is increasing in “volume, velocity and variety” the so called 3Vs (Laney, 2001; Boyd & Crawford, 2012). The increasing amount of research data, speed of data being created and varied forms of data are shifting how scholarship works. Not only is data increasing in volume, but also the ways that researchers access, manage, use and share data is changing. These changes around data have the potential to reshape how research can be done in a discipline, for example, increasing collaboration, creating very large collaborations, introducing new methods, new types of research and possibly altering funding sources (Graham & Shelton, 2013). New research questions may also arise from having data in digital form-and new interdisciplinary activity can be stimulated. The implications for research are potentially profound.

The changing of data is likely to be affecting departments, research groups and also individual researchers in complicated ways that are not yet understood. In addition, the effects of data changes are probably varied between different disciplines. It seems likely that the pre-existing nature and culture of a discipline may influence how changes in data might affect research. Research is needed to investigate the impact of these data changes in geography and how data practices shape and are shaped by discipline. It is for this reason essential to understand the nature of geography in order to understand the effect of changes in data on this subject.

The notion of a discipline is complex. According to Repko et al. (2013) and Chettiparamb (2007), discipline can be understood across three dimensions: epistemic, social and organizational. This chapter focuses on the organizational aspect of geography. The aim of this study is to examine how geography and its subfields are structured and organised within institutions. As schools, academic departments and research groups are basic academic units in universities, the study investigates the features of geography departments and research groups. In addition, a mapping exercise is performed to investigate the relationship between research groups and subfields and which are the dominant subfields of geography in the UK.

The outcome of the study helps answer the research sub-question (1): *‘What are the characteristics of geography as a discipline?’* particularly from an organisational perspective. Furthermore, combining with other parts of the study, the findings of this section contribute to identifying suitable

and interesting sub-fields for interviews and to inform interview questions. The intention is that the use of content analysis and bibliometrics in phase one will provide an insightful picture of geography as a discipline as a foundation for conducting interviews in phase two.

## 4.2 Data collection and analysis

The study described in this chapter examined websites of all UK-based geography departments (49) which participated in the Research Assessment Exercise (RAE) 2008 as part of the Geography and Environmental Studies unit of assessment, aiming to understand the current characteristics and organization of geography within institutions in the UK. Names of the 49 institutions were collected from the websites of RAE 2008 (<http://www.rae.ac.uk/results/>). A qualitative analysis of web content was carried out between December 2013 and February 2014. The process of this study was iterative, it moved between data collection, analysis, and interpretation. Yet, broadly it involved following steps. Firstly, the analysis consisted of reading the content of each departmental and research group website to gain a preliminary understanding on the characteristics of (1) geography departments and (2) geography research groups. Then, a mapping exercise of geography research groups and subfields, aiming to understand the relationship between them and the dominant sub-fields in geography. Research group names were collected from 49 institutions. coding categories, i.e. the sub-fields of both human and physical geography were derived from the literature (Cloke et al, 2013, Matthews & Herbert, 2008, Gregory, 2005) (See Appendix 7). Common subfields that are suggested from two sources were selected for the mapping exercise.

The sub-fields included in the initial mapping exercise were:

- Cultural geography
  - Development geography
  - Economic geography
  - Historical geography
  - Political geography
  - Rural geography
  - Social geography
  - Urban geography
  - Geomorphology
  - Geocryology
  - Hydrology
  - Biogeography
  - Climatology
- 
- The diagram uses blue brackets to group the sub-fields. The top seven sub-fields (Cultural, Development, Economic, Historical, Political, Rural, and Social geography) are grouped under the label 'Human geography'. The bottom five sub-fields (Urban, Geomorphology, Geocryology, Hydrology, and Biogeography) are grouped under the label 'Physical geography'. Climatology is listed at the bottom but is not grouped by a bracket.

Names of research group were imported into an Excel spreadsheet as row labels and selected subfields were added as column labels. By analysing research groups' descriptions on the web, the Excel table filled with 1 or 0 to indicate whether a research group has no relations [0] or positive relations [1] with subfields. Figure 4.1 shows a screenshot of the mapping table of human geography. From further analysis of research group websites it was apparent that the divisions between physical geography and human geography is strong, the mapping exercise was therefore also divided into human geography and physical geography. In addition, as a refinement of analysis, some initial subfield categories were updated and altered. For example, health geography, transport geography and quantitative geography were added to the initial categories, as these categories have appeared more than twice on research group websites throughout the analysis. Furthermore, in order to present data more clearly, visualisation tools were used such as TagCrowd (<http://tagcrowd.com>) to visualize word frequency for text and Gephi, an open source software, to visualise network graphs.

		Cultural	Social	Political	Economic	Urban	Development	Historical	Rural	Population	Health	Quantitative	Transpo
1													
2	Sustainable Development and Transformation	0	1	0	0	0	0	1	0	1	0	0	0
3	Cultural & Historical Geography	1	0	0	0	0	0	0	1	0	0	0	0
4	New Political Geographies	0	0	1	0	0	0	0	0	0	0	0	0
5	Communities and Social Identities	0	1	0	0	0	0	1	0	0	0	0	0
6	Society, Economy and Environment	1	1	1	1	1	1	0	0	0	0	0	0
7	Centre for Urban and Regional Studies	0	0	0	1	1	1	0	0	0	0	0	0
8	Historical and Cultural	1	0	0	0	0	0	0	1	0	0	0	0
9	Geographies of Political Economy	0	1	1	1	1	0	0	0	0	0	0	0
10	Personal Finance Research Centre	0	1	0	1	0	0	0	0	0	0	0	0
11	Spatial Modelling	0	0	0	0	0	0	0	0	0	0	0	1
12	Centre for Human Geography	1	1	1	1	1	0	0	0	0	0	0	0
13	Population, Health and Histories	1	1	0	0	0	0	0	1	0	1	0	0
14	Natures, Cultures, Knowledges	1	1	0	1	0	0	0	0	0	0	0	0
15	Contested Political Economies	0	0	1	1	1	0	0	0	0	0	0	0
16	Social Research	0	1	1	1	1	1	1	0	0	0	0	0
17	Pedagogic Research	0	0	0	0	0	0	0	0	0	0	0	1
18	Society Research	0	1	1	1	0	1	1	0	0	0	1	0
19	Culture-Economy-Life	1	1	0	1	0	0	0	0	0	0	0	0
20	Politics-State-Space	0	0	1	0	0	0	0	0	0	0	0	0
21	Geographies of Health and Wellbeing	0	0	0	0	0	0	0	0	0	0	1	0
22	Urban Worlds	0	0	0	0	1	0	0	0	0	0	0	0
23	Human geography	1	1	0	0	1	1	1	0	0	1	0	0
24	Environment & Society	0	0	0	0	0	0	0	0	0	0	0	0
25	Spatial Responsibilities	0	1	1	0	0	0	0	0	0	0	0	0
26	Geographies of Creativity and Knowledge	1	0	0	0	0	0	1	0	0	0	0	0

Figure 4.1 A screenshot of the mapping exercise

## 4.3 The Organisation of Geography at the Institutional Level

### 4.3.1 Schools, Institutes and departments

The findings show that geography as a discipline is generally found in the form of conventional disciplinary units, e.g. departments or schools within institutions. Of the 49 institutions which participated in RAE 2008 Geography and Environmental Studies, 80% of the institutions (39) have a department or school unit housing geography. It is clear that disciplinary divisions still play a

fundamental role in the organising structure of institutions. However, there are variations between institutions in how geography as a discipline exists in institutions, e.g. in some the disciplinary identity of geography is very strong, while in others geography is almost invisible as a discipline.

### *Single disciplinary unit*

An analysis of the websites shows that 26 institutions have a strong independent disciplinary identity for geography. Both teaching and research focus primarily on the discipline of geography and are organised within departments. They all have their own websites, detailing their research and teaching activities. Most of these 26 institutions have titles as ‘Department of Geography’, ‘Department of Geography and Environment’, ‘School of Geography and the Environment’, ‘School of Geography’ (Table 4.1). The title of the academic units seems reflects how a discipline is organised in the institutional context. Although these highly independent disciplinary characteristics appear among top, middle and low ranking institutions, the majority having these characteristic are in the higher ranking, e.g. the top 10 geography departments/schools. Of note is that most of the high ranking institutions are research intensive universities.

1. University of Cambridge	Department of Geography
2. University of Oxford	School of Geography and the Environment
3. University of Durham	Department of Geography
4. Queen Mary, University of London	School of Geography
5. University of Bristol	School of Geographical Sciences
6. University of Leeds	School of Geography
7. University of Sheffield	Department of Geography
8. King's College London	Department of Geography
9. University College London	Department of Geography
10. London School of Economics and Political Science	Department of Geography and Environment

Table 4.1 Top 10 geography departments/ schools

### *Closely related disciplinary unit*

There are 12 institutions where geography is organised with closely related disciplines. They have less strict disciplinary divisions. In these 13 institutions, geography is put with other disciplines into a department or merged into a bigger school (Table 4.2). For example:

Department of geography and planning in the University of Liverpool:

*“consists of two traditional disciplines: Spatial Planning (also known as Civic Design ..) and Geography (..).” (University of Liverpool, 2013).*

Newcastle University:

*“School of Geography, Politics and Sociology was formed in 2003, and brought together three formerly separate Departments.”(Newcastle University, 2013)*

University of Edinburgh School of GeoSciences:

*“was formed on 1 August 2002 by the merger of the former Institute of Ecology and Resource Management, the Department of Geography, the Department of Geology and Geophysics, and the Institute for Meteorology.”(University of Edinburgh, 2013)*

Rank	Institutions	Academic units	Disciplines
11	Aberystwyth University	Department of Geography and Earth Sciences	Geography, Geology
16	University of Manchester	School of Environment, Education and Development (formerly School of Environment and Development)	Architecture, Geography, Development Policy and Management, Education
19	University of Dundee	School of the Environment	Architecture, Geography, Environmental Science and Town & Regional Planning
22	University of Birmingham	School of Geography, Earth and Environmental Sciences	Geography, Geology
23	University of Hull	Department of Geography, Environment and Earth Sciences	Geography, Geology
24	University of Edinburgh	School of GeoSciences	Ecology, Geography, Geology, Geophysics, Meteorology, Oceanography and Environmental Sciences
26	Queen's University Belfast	School of Geography, Archaeology and Palaeoecology	Geography, Archaeology and Palaeoecology
25	University of Plymouth	School of Geography, Earth and Environmental Science	Chemistry, Earth Sciences, Environmental Sciences, Geography
27	University of Glasgow	School of Geographical and Earth Sciences	Geography, Earth Science/Geology, and Geomatics
28	University of Liverpool	Department of Geography and Planning	Spatial Planning and Geography
29	Newcastle university	School of Geography, Politics and Sociology	Geography, Politics and Sociology
37	Kingston University	School of Geography, Geology and Environment (teaching) / Centre for Earth and Environmental Science Research (Research)	Geography and Geology

*Table 4.2 12 Institutions where geography is structured with closely related disciplines*

Most of these 12 institutions are middle-ranking. From the analysis of web content, it is not possible to conclusively determine why these institutions merge disciplines into larger schools. However, one of the reasons could be to promote and facilitate interdisciplinary research across related disciplines, for example, by bringing together geology which is closely related to with physical geography, sociology and town planning are associated with human geography. As the University of Manchester web site states:



“SEED (School of Environment, Education and Development) is a working collaboration between five interrelated specialisms; as a school, we encourage each area to retain their own character while developing our combined, interdisciplinary presence.”

Although geography is put with other disciplines into a school, a lot of these institutions still have strict disciplinary divisions, a strong identity of geography as a discipline within the school, which can be noticed from the web content and the web structure (Figures 4.2 & 4.3).

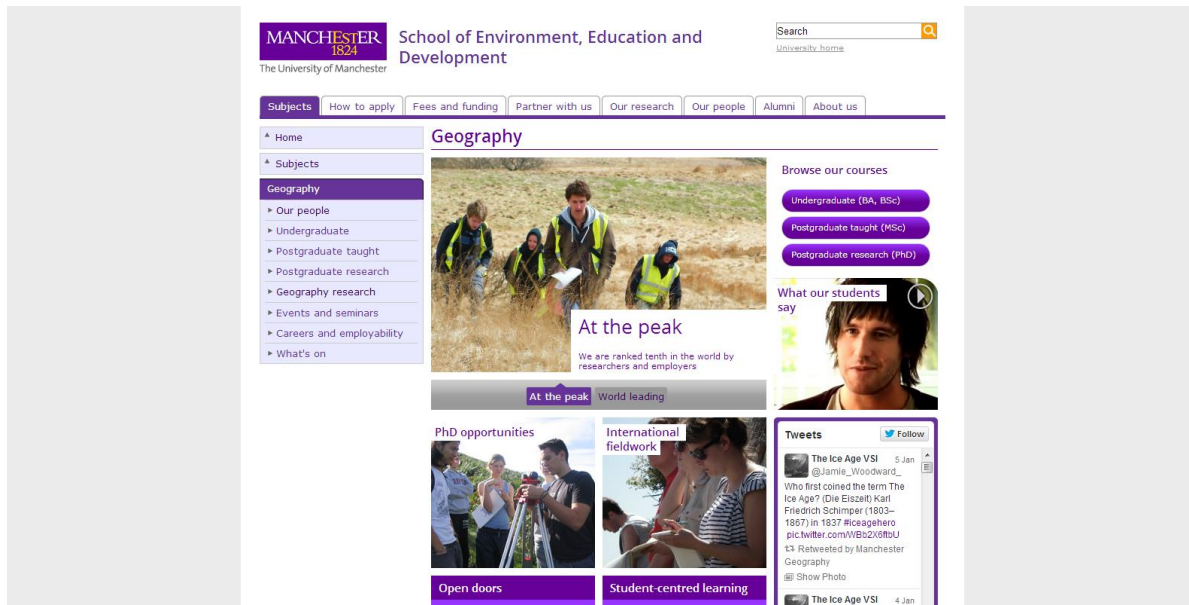


Figure 4.2 A screenshot of the School of Environment, Education and Development, University of Manchester



Figure 4.3 A screenshot of the School of Geography, Politics and Sociology, Newcastle University

However, some institutions which put geography and other closely related disciplines together, such as earth sciences/geology, seem to have less disciplinary division and are more integrated. For example, the University of Edinburgh School of Geosciences, University of Birmingham School of Geography, Earth and Environmental Sciences and University of Glasgow School of Geographical and Earth Sciences – although all these schools consist of geology and geography, it is difficult to distinguish the disciplinary division between geology and geography from their websites. The boundaries between geology and geography seem to be blurred in this institutional context. While some institutions in Table 4.1, such as, the University of Cambridge, the University of Oxford, and the University of Oxford, have geology (earth sciences) as a separate department.

#### *Broad disciplinary unit*

Of the 49 institutions which submitted in the RAE 2008 Geography and Environmental studies unit, there are 11 institutions where geography is difficult to be located as a discipline. None of them have an independent geography division within the institutions (Table 4.3). And most of them do not use the word “geography” in any academic unit’s title. Not all these institutions have courses in geography, although some offer geography related courses, such as, environmental science. Most of these 10 institutions are new teaching universities and low-ranking in RAE 2008 (Geography). There are two institutions where neither geography research nor geography education could be identified. The web study found that the organisation of these new institutions is a bit different from the traditional structure of universities which is department-based.

Most of these new institutions are school-based and do not have departments. It is probably because the size of these institutions is smaller. Disciplines are structured into multi-disciplinary schools, in a very broad disciplinary level, e.g. school of applied sciences, school of science and technology. Research activity of geography of these institutions is usually undertaken within a multi-disciplinary/inter-disciplinary research centre / group. These results are consistent with those of Hall et al (2015) who suggested that geography department are increasingly managed within multidisciplinary schools.

Rank	Institutions	Academic Units	Teaching areas	Research
30	University of Westminster	Policy Studies Institutes	(no geography course)	Policy Studies Institutes
35	Bournemouth University	School of Applied Sciences	BSc Applied Geography	Conservation Ecology and Environmental Sciences Group
36	Middlesex University	School of Science and Technology	(no geography course)	Flood Hazard Research Centre
37	Brunel University	Institute for the Environment, Centre for Human Geography	MSc Environmental Science	Centre for Human Geography Institute for the Environment
40	University of Salford	School of Environment & Life Sciences (Research)	Environmental Management	Ecosystems and Environment Research Centre
41	Anglia Ruskin University	Not found	-	-
43	University College Plymouth St Mark & St John	Not found	-	-
44	University of Gloucestershire*	School of Natural and Social Science	BSc Geography,	Centre for Environmental Change & Quaternary Research , Countryside and Community Research Institute
46	University of Worcester	Institute of Science and the Environment	Archaeology, Biology, Environmental Science, Nutritional Science, Geography, Horticulture, Forensic and Applied Biology, Animal Care	Centre for Rural Research, River Science Research Group
47	Bath Spa University (2005)	School of Society, Enterprise and Environment	Geography	Communities and Social Identities, Changing Landscapes Research Group
49	Swansea Metropolitan University	School of Built and Natural Environment (no geography department)	Environmental Conservation	Coastal and Marine Research Group

*Table 4.3 Institutions with broad disciplinary units*

This part of the study has shown that geography as a discipline is organised and managed in different ways in the institutional context. 49 institutions submitted geography research in RAE 2008, yet not every institution has a geography department. The organisational structure of geography is not straightforward at the institutional level. A discipline does not necessarily exist in the form of a department. Although, it should be noted that this study was conducted 5 years after the RAE, it is possible that some departments have disbanded or re-organised. The study has also shown that there are differences between top, middle and low ranking institutions in how geography is structured, e.g. some have more strong disciplinary identify, while some are more integrated with other disciplines. Thus, it is important that the organisational complexities of a discipline should not be overlooked when examining researchers' data practices.

### 4.3.2 Research groups

Moving from a departmental level to a finer level, geography research is commonly organised into research groups within departments. Different institutions use different terms for these groups, for example: “research groups”, “research clusters”, “thematic groups”, “research areas” or “research topics”. Although not all institutions have an independent geography department, all 49 institutions have research groups/clusters/ themes related to geography, except the two institutions in which neither geography research nor geography education could be identified. A total of 173 geography research groups were identified within the 49 institutions. The average is 3.5 groups per institution. The balance between physical and human geography is roughly even, the ratio is 1: 1.2. Most institutions have both. LSE, Open University and the University of Westminster are the only three institutions with only human geography; a few (5) have only physical geography. Institutions with higher ranking tend to have more research groups with a larger number of group members, while institutions in lower rankings have less groups and a less number of group members. For example, Oxford’s geography department have 5 research clusters, it’s ‘Climate Systems and Policy’ group have over 50 members and ‘Transformations: Economy, Society and Place’ group have 30 members, on the other hand, Edge Hill’s geography department only has one research group, which is ‘Environmental Processes and Change Group’ with only seven members.

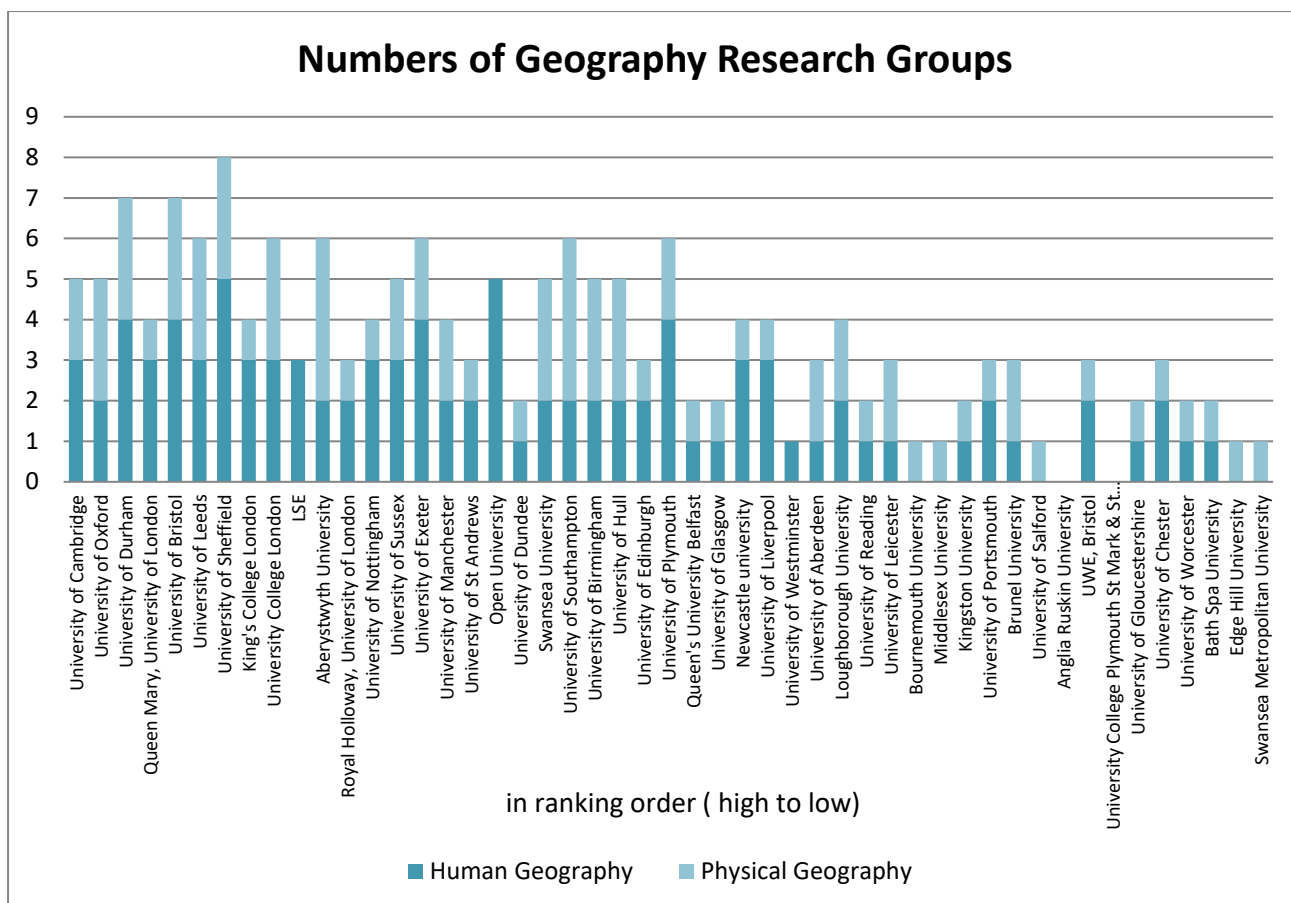


Figure 4.4 Numbers of geography research groups

#### The role of a research group

Several institutions' web sites pointed out the role and activities within geography research groups. Geography research groups are made up of a team of staff with shared interests and complementary expertise. As the geography department of the University of Durham suggested, research groups can act as 'ideas factories'. The purposes of research groups often seem to be to provide a forum to discuss, to coordinate current research, and to encourage collaboration within and outside the geography discipline.

As University of Exeter web site states,

*“We have a vibrant research environment, organized around 6 groups, which are the cornerstone of our research culture. These groups **act as hubs of activity by focusing and coordinating research activity, and by encouraging collaboration within and beyond the discipline.**” [Emphasis added by author]*

Similarly, the web site of Queen's University, Belfast says,

*“Environmental Change research and focus groups **provide a forum to encourage discussion of current research, promote collaboration and mentor younger members.** The focus groups meet informally, often with postgraduate students, to discuss papers, ideas and grant applications.”*

Similarly, Environmental Change Research Group (ECRG) of the University of St Andrew indicated some typical activities within a research group, such as, discuss research strategy, funding opportunities, share research experience, and make links with wider communities.

*“ECRG provides a regular forum for:*

- *Presentation and discussion of research problems, initiatives and results;*
- *Sharing of technical skills and research experience, opening the way to internal collaborative research and expanding the scope of individual research agendas;*
- *Discussion of strategy, appointments, technical provision and management of FEEA;*
- *Dissemination of funding opportunities and discussion of grant applications, research -management and knowledge exchange; and*
- *Interfacing with the Scottish Alliance for Geoscience, Environment and Society”*

### *Research groups & research strategy*

In order to find out the relation between research groups and the sub-fields of geography, a mapping exercise was carried out. UK Institutions have research groups covering major sub-fields of geography, such as, social geography, economic geography, cultural geography, urban geography, political geography, geomorphology, biogeography, hydrology, meteorology and climatology. Not every institution is active in research of all geography subfields. In general, institutions tend to focus on a few aspects of geography, for example, the geography department of the London School of Economics and Political Sciences emphasizes the social science aspects of geography. Departments have different research strategy and expertise. The organisation of group-based research could possibly reflect the research strategy of a department/school.

As University of Southampton stated,

*“Our research strategy fosters outstanding work in a set of five key areas, each represented by a research group: population, health, and wellbeing; globalisation, innovation and economic evolution; the observation of global environmental processes and change; the nature and implications of environmental change; and earth-surface dynamics.”*

London School of Economics and Political Sciences stated,

***“The Department's focus is on the social science aspects of geography. This is reflected in its research as it is in its teaching. Research is organised around three clusters of interest: Cities and Development, Economic Geography, Environmental Economics and Policy”***

### *Research groups and the branches of geography*

The web study found that the division between human and physical geography is still strong and clear at the institutional level. Yet, interestingly, the majority of geography departments did not label research groups as physical geography or human geography. Only a few departments specified the branch of geography that the research groups belonged to, such as, University of Exeter's and University of Hull's ( Figures 4.5 & 4.6).

#### Research groups

Exeter is a leading UK Geography Department and a centre for world-leading geographical research. We have a vibrant research environment, organized around 6 groups, which are the cornerstone of our research culture. These groups act as hubs of activity by focusing and coordinating research activity, and by encouraging collaboration within and beyond the discipline.

##### **Human Geography Research Groups:**

In human geography our research is established around the following 4 groups:

- **Geographies of Creativity and Knowledge (GeoCAK):** is a focal point for Exeter's internationally recognised excellence in cultural and historical geographies. The group's research positions Exeter at the leading edge of current geographical understandings of culture, matter, landscape, knowledge and creativity.
- **Spatial Responsibilities:** uses its international expertise in the geographies of justice, ethics and global responsibility to undertake agenda-setting research on the creation, performance, reconfiguration and impacts of key ethical and political spaces. Intellectually, the group develops accounts of these spaces through socio-cultural philosophies and methodologies to create more nuanced understandings of emergent spaces of ethics and responsibility.
- **Nature, Materialities and Biopolitics (NaMBIO):** a group of internationally recognized scholars whose research focuses on the geographies and politics of living and material systems, and who address key questions around geography, life science and bio-politics.
- **Environment and Sustainability:** a group comprises leading and emerging social scientists who undertake interdisciplinary research on frontier issues of environment and sustainability. The group bring spatial and geographical dimensions to research on sustainability theory and policy, and undertake cutting-edge research in distinctive areas such as place attachment, political economies of energy, climate justice and ecosystem services. The [Energy Policy Group](#), based at our Penryn Campus, is a core element of our wider environment and sustainability research and has a strong policy focus.

##### **Physical Geography Research Group:**

In physical geography our research is focused around the following 2 groups who work on past, present and future environmental change.

- **Environmental Change:** this group undertakes world-leading research, using cutting edge palaeoecological and palaeoclimatic analytical and dating approaches, to: a) document periods of major past climate and environmental change; b) to reconstruct and model glacial and ice-sheet dynamics over multiple Quaternary timescales; and c) to understand the effects of environmental change on contemporary terrestrial and marine environments.
- **Landscape and Ecosystem Dynamics:** this group builds on Exeter's established strengths in water, sediment and nutrient cycling, but integrates expertise in ecosystem functioning, carbon dynamics and climate modelling. A key focus is on the links between landscape processes and ecosystem responses to human activity and climate change.

*Figure 4.5 A screenshot from the University of Exeter geography department*



## Research themes

Our **human geography** research is divided into two themes:

- Environment and spaces of governance
- Knowledges, cultures and space

We separate our **physical geography** research into three themes:

- Sediment and flow dynamics
- System response to environmental change
- Aquatic systems and resources

Figure 4.6 A screenshot from the University of Hull geography

Yet research groups are clearly separated into human or physical geography groups by subject matter. For example, Figure 4.7 is the organisational structure of geography research at the University of Birmingham. It is clear that Environmental health science, Geosystems and Water Sciences belong to physical geography, whereas Society, Economy and Environment and Centre for Urban and Regional Studies belong to human geography. Only six (out of total of 173) research groups had cut across the traditional divisions between physical and human geography and brought physical geographers and human geographers together. It is clear that the human and physical geography division still exists.

### Environmental Health Sciences

Atmospheric Chemistry and Air Pollution | Environmental Nanoscience  
Meteorology and Climate | Persistent Organic Pollutants

### Geosystems

Palaeobiology | Palaeoclimates | Structure and properties of the Earth's subsurface

### Society, Economy and Environment

Diverse economies and communities | Energy, society and place  
Transitions and transformations | Urban resilience

### Water Sciences

Hydroecology (Ecohydrology) | Hydrogeology | Hydroclimatology | River Processes  
Urban Systems | Nuclear Decommissioning and Waste Management

### Centre for Urban and Regional Studies (CURS)

Strategic spatial planning | Economic development | Urban regeneration and resilience  
Place leadership

Figure 4.7 Research themes in the University of Birmingham

## Research groups and the sub-fields of geography

Research groups can often be closely related to a sub-field of a discipline. As shown in Table 4.4, research groups of the School of Geographical Sciences, University of Bristol are closely associated with established sub-fields of geography. For example, the Geographies of Political Economy group corresponds to political geography and economic geography, while the Bristol Glaciology Centre corresponds to glaciology, a subfield of physical geography. However, not every research group



exactly represents a sub-discipline. There is always a local decision about how research is organised at the institutional level. The relationship between subfields and research groups in human geography is complex. For example, Table 4.4 shows that most of the research groups are related to more than one subfields.

<b>Research Group Name</b>	Historical-Cultural Geographies	Geographies of Political Economy	Spatial Modelling	Personal Finance Research Centre	Hydrology Group	Bristol Research Initiative for the Dynamic Global Environment	Bristol Glaciology Centre
<b>Corresponding sub-fields of geography</b>	Historical geography, cultural geography	Political geography, Economic geography	Geographical Techniques	cross-cutting theme: social / economic geography	Hydrology	(Global Environmental Science)	Glaciology

Table 4.4 Geography research groups of the University of Bristol

### Human geography

Within the 49 institutions, 96 research groups were identified as related to human geography. Research groups in human geography appear to be easier to map to the sub-fields of geography than physical geography. This is because many human geography groups indicated their sub-fields of geography clearly through the group title. Figure 4.8 shows the word frequency of the research group titles in human geography. “Cultural”, “Social”, “Political”, “Development” which are associated with subfields, such as, cultural geography, social geography, political geography, development geography etc, are frequently used in the group titles.



Figure 4.8 Word frequency of the titles of human geography research groups

Some research group webpages also explicitly indicated sub-fields which the research groups are related to. For example,

The Population, Health & Histories group of the University of Cambridge stated,

*“The Population, Health & Histories thematic group has research interests that span a range of topics in **contemporary and historical demography, social geography, and cultural and historical geography.**”*

Similarly, the University of Durham stated,

*“The Culture-Economy-Life research cluster reflects the Department of Geography's commitment to generating dynamic and productive relations between **the fields of social, cultural and economic geography.**”*

The example above also shows that departmental research groups are commonly combinations of a number of subfields. This suggests that subfield boundaries are not neatly defined entities, they blur and overlap. As shown in Figure 4.9, 74% of the 96 human geography groups are related to more than one single subfield of geography. Research groups' grouping may reflect the research strategy of a department. Thus, combining different subfields together might indicate departments/ schools are increasingly promote collaboration between subfields through research groups.

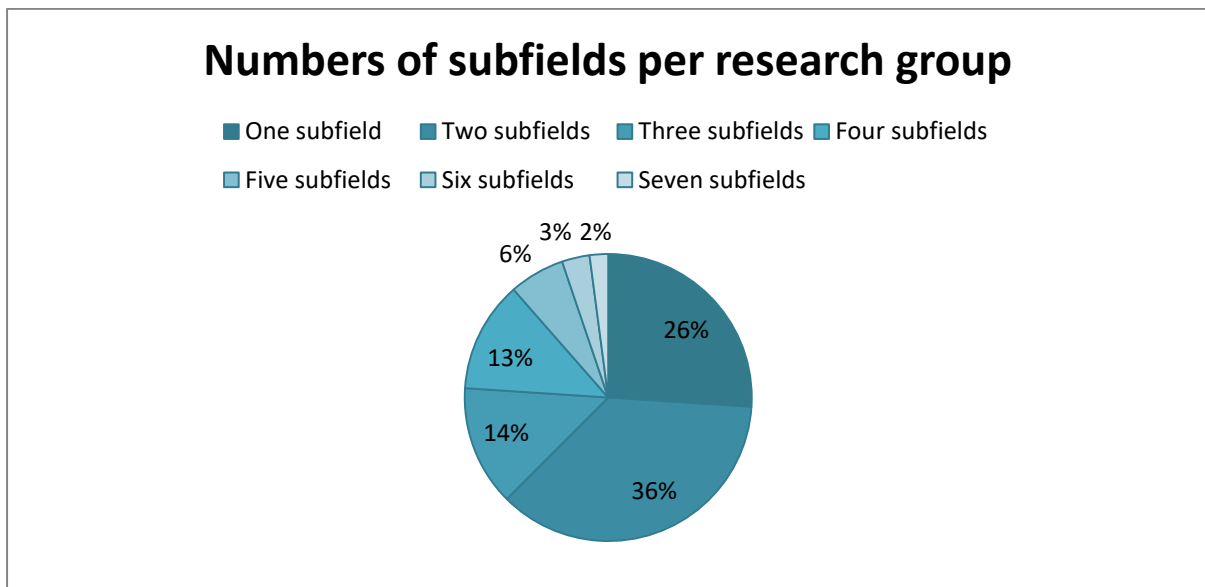


Figure 4.9 Numbers of subfields per research group

The study found that political geography, social geography and cultural geography are the most frequent subfields in the UK research groups (Figure 4.10).

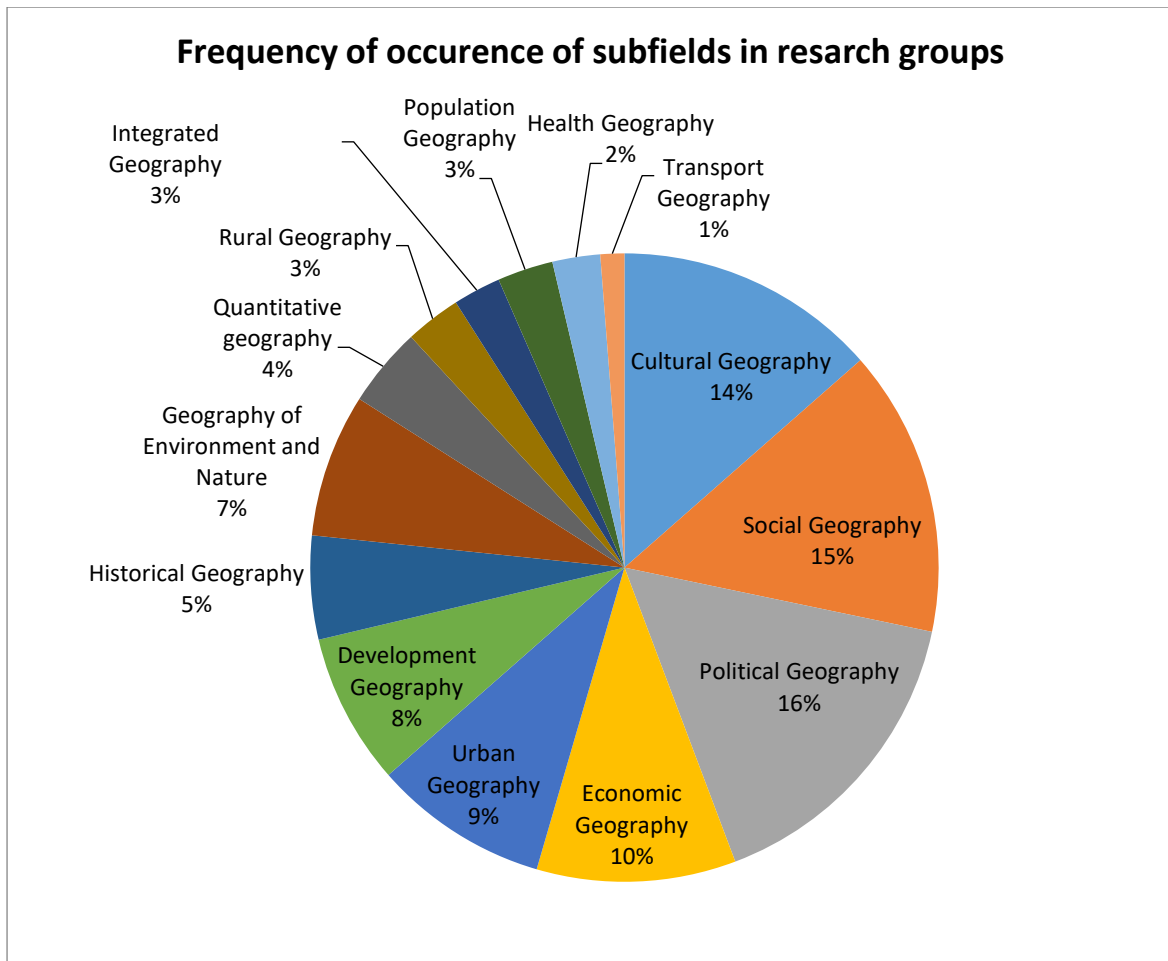


Figure 4.10 Frequency of occurrence of subfields in research groups

This finding seems to be consistent with Gregroy et al. (2009), who pointed out that economic, socio-cultural and political geography are the dominant broad divisions in academic geography. Yet, further analysis showed that these subfields have become more interconnected and integrated at the institutional level through research groups. Figure 4.11 presents the relationships between subfields in human geography. The web study finds that there is a difference in the relationship strength between subfields, as represented by the frequency in which they are clustered in research groups. The thickness of lines in the graph represent the frequency of occurrence of subfields in a same research group. e.g. historical geography is frequently linked to cultural geography in research groups while economic geography are frequently connected with political and social geography. Some subfields are rarely connected, for example, quantitative geography, rural geography and transport geography tend to be on their own. It is possible because these fields are more 'rural' (Becher & Trowler, 2001). They have a broader intellectual territory and have less overlapping concerns. This study also identified quantitative geography is an interesting subfields, using quantitative methods and big datasets in human geography, which can be highly relevant to this research.

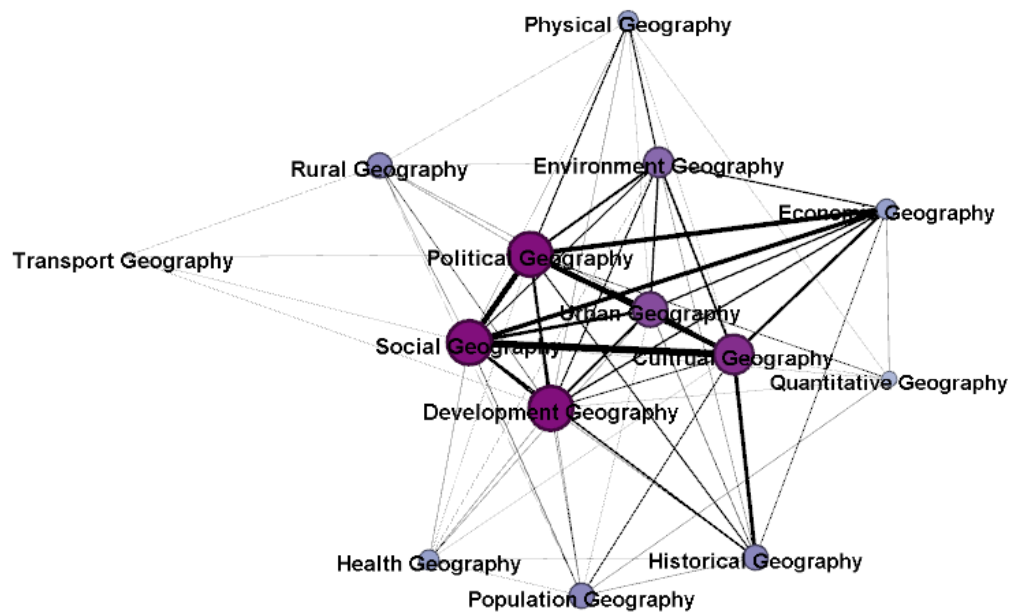


Figure 4.11 Relationships between subfields of human geography

The Science, Policy and Politics group from the geography department of the University College London stated,

***“Big data and applied geocomputation***

*Our shared understanding of large and complex datasets can make a vital contribution to the successful and practical application of science and the formulation of policy..... collection and analysis of big social media datasets is extending our understanding of daily activity patterns, extending geodemographic classifications far beyond night-time geographies of residence”*

*Physical Geography*

Within the 49 institutions, 77 out of 173 research groups (45%) were identified as related to physical geography. The average number of physical research groups per department is 1.6 . The number of research groups in physical geography is slightly less than human geography. Interestingly, in a quarter of the 49 institutions there is only one physical geography research group (17), in contrast, there are only 6 institutions that have only one had a single human geography group. It seems physical geography is less fragmented. The reason for this is not clear, but it might be related to the fewer numbers of physical geographers than human geographers, as pointed out in the RAE 2008 subject overview (Richards, 2008).

Physical geography research groups were more difficult to map onto subfields of physical geography than human geography. For example, in contrast to human geography, physical geography research groups did not indicate sub-fields of geography through the group title. Subfields such as geomorphology, climatology, biogeography, and soil geography etc., which described by Matthews & Herbert (2008), did not frequently appear in research group titles (Figure 4.12). It seems that physical geography research groups did not follow the conventional subfield divisions. As Gregory (2002) suggested physical geography research may be restructuring towards a more integrated physical geography.

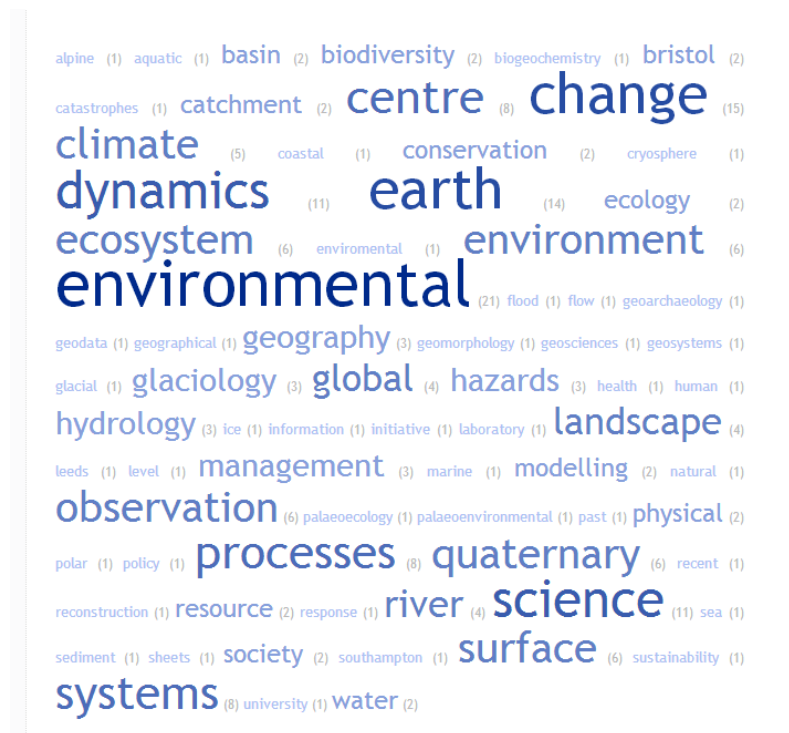


Figure 4.12 Word frequency of research group titles in physical geography

The web study found that institutions commonly integrated traditional sub-divisions of physical geography, such as climatology, geomorphology, biogeography, soil geography into one single research group. For example, as the department of geography at King’s College London stated,

*“Earth and Environmental Dynamics Research Group works to deepen the understanding of Earth’s hydrological, geomorphological, biophysical, atmospheric and ecological processes and their interactions”*

The University of Manchester web site also stated,

*“The Environmental Processes Research Group conducts pure and applied work research on the dynamics of contemporary earth surface processes. The group has expertise*

*in geomorphology, hydrology, freshwater environmental science, remote sensing, terrain analysis and geographical information science.”*

Another example is the department of geography at the University of Durham,

*“A key focus of the **Catchments and Rivers** cluster is on monitoring and modelling the interactions and feedbacks between **geomorphology, hydrology, vegetation, and water chemistry in hillslope and river systems.**”*

It is evident that departments are in support of integrated physical geography, which involves the study of whole geographic systems and processes. Thus all aspects of physical geography are brought together to recognise the need to integrate understanding of individual processes in a particular environment and landscape units.

Although the organisational structure of research groups in physical geography are not as closely associated with conventional subfield division as expected, some patterns were found. The following are some common topics that appeared in the physical geography research groups of UK institutions (Figure 4.13):

- Environmental systems/ processes/ change
- Earth Surface / Landscape
- Quaternary Science
- Ecosystems
- Glaciology
- Earth observation
- Rivers and catchments
- Climate systems
- Hazards
- Hydrology

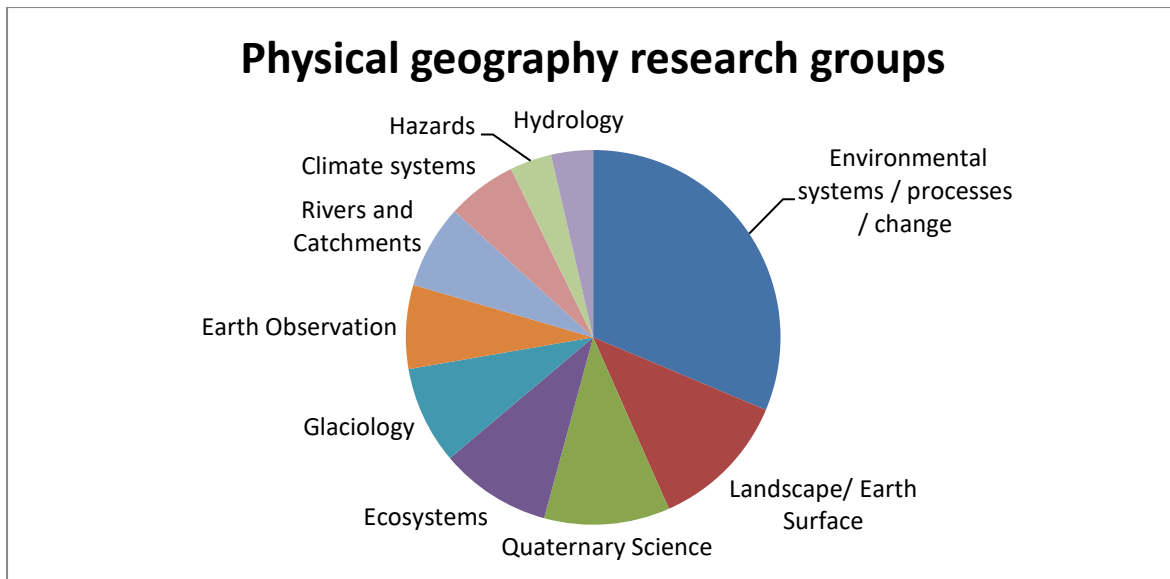


Figure 4.13 Physical geography research groups in the UK

Compared to Matthews and Herbert's (2008) suggestion about physical geography subfields (Figure 4.14), the present findings were unable to demonstrate the conventional physical geography subfield structure at research group level, however, it was broadly consistent with the interdisciplinary aspect of physical geography, e.g. global environmental change, landscape science, earth systems science and quaternary science.

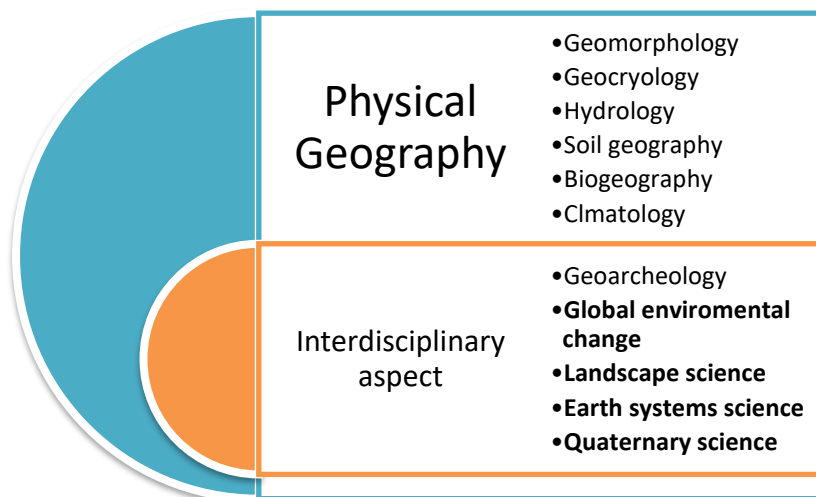


Figure 4.14 Sub-fields in physical geography (adapted from Matthews & Herbert, 2008)

### 4.3.3 Interdisciplinarity

More than half of the 49 institutions, especially the top and middle ranking ones, mentioned 'cross-disciplinary', 'interdisciplinary' or 'transdisciplinary' on their geography research webpages. (In lower rated departments, it possibly because departments are already likely mixed in with other subjects).

For instance, the geography department at King's College London stated,

*“our strategy is to strengthen this **inter-disciplinarity** further, working within and beyond the Department to address research questions of international significance, scale, uncertainty and complexity requiring ‘transdisciplinary’ perspectives, while at the same time, adding greater capacity in key **sub-disciplinary fields**”*

University of Cambridge stated,

*“we aspire for excellence through sustained **specialisation**, emergent exploration, and **interdisciplinary** collaboration within and beyond the department”*

Promoting interdisciplinary research is high on institutional agendas. Interestingly, the above two examples not only indicate their commitment to interdisciplinary research but also the prominence of specialisation and sub-disciplinary fields.

Previous sections showed that subfields are frequently put together into a research group, in both physical geography and human geography. There were some suggestions that a research group is united by a common topic or a theme rather than sub-disciplines. A number of institutions used terms such as ‘research theme’, ‘thematic research group’ to organise their research, instead of ‘research group’. For example, School of Geography, Earth and Environmental Science of the University of Birmingham used research themes to bring traditional sub-fields together:

*“Research in the School centres around five inter-disciplinary and multi-disciplinary research themes that cut across traditional discipline boundaries and respond to current international research agenda.”*

The use of the term ‘theme’ perhaps suggests that institutions are increasingly committed to interdisciplinary initiatives. Research themes enable researchers to explore connections across subfields and disciplines. A very few institutions also established cross-cutting themes which cut across different research groups within a department.

Geography research is not only undertaken by research groups/ research themes but also by other research units such as, research centres. This study found that nearly 50 % of the institutions have research centres alongside their research groups. The majority of these institutions are in the top and middle ranking. Table 4.5 shows the research centres in the top 5 geography departments.



University	Research Centre				
University of Cambridge- Department of Geography	Scott Polar Research Institute	Centre for Gender Studies	The Cambridge Group for the History of Population and Social Structure		
University of Oxford- School of Geography and the Environment	Environmental Change Institute	Smith School of Enterprise and the Environment	Transport Studies Unit		
University of Durham	Centre for Medical Humanities	Centre for Social Justice and Community Action	Institute of Hazard, Risk & Resilience (IHRR)	International Boundaries Research Unit (IBRU)	Sea Level Research Unit
Queen Mary, University of London	The Centre for Micromorphology	The Centre for Aquatic and Terrestrial Environments	The City Centre: researching city lives and connections	Centre for Global Security and Development	
University of Bristol	Cabot Institute	Politics and Matter Research Cluster	Coral Reef Research at Bristol		

Table 4.5 Research centres in the top 5 UK geography departments

Most of the research centres are organised under the geography departmental homes and separated from research groups. The major difference between research centres and research groups is research centre's research are cross departments and interdisciplinary in nature. As University of Sussex defined,

*"A University Research Centre is a large scale interdisciplinary grouping of academics undertaking research that is publicly recognised as world-leading in an established area of expertise. University Research Centres address significant international challenges and problems and redefine the parameters of one or more disciplines."*

#### 4.4 Summary and conclusion

The analysis of web content from 49 institutions which participated in RAE 2008 Geography and Environmental Studies provides an understanding of how geography as a discipline is structured and organised within institutions. Disciplinary and subfield structure plays an important role in the organisational structure of geography. More than half of the institutions have individual departments/schools with disciplinary labels such as, 'geography', 'geographical sciences', 'geography and environment'. However, there are clear differences between top, middle and low ranking institutions in how geography is structured. Top ranking institutions tend to have an individual geography department with strong independent disciplinary identity; middle ranking institutions tend to have departments/ schools which geography are joined with some closely related disciplines, e.g. geology, and in low ranking institutions, geography tends to be structured into multi-disciplinary schools, in a very broad disciplinary level, e.g. school of applied sciences, school of science and technology.

This study also explored the nature of research groups within institutions. All institutions have research groups related to geography. Some Institutions explicitly identified the purpose of research groups on their websites, e.g. providing a forum to discuss and coordinate research, encouraging collaboration within and outside the discipline. As not every institution is active in research of all geography subfields, the organisation of group-based research reflects the focus / research strategy of a department/school.

A mapping exercise was carried out to investigate the relationships between research groups and subfields. In general, this study has shown that research groups are closely related to a sub-field of a discipline. Yet, each research group does not represent a single sub-field of geography.

Departmental research groups are commonly combinations of a number of subfields. Some subfields are nearly always grouped with others, e.g. economic geography and political geography. On the other hand, some subfields tend to be on their own, mainly because they are relatively rare, e.g. transport geography, health geography. This study found that political geography, social geography and cultural geography are the most frequent subfields in human geography. Furthermore, an interesting subfield was identified, quantitative geography which applies quantitative methods and uses big datasets in human geography.

Compared to human geography, physical geography seems to be more integrated. Conventional sub-divisions of physical geography, such as climatology, hydrology, geomorphology, biogeography, soil geography commonly integrated into one single research group. Yet, this study identified some popular areas in physical geography research groups, including, environmental systems, processes and change, earth surface (landscape), quaternary Science and ecosystems. Furthermore, it was found that the division between human and physical geography is clear at the research group level. Even though most departments include both, only a very few groups, six out of total of 173, have both physical geographers and human geographers. Most of these research groups are related to the topic of environment.

It is also clear that interdisciplinarity has become an important agenda in institutions. This study has shown that institutions merge departments into bigger schools, cluster subfields into research groups, unite research groups by a theme rather than subject matter and establish interdisciplinary research centres. All these phenomena suggest that institutions are supporting interdisciplinary initiatives, encouraging collaborations and making connections within and outside disciplines and subfields. Yet, one of the findings to emerge from this study is that despite the increasing promotion of interdisciplinarity in the institutions, the divide between human geography and physical geography is still strong.

In conclusion, this web study investigated the organisational aspect of geography at an institutional level (Repko et al., 2013; Chettiparamb, 2007). It provided an overview of research structure and characteristics of geography within institutions. In addition, some key subfields of geography were identified. Academic departments and research groups provide institutional homes for researchers. The institutional structure of geography is likely to have implications for communications among researchers, which could affect researchers' data practices, such as, data sharing. For example, Parsons (2013) and Winn (2012) found that most researchers only share their data with researchers who create the data and researchers within their own research groups or departments. Combining the results of this section with material from the other studies in the following two sections, the results of the study are valuable in selecting suitable interviewees for in-depth interviews and informing interview questions. For example, the finding that there is a clear divide between physical and human geography suggests that interview participants should be sampled from both branches. And the findings of quantitative geography in human geography can be highly relevant to data changes. This section has also thrown up important questions in need of further investigation in the planned interviews, e.g. what is the meaning of research groups and subfields to researchers, what are their views on the divide between physical and human geography and how organisational structures affect their data practices?

Having examined geography at the level of departments and research groups, the study reported in the next section will be examining individual researchers. Looking at the individual level, the internal disciplinary characteristics of geography are more likely to be clear, i.e. the intellectual, e.g. research methods, research topics, and the social characteristics (collaborations and communications) (Becher & Trowler, 2001). Therefore, the next content analysis will examine geography researchers' web profiles, focussing on their research topics, research methods, publications and collaborations. The next study will continue to help answer research sub-questions (1) What are the characteristics of geography as a discipline, focussing on the intellectual characteristics of geography.

# Chapter 5 An Analysis of Web Content: Researchers' Profiles

---

## 5.1 Introduction

Each discipline has its own intellectual characteristics (Becher, 2001; Repko, 2008). Intellectual characteristics of a discipline are strongly related to data practices of researchers. As Repko (2008: 234) states a discipline “collects, organises and presents data in a way [that] is consistent with [the epistemology of a discipline]”. Thus, the work presented in this chapter focuses on the individual researcher level, aiming at answering research sub-question (1) ‘What are the characteristics of geography as a discipline?’ especially focusing on exploring the intellectual characteristics of geography, e.g. research topics and methods through an analysis of researchers’ web profiles.

## 5.2 Data collection and analysis

The sample of this part of the study consists of geography research group leaders (or the most senior staff if there is no indication of a group leader on the website) of the 15 out of 49 institutions which took part in the ‘geography and environmental studies’ unit of assessment in the Research Assessment Exercise 2008. Research group leaders (or the most senior staff) were chosen to be examined mainly because research group leaders and senior staff have extensive research experience and well developed research interests, their web profiles are more likely to have relatively more information, e.g. when comparing to PhD students’ profile.

The previous part of the study showed that there are differences between top, middle and low ranking institutions of how geography is organised. Thus, the 15 institutions chosen in this part of the study consist of 5 top, 5 middle and 5 low ranking institutions participating in the Research Assessment Exercise 2008 (Table 5.1). The locations of the institutions were also considered for practical reason when selecting the 15 institutions (in anticipation of interviewing for a later stage of the research).

Top	Middle	Low
2. University of Oxford	22. University of Birmingham	37. Kingston University
3. University of Durham	23. University of Hull	39. Brunel University
6. University of Leeds	28. University of Liverpool	45. University of Chester
7. University of Sheffield	29. University of Newcastle upon Tyne	46. University of Worcester
9. University College London	32. Loughborough University	48. Edge Hill University

Table 5.1 15 institutions that were chosen for this part of the study

The data collection for this part of the study took place from Nov 2013 to Dec 2013 and the qualitative analysis was mainly carried out between February 2014 and March 2014. The web profiles of research group leaders (or the most senior staff in a group) were accessed through the “members” page link on research group webpages which identified in the previous part of the web study. As with the previous part of the study, webpages of research group leaders or the most senior staff in the groups were captured as PDF using NCapture and then imported into Nvivo for coding.

### 5.3 Demographic results

The study examined a total of 65 geography research groups from the 15 institutions. The process of data collection found that different terms are used for research group leaders, such as ‘research group leader’, ‘cluster co-ordinator’, ‘cluster convenor’. Not all research groups (32) indicated their research group leaders on the webpages. Around half of the research groups (33) have no indication of their group leaders. Most groups that have indication have one leader per group. Only 3 research groups have co-leaders in a single research group.

Initially, of the 65 research groups studied, 38 were research group leaders and 30 were the most senior staff in the groups, most of them are professors and readers. Further examination found that 3 of the research group leaders are not in the field of geography (Table 5.2). As the previous part of the study showed that some institutions have departments which geography are joined with some closely related disciplines, e.g. geology, it was not surprised some research groups consist of researchers from different disciplines. As this study focuses on the nature of geography as a discipline, the 3 non-geography research group leaders are replaced by 3 most senior geographers of the groups. Overall, a total of 68 senior geography researchers were examined, in which 35 of them were research group leaders.

Research group	Title of the group leader	Research discipline of the leader	Research areas
Geosystem	Professor of Geology	Geology	Geology and geophysics- with interests in the use of geophysics to study geological problems, especially tectonics.
Environmental Health	Professor of Environmental Health	Environmental Chemistry	Air pollution, from emissions through atmospheric chemical and physical transformations to exposure and effects on human health
Visualities Research Group	Professor of Criminology	Sociology	cultural studies with the core concerns of criminology as a discipline

Table 5.2 Non-geography research group leaders

Out of the 68 senior researchers, the number of physical and human geographers is almost even (1:1.1), 33 physical geographers and 35 human geographers. This ratio is consistent with the balance between physical and human geography research groups which was showed in the previous part of the study. Interestingly, there is an imbalance between male and female geographers in this sample. Of the 68 senior geography researchers sampled in this study, 51 were male and only 17 were female. In addition, the number of females in physical geography in senior positions is considerably less than the number of females in human geography in senior positions (Figure 5.1). It seems that the role of research group leaders and senior staff in geography, particularly in physical geography, are dominated by males.

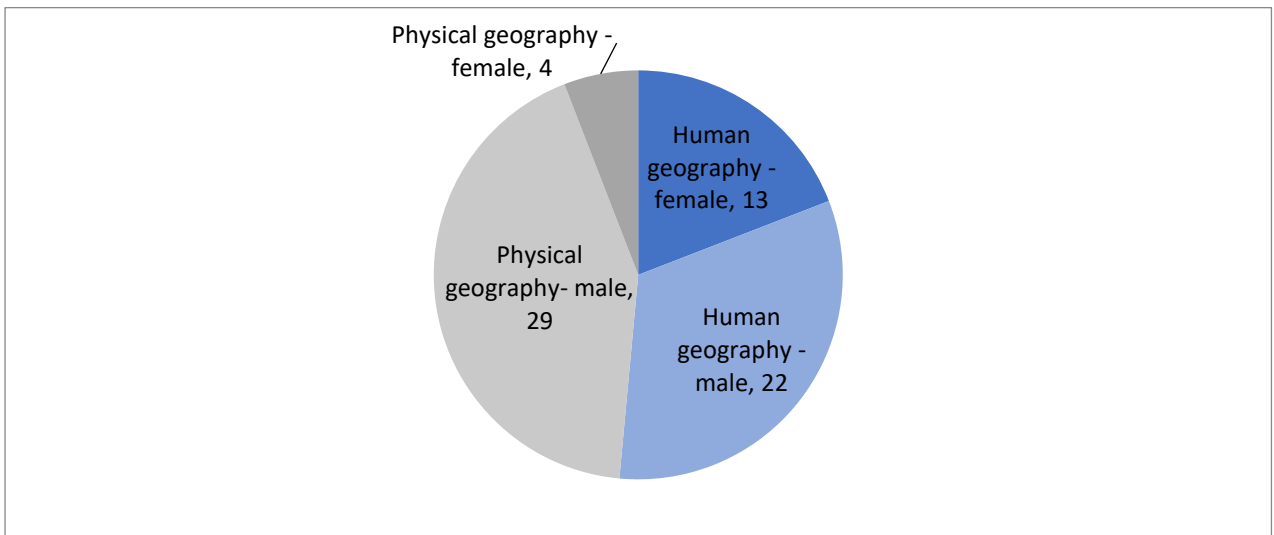


Figure 5.1 Gender of research group heads

## 5.4 Findings from the analysis of web profile content

### Overview

Researchers' web profiles reflect how researchers present and position themselves. Researchers' profiles also could reflect what researchers or institutions perceive as important within academic/disciplinary communities. The layouts of the 68 researchers' web profiles from the different institutions are in fact very similar. Although the amount of the web content varied between researchers, almost all of the researchers' web profiles consist of four parts. First, a short biography commonly includes qualifications, career history, achievements, internal and external responsibility and professional affiliations. Second, research interests, especially, their specialities, recent research projects and grants, occasionally methods, funders and collaborators were also mentioned. Third, teaching activities, including teaching interests, the modules they are involved in and postgraduate supervision. Finally, a publication list usually includes key publications of journal articles, books, presentations, and consultancy reports. These four elements appeared in all 68 researcher's web profile. This could suggest experiences, research, teaching, and publications are still the key elements of the value system within academic communities. In general, this analysis shows that apart from intellectual characteristics, such as research interests and methods which will be discussed later, the noticeable differences between the two branches in geography, human geography and physical geography, are publication format and key research funder.

### Publication format

Academic publication has always been important for researchers (Borgman, 2008). Of the 68 geography researchers, all of them have provided a publication list on their web pages. The analysis shows that the main publishing formats in geography are still traditional journal articles, books and reports. No data publication was mentioned on institutional web profiles. Evidently, physical geographers are less likely to publish research as a book or book chapters than human geographers. In this study, 28 human geographers have at least one book or a book chapter on the publication list, in contrast, only 7 physical geographers mentioned a book or a book chapter on their publication list (Table 5.3).

	Physical geography	Human geography
At least a book/ book chapters	7	28
Journal articles/ reports only	26	7

Table 5.3 Types of research output in physical and human geography

The findings are consistent with the report of the Excellence in Research for Australia (ERA) evaluation 2012, which also shows book and book chapter are more common in human geography than physical geography. The following charts illustrated the research outputs which were submitted to ERA 2012 by types in geography.

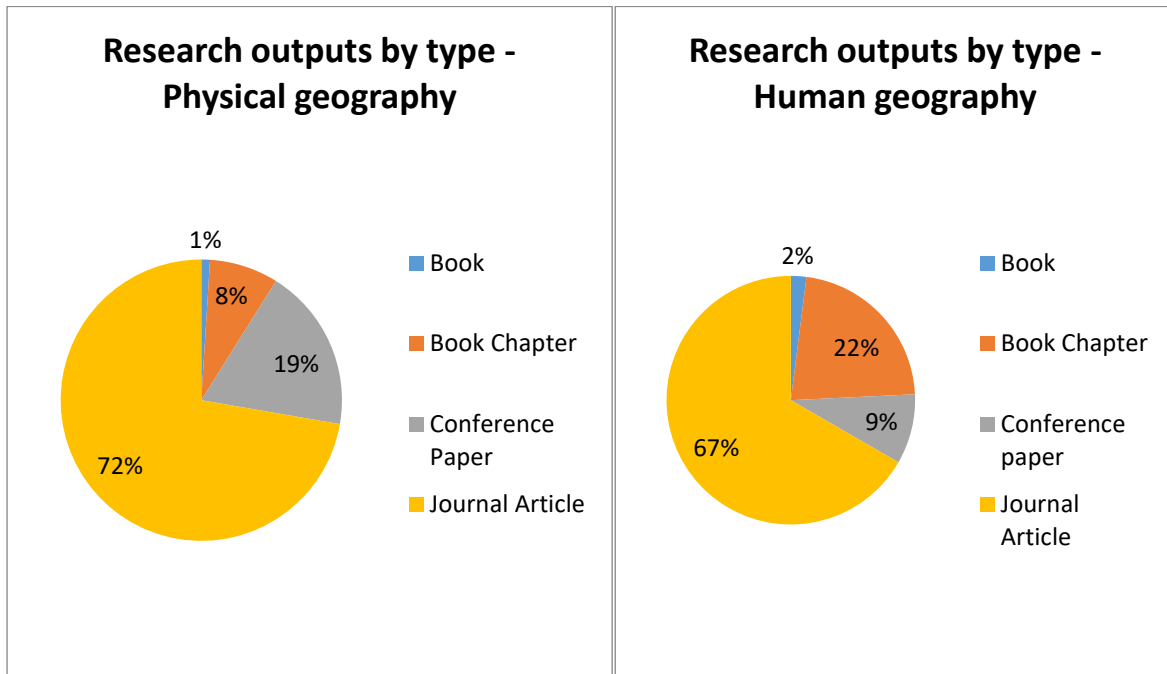


Figure 5.2 Research outputs which were submitted to ERA 2012 (Source: ERA 2012 report)

### Social media

Besides using traditional dissemination methods, such as publishing journal articles, books to communicate and share research ideas, some researchers also use personal research web blogs, twitter and videos (Figure 5.3). This study found that at least 7 geographers (1 human and 6 physical geographer) have personal research websites and 4 have twitter (Figure 5.4). In reality, there is probably a greater number of researchers use twitter. For example, this study found that some researchers do not mention twitter on their university web profile, yet they have twitter on their personal research website (Figure 5.5).

Furthermore, some researchers also mentioned media coverage of their research, embedded videos or included video and audio podcasts links on their web profiles, for example, an annual lecture video on YouTube. It seems that not only the way to disseminate research ideas is changing, but also the way researchers present themselves is changing. This could be influenced by research funders and the Research Excellence Framework (REF), which increasingly requested institutions and researchers to demonstrate the impact of their research, not only academic impact but also social and cultural impact (RCUK, 2011; Denicolo, 2013). In fact, the study found that researchers from a



top ranking institution have included research impact as a separate section on their web profile. The pressure and influence from research assessment and research funders cannot be underestimated. Ball (2014) suggests that as research funders' expectation on data sharing is increasing, the importance and impact of research data will be increasingly recognised by researchers, institutions and disciplines.

Environmental Sciences

- Research
- Distance Learning

Research in 60 seconds

- College Research Facilities

Dr Lee Chapman of the School of Geography, Earth and Environmental Sciences describes, in 60 seconds, his research in the impact of weather and climate on the built environment.

Print this page

Download as PDF

Send to a friend

Related Links

Research in the School of Biosciences

Research in 60 seconds - video gallery

I'm Dr Lee Chapman, a reader in climate resilience at the University of Birmingham and I am interested in the impact of weather and climate on the built environment. Cities are remarkable concentrations of both people and critical infrastructure which has evolved to serve the growing population. I specialise in improving the resilience of that infrastructure to extreme weather events such as heatwaves and flooding, so that society can continue to function effectively - whatever the weather.

My research team works with industry to monitor and adapt our energy, transport and ICT

Figure 5.3 Research in 60 seconds video

News and events

Research

Schools and Community

Alumni

Contact us

Vacancies

University of Hull, Cottingham Road, Hull, HU6 7RX

Tel: +44(0)1482 465343

Fax: +44(0)1482 466340

Email: D.Parsons@hull.ac.uk

Profile Teaching Research PhD Supervision Publications Twitter Projects

Tweets

Bondholders @hullbondholders

The UK's Energy Estuary in numbers

@Siemens\_Energy @greenporthull @SiemensUKNews pic.twitter.com/hxd5AoSfXo

Retweeted by Dan Parsons

Our Energy THE HUMBER IS A VITAL ENERGY PROVIDER

We support 1/3 of the UK's electricity supply and 2/3 of the UK's gas supply

We have 1/5 of the UK's gas supply

We have 1/3 of the UK's gas supply

Figure 5.4 Twitter on institutional web profile

Hello! I am a Professor of Physical Geography, specialising in fluvial geomorphology, at the Geography Department, University of Hull, U.K.

I am also Managing Editor & Co-Founder of **Earth Surface Dynamics**, an Open Access Geomorphology Journal. Links to publications and software are in the boxes at the top

Tweets

Tom Coultard @Tom\_Coultard

It must be a slow news day... BBC News - First bumps 'cleaner than handshakes' - bbc.in/Uw8CZE

Tom Coultard @Tom\_Coultard

Plotting the arrival of Geomorphology (I didn't know about the pre-war rise!) books.google.com/ngrams/graph?c...

Tom Coultard @Tom\_Coultard

Excellent Operational Surface water flood forecast model goes live for Glasgow: floodforecastingservice.net @NERC\_FFIR Expand

My research is focused (though not exclusively) in the following areas:

**Modelling fluvial geomorphology**

I am the developer of **CAESAR** and **CAESAR-Lisflood** landscape evolution models (Coultard et al., 2013). Caesar Lisflood is a geomorphological / Landscape evolution model that combines the Lisflood-2d hydrodynamic flow model (Bates et al, 2010) with the CAESAR geomorphic model to simulate erosion and deposition in river catchments and reaches over time scales from hours to 1000's of years. For more details click the tab at the page top

**Non linear processes in Geomorphology and earth systems:**

Geomorphic systems are often non-linear and chaotic and I have a long standing interest in how this reflects in the sedimentary record and basin behaviour. This includes system 'shedding' and Self Organised Criticality. Recent examples of this include publications in *Geology* and *ESurf*.

**Modelling continental scale palaeohydrology**

Our recent [PLOS](#) publication on North African Palaeohydrology showed how lining Earth System Model (ESM) outputs with hydrological and hydraulic models could be used to reconstruct Quaternary hydrology at continental scales. In this case we used it to reconstruct possible paths for human migration out of Africa

Figure 5.5 Personal research website

## Research Funders in geography

Research funders are a key part of the academic system. Of the 68 geographers, 70% of geographers (54= 27 human geographers + 27 physical geographers) mentioned their research funders. For example, the following are some quotes from the researchers' web profiles:

*"[My] research has received funding from the British Academy and the Economic and Social Research Council (ESRC)."*

*"My research in Botswana has been funded by the ESRC and RGS (HSBC grant), and I was recently currently Principal Investigator on a DFID funded research project in Namibia."*

*"[I have] been PI or Co-I on grants worth over £43 million (including 25 NERC, 2 EPSRC, 2 DFID, 2 Carbon Trust, 2 ESA, 3 Technology Strategy Board, Royal Society and DECC)."*

The study found that the key research funder mentioned on researchers' web profile is different between human and physical geography (Table 5.4). Economic and Social Research Council (ESRC) is the most mentioned human geography research funder. Of the 27 human geographers mentioned their research funders, 23 stated ESRC on their web profiles. Others funders mentioned from human geographers included, British Academy (mentioned by 9 human geographers), Engineering and Physical Sciences Research Council (EPSRC) (6), Leverhulme Trust (4), Joseph Rowntree Foundation (3), Nuffield Foundation (2), Wellcome trust (2), the European Commission (2), National Science Foundation (2), UK Office for the Deputy Prime Minister, (2) and university grants.

Human geography	Physical geography
<ul style="list-style-type: none"> <li>Economic and Social Research Council (23)</li> </ul>	<ul style="list-style-type: none"> <li>Natural Environment Research Council (24)</li> </ul>
<ul style="list-style-type: none"> <li>British Academy (9)</li> </ul>	<ul style="list-style-type: none"> <li>Royal Society (6)</li> </ul>
<ul style="list-style-type: none"> <li>Engineering and Physical Sciences Research Council (7)</li> </ul>	<ul style="list-style-type: none"> <li>Leverhulme Trust (6)</li> </ul>
<ul style="list-style-type: none"> <li>Leverhulme Trust (4)</li> </ul>	<ul style="list-style-type: none"> <li>Environment Agency (5)</li> </ul>
<ul style="list-style-type: none"> <li>Joseph Rowntree Foundation (3)</li> </ul>	<ul style="list-style-type: none"> <li>EPSRC (4)</li> </ul>
<ul style="list-style-type: none"> <li>Nuffield Foundation (2)</li> </ul>	<ul style="list-style-type: none"> <li>DFID (3)</li> </ul>
<ul style="list-style-type: none"> <li>Wellcome trust (2)</li> </ul>	<ul style="list-style-type: none"> <li>British Council (3)</li> </ul>
<ul style="list-style-type: none"> <li>European Commission (2)</li> </ul>	<ul style="list-style-type: none"> <li>EU (3),</li> </ul>
<ul style="list-style-type: none"> <li>National Science Foundation (2)</li> </ul>	<ul style="list-style-type: none"> <li>Tyndall Centre for Climate Change (2),</li> </ul>
<ul style="list-style-type: none"> <li>UK Office for the Deputy Prime Minister (2)</li> </ul>	<ul style="list-style-type: none"> <li>British Academy (2)</li> </ul>
	<ul style="list-style-type: none"> <li>ESRC (2)</li> </ul>
	<ul style="list-style-type: none"> <li>Nuffield Foundation (2)</li> </ul>
	<ul style="list-style-type: none"> <li>FP7 (2)</li> </ul>

Table 5.4 Common research funders mentioned on web profiles of human and physical senior geographers

In contrast to human geography, physical geography research is commonly funded by Natural Environment Research Council (NERC). Of the 27 physical geographers, 24 mentioned NERC on their web pages. Other funders mentioned are included Royal Society (6), Leverhulme Trust (6), Environment Agency (5), British Council (3), DFID (3), EPSRC (4), EU (3), Tyndall Centre for Climate Change (2), British Academy (2), ESRC (2), Nuffield Foundation (2), FP7 (2) and university grants. Different research funders would have different policy on research data. Thus, how funders' demand for good RDM and open data affect the two branches of geography will be explored in the interviews.

## 5.5 Intellectual characteristics

### Research topics

The study of the researchers' web profiles provides an understanding of the intellectual aspect of geography as a discipline, particularly at individual level. All 68 researchers included a statement of research interests on their web profile. Their research interests either fall into human geography or physical geography. In general, both human geographers and physical geographers indicated more than one area of research interest. Most of the geographers have few areas of research interests. For example, the following shows senior researchers listed their areas of research interests on their web profiles:

*"[My] research interests include modelling societies and their behaviour, the use of information and communication technologies in the public, industrial and commercial sectors, and the development of Geographical Information Systems for human and environmental analysis."*

*" Research interests:*

- *Speleothems and Climate Change, especially the development of palaeoclimate proxies and the understanding of karst processes including their hydrology. See also [www.speleothemscience.info](http://www.speleothemscience.info)*
- *Glaciation and carbonates in deep time - The Cryogenian. See GAINS project*
- *Aqueous Geochemistry in relation to weathering reactions and hydrology in glacial and riverine environments (Iceland, Himalayas).*
- *Experimental studies of mineral-water interactions"*

A mapping exercise, focusing on the individual level, was carried out to explore the relationship between subfields and researchers' interests/ research topics. The previous part of the study suggested subfields are usually grouped with others in the research group level research. Thus, this mapping exercise investigated whether researchers' research interests related to one or more than one geography subfield.

## Human geography

Generally, the five major subfields which related to the 35 human geographers are political geography, social geography, economic geography, urban geography and development geography (Figure 5.6). The result is broadly consistent with the previous results of the frequency of occurrence of subfields in research groups (Figure 4.10).

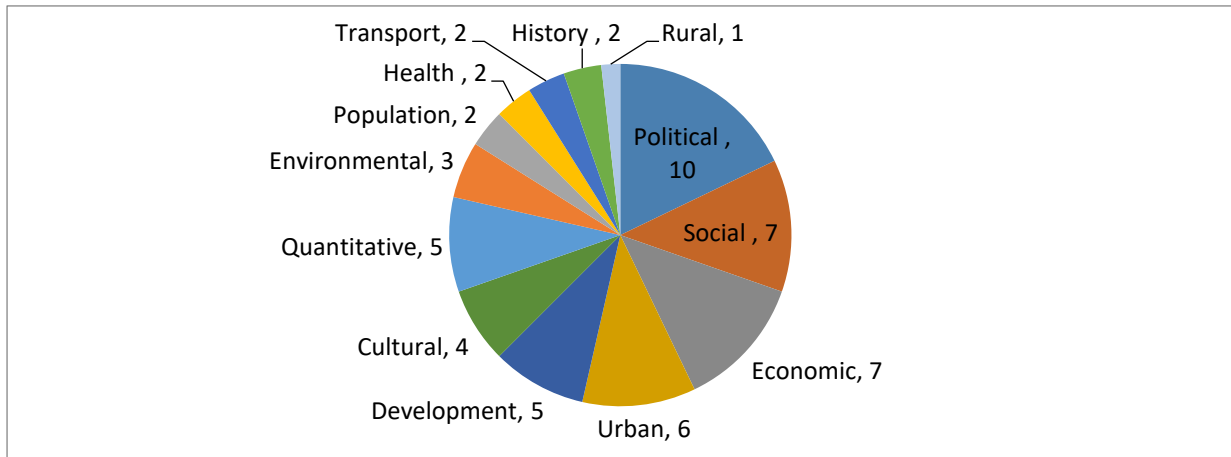


Figure 5.6 Subfields of the human geographers

Human geographers and their research do not always fit neatly into a single subfield of geography. The mapping exercise found that some geographers were involved in two research groups. In addition, around half (49%) (17 out of 35) of the human geographers' research interests are related to more than one subfield of geography. For instance, the following quotes from the web profiles of human geographers illustrate geographers have research interests related to several subfields:

*"I am a **social and cultural geographer**, with twin interests in 'Geographies of Children, Youth and Families' and 'Geographies of Social Difference'."*

*"I have research interests in **cultural and historical geography**, and in the histories of geographical knowledge"*

*"Tim's research can be positioned at the intersection of **urban, transport, cultural and political and economic geography**"*

*"My research covers a range of topics in **urban, political and economic geography**. I am particularly interested in the new regionalism, the urban politics of transition to a 'low carbon' economy, collective provision in new economic spaces, social enterprise and alternative spaces of finance."*

*“I teach and research in the fields of **cultural and historical geography**, I also have relatively unique research interests in the geographies of Italy and the former Italian empire.”*

These examples not only show human geography research is quite commonly related to more than one geography subfield. It also shows terms of traditional subfields, such as social geography, economic geography, are still in use. However, not all geographers use these traditional subfield terms. 54% (19 out of 35) human geographers did not use traditional subfield terms on their web profiles. Yet, less traditional terms, such as, Military geographies, Carceral geography, Geographies of Children, Geographies of ageing, Geographies of well-being, Geographies of airspace use and control, Geographies of cities and urban spaces, Geographies of gender and Feminist geography, are found on the human geographer web profiles. It is clear that these terms which were used for describing their own research are more focused and theme-based. As research topics of geographers are inherently narrower than subfields, these terms could probably be understood as specialities / specialisms of geography.

The finding seems to corroborate the ideas of Matthews & Herbert (2008), who suggested that traditional systemic geography subfields, such as social, cultural geography, are gradually substituted by more theme-based and integrated approaches. However, the findings suggest something more fundamental, which is the dynamic nature of a discipline. Klein (1996) and Dogan & Pahre (1990) used ‘specialization - fragmentation - hybridization’ to describe the changing processes of disciplines. This study shows geography subfields are not static. Human geography subfields, such as social geography, political geography, are constantly searching for ideas and concepts outside the discipline. They sub-divide, fragment and become new specialisms, e.g. geographies of children, military geographies. As these specialisms usually borrow concepts from other fields, they are hybrid in nature, and at the intersection of different subfields. For example, one of the geographer’s research area is military geographies, which including the politics of military land use (i.e. political aspect) and military identities and gender and the armed forces (i.e. social and cultural aspect). This shows military geographies can be connected to both political geography and social geography. In general, this study has found that not all human geographers and their research fit neatly into one subfield of geography. Another major finding was that non traditional terms such as carceral geography, geographies of children, are emerging. Rich range of topics and subject matters are introduced into human geography. All these could suggest human geography is constantly changing and evolving.

## Physical geography

Research topics and subject matter in physical geography are different from human geography. Comparing the research interests between human geographers and physical geographers, physical geography research is concerned with the surface of the earth, while human geography is concerned with people and communities. Interestingly, physical geography seems to be less diverse than human geography. Most physical geography research can be related to Biogeography & ecological processes; Geomorphology; Past environmental change; Hydrology; Climatology and Glaciology (Figure 5.7). Traditional subfield terms, such as hydrology, ecology, geomorphology, are still common in use. 17 out of 33 (52 %) senior physical geography researchers used subfield terms on their webpages.

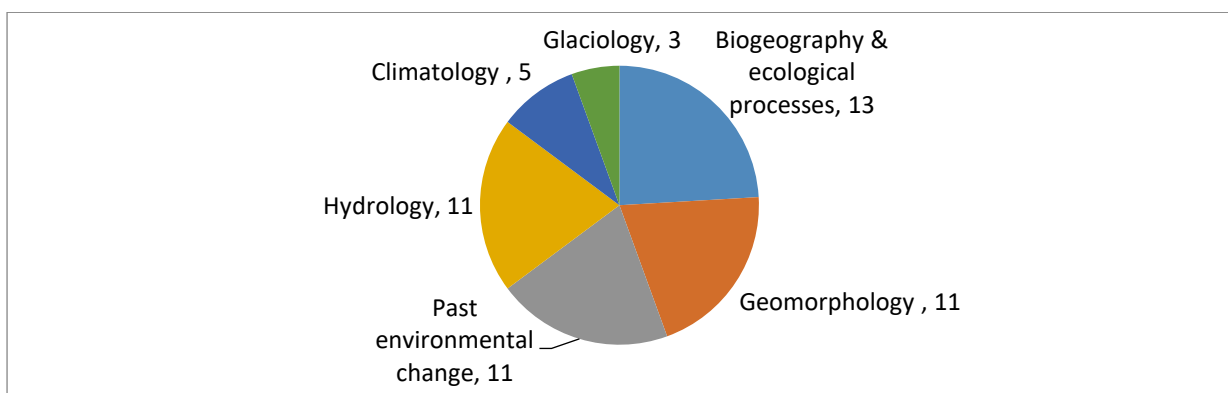


Figure 5.7 Subfields of physical geography

As with human geography, the research of physical geographers can be characterised by more than one subfield. For example, as one of the physical geographers' web profile states,

*“My research draws together elements of hydrology, ecology, geomorphology, environmental management and the conservation of freshwater ecosystems.”*

The analysis found that more than half (58%, 19 out of 33) of physical geographers' research interests can fit into more than one subfield. For example, one of the physical geography researchers is interested in the Interactions between geomorphology, hydrology and ecology in rivers. Furthermore, hybrid terms such as, hydroclimatology (interface between hydrology-climatology), hydroecologist, glacier hydrology, ecohydrology, ecogeomorphology are found on web profiles. These hybrid terms probably are indications of the changing nature of disciplines, i.e. the fragmentation of physical geography. However, compared with the fragmentation of human geography, such as, military geographies, carceral geography, geographies of children, geographies of ageing etc, the fragmentation of physical geography seems relatively less fragmented and diverse.

## 5.6 Methods in geography

The web profile analysis found that all of the 68 senior geography researchers did not have detailed descriptions of research methods they used. Yet, 49 (20 human + 29 physical) out of 68 senior researchers had mentioned methods of their research very briefly, e.g. when they described their research projects. As one of the human geographer web profiles stated:

*“..this project used **archival data and semi-structured interviews** to explore the place-specific nature of racialisation, and the ways in which processes of racialisation can produce highly spatialised understandings of difference in both an historical and a contemporary context.*

The content analysis found that there are differences between research methods in human geography and physical geography. Popular methods used in human geography research not only included qualitative methods, such as interviews, ethnography, visual methods, but also quantitative methods, such as, statistical analysis, spatial data analysis, GIS. The following list showed some research methods mentioned on the web profile of human geographers.

- Semi-structured interviews
- Ethnographic field work
- Oral narratives
- Visual methods ( e.g. photography, printed media)
- Use of archival data
- Iconography
- Policy analysis
- Multivariate statistical analysis
- Spatial data analysis
- Geographical Information Systems (GIS);
- Microsimulation
- Geodemographics
- Discrete choice modelling
- Mixed methods (qualitative and quantitative, GIS and spatial epidemiology, participatory GIS (field work))

On the other hand, physical geography methods are principally quantitative. The common methods used in physical geography are modelling, laboratory experiments and field-based work. The following list is research methods mentioned on the web profile of physical geographers.

- Qualitative mathematical approaches
- Numerical modelling (e.g. computer simulation)

- Physical Modelling
- Remote sensing (e.g. satellite imagery)
- Field based studies e.g. field observations, field measurements
- Field experiments
- Field spectroscopy
- Laboratory-based experimental work
- Statistical analysis of available data sets
- Luminescence dating
- Longitudinal research
- Environmental magnetism and pollen analysis
- Use of historical maps and aerial photographs
- Data–model comparisons
- Combines fieldwork, laboratory experimentation and numerical modelling

Research methods for geography research are diverse. The diversity of methods used in geography implies there are vast variety types of data.

## 5.7 Conclusion

In conclusion, this part of the study explored some characteristics of geography as a discipline which include the intellectual aspect, at an individual level. It provided some insights not only on research topics and methods but also publication types and key funders in geography. However, information about research methods, research data and collaboration pattern of geography are particularly limited in this web profile study. A bibliometric study will be used to investigate the social aspects of geography, e.g. collaboration patterns. The next study will continue to help answer research sub-question (1) What are the characteristics of geography as a discipline, focusing on the social aspect.



# Chapter 6 Bibliometric analysis

---

## 6.1 Introduction

Chapter six looks at research collaboration in geography through reporting the findings of a bibliometric study. The complex nature of disciplines can not only be understood from an organisational or intellectual perspective, but also from a social perspective e.g. communication and collaboration between researchers (Becher, 2001). Research collaboration is a central part of research processes (Day, 2008). How researchers collaborate has implications for their data practices, especially their data sharing practices i.e. the willingness to share data and with whom they are willing to share data. Thus, this chapter focuses on research collaboration patterns in geography (as represented by bibliometric data), aiming at answering the following research sub-question: What are the characteristics of geography as a discipline, particularly the social characteristics?

## 6.2 Bibliometric study—approach taken and techniques used

A small scale bibliometric study was conducted to investigate research collaboration patterns in geography. The study examined publications of the 68 senior researchers who were previously identified in the content analysis of researchers' web profiles. 'Senior researchers' are those academics who have been identified as either being a research group leader or the most senior member of a research group in an academic department or school in one of the 15 HEIs included in the web-based content analysis (see chapter 5). The publication year of outputs<sup>2</sup> examined were selected to map onto the most recent Research Excellence Framework (REF) period, 2008-2013. The study was conducted in two parts. The first part examined the number of co-authors per output focusing on the scale of collaborations in geography, as represented by the number of co-authors. The second part examined co-authors of the 68 senior researchers in further detail, e.g. the institutional affiliation and country associated with co-authors, so that the characteristics of collaborations in geography, such as, intra and inter-institutional collaboration; cross-sector; geography/ non-geography and UK/ international collaboration could be understood.

---

<sup>2</sup> Research outputs were retrieved from *Scopus*. Outputs such as, articles, book chapters, articles in press, conference papers, reviews and editorials were included in the analysis.

### 6.2.1 Sources of data

The *Scopus* database was selected as the data source for the bibliometric study. There were two main reasons for using *Scopus*. Firstly, *Scopus* tends to have a relatively good coverage of social science journals. According to Norris & Oppenheim (2007), *Scopus* has a better coverage of the social sciences than the *Web of Science*. Secondly, *Scopus* has more suitable searching tools and advanced functions relevant to this study than *Web of Science*. For example, *Scopus* allows for an author search using full first names, rather than initials only. Thus, it was easier to identify the correct author. Furthermore, detailed information from the database can be directly downloaded in excel format, so data can be collected efficiently.

Every data source has its own limitations. During the process of data collection, it was discovered that not all publication outputs were included in *Scopus*, when compared to some of the senior researchers' publication lists on their institutional websites. For example, grey literature such as reports are not usually included in *Scopus*. Thus, other sources of data for the bibliometric study were also considered, including *Web of Science*, publication lists on researchers' web profiles and the *International Bibliography of the Social Sciences (IBSS)* database. A small sample test was carried out to compare results from these sources; 4 geographers (2 human geographers and 2 physical geographers) were selected, and searches conducted in *Web of Science*, *IBSS*, and their institutional web sites. It provided comprehensive results and has relatively good coverage of both human and physical geography. For example, for publication lists on researchers' web profiles, not all universities have a complete publication list of researchers available on their website. The previous web content analysis found that some senior geographers only put key publications or selected publications on their university web page. *The Web of Science* had similar results to *Scopus*. Yet, it had less suitable tools and functions for the purpose of the bibliometric study, as mentioned earlier. Whereas *IBSS* had relatively less results than *Scopus*, yet occasionally had different results, e.g. when using an identical search strategy in both *Scopus* and *IBSS* for a particular historical geographer, one different publication was found in *IBSS*. In general, *Scopus* provided the most comprehensive results and the results were more consistent compared to the other data sources considered. As combining different data sources would not have made a huge difference, in particular there would have been a high-degree of overlap with the *Scopus* results, therefore, *Scopus* was used as the sole data source for this bibliometric study.

It should be noted that for one particular human geographer it was not possible to retrieve the publication list from *Scopus*. Thus, the total number of geographers included in the bibliometric

study was 67, instead of 68. In a follow-up interview it was discovered that the geographer was research inactive during the date range selected for the study.

### 6.2.2 Data collection and sampling

Data were collected and analysed in April 2014. In the first part of the bibliometric study, bibliographical and citation information of all 67 researchers' publications (2008-2013) covered in *Scopus* were downloaded and imported into an Excel spreadsheet, information included:

- Authors
- Title
- Year
- Source title
- Volume Issue
- Cited
- Authors with affiliations

The total number of articles of each researcher in 2008-2013 (range: 2- 114), the number of co-authors in each article were counted (range 1- 157) and the average co-authors per publication were counted, calculated and recorded in an Excel spreadsheet.

In the second part of the study, the top 10 co-authors of the 67 senior geographers (i.e. the 10 co-authors with the greatest number of joint papers with each of the 67 senior geographers) during the period of 2008-2013 for the articles included were examined. Part two focused on the top 10 co-authors instead of all co-authors of the 67 researchers, as this sampling approach made it possible to understand the most typical types of collaboration taking place among the 67 senior geographers in the period.

Part 1	All publications of 67 senior geographers in 2008-2013
Part 2	Top 10 co-authors of 67 senior geographers in 2008-2013

Table 6.1 Sampling strategy for the bibliometric study

*Scopus* was a useful tool for the study, as *Scopus* not only provides information of authors' affiliations but also tools which can generate a list of co-authors and provide the number of joint papers with each author within a specific period. Lists of co- authors of the 67 researchers in the period 2008-2013 were generated using the tools and imported into an Excel spreadsheet. The total number of co-authors of the 67 geographers (2008-2013) was counted, and the affiliations of top 10

co-authors of each geographer were analysed. The affiliation details of co-authors were obtained from *Scopus* and the internet, using Google search engine.

In order to understand the characteristics of collaboration, a coding frame was used in part 2. Initially, the coding frame only included 7 categories (categories 1-7 in Table 6.2). However, after a pilot in which 10 co-authors were examined, the coding frame was expanded to 9 categories which included collaborations outside academia. Co-authors' affiliations were checked by searching for the institution name using (whatever browser you used) to identify whether this was in academia where it was not clear. The coding categories are shown in Table 6.2. Figure 6.1 shows a screenshot of the analysis on an excel spreadsheet. Names have been obscured to maintain participants' confidentiality.

### Categories

1	The same research group
2	The same department (but different research group)
3	The same university (but different department)
4	Other UK universities, geography department
5	Other UK universities, non-geography department
6	International, non UK universities, geography/ geography related department (e.g. geosciences)
7	International, non UK universities, non- geography department
8	Non-academia, UK
9	Non-academia, global

Table 6.2 Categories for co-authors

Researcher	University	Subfield	Research group	Document	Average n	No of coll	Co authors	Document	Group	Department	University	UK-Geog	UK-Non-Geog	International	International	outside academic	UK	
[Obscured]	[Obscured]	Cultural	Social and cultural	10	0.40	3	[Obscured]	2	0	0	0	0	1	0	0	0	0	University of Liverpool, Sc
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	0	0	0	1	0	0	0	0	0	University of Wales Aberys
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	0	0	0	1	0	0	0	0	0	University of Liverpool, De
Total								0	0	0	2	1	0	0	0	0	0	
[Obscured]	[Obscured]	Cultural, Historical	Knowledges, culture and space	0														
[Obscured]	[Obscured]	Development	Transformations: Economy, Society and	18	0.61	6	[Obscured]	3	0	0	0	0	1	0	0	0	0	University of Edinburgh, Ec
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	3	0	0	0	0	1	0	0	0	0	University of Edinburgh, Sc
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	2	0	0	0	0	0	1	0	0	0	University of Wisconsin, M
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	1	0	0	0	0	0	0	0	0	University of Oxford, Scho
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	0	0	0	0	0	0	1	0	0	Simon Fraser University, Bi
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	0	0	0	1	0	0	0	0	0	University of Durham, Dep
Total								1	0	0	1	2	1	1	0	0	0	
[Obscured]	[Obscured]	Development, Environment	Society and Environment	11	3.18	20	[Obscured]	5	0	0	0	1	0	0	0	0	0	University of Oxford, Scho
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	4	0	0	1	0	0	0	0	0	0	University of Leeds, Scho
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	3	0	0	1	0	0	0	0	0	0	University of Leeds, Sustai
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	2	0	0	0	0	1	0	0	0	0	University of Reading, Agri
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	2	0	0	1	0	0	0	0	0	0	University of Leeds, Scho
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	2	0	0	0	1	0	0	0	0	0	University of Oxford, Oxfo
[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	[Obscured]	1	0	0	0	0	1	0	0	0	0	University of Reading, Scho

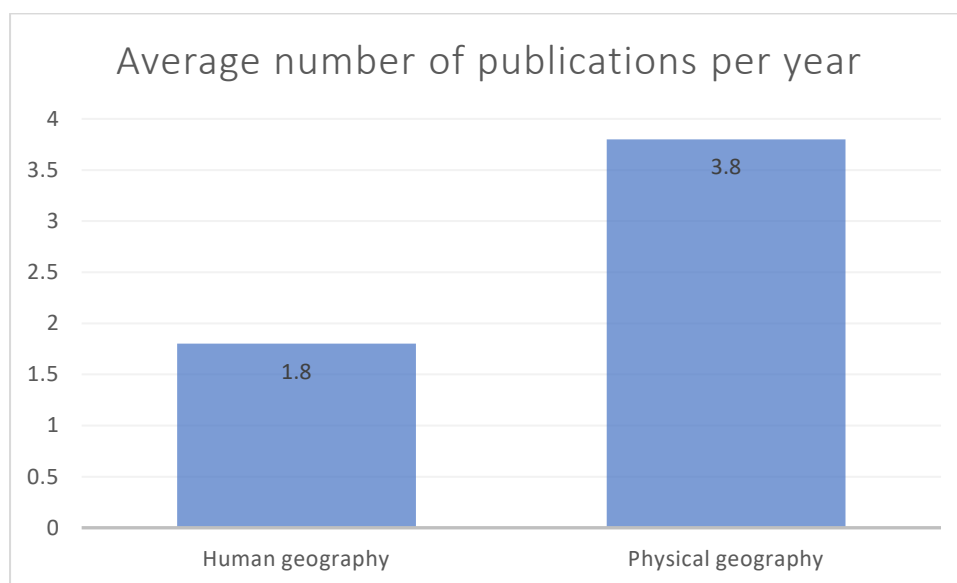
Figure 6.1 A screenshot of the analysis of co-authors

## 6.3 Results of Part 1

The first part of the bibliometric study examined the average number of publications per year, the number of co-authored publications and the average number of authors per publication in geography.

### 6.3.1 Number of publications per year

A total of 1492 publications were examined. 493 publications were from 34 human geographers and 999 publications were from 33 physical geographers. The average number of publications per human and physical geographer per year was 1.8 and 3.8 respectively (Figure 6.2). However, not every researcher published every year during the period 2008-2013.



*Figure 6.2 Average number of publications in human and physical geography (2008-2013)*

A T-test was conducted to show whether the mean of publications in human geography and physical geography are statistically significantly different. There are various types of T-test (Vaughan, 2001). A preliminary test for the equality of variances was used, i.e. F-test. The test was used to make sure the correct form of statistics is applied for the T test.

The results of the F-test indicate that the variances of the two groups were significantly different  $F=.210, p=.00001$  (Table 6.3). Therefore, a T-test was performed that does not assume equal variances (Table 6.4).

### F-Test Two-Sample for Variances

	<i>Variable</i>	
	<i>1</i>	<i>Variable 2</i>
Mean	1.816176	3.784091
Variance	1.593973	7.58718
Observations	34	33
df	33	32
F	0.210088	
P(F<=f) one-tail	1.15E-05	
F Critical one-tail	0.557591	

Table 6.3 F-test

### T-Test: Two-Sample Assuming Unequal Variances

	<i>Variable</i>	<i>Variable</i>
	<i>1</i>	<i>2</i>
Mean	1.816176	3.784091
Variance	1.593973	7.58718
Observations	34	33
Hypothesized Mean Difference	0	
df	45	
t Stat	-3.74047	
P(T<=t) one-tail	0.000259	
t Critical one-tail	1.679427	
P(T<=t) two-tail	0.000518	
t Critical two-tail	2.014103	

Table 6.4 T-test

The mean for the human geography group (M=1.82, N= 34) was significantly smaller than the mean for physical geography (M=3.78, N=33) using the two-sample t-test for unequal variances,  $t(45) = -3.74$ ,  $p < 0.001$ . This result (as shown in Table 6.4) shows that the average number of publications in human geography is significantly lower than in physical geography.

### 6.3.2 Number of co-authored publications

In human geography, 126 out of 493, 25.6%, of the publications were published by a single author, 74.4% of the publications were co-authored. In contrast, in physical geography, only 3.2% of publications, 32 out of 999, were published by a single author, 96.8% of the publications being co-authored (Figure 6.3).

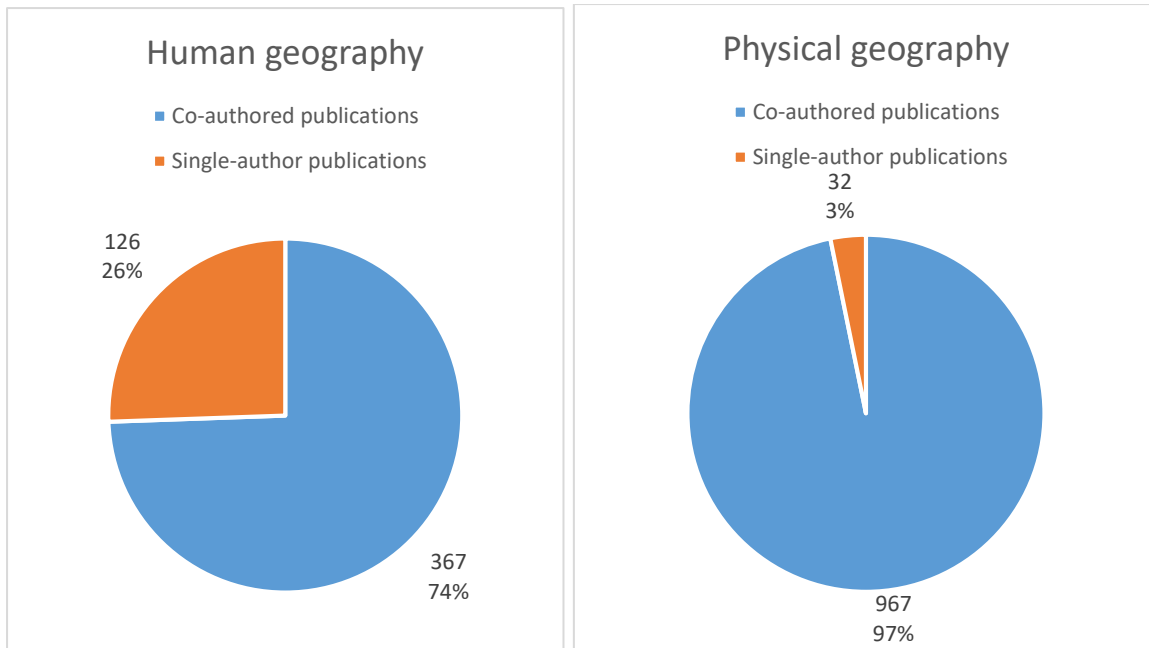


Figure 6.3 Number of co-authored publications

### 6.3.3 Number of authors per publications

The average number of authors per publication in human geography is 2.9 (mean), 2 (median), range (1-31) while the average number of authors per publication in physical geography is 7.8 (mean), 5 (median), range (1-134).

In order to show whether the average numbers of authors per publication in human geography and physical geography are statistically significantly different from each other, a T-test was carried out. A preliminary test for the equality of variances indicates that the variances of the two groups were significantly different  $F=.038$ ,  $p=0$  (Table 6.5). Therefore, a two-sample t-test was performed that does not assume equal variances (Table 6.6).

#### F-Test Two-Sample for Variances

	<i>Human geography</i>	<i>Physical geography</i>
Mean	2.87247	7.767768
Variance	5.498916	145.6614
Observations	494	999
df	493	998
F	0.037751	
P(F<=f) one-tail	0	
F Critical one-tail	0.8783	

Table 6.5 F-test

A T test was conducted:

T-Test: Two-Sample Assuming Unequal Variances

	<i>Human geography</i>	<i>Physical geography</i>
Mean	2.87247	7.767768
Variance	5.498916	145.6614
Observations	494	999
Hypothesized Mean Difference	0	
df	1143	
t Stat	-12.357	
P(T<=t) one-tail	2.58E-33	
t Critical one-tail	1.646188	
P(T<=t) two-tail	5.15E-33	
t Critical two-tail	1.962042	

Table 6.6 T-test

The mean for the human geography group (M=2.87, N= 494) was significantly smaller than the mean for physical geography (M=7.68, N=999) using the two-sample t-test for unequal variances,  $t(1143) = -2.36$ ,  $p < 0.001$ . The data in Table 6.6 shows that the average number of authors per publication in human geography is significantly lower than in physical geography.

#### 6.4 Discussion of Part 1 results

The results of the first part of the study showed that, on average, physical geographers had more publications, more co-authored publications and more authors per publication than human geographers.

	Human geography	Physical geography
No. of publications per year (per geographer)	1.8	3.8
No of co -authored publications (2003-2008)	367 out of 493 74%	967 out of 999 97%
No of single authored publications (2003-2008)	126 out of 493 26%	32 out of 999 3%
No of authors per publication	2.9 (mean) 2(median)	7.8 (mean) 5 (median)

Table 6.7 Results of part 1



These results suggest that physical geographers have a more collaborative research culture than human geographers. The collaborative nature of physical geography may be one of the factors resulting in a higher volume of research output. For example, researchers' workload can be more distributed and reduced when there are more collaborators in a research project therefore more publications can be produced. The number of authors per publication were higher in physical geography than human geography, i.e. 2.8 vs 7.8, indicating that research teams or groups in physical geography are larger than those in human geography. Although collaborative research in human geography is not uncommon, i.e. 74% were co-authored publications, the results show that single authored publications in human geography were significantly higher than in physical geography. When comparing the results with other disciplines (Figure 6.4), the percentage of publications written by one author in human geography (i.e. 26%) was close to the social sciences, such as economics and business, while physical geography (i.e. 3 %) was close to science subjects such as, neuroscience, chemistry and biology.

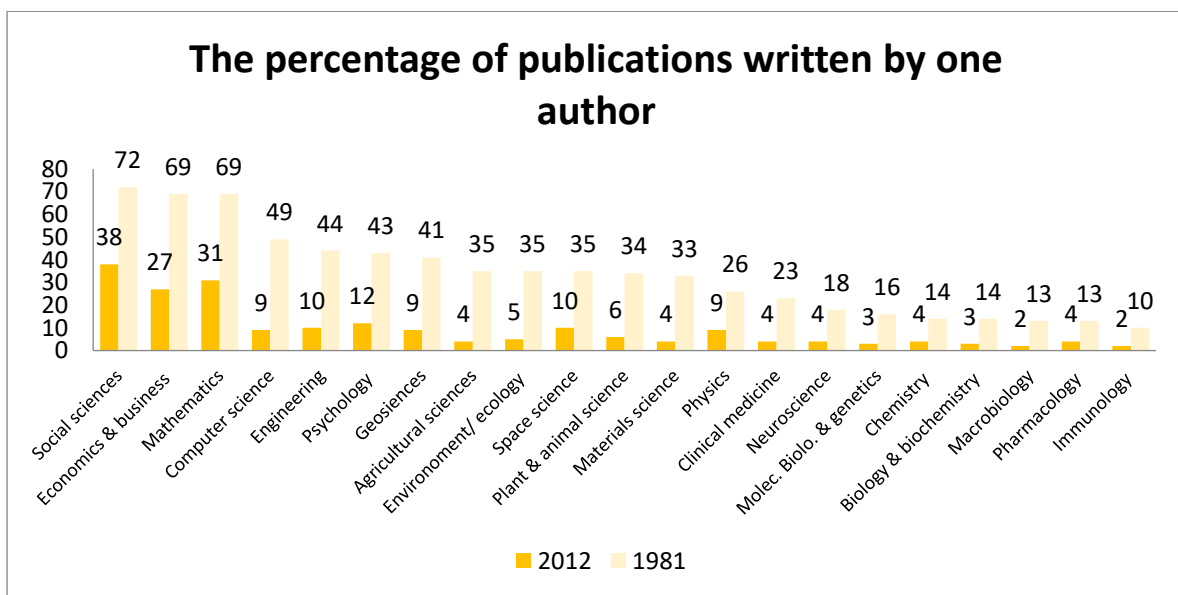


Figure 6.4 The percentage of publication written by single author across fields. Data from Thomson Reuters Web of Science. Source: adapted from Nagurney, 2013

It is interesting to note that the publication with the highest number of authors in human geography (2008-2013) has 31 authors and is related to the field of quantitative geography. In physical geography, the publication with the highest number of authors has 135 authors and is related to biogeography. Numbers of authors per publication in geography compared with other research fields, such as medicine, high energy physics and sociology, are shown in Table 6.8. However, the literature suggests that the number of authors per paper tends to be constantly increasing (Ioannidis, 2008; Ossenblok et al, 2014), thus it is important to bear in mind the figures for the average authors per paper from the older studies may under-represent the current position.

Subject	Source	Authors per paper
Medicine	Newman, 2001	3.75
Physics	Newman, 2001	2.53
High energy physics	Newman, 2001	8.96
Chemistry	Velden & Lagoze, 2009	3.3- 4.0
IT	Newman, 2001	2.22
Mathematics	Newman, 2004	1.45
Sociology	Moody, 2004	2.7
Management and organizational studies	Acedo et al, 2006	1.88
Physical geography	This study, 2014	7.8
Human geography	This study, 2014	2.9

Table 6.8 Number of authors per publication in different research fields

## 6.5 Results of Part 2

The second part of the bibliometric study examined in more detail the co-authors of the 67 senior geography researchers during the 2008-2013 REF period, including the frequency with which a geographer collaborates with other authors and where they are from.

### 6.5.1 Number of collaborators per geographer

The results show that the human geographers worked with a total of between 1 and 76 collaborators and the physical geographers worked with a total of 5 to more than 156 collaborators<sup>3</sup> over the 8-year period studied. All geographers have at least one co-authored paper. The average number of collaborators per geographer in the 2008-2013 REF period is 14.1 (mean), 9 (median) in human geography and 77(mean), 78 (median) in physical geography.

A preliminary test for the equality of variances indicates that the variances of the two groups were significantly different  $F=.10$ ,  $p=7.37E-10$  (Table 6.9). Therefore, a two-sample t-test was performed that does not assume equal variances (Table 6.10).

---

<sup>3</sup> The maximum number of collaborators which *Scopus* was able to display was 156.

### F-Test Two-Sample for Variances

	Variable 1	Variable 2
Mean	14.28571	76.63636
Variance	232.0336	2308.551
Observations	35	33
df	34	32
F	0.10051	
P(F<=f) one-tail	7.37E-10	
F Critical one-tail	0.560839	

Table 6.9 F-test

### T-Test: Two-Sample Assuming Unequal Variances

	Variable 1	Variable 2
Mean	14.28571	76.63636
Variance	232.0336	2308.551
Observations	35	33
Hypothesized Mean Difference	0	
df	38	
t Stat	-7.12471	
P(T<=t) one-tail	8.34E-09	
t Critical one-tail	1.685954	
P(T<=t) two-tail	1.67E-08	
t Critical two-tail	2.024394	

Table 6.10 T-test

A t-test was conducted. The mean for the human geography group (M=14.286, N= 39) was significantly smaller than the mean for physical geography (M=76.636, N=33) using the two-sample t-test for unequal variances,  $t(38) = -7.12$ ,  $p \leq 1.67E-08$ . The data in Table 6.10 shows, therefore, that the average number of collaborators per geographer in human geography is significantly less than in physical geography.

### 6.5.2 Collaborators of geographers

For the 67 senior geography researchers, there was a total of 565 top collaborators during the REF period; 240 were from 34 human geographers and 325 were from 33 physical geographers. It was the 565 top collaborators instead of 670, as some geographers had less than 10 collaborators over the period of 2008-2013.

The study shows that the 67 geographers published outputs with some of their collaborators more than once. On average, the human geographers collaborated with their top 10 co-authors twice, 2.42 (mean), 2 (median), whereas physical geographers collaborate with their top 10 co-authors around 5 times, 4.87 (mean), 3 (median) in 2008-2013. It is interesting to note that the highest frequency of a geographer to collaborate with a single author is 32 times out of 33 publications in human geography (political geography) and 32 out of 63 publications in physical geography (biogeography) during 2008-2013.

The bibliometric study examined the affiliations of 565 co-authors and categorised them into 9 categories. The results are presented in Table 6.11 and Figure 6.5.

	Categories	Human geography		Physical geography	
		No of collaborators	%	No of collaborators	%
1	The same research group	34	14%	38	12%
2	The same department (different research group)	12	5%	19	6%
3	The same university (non-geography department)	19	8%	7	2%
4	Other UK universities, geography department	63	26%	80	25%
5	Other UK universities, non-geography department	55	23%	43	13%
6	International, geography/ geography related department (e.g. geosciences)	28	12%	88	27%
7	International, non- geography department	17	7%	9	3%
8	Non-academia, UK	6	3%	16	5%
9	Non-academia, global	6	3%	25	8%

Table 6.11 Results of collaboration pattern in geography

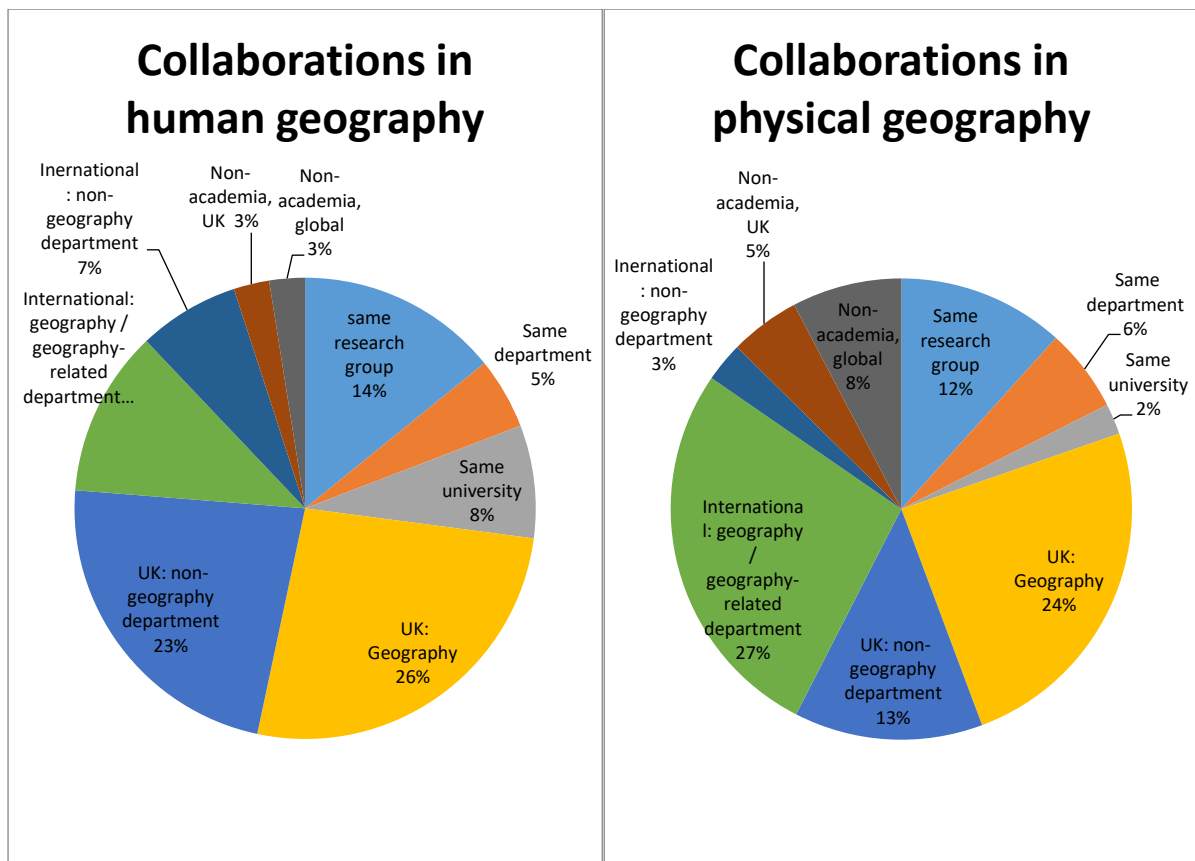


Figure 6.5 Collaboration patterns in human and physical geography

The results show that the most common type of collaboration in human geography is collaborating with researchers from other UK geography departments, 63 out of 239, 26%. Collaborating with researchers from other UK universities, non-geography departments, is the second most common, 55 out of 239, 23%. The least common types of collaboration are ‘non academia’ both UK and international, with only 6 out of 239, 3% each.

For physical geography, the results show that collaborating with authors in geography or geography related departments outside the UK is the most common. More than a quarter of their collaborators, 88 out of 325, 27% are from international geography/ geography related departments. The second most common type of collaboration in physical geographers is co-authoring with authors in other UK geography departments. 80 of the collaborators, 25%, are from other geography departments in the UK.

It shows that physical geographers collaborated with geography/ geography related departments the most. Nevertheless, the analysis found that physical geographers had some collaborations with non-geography departments, such as, Computer Science, Engineering and the Environment, Zoology, Agriculture, Botanist, Physical Sciences, Applied Mathematics, Civil/ Mechanical Engineering, Biology; Arctic Biology; Animal and plant sciences. Furthermore, the data shows that the least

common type of collaboration for physical geographers is collaboration with other department within the same university. Only 7 out of 325, 2 %, of the physical geographers' collaborators are from the same university, different departments.

A chi-squared test was carried out manually using Excel to investigate the differences in the distribution between human and physical geographers. The difference was found to be statistically significant ( $\chi^2= 48.84$ , 8 degrees of freedom,  $p<0.001$ ). Table 6.12 shows the categories which have relatively greater differences between human and physical geographers:

	Categories	Human geography	Physical geography
3	The same university (non-geography department)	8%	2%
5	Other UK universities, non-geography department	23%	13%
6	International, geography/ geography related department (e.g. geosciences)	12%	27%
7	International, non- geography department	7%	3%
9	Non-academia, global	3%	8%

Table 6.12 Categories which have significance differences between human and physical geographers

The results show a key difference between human and physical geography is human geographers had significantly more collaborations with non-geography departments anywhere (same university, other UK universities or international) than physical geographers. These non-geography collaborators of the human geographers were from a wide range of departments, such as, anthropology, sociology, social policy, political & international studies, business, psychology, criminology, food security, agriculture and development, medical, population health science, computing, built environment and town planning.

On the other hand, physical geographers had significantly more collaborations with international geography/geography related department or international non academics. The international collaborators of the physical geographers were from geography related departmental units, such as, Department of Geosciences, Department of Ecohydrology, Department of Cryolithology and Glaciology, Remote Sensing Division, Department of Stratigraphy, Paleontology and Marine Geosciences, School of Earth and Environmental Sciences, Earth System Science Interdisciplinary Center and School of Environmental and Life Sciences.

## 6.6 Discussion of Part 2

The results of the second part of the study shows that physical geographers have a significantly higher number of collaborators (Table 6.13).

	Human geography	Physical geography
No. of collaborators per geographer in 2008-2013	14.1 (mean), 9 (median)	77 (mean), 78 (median)
The most common type of collaboration	1. UK, geography 2. UK, non-geography	1. International, geography related 2. UK, geography
The least common type of collaboration	1. Outside academia, UK 2. Outside academia, international	1. Same university, non-geography 2. International, non-geography
The major difference between human and physical	likely to collaborate with non-geography departments anywhere	more likely to collaborate with international geography department or non-academic

Table 6.13 Results of part 2

As suggested earlier, more co-authors may reduce workload and produce more publications. This study supports this idea. The number of collaborators and the number of publications are related (Fig 6.6).

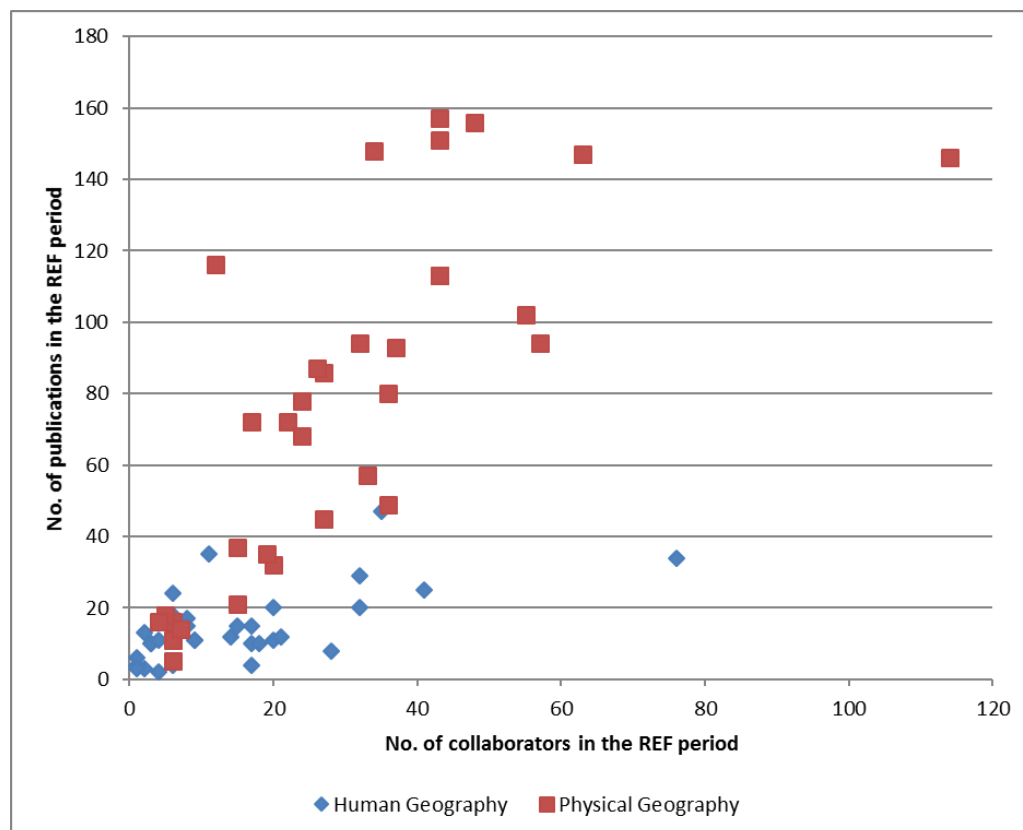


Figure 6.6 Number of collaborators and publications in geography

Pearson correlation tests were conducted. The data concludes that the number of collaborators and publications are positively correlated in both human geography ( $r^2=0.62$ ) and physical geography ( $r^2=0.73$ ). It shows that the greater the number of co- authors, the more number of publications are produced.

The second part of the results further revealed that there are differences in collaboration patterns between human geography and physical geography. It was found that physical geographers are more likely to collaborate with international geography departments or non-academics; human geographers are more likely to collaborate with non-geography departments anywhere. The results could indicate that physical geography is urban in nature, while human geography is rural in nature.

Looking more deeply into the 'international collaboration with geography and geography related' department in physical geography, many of the international collaborations are not from traditional geography departments. Out of 87 co-authors from the "international geography/ geography related" category, only 18 co-authors are from a department labelled as 'Department of Geography'. Many of these academic departments/ centres/ schools are not traditional disciplinary departments. They are relatively hybrid, integrated. For example, as mentioned earlier, Department of Geosciences, Earth System Science Interdisciplinary Center. It was not surprising, as the findings of the web study of UK geography departments showed (Chapter 4), not all UK geography academic units are named geography, e.g. School of Geosciences, School of Geographical and Earth Sciences, Department of Geography, Geology and Environmental Science. The analysis of this chapter showing that physical geographers tend to collaborate with integrated geography departments further suggests that the nature of physical geography is urban. As Becher and Trowler (2001) pointed out, the characteristics of an urban discipline have a high-people-to problem ratio, are competitive, with intense interaction and collaboration. Because of the urban nature of physical geography, it is possible that data in physical geography can be more easily shared within an urban environment, e.g. shared data can be useful not only for physical geographers but also for earth and environmental scientists.

In contrast, the results showed human geographers tend to collaborate with non-geography departments, such as, anthropology, sociology, political & international studies, business, health science, and town planning. This can indicate that the nature of human geography is rural. As Becher & Trowler (2001) stated, research problems of rural disciplines are 'thinly scattered' across a broad area (Becher & Trowler, 2001:185). Rural disciplines have a lower people-to-problem ratio and less interaction. Therefore, it is possible that data in rural human geography can be less easy to share.



In general, the results showed different collaboration patterns between physical and human geography (Figure 6.7). Nevertheless, there are some similarities on a broader level. For example, as shown in Figure 6.7 the majority of geographical research collaboration, both physical and human, is within the UK. Relatively few co-authors are outside academic institutions.

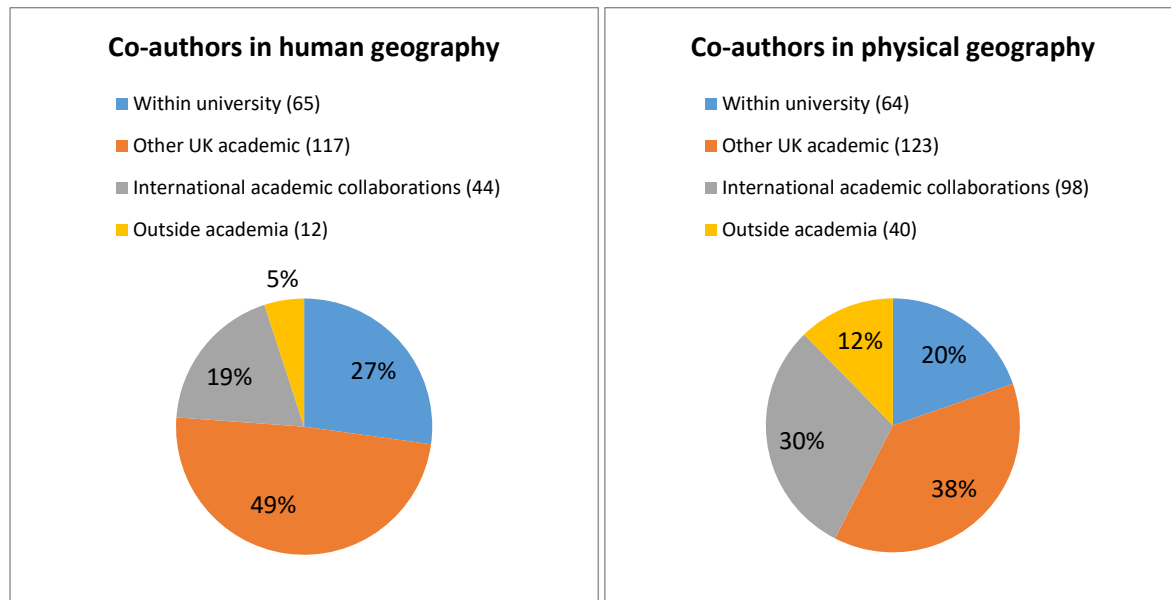


Figure 6.7 Results comparison on a broader level

## 6.7 Summary and conclusion

This chapter investigated research collaboration in geography through co-authorship. The small scale bibliometric study examined publications and co-authors of 67 senior geography researchers in the REF period, 2008-2013. The study showed that the average number of publications per physical geographer is significantly higher than that of human geography per year (1.8 vs 3.8). The number of co-authors per publication in physical geography (6.8) is also higher than human geography (1.9). The correlation analysis showed that the greater the number of co-authors the more publications are produced. Researchers in physical geography not only have relatively larger research teams or groups, but also have more collaborative research than human geography. This result could suggest physical geography has a higher people to problem ratio and intense interaction, while human geography has a lower people to problem ratio and less interaction. Using Becher's (2001) taxonomy physical geography could be considered as urban in nature and human geography can be considered as rural in nature. This rural / urban nature of research organization could have major implications for researchers' data practices, especially data sharing and data re-use. For instance, urban organisations feature a high people to problem ratio field, i.e. narrower field of study in urban organisation may imply data are more re-usable, yet the intensive competitive culture of urban

organisations can be a barrier for researchers to share data beyond a research team. On the other hand, rural research organisations are more isolated and experience more difficulties of finding audience for their data. This may imply data could be less likely to be re-used.

This study further examined co-authorship in geography in more detail, particularly the affiliations of the 564 top co-authors of 67 senior geography researchers. The findings showed that there are differences between the collaboration patterns among human geographers and physical geographers. One of the major differences is physical geographers more likely to collaborate with international geography department or non-academic colleagues; human geographers are more likely to collaborate with non-geography departments anywhere. This again suggested physical geography is urban in nature whereas human geography is rural in nature.

Collectively these cultural differences could indicate that there will be differences between human and physical geography in storing, sharing and re-using research data. For instance, sharing and reusing research data can be easier in an urban disciplinary environment than a rural environment. Furthermore, human geographers may tend to store and share data within the UK, while physical geographers are more likely to share data internationally. There may also be a greater need to promote international data sharing in physical geography.

In conclusion, the bibliometric study provides an understanding of geography as a discipline from a social perspective, especially collaboration patterns within human and physical geography. The study identified differences between human and physical geography collaborations, which could have important implications on how researchers store, share and re-use data.

# Chapter 7

## Intellectual Characteristics in Geography

---

### 7.1 Introduction

This chapter and next three chapters present findings from the Phase 2 interviews. The Phase 2 of this study investigated research practices and attitudes to data sharing of researchers, aiming to answer the key research question, i.e. in what ways are the data practices of researchers in geography shaped by the nature of the discipline? This chapter focuses on the intellectual characteristics of geography by looking at research being conducted by interview participants. Intellectual characteristics of disciplines are associated with intellectual ideas, cognitive norms and procedures in an academic field, for example, research methods and topics (Whitley, 1974; Holland, 2014). They are often considered as a key foundation for exploring the nature of academic fields, for example, according to Whitley (1974), Ziman (1984), Becher (1989), Fry (2007), Repko (2014), Holland (2014). Intellectual characteristics of disciplines and research data are closely related. This chapter provides a deeper insight not only into research practices but also research data in geography.

### 7.2 Selecting interview participants

A total of 23 semi-structured interviews with researchers from 11 institutions were carried out between September 2014 and November 2014. The following described how suitable interviewees were selected.

#### *Sampling of context*

As the findings of Phase 1 showed that there are clear differences between human and physical geography, interviewees in phase 2 were drawn from subfields of both human geography and physical geography. Table 7.1 shows suitable geography subfields for interviews that were identified. Most of the subfields were chosen on the basis of being common subfields in the UK. As identified in Phase 1, these common subfields included political geography, social and cultural geography, geomorphology, and hydrology. Some less common sub-fields were also included, such as, quantitative geography and environmental geography in human geography, as they could be highly relevant to the latest research data practices, e.g. involving big datasets.

Human geography	Social & Cultural geography	Political geography	Urban & development geography	Quantitative geography	Other subfields (e.g. health / environmental geography)
Physical geography	Biogeography & ecological processes	Geomorphology & hydrology	Past environmental change ( e.g. glaciology )	Climatology	

Table 7. 1 Sample subfields

### Sampling of participants

The initial sample of phase 2 interviews was chosen from the 68 senior researchers of the 15 institutions from Phase 1. However, there were not enough responses from this group, thus the sample was expanded to researchers from the 15 institutions. 72 email invitations (Appendix 8) were sent. The main selection criteria of participants were: 1) the willingness to participate, 2) academic staff from suitable geography subfields, as stated above, and 3) being members of the 15 chosen institutions. As the phase 1 study found that there are institutional differences between top, middle and low ranking geography departments, in phase 2 interviews, it was attempted to recruit a balanced number of participants from top, middle and low ranking institutions in both human and physical geography. A profile of interviewees is contained in Appendix 9.

Top	Middle	Low
2. University of Oxford	22. University of Birmingham	37. Kingston University
3. University of Durham	23. University of Hull	39. Brunel University
6. University of Leeds	28. University of Liverpool	45. University of Chester
7. University of Sheffield	29. University of Newcastle upon Tyne	46. University of Worcester
9. University College London	32. Loughborough University	48. Edge Hill University

Table 7.2 Sample Institutions

### 7.3 Research topics and subject in geography

Looking at interview participants' research topics and subjects offers fundamental understandings of intellectual characteristic of a discipline. The interview analysis shows that research topics of interview participants varied widely. Interview participants were researching a range of different topics such as, glaciology, limnology, GIS, aquatic vegetation, voting system, young people and gender. Table 7.3 shows the exact keywords which interviewees used for describing their own research interests. As can see seen from Table 7.3, a division between physical geography and human geography is noticeable from their topics of research. The subject matter of physical geography interviewees' research is related to the physical and natural environment, e.g. lakes, glacier, climate, whereas for human geography interviewees' research is related to the human

activities, e.g. economics and education. This division of subject matter is one of the most noticeable and important intellectual character of geography as a discipline.

Physical			Human		
Interviewee	Topics	Methods	Interviewee	Topics	Methods
Matthew	"Climate change; Geo-chemistry"	Quantitative -field measurements	Brian	"Urban modelling"	Quantitative -modelling, simulation
Patrick	"Biogeochemistry; Paleo-ecology"	Quantitative -field measurements	Paul	"Town planning; Community based learning"	Qualitative -interviews, Focus group ethnographic kind
Ryan	"Climate change; Hydrochemistry; Quaternary sciences;	Quantitative - field measurements	Samuel	"Political geography"	Quantitative - quantitative analysis model
Christopher	"Ocean modelling; Climate change"	Quantitative -modelling	Anthony	"Transport; climate change"	Quantitative &Qualitative (quantitative data +interviews)
William	"Hydro-climatology"	Quantitative -modelling	Sarah	"Development, social , urban geography; Global south"	Qualitative (questionnaire survey +interviews, focus group)
Kevin	"Water dynamics"	Quantitative -modelling	*Jessica	"Sustainability; Climate change"	Qualitative -interviews
Michael	"Glaciers; Remote sensing"	Quantitative -remote sensing	Joshua	"Big data"	Quantitative - spatial, network analysis, statistical modelling
Kyle	"Geo-visualisation"	Quantitative -GIS	Adam	"Development geography"	Qualitative - secondary sources -interviews - workshops
Rachel	"Remote sensing; Vegetation"	Quantitative -remote sensing	Eric	"Development and political geography"	Qualitative - secondary sources - interviews
*James	'Geographical Information system"	Quantitative -GIS	Kimberly	"Gender; Environment"	Qualitative Secondary – interviews-ethnographies
*Mark	"Community energy"	Quantitative - remote sensing, field measurements  Qualitative Interviews + Questionnaire	Amanda	"Children geographies; Social differences; Gender"	Qualitative - interviews, ethnographic observation
*	Interviewees who research subjects span across physical and human geography		Thomas	"Historical and cultural geography"	Qualitative - Archival work + interviews

Table 7.3 Research topics and methods of interviewees

### *Changes in research interests*

Research ideas or topics of researchers are not static. Human geographer Adam (development) found it difficult to describe his research area, as he said *“That’s quite tricky in the sense of summarising a kind of twenty years’ career”*. The interview analysis reveals that participants’ research career can be quite complex which leads to changes in their research topics and subject matter. Physical geographer Patrick (quaternary, biogeography) explained the change of one’s research topics can link to following funding sources, new research ideas and opportunities. As he said,

*“Many years ago, I was given the opportunity to work in Greenland. So I started to develop an interest in lakes and climates directly. And then I changed jobs; I came back to the UK. [...] I mean one’s research career evolves. You start asking different questions. So part of it is you go after funding sources. [...] partly it’s because you generate new research questions. And in my sort of example, [...] my job situation changed in the sense that I was allowed to work in Greenland. So I went from doing lowland lakes to arctic lakes. So that’s serendipity. [...] Everybody has a large element of serendipity in their research career.”*

The analysis found that some interviewees’ research topics were built from their PhD research, whereas some interviewees’ research interests have shifted or expanded to different areas. For example, Rachel (remote sensing) was a geomorphologist who moved to the area of remote sensing, as she said:

*“I (was) more a geomorphologist originally. I did a PhD looking at soil erosion [...] I don’t think I’d call myself geomorphologist anymore. I might now call myself a remote sensing scientist.”*

Physical geographer Ryan (quaternary) also shifted his research from geology to environmental science and geography. As he said,

*“My research is really around the boundaries of physical geography, environmental science and earth science. So my original training was in geology. And then I moved to, I started looking more at the environment [...] and then after a few years I switched university [...] So I changed from geology to geography.”*

Similarly, human geographer Eric (political and development) expanded his research from African studies to development geography. As he said,

*“My PhD was in African studies. And I was supervised by a professor in Education and a professor in Anthropology. So my training as a postgraduate was interdisciplinary. And [...] I try and hope to bring those ideas into what I am doing in development geography.”*

### *Crossing the physical and human divide*

Researchers’ research topics evolve over time, yet, as the above examples show, the shift of topics usually are within the boundary of physical geography or human geography. Nevertheless, interestingly, three interview participants shifted their research topics further than others. They crossed the division between physical and human. All three participants, Jessica (Climate change), James (GIS) and Mark (sustainability), started as physical scientists, and they gradually moved their research topics towards the social, human side of the discipline. Interestingly, none of the interview participants shifted from human to the physical side. Jessica (climate change) did a PhD in climate and moved into the area of social sciences, such as public interests in climate change, the discourse of climate change. As she described it:

*“I started as a physical scientist. [...] I have gradually moved more into social sciences [...] I became more engaged myself in social sciences. So very late in [...] my research life, I did then start to do my own project which is very very different from anything to do with climate change which is my real background. That was very different and a very small project where I actually went out to interview people and set up myself.”*

James (GIS) and Mark (sustainability), also started as physical geographers, but described their current research as situated somewhere between physical and human geography. As they said:

*“My background is in physical geography. I still like to put on my Wellington boots and get involved but mainly most of my work is more GIS focused. I wouldn’t say it’s physical geography; I wouldn’t like to say it’s human geography either [...] It is somewhere in the middle.” (James, GIS)*

*“I certainly started as a physical geographer. I still have physical geography interests. But I am a geographer, so my interests span physical and human geography, so I have got into the dark side of human geography in recent years.” (Mark, sustainability)*

Mark’s (sustainability) comment shows that there can be an identity which transcends the human/physical divide in geography. In the interview, James (GIS) explained why his research was between human and physical geography, saying that it is because his research topic is more of a

notion of human geography, nonetheless, the method he used is from physical geography. As he commented,

*“I think when you are dealing with environmental issues, let’s say wilderness quality issues- wilderness is a human construct, a human idea we apply to the natural world. [...] (It) is very much of a human social geography question. But in terms of the kind of work I’m doing modelling landscape, land cover, natural processes, that’s very much physical geography bringing in and out the element of ecology and biology possibly zoology all of those sciences which are physical sciences as opposed to the human sciences but is very much of a human question. And I think that’s applied to a lot of environmental topics, for example, sustainability.”*

The above comment also suggested the connection between human and physical geography is more common in environmental topics. In fact, Mark’s research is in the field of environment and sustainability. His comment below provides some hints why such a connection is typical to environmental topics. As he said:

*“[...] because there is no natural environment in this world. Humans interfere everywhere. So there is no such thing as a natural environment and therefore if you are going to study the environment and landscapes, you have to understand the human impact and input on those landscapes. And equally the (natural) environment does set a limit to what humans can do and therefore you need to understand the (natural) environment.”*

Mark (sustainability) believes it is essential to consider the elements of both human and physical geography when studying environment topics. In his research in energy and sustainability, he combined technical and scientific knowledge, such as thermal imaging, insulation, with methods such as interviews and questionnaires which are associated with human geographers. As he said,

*“I think it is important to have a well-rounded view and be open to using a range of different methods to suit different situations. [...] for instance, community carbon reduction, community energy behaviour clearly the methods one would use and study those would be typical methods used by human geographers- qualitative methods, but also quite a bit of quantitative work. [...] But the need to understand the technology that underlies what people are doing [...]and how they can generate other energy is equally important. So I think you need to be a scientist and a social scientist to do that job properly.”*

Although research between human and physical geography evidently exist, the division between human and physical geography is still deep in the UK. As human geographer Joshua (economic)



described, “Yes there is a big division at least in the UK”. Jessica (climate change) who did try to do research to bridge the divide, also commented that research between human and physical geography is unusual in her geography department. She said,

*“It is quite rare. [...] Part of the reason I started was because when I came into the department many many years ago, I found it very very strange how little linkage there was between the human and physical geographers. [...] I thought that was a shame really. [...] I suppose it was my attempt that [...] 20 years to try to get that going. To say it failed is perhaps a bit of an understatement but I think it was viewed with bitterness. It was viewed with scepticism. [...] I can remember the professor of physical geography at the time said ‘This is never going to be a main part of geography and so you know this is always going to be a side dish.’”*

In general, examining research topics and subjects of interview participants provided insights into disciplinary characteristics of geography. The fundamental differences between physical geography and human geography can evidently be found through intellectual topics and subjects with which researchers were interested. The analysis showed that changes in research interests usually are within the boundary of physical and human geography. Yet, the analysis revealed some interviewees’ research has crossed the human / physical divide in geography. Research situated between human and physical geography exists.

#### 7.4 Research methods in geography

Topics, subjects and methods of research are key intellectual elements of a discipline (Whitley, 1974; Becher & Trowler, 2001). Looking at research methods used by interview participants offers important understandings to intellectual norms of geography and the nature of the data of the discipline. As Holland (2014:213) stated, research method principally is “the construction and interpretation of data”. The interview analysis shows that certain research methods are associated with particular research fields. For instance, qualitative methods are commonly associated with human geography, whereas quantitative methods are associated with physical geography. The interview analysis found that this perception is true. As table 7.3 shows physical geographers only use quantitative methods, in contrast the majority of human geographers use qualitative, only 4 out of 12 used quantitative methods in their research. It is clear that physical geographers and human geographers use different research methods.

## 7.5 Research methods in physical geography

The interview analysis found that field measurements, numerical modelling and remote sensing were the major research techniques used by physical geography interview participants. These findings are in line with previous literature, such as, Clifford (2010), Gomez and Jones (2010). As physical geographer Michael (glaciology), who was talking about the consensus of research techniques in physical geography, said:

*“In terms of the techniques [...] you know the three main approaches I would say are numerical modelling, where you are essentially writing a series of code and equations to explain the process that you are trying to understand. And then the field work, people actually go and visit places, collect data in the field and probably remote sensing where people can look at satellite imagery. So they are the three main techniques.”*

### 7.5.1 Field measurements

Field measurement is one of the key research techniques in physical geography. Noticeably, interview participants Ryan (quaternary), Matthew (quaternary) and Patrick (quaternary), who used field measurements as their primary research method were in the field of geomorphology and quaternary sciences. In general, field measurements consist of collecting and analysing samples from the field. Field measurements can be illustrated by the following examples from the research of Ryan (geomorphology, quaternary), Patrick (quaternary) and Matthew (quaternary).

Ryan (geomorphology, quaternary) is a Geomorphologist. He was involved in a variety of research, yet all of his research involved taking field measurements. For example, in his geological research, he collected samples and made measurements of rock layers, as he said:

*“we have to look at the different kinds of research. So in the more geological research, it involves fieldwork and collecting samples and making measurements of successions of rock layers- rock strata. And then followed by chemical analysis and study by microscope- so microscopy.*

For his cave research, he took stalagmite samples and measured different parameters from the cave environment:

*“(he was showing some photos taken from the inside of a cave) The cave work [...] involves field work and monitoring the cave environment. So that includes measuring physical and chemical parameters of the environment. So it may mean measuring the carbon dioxide content of the air, the air movement, the humidity of the air, the temperature of the*

*air, the amount of dripping water, the chemistry of the water. And so this measurement can be very quantitative and they may be continuous log measurements. And then working on the stalagmite samples themselves (he was opening a box and took out a stalagmite) is more like geological research that involves chemical analysis, and involves microscopy. When we look at the water, we do chemical analysis on the water as well.”*

For his glacial research, he collected and analysed water samples from glacial melt streams:

*“And when I work in glacial environments looking at glacial melt streams, for example, is a similar thing which involves collection of water, measurement of the discharges of the stream, the amount of water and measuring the chemical composition of the water.”*

The interview shows that field measurements are not only used in geomorphology but also in quaternary and biogeography research. In a similar way, Patrick whose research on quaternary, biogeography, took field measurements but in a different environment. He described methods he used in his research:

*“I work on lakes. [...]if we go limnology, the contemporary...we take or measure water samples and we analyse the water samples. That’s for a variety of different things, whether its chemical, the neutron concentration, the PH, the alkalinity. [...] We also take sediment cores. We go to the mud that collect in the bottoms of the lake and you take the sediment core, mud and chop it up and analyse it and then you have to date it. So you know how old it is. And then you analyse the mud for different components. One of which is diatoms, carbon content.”*

As the above examples show, his research methods involve taking measurements from, air, water, rock layers, sediment cores etc. and supplemented by chemical or biological analysis.

### *Data*

Quite similar to Patrick’s research, Matthew’s (quaternary) research involved analysing sediment cores from lakes. What is interesting is that, when Matthew was asked how he processes his data, his answer was about how he process sediment cores. As he said:

*“You start off with, basically some mud. (he was looking for a lump of dry mud and a metal sieve) And to get these things out (he was pointing to a microfossil slide from a small wooden cabinet), you wash the mud dirt in sieve. And that’s all the fine stuff comes out and you end up with basically some dry mud. You should than look at the microscope, pick up the micro fossils with paint brush. You identify them what kind of species they are and then if you*

*are looking at the geo-chemistry of them. It would be chemical composition. You then... you have various processes you put them through and you analyse using, depending on what you actually measure. [...] these shells are made of calcium carbonate, and I am interested in trace metal in there. So there are different metals that do the calcium carbonate which can tell us things about the lake, what the bottom of the lake was like when these things were living in there."*

Matthews seemed to see the sediment cores he collected as his data. Yet, Matthew (quaternary) mentioned later in the interview, the data could also be "*numbers in the spreadsheet*". This shows that the concept of research data can have different, multiple meanings even in one person's research. It can refer to physical objects collected from fieldwork, such as mud samples from a sediment core, to recorded measurements in numerical form.

In the interview, Matthew further mentioned there are two types of research data he usually produced, i.e. paleoecological data and geo-chemistry data, as he described:

*"One is what we call paleoecological data, which is information on microfossils in lake sediments. So basically, [...] if I pick these things out of sediments. These are calcium carbonate shells. These are the pictures under microscope. [...]The different species can tell us lots of things about the characteristics of the lake which in turn tell us information about climate. So that's one sort of data. And the other thing I do is look at the chemical composition of these shells, using a range of techniques that generates geochemical data. So a big spreadsheet full of numbers really."*

This suggested that the notion of data can be understood from two aspects. The first aspect is the form, for instance, numerical form, digital form, physical form, the format of the data, such as "*spreadsheets full of numbers*". And the other aspect is the content, such as, "*paleoecological data*", "*geochemical data*", where the content of data is about paleoecology, geochemistry. The content and the form of data can be closely associated. For instance, as the above comment showed, Matthew (quaternary) associated geochemical data closely with '*big spreadsheets*'.

### *Characteristics of field measured data*

Field measurements can be seen as a description of the natural landscape and environment. The interview analysis found that there are three main characteristics of this field measured data. One of the key characteristics of this data is, "*full of numbers*" (Matthew, quaternary), "*very quantitative*" (Ryan, quaternary). The analysis showed that these numerical descriptions of the field are commonly managed and processed through spreadsheets. Patrick (quaternary, biogeography),

Matthew (quaternary) and Ryan (geomorphology, quaternary) all used spreadsheets for processing their data. For example,

*“I use a spreadsheet to process the data. There are some specialised software, but I don’t use that too much, so mostly spreadsheet.”* - Patrick (quaternary, biogeography)

*“We tend to manage the data through spreadsheets not through databases. [...] So most of it is stored on a spreadsheet and coded by site and date.”* - Ryan (geomorphology, quaternary)

Interestingly, Patrick (quaternary, biogeography) used the term ‘process’, this suggests he sees the data is processed, actively analysed, rather than simply recorded.

Besides having a quantitative nature, the other characteristic of field measured data is that it is continuous. For example, Ryan (geomorphology, quaternary) used *“continuous measurements”* to describe his data. This is because physical geographers are concerned with physical environmental changes and processes, thus in order to examine how things change, data is commonly measured on a continuous time scale (Clifford, 2010). It can be hourly, daily, seasonal or even by millennia.

Matthew’s (quaternary) research looks at seasonal time series data, as he said:

*“We also do seasonality, so we go back to the lakes, look at how things change over the summer.”*

The following comment further reveals how this continuous data is presented and analysed in standardised ways, e.g. using time series and other statistical approaches, as Matthew (quaternary) said:

*“We can present the data as time series. [...] You could also look at the data spatially. Because I work on multiple lakes, so you can look at gradients, how it changes along the gradients, the mud. That’s normally presented against lake age, how old the sample is. Then again that’s time series, just a longer time series you know, cover thousands of years. And then on top of that, that’s data presentation but then there can be whole range of statistical tools that we use to analyse the data. Some of them are very simple things, just simple regressions but other more complex multivariate statistical approaches. So there’s a full range of different approaches.”* – Patrick (quaternary, biogeography)

The third characteristic of field measured data is that those datasets can have a large number of different variables which increase the complexity of the data. For instance, as Patrick (quaternary, biogeography) said,

*“Actually one of the problems we have is because we deal with large datasets with a large number of different variables. So we have quite complex data and it can be quite... that’s why we use multivariate statistical packages.”*

The three data characteristics of the research technique, quantitative, continuous and large number of variables, can be illustrated by following example, from Matthew (quaternary)’s data. In the interview, quaternary scientist Matthew showed some of his own research data. Most of his data is in digital format yet some of his data was in printed format. He presented one of his thick files on his shelf, which each have hundreds sheets of paper in it. Figure 7.1 provides an impression of what the content of these files looks like:

3	34	0.45	44.18	89.77	56.86	274.57	34.65	55.12	0.63	26.9	5.5	186	0.45	36.56	9.99
5	36	0.45	41.63	78.56	50.93	212.12	55.15	23.41	2.36	12.3	10.5	271	0.48	41.47	11.36
9	46	0.38	41.89	55.60	38.89	110.17	27.91	27.69	1.01	19.9	12.5	307	0.57	35.19	11.00
17	73	0.39	30.48	31.03	23.99	46.97	13.10	17.94	0.73	24.5	15.5	363	0.56	30.86	9.23
25	108	0.37	33.94	25.93	21.33	34.54	13.12	12.81	1.02	19.7	20.5	460	0.53	37.94	10.06
29	134	0.36	22.43	14.88	12.81	19.40	5.89	8.99	0.66	25.8	25.5	562	0.39	39.22	7.30
40	195	0.39	24.84	14.76	13.05	17.45	6.69	8.08	0.83	22.3	30.5	670	0.50	41.63	9.35
50	268	0.40	31.24	16.06	14.62	18.18	6.58	9.47	0.69	25.0	32.5	714	0.49	38.32	8.41
60	352	0.34	27.61	10.67	9.93	11.71	4.56	6.12	0.75	24.0	35.5	783	0.49	36.45	7.70
70	448	0.31	27.40	8.63	8.17	9.26	3.52	5.11	0.69	25.3	37.5	830	0.43	39.60	7.31
80	554	0.33	31.87	9.53	9.15	10.07	4.07	5.46	0.75	24.3	40.5	901	0.49	45.48	9.23
90	670	0.37	15.54	4.82	4.68	5.04	1.78	3.04	0.58	27.2	45.5	1024	0.43	36.00	6.11
100	796	0.40	19.75	6.15	6.00	6.37	2.21	3.94	0.56	27.9	47.5	1075	0.41	38.22	6.17
108	897	0.39	31.94	9.18	8.96	9.43	5.31	3.87	1.37	16.6	50.5	1152	0.53	44.81	9.12
109	914	0.39	27.69	7.89	7.72	8.12	3.76	4.13	0.91	21.1	55.5	1284	0.42	33.82	5.31
111	932	0.39	37.57	10.62	10.35	10.86	5.44	5.18	1.05	19.4	57.5	1339	0.39	33.57	4.81
112	949	0.30	21.65	4.59	4.48	4.70	1.95	2.64	0.74	23.9	60.5	1421	0.39	45.02	6.24
113	967	0.30	27.33	5.75	5.62	5.89	2.80	2.94	0.95	20.6	65.5	1563	0.27	36.56	3.49
114	985	0.30	20.05	4.18	4.10	4.30	1.40	2.78	0.50	29.2	67.5	1620	0.30	33.49	3.51
116	1003	0.30	18.73	3.88	3.78	3.96	1.32	2.56	0.52	28.9	70.5	1708	0.29	48.25	4.78
122	1091	0.40	32.73	8.85	8.66	9.05	3.91	4.94	0.79	23.0	75.5	1857	0.29	47.69	4.64
130	1221	0.34	24.25	5.31	5.21	5.44	2.09	3.22	0.65	25.4	80.5	2011	0.36	36.10	4.24
139	1357	0.37	21.07	4.86	4.75	4.98	2.05	2.81	0.73	23.9	82.5	2073	0.32	33.64	3.43
147	1500	0.39	28.29	6.55	6.42	6.71	2.44	4.10	0.59	26.9	85.5	2168	0.32	43.97	4.41
156	1647	0.36	32.76	6.67	6.52	6.83	3.37	3.29	1.02	19.7	90.5	2328	0.33	42.51	4.30

Figure 7.1 Field measured data

The papers in these files were full of numbers and with different columns. Matthew (quaternary) explained the first column usually presents the depth of core sediments, across the rows are different variables, such as various chemical contents in shells. This clearly confirms the three characteristics of research data that we have identified, i.e. quantitative, continuous and with different variables.

In the following comment, Matthew (quaternary) explained his printed data in the interview. What is surprising is that he pointed out that not all raw data but only part of the data is usually put into archive, as he said,

*“So just as an example, I’ve got a lot of this is kept digital... This is all the geochemical data from the analysis. You could see this is the depth in the sediment core (the first column) and they’ve got lots of values across here but ultimately the only value would go into the archive would be these two over here (the last two columns). But I know if I want to go back to do things with the data, it is really important to have everything come in here. Even these spreadsheets [...] they have other values come up the machine. I would call the*

*values that come out from the machine the raw data. There is some level of processing done before that data becomes useful and also before they could be put into any kind of archive."*

He gave an example of why his data need to be semi-processed before putting it into an archive:

*"Usually what the archive was interested in having is values and their related age. So if again you thinking about the case of the sediment core, the primary, the raw data would be depth in the core. But usually in the archive what you would put is, you might put the depth but more interesting is the age. Because 10 cm down the sediments could be 5 years ago in one sediment core, it could be a thousand years ago in another core, so it depends on circumstances. The depth itself isn't, although is a primary measure, it is not very very useful to have in an archive. So sometimes in the archive, the primary record would be the age of the sample. And then very often something would be done with the data, they would be manipulated in some way. So if you are making geo-chemical measurements on these things, each geochemical instrument will give you lots of data. But ultimately you will manipulate the data into one value. And that one value would go into the archive."*

This suggests that for putting this type of data into an archive the dataset does not necessarily need to be a complete set or raw data but the most useful. The following comment seems confirm this, as Ryan (quaternary) said the data has to be the most useful especially when putting into a journal archive.

*"so often this would normally have to be simplified from the version, a working version. Just to summarise the key information otherwise it is too complicated. What you put on there must be something that is useful, the most useful thing."*

In general, field measurement is a key research method in physical geography. The interview analysis showed that data can refer to different things in research, e.g. physical objects collected from fieldwork, such as mud samples, a sediment core, through to recorded measurements. It was suggested that the complex notion of data can be understood from two aspects, e.g. through the form and content of data. Some characteristics of data of this research method are also presented, such as numerical, continuous and with large number of variables. Furthermore, the interviews revealed that raw data are not necessary the most useful type of data. In fact, it is common to put processed data into a data archive.

## 7.5.2 Remote Sensing

Besides field measurements, remote sensing and GIS are also common techniques used in subfields of physical geography. Interviewees Michael (Glaciology) and Rachel (remote sensing) used remote sensing as their main research technique. Broadly speaking, remote sensing involves obtaining information of features of the earth's surface or atmospheric environment from a distance, using sensors to record or observe the reflected or emitted electromagnetic energy (Clifford, 2010; Mather & Koch, 2011). These sensors could be mounted on satellites, aircraft etc. Remote sensing techniques are used in different areas of research in physical geography. In Rachel's research, she was developing remote sensing techniques for mapping and monitoring vegetation. As she said:

*"I look at using remote sensing techniques and to try to detect, map and monitor [...] vegetation. [...] I am trying to develop methods to make that possible, to make it a quick and easy job really, because nowadays people first need to go out into the river and take out the plants. It's a very labour-intensive way of doing it. And I thought maybe we can map it from above by using remote sensing techniques. It may be easier to do."*

For Michael's (Glaciology) research, he used remote sensing to examine glaciers and changes in glacier landform, as he said,

*"I study glaciers in the present day and on the earth now and I study the changes of them and how that may or may not be linked to climate change. [...] I use the technique called remote sensing as well which is using satellite observations to collect data."*

### Collecting data- images

The interviews show that information or data acquired by remote sensors are essentially in the form of images, which is different from numerical data of field measurements. The research technique is extensively involved with image data and image analysis. As Rachel (remote sensing) said, *"what I do is image analysis and I work with image data"*. The interview analysis shows that there are different ways to acquire these remotely sensed images. According to Clifford (2010), there are three types of remote sensing-platforms where the sensor can be installed: firstly, ground –based devices, such as handheld devices; secondly, aircraft, such as helicopters, airplanes; and thirdly, satellites. Rachel (remote sensing) usually uses the first two types. Similar to the field measurement technique, she went to the field using UAVs (unmanned aerial vehicles) or cameras to take images, as she said:

*"Sometime I use satellite data but most of the time I use data collected from the UAVs or just a camera on a pole."*



She then explained what kinds of data she usually collected and used in her research,

*“So they are always photographs but sometime they are multispectral data. [...]The normal colour photo consists of red, green and blue. So there are basically three layers of information. So basically you’ve got three on the same photo, one layer contains information about red light, one about blue light, one about green light. If you combine those you get a colour photo. But you can also add more layers, more dimensions to your data, like infrared or any other part of the electromagnetic spectrum that you can photograph with your sensor[...] So with hyper spectral data you can have up to 250 layers of information, different parts of the spectrum that record different bands. And so because different surfaces reflect light in different ways, you may be able to get information .. you may be able to detect different surfaces based on that different data layers that you got. [...]I now tend to mostly use normal colour cameras RGB, red green and blue and I’ve got converted camera as well with infrared. So I basically use multispectral data, colour, visual light and infrared light. That’s my most commonly used data.”*

As the comment showed the characteristic of data from remote- sensors is visual, image-based (such as photographs, images). The following is an example of multispectral data (Figure 7.2):

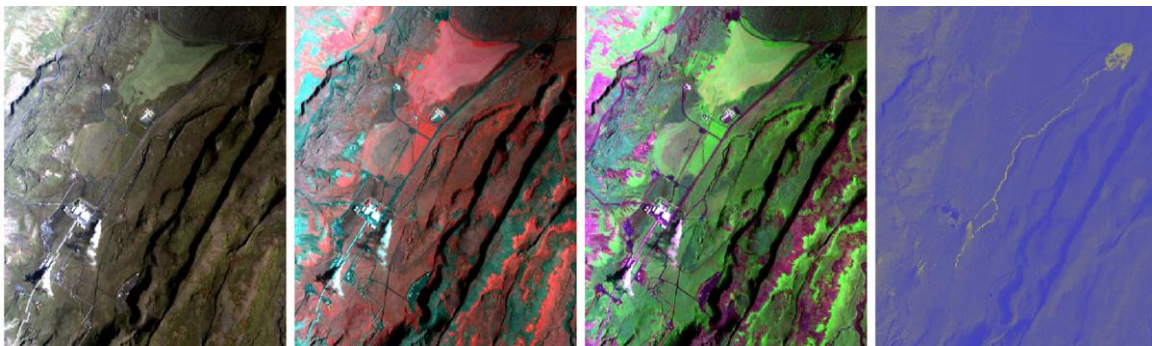


Figure 7.2 Multispectral infrared data (from left to right: visual, near infrared, short wave infrared, and thermal infrared)  
Source: Stevenson, J (2013)

As mentioned earlier remote sensing images not only can be taken using ground based devices or aircraft, they can also be taken by satellites. For example, Michael (Glaciology) uses satellite images for his research. As he said,

*“The main research technique I use is the images of the earth taken by satellites. So I guess most people these days are familiar with google maps and google earth. And everything on there are either aerial photographs or they are satellite images taken from space. And that’s kind of imagery that I would use. I don’t use google earth. But I get satellite images from other sources.”*

### Primary data vs secondary data

Remote sensing images which Michael (Glaciology) uses can be seen as secondary data sources. This means Michael does not collect data by himself. In contrast to Rachel (remote sensing) who usually goes to the field and takes images using cameras or UAVs, Michael (Glaciology) uses existing satellite sensory imagery available on websites. For instance, from the US Geological Survey's Landsat Look Viewer<sup>4</sup> or Global Land Cover Facility<sup>5</sup>. It is noted that a lot of the organisations which provide data are government or private agencies. Michael (Glaciology) explained where he gets satellite images and pointed out that while some images are freely available, some high-resolution images can be quite expensive. As he said,

*“You can get them from websites. They are all around the world. There are organisations [...] effectively download the imagery from the satellite to a ground station on the earth and they archive them and process them and make them publicly available. Mostly Landsat imagery are free. And anybody with a little bit of training can easily access them and display them. Other imagery sometimes you have to pay for it and it can be quite expensive, usually the higher-resolution imagery can be quite expensive and require some more sophisticated expert processing and steps.”*

### Processing data

Collecting images is only the first stage of applying remote sensing techniques. Collected remote sensed images as raw data requires processing and further analysis. For example, Figure 7.3 shows the workflow of producing a map using satellite imagery. It shows that lots of steps are involved in order to develop a research output from raw data.

---

<sup>4</sup> US Geological Survey's Landsat Look Viewer <http://landsatlook.usgs.gov>

<sup>5</sup> Global Land Cover Facility (<http://www.landcover.org/>)

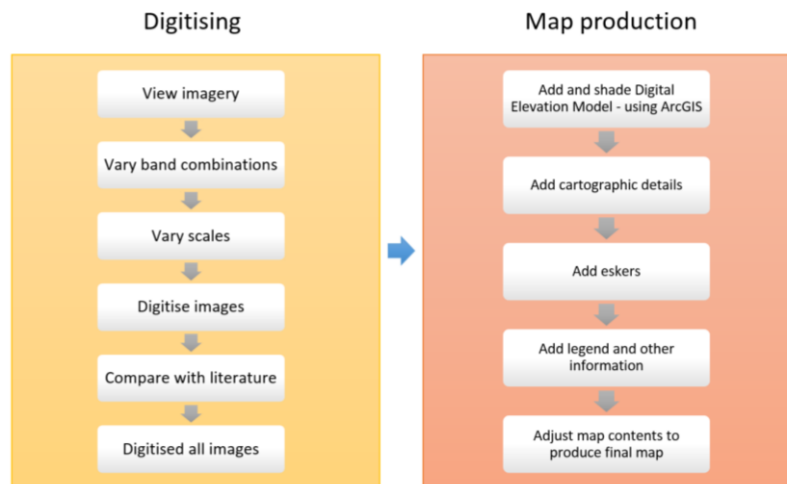


Figure 7.3 Workflow of producing a map using satellite imagery data (adapted from Storrar et al, 2013)

The interview analysis found that software such as ERDAS<sup>6</sup>, ENVI<sup>7</sup>, eCognition<sup>8</sup> were used to analyse and manipulate remote sensed images and ArcGIS<sup>9</sup> was used to produce maps. Rachel (remote sensing) also used eCognition to analyse her data. In the interview she described how she extracts information from her data using object based image analysis, she said,

*“I use computer software to do image analysis. [...] I use ENVI most of the time. I also use eCognition which is object based image analysis [...] So I look at groups of pixels and what sort of shape they formed in an image. So think of an aerial photograph of a city, you may be able to detect by looking at pixels that form a rectangle together. They could be the roof of a house. [...] By grouping pixels that are similar [...] you get objects. And you try to analyse those objects to look at their shape, their size, how they are related to each other. So that’s object based images. Apart from the spectral information-how much light is reflected from the surface, you also use sort of groups of (pixels) and get in extra information, another layer of information out of that (the images).”*

<sup>6</sup> ERDAS Imagine <http://www.hexagongeospatial.com/products/remote-sensing/erdas-imagine/overview>

<sup>7</sup> ENVI <http://www.exelisvis.co.uk/ProductsServices/ENVIProducts/ENVI.aspx>

<sup>8</sup> eCognition <http://www.ecognition.com/>

<sup>9</sup> ArcGIS <http://www.arcgis.com/features/>

## Data and information

Similarly, Michael (Glaciology) also extracts data from remote sensed images, as he said:

*“usually most of my research involves extracting data from the satellite image. Sometimes that’s very simple and is a question of on the computer drawing around features and mapping the glaciers and sometimes you have to do some a little bit of processing of the image to highlight ice or snow or you can do other techniques, like you can compare different sets of images to extract the movement of glaciers and things like that. Or elevation changes of the glaciers as well. Most of the time I’m looking at satellite imagery and I am extracting data, mostly in a map form but sometimes in other information as well.”*

It is noted that Rachel (remote sensing) used the term ‘information’, whereas Michael (Glaciology) used ‘data’ to describe what was extracted from remote sensed images. Here, it seems the terms ‘information’ and ‘data’ can be used interchangeably. Despite some literature suggesting it is essential to differentiate between data, information and knowledge (Zins, 2007), this seems to suggest the difference between the notion of data and information is not always noticeable. This shows the complex nature of data. Research data is not easy to define, especially as showed, data evolve throughout the process of research.

## Research output

The extracted information from remote sensed images can be seen as semi-processed or processed data of research. Similar to field measurements, the interview analysis suggests that data put into a data archive/ repository is not always raw data but processed data. As Michael (Glaciology) explained the data he put into data repository were maps and measurements from the image analysis rather than the raw images, as he described:

*“Some of the data may just be very very big tables of information of say if we measured 200 glaciers in Greenland from 1970s to 2010, for each glacier you may have the name of the glacier and then the measurements for each and then the next glacier and all the measurements. So often you may see the data like that in a table. But also you may provide the actual map with your lines where you map the glaciers, you may provide that electronically.”*

Interestingly, in the interview, Michael (Glaciology) also pointed out that the maps as supplemental materials on journals are not necessarily supplemental but key research outputs. In fact, Figure 7.4 was from the supplemental content of an open access article. Yet, as the title of the article, “A map

of large Canadian eskers from Landsat satellite imagery”, suggests the ‘supplemental’ map actually is a key output of the research.

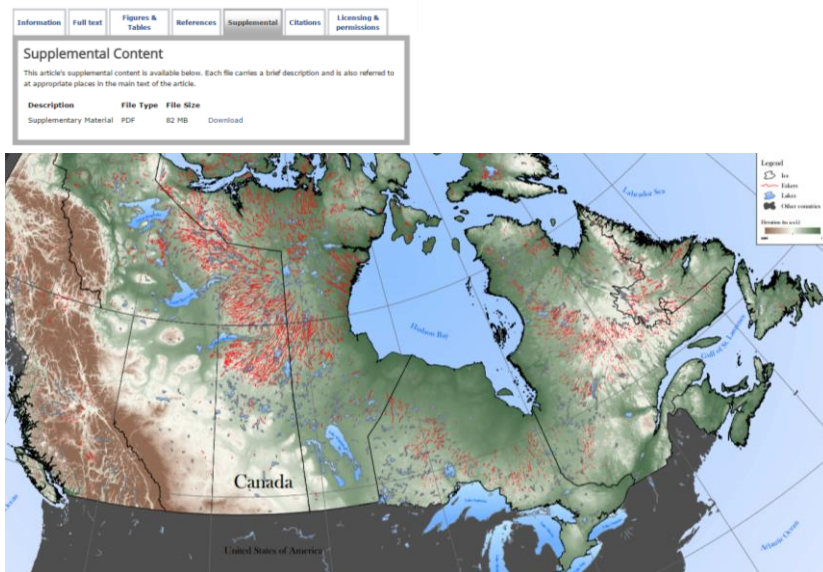


Figure 7.4 Supplementary material (Source: Storrar et al. 2013)

In general, remote sensing is another common method use in physical geography. The interview analysis showed that data from this research technique is commonly in the form of images. Researchers who use this research method either collect primary data, e.g. collecting remote sensed images using ground based devices, or secondary data, e.g. using satellite images that are available online. Raw images are continuously processed and analysed. Further information / data is then extracted from the remote sensed images. An interesting subject that emerged from the interviews was that it is not clear that the processed data is information or data. This confirms the nature of data is abstract and complex. And it seems that the perception of what is data is not strong among interviewees. The interviews further showed data being put in data archive is not usually raw images but processed data or even key research outputs.

### 7.5.3 Numerical modelling

Numerical modelling is another popular method used by physical geography interviewees. According to Clifford et al (2010), numerical modelling has been used extensively in physical geography. In this study, physical geography interviewees, William (hydro climatology), Christopher (climatology, quaternary) and Kevin (fluvial geomorphology) used numerical modelling in their research. Kevin (fluvial geomorphology) was interested in river modelling, as he said:

*“I make numerical models of how weather systems change over time. I’m interested in how rivers shape the surface of the earth, how they change the landscape.”*

Christopher (climatology, quaternary) was interested in ocean modelling and climate change, as he said:

*“I have a physics and maths background [...] So what I do is very much oriented to looking at modelling of the oceans and how the climate affects the ocean [...].”*

And William (hydro climatology) was interested in climate modelling and data analysis. As he described his research:

*“..essentially it would typically involve some form of modelling or data analysis. [...]and sometimes I ‘ll have to write my own computer code to process raw dataset or to develop and run my own models.”*

Broadly speaking, numerical models involve “a series of quantitative relations among variables that represent the underlying theory of the system being studied” (Gomez & Jones III, 2010:360). In simple words, as Michael (glaciology) described,

*“numerical modelling (is) where you’re essentially writing a series of code and equations to explain the processes that you are trying to understand”.*

He described two aspects of modelling, one is to develop a model and the other is to answer specific questions. As he described it,

*“there are two sides to the sort of modelling side of things. So there is technical development, so if you are developing a new aspect or facet of a process to model or new model themselves to study something new. [...] Then the second research method or practice is answering this specific question, having a research question that you wish to answer.”*

### *Ways to obtain data sources*

In order to create, develop or evaluate a numerical model, suitable data need to be obtained and processed (Gomez & Jones III, 2010). Modelling is a method which quite often requires a lot of data, including data from secondary sources. William (hydro climatology) and Christopher (climatology, quaternary) pointed out that data can usually be obtained from a different range of sources. The analysis found that one of the key sources of data is the public domain. As William (hydro-climatology) said:

*“So it may come from public domain datasets. In the US for example, any research which is being paid for by the tax payer is put into a public domain. So in the US, there is a*

*huge amount of public data on weather and climate. And I have been looking at today fire information, river flood data all sorts. So there's lot of that."*

Kevin (fluvial geomorphology) who usually obtained data such as rainfall data, temperature data or climate data explained that government data are increasingly available for free. As he said:

*"The UK is becoming free with its availability of government data. United States of course has a very long history of all government data is free. And this country is slowly catching up with that. Things such as the release of LIDAR Data for now is free, very useful for academic purposes...hydrological data, climate data, these types of things. "*

When Christopher (Climatology, Quaternary) was asked places where he usually looked for dataset, he answered:

*"probably not one place... because I deal with the number of different areas. But there are sort of set locations I tend to look so the global geological database, I will go look there and there is a couple of specific data centres for environmental data of the current date or the last century. So there's one at Columbia University<sup>10</sup> I tend to go to quite often or the IPCC their climate data centre<sup>11</sup>, which also has modelling data."*

Significantly, besides obtaining data from the public domain, interviewees also can gain access to data by direct contact or even by collaborating with someone who has access to some specific data. This shows that the intellectual and social aspect of an academic field are interrelated. In order to obtain research data, researchers collaborate with people who have mutual research interests and have the right sort of intellectual content of data. As interviewees said:

*".. in specific areas you might seek out people who work in that area and ask if you can use some of their data" - Kevin (fluvial geomorphology)*

*"Sometimes you can gain access to data through who you are collaborating with. So some of the work I've done in central Asia has involved working with water ministries who hold the data and we work with them, so we have gained access to the data. So there's lots of different ways that can happen." -William (hydro climatology)*

---

<sup>10</sup> International Research Institute for Climate and Society / Lamont-Doherty Earth Observatory (IRI/LDEO) Climate Data Library - <http://iridl.ldeo.columbia.edu/>

<sup>11</sup> Data Distribution Centre of the Intergovernmental Panel on Climate Change (IPCC)- <http://www.ipcc-data.org/>



This need of having access to others' data may contribute to the culture of informal data sharing in physical geography which will be revealed in chapter 9.

The analysis found that similar to the remote sensing method, raw datasets can not only be obtained from secondary sources but also from field work. For example, William (hydro climatology) not only obtained data from the public domain but also collected his own data from field work. As he said:

*"..there's also an element of research which involves installing equipment in the field and collecting raw data. So on the wall there, you can see a map of two rivers in the Peak district where we installed about 70 devices to measure air and water temperature, every 15 minutes."*

### Challenges

The interviews show modelling researchers face some challenges for handling their data. As interviewees commented, one of the issues of using other's data is related to format of the data. For example, William (hydro-climatology) said,

*" There is a huge amount of data out there. It is not always in a user-friendly format, often one has to apply post-processing to the data."*

Similarly, Christopher (climatology, quaternary) said:

*"When you have a lot of data, sometimes the datas come in a very different format and so actually being able to read data can be a challenge to start with."*

In fact, Christopher (climatology, quaternary) further commented not only on the different format of data but also the huge volume of data poses challenges to a modelling researcher. He explained, while some obtained data, such as field measurement data, can be simply analysed using spreadsheets, some obtained data is in huge volume, large files. It requires huge computer power, storage space and special software to process. As Christopher commented,

*" ..you have so much data, being able to store it and having enough computer power to process it can sometimes be an issue. [...] you have to analyse it by computer, using programming rather than... is too big to fit to spreadsheets"*

He then mentioned the software he usually used for processing and modelling. He said:



*“I often use MATLAB<sup>12</sup> as a processing and visualising tool and to do some modelling. But also use Fortran programming to do analysis of particular sorts of files.”*

### *Sharing model data and output*

One of the characteristics of model data is its huge volume. The analysis found that this characteristic of model data shapes data sharing practices of researchers. Because of its volume, model data is not easy to share. As Christopher (climatology, quaternary) said, despite the fact that journals are increasingly requiring researchers to share their research data, model datasets are not often shared in journals’ archives because of its volume. Interestingly, model data outputs in the form of images are usually shared instead. As Christopher (climatology, quaternary) said,

*“(journals) require less for model data because is too large. If I am collaborating with people who collect field data, often the field data will end up in a data repository or supplementary material online. But the model data tends to end up being images rather than raw numbers in many cases because it’s just too large amounts.”*

This again shows data is actively processed and the data being put in data archives is not always raw data but selected and processed data or output. Kevin (fluvial geomorphology) described what model data output can be, as he said,

*“a model output, it may be an image of a landscape, it may be a map of whether it’s been say changes or erosion or whether there’s water or no water or the speed of the water, often it is quite graphically output, as a visual as sort of map type format. You are looking at [...] landscape developed over time, so you need a way of squashing a lot of data into something you can use to analyse the results quickly and that’s using your eyes and your brain.”*

Output of models are often very visual, i.e. in the form of images. An example of image output of models<sup>13</sup> can be found in Coulthard & Van De Wiel (2012). In the interviews, William (hydro-climatology) and Kevin (fluvial geomorphology) mentioned they shared their actual model software

---

<sup>12</sup> MATLAB <http://uk.mathworks.com/products/matlab/>

<sup>13</sup> Examples of model output <http://rsta.royalsocietypublishing.org/content/370/1966/2123.figures-only>

on their personal websites. For example, when Kevin (fluvial geomorphology) was asked whether he had ever made his 'data' openly available, he answered,

*“Yes! The model that I work with, I developed, that’s freely available [...]So it is released with [...] a general license so anyone can use it or modify it.”.*

Strictly speaking, the model software which Kevin shared on his website is not data but the product/output of research. This again shows the concept of data is ambiguous.

In general, modelling researchers commonly collect huge amounts of data and quite often their work involves using secondary data. The interviews found that the characteristics of model data, i.e. huge volume and diverse data formats, challenge researchers to deal with the data. As the interviews revealed, using this data often requires high computer power, large storage space and special software to process, it is therefore likely that researchers who reuse this type of data will face similar challenges. Furthermore, the interviews showed that the concepts of data and research output are ambiguous among modelling researchers.

## 7.6 Research methods in human geography

While physical geography interviewees use solely quantitative research methods, human geography interviewees use both qualitative and quantitative methods.

### *Qualitative methods*

The dominant use of qualitative methods is one of the intellectual characteristics of human geography. In this research, 9 out of 12 human geography interviewees used qualitative methods in their research; nonetheless 5 out of 12 participants used quantitative approaches (see Table. 7.3).

The interviews show that almost every physical geography interviewee had a key research technique such as, physical measurements, GIS, remote sensing or modelling, which they are specialised in. In contrast, the analysis found that qualitative interviewees do not usually specialise in one single qualitative method but often apply a range of qualitative methods, such as, interviews, observations, secondary analysis, in their research. For example, human geographers Kimberly (environment, gender) and Brian (urban) mentioned a range of different methods, such as, interviews, ethnography, focus group and document analysis, when they were asked what research methods they usually use. As they said,

*“I am a qualitative researcher. And most of my data will come from documents [...], interviews, and also from ethnographies, kind of. They go together really.” - Kimberly (environment, gender)*

*“I tend to be a sort of very qualitative person, you know, case studies, usually interview based, semi-structured interviews, perhaps focus groups, [...] a little bit but not completely sort of ethnographic kind [...] of approach.” – Brian (urban)*

Interviews, focus groups, observations and secondary analysis are the most commonly used research techniques by qualitative human geography interviewees. These qualitative research methods were quite often used together, for example, interviews supplement archival work, interviews supplement observations. Yet sometimes these methods can be used independently. As development and political geographer Eric said, it depends on the project and research questions. Eric (development and political) used various methods, such as, textual analysis, interviews, participant observation, and archival work in different research projects, as he said:

*“It depends a lot on the project. So with things, work I’ve done on [...] protests that’s a lot to do with a kind of analysis of text or imagery. So it’s quite a discursive process, discourse analysis and so on, textual analysis and that. Other projects, so more around the citizenship projects often involves interviews, [...] focus groups, participant observation, so spending time in a community school [...] Other projects that have been based in communities- observing, participating, understanding what’s going on. So there’ve been bits of work I have done involving archival works going back through government papers in the archives and pulling out, kind of analysing what’s in those.”*

These findings that qualitative human geographers use a variety of methods suggest that different types of data would be created from qualitative human geography research. The following section will look at some common methods used in qualitative human geography and their data.

### 7.6.1. Interviews

Interviews are one of the most commonly used qualitative methods, particularly in social sciences (Clifford et al, 2010). As Thomas (historical) commented,

*“Social scientists mainly interview people like this all the time” .*

The findings of this study found that all human geography interviewees who employed qualitative methods used interviews. For example, Eric (development and political), Thomas (historical), Brian (urban) and Amanda (children) all used interviews (see Table 7.3).

### *Human participants*

Amanda (children) described interviews and focus group interviews are “*talking basically*”. The interview is “*a form of conversation between the researcher and the research participant*” (Secor, 2010: 195). Thus, one of the key features of this method is it fundamentally involves human participants. In contrast, physical geography research methods usually involve non-human research objects, i.e. physical research objects, such as air, mud, water. The characteristic of involving human participants may seem obvious yet it is significantly related to researchers’ practices and attitudes to data sharing. This could explain the contrasting data sharing practices between physical and human geography. For example, human geographers shared less data as they face more challenges e.g. ethical issues, to share data which involves human participants (see chapter 10).

### *Subjectiveness*

One of the findings to emerge from the analysis is that the nature of the method of interview is relatively subjective, especially, when compared to physical geography methods such as physical measurements. The subjectivity is mainly from two aspects: the content of the interview and the way the interview is constructed. The following will look at an example from Sarah (development, social, urban) on how interview research is conducted.

Interviews are usually used to gather people’s views, or experiences (Clifford et al, 2010). For example, human geographer Sarah used interviews to find out people’s life stories for her research on youth employment in Africa. As she described,

*“So a lot of life story interviews. So finding out particularly from the young people and trying to understand how do they end up in this employment they are doing today that involves finding out the facts of their life stories in terms of where they got to where they are today”*

Although Sarah (development, social, urban) used the term ‘*facts*’, the content of interviews, such as people’s experiences or life stories, is essentially views on opinions from interviewees. Thus, the nature of the interview content is subjective. The content can be difficult to compare, in contrast to physical geography methods, such as physical measurements which usually have standard units of measurement to measure things like air temperature or the chemical content in micro fossils.

### *Socially constructed*

Interviews are not only constructed by the responses of interviewees but also by researchers. The analysis found that all the qualitative human geographers used semi -structured or in depth

interviews for their research. None of them said they used structured interviews. For example, the following comment shows Sarah (development, social) used semi – structured and in- depth interviews. As she said,

*“We do face to face interviews, often do a lot of in-depth interviews, or semi-structured stroke in-depth interviews. But I tend to have just [...] a few keywords, few key questions I want to ask somebody and then try to make it as much as conversations.”*

Researchers gather data or information by talking and asking questions. Researchers usually design their own research instrument before collecting data, i.e. an interview schedule which has questions and/ or topics to use in the interviews. As the above comment showed, Sarah (development, social) used semi-structured/ in-depth interviews for her research. She had a few planned key questions and keywords for the interviews. Compared to structured interviews which have standardised questions for each interviewee, semi - structured interviews involve a predetermined set of questions yet with the flexibility which allows interviewers to explore responses by asking follow up questions (Secor, 2010). Data created from semi –structured / in depth interviews, therefore, is constructed by the presence of the researcher, for instance, what the researcher chose to ask, how it was asked, how framed or led in the interviews. This shows that the nature of the method is subjective. Each interview is constructed and unique. The data from semi- structured interviews, in-depth interviews could be more difficult to make comparisons than structured interviews, for example, comparison between different interviewees.

Besides the research instrument the setting and arrangement of interviews would affect interviewees’ responses, i.e. the interview data. For example, the following showed Sarah (development, social) arranged focus groups by gender or age in order to create a comfortable environment for participants to talk about their views. As she said:

*“We do focus groups and discussions as well so getting a group of people together. [...] -So we tend to do men and women separately and older and younger separately. So that you have... because often women won’t speak when men are present in certain cultures and also younger people won’t speak very much when older people are present. So we tend to have four (groups) within a community .. older men, older women, younger men, younger women.”*

Interviews involve interaction between interviewees and the interviewers and are constructed by both interviewees and researchers. The degree of subjectivity and flexibility is high in the interview method. As the content of data from interviews are shaped by both interviewee and the researcher,

interview data unlike experiments, cannot be easily reproduced. It is very unlikely to produce the same interview data, even with the same researcher, participants, research instruments and arrangements. As Holiday (2007: 61) pointed out, the interviewee's *"words belonged only to the moment of the research event"*.

### *Interview participants*

Another finding from this study suggests that data sharing practices can possibly be influenced by who are the interview participants. This could be one of the reasons why sharing data is not common within human geographers.

Interview participants are chosen in relation to one's research questions. The analysis found that qualitative human geographers interviewed a wide range of people, including, academics, protest groups, protesters, children, adults and 'old Italians' etc. For example, Eric (political) said his research involves:

*"..interviews with elites, interviews with stakeholders, interviews with general public, interviews with teachers, with students, with a range of people."*

Interestingly, there were quite a few qualitative geographers who mentioned interviewing elites. For example, development geographer Adam said,

*"interviews with key stakeholders and actors and actually that's generally been the way that I use interviews, rather than doing detailed interviews with community members [...] it's been much more about the interviewee being representative of political organisations or government departments. So it's been elite interviews or representative interviews that kind of sense."*

Furthermore, Kimberly (environment, gender) and Anthony (transport) also mentioned they have interviewed elites, such as, interviews senior women in oil companies, senior waste managers, leading academics in the UK. As Hochschild (2009:124) stated, the purpose of elite interviews is *"to acquire information and context that only that person can provide about some event or process"*. Therefore, because of the position of the person in the organisation or community, releasing the content of the interview would pose some ethical challenges for researchers. In fact, it is known that there are issues of managing confidentiality with professional and elite interviews (Dexter, 2006; Lancaster, 2016). Thus human geographers who conduct elite interviews could be less willing to share their interview data. This shows that how research methods can shape data sharing practices of researchers.

## Interview data: audio recordings, written transcripts

So far the previous section has focused on the characteristics of interviews as a research method. This section looks more into the actual data of interviews. In the following comment, Eric (political and development) described how he collected and processed his interview data in his research. He said,

*“Increasingly digitally, I suppose. So it is a bit like your digital recorder here. I would take my digital recorder, upload the files into a secure folder on my computer with a password protection and then once you’ve got time, either try to transcribe it yourself or if you have got the research funds [...] pay a transcription service to do the transcribing and then listen to the transcripts and read it and make sure it matches.”*

Interview data are commonly in forms of audio recordings and transcripts. The comment above showed Eric (political and development) treated his audio data, i.e. interview recordings, securely, as sensitive materials. The content of the interview is recorded, securely stored and then transcribed. The following is an example of an interview transcript from UK Data Archive<sup>14</sup> (Figure 7.5):

A: OK, it's recording  
B: OK. [xxx]  
A: the interview is divided into parts the first part is just a way for us to gather general  
opinions on GM, and the second part of the interview is a text, produced by the  
University and we would like to get your reactions to reading the first three  
paragraphs  
B: /OK/  
A: OK, can I ask you, as the first question what's your, what's your interpretation  
what is in your opinion the, what is behind the public response to GM in Britain?  
B: (S) mistrust, I think, of (S) organisations and then, people who are, seem to be  
in authority, and in a position of special knowledge, and I think it's, stems from, a  
number of, origins.  
A: mhmh, in the field of, in the, related field of food, scares or  
B: well I think yes I think, things like BSE and so on had a, an impact but I think also  
there's a deeper, deeper resistance to, the way in which, society works in the UK to  
some extent I think, there's a perception that, what was MAFF, wasn't very open  
wasn't very clear had it's own agenda it was on the side of agriculture if you like not,  
what not thinking necessarily in the public interest, I think those are really the, things  
that, make people, nervous or made people nervous about GM  
A: mhmh so there would be maybe a lack of trust in the government do you think also  
a lack of trust in, the people who are actually involved in the research? [someone  
knocks at the door]  
B: yes, can I? [goes to open the door]  
A: would you like me to switch this off [talking about the microphone] [they resume  
the interview]  
B: so the question was?  
A: so there is a, there seems to be a lack of trust you were mentioning in the  
government probably, or on the part of the bodies that [xxx] do you think this lack of  
trust extends also to, people who are actually involved in the research? or  
mhmh so there would be maybe a lack of trust in the government do you think also  
a lack of trust in, the people who are actually involved in the research? or

Figure 7.5 An example of an interview transcript

As the comment suggested interviews are transcribed into written form. Interview data is essentially words. It could be argued that the audio recording of interview is the raw data rather than the transcripts. As Bailey (2008) states transcribing is an interpretative process, it involves judgement. For instance, a transcriber needs to determine how much detail goes into interview

<sup>14</sup> <http://ukdataservice.ac.uk/teaching-resources/interview/structured/extract-two>

transcripts, e.g. non-verbal features, pauses, laughter etc. Thus, strictly speaking, an interview transcript is processed data.

### *Processing data: transcribing interviews*

The findings found that interviews usually are conducted and transcribed by researchers themselves, except for a very few researchers who had research grants, such as development geographer Sarah who had post-docs to collect both interview and questionnaire data. As she said,

*“I also have post-docs who do data collection as well. Because.. I like doing my own interviewing but yeah it is simply that I haven’t got time. I mean I am going for a sample to Rwanda on Sunday, um but I am only going for a week. [...] I was in Ghana for 3 weeks [...]. But as much as nowadays to go out to do the settlement selection, do the introduction, participate in a few interviews, and then leave a post-doc there collecting the data.”*

Yet, Sarah (development) pointed out that she prefers doing her own interviews. Furthermore, for transcribing, she believed it is the best if researchers who do the interview transcribe the interview recording. As she said,

*“[...] Well the best is if the person who has done the interview or questionnaire does the data transcription or data entry themselves. But that is not always possible. But where if it is possible that’s what we do because they know the data.”*

Sarah’s comments showed she would like the interview transcripts to be able to capture the moment of the interview as close as possible. As she said it is the best for interviewers who ‘*know the data*’, who were present at the moment of the interview to transcribe interviews. Both the form of data, interview recordings and interview transcripts are a kind of a reduction of interviews, they cannot entirely represent exactly what happens at the moment of the interview.

Research practices of transcribing seem to be varied between human geography interview participants. It is interesting to note that not all geographers always transcribe the whole of interviews, especially researchers who do their own interviews. Anthony (transport) and Thomas (historical) both mentioned not all interviews are fully transcribed. For example, Anthony said,

*“..part of my other work is interviews...so those are usually transcribed in whole or partially.”*

Thomas (historical) explained one of the reasons of not transcribing in full is time. As he said,



*“other people probably told you as well, there’s such a lot to do that the easiest thing will be to not write up a transcript, just kind of listen to it and take out the quote if you need, it is naughty but.... that’s what people do and then also to not therefore kind of archive it properly. [...] I am sure that I know that’s a short cut that other people take as well.”*

He also pointed out that as interviews are not fully transcribed, interview transcripts i.e. the data, could not be archived or shared appropriately.

### *Technological influences*

Evidently, technology has changed how researchers do interview research and hence the data. The following comment showed how Thomas (historical geography) did his interview research in the past. He said,

*“We didn’t have recorders in those days. So I just made notes after the interview”*

There was no audio recording, and no data such as interview transcripts but only some notes from interviews. Evidently, the form of the data from interviews has changed. In turn the role of interview data and research practices has also changed. For instance, the presence of interview transcripts has possibly shaped the social characteristics of the discipline, for example, interview transcripts enabling more collaboration with others in human geography.

Technologies have had influences on research practices and data. Besides the usual audio recorded interview data, the findings found that two qualitative human geographers also have data in the form of video. For example, children geographer Amanda (children) videoed focus group interviews. As she said,

*“We did video things [...] we videoed the focus groups and the activities then we analysed the video of them.”*

Adam (development) also produced some video data from his research, as he said,

*“The major technique of the project we are doing on energy [...] at the moment is convening workshops with a range of different stakeholders. [...] We actually used video and then transcribed from video. We’ve used photos of post-it notes and then interpreted those post-it notes. So we’ve tried to record the interaction between people and then pull out the major features of that from those processes.”*

Adam’s comment showed that he used less typical methods among other interviewees. He not only used interviews and secondary material but also conducted workshops as part of his data collection.

Various kinds of data, such as interview transcripts, video transcripts, photos of post-it notes and secondary materials, were collected or produced in one single research project. This shows one single research can have data with different degree of processing but also different types of data.

### *Analysing interview data*

Interview data are generally analysed by using codes. The following comments showed how qualitative human geographer Kimberly (environment, gender) processes and analyses her qualitative data. When was asked how she processes her data, she answered,

*“by reading again [...] I either use my research questions as themes or look inductively for themes that arise and then I code and kind of group the themes [...] And I don’t use computer programs, I do it myself.”* – Kimberly (environment, gender)

What is interesting is that almost half of the interviewees said they prefer to analyse their data manually, i.e. on paper, not using software such as Nvivo.

*“I printed out a paper copy and go through and coded by hand. I know colleagues who use Nvivo and coding software. I have never learned. I quite like having my coloured pens and going through and marking things up. So that would be how kind of general deal with interview data and focus group data.”* --- Eric (political and development)

*“In terms of interviews, transcribing and then kind of formally going through those transcriptions, looking for codes then going through that. I tended to do that long handed with colour pens that kind of stuff.”* Adam (development)

*“.. I kind of record interviews and things like that and kind of coded in a kind of more manually .... I am afraid I am a bit old fashioned ... I know there are packages you can use to help with that.. But I tend to not do that.”* Brian (urban)

This is in contrast with physical geographers who typically use computer software, such as statistical packages to analyse data. Unlike data in physical geography, processed data of interviews is not always available electronically.

In general, the analysis showed some characteristics of interviews and their data. The method is the most commonly used among the qualitative human geographers. Interviews are essentially ‘talking’, thus it is impossible to avoid human participants. The analysis showed that the method of interviews is relatively subjective. Its subjectivity is from two aspects: the content of the interview and the way the interview is constructed. Interviews are not only constructed by the responses of interviewees

but also by researchers, e.g. how questions are designed and asked. The analysis also showed interviewing elites are common in human geography. Interview data are commonly in forms of audio recordings and transcripts. Interestingly, not all geographers always transcribe the whole of interviews. The characteristics of this methods, e.g. subjectivity, involving elites, all these can pose challenges for qualitative human geographers to share their data.

### 7.6.2 Observation

Observation is another popular qualitative method used by human geography interviewees. The majority of the qualitative human geography interviewees used observation in their research (6/10). Observation involves watching and listening to human behaviours. Interestingly, among the interviewees who did observation, all but one said they did “*ethnographic observation*” or “*participant observation*”. Participant observation or ethnographic observation is a form of observation where researchers observe and participate in real world activities. As Cloke et al. (2004:200) described, participant observation is:

*“research that involves ‘being there’ and ‘stepping into others’ shoes,’ as much as this is possible. It involves participating in and observing social life, and conveying this to others mainly through writing.”*

It is noted that the research output of observation is in a form of writing.

Participant observation is closely associated with ethnography. As Allsop et al. (2010) stated it is a core method of ethnography, yet, the majority of human geography interviewees do not do pure ethnography. As Paul (urban) said,

*“I wouldn’t like to describe it is pure ethnography but is that kind of approach.”*

Sarah (development, social) provides a possible explanation of why human geography interviewees tend to do ethnographic observation but not ethnography. She explained that her time spent in the field was too short to be counted as ethnography. She said,

*“I am not in the field long enough to said it’s really ethnography. [...] Nowadays I tend to be out in the field for maybe two weeks at a time”*

Surprisingly, although six of the interviewees mentioned they used observation or participant observation, when they asked what kinds of data they generated from their research methods, almost none of the interviewees mentioned observation data. Most of them only mentioned data from other methods, such as interview transcripts, documents from secondary analysis. This might

suggest that interviewees are not aware of or do not recognise observation data. It could be because researcher's research is not informed by formalised, textual data but informed by what they observe and learn in the observation.

The only two interviewees to mention observation data were Kimberly (environment, gender) and Amanda (children). They pointed out that their observation data were kept in their research diaries. As they said,

*"I'm interested in [...] ethnographic observation which was recorded in a research diary" - Amanda (children)*

*"if I interview people I am also kind of interviewing myself as I go along, so I keep a research diary. So all my observations and about settings and contexts and things"- Kimberly (environment, gender)*

From the interviews, it was not very clear what exactly geographers put in their research diaries. According to Allsop et al. (2010:209), participant observation data are

*"pieces of creative writing based on what you see, hear, and feel under specific and often unpredictable circumstances. This data isn't collected. It takes hours to write".*

Thus, a research diary can be written after the observation. Yet, it can also have fragmented field notes which were written during observation. As Montello & Sutton (2006) and Laurier (2010) described observation data are often built from field notes which are not often well written and cryptic. Human geographer Kimberly (environment, gender) suggested research diaries are not only records about settings, context or things which researchers observed, but also researchers' reflections, personal thoughts on the process of a research. Thus the nature of observation data can be quite sensitive and non- objective. And thus she would not want to share her research diaries with anyone. As she said,

*"I wouldn't be comfortable sharing my own research diary because I can be open and honest with myself in a research diary which I know nobody even my closest colleagues is not going to read the whole thing. But it makes me reflect on sort of quite sensitive and difficult areas. And if I knew it is going to be publicly available or more widely available, I wouldn't write it in there. And then I wouldn't be .. I wouldn't have that resource to draw on."*

In general, participant observation is popular among human geography interviewees. Data of observation is in a form of writing. However, the analysis showed that most interviewees did not recognise observation data as data. This could be because the nature of observational data is unlike

traditional data. These data are writings in research diaries. They are not only records about settings and things which a researcher observed, but they also consist of researcher's reflections and personal thoughts of a research. Thus, qualitative observational data can be quite personal, sensitive and non- objective in nature.

### 7.6.3 Documentary and archival research

Besides interviews and observation, another common method used by human geography interviewees is documentary research. Human geographers Adam (development, political), Eric (political and development), Kimberly (environment, gender) and Thomas (historical) used documentary materials for their research analysis. For example, Kimberly (environment, gender) said,

*"Most of my data will come from documents and so analysing documents."*

Development and political geographer Adam examined policy documents in his research on governance. He said,

*"..the research I am involved with at the moment, [...]one is re-interpretation of secondary materials, so there will be deconstructing or interpreting official policy documentation, comparing it across different types of secondary materials. A lot of my work-policy work over the years has been that using secondary materials and sources and interpreting sector."*

Examining documents seems popular in political geography. Eric (political and development), also a development and political geographer, used similar methods, he said,

*"with work I've done [...] that's a lot to do with a kind of analysis of text or imagery [...].- So there've been bits of work I have done involved archive works going back through government papers in the archives and pulling out, kind of analysing what's in those."*

The above comment is in line with Gilbert (2014) which supports the idea of documentary research as not only limited to analysing texts, such as policy or government documents but also includes analysis of visual documents such as images. Historical geographer Thomas also analysed visual materials in his research. He pointed out that because of the internet, more images are available for research use. As he said,

*“I did do some propaganda maps, working with propaganda maps [...] It’s more visual. [...] When I started not many people used images, these days because of the internet ... to find an image perfect, is anywhere, anybody use images all the time.”*

### *Documentary Data*

In documentary research, researchers collect and analyse documentary materials, thus the ‘documents’, such as text or images are principally the data of the research. Eric (political and development) described a wide range of data he used in his documentary research, as he said,

*“... Documents, pamphlets, lots of grey literature, documents produced from societies, organisations, pamphlets, posters, budget that kind of ... [...]. Images normally those are kind of cartoons kind of published press or a while ago ... a pack of playing cards that have cartoons, political cartoons on that..”*

For Thomas’ (historical) research, his documentary data included old newspapers, books and articles. As he described,

*“My research method is mainly archival [...]. I do use a lot of old newspapers, lots of old journals and old books, tracing ideas in archives. [...] Mine mainly deal with dead people, so I was just looking at what they’ve written.”*

It was found that documentary materials come from different sources. For example, Kimberly’s (environment, gender) documentary sources are mainly from the public domain. As she said,

*“Most of the documents that I used are in the public domain.”*

There are also other data sources mentioned by interviewees, such as, documentary materials from local authorities, private companies and archives. The sources of documents could be significant to sharing data. Depending on the source of documents, concerns related to sharing data could be raised, for instance, issues around copyright, and the ownership of data.

It was found that the analysis of documentary data seems more straight forward and simpler than other methods, such as interviews. As Adam (development) said,

*“I don’t use any formal coding system in the way that I would with interview data.[...] You tend to go through materials looking ... in a sense you’re looking for key themes, you are pulling out the key information related to those themes, you are comparing to other documents and so on.”*

Similarly, Thomas (historical) *'to put these ideas together and make sense of them' and create a piece of writing.*

*"[...]so how to put these ideas together and make sense of them. Is by going back and looking at what they have written and put these things together. So is mainly about the data. I produce writing about these things."*

This shows that document data do not require stages of data processing as in other methods as seen in physical geography.

In general, analyses of documents are not limited to texts but also imagery. These text or images are essentially the data. The analysis showed documentary materials can be collected from different sources, including, the public domain, local authorities, private companies and archives. Depending on the source of documents, there may be issues related to sharing data e.g. copyright and data ownership issues. Furthermore, the analysis showed that documentary data do not require a lot further processing. Thus, documentary data is more likely to remain its raw form throughout the research process.

### *Quantitative methods*

Although the majority of human geography interviewees used qualitative methods (9/12), five human geography interviewees used quantitative methods. Sarah (development, social) and Anthony (transport) used mixed qualitative and quantitative methods, whereas Samuel (political), Brian (urban) and Joshua (economic) used solely quantitative methods in their research. What is interesting is that human geography interviewees who use solely quantitative methods tend to use secondary data, i.e. data that have been collected by others, whereas interviewees who use mixed qualitative and quantitative methods tend to collect their own quantitative data. In general, research methods which were employed by human geography interviewees can be categorised into surveys and statistical modelling.

#### *7.6.4 Survey*

Human geographer Sarah (social, urban and development) and Anthony (transport) used surveys to collect primary data. Quantitative survey research tends to be large scale, involves collecting information from a large number of people. Although the data collected from surveys is less in-depth than qualitative in-depth interviews, the quantitative data would be able to make statistical generalisations (Bryman, 2016).

Human geographer Sarah (development, social) collected not only qualitative interview data but also quantitative questionnaire survey data. As she said,

*“I use quantitative and qualitative methods. [...] I prefer the qualitative in terms of methods.[...]But we also do questionnaire surveys. I try to keep them very short and precise”*

One of the characteristics of questionnaire survey is that it is very structured. A questionnaire survey and structured interviews are quite similar in some ways, for example, both have standardized questions. As Secor (2010: 145) states, a questionnaire survey *“can also be thought of as a conversation – but a very highly scripted one”*. The structured, standardised questionnaire allows researchers to compare across respondents. Thus, the reliability and reusability of data could be higher when compared to less structured interviews, such as semi structured interviews. The analysis also found that researchers do not necessarily collect questionnaire data in person. For example, Sarah (development, social) had research assistants to collect questionnaire data, as she said,

*“What I tended to do nowadays is do as many qualitative interviews as I can then have an assistant to do the quantitative questionnaires.”*

It is possible that because of the structured nature of a questionnaire, researchers do not carry out the questionnaire themselves. Yet, it could be also due to the sample size of the questionnaire, as sampling size of a questionnaire survey is usually larger than interviews (Bryman, 2016). Therefore, researchers are not able to collect all their own data. As Sarah (development, social) said,

*“if you do a questionnaire survey it is quick is fast [...]But the chances are ... that often you can't conduct them all yourself. So you got to use research assistants. And there is also risk that research assistants are just doing it quick and just ticking some boxes. You know there is a lot of insecurity and a lot of unreliability. But sometimes it is nice to know how many people do something or think of something so I do combine.”*

The above comment shows Sarah (development, social) is aware of some issues of the method. She was not sure of the reliability of the data. This lack of certainty in the method seems to affect her attitude towards reusing other's data. As asked whether she used data from other researchers later in the interviews, she answered,

*“I mean I use occasionally. But I don't really trust them. Occasionally some quantitative, like World Bank data or national level data GDP that sort of stuff, only limited because I don't have much faith in it.”*



This shows that researchers' perception of a method and the limitations of a method and could affect one's willingness to reuse data.

### *Diary record*

Besides questionnaire surveys, diary records can also be used as a research instrument in surveys (Wiseman, 2005). For example, Anthony (transport) used diaries to collect quantitative data. He collected records of how participants use transport. As he described,

*"A lot of the quantitative data I've worked with as been generated through diaries. Effectively, you ask people to keep a log of everything they have done on a given sort of time, years, all their activities. Where they have come down, with whom, for how long etc."*

The method of diary record is in fact quite similar to the method of physical measurements in physical geography. While the method of physical measurement measures physical objects, Anthony used diaries to measure human activities, such as transport uses. The key difference is that the former is about physical objects and the latter is about human subjects. As mentioned earlier, when topics are related to human subjects, privacy issues are likely to arise. Issues of confidentiality are commonly discussed in qualitative research, e.g. Corti (2000) and Kaiser (2009). Yet, the interview analysis of this study shows that issues of confidentiality are not only found in qualitative data but also in quantitative data. As when Anthony (transport) further described his quantitative record data, he said:

*"That data is quite complex to analyse but obviously is very sensitive and there's lots of privacy issues."*

Anthony (transport) did not only collect his own quantitative data but he also used survey data from secondary sources, such as, from government agencies. His comments below highlight some differences between secondary data and primary data of surveys. He said,

*"So when we use data that have been collected by others for instance, government agencies, those data will have been anonymised in some sense. And often what you see is that, we only get information on the geographic... sort of the locations of activities and origins and destinations of trips that are fairly aggregated. So in the UK context, we only get data for instance, the [...] post code level. So it is a very large zone. And not at the level of the actual street or blocks. If we collect the data ourselves, we do have that very detailed information but that also means we don't make it publicly available and we don't share it with others... We share with the research team .. we wouldn't share it otherwise."*

Anthony's comment suggests that primary data contains more detailed information, while secondary data are usually processed and aggregated. Due to anonymization, some information in the secondary survey data could be missing. While having more details in data will benefit researchers, Anthony pointed out he would not make his own data publicly available because of the rich details in the data. It seems that for human related data, the more detailed it is, the less likely the data could be able to be shared.

In general, the analysis showed that due to research topics of human geography are related to human subjects, not only qualitative data, such as interviews, observational data, but also quantitative data, such as, quantitative surveys, have confidentiality issues. The analysis also revealed some details of survey data are useful but are not sharable. It also showed one's perception of a method and the limitations of a method and could affect one's willingness to reuse data.

### 7.6.5 Statistical analysis and modelling

Economic geographer, Joshua, political geographer, Samuel and urban planner, Brian are the human geography interviewees who primarily only use quantitative methods. All three interviewees were involved in some sort of quantitative statistical analysis or modelling. For instance, as Samuel (political) said, his methods involve

*“taking quantitative datasets and apply statistical methods, [...] Commonly the family of methods around regression models where you are applying linear models to try to explain what 's going on in the population”.*

Whereas Joshua (economic) uses a diverse range of quantitative methods for his research on the digital economy, including econometric modelling and statistical modelling, as he said,

*“I use methods such as such as [...] spatial analysis, network analysis, econometric and statistical modelling.”*

Brian (urban) who is interested in urban and regional geography works with methods such as “computer simulation models”.

Methods of quantitative data analysis or modelling generally require input data. Interestingly, all three human geographers predominantly use data that already existed, i.e. secondary data. For example, as Joshua (economic) said,

*“I am using mostly secondary data, either from the usual statistical sources and also what we call nowadays big data. So I have done research on using data from mobile phone operators. I’ve done research focusing on internet infrastructure- IP links.”*

### *Data sources*

The sources of data which the three interviewees used for analysis and modelling are varied. It depends on their research topics. Brian (urban) mentioned some common sources of his data, he said:

*“Some datasets are routinely made available now to anybody. Some of them are made available to say the academic community. The ESRC for example, have lots of agreements with the Office of National Statistics and Ordnance Survey. So that data is routinely available to academic users. Some data, normally big data, we get special arrangements, special agreements with Transport for London with the Oyster card data. So it really varies in a sense in terms of getting the data. We do normally have to develop quite a lot of very specific links as it works for some of these data providers. “*

As the above comment suggested, the sources of data used by quantitative human geographers can be grouped into three main types: 1) open, publicly available data, 2) semi- open data, e.g. data specifically available to academic communities, and 3) private data, e.g. data from private companies. In fact, the sources of data are related to data sharing practices. For instance, Joshua’s (economic) data is mainly either from private companies, such as internet providers or from publicly available data. When he asked whether he shares his data, he answered:

*“Most of the data I use, with exception of the communications data, was [...] open data. So everybody could download that data. So for that reason, [...] for this kind of data there is no need to do it again, as it’s already available and so many are available online. [...] So the actual data is usually either open data or data that I am not able to share with other people.”*

Publicly available data is evidently a key source of data for the three interviewees. Joshua (economic) also pointed out data which is already openly available no need to be shared again. For Samuel (political), the two main types of data sources he used are from election results and survey

data, such as, census and British election study<sup>15</sup>. Both are openly available. In the following comments, Samuel explained one of his data sources, election results, used for his political research. He said:

*“Obviously, elections themselves generate data, through election results and so forth and a lot of that information is publicly available and is open access. [...] For instance, [...] the election results themselves down to the level of parliamentary constituencies but also information on the social composition of constituencies with census give you that. And for information on how hard parties campaigned -because under UK election, all parties must return how much it spends in each constituency and that has been publicly available. So for anyone who wants to get this you can actually get that data and ... there are lots on that campaign data. So that's information available for anyone. The election results and census lots and lots of people have access to.”*

As he suggested, election data are easily accessible, generally are open access and available for everyone.

What is interesting is that the above comment also shows Samuel (political) used the term ‘data’ and ‘information’ interchangeably. This confirms what was observed in the earlier section on physical geography, that researchers perceive data and information as the same or very similar. This might be explained by the fact that, especially in research methods involving secondary data, from what point data or information is considered as researcher’s data is complex and is not commonly explored. As Bryman (2016) stated, it is not always clear where the start and finish of primary data and secondary data is.

### *Data*

The following explores how election results became Samuel’ research data. As Samuel (political) explained, elections generate election results and electoral data. It should be noted that election results are not originally generated for academic research purposes. Under the open government data initiative, election results are available to everyone. Thus, Samuel was able to obtain the election results and use them as his key data source, as his research data. Samuel’s data of election results would be usually referred as ‘secondary data’ as they are not collected by Samuel himself. In general, secondary data used by academic researchers are not necessary created by academic

---

<sup>15</sup> British Election Study <http://www.britishelectionstudy.com>

researchers nor for academic purposes. Therefore, in a sense, everything, every piece of information that is available publicly can potentially be useful for academic research and potentially become researcher's data. Yet, not every researcher would be interested or find it useful. Samuel continued commenting on the election data, saying,

*“The campaign spending data fewer people have access to [...] because fewer have actually search it down. So I rather tend to be one of the people who get that first each time when it becomes available because we are really interested in it. And then supply to other people who ask for it. Having that data file, the electronic data file, if other people want it we would share with them without any requirements or questions. It's public data, there's no confidentiality issue.”*

Samuel is really interested in the data. He would like to be the first to be able to get the data. The above comment shows there can be different degrees of interest in publicly available data. As Samuel pointed out fewer people have access to the campaign spending data, not because the data is difficult to access but not many people have searched for it. It could suggest that good accessibility and availability of the data does not necessarily increase the use of data. As different researchers have a different degree of interest in particular types of content.

The quantitative data which human geographers collected can be generally divided into two types: data that is collected routinely and data which is not collected repeatedly. Brian (urban) gave some examples of these two types of data, saying:

*“Some of it (the data) is routine, like the census. [...] Census data which is taken in frequently, meaning every year or every 10 years. This is very important data. Some of it is one-off. So a transport survey, for example, how many people are using a bus system? Often they were asking of questions of bus travellers. That kind of data is one-off. It isn't repeated routinely.”*

Besides routine data and one off data, researchers also use a type of data which is relatively new, i.e. big data. Big data is commonly described as high-volume, high-velocity and/or high-variety (Laney, 2001)

Interviewee Joshua (economic) and Brian (urban) both used big data. Joshua used large datasets from third parties, such as internet providers and mobile phone operators. Interestingly, he explained he had always been interested in using relatively large datasets, as he said,

*“Nowadays big data is a very trendy term, right, everybody is using it. So if a few years ago, I was using.. let’s say quite large datasets with IP links from around the world. In terms of size its not that big as we called today, mostly 5, 10 million observations. But most importantly it was user generated data. So yes I always have an interest in digital data in order to understand cities and regions.”*

Joshua (economic) had already been using large datasets. The term big data seems to be more like a trendy term for him. Yet evidently, advanced technology is shaping and shifting the nature of data and researcher’s research practices. Brian (urban) described how the nature of data he uses is changing, saying,

*“We always use data but recently the data has become very big because computers have begun to spread out into the environment. [...] And they’ve been used in particular to control devices which sense information around us. So mobile phones, for example, [...] are sensing where people are travelling etc. [...] So if you took the tube and you use an oyster card then all that data is captured you know second by second basically [...] Normally these datasets are delivered to us by the people who collected them, third parties for example.”*

This new type of data, i.e. big data, has become accessible for researchers to use. In the interview, Brian (urban) described the big data he collected:

*“Oyster card data from the tube system [...] we have 3 months oyster card data for [...] the 2012 year the Olympics. We’ve got 2 weeks of oyster card data for a few weeks ago when there was a tube strike.”*

Although Brian (urban) collected big data set, he stated he doesn’t use real time data analysis. He said,

*“a lot of the real time data is not very meaningful to what we are interested in”*

He then further explained that the challenges of using real time data is due to the unstructured nature of the data, as he said,

*“The big data that’s been coming out on stream in the last 3 or 4 years, one of the central issues involving in getting it and understanding it, is visualisation as in visualising the structure of the data really. It’s so big, that you need to get your mind around it. [...]The big data tends to be very unstructured, meaning it is very flat. When we get big data, it’s not got any real structure in it.”*

The following comment shows the challenges of using unstructured data in more detail, he said:

*“.. We kind of know where it’s made from. We might know who made it and the location and the time etc. But any other attribute data is not connected to the data itself. If there is attribute data, then we need to take it from characteristics about the individuals that come from other data, other dataset. For example, somebody makes a mobile phone call. If we need to know what income group they are then we need to know if the mobile phone company has information on the user then we have to tap that to that data. Because we are not able to do that, because of confidentiality most of the time but .. So in some sense big data is very flat data, very unstructured and it needs a lot more work to make sense of it then some of the traditional datasets, such as census.”*

In general, human geographers do not only use qualitative methods; some geographers use quantitative methods, such as statistical analysis and modelling methods. The interviews showed the researchers using statistical analysis and modelling methods often use secondary data. The sources of secondary data were varied. While some of the data was publicly available, some was from private organisations. The source of the data essentially affects data sharing. For example, as one of the interviewees pointed out, there is no need to share data if the data is already publicly available. The analysis also shows good accessibility and availability of data does not necessarily increase the use of data. This is because different researchers have a different degree of interest in particular types of content. Furthermore, some challenges of using big data for geographers were revealed. The challenge is mainly due to the unstructured nature of big data. Generally, the interviews found that characteristics of secondary data, such as, the source of the data, the frequency of the data that is collected, the confidentiality and the structure of the data, all these characteristics can have influence on researchers’ reusability of data.

## 7.7 Conclusion

This chapter examined the intellectual aspect of geography as a discipline, such as, research objects and research methods of physical geography and human geography. The findings from interviews showed that both physical geography and human geography have their own research objects and methods. Key methods which physical geography interviewees used were all quantitative, such as, physical measurements, remote sensing and numerical modelling. In contrast, methods used by most human geography interviewees were qualitative, such as, interviews, observation and documentary research. Nonetheless, few human geography interviewees used quantitative methods, including, statistical modelling and quantitative survey. Different research methods and

objects used in different knowledge communities influence data characteristics, as they influence how data are collected, used and analysed. E.g. Numerical modelling data features requiring high computer power and large storage space, while the characteristics of interview data features requiring less storage space and relatively subjective content.

This chapter not only showed the intellectual characteristics of geography and characteristics of different types of geography data but also provided some insights of how the intellectual characteristics shape the nature of data and more fundamentally, what is data. The concept of data is interpreted widely differently among interviewees. For instance, interviewees interpreted research data as from mud samples, record measurements, images, fieldnotes to research output. The chapter, nevertheless, revealed data can be understood from two different perspectives: 1) the physical form of data, e.g. data in numerical or textual form and 2) intellectual content of the data, e.g. data about palaeography or data about town planning.

The chapter also showed that research data is not static. The nature of data change and evolve over time during the cycle of research. For example, quantitative data in physical geography, such as, remote sensing data, data from field measurements, is commonly actively analysed and processed during research. It can be raw data, semi- processed data or processed data. Physical geography interviewees explained that the raw data is not necessarily is the most useful form for data sharing, and that more processed forms can be preferable. The transformation of data during the cycle of research can also explain why the concept of data is interpreted differently among interviewees. And this leads to questions of what 'data', e.g. what form, content or stage of the data should be shared and would be useful to share. The interview analysis continually revealed some elements which could affect data sharing, such as, the subjectivity of data and the divisibility of data, e.g. human geography interviewee mentioned interview transcripts need to be shared as whole without being divided in order to be useful. It should be noted that the two characteristics are related to the content of data. One of the interesting subjects that appeared is that better accessibility of data would not necessary increase the reuse of the data. Researchers' interests are actually quite limited to their research fields. Thus, not all data are useful for all researchers. The characteristics of a field can become the border of one's research and shape one's data practices. It is noted that researchers' interests and methods are profoundly connected to both the intellectual characteristics and the social characteristics of their research field, such as, disciplinary consensus about methods and research problems. Thus, it is essential to look at the social aspect of a discipline in order to further understand researchers' data practices and their attitudes on data sharing.



# Chapter 8

## Social Characteristics of Geography

---

### 8.1 Introduction

This chapter explores the social characteristics of geography as a discipline. The social dimension is associated with social interactions and collaborations within knowledge communities. Social features, such as, research collaborations, have long been known to influence the development of knowledge and the culture of a discipline (Crane, 1972 ; de Sollar Price, 1963). Along with intellectual characteristics, social characteristics are an important dimension to the understanding of disciplines (Whitley, 1974; Becher, 1989; Fry, 2007, Repko, 2014; Holland, 2014). The aim of this study is to investigate how data practices of researchers are shaped by the nature of discipline. Thus, this chapter looks at the social characteristics of geography, including, the nature of interviewees' collaboration and their internal data sharing practices, i.e. how researchers share their research with their research teams and collaborators, in order to explore the interrelationship between social culture of geography and data practices of researchers.

### 8.2 Social activities in geography

Becher and Trowler (2001) used the metaphor of academic tribes and territories to describe the notion of disciplines. Academic researchers are not individually isolated but are members of academic tribes. Each research field has its own culture. Although the boundaries of academic territories are difficult to locate and define, disciplinary boundaries can be identified in social activities that researchers participate in. For instance, as human geographer Paul (urban) said

*"I tend to go to planning research conferences rather than geography based."*

The comment shows that he differentiates planning research from geographical research. There is a boundary between geography and planning.

#### *Physical geography*

Interviewees were mostly involved in activities in their own territories, their own research areas. The analysis found that all interview participants were involved in social activities which related to their research fields, such as, attending and presenting at conferences, being a member of international

associations and departmental research groups. For example, physical geographer Michael (glaciology) described some of his research activities:

*“The meeting I was at last week was the British branch of International Glaciological Society. So I had attended as many meetings as possible. [...] I would also try to get to the American Geophysical Union<sup>16</sup>- the AGU and sometimes the European Geophysical Union- EGU.”*

Attending conferences of the AGU and EGU are apparently popular among physical geographers. In the interviews, physical geography interviewees, including, Kevin (fluvial geomorphology), Matthew (quaternary), Michael (glaciology) and William (hydro climatology), stated that they attended these conferences and were members of the two societies. In the interview, Matthew (quaternary) described the conferences of AGU and EGU, he said,

*“The biggest conferences would be AGU, American Geophysical Union, which is yearly in San Francisco. That’s massive- that’s about 10 -11 thousand people go there. [...] It covers the whole area of geo- sciences. [...] And then there is a sort of European equivalent, called the EGU, the European Geosciences Union, which is held yearly in Vienna. It is not as big as AGU. And again it covers the whole of geosciences but it does have a very strong representation in environmental change.”*

The comment not only shows that the scale of the conference is huge but also that physical geographers participated conferences beyond geography, such as geophysical sciences.

### *Human geography*

The analysis found that conferences which human geography interviewees attended seem to be more specialised or more geography specific. For example, historical geographer Thomas attended specialist conferences, such as ‘*Holocaust studies conferences*’. Interestingly, none of the human geography interviewees mentioned attending big conferences in a broader field. The common important big conference attended or mentioned by human geography interviewees, such as Kimberly (environment, gender), Amanda (children) and Thomas (historical), was the Royal Geographical Society conference. For instance, when Kimberly (environment, gender) was asked what conferences she usually goes to, she answered,

---

<sup>16</sup> American Geophysical Union- an international scientific community of Earth and space scientists <https://sites.agu.org/>

*“The Royal Geographical Society with the RGS IBG which is kind of professional association of geographers in this country [...] And then after that, more specialist conferences. [...] A conference where it was mostly feminist geographers”*

Looking at conferences interviewees attended, it appears that researchers in physical geography have more presence in the wider field geo-sciences, whereas researchers in human geography seem to be more isolated in their own specialism. Figure 8.1 represents the social differences between physical and human geography. Physical geographers have interaction not only within geography but also with wider area such as, geo-sciences, while human geographers focus interacting with individual specialisms.

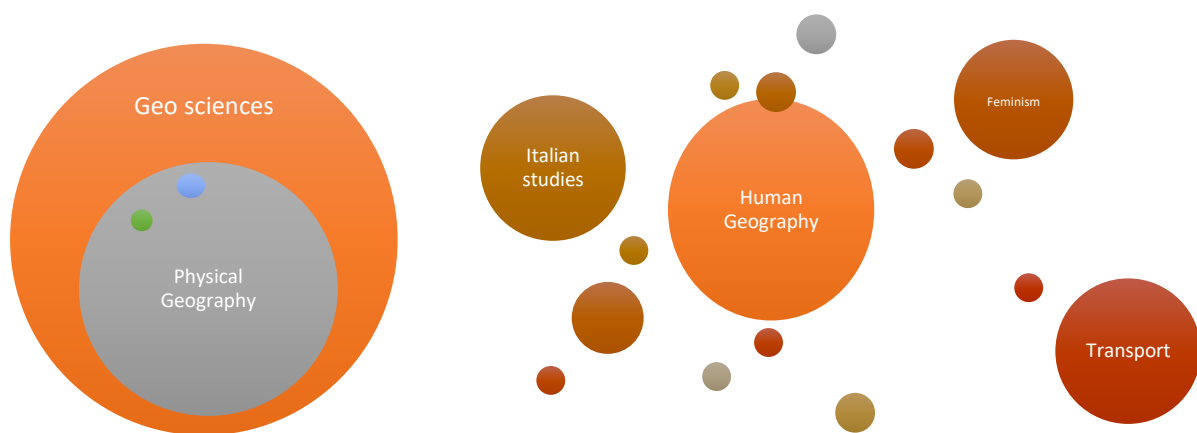


Figure 8.1 The social differences between physical and human geography

### *Urban and rural disciplinary environment*

The social differences could be explained by the fact that fields in physical geography are relatively ‘urban’, whereas fields in human geography are ‘rural’ using Becher’s and Trowler’s concepts (2001). As Becher and Trowler (2001) described, the people-to-problem ratio is high in urban academic fields. Urban fields are clustered around few central topics (Becher & Trowler, 2001). Thus, they are more likely to have overlaps with other fields and the boundaries between fields can be more permeable. In contrast, the people-to-problem ratio is relatively low in human geography. The scale of a rural field is relatively small. The following comments show that the scale of historical geography is quite small. As human geographer Dave (historical) described his research field,

*“Historical geographers are everywhere, it’s small-ish in the United States but there are historical geographers from France, Italy, Germany, Netherlands and Scandinavia, Japan,*

*India, Australia. Not that many. Around the world, there may be two or 300 perhaps. It is not a huge number.”*

Rural academic fields tend to be isolated, one of the reasons could be because there is not much commonality with other fields, thus less overlap with other fields. As Becher and Trowler (2001:185) stated, research problems of rural areas are “thinly scattered” across a broad area. The following comment illustrates some research problems can be very rare and specialised, as Dave said,

*“We need to do more work (on the research topic) [...] we haven’t done anything in the last 70 years on this[..] Hopefully someone may [...] do some work with the old history about them[...] you know in 100 years time [...] my work and their work can come together ..”*

This shows that Thomas’ (historical) research is not a common research subject.

Rural fields have less overlap with other fields. The distinctiveness of research topics, the specialism of the language, the isolated nature of rural field can influence rural researchers to be less able to engage with wider or different networks. For example, physical geographer Kevin (fluvial geomorphology) commented that the language of human cultural geography makes the field quite impermeable. As he said,

*“We tease each other all the time.. but I mean I have worked with social scientists here and in other places. I found a lot of the ideas and the concepts quite interesting. But [...] I think if you work with people ... particularly with human geography in some aspects such as cultural human geography.. [they] have a very bespoke language which I think does themselves no favours really. I think it makes it very impenetrable as a discipline to get into.”*

### *Conferences and research audiences*

Conferences which interviewees go to are an indication of the urban/ rural nature of fields, but also indicate the nature of interviewees’ research audiences. The analysis shows that interviewees participated in conferences not only for socialising, in search of collaborators, but also for disseminating research. Audiences for research and conferences are interrelated. For instance, when physical geographer Patrick (quaternary, biogeography) was asked who his research audiences were, he described conferences that he can go to. He said:

*“because my research covers different areas. I can choose which audience I want to put my work to. I can talk to different people [...] I can go to a conference on environmental change. Or I can go to conferences on diatoms- I’ve just been to a conference purely on diatoms. Or I can go to more big multi-disciplinary meetings.”*

Patrick (quaternary, biogeography) sees presenting at conferences as a way to disseminate his research to his audiences. Thus, conferences which a researcher presents at reflect their research audiences to some degree. Whitely (2000) pointed out that audience structures are closely associated with the nature of research fields. The urban / rural nature of a research field could shape audience structures, such as the diversity of research audiences. As suggested earlier, physical geography is urban in nature. Boundaries of urban fields tend to be more permeable. It is therefore likely that researchers in urban fields are able to present to a wider range of audiences. In fact, the above comment of the physical geographer Patrick (quaternary, biogeography) confirmed this idea. As the comment showed, Patrick as a researcher in an urban field, can present his work not only to specialists but also to people in wider areas. In contrast, human geographers in relatively isolated rural fields do not have much overlap with others. Thus, their research audiences can be quite limited to a relatively small group of people.

The urban and rural nature of fields and the range of audiences of research can have important implications for data sharing and re-use. As audiences of one's research, who is interested in one's research are probably the most likely potential users of one's research data. Thus, researchers in urban fields would have more potential users of the data they shared. For researchers in a rural field where audiences are limited, it is likely that less people would be interested in using their data and this could also influence their incentive for sharing data.

### 8.3 Collaborations in geography

Besides participating in conferences or being a member of international associations, research collaboration possibly is the most important part of a researchers' social activities. All interviewees collaborate to some degree. While the smallest formalised unit of collaboration in academic institutions is most likely the departmental research group as explored earlier in chapter four, the interview analysis shows that geography interviewees actually collaborate more closely with their own research teams rather than the formalised departmental research groups. Physical geographers Kevin (fluvial geomorphology) and Patrick (quaternary, biogeography) both described what is a true research group in the interviews. Kevin described,

*"From a science point of view, I think it ( a research group) doesn't have to be any sort of form or structure, I think it can be about a group of like-minded people who have common interests and are keen on working together. That's where the real synergies comes from."*

Patrick also commented,

*“I am a member of different associations, societies [...] that’s not the same as being part of a research group. We have a research group cluster here [...] But my research group as such I think is these people I work with, this multidisciplinary group, my colleagues in America, [...] I work quite closely with two people in the department. But my research contacts, my research capability is maintained by my external links. [...] but then again they are not formalised through any research groups in the sense of what’s a society.”*

As the above comments show, a true research group for interviewees is not the departmental research group but the group of people they collaborate closely with in research.

## 8.4 Collaboration in physical geography

### *Collaborative in nature*

The interview analysis shows that physical geographers are very collaborative. When physical geography interviewees were asked whether they often do collaborative research, all physical geographers’ response were: “yes”, “nearly all”, “everything”. For instance, physical geographer Matthew (quaternary) described his work, saying,

*“Many of the human geographers work completely by themselves, as you probably established. I would say virtually all the work we do is collaborative. Even small projects that I have got. I’ve got a number of small projects that are not funded. There all sorts of collaborative [work] with people in the department and other colleagues. Very Very little work is just me and nobody else.”*

Similarly, Michael (Glaciology) and Ryan (quaternary) also said most of their research is collaborative. The following comments show their collaborators are not only from their own institutions, but also from other UK institutions and abroad. As they said,

*“I’ve only written a few single authored papers. Most of my work is with multi-authors, both from the UK and further afield.”* Michael (glaciology)

*“Yes nearly all of it is collaborative. They are mostly from other universities but they do include colleagues from within this university as well. [...] So certainly in terms of large projects that I have been running they have been a lot of people overseas and a lot people from many other UK universities.”* Ryan (quaternary)

The interview findings are consistent with the results from the bibliometric study. Research in physical geography is evidently collaborative in nature and collaborations across institutions and countries are quite common. The implication of this can be physical geographers need institutional support to share data within research teams, especially with inter-institutional and international collaborations.

### *Collaborations and publications*

Collaborations, co-authors and publications are closely related. When interviewees were asked how big their research team is, interestingly, they quite often used the number of co-authors on publications to indicate the scale of their collaboration. For example, when William (hydro-climatology) was asked how big his research team, he answered,

*“If you look at the .. let’s say the number of authors listed on a publication, as an indicator of the size of the collaboration, then you know [...] collaborating with two or three other people. So that’s four names on the paper. That wouldn’t be unusual. There’s been several occasions when there’s been up to 20 people collaborating on a paper or even more than 20.”*

Similarly, Michael (glaciology) answered,

*“There are some people that I have published several papers with them. We have a good relationship and we keep working on the same problems. [...]so I think I have a core 10-15 people who I would publish with semi-regularly and a broader up to 50 other people who I have just published once or twice with.”*

The close link between collaboration and publications suggest that producing publications is essential for research collaborations. Publications are a fundamental end product of a collaboration. It seems that the key objective of researchers to collaborate is to publish. Researchers may see the value of publications, i.e. the end product of research, more than research data, i.e. the building blocks of research.

### *Reasons to collaborate*

The interview analysis also shows physical geographers collaborate with a wide range of collaborators from different disciplines. For example, as Christopher (climatology, quaternary) said he collaborated *“with geographers, with oceanographers, geologists, engineers”*. Similarly, William (hydro-climatology) said,

*“I have collaborated with all sorts of many many disciplines, possibly too many to list. [...] I collaborate with a geologist, ecologists, climatologists, engineers, [...] I think in the past I worked with people in building design, so architects yeah and planners of all sorts.”*

In the interviews, physical geographers explained why they do collaborative research. Interestingly, most interviewees stated that they collaborate because different skills are required for their research. A possible explanation for this might be that physical geographers tend to specialise in one key research technique, as revealed in chapter 7. This shows the collaborative nature of physical geography is linked to its intellectual characteristics. Physical geographer Matthew (quaternary) explained that the collaborative nature of physical geography is inherent to the nature of research problems which physical geographers explore. He explained that in order to solve the research problems, different research techniques and skills are always needed. He said,

*“To address any particular problem you need a range of different techniques and skills. And it is very unlikely, one person would have those techniques [...] You may be interested in the geochemistry, you may be interested in micro fossils and different types of micro-fossils. Some have to date the sediments. And all of those are different skills require different people. It would be very uncommon for one person to have all those skills. So the idea is you collaborate with people who’ve got the skills.”*

Requiring different skills clearly is a key reason for physical geographers to collaborate. As it was not only mentioned by Matthew (quaternary), physical geographers in other subfields, such as, Christopher (climatology, quaternary), James (GIS), Kevin (fluvial geomorphology), William (hydro-climatology), also stated something very similar. For instance, the following comment by William explained why he did collaborative research,

*“...because you are drawing on lots of various specialists, disciplinary expertise. So someone may be an expert on river water quality, someone on plants in river, someone on bugs in river, someone on fishing rivers, someone in sediments in river. And you bring all of those disciplinary areas together.”*

Matthew and William’s comment not only indicated that the nature of research problems in physical geography are holistic but also explained why physical geographers tend to collaborate with a wide range of people. As Christopher (Climatology, Quaternary) said, the nature of collaboration in physical geography tends to be *“a combination of different skills and different ways of looking at the problem”*.



Further analysis shows that there is another reason for physical geographers to do collaborative research. Michael (glaciology) explained that the reason he collaborated with others is to get access to someone's data, as he said,

*"Sometimes because people have datasets you want to work with and they know they are more familiar with the data and the intricacy. So it's rather just sent the new data and you are trying to work it all out, it makes sense to collaborate."*

Interestingly, the analysis shows that physical geographers who mentioned this reason, including, Michael (glaciology), Kyle (GIS) and Rachel (remote sensing) are mostly using techniques which involve secondary data sources. Yet, Rachel (remote sensing) pointed out that people who work with secondary sources of data, such as satellite data, which researchers do not need to collect from the field, do not have to collaborate with others. As she said it would depend on the nature of the research and the volume of data, she said:

*".. it's nice to collaborate with people but you don't need to. If there is lots of computer processing, if you work with a lot of large amounts of data then you will need people to help you to do the computer processing but if you are trying to develop an analysis technique, image analysis techniques is something you can very well do on your own."*

She explained that her research is to develop a technique, thus it does not require a huge volume of data. This shows that not all types of research are data intensive, requiring large amounts of data.

Kyle (GIS)'s research is another example of collaboration because of other datasets. When he was asked does he always do collaborative research, he answered,

*"Almost always...because I need to work with people who collect the animals and plants to do the DNA analysis"*

Kyle's research is to understand the evolving patterns of animals and plants using GIS and geo visualisation. As the above comment shows, he depended on other researchers' data for his research. While one may consider research collaboration is always a close collaboration, yet the below comment shows it is not necessary. As Kyle (GIS) said,

*"So there are collaborations but what has tended to happen is that they give me the data and I do all the work. So it is a collaboration.. where they are sharing data with me, [...] they are unable to do when you are writing practical paper...they don't know the computational background stuff."*

This finding shows that collaborative activities within physical geography are not necessarily 'tightly coupled' (Whitley, 2000), e.g. collaboration can take place without collaborators sharing datasets. This could be explained by the feature that was shown earlier, i.e. physical geographers collaborate with people who have different skills. As a result, it is not surprising that each collaborator could only contribute within their own expertise. For example, as Kyle (GIS) said,

*"They write most of their papers on those (their) subject matter and then when they come to the geographical component we are trying to collaborate on that... But because that's not their area of expertise, they can only put a limited amount of time and knowledge into it."*

The findings suggest that collaboration in physical geography consists of different separated tasks. Each team member has a specific defined role. This characteristic has implications for research data, for example, what research data is shared among collaborators.

#### 8.4.1 Internal data sharing in physical geography

The interviews not only provided a deeper insights of collaboration in physical geography but also what and how data were actually shared among the collaborators. Surprisingly, the analysis found that despite the collaborative nature in physical geography, physical geography interviewees did not share all of their research data with their collaborators. They usually only shared a part of their data. For example, in the following comments Christopher (Climatology, Quaternary) talked about the 'data' he shared in his collaboration. He said,

*"Sometimes [...] exchanging images that the different groups have produced, so not so much the raw data as the results. Sometimes there would be some data transfer as well. So I may have a model run and transfer some of that data to my collaborators for them to study an aspect of it."*

The 'data' which physical geographers shared is not necessarily raw data or even data, but research results or outputs, such as, images, maps, diagrams. It shows that physical geographers shared things which were most relevant to their collaborators. As Rachel (remote sensing) said, she shared her data only *"if it is relevant to them"*.

The interview analysis found that raw data is not always shared within collaborators, this could be because physical geography collaboration consists of diverse skills and separated tasks. As Ryan (quaternary) said, it is not necessary for collaborators to know all the details, he explained,

*“For example, at the moment, we are collaborating with the group who do climate modelling. So they have done something quite complex climate model runs, where all they share is the output in terms of some maps of the climate. So you know they say ‘here are the maps’. But the detailed program and so on they just keep themselves. Because it is too specialised. Similarly, the spreadsheet I generate with many of my collaborators I may just show them the diagram that results from it. Because they don’t need to know the details of the spreadsheet.”*

Not only Ryan but Mark (sustainability) also mentioned sometimes that data can be too specialized to understand for external collaborators. Thus, Mark tended to share his output rather than the data with his external collaborators, such as local councils. Furthermore, physical geographer Patrick (quaternary, biogeography) explained that raw data is not always needed to be shared within collaboration because the raw data has little use to collaborators before processing and synthesising.

In general, the interviews show physical geographers are collaborative in research. The findings are consistent with the results of bibliometric analysis. The interviews showed some insights into why collaborative research is popular in physical geography and some data sharing practices of physical geographers. The collaborative nature of physical geography can be explained by the fact that research problems in physical geography are holistic in nature and physical geographers tend to specialise in one key research method. Therefore, it is essential for them to collaborate with other researchers who have different skillsets in order to solve research problems holistically. The interviews further revealed that as physical geographer tend to have specialised and separate tasks in collaboration. Raw data is not easy to be understood and used by collaborators. Thus, it is more common to pass or share processed data or output with collaborators.

## 8.5 Collaboration in human geography

The nature of collaboration in human geography is different from physical geography, in terms of scale and data shared with their collaborators. The interviews show all human geographers did some collaborative research, however, they do not do collaborative research as frequently as physical geographers. When interviewees were asked whether they often do collaborative research, while almost all physical geographers’ responses were ‘yes, nearly all the time’, most human geography interviewees’ answers were, “yes, but not always”, “it varies”. For example, historical geographer Thomas and political geographers Eric answered,

*“Not very often. Probably about a quarter of my papers may be collaboratively authored.”- Thomas (historical)*

*“Yes I do quite a lot of collaborative work. If you measure in terms of publications, probably half my publications are co-authored collaboratively.”- Eric (political and development)*

As Eric’s (political and development) comment suggested, half of publications being co-authored can be considered as a lot of collaboration in human geography. Interestingly, the above comments also show, similar to physical geographers, human geographers relate collaborations and publications together. As suggested earlier, this indicates that the main objective of researchers to collaborate is to publish and a publication is seen as a key product of collaboration.

Evidently, human geographers do not collaborate all the time. Yet, there were a few exceptions: Samuel (political) and Brian (urban) answered, ‘almost all’ and ‘all the time’. Interestingly, Brian (urban) and Samuel (political) both used quantitative research methods. This seems to suggest that qualitative researchers are less collaborative, whereas quantitative researchers tend to be more collaborative. The nature of the method used by researchers shapes the social nature of their research. In fact, Joshua (economic) thought methods and collaboration are interconnected. As he commented,

*“I think it (collaboration) has to do with the methods..So it is more common for people who do quantitative research to have more collaborators than people who do qualitative research who usually [...] have single authored paper or less amount of co-authors.”*

And Joshua (economic) believed quantitative research is generally more collaborative is because, as he said, it could *“be easier to split the task of things that need to be done”*.

Although human geographers do not collaborate very often, they collaborate not only with geographers but also people from outside from the disciplines. For example, Sarah (development, social) said,

*“Of course you can get geographers just working in the area. But I work, I collaborate a lot with people outside the geography [...] with anthropologists, business studies, engineers, architects and people from education studies, media studies, development studies...”*

The interview analysis found that human geographers usually collaborate with only a few

collaborators in a project. For instance, when Samuel (political) and Mark (sustainability) described their collaborations, they said,

*“My main collaborator is...(Amy).. who’s a geographer but at the University of (Exeter).[...] probably 80, 90 % of my work is with (him).”- Samuel (political)*

*“Not very big- is me and my research assistant. But occasionally draws in some of my other colleagues in the department to get involved in parts of the project.”- Mark (sustainability)*

The interview findings confirm the results of the bibliometric findings. As shown in the bibliometric chapter (chapter 6), human geographers tend to collaborate with researchers outside geography and have a relatively small scale of collaboration than physical geographers.

The interview analysis also found that some human geographers prefer to work or write on their own. For example, historical geographer Thomas explained that he did not often collaborate very often and *“that’s partly deliberately because I like working by myself”*. Cultural geographer Amanda (children) also mentioned she prefers to write on her own. She doesn’t like writing articles collaboratively, such as each academic writes a few paragraphs. This is interesting, as this could suggest human geographers face more difficulties to separate tasks when collaborate with other researchers. There are less separate, defined roles in qualitative human geography collaboration than in physical geography.

#### *Different collaboration practices in quantitative human geography*

The interview showed that most human geography interviewees worked as a small team when they collaborated. Yet there is an exception. As mentioned earlier about Brian (urban) who uses quantitative methods, most of his research is collaborative. He worked with a team of 10 -20 people. As he described,

*“out of 10 people, 2 geographers, 1 planner, and then maybe 2 physicists, 2 mathematicians, computer scientists, maybe an engineer. So quite varied.”*

It is clear that similar to physical geographers, members of Brian’s group contributed different skills. Each member has well defined roles. For example, Brian said,

*“So the programmers would be involved in heavy programming, quite a few people involved in visualisation which is involved with some programming. And others will be involved in more the mathematical structure of things. Some people would be involved in*

*much more applied work. But all of the applied work would involve some programming, some data mining, some facilities with standard software such as geographical information systems. Many people here have different skills and different types of software basically.”*

As Brian (urban) said, it is very rare to have such a big research team. While the nature of the method, i.e. its quantitative nature, could be one of the factors that to explain the large collaboration, the analysis found that funding can be another possible factor. Brian explained that the research team was almost entirely funded from research grants. He said,

*“(It) is a group that really runs on research grants, soft money.. meaning research grants from the research councils, European community. So it depends almost entirely on getting research grants this group.”*

Evidently, the availability of funding has influences on collaborations in human geography. Anthony (transport) also stated that there is more collaboration in his field than in cultural geography because of the availability of funds. For example, when he talked about collaboration, he said,

*“my field, transport geography is different from cultural geography or to some extent political geography but I think particularly cultural geography where still it is often the case that people do their own research. And I think part of that is because in transport geography most of the research will be funded through projects.”*

The comment also shows that the availability of research funds is different in different fields. Research funders seem to prefer bigger collaborative projects.

### *Reasons to collaborate*

The analysis found that the reasons to collaborate are different between physical geographers and human geographers. As mentioned earlier, one of the main reasons for physical geographers to collaborate is because different skill sets are required for the nature of their research. In contrast, none of the qualitative human geography interviewees mentioned this reason. The interview analysis revealed some of the reasons for human geographers to collaborate, including, to *“share and test out ideas”*, *“get inspirations from others”*. Yet, what is interesting is that one of the key drivers for human geographers to collaborate is funding agencies, which was not mentioned by physical geographers. E.g. Eric (political and development) stated some reasons for human geographers to collaborate, saying,

*“there’s various reasons for that. For some people, it is because [...] they hold a research grant [...] For others, it would be like the project with (Amy) and (John), it is a big EU*

*funded project. It is just so big; you need to have collaborators. For other people, collaborations might be because of not being able to get to the field, got a young family or other factors to do with travel [...] but one of the big drivers in the field at the moment and this is partly driven by the funding agencies.”*

As mentioned in section 8.4, the nature of collaboration in physical geography is largely shaped by its epistemic nature. None of the physical geographers mentioned funding agencies are the drivers of their research collaboration. In contrast, collaboration in human geography is more driven by external factors, such as, research funding and university strategies. Adam (development) and Thomas (historical) both mentioned the number of collaborations had been increasing in the past few years. The following comment indicates researchers have pressure to do more collaborative research from institutions. As Thomas (historical) said,

*“Nobody really works, writes by themselves so much these days. [...] The message we are getting from management is collaborating, work with other people”*

The analysis found that human geography interviewees mentioned research funders’ influences on their collaborations much more than physical geography interviewees. Funding agencies clearly play a significant role in driving collaborations in human geography. As children geographer Amanda explained, it is essential to collaborate in order to get a big grant nowadays.

Research funders and institutions not only drive human geographers to collaborate more but also have influences on who they collaborate with. For example, Sarah (development, social) collaborated with non-geographers because of funders favouring of multi-disciplinary projects. As she said,

*“so most of my funding has been from DNIDA and they like interdisciplinary, multidisciplinary projects. So the youth employment project [...] when I was applying for that money we looked for someone else to collaborate with.”*

Although increasingly human geographers are driven to be collaborative, the analysis shows human geographers do not collaborate very often. The analysis shows that not all research is funded. For example, Francis research was not funded, she said,

*“As in my case I went out and collected the data having got no money. So you know I went out and spent my own money getting the data. So I don’t see why I shouldn’t have some ownership of that data really”.*

The comment shows that there are human geographers who do non-funded research and this could affect researchers' willingness to share data.

### 8.5.1 Internal data sharing in quantitative human geography

Internal data sharing takes place not only in physical geography but also in human geography. As human geographer Samuel (political) said, data sharing with collaborators *"usually is fairly open"*. As mentioned earlier, physical geographers tend to share part of their data in collaborations. The analysis found that this also applies to human geography interviewees, particularly interviewees who use quantitative methods, such as, Anthony (transport) and Samuel (political).

Deeper understandings of the practices of collaborative research provide some hints as to why quantitative researchers do not share all data even with collaborators. In the following comment, Anthony (transport) revealed how his collaborative research works. He said,

*"In some of the projects [...] we are using data from the other partners, so they own the data. [...] In most cases, [...] most of us collaborate with quantitative data, they would run the analysis and I would see the results and would discuss them. And I would see sort of the files with the commands. [...] But the data are not physically here. I don't have them on my computer. [...] actually the actual data are not being shared."*

The comment shows that there are two reasons which can explain why raw data are not shared within a research team. Firstly, the ownership of the data, is a factor. The above comment shows that the raw data of the project were not shared. Anthony (transport) did not own the data or have access to the raw data. This shows even research partners who collaborate together do not mean they have equal ownership of the data.

Secondly, the practice of division of labour in quantitative research, is a factor. While Anthony's collaborators who owned the data did the data analysis, he only examined the derivative data and results. This can suggest that common practices of quantitative collaborative research are that each team member shares their semi-processed analysis and results with each other rather than their raw data.

Human geographer Samuel (political) who also does quantitative research explained that raw data is not always needed in collaborations. Similarly, he explained that he and his collaborator had different roles. As he said,

*"most of the collaborations I am involved in, the way we tend to work is one of us will have the idea for the paper for the project. That person will do the bulk of the analysis"*



*and the bulk of the writing of the first draft of the paper. And the collaborator is primarily operated as a check if you like, and as a source of comments and ideas.”*

Samuel (political) further explained that most of time he does not need the data when he comments. He said,

*“most of the time no (there is no data sharing) because the results are quite easily replicated. So you feel you are reassured, no problem in the analysis itself. [...] We started to share things or we went to analysis in a bit when something unexpected happens [...] And at that point, your other partner would start to dig into the data and try to run the analysis himself and see if he can spot the problems.”*

The above comment suggests that researchers do not usually go back to the data unless it is necessary, e.g. when some problems occur. Samuel (political) also shows a lot of trust is involved in collaborators, as he said,

*“Most of the time, we know each other well enough to be fairly confident about each other’s ability to do the analysis competently.”*

In general, it seems that raw dataset is fundamental but is not the most important; researchers tend to focus more on research analysis and research output.

### 8.5.2 Internal data sharing in qualitative human geography

The analysis found that collaboration and internal data sharing practices of qualitative human geographers are quite different from quantitative human geographers. While researchers tend to have very clear and different roles in quantitative collaborations, qualitative researchers seem to have less clearly defined roles in collaboration. For example, Thomas (historical) described how he works with collaborators, he said,

*“..find some archival material, you can show it to your collaborators and then you write together about it.”*

The analysis found that qualitative interviewees usually “*work together*” or “*write something together*”. None of the qualitative interviewees mentioned division of labour in collaborations as quantitative interviewees.

The ways of collaboration in qualitative human geography seem to have an influence on internal data sharing. The analysis found that qualitative human geographers emphasised that each team member should have a complete set of data. For example, when interviewees were asked about

sharing data within research teams, they said that they made sure collaborators “*work on the same dataset*”, “*have a complete set of data*” or “*have everything*”. This is interesting because none of the physical geography interviewees have stated that all members in the research team should have a complete set of data. This shows there are differences in internal data sharing practices between human and physical geography.

The following comment of historical geographer Thomas showed that he found it difficult to collaborate when someone does not share a complete set of data. He said,

*“I am doing some writing at the moment [...] with someone who is a research associate, [...] she was not very keen on share her research. [...] So I am now in a situation whereby working with this person, I don’t see the whole of the data. And so I don’t feel as confident about doing the right thing. And because I don’t feel as confident about doing the writing, I put it off [...]it is frustrating”*

This suggests that a whole set of data is required in order to have a full understanding to complete the research for human geographers.

## 8.6 Methods and issues of internal data sharing in geography

The previous section showed some differences in collaboration and internal data sharing practices between physical and human geography. This section looks at how geographers share their data and some of their internal data sharing issues in general. Surprisingly, despite the intellectual differences in different fields, such as different methods and the different nature of data, both physical geography and human geography interviewees use similar ways to share data with their collaborators. The interview analysis shows that email and dropbox are the most common way for interviewees to share data internally. For example, when interviewees were asked how they share data with their collaborators, human geographer Eric (political, development) answered,

*“Email, dropbox, google drive, those are the main platforms.”*

Similarly, physical geographer Patrick (quaternary, biogeography) said,

*“traditionally [...] we just used to send to each other by email but now of course we start using dropbox and things like that... it centralised it.”*

This comment also shows technologies are subtly changing the way people sharing data.

Most interviewees found that using email and dropbox to share data with collaborators works well for them. Yet, a few interviewees mentioned they had concerns when they share data with their collaborators. The following comment illustrates common concerns of geographers, especially human geographers. As Eric (political and development) said,

*“Generally you are [...] concerned about files become corrupted or if you are working on projects, do you remember to look at the right folder for that document... then when collaborating on a report from something, are you both working on it at the same time and do you end up with two different documents but other than that generally it is fine.”*

The analysis shows that versioning, data corruption and loss are the most common concerns among geography interviewees. Yet, for human geographers, these are not their only concerns. There are also other concerns such as security issues. As Eric (political and development) pointed out,

*“There are always issues around other questions over ethics and security at times.”*

The analysis shows that qualitative human geography interviewees evidently had more concern with security issues than physical geographers. For example, Eric (political and development), Kimberly (environment, gender), Amanda (children) and Anthony (transport) all mentioned security concerns when they talked about sharing data internally. In contrast, none of physical geography interviewees mentioned security issues. As mentioned in chapter 7, one of the characteristics of data created by qualitative human geographers is the confidentiality of interview data. Thus, it is not surprising that qualitative human geographers have concerns of keeping confidential data secure. This shows that interviewees’ concerns about internal data sharing were shaped by the characteristics of the data they produced.

### *Dropbox*

Dropbox is another popular way to share data within a research team. Although all interviewees used dropbox, there are mixed perceptions as to whether dropbox is ideal for internal data sharing. For example, human geographer Kimberly (environment, gender) said dropbox works ‘*well enough*’ and believed it is secure for sharing data with collaborators. As she said,

*“We’ve done it by email and working on one (project) now with the dropbox facility which is a secure dropbox. [...] I trust the person who organised it. I haven’t set up my own, but I am told it is a secure facility.”*

On the other hand, Anthony (transport geography) was aware of the risk of using dropbox, as he said:

*“We use a dropbox ... I don’t think our IT department here would be very pleased with it if they knew. [...] I know actually know other universities where IT departments actually forbid people to use google drive or use dropbox. [...] And obviously when data are not anonymised, using that kind of infrastructure dropbox or goggle docs or google drive is very very risky. And I would rather do it in another way but the thing is I don’t really.”*

What is interesting is that, as revealed earlier, human geographers are evidently concerned with the security issue, yet the contrasting comments of Kimberly (environment, gender) and Anthony (transport) show that there are inconsistent understandings of what security level infrastructure provides, e.g. how secure is dropbox in reality.

Physical geographer Rachel (remote sensing) also did not think using dropbox for sharing data internally is an ideal way. As she said,

*“We mostly use dropbox actually. [...] I don’t have a pay for account though, so I get full all of the time which is annoying [...] I hate the program to be honest. I actually lost stuff. [...] Because I didn’t really understand how it worked. So I don’t like it but I don’t know any the other way of doing it.”*

Instead of the security issue, physical geographer Rachel (remote sensing)’s concern is related to the volume of storage and data loss. Nevertheless, similar to Anthony (transport), she felt there was no better way to share data with her collaborators. This seems to suggest that there are limited ways for geographers to share data internally.

Surprisingly, the analysis shows that none of the interviewees had much institutional support for internal data sharing. For example, when Matthew (quaternary) was asked whether the university supported data sharing with collaborators, he answered *“Not particularly.”* The first reason could be because currently universities are focusing on developing infrastructure and services for funding requirements, i.e. external data sharing, making data publicly accessible. The following comment suggests that the objectives of internal and external data sharing are different, e.g. Matthew (quaternary) said:

*“Within our group, we have tried to set up an internal database for data.[...] The difficulty is it’s not really publicly accessible. So it is not very suitable for fulfilling the requirement for the research councils to make your data publicly available. [...] that’s not really supported by the institutions.”*

The objective of internal data sharing is to facilitate research, enable things to be done more effectively within a research team; while external data sharing quite often is to meet funders' or journals' data policies, to make data publicly accessible.

The second reason, as Anthony (transport) commented, is that it is difficult for universities to provide support for sharing data within research teams due to their own secured network and protected systems. As Anthony (transport) said,

*“it is quite difficult because our university and many other universities have sort of protected systems. So is very difficult to set up something secure to exchange with researchers in other universities.”*

Anthony (transport) also mentioned *“researchers are more and more collaborating”*, in addition, as revealed in the bibliometric chapter (chapter 6), geographers quite often have multiple collaborators from different institutions, these all suggest there is an increasingly challenge for universities to provide support to researchers.

## 8.7 Summary

This chapter presented the interview findings of the social characteristics of geography. Geographers took part in a range of social research activities, such as attending conferences, being a member of research groups and national associations. One of the interesting findings is that physical geography interviewees tend to be able to go to conferences in wider areas, not only conferences purely on specialised subjects but also multi-disciplinary conferences, for instance, not only going conferences on diatoms but also conferences on geo-physical sciences. By contrast, human geography interviewees tend to only attend specialised conferences. This indicates there are differences in the range of audiences and potential users of data between physical and human geography. This could also influence researchers' inclination to share their data with others.

The chapter further examined the nature of research collaborations in geography. The analysis showed that collaborations between physical and human geography differs, not only in terms on the frequency they collaborate, the scale of collaboration but also the reason to collaborate. In physical geography, the culture of collaboration is mainly driven by its intellectual characteristics. Different specialized skills are required for the holistic nature of physical geography research. Members in the research collaboration have clearly defined, separated roles. On the other hand, in human geography, the culture of collaborations is more driven by external, e.g. research funders. Yet it was

found that that not all subfields are equally funded. Furthermore, the feature of clearly defined tasks in collaboration in physical geography was not found in human geographers' collaboration.

Internal data sharing practices in physical and human geography are shaped by their nature of collaboration. These interviews revealed that the nature of data which researchers shared with collaborators differed between physical geography and human geography. Physical geographers did not always share raw data within research teams but they shared research outputs or only part of their data with collaborators. Clearly separated roles in collaboration lead researchers to share only part of the data with collaborators rather than raw data, as raw data can be too specialized for others to understand. In contrast, human geographers, specifically qualitative human geographers, who tend to work together, emphasized on sharing raw and whole set of data, e.g. a complete set of interview transcripts. They made sure research team members had the same datasets as each other. This shows that what and when data is shared in collaboration in the course of research cycle are different between physical geography and human geography.

The interviews also showed some common ways researchers use to share their data with collaborators. The most common way to share data within collaborations is by email. Technologies have been subtly changing the way researchers sharing data. Increasingly, researchers not only use email but also Dropbox, even though some interviewees recognized that Dropbox is not the best platform to share data internally, for example, in terms of security. The findings suggested that there were limited ways for researchers to share data with collaborators; institutional support to internal data sharing was also limited.

# Chapter 9

## Data Sharing Practices in Geography

---

### 9.1 Introduction

This section explores interview participants' external data sharing practices, i.e. the practices of sharing data beyond their research collaborators. The term 'practice' refers to actual acts rather than interest or intentions to act. This notion of practice was used by Knorr Cetina (1999:10) and Pickering (1992). This section will focus on interviewees' acts of sharing data, such as how they share data, while their attitudes towards data sharing will be explored in chapter 10.

### 9.2 Diverse data sharing practices within one discipline

One of the important findings to emerge from the interviews is that the practices of sharing data are diverse within one single discipline. While some interviewees have never shared research data beyond their research team, some interviewees have extensive experiences of making data available to others. When the participants were asked whether they have ever shared their own data or made the data publicly available, some commented: *'quite a lot'* (Michael, physical geography, glaciology), *"Yes"* (Christopher, physical geography, climatology). In contrast, some interviewees commented: *"I haven't done it yet"* (Joshua, human geography, economic); *"No not yet"*; Adam (human geography, development). It is interesting to note that most of the interviewees who have experience of external data sharing were physical geographers. By contrast, there was not much data sharing taking place among human geography interviewees. Interviews show that there are different ways to share research data in geography, such as sharing data via journals and national data repositories. The analysis also shows that data sharing practices between human and physical geography are quite different, thus this chapter separately explores the practices of data sharing in physical geography and human geography.

### 9.3 Data sharing practices in physical geography

The interviews analysis found that there are four key ways that physical geographers share their data beyond research teams, including, sharing data through journals, data repositories, personal requests and personal websites.

### 9.3.1 Sharing data through journals

For physical geographers, sharing data through journals seems to be a well-known method. More than half of the physical geography participants shared some of their own data via journals. In fact, all physical geography interviewees said that they were aware that datasets can be put included in some journals. As physical geographer Patrick (quaternary, biogeography) explained *“That’s become increasingly a standard.”*

Data sharing via journals is increasingly common in physical geography, but the practices of making data available via journals, such as submitting data as supplementary information with a paper, putting data on journals’ websites, has only become common in the last few years.

*“I think maybe 5 years ago or older it would have been quite hard to find that data. But now most of the papers that I publish, the data would be in the paper or publicly available on a supplementary file, an open access paper. And somebody can look at the data.”* (Michael, glaciology)

The ways of disseminating research data in journals is changing. Putting data into journals’ archives or submitting data as supplementary information with publication seems a relatively new development; nevertheless, disseminating data through journals has existed for some time. Some interviewees provided insights into how data was disseminated in journals in the past. For example, Ryan (quaternary) said:

*“..that is only recently journals have started to offer the opportunity to link your data. In former times, you could quite often publish a table of information within the paper. So there’s been number of occasions I have to have this table of data. But this table can’t include very much information. So there still a lot of information gets lost.”*

The comment shows Ryan (quaternary) was “sharing’ or disseminating his data concisely in the form of a table within his journal article in the past. Talking about how the data has changed in journals, Michael (glaciology) also commented:

*“One of the interesting things in some of the fields that I’m working is historically you do some mapping that effectively that map is your data. Up until recently, journals sometimes are very happy .. {he was unfolding a big paper map approximately larger than A1 size from a journal- i.e. Journal of Glaciology }... In the inside back cover inside a journal, you may have a journal, for example, where you can put the maps. You can fold the map up and you can put it in the journal. But now because this is very expensive to do most of the*



*publishers stopped allowing to put the map on but they would say you have to put it online, like a pdf..”*

Map data was shared as physical objects in the past and is now shared digitally in pdf format as digital objects. Technology has enabled researchers to make their data available to others in different ways. The amount of data that can be shared has increased. Furthermore, data in digital formats has become more useful, especially in some research fields. For example, Michael (glaciology) commented on the advantages of map data becoming digital:

*“A map like this is a research output but if somebody wanted to use it, it is quite difficult [...] This is an outline of an ice sheet over Canada. And as a hard copy it is not as useful. But of course now you can actually get this electronically which means you get the actual outlines of all the features. And I can place it on my screen and I can compare them to my data or other people’s data. [...] When these were done back in 1997 this is all you had. And you know you cannot manipulate it at all or interrogate it. But now somebody who did this will make sure it is digitally available for other people to interrogate the map and extract properties from it.”*

### *Journals’ Requirements*

The change in data sharing practices of physical geographers is not only driven by the advancement of technology but also the expectations of journals. Talking about this issue Ryan (quaternary) said:

*“There’s been more of an expectation that you should make available your raw data at the point when you publish the data. For example, if you look at the journals Nature and Science, they will expect to see the raw data.[...] I don’t think it is just in my field, is of the whole of science now.”*

The analysis of interviews confirms that journals are making data sharing as a requirement of publishing paper, especially in glaciology and quaternary science in physical geography. Data sharing is becoming compulsory in some journals, particularly in high impact journals. As Michael (glaciology) and Matthew (quaternary) commented:

*“Most of the journals I publish in they would require you to deposit your data either on the journal website or insert a sentence in the acknowledgment saying ‘ Data can be provided’ [...] Recently, we have a paper [...] in the journal Nature and they required us to put all our data into a supplementary file which people can access online. So they can really look at the data.” - Michael (glaciology)*

*“Yes increasingly a journal will either ask you to put the raw data into an electronically annex which is held by the journal or they ask you to put the data into public archive...” - Matthew (quaternary)*

Similarly, physical geographer Patrick also from quaternary science commented:

*“...[pointing to a journal on his table] high impact journals demand that you put the data into the public domain.”*

Evidently, journals' data sharing requirements are common in the field of glaciology and quaternary science in physical geography. However, the interviews show that these requirements are not common for every journal and every field of physical geography. There are some differences between sub-fields. For example, Rachel who is in the field of remote sensing said that depositing data to journals is optional. As she said:

*“Not that you have to. But I know that is possible. [...] I think it may be Remote Sensing actually. There's a journal called Remote Sensing that I saw that you could do that.”*

Furthermore, the interviews reveal that it is not common to share modelling data through journals. Two physical geography interviewees who use modelling methods when asked whether it is common to share data, they answered: “no, with modelling” (Christopher, quaternary); “not really, there are some and it is a kind of mixture” (Kevin, fluvial geomorphology). Christopher further explained,

*“There are a number of journals that do that now but they require less for model data because it's too large. If I am collaborating with people who collect field data, often the field data will end up in a data repository or supplementary material online. But the model data tends to end up being images rather than the raw numbers in many cases. Because it's just too large amounts.....”*

This shows that research methods of researchers used could affect how and what data could be shared.

### 9.3.2 Sharing data through data repositories

The second way to make data available in physical geography that emerged from the interviews is putting data in data repositories or archives. But it is clear that this is less common than putting data onto journal websites. Less than half of the physical geographers have ever put their data in data repositories.

## Data centres

The places where the physical geographers have put their data include the American National Oceanic and Atmospheric Administration (NOAA), British Atmospheric data centre, British Oceanographic data centre and PANGEA. The following comments illustrate where the physical geographers have put their data. Christopher (Climatology, Quaternary) said:

*“There is a British Oceanographic Data centre and a British atmospheric data centre that I deliver data to. Some of my collaborators who are geologists put data in an international geological database which has got a German and American wings.” -*

This shows different disciplines use different data centres/ databases. Data centres are actually quite discipline specific. This was echoed by comments from quaternary scientists Matthew, Ryan and Patrick, as they said,

*“There are specific websites- NOAA in America and the USGA. They’ve collected a lot of data [...] certainly paleo climate data which you could put your data into that and so it could be accessed by anybody.”- Matthew (quaternary)*

*“The NOAA is one place and there is another place called PANGEA which is a German site where paleo climate data is. So they have many the same data as NOAA [...]”- Ryan (quaternary)*

*“... most of my other data is held either in electronic annexes with journals or at the national geophysical data centre in the US which is part of NOAA, which is the National Oceanic and Atmospheric Administration. So you have heard of NASA obviously, you know the space agency. NOAA is the organisation in the America that looks at the ocean and the atmosphere. And they have a very very good publicly accessible database, which is free to deposit data. So I got one major dataset there and one small dataset. And I am gradually working on another dataset to try to get them ready to deposit there.” - Patrick (quaternary)*

NOAA is clearly a popular data deposit place especially for paleo-climate data, while the British Atmospheric data centre, British Oceanographic data centre are more for projects which are funded by the UK research councils, such as NERC. As Matthew (quaternary) commented:

*“There is one project I had finished a few years ago where we were we had to sign a data agreement. So all that data was made publicly accessible. So I have no option I have to do that, it kinds of forcing me . [...] So that particular project, the data for that is held in the UK British oceanographic data centre, database. Even though it’s not oceanographic data, it*

*is terrestrial data. The way in which the data agreement was set out said that that is where we have to put it .”*

Furthermore, Ryan (quaternary) also mentioned depositing data from a NERC funded project to British Oceanography Data Centre. As he said,

*“ .. the work that we’ve done is funded by the NERC- the Natural Environmental Research Council. And they are now able to offer an archive of data for you when the project finished. So we have the project finished in 2008 and there are people, I think it is from the BODC which is the British Oceanography Data Centre who were designated by the NERC, [..] who would archive our data.”*

Similar to journals, data sharing has become compulsory for government and research councils funded research. As Michael (glaciology) said, *“actually under the new open access guidelines from the research council data has to be publicly available.”*

#### *Challenges of depositing data*

The analysis of interviews shows that the actual submission processes are relatively simple, but there still are challenges for geographers. The key challenges of depositing data are related to the format of the data and the time required for getting the data into that format. When interviewees who have experience of depositing data into repositories were asked whether is it difficult to archive the data, all commented “Yes!” Christopher (Climatology, Quaternary) found it problematic because of the format which data repositories use is different from what he usually uses. As he said:

*“They require a particular format and for data and that’s not the format that I tend to use .. so there are lots of issues in getting the data in the right form to go into the repository.”*

Ryan (quaternary) also commented during the process of archiving a lot of the difficulties are related to the format of data:

*“So we have a lot of emailing exchange where they asking for the data and then they kept on asking many many questions about the format of the data. And it shows you how difficult to these to archive the data properly.”*

Data formats normally used by researchers are not necessarily acceptable to data repositories, as these formats are not always best for long-term preservation. It is clear, however, that for researchers, the challenge of depositing data into a repository is not only putting the data into a

form that is acceptable by the repository but also the time required for the whole process is huge. As Matthew (quaternary) said:

*“it takes ages you know, you’ve got a dataset spread all over different places in your computer to get them all together, to get them into common format, to make them so other people when they download them would be able to understand what they are all about would take quite a lot of time. [...]I couldn’t even quantify it. [...] Getting the data into a form that is acceptable for the repository and that’s just takes time.”*

The time required and the format of the data appear as the challenges of depositing data, yet there is a deeper, more fundamental challenge, i.e. to make the data understandable to others. Depositing data into data centres is not only for archiving and meeting funders’ requirements but also for others to use. As in Matthew’s comment above, to make the data so other people would be able to understand is not easy and it takes time. It is not surprising because the nature of data is inherently unorganised (Rowley, 2007). In order for people who did not generate or collect the data to understand and be able to reuse the data, further information, such as metadata, is needed to accompany the deposited data. This becomes another challenge of depositing data for researchers. Ryan (quaternary) described the difficulty when he was depositing data into the British Oceanographic Data Centre:

*“The difficulty is the metadata. Because when the BODC was saying to us we need a bit more information-they were talking about metadata. So they need to make sure that other workers would know how to reproduce your data and they know, for example, if you were to put a time series that’s a set of data from youngest to oldest through a stalagmite sample. The metadata should include where is the cave, where is the sample collected within the cave, what is the temperature of the cave and that kind of thing. [...]to do it thoroughly can be very time consuming. [...] And I think this is one reason why people haven’t archived the data in a way that would now be expected. So I have a huge back log of stuff that hasn’t been archived to that standard.”*

### 9.3.3 Sharing data by request

Besides sharing data through journals and data centres, the analysis of the interviews showed that there is a less formal type of data sharing take place in physical geography, i.e. informal data sharing. Informal data sharing refers to data sharing by requests directly from primary researchers or through disciplinary networks, as opposed to formal data sharing which is sharing data officially

through archives or publications (Goodwin, 2012:11). What is surprising is that the analysis shows informal data sharing by request is popular among physical geography interview participants.

Sharing data by request is in fact more common than sharing data via journals and sharing data through data centres. The analysis of interviews shows all but one of the physical geography interviewees stated that they had received data requests. For example, as William (Hydro-climatology) said:

*“Yeah, typically. So just in the last few day has been dealing with a couple of requests, no it’s three requests: one from Ethiopia, one from a researcher in Australia and other one from a researcher from India or China.”*

The following comments further show informal data sharing by request can take place in physical geography, as interviewees said:

*“Yes I got one (a data request) last week [...] I have shared data even old data. People are writing to me. I published it. And I’ve given it to them. I can think of couple of examples.”*  
– Patrick (quaternary, biogeography)

*“.. a few weeks ago, someone asked if they could have the raw data related to a paper we published in 2005. And at that time, the journal didn’t offer the opportunity to archive the data. And so I was able to send him data. And he also asked if he can use some of the data that haven’t been published. Because that study was completed, then I said ‘Please go ahead and use that information’.”* - Ryan (quaternary)

*“For some of my data that had not been publicly accessible, sometimes have people email me and said ‘Are your data available?’ and I said ‘No, they are not but here you can have them.’ They are published, there is no problem.”* - Matthew (quaternary)

These comments suggest receiving informal data requests is common in physical geography and physical geographers are willing to share their data when they receive requests. Yet all the above comments indicate that the interviewees were willing to sharing their data because the study was “completed” or “published”. It is possible, therefore, that research has to be completed or published as a condition for physical geographers to share data informally. This is in agreement with Wallis et al.’s (2013) findings which showed the most common condition of data sharing placed by researchers is reserving the right to be the first to publish.

### Source of Data requests

The analysis of interviews provides some insights of who the physical geography interviewees shared their data with informally. Even though physical geography interviewees' data requests mostly were from individual researchers, some interviewees also shared their data with other parties. For example, Mark (sustainability) shared his data with local authority and local organisations, as he commented:

*"..the local authority and other organisations locally have come to us and ask for what's the potential for wood felling in this area. We can immediate tell them "well best of our knowledge from our sample is something like so many wood burning stoves in the village etc.' So we selected part of the data with a range of uses."*

Also he shared his data with a research student who was doing something related. As he said:

*"We have shared some of that information with a research student who is doing a related piece of work [...] so we share some of that information with him to inform his piece of research. "*

Michael (glaciology) received data requests from science organisations, such as National Snow and Ice Data Centre<sup>17</sup>. As he said:

*".... I have from scientific organisations. Yes. So I 've been involved in a couple of collaborative projects. For example, one of the things I told you about was we often draw around glaciers and I have done that in some small parts of the world and lots of other people have different parts of the world. And there was a project to bring all of this together into one big data inventory . And so some people emailed and said ' could you send... would you be happy for some of your data to be part of this global dataset?' And I have no problem with that , as I have already published the data. Is was just sitting on my computer and I've just sent it (C.S)"*

It is interesting to note that the intellectual nature of Mark's and Chris's research are different. Mark (sustainability) described his nature of research as "*majority applied*", whereas Chris described his nature of research as "*pure*" and "*theoretical*". The intellectual aspect of pure/applied research may have some influence on who will be interested in researchers' data. For example, the above comments show that Michael's data from pure research gained interests from academic, non-

---

<sup>17</sup> National Snow and Ice Data Centre (<http://nsidc.org/>), part of the University of Colorado Cooperative Institute for Research in Environmental Sciences.

industrial scientific organisations. On the other hand, Mark's (sustainability) data from applied research was shared with and used by organisations outside academia, such as local communities and local authorities. This could suggest that the audiences of data from pure research tend to be from academia, whereas data audiences from applied research can be more from industrial and non-academic organisations.

#### *Why sharing data by request is more common*

Data sharing by request is relatively more common than formal data sharing among physical geography interviewees. A probable explanation to emerge from the interviews is that data sharing by request is relatively easy. As presented in the 9.3.2 of this chapter, there are challenges of depositing data, such as, getting data into an acceptable standard and the vast amount of time that takes. By contrast, sharing data by request appears relatively easier and flexible. Christopher (climatology, quaternary) who has both experiences of formal and informal data sharing by request explained:

*"[...]I had an email request for specific data [...] Normally just some small subset of that so is relatively easy to provide."*

Physical researchers can have requests for a particular part of the data. They can divide up datasets and share only part of the data informally. This is echoed by Mark (sustainability) who said he "selected part of the data" for informal sharing, to "allow people to have aspects of it".

The analysis further shows sharing data by request is more common not only because it is relatively easier but also it benefits from the advantage of direct communication. Physical geographer Christopher (climatology, quaternary) explained the advantages of sharing data by request:

*"I prefer direct contact. [...] You can have an exchange with the person who is using your data then and it is interest trying to do but also you can advise them more about they doing something sensible with the data. So you can kind of interact with the way they use it better to help them get a sensible answer."*

One of the advantages of sharing data through request is that, as Christopher commented, it allows direct interactions with the person who asks for the data. So researchers who share the data are able to know how their data is used by others. Furthermore, research data is not self-explanatory, thus direct communications would allow researchers to offer advice to the user, make sure the data is understandable and used appropriately.



Communication channels should not be overlooked in data sharing. The popularity of sharing data by request among physical geography interviewees suggests invisible communications exist, researchers communicate within and outside academia privately. The interviews show the majority of physical geography interviewees were contacted and received data requests through emails. Interestingly, there is an exception. When Christopher (climatology, Quaternary) was asked whether he received any request for his data, he said: *“Yes! Through the ResearchGate website”*. ResearchGate is a social networking site for researchers. The comment seems to be consistent with the findings from the web content analysis which showed some geographers are adopting new technologies, e.g. using social media and websites to communicate and share research ideas.

### 9.3.4 Sharing data through websites

Sharing data through websites is another informal way of sharing data which emerges from the interview analysis. Nearly half of physical geography interviewees shared their data through websites. These researchers have put their data on websites and make the data freely available to the public, beyond academia. As William (climatology) said,

*“We’ve created a web portal for anyone to go and access these datasets. So we are sharing them, not only with the rest of the scientific community but with stakeholders, fishermen, schools who may also want to use the data.”*

William has extensive experiences on sharing data through websites. He shared a lot of data through various websites. For example, he has put river water temperature and air temperature data on a website which has been running for four years. The following comment illustrates how the data is collected, which he described as there is *“a strong field element to it”*:

*“We installed about 70 devices to measure air and water temperature, every 15 mins [...] trying understand the physical controls on river water temperature which really matters for fish and bugs and plants. We’ve been running this network for nearly four years now and that involve collecting, going to the sites, collecting the data, downloading it, checking the equipment and so forth.”*

William also shared data on another website. As he described:

*“This is another website where we created an index of flood risk of the UK for all these sites. And this is updated every single or more or less every day. And produced in file like this in zero time... So people can access... So there is a script processing the data showing how flood risks have changed. [...] We also put the data here. So someone can just collect and*

*download the raw data... (he opened a csv file from the website) so you see that's right up to the 29 Oct... so it's just few days old."*

The above comments show that the data William put on websites were time series data where same measurements are obtained on a regular basis. Besides sharing time series data, he also shared software he developed. He made his statistical model, a climate scenarios tool, available to the public to download. As he said,

*"We produced not only the tool but also the datasets to run the tool for anywhere on the planet."*

Interestingly, from the interview analysis, most of the physical geography interviewees who shared data informally through websites either develop models or databases as part of their research. William is not the only interviewee who developed a model and shared it on internet; similarly, physical geographer Kevin (fluvial geomorphology) shared software he developed on internet. The software he developed is a physically based numerical model to simulate geomorphic development in river catchments. Kevin described:

*"the model that I work with, I developed, that's freely available. That's been freely available since 2001 or 2002, maybe 2002. It's been on a website. It is presently on google code pages. So it is released with a licence, a general license so anyone can use it or modify it."*

Furthermore, Kyle (GIS) also shared his experimental database on the internet, a database which put different datasets together, to allow people to query connected datasets. When Kyle (GIS) was asked whether his database is freely available on a website, he answered:

*"Yes! Not the most elegant thing in the world. But it's there and it works. There is a demonstration for what could be done."*

### *Pros and cons*

There are advantages and disadvantages of sharing data informally through websites. One of the main advantages of sharing data informally through websites is the flexibility, for example, as William (hydro-climatology) explained:

*"I personally prefer my own site because I can customise it to make it as appealing as possible. So if we go back to the data... here is just a simple data table [..]But here there's a gallery showing pictures of the sites and background information on the publications and all*

*sorts... you know it gives us far greatest flexibility to tell the story in our own words without you know having a standard template.”*

It is clear, however, that there are disadvantages. William also realised the websites he and his collaborator create may have less longevity than a formal repository. As he said:

*“But I do see that there are no guarantee, if I were to leave (the university) then this would be maintained. So whereas in the institutional repositories has probably greater longevity. So I can see the pros and cons.”*

In fact, William mentioned one of his datasets is not openly available anymore due to a third party stopping support for the portal. As he explained,

*“that’s actually a problem because we set this initial dataset portal up with an organisation called Environment Canada about 10 years ago. And they subsequently have withdrawn support for this portal. So now all of these users are asking us, how do we get the data that we run [...] I’ve actually got some research assistant coming to see me on Thursday to talk about how we build our own portal that we can control to remove the risk of our third party supplying the data for us.”*

On the other hand, at the time of writing, it was found that Kevin (fluvial geomorphology) who indicated his software were on Google Code, needs to be moved to other open source websites, such as, Source Forge, GitHub. This is due to Google Code being phased out. The issue of sustainability does exist in sharing data through websites.

## 9.4 Data sharing practices in human geography

The practices of sharing data in human geography are significantly different from physical geography. The interviews revealed that there was not much sharing data in human geography. More than half of the human geography interview participants did not have any experience of sharing data.

### 9.4.1 Data sharing through journals

Data sharing through journals is not common among the human geography interviewees. Only one shared data through journals. This is probably because, as human geographer Adam (development) explained, *“I’ve never asked to do it”*. The practice is evidently uncommon not only among the human geography interviewees but also in their fields. When the human geography participants

were asked whether they noticed any journals in their field requiring them to link their data with their publications, the majority commented “No”, “I haven’t come across that”.

For example, Kimberly (environment, gender) whose research area is in gender and environmental and planning commented:

*“No, nowhere I publish has asked that. I am not aware of any geography, sociology journal that does that or planning.”*

Eric (development) pointed out that the practice probably is related to the intellectual characteristics of research. He suggested that the practice of sharing data through journals is more relevant to quantitative research. As he answered,

*“Not really. Because I do a lot of qualitative research, so is norm in this ... I can imagine people working more with quantitative data they will find it more.”*

Yet, Brian who does quantitative research in urban planning and Joshua who does quantitative research in economic geography, both commented that sharing data through journals is not common in their fields either. They suggested that the practice is probably more common in science and medicine disciplines. As they said:

*“I can’t remember which I’ve seen but It wasn’t like the standard economic geography journals. More sciences’ I would say.”- Joshua (economic)*

*“No. It’s been talked about it a lot [...] There is nothing at all in our field that mandates us to share the data. I am not aware of any journal in our field. But I am aware that in some areas, is possibly medicine this is beginning.” – Brian (urban)*

The only example of sharing data through journals to emerge from the interviews is from Samuel (political) who is a quantitative political geographer. He mentioned the practice of sharing data in journals is becoming more common, especially in American journals, as he said:

*“Yes. It’s probably more common, such as the major American journals increasingly demand, certainly would normally expect some online appendices (supplementary material) to actual data.”*

He suggested the practice is more common in the discipline of political science than geography, as he said:

*“That’s much more the case in political science and less in geography. I have yet to encounter a geography journal that insists on that. Well I only encounter one political science journal who absolutely insist. Most political science journals would encourage you to provide at least your computer code on an online appendix. It is increasingly a little bit odd if you don’t. But I have yet to encounter a geography journal that asks me to do that.”*

The research methods which Samuel (political) usually uses are *“taking quantitative datasets and apply statistical methods”*. He explained that secondary datasets he used usually are already publicly available, thus he does not need to put them on appendices. As he said:

*“Very often if you are doing something like a British election study or the American national election study, any academic, anyone in the world can access that data.”*

Yet the changing practice in his field is increasingly for journals to require computer code or datasets which are not available, as he explained:

*“But what journals are increasingly asking people to supply for online appendices is your computer code, so how do you actually program your analysis, how do you construct variables. So the code becomes checkable. Because the data is available to analyse, but the code isn’t. That is the code you often work for replication purposes. That occasionally matters [...]People are working on completely new datasets they gathered which aren’t publicly available at that point, journals are often quite insistent on how to make sure people can access the data.”*

#### 9.4.2 Data sharing through data centres

Compared to sharing data through journals, there are more human geography interviewees that have shared data thorough data centres. For human geography interviewees, data sharing through data centres was the most common form of sharing. Yet, still less than half of the interviewees shared their data through data centres. The majority of the interviewees who shared data through data centre were funded by ESRC, for example, as Amanda (children) said,

*“I got my data from the ESRC project [...]and is all gone to the ESRC data archive.*

Samuel (political) also was funded by the ESRC. He explained depositing data to the UK data archive in Essex was a condition of the funding:

*“Most of the major datasets are archived in national repositories like the Essex data archive. And certainly in the past, I was involved in some projects a few years ago [...] all those we gathered new datasets. They’re all funded by Economic and Social Research Council*

*and all those as a condition of funding the data had to be deposited into the Essex archive. Actually we were quite happy to do that. That data is now sitting in a national repository for anyone to access.”*

Furthermore, Sarah’s (development, social) research is also partly funded by the ESRC:

*“ So we have a project with ESRC/ DIFC , where everything has to get has to be deposited centrally. So all the quantitative data which was entered into SPSS and all the qualitative data which was transcribed. All's been sent to the PI of the project [...] and all that have been submitted to ESRC.”*

It seems that data sharing in human geography is mainly driven by funders’ policies. What is surprising is that Sarah (development, social) indicated that she did not know where the deposited data would be. As she commented:

*“I don’t actually know what happens when you give it to them. I don’t know who have access to it. [...] I just sent off everything I done to the PI.”*

This indicates that sharing data through archives is still novel to human geographers. Even through the practice is driven by policies, geographers do not have clear idea of why are they doing it.

### 9.4.3 Informal sharing

In contrast to physical geography, the interview analysis found that there is not much informal data sharing in human geography. For example, Samuel (political) and Paul (urban) are the only human geographers who shared their data informally through request.

Political geographer Samuel (political) who shared data “even before journals were asking to”, explained the reason he was able to share the data informally is because there is no confidentiality issue with his data. As he said,

*“Almost all the data that I have and almost all the data I use doesn’t raise issue of huge confidentiality. So because of that, people ask for copy of it. I give them a copy of it and I don’t see why shouldn’t.”*

Paul (urban) who does applied research also shares data:

*“Yes. People have. We have done reports that one of the political parties wanted to use. So yes , we have done some work recently about the state of the development plans , the production and development plans in the west midland and we receive a number of requests for access to that data. ...Ussally with the caveat , saying this particular research*

*wasn't finished so we did sent them a bit of health warning to say that this is where we've got to with this [...] and the work was not finished then there may be some inaccuracies in there. So yeah we shared it."*

## 9.5 Summary

The chapter presented the findings of data sharing practices in geography. One of the key findings of this chapter is that there are noticeable differences in data sharing practices within geography, especially between its two branches, physical geography and human geography. For external data sharing, i.e. sharing data beyond research collaborations, the analysis found that almost all of the physical geographers had extensive experiences, whereas there were only very few human geographers who had experience of sharing data beyond research teams. Physical geographers shared their data by various means. Interestingly, physical geographers share their data more on an informal basis, e.g. sharing data through websites, than sharing data formally, e.g. through journal repositories or national repositories. This showed data sharing practices in physical geography were not only driven by journal policies or research funders policies but also by the culture in their own communities.

# Chapter 10

## Data Sharing Attitudes in Geography

---

### 10.1 Introduction

This chapter focuses on geographers' attitudes towards external data sharing, i.e. sharing research data beyond research teams. It not only looks at the reasons why geographers support the idea of data sharing but also their concerns about sharing data with others. This chapter provides insight to further understand the diverse data sharing practices of geographers presented in Chapter 9. It again shows that there are significant differences between physical and human geographers.

### 10.2 Data sharing attitudes

The interview analysis shows that most interviewees' attitudes to data sharing are positive (20/23). For example, when interviewees were asked what they think about the idea of sharing data beyond their research team, they answered,

*"In general, I am supportive."* – Christopher (physical geography: climatology, quaternary)

*"Yeah I'd be happy to do that. I think it's good practice."* - Rachel (physical geography: remote sensing)

*"I am absolutely a supporter of this."* – Joshua (human geography: economic)

*"In principle, I think the more we share the better."* - Kimberly (human geography: environment, gender)

Interestingly, the findings show that not only physical geographers but also most human geographers had positive attitudes towards sharing data. This is somewhat surprising, as the previous chapter revealed that data sharing practices in physical geography and human geography are quite different. While most physical geography interviewees had experiences of making data available, there was not much data sharing beyond research teams in human geography.



### 10.3 Reasons behind supporting data sharing

The interviews revealed the reasons why geographers support the idea of data sharing (Table 10.1).

	Physical Geography	Human Geography
Data sharing attitudes	<ul style="list-style-type: none"> <li>• Positive</li> </ul>	<ul style="list-style-type: none"> <li>• Positive</li> </ul>
Reasons behind supporting data sharing	<ul style="list-style-type: none"> <li>• Contribute to knowledge</li> <li>• Increasing transparency, replicability and accountability of research</li> </ul>	<ul style="list-style-type: none"> <li>• Contribute to knowledge</li> <li>• To use resources efficiently</li> <li>• To avoid repetition of effort and replication</li> </ul>

Table 10.1 Reasons behind supporting data sharing

In general, both physical and human geographers are happy to share data if it can contribute to knowledge or science. For example, physical geographer Matthew (quaternary) said,

*“If data I have got helps to solve a particular scientific problem, then I am more than happy to share it. That’s really important.”*

Similarly, human geographer Adam (development) also supported the idea of data sharing said:

*“I think my general reaction is if we can add to the sum of knowledge by being open with data then that is a good thing.”*

Yet, further analysis shows that human and physical geographers have slightly different reasons for supporting the idea of data sharing. For physical geography interviewees, increasing validity of research is evidently one of the key drives to support data sharing. Most physical geography interviewees, including Mark (sustainability), Michael (glaciology), Rachel (remote sensing), William (Hydro climatology), Kevin (fluvial geomorphology) and Matthew (quaternary), expressed that they support data sharing because others can validate their results. For instance, in interviews, Rachel (remote sensing) and Mark (sustainability) explained,

*“I prefer to be as open as possible [...] so that anyone could check what I’m doing and validate what I have done as well.” - Mark*

*“That makes sense because then people can check what you’ve done to make it credible what you do.” - Rachel*

An unanticipated finding was that scepticism and scandals surrounding research has also driven physical and human geographers to support the idea of data sharing. For example, physical geographers Michael (glaciology) said,

*“I think what motivates me to share my data is...there’s quite a lot of scepticism about climate change and possibly because of people don’t have the access the data. So I am very happy to share data to non-experts to show them what’s happening to glaciers and convince them. [...] There’s some well publicised problems with maybe climate scientists not sharing their data and not letting people see how they have done the analysis. A lot of that is un-true but I think it is very clear that scientists should be accountable and should be transparent in terms of their data and sharing it.”*

The analysis indicated that the scandal around climate sciences in 2009, Climategate<sup>18</sup> had had an influence on the practices of data sharing and attitudes towards data sharing, especially in the field of climate sciences. Both physical geographers Ryan (quaternary) and Mark (sustainability) pointed out that Climategate was a significant episode in openness in climate science. Ryan described the effect of the climate gate scandal in the interview, he said:

*“This was in at the time of the Copenhagen Climate Summit. I think it was about 2009 or 2010. Some hackers hacked into the University of East Anglia websites and took many thousands of emails from climate researchers in that university. And then they took very selective extracts from them to try to discredit climate scientists. There’s been various inquiries since then. The scientists haven’t been found to do anything wrong. Nobody knows who did the hacking. But that did draw attention to the fact that those climate researchers were keeping their raw data to themselves and not sharing them. And since that time, it’s been thought it [...] is better to share all your primary data. And so scientific journals are now increasingly wanting you to publish your data with the paper or to log it to separate database.”*

Research scandals were not only in climate sciences, on the physical geography side, but also in political sciences, the human geography side. For example, human geographer Samuel (political) said that there had been a

*“... political data analysis scandal out there where people have fudged the analysis, fudged the data and fudged the results, to trying to get the answer they want. We don’t need that. Things should be replicable. And a price of that replicability is that is the software should be available, the programme should be available, data should be available. Unless there is an incredibly good reason not to, and that reason has to be with individual*

---

<sup>18</sup> Climategate <https://www.theguardian.com/environment/2010/jul/07/climate-emails-question-answer>

*confidentiality. There has to be some sensitivity about data to make it something they wouldn't want to share.[...] I think it is essential data is shared because results should be transparent, should be replicable, it should have accountability."*

What is interesting is that Samuel was the only human geographer that supported data sharing because of science agendas such as transparency, replicability and accountability. Charles' research is quantitative in nature. As explored earlier, most physical geography interviewees support data sharing because of validity, an important science agenda. It is possible that quantitative geographers value more the importance of transparency, replicability and accountability than qualitative geographers.

The interview findings show that most human geography interviewees have different reasons to support data sharing from physical geographers. For human geographers, the reasons to support data sharing are not to increase transparency, validity or replicability but to use resources efficiently and to avoid repetition of effort replication. As human geographer Amanda (children) said,

*"I think is a good idea in terms of the most efficient use of resources."*

Similarly, human geographer Adam (development) also believed that one of the key advantages of sharing data is to avoid replication, as he said,

*"I think the principle is a good one because it enables people to get access to that raw data. It avoids replication [...]for example, there are certain communities in the area that I am working with [...] which have been totally over researched. So it is kind of being aware of that and being able to get access to information [...] is very useful. Although you will then always have to look at how that research has been conducted .. and you would have done it in a different way."*

Furthermore, for human geographer Thomas (historical), data sharing is a way to reaching out to wider audiences, as he said,

*"I think it (data sharing) is a good idea. Anything that democratises and broadens the reach of academic research is a good idea."*

He further explained the way the culture of disseminating research has changed. Today, it is important to make research, including research data, accessible to others. As he commented,

*" If you think about it 20 years ago, when an academic could sit in an university library [...] making themselves a world expert [...]. But then they never published anything*

*about it. I think it is quite scandalous really, is a waste of public money. And that person if then they go to their death, never have them written it up. [...] if they haven't published it and they haven't made their materials accessible then that's the problem. So I think, I agree absolutely we should be reaching out to different groups in society. [...] I think we do have a responsibility to talk to different audiences. I feel quite strongly about that actually. I am in a position of someone whose work is really not as relevant to many people at all, the kind of historical geography [...] But I still kind of try to put it out there."*

Historical geographer Thomas sees data sharing is a way to disseminate research. His enthusiasm for reaching out to wider audiences was somewhat surprising, especially, as he is in the field of historical geography which is rural in nature.

In general, physical and human geographers have different sets of reasons of supporting data sharing. For physical geographers, one of the key reasons to support data sharing is because it facilitates good scientific practices, such as increasing transparency, replicability and accountability of research. On the other hand, none human geographers mentioned this reason. The interviews showed human geographers support data sharing mostly because of avoiding repetition and efficient use of resources.

#### 10.4 Barriers and concerns about data sharing

This section explores barriers to data sharing in geography. The previous section showed that both human and physical geographers support the idea of data sharing beyond the research team. However, as shown in chapter 9, there were contrasting differences in data sharing practices between physical and human geographers. Only a very few human geographers had the experience of sharing their data. Interviewees expressed the idea that data sharing is not always straightforward. As human geographer Eric (political and development) said,

*"In many ways it's something that I think is good and can be encouraged but there are various barriers."*

The interviews reveal that there are six barriers and concerns to data sharing in geography (Table 10.2). Three of the barriers are specific to human geographers, namely ethics, data access and copyright.

	Physical Geography	Human Geography
Barriers to data sharing	<ul style="list-style-type: none"> <li>• Competition</li> <li>• Time</li> <li>• Usefulness of sharing data</li> </ul>	<ul style="list-style-type: none"> <li>• Competition</li> <li>• Time</li> <li>• Usefulness of sharing data</li> <li>• Ethics</li> <li>• Data access</li> <li>• Copyright</li> </ul>

Table 10.2 Barriers to data sharing

### 10.4.1 Ethical issues

Human geography interviewees have more concerns on data sharing than physical geography interviewees. This is due to the intellectual characteristics of human geography, such as, involving human participants and the qualitative nature of research, as showed in chapter 7. Research in human geography research commonly involves human participants, so it was not surprising that one of their concerns is breaches of research ethics. As human geographer Eric (political and development) said,

*“one is to do with the ethics of what somebody has consented to in providing your data in the interview and if they don’t say yes you can use it in the project and then you show someone else[.]. You are breaching the ethical agreement with that participant so that’s not so good at all.”*

Some human geographers were also concerned that changes on ethical and confidential agreements with participants could discourage people from participating in their research. As development geographer Adam said,

*“I guess the issue of confidentiality and of what you promise to your respondents in terms of the inability to track down who someone is from it and the uses [..] So you can guarantee how you are going to use the data but if others are then going to have access to that data. There are some types of research where people much actually said - I don’t want to do that... you know.”*

#### Anonymisation

Anonymisation is evidently a key ethical concern among qualitative human geographers. Most interviewees who use qualitative interviews in their research, including Eric (political and development), Adam (development), Amanda (children), Kimberly (environment, gender) and Anthony (transport) expressed this concern. The following comments shows Kimberly (environment,

gender) recognised that anonymizing data in a whole set of interview transcripts is much more difficult than anonymising data that just appears in a publication, such as selected quotes from interviews. As she said,

*“So the research I have done on [...]women is with a very small group. In writing about it we’ve been able to anonymise. Looking at the original data, I’m not sure how to anonymise.”*

Amanda (children) who has the experience of making data available, stated that it is very difficult to anonymise participants. As she explained in detail how difficult the process was:

*“Not just anonymised it because in terms of changing names and things that was really quite easy. But it was places and things that could identify people. So changing a name is fine and to replace a job is quite easy. [...] I mean we had hundreds and hundreds and hundreds of pages of data.. we have research diaries from each of our schools and we did (many) schools and each school had at least a hundred pages of research diaries. And then we had (250 +) [...] interviews. So I mean it was a lot of data. And it was literally.. it was months. And it was not just my work. (Charles) did some work on it as well. And it wasn’t the changing and it was the checking. [...] I think it was probably particularly problematic for our data. Because it was very specific, so you could actually identify individual people by their characteristics. And then they got these conditions and they live there. They would know who that is. So we have to change it[...]changing is one thing but then reading through them all to make sure you had changed everything was quite a different thing. And then when you did that, then you think ‘oh no...’ it said.. [...] ‘I am in class ..I am in the elephant class’. You know, there are not every school call that classes by animals. So you know somebody reads that and realised that you are in the elephant class [...] then they know that school did participate in the research. [...] so that was a very big job”*

Not only anonymising whole interview transcripts is challenging but also images. Amanda (children) explained she did not share her image data from her funded research because it is impossible to anonymise the data to a high standard. As she said:

*“There was no way that we could have possibly have anonymised everything to a standard that anyone accessing it might not be able to identify. [...] There is a lot of work to make sure that any identifying features are removed, you know, street names or even what houses are like and stuff..”*

## 10.4.2 Concerns over data access

The second barrier is related to data access. Besides ethical issues, human geographers also had concerns about who would have the access to the data. For example, when human geographer Sarah (development, social) was asked whether she will share her data in the future, she answered:

*“It depends on who they are. That’s a tricky one.”*

It is worth noting that most human geography interviewees did not have the experience of data sharing, thus most of them did not know how data is actually shared. For example, the following comments suggested Adam (development) is unsure about what would happen in terms of data access if the data is shared through a repository. As he said,

*“I don’t know. Because I think it would depend on who was asking for the access. So for example, I might have interview transcripts that were relatively commercially sensitive or it could be sensitive in some sense. So I suppose you will ask questions about why people wanted to access and what they would get, what use of it going to make them [...] I suppose what I was saying is I ‘d want to have some control over it”*

The barrier is not only because of the sensitive nature of human geography data, but partly also because of the lack of understanding of how data can be shared, such as, who will have access to the shared data. Besides concerns about who might access the data, human geographers were also concerned about how data might be used. For example, Sarah (development, social) and Eric (political and development) said,

*“You don’t know your data may get misused.” - Sarah*

*“There’s always concerns about how other people would use your data and whether they would use it appropriately or inappropriately.” - Eric*

The follow comments Kimberly (environment, gender) illustrated how data can be misused, she said:

*“..Somebody can use your data for purposes I wouldn’t be happy about[...] It is a bit like the person who develops atomic fusion and then somebody else then develops the bomb and they ended up committing suicide.”*

The interview analysis found that physical geographers had quite a different view from human geographers regards of data misuse. Physical geographers did not care about how their data might be used. For example, as Matthew (quaternary) said,

*“ I think in a way, if somebody misunderstands it that’s up to them. All you are doing is presenting them the data, obviously, if you published the data, they can see how you interpreted them. If they wish to interpret them in a different way, that’s up to them. um I don’t think that (data misuse) in itself is a massive worry. [...] Maybe sometimes they will come up with a better interpretation: that’s always possible. You know, that’s how science works. You can only do the best you can with the dataset base on what you know at any one time. Knowledge moves forward, so somebody can come back to a dataset that you produced maybe 10 years ago with better improved knowledge and said you interpreted in this way, but actually what we know now suggested that that’s possibly incorrect. You know, that’s fine. They make their argument correctly. That’s not a problem.”*

Where physical geographer William (hydro-climatology) thought that it is very rare for people to misuse the work,

*“I can only think of few instances really ... in my entire career where someone has done something naughty either with my work [...] is very very rare. So that’s not something I worry about.”*

Physical geographer had less worries of data misuse probably because they were used to sharing data with academics. It can be seen that both comments above were in the context of where data are used scientifically. Yet, as Mark (sustainability) whose research spans across both physical and human geography pointed out data now is open to everyone. He said,

*“In theory, it should be open to anybody. But that then potentially allows mis-use of the data and particularly anything to do with climate change of course which is politically a very very hot topic [...] So you have to be very careful with who gets access to data and what they do with it. Because they’re not necessarily going to use it in straight scientific terms, they can use it for other sorts of reasons.”*

It shows shared data is not necessarily limited for academic use, this could possible increase geographers’ concern over the misuse of data.

### 10.4.3 Copyright

Copyright and intellectual property appears as another barrier to human geography interviewees, such as, Eric (political and development) and Joshua (economic). As expected, the copyright concern is particularly related to researchers who use secondary data, i.e. data that already existed. For example, Joshua (economic) described the data he used, he said,



*“I use secondary data. You analyse data, so you have some products based on this analysis. But these are the products. [...] the analysis I do, that is not the actual data. The actual data is usually either open data or data that I am not able to share with other people.”*

What is surprising is that there were quite a few physical geography interviewees who used secondary data, yet none of them mentioned this concern. It is possible that, as revealed in chapter 7, physical geographers tend to share research output as data, instead of sharing raw data which would have copyright issues.

The interviews reveal that some human geographers believe there is no need to share copyright materials, as they are already available in public. For example, historical geographer Thomas said,

*“If it is in an archive, is there already, is already stored, or is public so is already stored somewhere. So you know, is in the library.”*

Similarly, Economic geographer Joshua said,

*“Most of the data I use [...] is open data. So everybody could download that data. So for that reason, there’s no need you know for this kind of data.”*

In general, ethical issues, anonymization, data access and use and copyright are the main concerns for human geographers to share data beyond research teams. The analysis shows that there are also other concerns, such as competition, time and usefulness of sharing data. Nevertheless, these are not concerns exclusive to human geographers but also to physical geographers.

#### 10.4.4 Competition

Competition is a common concern in both physical and human geography. As mentioned in chapter 9, one of the conditions for physical geographers to share data is that the research project has to be completed or published. The analysis found that human geographers shared very similar views. For example, Jessica (Climate change) and Mark (sustainability) who did qualitative research both believed that researchers should have the first right to publish before the data is shared. As they said,

*“I think I would like perhaps to have a time limit. If you’ve gone out, you have collected data. I think it is reasonable to expect that you should have a certain amount of time to use that data before everybody else gets dips on it . [...]. So I don’t see why I shouldn’t have some ownership of that data really, at least for some length of time before is then free for other people to use. [...] But then after that, you know why not”- Jessica*

*“ultimately because it’s been done at public expense then the data must be made available after a certain amount of time. [...] I think that has to be a period of time when the person doing the work and putting the effort has to have the opportunity to work on the data themselves. [...] Otherwise, you won the grant you collect the data, someone else grabs it, writes it up and you can’t publish it because they already published the results.” - Mark (sustainability)*

It can be seen that there is competition for publication of results. Researchers see research results as highly important. The above comments suggests that the recognition of publication is much higher than of data. This could have an implication for incentives to share data. Further analysis shows that researchers not only would like to have the first right to publish but also they have expectations on co-authorship when they share data. For example, as human geographer Sarah (development, social) said:

*- “If it was going to get published it would mean everybody who participated in that data collection would have to be acknowledged and not just an acknowledgment but should be co-author if something is published.” -Sarah*

#### 10.4.5 Time

Time is also a key concern to all geographers. While most interviewees supported the idea of data sharing beyond research teams, many of them expressed time as an issue for them. For example, in the following comment Eric (political and development) explained that he would not share his data unless he is required to:

*“it will be very very unlikely I’d do it unless you are required to. For a number of reasons, firstly, the time and effort you need to spend to strip out from the interview transcripts, potential identifying information about place, say ways people could be identified... you would be very very careful with that for ethical reasons. The time taken doing that is tremendous.”*

And Amanda (children) who had the experience of shared qualitative data said,

*“it’s very time consuming. It is really hard to get all your data totally anonymised and ready and uploaded into the archive. And you think you anonymised things but then you realise that there is a place in them or something like that which could identify somebody.”*

Not only human geographers but also physical geographers expressed that time is an important issue. For example, as physical geographer Rachel (remote sensing) said,

*“ I would be happy to do that but then again it would be quite time consuming to put it all, to make it all available because I collect so many images and lots of them don't ever get used. But if you make them available they should be available with good metadata so that people can know what they are looking at. And to do that is quite a time consuming business really.”*

As revealed in chapter 9, time is one of the challenges for physical geographers to share data. Yet, the above comment shows physical geographers can select part of their data to be shared, which is different from qualitative geographers who usually need to share all of their data, e.g. all interview transcripts. This could contribute to the fact that more data sharing take place in physical geography than human geography.

#### 10.4.6 Usefulness

Besides competition and time, the analysis found that there was scepticism about the usefulness of shared data, including, scepticism about the quality of data researchers put into the archive, whether others can understand their data, the relevance of the data to others etc. As regards the quality of shared data, physical geographer Patrick (quaternary, biogeography) stated it is actually not difficult to make shared data difficult for others to use. As he said,

*“If you really want to.. you can make it ( shared data) difficult for people to use it. Of course it is not difficult to do that. [...] because of course most papers now have to be get shorter than it used to be, so you have to be very specific about what the point is trying to make. So you only use(share) some of the data [...] which means it can be difficult for people to use it properly.”*

Kevin (fluvial geomorphology) who used funded project data said, it is not always easy to use someone's data. He explained,

*“Yes, but it's difficult. I think despite the fact you can get it for free. I think it still ultimately comes down to the personality of the person who collects it. Because they can hang on to it. [...] They could just hand the spreadsheet that full of numbers. So .. not very useful. I think sometimes you need to have an explanation or interpretation of what's that. To some research the data is their power, [...] so they see it as being their way of holding .. leverage over someone or something.”*

Furthermore, some human geographers were unsure whether the data will be relevant to or understood by others. For example, human geographers Brian (urban) and Joshua (economic) said,

*“If you think computer programs are data. Often the program is not in a form which is very easy for another user to actually understand them really in that sense.” - Brian*

*“My code is mostly implementation of statistical modelling which makes sense for my own data. It wouldn't be that useful for somebody else to take my code and use it for something else.”- Joshua (economic)*

Eric (political and development) also stated that the interview data he collected was specifically for his own research, he was not sure it would be relevant to people even whose research was in the same research area or topic. In the following comment Eric (political and development) explained why data sharing is not popular with qualitative data,

*“Say my interviews on citizenship, education, the questions we designed with that project in mind [...] in terms of sub-questions to do with say dignity, respect whatever else. For somebody else working on citizenship education that might not be relevant to all ...we may all be doing very different things.”*

Furthermore, the subjectivity of data and research method also shape researchers' view on shared data. The analysis found that qualitative interviewees, including Kimberly (environment, gender), Anthony (transport) and Jessica (Climate change), were sceptical about the usefulness of data because of the subjectivity of the data. They explained in the following comments qualitative data are subjective in nature. It involves interpretation, as mentioned in chapter 7. Thus it is not suitable for reuse. For example, Kimberly (environment, gender) said,

*“ I mean in theory the findings that you draw if they are replicable that is going to increase their trust worthiness. It verifies the data[...]If you are talking one to one with somebody, the interview is as much about you as it is about them and therefore the questions I asked there is a lot of 'me' in it. [...] So I think philosophically I don't think it is appropriate for somebody to use that data to do anything with. Because you are only looking at really half of the data. Because they are then bringing themselves into it. So it raises questions about what data is really and I think it's assuming data is something that kind of independent to the research.”*

Human geographer Anthony (transport) also explained the social constructed nature of data, as he said,

*“I am also a little bit sceptical . I've worked with interview transcripts that someone else did the interviews. And then analysed them myself. But it is so much more difficult to*

*analyse interviews that someone else has conducted because you only see the words. And all the rest is gone. [...] I mean it is a little bit different with quantitative data , for instance, diary data that we collected- yes we could make them available publicly for others to use [...] there would be lots of issues because there are so many decisions that you make along the way. I mean data are not given. Data are created. [...] I would want to work with data where I understand that construction process. So I have a very strong preference for working with my own data.”*

What is surprising is that not only qualitative data but also quantitative data is not necessarily seen as objective. Physical geographer Kevin (fluvial geomorphology) recognised that his quantitative modelling data are actually constructed in nature. It involves a lot of qualitative data judgements. As he explained,

*“A huge chunk of my analysis is qualitative, is not quantitative, is not necessarily plotting a graph and saying numbers of that sort of things. It is observing is watching is seeing what happens looking for behaviours[...] And a lot of my analysis ultimately comes down to experiences and hunches about whether the model is generating something that looks sensible that looks right that is correct, is what we might expect. And I think there is enormous similarity there between someone working in laboratory or to some extent someone working in the field, will be looking at landscape or situation or scenario and they will be continue making qualitative judgements in their head whether something is right or wrong or what they expect or they don't expect. They may then choose to ignore certain things. So it is a similar process because everything does generate numerically then you do have a data store there, you can go back and you can look at things they are more... you can look at things in a quantitative way. You can plot data and distribution use statistics and that sort of things if you want to. But I think to be honest, both of the analysis and interpretation comes from qualitative analysis.”*

Chapter 7 showed that a lot of data in physical geography is semi- processed or processed. This shows processed quantitative data is not necessary totally objective, as there were interpretive elements in the processing of the data and this can affect the usability of the data if it is shared.

## 10.5 Summary

This chapter presented the data sharing attitudes of geographers. One of the interesting findings is there were differences between attitudes and practices in data sharing, especially among human geographers. As mentioned earlier there was not much data sharing among human geography

interviewees, yet, surprisingly, human geographers' attitudes towards sharing data were very positive. Both physical geographers and human geographers recognized sharing data is a good idea, as it can contribute to more knowledge. Other reasons for geographers to support data sharing were also revealed. While Physical geographers tended to support data sharing because it facilitates good practices in science, human geographers supported data sharing because of avoiding repetition and reaching out to wider audiences. The different reasons for sharing data between researchers indicate the incentives and motivations for sharing data would vary across geography. The chapter further examined interviewees' concerns about data sharing, which could be seen as barriers to share data. The common concerns mentioned by interviewees were competition, the amount of time needed to prepare data for sharing and the usefulness of shared data. In addition, for human geographers, there were concerns on ethical issues and copyright issues, which are evidently related to both social constructs and intellectual characteristics of human geography.

# Chapter 11 Discussion

---

## 11.1. Introduction

The aim of this research is to look at the mutually shaping relationship between data practices and the disciplinary culture in geography. A number of empirically-based exploratory studies, including analyses of department websites, researchers' web profiles and publications, a bibliometric study and interviews, were conducted to understand the disciplinary characteristics and data sharing practices in geography. This chapter is organised around themes which bring together some significant findings from each of the studies and relate them to the literature on disciplinary cultures research data and data sharing. It starts by discussing the characteristics of geography. Then it explores the conceptions of data from a disciplinary perspective connecting this to the data sharing practices in geography.

## 11.2 Characteristics of geography

### *Intellectual and social differences between physical and human geography*

The long-standing divide between physical and human geography has been well documented (e.g. Lave, 2015; Cresswell, 2013; DeLyser & Sui, 2013; Jones & Macdonald, 2007; Harrison, 2004).

Although the increasing development in open science, big data and digital humanities can potentially bridge the divide (DeLyser & Sui, 2014), the findings of this research showed the divide still exists. It showed that there are intellectual differences between the two branches of geography. For example, analyses of interview data and researchers' web profiles both indicated that physical and human geographers undertake research about quite different research objects and apply different types of methods to their research. The objects researched in physical geography are physical in nature, e.g. lake sediments, glaciers and climate, whereas human geography is human in nature, e.g. young people, gender and the economy. Table 11.1 shows some of the key methods in geography which were observed in this research and are also mentioned by Clifford (2010).

Methods in physical geography	Methods in human geography
<ul style="list-style-type: none"><li>- Field observation and measurements</li><li>- Numerical modelling</li><li>- Remote sensing</li></ul>	<ul style="list-style-type: none"><li>- Questionnaire surveys</li><li>- Historical sources</li><li>- Observations</li><li>- Interviews</li></ul>

Table 11.1 Different methods used by physical and human geography

DeLyser & Sui (2013) have pointed out that methodological differences, particularly the quantitative/ qualitative divide is a major division between physical and human geography. Physical geography is characterized by using quantitative methods, whereas human geography is traditionally characterized by engaging qualitative approaches. Using Becher and Trowler's (2001) conceptual terms (see chapter 2.4.1), the knowledge structure of physical geography can be considered as relatively 'hard' whereas the knowledge structure of human geography is relatively 'soft'. Nevertheless, the quantitative element has been growing in human geography (Kwan, 2010; Murray, 2010). Quantitative methods are not mainstream approaches in human geography, yet big data analysis is growing in human geography (Miller & Goodchild, 2014; Kitchin, 2014). As the interviews showed, quantitative approaches, including spatial analysis using big data, appeared in some particular subfields of human geography, e.g. economic geography. Thus, the gap between the two branches may be narrowing, as the range of methods in human geography broadens.

The divide that exists between physical and human geography is not only epistemological but also social. Yet, there is less emphasis on social differences in general literature. This research showed collaborations in physical geography and human geography differ, such as, in terms of the frequency, the scale of collaborations and the reason to collaborate. The social differences between physical and human revealed from this research are highlighted in Table 11.2.

Physical geography	Human geography
<ul style="list-style-type: none"> <li>- Collaborate all the time</li> <li>- Larger scale of collaboration               <ul style="list-style-type: none"> <li>o co-authors in a publication are around 6</li> </ul> </li> <li>- Clearly defined tasks within collaborations</li> <li>- Researchers not only go to conferences in specialized areas but also in wider areas</li> <li>- the culture of collaborations is more driven by its intellectual features. i.e. diverse specialized skills are required for the holistic nature of physical geography research</li> </ul>	<ul style="list-style-type: none"> <li>- Collaborate, but not always</li> <li>- Smaller scale of collaboration               <ul style="list-style-type: none"> <li>o co-authors in a publication are around 2</li> </ul> </li> <li>- Less defined tasks within collaborations</li> <li>- Researchers tend to go only to specialized conferences</li> <li>- the culture of collaborations is more driven by external influences, e.g. research funders</li> </ul>

Table 11.2 Differences in social organisation of physical and human geography

As shown in the bibliometric analysis, physical geographers generally have more co-authors and more collaborative publications than human geographers. The results are in agreement with Gilbert (2014) who commented that multi-authorship is the most common form in physical geographers and thus they have higher numbers of publications. Furthermore, interviews showed geographers took



part in a range of social research activities, such as, attending and presenting at conferences, being members of international associations and departmental research groups. One of the interesting findings is that physical geographers seem to be in a more interdisciplinary environment. For example, findings of interviews suggested that physical geographers tend to be likely to go to conferences in wider areas, not only conferences purely on specialised subjects but also multi/ interdisciplinary conferences. By contrast, human geography interviewees tend to only attend specialised conferences. In fact, these findings match those observed in earlier literature. For example, Thrift (2002) commented that physical geographers are in an environment which is highly interdisciplinary and publish in journals outside of geography, while human geographers publish in a narrower range of geography journals. These social differences could be explained by the fact that, in Becher's (2001) terms, physical geography is relatively 'urban', i.e. there is a high people-to-problem ratio, whereas human geography is relatively 'rural' which has relatively low people-to-problem ratio. For example, in physical geography there are lots of people working on relatively few problems, whereas, in human geography there are only few people working on a broad range of problems. In general, the intellectual and social divide between physical geography and human geography still exists.

### *Data differences*

This section discusses data differences between the two branches of geography and the interrelationship between data and the culture of the discipline. This research showed that physical geography and human geography not only have differences in terms of intellectual and social disciplinary characteristics but also the nature of data. These findings support the idea that different disciplines create different types of data, for instance, as Akers & Doer, (2013) pointed out data can differ in size and content between disciplines. In fact, the findings also confirm data differences not only appear at disciplinary level, but also at the sub disciplinary level (Cragin, 2010; Borgman, 2015).

The analysis of interviews in this research showed that there is a strong relationship between research methods and the nature of data. It is evident that different research methods involve the creation (or use) of data of very different types. The following table summarises some of the data differences of each method revealed in the interviews.

	Method	Characteristics of its data
<b>Physical geography</b>	Physical measurements	<ul style="list-style-type: none"> <li>- Numbers, quite often time series data</li> <li>- Usually collected from the field by researchers themselves</li> </ul>
	Remote sensing	<ul style="list-style-type: none"> <li>- Remote sensing images (multi layers)</li> <li>- Collected through field work or using secondary sources</li> </ul>
	Numerical modelling	<ul style="list-style-type: none"> <li>- Require input data, mostly from secondary sources</li> <li>- Big in volume</li> </ul>
<b>Human geography</b>	Statistical modelling	<ul style="list-style-type: none"> <li>- Input data- mostly from secondary sources or tertiary data</li> <li>-</li> </ul>
	Quantitative Survey	<ul style="list-style-type: none"> <li>- Structured research instruments</li> <li>- Sensitive data can be collected, such as, street names, post codes</li> <li>-</li> </ul>
	Interviews	<ul style="list-style-type: none"> <li>- Recordings, transcripts</li> <li>- Involves human participants, sometimes involves elite people</li> <li>- Relatively subjective</li> </ul>
	Observation	<ul style="list-style-type: none"> <li>- No standard form of data</li> <li>- Can be a diary or notes, fragmented</li> <li>- With researchers' reflections- sensitive materials</li> </ul>
	Documentary research	<ul style="list-style-type: none"> <li>- Documents, copyright materials</li> </ul>

Table 11.3 Methods and data characteristics in physical and human geography

The nature of data and disciplinary culture are closely related. In fact, as with research methods and research objects of a discipline, research data is an essential component of the knowledge structure of a discipline. Jones & Gomez (2010: 1) pointed out that research method essentially “*is a way of collecting and analyzing data*”. Methods are an integral part of the knowledge structure of a discipline, hence the data derived from the methods. In the section that follows, it will be discussed further how disciplinary cultures and the notion of data shape each other.

### 11.3 A disciplinary conceptual framework for research data

This research, especially, the Interviews showed that the notion of data is very complex. This is in line with the findings of previous studies, such as Borgman (2012) and Lyon (2007). One of the interesting themes emerging from the interviews is about what is the nature of data. In this

research, different interviewees interpreted data differently, anything from mud samples, record measurements to research outputs were all seen as data by different interviewees. As Borgman (2012) pointed out data is not easy to define. Based on the empirical findings from this research, a conceptual framework for research data is developed (Figure 11.1).

### 11.3.1 Data differences

This research found that the notion of data can be understood along three dimensions: physical form, intellectual content and social construction. Disciplinary differences were found in these three dimensions. And in each dimension, different characteristics can be observed.

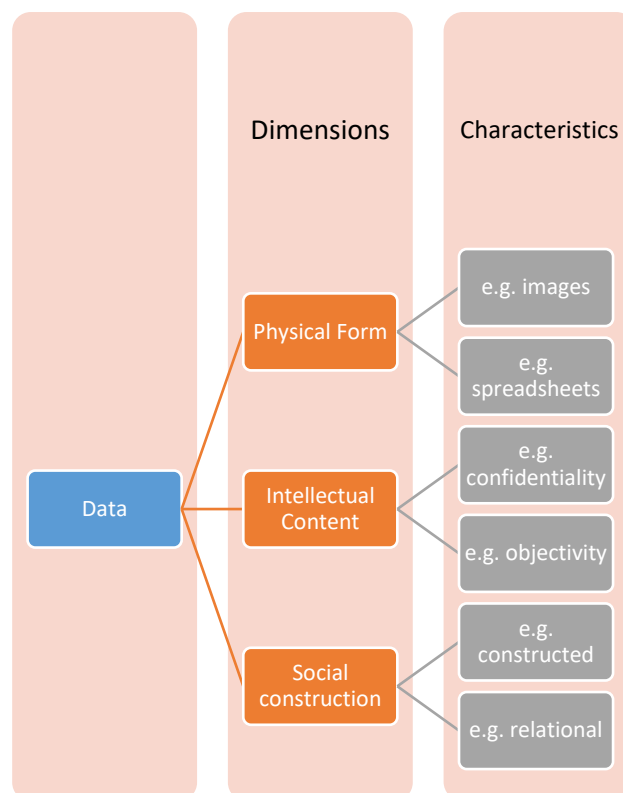


Figure 11.1 Conceptual Framework for Research Data

#### Physical form

The first dimension is the physical form of data. All data takes some characteristics of form. Orlikowski and Yates (1994:544) refers ‘form’ as ‘readily observable features’. The ‘form’ of data can be seen as the physical characteristics of data. E.g. Data can be in the form of digital or physical objects. It can be in form of numbers, text, images; in different format or standards, such as, a spreadsheet or a word document. Different disciplines create different forms of data. Fecher et al (2015) pointed out that “formatting standards” is one of the features influencing data sharing and reuse in academia, as lack of standards can hamper interoperability. This data feature can be seen as related to the physical form of data. Physical entity probably is one of the easiest and most

observable way to approach data in research. There were definitions of research data in the literature describing the form of data. For example, Pryor (2012) refers data as

*“The primary building block of information, comprising the lowest level of abstraction ..., where it is identifiable as collections of numbers, characters, images or other symbols ...”*

University of Southampton’s (2015) research data management policy also identified some possible forms of data in their definition, as it said,

*“1.3 “Research Data” means information in digital, computer-readable format or paper-based that:*

*1.3.1 is contained or presented in various ways including notes, facts, figures, tables, images (still and moving), audio or visual recordings.....”*

This research suggested that the form of data is primarily shaped by research methods, i.e. the way that data is created, an intellectual feature of a research field. There are differences between physical geography and human geography in terms of the form of data. E.g. The interviews showed, data is frequently presented in tables and graphs in physical geography publications but not in human geography publications. Furthermore, the form of data differs between different subfields. As showed in interviews, for the quaternary geographers who use the method of field measurements, their data is commonly in the form of spreadsheets and numbers. For remote sensing geographers, their data is commonly in the form of images. For social geographers who use interviews, their data is generally in the form of recordings and transcripts. It is noted that researchers quite often use more than one method; therefore, a researcher commonly has data with a range of forms.

Although the form of data is one of the most observable features of data, what is surprising is that not all research methods produce data that has clear and recognizable form. For example, this research showed that it was not clear what the physical form of data from qualitative observation is. As shown, although interviewees stated they used qualitative observations, they did not acknowledge data from the method but from other methods, such as interview transcripts or documents from secondary analysis. Only two interviewees indicated that data from observation can be some personal, fragmented notes. This may suggest that not all research methods generate recorded data. In fact, in the discipline of philosophy, as Daly (2015) have pointed out, philosophical research methods are distinct from methods in empirical sciences. Traditionally, philosophical researchers use a priori methods, i.e. without using empirical data (Knobe, 2015). Therefore, researchers in these disciplines may consider they do not generate data with physical entity or any

form of data. Nevertheless, physical form of data is a good starting point to understand the complex and abstract notion of data.

### Intellectual content

The second dimension of data is the intellectual content. The notion of data can be understood from its content. In this research, it was observed that data was quite often described according to its content. For example, physical geographers described their data as “paleoecological data”, “geochemical data”, i.e. data is about paleoecology, geochemistry. The content of data varies between different disciplines and research fields. This has been mentioned by Akers & Doer (2013). The current research showed that the content of data in physical geography and human geography are different. The intellectual content of physical geography data is mostly about the physical environment, such as chemical content in lakes, temperature in caves, whereas the content of data in human geography is associated with human activities, such as, conversations of lived experience from people in Africa, surveys on political and economic behaviours. This shows that research objects and topics, i.e. the intellectual structure of a discipline, are the key elements shaping the intellectual content of data.

#### Characteristics of the content

Intellectual characteristics of a discipline influence what the content of data is, and in turn shape characteristics of the content, e.g. confidentiality of data. Data in different disciplines have different characteristics of content. The following are some characteristics which were found in this research.

##### 1. Confidentiality / non-confidentiality

Confidentiality issues with data collected in research have been well known particularly in social sciences, (e.g. Ritchie et al, 2014; Corti, 2000; Boruch & Cecil, 1979). The findings of this research are consistent with this. This research showed that human geographers, i.e. in the realm of social sciences, are more likely to have confidential or sensitive data than physical geographers. This can be explained by the fact that human geographers’ topics are related to human nature, whereas physical geographers’ research topics are related to physical nature, e.g. climates. It is, therefore, to be expected that human geographers will collect data with sensitive information, such as, names brought up in qualitative interviews, address postcodes collected in quantitative survey. This suggests confidentiality of the content of data is largely shaped by researchers’ research objects rather than qualitative/quantitative methods.

## 2. Objectivity/ subjectivity

This research showed that the content of data can be relatively objective or subjective. In fact, subjectivity and objectivity of data has been well recognized in other disciplines, such as, nursing (Rosdahl & Kowalski,2011). For example, data based on what patients say about their feelings and perceptions, e.g. patients' complaints about discomfort versus data that are measurable e.g. high temperature readings. The objectivity/ subjectivity of the data is shaped by research methods, is dependent on how the data is collected. Data from interviews and observations which are commonly used by human geographers tend to be relatively subjective, especially, when compared to data, such as physical measurements or remote sensing images in physical geography. For example, this research showed that data content of interviews, i.e. the conversation, is actively constructed by the researcher and the participant. As Holstein & Gubrium (2011) and Holliday (2007) pointed out, it is impossible to reconstruct the same data even with the same researcher, participants, research instruments and arrangements. Moore (2006) suggested that it is important to understand the context of how qualitative data was produced when re-using data. Unlike scientific experimental data, interview data cannot be reproduced. In general, the data of human geography is relative subjective and unique, whereas the data of physical geography is mostly objective and could be captured again if needed.

It may seem that subjectivity of data only appears in human geography. Especially as Cutis (2012) pointed out that subjectivity in science is often less emphasised in science literature. Yet, this research indicated that scientific research data are not totally objective. As one of the quantitative modelling interviewees said there are a lot of decisions involved when collecting and processing data, and a lot of the interpretation of data depends on "experiences and hunches". Therefore, as Press & Tanur pointed out (2001), scientific data are not always absolutely objective.

## 3. Divisibility / non-divisibility of data

This research suggests some content of data is more divisible than others. For example, some data can be divided up into pieces while some data cannot be divided and need to be taken as a whole. Compared to the previous two characteristics discussed, the concept of divisibility has not been explored much in the literature, especially in the literature around data sharing. Nevertheless, it was found that quantitative data which are more structured tends to be more divisible than qualitative data which are relatively less structured in nature. For example, the interviews showed that the content of quantitative data from sediment cores collected by physical geographers can be easily separated by a range of variables, such as, years and different chemical components in the sediment

core. By contrast, for some content of data like a qualitative interview data, it is much difficult to divide up the data. As Moore (2006) pointed out it is important to understand the context in qualitative research, thus a complete interview transcript is essential. It is suggested that the divisibility of data is shaped by research methods.

### Social construction

The third dimension of data is social construction. The concept of data not only can be understood physically and intellectually but also socially. Research data is a socially defined and constructed concept. The notion of data is socially constructed by disciplines. Different disciplines define data differently (Gitelman,2013). As shown earlier, data of physical geographers are different from data of human geographers. The perception of data is evidently varied from disciplines and sub-fields. For instance, while physical geographers are very familiar with the term and the concept of data. E.g. physical geographers used data as an input for computational or chemical analysis. In contrast, the interviews indicated that historical geographers are less familiar with the idea of “data”. The term ‘sources’ is used more than ‘data’ in historical geography. In fact, in the discipline of history, data is commonly referred as sources. Typical research methods of historians involve using historical sources, e.g. primary sources, secondary sources (Danto, 2008). Thus it can be suggested that different perceptions of research data are shaped by different research methods used in different disciplines.

The concept of data is relational and constructed. This research showed that data is not simply a fixed thing. For example, for some quantitative human geographers the census was a key data source. By contrast, for some qualitative human geographers the census was not their data. It was only used for informing their research context. Thus, the notion of data is relational. What can be regarded as data is constructed by researchers. It depends on the purpose and how it is used by researchers. Edinburgh University Library (2011:5) suggests that research data “..is collected, observed, or created, for purposes of analysis to produce original research results.” This shows that things become data when a researcher conducts an analysis on them. Thus, the concept of data is situational. This is in accord with the situational view of information. As Hjørland stated (2007: 1449),

*“What is information for one person in one situation needs not be information for another person or in another situation.”*

This research showed that this can be same for research data. Data for someone does not necessarily mean that it is data to others.

Perceptions of research data are not only shaped by research methods but also shaped by the social characteristics of a discipline. For example, disciplines with a higher degree of consensus, such as physical geography, will have a clearly defined conception of what they consider to be data. In contrast, disciplines with a lower degree of consensus and multiple paradigms, such as, human geography, will have more different views on data. As this research showed data in physical geography tends to be quantitative, while data in human geography ranges from quantitative datasets for modelling to interview transcripts, photos, fragmented notes or objects, e.g. stamps collected from observation.

In conclusion, a conceptual framework for research data is developed. The notion of data can be understood from three dimensions, physical forms, intellectual content and social perceptions. These dimensions are strongly shaped by disciplinary cultures. For example, the physical form and intellectual content of data are shaped by intellectual characteristics of disciplines, i.e. research methods and objects. The social construction of data is shaped by both intellectual and social characteristics of disciplines, such as methods and consensus in a discipline. Understanding the notion of research data is an important aspect to understand disciplinary differences in research data and data sharing. However, this aspect has not yet been fully explored. For example, Fecher et al (2015)'s conceptual framework for academic data sharing examined some factors influencing data sharing, yet the only factor related to the notion of data within the framework is "formatting standards", which is related to the physical dimension of data. The three-dimensional framework for research data developed in this research might not be the only way to understand the notion of data, yet it is a useful framework to further understand disciplinary differences in data and data sharing which will be discussed later (Section 11.4).

### 11.3.2 Research / data life cycle and the notion of data

The concept of the research cycle and the notion of data are significantly related. Pryor (2012) suggested that researchers are involved in data activities throughout research cycle, e.g. collecting data, interpreting data and synthesizing data. One of the interesting findings from this research was that data changes throughout the research lifecycle. The three dimensions from the conceptual framework for research data can be evolved throughout research lifecycle (Figure 11.2).



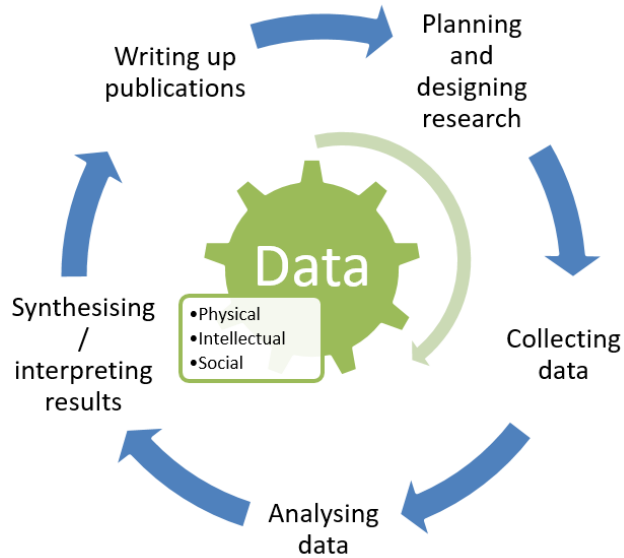


Figure 11.2 Different data aspects changes throughout research process

This includes changes in its physical form, intellectual content and in how it is perceived. Interviews suggested that data are actively processed throughout the research cycle. Data can be raw data, semi-processed or processed. Borgman (2015) suggests that degrees of processing can be a useful category for operational and general purposes. One of the well-known example is NASA’s Earth observing system data information system (2010) which divided data products to different levels, ranging level 0 to level 4, from raw, high resolution data to model output, results from analyses of lower-level data. It seems that the degree of processing has been identified as a feature of scientific data, yet, it can also be relevant to social sciences or humanities data. This research showed quantitative and qualitative data changed throughout the research lifecycle (Figure 11.3).

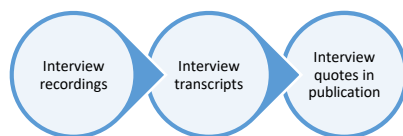


Figure 11.3 An example of interview data changes throughout research process

The changing nature of data during the cycle of research can explain why the concept of data can be interpreted so widely differently among interviewees.

### Data sharing and disciplinary culture

Disciplinary cultures not only shape the notion of data but also data sharing practices of researchers. This section discusses the mutual shaping between disciplinary cultures and data sharing. It explores

what disciplinary cultures and data characteristics from the conceptual framework are more likely to encourage data sharing, using Becher's taxonomy of disciplinary groupings and the three dimensions of the notion of data discussed earlier.

#### 11.4 Applying Becher's taxonomy of disciplinary cultures

Becher's (2001) four dimensional taxonomy, hard-soft, pure-applied, urban-rural and convergent-divergent is one of the classifications that have been developed to explain disciplinary differences. In this research, Becher's framework was adopted to be the conceptual basis for the design and analysis of the research. As mentioned earlier, the framework was chosen because it can not only be applied at a broad disciplinary level but also to narrower levels such as at a sub-disciplinary level. In addition, the framework concerns both intellectual characteristics and social characteristics of disciplines. Thus, it provides good levels of dimensions and terminologies for exploring disciplinary differences within geography for this research. Compared to some other classifications, such as Whitley's, Becher's classification is not too complex to be operational/functional in practice.

From examining patterns between data sharing practices of geographers and Becher's 4-dimension taxonomies, i.e. hard/ soft, pure/ applied, urban/ rural and convergent/ divergent, it is clear that some dimensions have positive relations with data sharing practices. The following presents 4 diagrams which show how each of the Becher's dimension is related to data sharing practices. As particular interview questions were developed to find out researcher's data sharing practices and research characteristics using Becher's framework, it is possible to fit physical geography and human geography interviewees relatively to scales to the 4 dimensions (hard-soft, pure-applied, urban-rural and convergent-divergent) and data sharing practices based on interviewees' answers in the interviews. For example, when an interviewee indicated their research is mostly pure research and with high degree of data sharing, the participant will be placed at the top right corner of the diagram. Thus, 4 diagrams were developed showing the relationship between data sharing practices of geographers and Becher's framework.

##### Hard/ Soft

Among the four dimensions, Hard / soft shows the strongest association with data sharing practices. The hard/ soft concept, originally from Biglan (1973) and Kolb (1981), refers to different levels of paradigm development in different disciplines. The dimension is related to intellectual characteristics of a discipline (Becher & Trowler, 2001). It is common that hard knowledge fields are associated with quantitative, whereas soft knowledge fields are associated with qualitative methods (Becher & Trowler (2001)). The interviews showed that physical geography is predominantly hard,

whereas human geography is predominately soft. This confirms Becher and Trowler (2001: 186) where they described Geography as a ‘borderline discipline’ in-between hard/soft dimension.

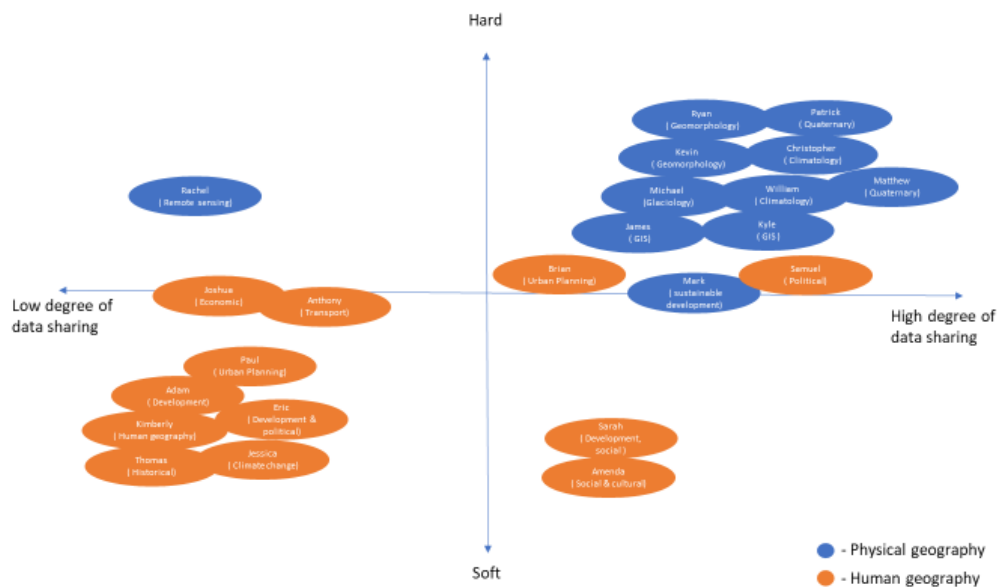


Figure 11.4 Degree of data sharing and hard /soft disciplinary characteristics

This research found that intellectual hard / soft dimension is strongly associated with data practices of geographers. Figure 11.4 illustrates that interviewees in relatively soft fields shared less data than interviewees in hard fields. Such a connection may be explained by the fact that the hard / soft dimension is associated with the intellectual disciplinary characteristics, such as quantitative/ qualitative, impersonal / personal (Becher& Trowler, 2001). As discussed earlier, intellectual characteristics of a discipline shape the notion of data, including physical form and intellectual content of data. In general, quantitative methods in geography produce data in the form of numbers, spreadsheets and images. These contents are usually easily divided by variables and relatively objective. All these intellectual disciplinary characteristic shaped by research methods can, therefore, contribute to hard fields sharing their data more. In contrast, qualitative methods in geography usually produce data in the form of interview transcripts, documents, personal notes from observations. These content of data are not easily divided, are relatively subjective and contain personal content. It is possible, therefore, that these intellectual disciplinary characteristics contribute to less data sharing in soft research fields.

Intellectual disciplinary characteristics, hard/ soft, are not the only factor influencing data sharing practices of geographers. There are some other factors. For example, as the above graph shows (lower left corner), two human geographers in soft research fields shared their data. This can be

explained by the fact that both of their research projects were funded. Thus, their data needed to be made publicly available in order to comply research funder's data policy. Furthermore, the figure also shows that a physical geographer in a hard field did not share data (right hand top corner of the above graph). Possible explanations for this could be this geographer is from a knowledge field where the people-to problem ratio is low, i.e. a rural field and a lower- ranking institution which usually features a smaller sized department. She expressed in the interviews that she is happy to share data yet she did not share data with others as no one had asked. In general, it shows that hard/ soft is not the only dimension influencing data sharing practices. In reality, data sharing practices can be influenced by other factors, such as external factors, e.g. research funder's policies, and social factors, e.g. the people-to-problem ratio.

### Urban / rural

Among the four dimensions, urban / rural also shows a connection with data sharing practices of geographers. Urban / rural dimension concerns the problem-to-people ratio (Becher & Trowler, 2001). An urban field is characterized by a dense population, i.e. a substantial number of researchers pursue one research problem. By contrast, in rural areas, only a small number of researchers work on a research problem. Findings from this research, including the web study, bibliometric and interviews, suggested that physical geography is more urban, whereas human geography is more rural along the urban/ rural continuum.

Evidently, the rural /urban nature of a field has a significant relation to researchers' data sharing practices. Figure 11.5 graphically illustrates that there is a clear pattern of most urban researchers sharing data, whereas the majority of rural researchers did not share data. This may be explained by the fact that there are tighter communication and networks in urban communities (Becher and Trowler, 2001), therefore it is easier and there are more opportunities to share data, and in turn these increase researchers' motivations to share data. As this research showed, researchers tend to share or reuse others' data when it is related to their own research fields. Urban researchers tend to receive a higher level of interest for their data. Furthermore, physical geographers are more collaborative. They have had experiences of sharing data within collaborators. Thus they are already familiar with the culture of sharing data.

By contrast, rural research communities are spread out sparsely across a range of topics (Becher & Trowler,2001). There are less opportunities and it is more difficult to share data. As shown in the interviews, a few researchers mentioned that they do not oppose the idea of data sharing. Yet they did not share data because no one has asked. This suggests that the level of interest in rural researchers' data is relatively low. This can affect researchers' incentive to share data. An

implication of this is that the urban/ rural dimension can be a significant factor which influences researchers' incentive to share data.

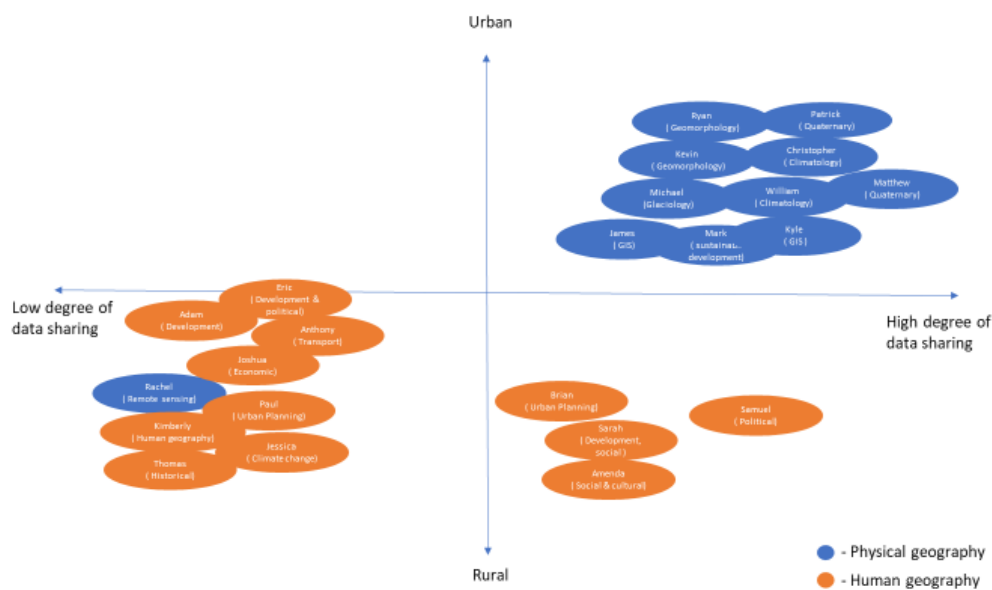


Figure 11.5 Degree of data sharing and urban/ rural disciplinary characteristics

Figure 11.5 shows that there are again a few exceptions. Four geographers in relatively rural fields shared data. It is unclear why one of the rural geographers shared data. A possible explanation for this might be that he does not collect data himself but uses quantitative data that is already publicly available, as his research method mainly involves using secondary data. However, that the other three geographers shared their data could be explained by the fact that they are funded by research funders. Thus, they shared data in order to comply with research funders' policies. This shows again research funders have a significantly influence on researchers' data practices.

### Convergent/ divergent

Convergent/ divergent is another dimension that shows a strong connection with data sharing practices of geographers. The convergent / divergent dimension is related to the level of consensus in a discipline. Convergent communities have strong consensus to intellectual styles, e.g. they have dominant subject matters and highly developed methods and techniques, whereas divergent communities allow greater differences in research topics and methods (Becher & Trowler, 2001). Thus there are less uniform standards and procedures in divergent communities.

This research found that this social dimension of convergent/ divergent is strongly associated with data practices of geographers. This can be explained by the fact that the dimension is closely associated with the social perception of data. As the figure below shows geographers in convergent

communities share data, while most geographers in divergent communities do not share data. This could be because convergent communities have strong intellectual consensus and a sense on commonality, e.g. based on research techniques, subject matter and data. As discussed earlier, the notion of data is strongly associated with knowledge structure. In general, convergent communities have a strong consensus on techniques, standards and what data is, these all encourage data sharing in their communities. In contrast, divergent disciplines have less uniform standards, diverse of data. The lack of a sense of commonality can pose difficulties for researchers in divergent communities to share data.

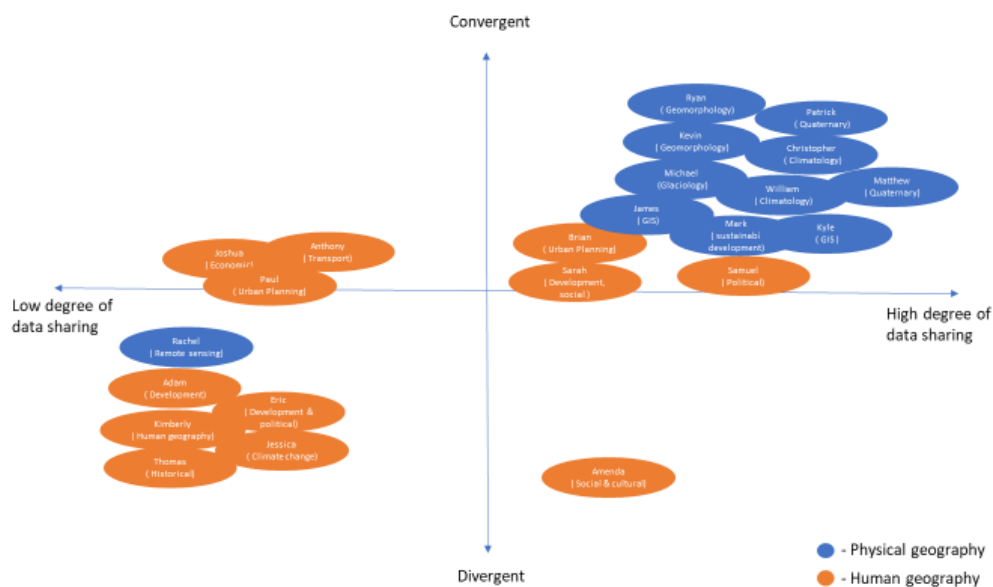


Figure 11.6 Degree of data sharing and convergent/divergent disciplinary characteristics

### Pure/ applied

Pure/applied is the only dimension that does not demonstrate a strong connection to data sharing experiences among the four dimensions. This dimension concerns application of research to practical issues, which is also related to the intellectual nature of a discipline (Becher & Trowler, 2001; Biglan, 1973). As Figure 11.7 showed, there is not much pattern between pure / applied and experiences of data sharing.

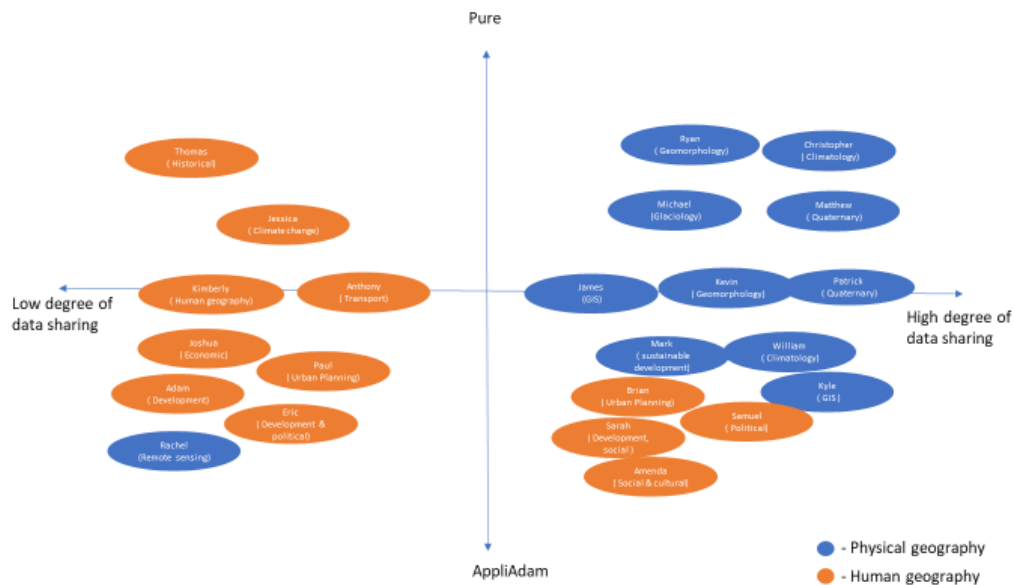


Figure 11.7 Degree of data sharing and pure/applied disciplinary characteristics

It is unclear why there is no strong connection. This could be because increasingly research would have both pure and applied element. Especially research funders have been increasingly focused on applied research (Else, 2014). As this research showed, almost half of the interviewees said their research had both pure and applied elements. Therefore, pure / applied can become a less important dimension as an intellectual feature of a discipline.

In general, hard /soft, urban/rural and convergent/ divergent are useful dimensions for understanding data sharing differences across geography.

## 11.5 Conclusion

This chapter discussed three significant themes from the findings of this study: the characteristics of geography, the conception of data and data sharing practices in geography.

### *The Characteristics of Geography*

The data collected and analysed in the web content analysis, the bibliometric study and interviews enhanced our understanding of geography as discipline. The findings described the characteristics of geography in detail. For example, it showed that physical geographers solely used quantitative methods whereas human geographers used both qualitative and quantitative methods, including quantitative survey and statistical modelling. Furthermore, it showed both human and physical geography are collaborative: most papers are jointly authored, though physical geographers have

more co-authored papers and more co-authors per paper. While some of the literature suggests that the emergence of open science and big data may bridge the divide between physical geography and human geography, the findings showed that the divide unquestionably still exists. As Zelenkauskaitė and Bucy (2015) suggested, big data can induce a 'scholarly divide' between researchers who have the technical skills to acquire and analyse data sets and those who do not. This research showed that differences not only appeared between human and physical geography but also within them. For example, in human geography, while quantitative methods are commonly used by economic geographers, qualitative methods are more common in social geography. This study showed the nature of discipline is complex and fragmented in nature, which is opposed to what people tend to think discipline as static and uniform. As Dogan & Pahre (1990), Klein (1996) and Abbott (2000) described, disciplines tend to grow and sub-divide. Understanding this complex and fragmented character of disciplines is vital, as the complexity of disciplines is reflected in the nature of data, researchers' data sharing practices and attitudes.

The understanding of the characteristics of geography from this study was not only empirical but also theoretically based. A conceptual framework differentiating intellectual and social aspects of disciplines was used to guide the research design and analysis throughout this study. This study showed that examining disciplines through intellectual and social aspects is a good way to acquire a deeper disciplinary understanding. The intellectual aspect is commonly associated with intellectual ideas and cognitive norms in a research field, e.g. research topics and methods. The social aspect is associated with communications and collaboration. For example, the bibliometric study examined the co-authorship pattern in geography. The intellectual aspect is commonly given more attention, whereas the social aspect of a discipline is overlooked when examining disciplinary differences. Nevertheless, this study suggested both aspects are important for further understanding of data practices and data sharing.

### *The Conception of Data*

The second theme discussed in this chapter is the conception of data. This study suggested a new framework for understanding the notion of data from a disciplinary perspective. The findings showed that data is very diverse in both physical geography and human geography. For example, data can be digital and non-digital. It can be interview data, documents, physical measurements, remote sensing images etc. This study found that these data differences can be categorised along three key dimensions: physical form, intellectual content and social construction.



The first dimension is the physical form of data. It is associated with the physical characteristics of data. For instance, the findings showed data can be in the form of digital or physical objects; can be in form of numbers, text, images; can be in different formats or standards.

The second dimension of data is the intellectual content. Data can be understood from its content. For example, from the web study and interviews, some of the data content in physical geography were chemical content in lakes, or temperatures in caves, whereas data content in human geography were about lived experience from people in Africa from interviews, political behaviours discovered through surveys.

The third dimension is social construction. This dimension is based on an understanding that data is a socially defined and is a relational concept. For example, the interview findings showed while the census can be a key data source for a quantitative human geographer, another geographer who only used census for informing research context did not see census as his data. Thus, research data is relational.

Importantly, these three dimensions are not simply empirical findings but grounded in theories of discipline. The three dimensions of data are strongly shaped by disciplines. The discussion showed intellectual characteristics of disciplines shape all three dimensions of data. For example, research objects and methods of a researcher not only determine the format and content of the data but also influence a researcher's interpretation of data. Furthermore, social characteristics of disciplines, e.g. consensus in a field, shape the social construction of data. It affects what can be considered as data in a research field.

### *Data sharing practices*

The third theme discussed is data sharing practices in geography. The findings of interviews showed that data sharing practices differ within geography, especially between physical geography and human geography. For instance, physical geographers had extensive experiences on sharing data beyond research teams, whereas most human geographers did not have data sharing experiences. Yet, this chapter showed the differences in data sharing practices are grounded in something more fundamental, i.e. the characteristics of disciplines. Becher's 4-dimensional taxonomies i.e. hard/ soft, pure/ applied, urban/rural and convergent/ divergent, was used to characterise data sharing practices of geographers. The section of the chapter demonstrated there is a relationship between the actual data sharing behaviour and Becher's taxonomies.

The hard/ soft dimension which is associated with intellectual disciplinary characteristics had the strongest connection with data sharing practices of the four dimensions. Researchers in soft fields

tend to share less data than those in hard fields. It is highly probable that the qualitative and personal nature of soft research fields (Becher & Trowler, 2001) generate more barriers for sharing data, such as confidential and anonymity issues than a hard research field which is quantitative and impersonal in nature (Becher & Trowler, 2001). Therefore, less data is shared in soft disciplines.

The urban / rural dimension also has a connection with data sharing practices of geographers. This dimension concerns the problem-to-people ratio (Becher & Trowler, 2001). It was shown that most urban researchers shared their data, while the majority of rural researchers did not share data.

There is a strong possibility that tighter communication and networks in urban communities provide more opportunities to share data. On the other hand, rural researchers spread out sparsely across a range of topics, thus had less opportunities and it was more difficult to share data.

Convergent/ divergent is another dimension which shows a strong connection with data sharing practices of geographers. The dimension is related to the level of consensus in a discipline. It was shown that geographers in convergent communities shared data, while most geographers in divergent communities did not share data. Convergent communities have a strong consensus on techniques, standards and what data is. Thus, all these can enable data sharing in their communities. In contrast, divergent disciplines have less uniform standards and diverse of data. Therefore, the lack of a sense of commonality in divergent communities can pose challenges for researchers to share data.

Pure/applied is the only dimension that did not demonstrate a strong connection to data sharing experiences among the four dimensions. This could be because increasingly research has both a pure and applied element (Else 2014). Therefore, it becomes a less significant dimension as an intellectual characteristic of a discipline. In general, broad patterns of data sharing practices were revealed by using Becher's framework. Hard, urban and convergent disciplinary characteristics encourage data sharing, whereas disciplines with soft, rural and divergent characteristics experience more challenges in data sharing.

To conclude, this chapter presented three important themes, the characteristics of geography, the conception of data and data sharing practices in geography. All the three themes were discussed and conceptualised using theories in disciplines. It demonstrated the value of theories of discipline are not only useful to understand the notion of disciplines but also research data and research data sharing practices.

# Chapter 12 Conclusion

---

## 12.1 Research questions and main findings of the study

The purpose of the current study was to investigate the relation between data practices of researchers and the culture of a discipline, focusing on the discipline of geography. In order to achieve the aim of this research, three research sub-questions were set up at the beginning. This section shows how each of the research sub-questions has now been answered.

### 1. What are the characteristics of geography as a discipline (i.e. intellectual, social and organisational aspects)?

This study examined the characteristics of geography as a discipline from three different aspects, intellectual, social and organisational aspects. In general, this research showed that there are clear disciplinary differences between physical geography and human geography in all three aspects.

#### Organisational aspect

The first aspect investigated was the organisational aspect of geography as a discipline. In chapter 4, the findings of the web based content analysis on institutional departments & research groups revealed some organisational characteristics of geography in institutions. The analysis showed that there were clear differences in how geography was organized between institutions, such as, top ranking institutions tended to have individual geography departments; whereas in lower ranking institutions, geography as a discipline tended to be without an independent department and was integrated into multi-disciplinary schools, e.g. school of applied sciences. The analysis also showed that there is a clear divide between human geography and physical geography at the research group level. It was very unusual to have a research group with both physical and human geography elements.

#### Intellectual aspect

Research topics and methods are important intellectual elements of a discipline. The fifth and seventh chapter of the thesis examined these intellectual elements of geography. For example, the findings of the web based content analysis of individual researchers from chapter 5 identified some key subfields in physical geography and human geography. The major areas in physical geography were geomorphology, past environmental change, hydrology, climatology and glaciology whereas human geography had political geography, social geography, economic geography, urban geography and development geography, in addition, military geographies, carceral geography, geographies of

children, geographies of ageing etc. The findings suggested that topics in human geography is more fragmented than physical geography. Furthermore, the analysis showed that human geography research methods are relatively wide-ranging than physical geography. For example, while physical geography focuses on quantitative methods, human geography not only included qualitative methods, such as interviews, ethnography, visual methods, but also quantitative methods, such as, statistical analysis, spatial data analysis. Using Becher and Trowler's (2001) conceptual terms, the knowledge structure of physical geography can be considered as relatively 'hard', whereas the knowledge structure of human geography is more varied, range from 'hard' to 'soft'.

The seventh chapter of the thesis further revealed more intellectual features of geography. Certain research methods are associated with particular research fields. Research methods discussed in detail included methods of physical geographers, i.e. physical measurements, remote sensing, numerical modelling, and methods of human geographers, i.e. interviews, observation, documentary/archival research, statistical modelling and quantitative survey. The findings showed that research data and methods are mutually shaped. Data practices and research methods are unquestionably inter-related, as research methods essentially is about how data is constructed and interpreted.

### Social aspect

The social characteristics of geography as a discipline, including co-authorship patterns, research audiences of geographers and how geographers collaborate were discussed in the sixth and eighth of this thesis. This research showed that the social differences between physical geography and human geography are noticeable.

Chapter 6 presented the findings of the bibliometric study. It investigated collaboration patterns of geographers through co-authorship. The analysis suggested that physical geographers published almost double the quantity of publications compared to human geographers. The results also indicated that physical geographers were more collaborative, e.g. in physical geography, more than 95 % of publications were co-authored, while in human geography, almost a quarter of publications were single authored. Furthermore, it was suggested that physical geographers had bigger collaboration teams than human geographers, as co-authors per publication in physical geography was around 6 while in human geography was around 2.

The eighth chapter of the thesis revealed other social characteristics of geography, such as, the nature of conferences in which geographers were involved and their audiences. Physical geographers not only went to specialized conferences but also to multi-disciplinary conferences. By

contrast, human geographers tended to only attend specialized conferences. It was suggested that physical geographers had a wider range of audiences and hence more potential re-users of their data.

Furthermore, how geographers collaborate are different between physical and human geography. This research showed that, in physical geography, members in a collaboration team often had clearly defined, separate roles. Physical geographers usually shared some part of the processed data rather than whole set of raw data with team members. It was explained that this is because of the holistic nature of physical geography research. On the other hand, collaborators in human geography tended to work together and use the same set of raw data. The evidence from this study suggested that social disciplinary characteristics shaped data sharing practices of researchers.

In general, the findings of this research suggested, physical geographers are more collaborative, have a wider range of audience. Using Becher's (2001) terms, physical geography is considered as relatively 'urban'. In contrast, human geography is relatively 'rural', which has relatively low people-to-problem ratio.

## 2. What are the data practices (i.e. collecting, managing, using and especially sharing data) of researchers in geography?

Data practices of researchers are closely related to research methods. Thus, in this research, researchers' data practices in geography are illustrated in the seventh chapter of the thesis by different methods, including physical measurements, remote sensing, numerical modelling, interviews, observation, documentary/archival research, statistical modelling and quantitative survey. This research found that data practices of researchers, such as, collecting, managing and using data are largely shaped by intellectual elements of a discipline. Research methods and objects not only shape researchers' data practices, but more fundamentally they shape the characteristics of data, including the format of the data, the volume of data and the confidential nature of data. These influence different researchers interpreted the concept of data differently, for instance, an interviewee perceived mud samples as data, whereas another interviewee perceived research output as data. Thus, not only data practices but also the meaning of data vary across different subfields in geography.

This research also presented data sharing practices in geography. One of the key findings was there were clear differences in both external and internal data sharing practices within geography. While almost all physical geography interviewees had extensive experiences of sharing data, only very few human geographers had experiences of sharing their data beyond research teams. Furthermore,

interestingly, the nature of data which researchers shared with collaborators differed among geographers. Sharing raw data within research teams was not common among physical geographers, they tended to share research output or only part of the data with collaborators. In contrast, human geographers, specifically qualitative human geographers, emphasized sharing raw data, such as interview transcripts. And they made sure research team members had the same datasets as each other.

### 3. Do patterns of data practices, especially data sharing vary across in different sub-disciplines geography and, if so, why?

Significant differences in the nature of data and data practices between physical geography and human geography were found in this research. Research data from different disciplinary fields has different characteristics. Based on the empirical findings from phase 1 and phase 2, this study developed a conceptual framework for understanding the notion of research data from a disciplinary perspective. The study identified data characteristics can be understood from three dimensions: physical form, intellectual content and social construction. For instance, data can be digital/non-digital in terms of its physical form. It can be confidential or non-confidential in terms of its intellectual content. And in terms of social construction, data can have different meanings to different people, e.g. for a same object, some researcher see it as data, while some do not. Furthermore, the nature of data is dynamic. It changed constantly during research. The characteristics of the data of the three dimensions can change during research cycles, e.g. the form of data can be changed from numerical raw data to visual graphs during research processes. Another key theme discussed in this research was using Becher's 4-dimension taxonomies to explain the relation between data sharing practices of geographers and disciplinary cultures. The research found that researchers in a research field which were hard, urban or convergent shared more data. While researchers in a research field which were soft, rural or divergent tended to not share much data. In general, the developed conceptual framework of research data not only enhances our understanding of the notion of data but also provides a deeper understanding of why disciplinary differences in research data exist from a more fundamental perspective. Furthermore, this study has demonstrated that Becher's 4-dimension taxonomies are useful to explain why data sharing practices vary across in different disciplinary fields.

## 12.2 Contribution to knowledge

This section addresses the empirical contribution, theoretical contribution and practical implications of this research.

### *12.2.1 Empirical contribution*

This study makes an empirical contribution to the field of scholarly communication in Information Science, specifically the newer area of research data management. Scholarly communication is concerned with how scholars use and disseminate information formally and informally (Borgman, 1989). The findings of the study provide a better understanding of disciplinary cultures and research processes which are important to the support for the advancement of scholarly knowledge.

Research data management is a relatively new area of interest within information research. In the past ten years, partly driven by policy change, research data has been increasingly recognised as a key product of scholarship and emerged as a new form of scholarly communication. In this context, this current study is the first in-depth study of research data practices in UK geography. The study contributes to the field by providing a greater understanding of diverse research and data practices in UK geography.

### *12.2.2 Theoretical contribution*

This study makes not only an empirical contribution but also a theoretical contribution to Information Science. The research area of research data management is still at the beginnings of developing conceptual theories to understand topics around research data. The first theoretical contribution of this study is it proposes a conceptual framework to characterise the notion of research data. This study investigated the complex and dynamic nature of research data. One of the significant findings to emerge from this study is that the characteristics of data are so diverse even within one discipline. Drawing on the empirical findings, this study suggests that disciplinary differences in data can be understood through three dimensions, physical form, intellectual content and social construction. This conceptual framework was developed not only by drawing on empirical findings but also grounded in theories of discipline. It was suggested that the three dimensions are strongly shaped by intellectual and social characteristics of disciplines. Furthermore, the conceptual framework can use to illustrate the dynamic nature of data, i.e. the three dimensions evolve during research processes. For example, the physical form of data can change from sound recordings to interview transcripts in written form during research process. Thus, the conceptual framework is a useful way to understand the notion of research data. The framework assists in our understanding of disciplinary differences in data and explain the dynamic nature of data during the research process.

Although the framework was developed through findings from the discipline of geography, it seems reasonable to suppose that it could be extended to other research disciplines.

The second theoretical contribution of the study to demonstrate the usefulness of disciplinary theories in information science. This study is an early work in the field to apply disciplinary theories to research data management. It applied Becher's disciplinary framework to explain differences in data sharing practices in geography, in a way which has not been previously attempted. While Becher's work has been employed in information science, e.g. Fry et al. (2015), this study is the first major application to the study on research data sharing practices. This study uses Becher's framework to map the complexity of data sharing practices in one discipline. While disciplinary labels such as, human geography, physical geography are useful broad categories, it was found that Becher's typology provides a more fundamental perspective and a better explanation to disciplinary differences in data sharing practices. The findings of this study suggest that not only different knowledge structures but also social characteristics of a field shape data sharing practices of researchers. The findings show some patterns in data sharing practices, e.g. disciplines with hard, urban and convergent characteristics are most likely to share data, whereas disciplines with soft, rural and divergent characteristics are less likely to share data. Although a lot of studies have highlighted disciplinary differences in data sharing, using the framework in this study provides insights and fundamental understanding of how and why these differences exist. It shows that understandings of issues can be enriched by employing theories and exploring theoretical connections. This study shows Becher's framework is useful and valuable to inform information science work.

The framework not only explored intellectual but also social characteristics of disciplines. It can be used to compare not only differences within geography but also in other disciplines or across disciplines. Furthermore, Becher's framework can be employed to other research areas in scholarly communication, in information sciences, including, issues in data-reuse, attitudes towards data sharing and open access.

### 12.2.3 Methodological contributions

This study used a combination of methods, including, bibliometric analysis, web based analysis and interview data for an in-depth understanding of a discipline and its data practices. This study demonstrates that the use of multiple methods is an effective way to investigate different disciplinary characteristics in the context of RDM. Furthermore, the methods enabled the study to offer a comparison of diverse disciplinary differences in data in one single discipline. This study is



different from the common approach which usually compares disciplinary differences at a disciplinary level, for example, RIN (2008) DCC SCARP (2010) and Van den Eynden & Bishop (2014)

### 12.3 Practical Implications of the study

The findings of this study provide a deeper fundamental understanding for developing services, infrastructure and tools to support academic scholarship.

This research has shown that researchers' data sharing practices are grounded in the underlying culture of disciplines. The fundamental understanding of the concept of data and its relationship with disciplinary cultures have a number of implications to the practical implementation of good research data management, especially at institutional level. Higher Education Institutions are increasingly giving greater priority to the management of research data. Top-down approaches are commonly used for strategies around institutional research data management (Pryor et al, 2014). The present study provides a bottom- up approach to help to developed researcher- focussed data management services.

#### The concept of data

Looking at some definitions of research data developed by institutions, data is commonly referred to as a thing, as a static identity. In reality, research data is more complicated than usually thought. As the framework developed in this study established, data can be understood from three aspects, physical, intellectual and social, and all these three aspects can evolve during the life cycle of researchers' research. The conceptual framework, therefore, helps to understand not only the notion of data but also the dynamic nature of data. The conceptual framework to explain research data practices developed in this thesis provides a more detailed understanding of the concept of data and will support the development of data management services and tools in order to help researchers decide what are the best data to be kept.

#### Rationale for managing research

There are some common rationales for managing research data that are being used, for example by librarians, to advocate good research data management practices and data sharing. Rationales such as: good research data management avoids data loss, is a good professional practice, is compulsory for funded research, increases reproducibility of research and increases research impact if data gets shared and reused. This study found that while most researchers support data sharing, there are marked disciplinary differences in reasons researchers support the idea of data sharing. For example, for physical geographers, increasing validity of research is a key motive for them to support data sharing. By contrast, for human geographers, the reasons for them to support data sharing are to

use resources efficiently, to avoid repetition of effort in data collection and to fulfil funders' requirements. None of the human geography interviewees mentioned that they support sharing data is to increase transparency, validity or replicability. It follows that in advocating RDM, institutional services, such as libraries should be aware that different disciplinary fields have different arguments for managing research data and sharing research data. Not all rationales mentioned by policy makers are relevant to all researchers. Thus, in terms of delivery of research data management training course and advocacy of good research practices, it is important to aware of disciplinary differences. Understanding these differences, will make it easier for information professionals to promote research data management services among researchers.

#### Differences within one discipline

One of key findings from this research is that there are differences within one single discipline. Both intellectual characteristics, e.g. methods, data, and social characteristics, e.g. collaboration patterns, are different within a discipline. This implies there are different needs for data support even within one single discipline or academic department. Academic departments themselves are diverse. Thus, it is essential not to assume disciplines or departments are unified units when providing research data services.

#### Becher's taxonomies

This study showed that Becher's taxonomies can be a useful tool for information professionals to identify disciplinary and data differences. In this research, it was established that knowledge structures are strongly related to data sharing. Researchers in 'hard' research fields tended to share more data than researchers in 'soft' research fields. Even within one department, there can be many different sub disciplines with different cultures. Thus, in terms of providing research data management training courses and support websites, it can be more effective to deliver training by knowledge structures, such as by methods or research topics, rather than by department. Those using the same method are more like each other than those who study the same subject. This approach will also go beyond disciplinary boundaries and encourage interdisciplinary collaboration.

#### Internal data sharing

In this research, interviewees mentioned that there is not much institutional support for internal data sharing, i.e. data sharing with collaborators, when they are doing research. Most interviewees used dropbox to share their data with collaborators even some interviewees who were aware of security issues of dropbox. This suggests that there is a need to provide more guidance and infrastructure to researchers on internal data sharing and management.

### Informal data sharing

While physical geographers have experience of data sharing, a lot of the interviewees said they prefer informal sharing. The findings of this study indicated that researchers prefer to have direct contact with the re-user, so they understand how data is re-used and able to provide suitable help. The implications can be institutions can encourage researchers to share a record that shows the data exists, especially for unfunded research, and the data can be obtained by direct contact. It is possible that this will help developing a culture of open data gradually.

## 12.4 Future Research

There are a number of potential avenues for future research, building on this study.

### *1) applying Becher's framework to other disciplines*

In this study, Becher's framework was introduced to understand research data and data sharing practices. It revealed some patterns in data sharing practices. The characteristics of hard, urban or convergent tend to encourage sharing data, whereas soft, rural or divergent characteristics tend to pose more challenges for sharing data. However, the findings were based on one single discipline, the discipline of geography. Therefore, it is suggested applying Becher's framework to other research fields in future research in order to validate the data sharing patterns found in this study. For example, comparing research fields with hard and convergent characteristics to research fields with soft and divergent characteristics, using methods such as quantitative surveys.

### *2) employing Becher's framework to investigate other research data related issues*

In this study, Becher's framework was used to conceptualise differences in data sharing practices. Future research can benefit from applying Becher's framework to conceptualise other data related issues and compare the findings with this study, for example, to examine whether similar patterns in data sharing practices will be found in disciplinary differences in data re-use and data sharing attitudes.

### *3) further investigations and applications of the conceptual framework of data*

This research proposes a conceptual framework to characterise the notion of data. It categorised data differences into three key dimensions, physical form, intellectual content and social construct, however, further research is required to prove the generalisability and usability of the concept. It is suggested that this concept can be investigated by applying it to data in other disciplines or the concept of how the three dimensions evolve throughout research cycle can be explored in more

detail. The framework can also be used to evaluate research data in data repositories across disciplines.

#### *4) further investigating data and disciplinary differences using different methods*

Mixed methods used in this research were effective to examine the relation between disciplines and data. Thus, it is suggested future research to use multiple methods to further investigate the complex nature of disciplinary cultures and research data. The bibliometric method used in this study was evidently a useful way to examine collaboration patterns in a research field. However, the scale of the bibliometric study in this research was quite small. Thus, it is suggested that future research can be done on a larger sample and a longer period, for example, across two REF periods instead on one. Other methods might also be considered in future research, for example, using citation analysis to examine the effect of data publication on citation impact across disciplines, or applying social network analysis to understand data sharing behaviours within and across institutions.

Furthermore, disciplinary cultures and research data practices are not easy to examine through qualitative interviews. Thus, it is suggested further insights into disciplinary cultures and data culture could be achieved through using ethnography to gain a richer perspective on disciplinary culture data practices. This will give insights of day to day practices on research data, and how what people say they do relates to what they actually do.

## 12.5 Closing Comments

The data sharing agenda has been driven by research funders, journals and institutions. For some researchers, data sharing is a new and unfamiliar idea. A significant cultural change will be needed for it to become a common practice. To inform such changes, we need to know more about current research cultures and data practices. Theory-driven empirical studies are a powerful way of analysing such cultures and practices. This study has made an initial attempt to construct a framework for such analyses based on theories of disciplines. More theory-driven research is needed to give deeper insights on fundamental issues around data sharing and research data.

# Bibliography

---

- Abbott, A. (2001). *Chaos of Disciplines*. Chicago: University of Chicago Press.
- Aboelela, S. W., Larson, E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S. A., Haas, J. and Gebbie, K. M. (2007), Defining Interdisciplinary Research: Conclusions from a Critical Review of the Literature. *Health Services Research*, 42, 329–346. <http://doi.org/10.1111/j.1475-6773.2006.00621.x>
- Abreu, A. & Acker, A. (2013). Context and Collection: A Research Agenda for Small Data. iConference 2013 Proceedings, 549–554. <http://doi.org/10.9776/13275>
- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16(1), 3–9. <http://doi.org/citeulike-article-id:6930744>
- Aditomo, A., Goodyear, P., Bliuc, A.-M. & Ellis, R. A. (2013). Inquiry-based learning in higher education: principal forms, educational objectives, and disciplinary variations. *Studies in Higher Education*, 38(9), 1239–1258. <http://doi.org/10.1080/03075079.2011.616584>
- Akers, K. G. & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, 8(2), 5–26. <http://doi.org/10.2218/ijdc.v8i2.263>
- Alexogiannopoulos, E., McKenney, S. & Pickton, M. (2010). *Research data management project: a DAF investigation of research data management practices at the University of Northampton*. Northampton: The University of Northampton.
- Allsop, D., Allen, H., Clare, H., Cook, I. & Raxter, H. (2010). Ethnography and participant observation. In *Research Methods in Geography: A Critical Introduction*. Chichester, UK: Blackwell.
- Association of Internet Researchers. (2012). Ethics guide. <http://aoir.org/documents/ethics-guide/>
- Bailey, J. (2008). First steps in qualitative data analysis: Transcribing. *Family Practice*, 25(2), 127–131. <http://doi.org/10.1093/fampra/cmn003>
- Baker, S. E. & Edwards, R. (2012). *How many qualitative interviews is enough?* National Centre for Research Methods.
- Ball, A. (2012). *Review of Data Management Lifecycle Models*. Bath, UK: University of Bath. <http://opus.bath.ac.uk/28587/>
- Beagrie, N. & Pink, C. (2012). *Benefits from Research Data Management in Universities for Industry and Not-for-Profit Research Partners*. Charles Beagrie Ltd and University of Bath. <http://opus.bath.ac.uk/32509/>
- Beagrie, N., Beagrie, R. & Rowlands, I. (2009). Research Data Preservation and Access: The Views of Researchers. *Ariadne*, 60. <http://www.ariadne.ac.uk/print/issue60/beagrie-et-al>

- Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, 19(2), 151–161. <http://doi.org/10.1080/03075079412331382007>
- Becher, T. (1987). The disciplinary shaping of the profession. In B. R. Clark (Ed.), *The Academic Profession*. Berkeley: University of California Press.
- Becher, T. & Trowler, P. R. (2001). *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. Philadelphia, PA: Open University Press.
- Berg, B. (2009). *Qualitative Research Methods for the Social Sciences* (7th ed.). Boston: Allyn & Bacon.
- Berger, G. (1970). Introduction. In *OECD-CERI Indisciplinarity: Problems of Teaching and Research in Universities*. Nice: CERI/ French Ministry of Education.
- Berlin declaration on Open Access to Scientific Knowledge. (2003). <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>
- Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57(3), 195–203. <http://doi.org/10.1037/h0034701>
- Borgman, C. L. (1989). Bibliometrics and Scholarly Communication: Editor's Introduction. *Communication Research*, 16(5), 583–599. <http://doi.org/10.1177/009365089016005002>
- Borgman, C. L. (2015). *Big data, little data, no data*. Cambridge, Massachusetts: MIT Press.
- Borgman, C. L. (2009). The Digital Future is Now: A Call to Action for the Humanities. *Digital Humanities Quarterly*, 3(4). <http://www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html>
- Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, Massachusetts: MIT Press.
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078. <http://doi.org/10.1002/asi.22634>
- Boruch, R. F. & Cecil, J. S. (1979). *Assuring the Confidentiality of Social Research Data*. Philadelphia: University of Pennsylvania Press.
- Bose, R. & Reitsma, F. (2005). *Advancing Geospatial Data Curation*. Institute of Geography. The School of Geosciences. The University of Edinburgh. <https://www.era.lib.ed.ac.uk/handle/1842/1074>
- Boyatzis, R. E. (1998). *Transforming Qualitative Information: thematic analysis and code development*. Thousand Oaks, CA: Sage Publications.
- Boyd, D. & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679. <http://doi.org/10.1080/1369118X.2012.678878>
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <http://doi.org/10.1191/1478088706qp063oa>

- Braxton, J. & Hargens, L. (1996). Variation among academic disciplines: Analytical frameworks and research. *Higher Education: Handbook of Research and Theory*, 11, 1–46.
- Brown, S., Bruce, R. & Kernohan, D. (2015). *Directions for Research Data Management in UK Universities*. JISC. <https://www.fosteropenscience.eu/sites/default/files/pdf/1240.pdf>
- Bryman, A. (2016). *Social Research Methods* (5th ed.). New York: Oxford University Press.
- Budapest Open Access Initiative. (2002). <http://www.budapestopenaccessinitiative.org/>
- Cetina, K. (1999). *Epistemic Cultures*. Massachusetts: Harvard University Press.
- Chettiparamb, A. (2007). *Interdisciplinarity: a literature review*. Southampton: The Interdisciplinary Teaching and Learning Group, University of Southampton.
- Clark, B. (1983). *The Higher Education System: Academic Organization in a Cross-National Perspective*. Berkeley: University of California Press.
- Clifford, N., French, S. & Valentine, G. (2010). *Key Methods in Geography*. London: Sage Publications.
- Cloke, P., Cook, I., Crang, P., Goodwin, M., Painter, J. and Philo, C. (2004). *Practising Human Geography*. London: Sage.
- Cloke, P., Crang, P. & Goodwin, M. (Eds.). (2013). *Introducing Human Geographies* (3rd ed.). Oxon: Routledge.
- Conant, J. B. (1951). *Science and Common Sense*. New Haven, Conn.: Yale University Press.
- Corrall, S., Kennan, M. A., Afzal, W. & Sen, B. (2013). Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research. *Library Trends*, 61(3), 636–674. <http://doi.org/10.1353/lib.2013.0005>
- Corrall, S. (2012). Roles and responsibilities: libraries, librarians and data. In G. Pryor (Ed.), *Managing Research Data*. London: Facet.
- Corti, L., Day, A. & Backhouse, G. (2000). Confidentiality and Informed Consent: Issues for Consideration in the Preservation of and Provision of Access to Qualitative Data Archives. *Forum: Qualitative Social Research*, 1(3). <http://doi.org/ISSN 1438-5627>
- Coulthard, T. J. & Van De Wiel, M. J. (2012). Modelling river history and evolution. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1966), 2123–2142. <http://doi.org/10.1098/rsta.2011.0597>
- Cox, A. M. & Pinfield, S. (2013). Research data management and libraries: Current activities and future priorities. *Journal of Librarianship and Information Science*, 46(4), 299–316. <http://doi.org/10.1177/0961000613492542>
- Cox, A., Verbaan, E. & Sen, B. (2014). Rising to the research data management challenge. *SCONUL Focus*, 60, 42–44. [https://www.sconul.ac.uk/sites/default/files/documents/13\\_13.pdf](https://www.sconul.ac.uk/sites/default/files/documents/13_13.pdf)

- Cox, A. & Williamson, L. (2015). The 2014 DAF survey at the University of Sheffield. *International Journal of Digital Curation*, 10(1), 210–229. <http://doi.org/10.2218/ijdc.v10i1.362>
- Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 368(1926), 4023–4038. <http://doi.org/10.1098/rsta.2010.0165>
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago: University of Chicago Press.
- Cresswell, T. (2013). *Geographic Thought: A Critical Introduction*. New York: Wiley-Blackwell.
- Creswell, J. W. (2013). *Qualitative Inquiry and Research design: Choosing among Five Approaches*. Los Angeles: Sage Publications.
- Curtis, A. (2012). The Science of Subjectivity. *Geology*, 40(1), 95–96. <http://doi.org/10.1130/focus012012.1>
- Daly, C. (Ed.). (2015). *The Palgrave Handbook of Philosophical Methods*. London: Palgrave Macmillan UK. <http://doi.org/10.1057/9781137344557>
- Danto, E. A. (2008). *Historical Research*. Oxford University Press. <http://doi.org/10.1093/acprof:oso/9780195333060.001.0001>
- Data Asset Framework. (2009). *Implementation Guide*. [http://www.data-audit.eu/docs/DAF\\_Implementation\\_Guide.pdf](http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf)
- Davis, L. (2007). A Grand Unified Theory of Interdisciplinarity. *The Chronicle Review*, 53(40), 49.
- De Paul Obade, V. & Lal, R. (2013). Assessing land cover and soil quality by remote sensing and geographical information systems (GIS). *CATENA*, 104, 77–92. <http://dx.doi.org/10.1016/j.catena.2012.10.014>
- De Roure, D. (2014). The future of scholarly communications. *Insights*, 27(3), 233–238. <http://dx.doi.org/10.1629/2048-7754.171>
- De Solla Price, D. J. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Del Favero, M. Del Favero, M. (2004). Academic Disciplines: disciplines and the structure of higher education, discipline classification systems, discipline differences. <http://education.stateuniversity.com/pages/1723/Academic-Disciplines.html>
- DeLyser, D. & Sui, D. (2014). Crossing the qualitative-quantitative chasm III: Enduring methods, open geography, participatory research, and the fourth paradigm. *Progress in Human Geography*, 38(2), 294–307. <http://doi.org/10.1177/0309132513479291>
- DeLyser, D. & Sui, D. (2013). Crossing the qualitative-quantitative divide II: Inventive approaches to big data, mobile methods, and rhythm analysis. *Progress in Human Geography*, 37(2), 293–305. <http://doi.org/10.1177/0309132512444063>



- Denscombe, M. (2010). *The Good Research Guide: for small-scale social research projects* (4th ed.). Berkshire: Open University Press.
- Dexter, L. A. (2006). *Elite and Specialized Interviewing*. Colchester: ECPR.
- Diener, E. & Crandall, R. (1978). *Ethics in Social and Behavioural Research*. Chicago: University of Chicago Press.
- Digital Curation Centre. (2006). Curating Geospatial Data. In *DCC Briefing Papers: Introduction to Curation*. Edinburgh: Digital Curation Centre. <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/curating-geospatial-data>
- Digital Curation Centre. (2008). Digital curation 101- What is digital curation. <http://www.dcc.ac.uk/sites/default/files/documents/DC>
- Dogan, M. & Pahre, R. (1990). *Creative Marginality: Innovation at the intersections of social sciences*. Boulder, CO: Westview Press.
- Donald, J. (2002). *Learning to Think: Disciplinary Perspectives*. San Francisco: Jossey-Bass.
- Dutton, W. H. & Jeffreys, P. W. (Eds.). (2010). *Worldwide Research: Reshaping the Sciences and Humanities*. Cambridge, Mass.: MIT Press.
- Economic and Social Research Council (ESRC). (2013). *International Benchmarking Review of UK Human Geography*. <http://www.esrc.ac.uk/research/evaluation-impact/research-evaluation/international-benchmarking/uk-human-geography.aspx>
- EDINA & University of Edinburgh Data library. (2013). Research Data MANTRA [online course]. <http://datalib.edina.ac.uk/mantra/datamanagementplans.html>
- Edinburgh University Library. (2011). Edinburgh University Data Library Research Data Management Handbook. [http://www.docs.is.ed.ac.uk/docs/data-library/EUDL\\_RDM\\_Handbook.pdf](http://www.docs.is.ed.ac.uk/docs/data-library/EUDL_RDM_Handbook.pdf)
- Else, H. (2014). Research funders must club together to address grand challenges. <https://www.timeshighereducation.com/news/research-funders-must-club-together-to-address-grand-challenges/2017634.article>
- Elwood, S. & Cope, M. (2009). Introduction: Qualitative GIS: Forging mixed methods through representations, analytical innovations, and conceptual engagements. In *Qualitative GIS: A mixed methods approach* (pp. 1–13). London: Sage Publications.
- Engineering and Physical Sciences Research Council (EPSRC). (2016). *EPSRC policy framework on research data: Scope and benefits*. <https://www.epsrc.ac.uk/about/standards/researchdata/scope/>
- Engineering and Physical Sciences Research Council (EPSRC). (2011). *EPSRC policy framework on research data*. <http://www.epsrc.ac.uk/about/standards/researchdata/Pages/policyframework.aspx>
- European Commission. (2016). *European Cloud Initiative to give Europe a global lead in the data-driven economy*. [http://europa.eu/rapid/press-release\\_IP-16-1408\\_en.htm](http://europa.eu/rapid/press-release_IP-16-1408_en.htm)

- Eynden, V. Van Den & Bishop, L. (2014). *Sowing the seed: Incentives and motivations for sharing research data, a researcher's perspective*. [http://www.data-archive.ac.uk/media/492924/ke\\_report-incentives-for-sharing-research-data.pdf](http://www.data-archive.ac.uk/media/492924/ke_report-incentives-for-sharing-research-data.pdf)
- Fecher, B., Friesike, S. & Hebing, M. (2015). What drives academic data sharing? *PLoS ONE*, 10(2), 1–25. <http://doi.org/10.1371/journal.pone.0118053>
- Ferguson, R. (2003). Publication practices in physical and human geography: a comment on Nigel Thrift's "The future of geography." *Geoforum*, 34(1), 9–11. [http://doi.org/10.1016/S0016-7185\(02\)00082-9](http://doi.org/10.1016/S0016-7185(02)00082-9)
- Floridi, L. (2008). Data. In W. A. Darity (Ed.), *International Encyclopaedia of the Social Sciences*. Detroit: Macmillan.
- Frankfort-Nachmias, C. & Nachmias, D. (2000). *Research Methods in the Social Sciences* (6th ed.). New York: Worth Publishers.
- Freiman, L., Ward, C., Jones, S., Molloy, L. & Snow, K. (2010). *Incremental scoping study and implementation plan: a pilot project for supporting research data management. Project report*. [http://www.lib.cam.ac.uk/preservation/incremental/documents/Incremental\\_Scoping\\_Report\\_170910.pdf](http://www.lib.cam.ac.uk/preservation/incremental/documents/Incremental_Scoping_Report_170910.pdf)
- Frické, M. (2009). The knowledge pyramid: a critique of the DIKW. *Journal of Information Science*, 35(2), 131–142. <http://doi.org/10.1177/0165551508094050>
- Frodeman, R., Klein, J. T. & Mitcham, C. (Eds.). (2012). *The Oxford Handbook of Interdisciplinarity*. Oxford: Oxford University Press.
- Fry, J., Lockyer, S., Oppenheim, C., Houghton, J. & Rasmussen, B. (2009). *Identifying benefits arising from the curation and open sharing of research data produced by UK Higher Education and research institutes*. <http://repository.jisc.ac.uk/279/>
- Fry, J., Spezi, V., Proberts, S. & Creaser, C. (2015). Towards an understanding of the relationship between disciplinary research cultures and open access repository behaviours. *Journal of the Association for Information Science and Technology*, 14(4), 2710-2724. <http://doi.org/10.1002/asi.23621>
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies*. London: Sage Publications.
- Gilbert, D. (2014). *Disciplinary norms and expectations: Geography*. Royal Holloway University of London.
- Gillham, B. (2000). *The Research Interview*. London: Continuum.
- Gitelman, L. (Ed.). (2013). *Raw Data Is an Oxymoron*. Cambridge, Massachusetts: MIT Press.
- Glanzel, W. (2014). Analysis of co-authorship patterns at the individual level. *Transinformacao*, 26(3), 229–238. <http://doi.org/10.1590/0103-37862014000300001>

- Glänzel, W. (2012). Bibliometric methods for detecting and analysing emerging research topics. *El Profesional de La Información*, 21(1), 194–201. <http://eprints.rclis.org/16947/1/Bibliometric>
- Gomez, B. & Ill, J. P. J. (Eds.). (2010). *Research Methods in Geography. International Planning Studies* (Vol. 16). Blackwell. <http://doi.org/10.1080/13563475.2011.561066>
- Goodlad, S. (1979). What is an academic discipline? In R. Cox (Ed.), *Cooperation and Choice in Higher Education* (pp. 10–20). London: University of London Press.
- Goodwin, J. (2012). *SAGE Secondary Data Analysis*. London: Sage. <http://doi.org/10.4135/9781446268544>
- Grace, S. (2013). Results of University of East London survey on managing research data.
- Graham, M. & Shelton, T. (2013). Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography*, 3(3), 255–261. <http://doi.org/10.1177/2043820613513121>
- Graham, M. & Shelton, T. (2013). Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography*, 3(3), 255–261. <http://doi.org/10.1177/2043820613513121>
- Gregory, D., Johnston, R., Pratt, G., Watts, M. J. & Whatmore, S. (Eds.). (2009). *The Dictionary of Human Geography* (5th ed.). Chichester: Wiley-Blackwell.
- Gregory, K. J. (Ed.). (2005). *Physical Geography*. London: Sage.
- Gregory, K. J., Gurnell, A. M. & Petts, G. E. (2002). Restructuring physical geography. *Transactions of the Institute of British Geographers*, 27(2), 136–154. <http://doi.org/10.1111/1475-5661.00046>
- Gubrium, J. F. & Holstein, J. A. (2003). *Postmodern Interviewing*. Thousand Oaks, CA: Sage.
- Hakim, C. (2000). *Research Design: Successful Designs for Social and Economic Research*. Oxon: Routledge.
- Hall, T., Toms, P., Mcguinness, M., Parker, C. & Roberts, N. (2015). Where's the Geography department? The changing administrative place of Geography in UK higher education. *Area*, 47: 56–64. <http://doi.org/10.1111/area.12154>
- Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S. & King, C. J. (2010). *Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines*. University of California: The Centre for Studies in Higher Education. <http://escholarship.org/uc/item/15x7385g>
- Harries, G., Wilkinson, D., Price, L., Fairclough, R. & Thewall, M. (2004). Hyperlinks as a data source for science mapping. *Journal of Information Science*, 30(5), 436–447. <http://doi.org/10.1177/0165551504046736>
- Harrison, S., Massey, D., Richards, K., Magilligan, F. J., Thrift, N. & Bender, B. (2004). Thinking across the divide: Perspectives on the conversations between physical and human geography. *Area*, 36(4), 435–442. <http://doi.org/10.1111/j.0004-0894.2004.00243.x>

- Hedges, L. V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, 42(5), 443–455. <http://doi.org/10.1037/0003-066X.42.5.443>
- Herold, P. (2015). Data Sharing Among Ecology, Evolution, and Natural Resources Scientists: An Analysis of Selected Publications. *Journal of Librarianship and Scholarly Communication*, 3(2), eP1244. <http://doi.org/10.7710/2162-3309.1244>
- Hewson, C., Yule, P., Laurent, D. & Vogel, C. (2003). *Internet Research Methods: a Practical guide for the Social and Behavioural Sciences*. London: Sage Publications.
- Hey, T. & Trefethen, A. (2003). The Data Deluge: An e-Science Perspective. In *Grid Computing* (pp. 809–824). Chichester: John Wiley & Sons Ltd. <http://doi.org/10.1002/0470867167.ch36>
- Higgins, S. (2008). The DCC Curation Lifecycle Model. *International Journal of Digital Curation*, 3(1). <http://doi.org/10.2218/ijdc.v3i1.48>
- Higgins, S. (2012). The lifecycle of data management. In *Managing Research Data* (pp. 17–45). London: Facet.
- Hjørland, B. (2007). Information: Objective or subjective/situational? *Journal of the American Society for Information Science and Technology*, 58(10), 1448–1456. <http://doi.org/10.1002/asi.20620>
- Hochschild, J. L. (2009). Conducting Intensive Interviews and Elite Interviews. *Workshop on Interdisciplinary Standards for Systematic Qualitative Research*. Washington D.C.: National Science Foundation. [http://www.nsf.gov/sbe/ses/soc/ISSQR\\_workshop\\_rpt.pdf](http://www.nsf.gov/sbe/ses/soc/ISSQR_workshop_rpt.pdf)
- Hochschild, J. L. (2009). Conducting Intensive Interviews and Elite Interviews. In *Workshop on Interdisciplinary Standards for Systematic Qualitative Research*. [http://www.nsf.gov/sbe/ses/soc/ISSQR\\_workshop\\_rpt.pdf](http://www.nsf.gov/sbe/ses/soc/ISSQR_workshop_rpt.pdf)
- Holiday, A. (2007). *Doing Qualitative Research: Second Edition (Second)*. London: Sage.
- Holland, D. (2014). *Integrating Knowledge Through Interdisciplinary Research*. Oxon: Routledge.
- Holloway, S. L., Rice, S. P. & Valentine, G. (Eds.). (2003). *Key concepts in geography*. London: Sage.
- Holstein, J. & Gubrium, J. (2011). Animating interview narratives. In D. Silverman (Ed.), *Qualitative Research* (3rd ed.). London: Sage.
- Hood, W. & Wilson, C. (2001). The Literature of Bibliometrics, Scientometrics, and Informetrics. *Scientometrics*, 52(2), 291–314. <http://doi.org/10.1023/A:1017919924342>
- Horton, L. & (DCC), D. C. C. (2016). Overview of UK Institution RDM Policies. <http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies>
- Houghton, J. & Gruen, N. (2014). Open Research Data. Report to the Australian National Data Service (ANDS). [http://ands.org.au/\\_\\_data/assets/pdf\\_file/0019/393022/open-research-data-report.pdf](http://ands.org.au/__data/assets/pdf_file/0019/393022/open-research-data-report.pdf)

- Huang, H., Stvilia, B. & Jørgensen, C. (2012). Curation roles and perceived priorities for data quality dimensions and skills in genome curation work. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–3. <http://doi.org/10.1002/meet.14504901312>
- Illinois Informatics Institute, University of Illinois (2011). *What is informatics?* <https://rhino.ncsa.illinois.edu/display/infominor/Static+content+page>
- Ingram, C. (2016). *How and why you should manage your research data: a guide for researchers*. JISC. <https://www.jisc.ac.uk/guides/how-and-why-you-should-manage-your-research-data>
- International Geographical Union. (2013). <http://www.igu-online.org/>
- Jacobs, J. A. (2014). *In Defence of Disciplines: Interdisciplinarity and Specialization in the Research University*. Chicago: University of Chicago Press.
- Janet Boddy, Tim Neumann, Sean Jennings, Virginia Morrow, Priscilla Alderson, R. R. and W. G. (n.d.). Limits of confidentiality: professional and elite interviews. <http://www.ethicsguidebook.ac.uk/Limits-of-confidentiality-professional-and-elite-interviews-232>
- Jankowski, N. (2009). *E-Research: Transformation in Scholarly Practice*. New York: Routledge.
- Jennex, M. E. & Diego, S. (2013). A Revised Knowledge Pyramid. *International Journal of Knowledge Management*, 9(3), 19–30. <http://www.irma-international.org/viewtitle/99641/>
- JISC. (2011). Research data management training materials (RDMTrain). <http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmtrain.aspx>
- JISC. (2016). *Managing research data in your institution*. <https://www.jisc.ac.uk/guides/research-data-management>
- Jones, P. & Macdonald, N. (2007). Getting it wrong first time: Building an interdisciplinary research relationship. *Area*, 39(4), 490–498. <http://doi.org/10.1111/j.1475-4762.2007.00767.x>
- Jones, W. A. (2011). Variation among academic disciplines: an update on analytical frameworks and research. *Journal of the Professoriate*, 6(1), 9–27.
- Kahn, R. & Cannell, C. (1957). *The Dynamics of Interviewing*. New York: John Wiley.
- Kaiser, K. (2010). Protecting Respondent Confidentiality in Qualitative Research. *Qualitative Health Research* 19(11), 1632 – 1641. <http://doi.org/10.1177/1049732309350879>
- Katz, J. S. & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18. [http://doi.org/10.1016/S0048-7333\(96\)00917-1](http://doi.org/10.1016/S0048-7333(96)00917-1)
- Key Perspectives. (2010). *Data dimensions: disciplinary differences in research data sharing, reuse and long-term viability*. SCARP Synthesis Study. Digital Curation Centre. <http://www.dcc.ac.uk/projects/scarp>

- Kim, Y. (2012). Factors influencing STEM researchers' data sharing behaviours. *Proceedings of the American Society for Information Science and Technology*, 49(1).  
<https://www.asis.org/asist2012/proceedings/Submissions/348.pdf>
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography*, 3(3), 262–267. <http://doi.org/10.1177/2043820613513388>
- Kitchin, R. (2014). *The Data Revolution*. London: Sage.
- Kitchin, R. & Lauriault, T. P. (2015). Small data in the era of big data. *Geo Journal*, 80(4), 463–475.  
<http://doi.org/10.1007/s10708-014-9601-7>
- Klein, J. T. (1996). *Crossing Boundaries: knowledge, disciplinarity, and interdisciplinarity*. Charlottesville: University of Virginia.
- Klein, J. T. (1993). Blurring, Cracking, and Crossing: permeation and the fracturing of discipline. In E. Messer-Davidow, D. Sylvan & D. Shumway (Eds.), *Knowledges: Historical and Critical Studies of Disciplinarity* (pp. 185–211). Charlottesville: University of Virginia.
- Knight, G. (2013). *Research Data Management at LSHTM: web survey report*. London: London School of Hygiene and Tropical Medicine. <http://blogs.lshtm.ac.uk/rdmss/files/2013/04/LSHTM-RDM-Web-Survey-Report.pdf>
- Knobe, J. (2015). Philosophers are doing something different now: Quantitative data. *Cognition*, 135, 36–38. <http://doi.org/10.1016/j.cognition.2014.11.011>
- Kolb, D. A. (1981). Learning styles and disciplinary differences. In A. Chickering (Ed.), *The Modern American college*. San Francisco: Jossey-Bass.
- Kometa, S. (2011). *Iridium: research data management online requirement survey*. Newcastle: Newcastle University.
- Krishnan, A. (2009). *What are academic disciplines? Some observations on the disciplinarity vs. interdisciplinarity debate*. National Centre for Research Methods.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Kumar, R. (2011). *Research Methodology* (3rd ed.). London: Sage.
- Kvale, S. & Brinkmann, S. (2009). *Interviews: learning the craft of qualitative research interviewing*. Los Angeles: Sage.
- Kwan, M. P. (2002). Feminist Visualization: Re-Envisioning GIS as a Method in Feminist Geographic Research. *Annals of the Association of American Geographers*, 92(4), 645–661.  
<http://doi.org/10.2307/1515293>
- Kwan, M.-P. (2010). A Century of Method-Oriented Scholarship in the Annals. *Annals of the Association of American Geographers*, 100(5), 1060–1075.  
<http://doi.org/10.1080/00045608.2010.523341>

- Lämmerhirt, D. (2016). Briefing Paper: Disciplinary differences in opening research data. *PASTEUR4OA*, (June 2013), 1–8. <http://apo.org.au/node/63531>
- Lancaster, K. (2016). Confidentiality, anonymity and power relations in elite interviewing: conducting qualitative policy research in a politicised domain. *International Journal of Social Research Methodology*, 1–11. <http://doi.org/10.1080/13645579.2015.1123555>
- Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
- Lave, R. (2015). Introduction to special issue on critical physical geography. *Progress in Physical Geography*, 39(5), 571–575. <http://doi.org/10.1177/0309133315608006>
- Leedy, P. D. (2010). *Practical research: planning and design*. (J. E. Ormrod, Ed.). Boston, Mass: Pearson Education International.
- Lewis, M. J. (2010). Libraries and the management of research data. In Sue McKnight (Ed.), *Envisioning Future Academic Library Services* (pp. 145–168). London: Facet Publishing. <http://eprints.whiterose.ac.uk/11171/>
- Lindblom Ylänne, S., Trigwell, K., Nevgi, A. & Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education*, 31(3), 285–298. <http://doi.org/10.1080/03075070600680539>
- Lodahl, J. B. & Gordon, G. (1972). The Structure of Scientific Fields and the Functioning of University Graduate Departments. *American Sociological Review*, 37(1), 57–72. <http://doi.org/10.2307/2093493>
- Lord, P. & Macdonald, A. (2003). *e-Science Curation Report*. Twickenham: Digital Archiving Consultancy. [http://www.jisc.ac.uk/uploaded\\_documents/e-ScienceReportFinal.pdf](http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf)
- Lord, P., Macdonald, A., Lyon, L. & Giaretta, D. (2004). From Data Deluge to Data Curation. In *e-Science All Hands Meeting 2004*. <http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/150.pdf>
- Lyon, L. (2007). *Dealing with Data: Roles, Rights, Responsibilities and Relationships. Consultancy Report*. UKOLN. [http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing\\_with\\_data\\_report-final.pdf](http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dealing_with_data_report-final.pdf)
- Lyon, L., Rusbridge, C., Neilson, C. & Whyte, A. (2010). *DCC SCARP: disciplinary approaches to sharing, curation, reuse and preservation - final report*. <http://www.dcc.ac.uk/sites/default/files/documents/scarp/SCARP-FinalReport-Final-SENT.pdf>
- Marshall, C. & Rossman, G. B. (2006). *Designing Qualitative Research* (4th ed.). Thousand Oaks: Sage.
- Mather, P. M. & Koch, M. (2011). *Computer Processing of Remotely-Sensed Images* (4th ed.). Chichester, UK: John Wiley & Sons, Ltd. <http://doi.org/10.1002/9780470666517>
- Matthews, J. A. & Herbert, D. T. (2008). *Geography: a very short introduction*. Oxford: Oxford University Press.

- May, T. (2011). *Social Research: issues, methods and process*. Maidenhead: Open University Press.
- McGarva, G., Morris, S. & Janée, G. (2009). *Technology Watch Report: preserving geospatial data*. Digital Preservation Coalition. <http://www.dpconline.org/docman/technology-watch-reports/363-preserving-geospatial-data-by-guy-mcgarva-steve-morris-and-gred-greg-janee/file>
- Melin, G. & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, 36(3), 363–377. <http://doi.org/10.1007/BF02129600>
- Meyer, E. T., Bulger, M., Kyriakidou-Zacharoudiou, A., Power, L., Williams, P., Venters, W., Wyatt, S. (2011). *Collaborative yet Independent: Information practices in the physical sciences*. The RIN, the Institute of Physics (IOP), Institute of Physics Publishing (IOPP) and the Royal Astronomical Society (RAS). <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/physical-sciences-case-studies-use-and-discovery->
- Michener, Wi. K., Brunt, J. W., Helly, J. J., Kirchner, T. B. & Stafford, S. G. (1997). Nongeospatial metadata for the ecological sciences. *Ecological Applications*, 7(1), 330–342. <http://lits.bio.ic.ac.uk:8080/litsproject/Micheneretal1997.pdf>
- Miller, H. J. & Goodchild, M. F. (2014). Data-driven geography. *GeoJournal*, 80(4), 449–461. <http://doi.org/10.1007/s10708-014-9602-6>
- Monastersky, R. (2013). The library reboot: as scientific publishing moves to embrace open data, libraries and researchers are trying to keep up. *Nature*, 405, 430–432. [http://www.nature.com/polopoly\\_fs/1.12664!/menu/main/topColumns/topLeftColumn/pdf/495430a.pdf](http://www.nature.com/polopoly_fs/1.12664!/menu/main/topColumns/topLeftColumn/pdf/495430a.pdf)
- Moore, N. (2006). The Contexts of Context: Broadening Perspectives in the (Re)use of Qualitative Data. *Methodological Innovations Online*, 1(2), 21–32. <http://doi.org/10.4256/mio.2006.0009>
- Morse, J. M. & Coulehan, J. (2015). Maintaining Confidentiality in Qualitative Publications. *Qualitative Health Research*, 25(2), 151–152. <http://doi.org/10.1177/1049732314563489>
- Murillo, A. P., Thompson, C. A., Carver, N., Robertson, W. D., Greenberg, J. & Anderson, W. L. (2012). The data-at-risk initiative: Analyzing the current state of endangered scientific data. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–3. <http://doi.org/10.1002/meet.14504901346>
- Murray, A. T. (2010). Quantitative geography. *Journal of Regional Science*, 50(1), 143–163. <http://doi.org/10.1111/j.1467-9787.2009.00642.x>
- Myers, M. D. (2008). *Qualitative Research in Business & Management*. Los Angeles: Sage.
- Nassiri, S. & Worthington, B. (2012). *Research Data Assets Survey*. Hertfordshire: University of Hertfordshire. <http://research-data-toolkit.herts.ac.uk/2012/08/data-asset-survey-results/>
- National Science Foundation (NSF). (2005). *Cyberinfrastructure Vision for 21st Century Discovery*. Arlington: National Science Foundation. <http://www.nsf.gov/pubs/2005/nsb0540/>
- Nelson, B. (2009). Data sharing: Empty archives. *Nature*, 461(7261), 160–163. <http://doi.org/10.1038/461160a>



- Nielsen, H. J. & Hjørland, B. (2014). Curating research data: the potential roles of libraries and information professionals. *Journal of Documentation*, 70(2), 221–240.  
<http://doi.org/10.1108/JD-03-2013-0034>
- OECD. (2004). *Declaration on Access to Research Data from Public Funding*.  
<http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=157&Lang=en&Book=False>
- Open Data Charter. (2013). *G8 Open Data Charter*. UK Cabinet Office.  
<https://www.gov.uk/government/publications/open-data-charter>
- Open Data Institute. (2015). *Open Data Roadmap for the UK*. <https://theodi.org/roadmap-uk-2015>
- Open Data Strategy. (2014). *Open Data Strategy 2014-2016*. Department for Business Innovation and Skills.
- Open Exeter Project Team. (2012). *Open Exeter Project: summary findings of the open Exeter data asset framework survey*. Exeter: University of Exeter.
- Open Source Initiative. (1998). <http://opensource.org/>
- Organisation for Economic Co-operation and Development (OECD). (2015). *Making Open Science a Reality*.
- Organisation for Economic Co-operation and Development (OECD). (1972). *Interdisciplinarity: Problems of teaching and research in universities*. Paris: OECD.
- Orlikowski, W. J. & Yates, J. (1994). Genre Repertoire: The Structuring of Communicative Practices in Organizations. *Administrative Science Quarterly*, 39(4), 541–574.  
<http://doi.org/10.2307/2393771>
- Ostreng, W. (2007). Crossing Scientific Boundaries by Way of Disciplines. In W. Østreng (Ed.), *Complexity. Interdisciplinary Communications 2006/2007*. Oslo: Centre for Advanced Study at the Norwegian Academy of Science and Letters.
- Packer, M. (2011). *The Science of Qualitative Research*. Cambridge: Cambridge University Press.
- Pantin, C. F. A. (1968). *The Relations between the Sciences*. Cambridge: Cambridge University Press.
- Parker, J. (2002). A New Disciplinarity: Communities of knowledge, learning and practice. *Teaching in Higher Education*, 7(4), 373–386. <http://doi.org/10.1080/135625102760553883>
- Parsons, T., Grimshaw, S. & Williamson, L. (2013). *Research Data Management Survey*. Nottingham: University of Nottingham.
- Patel, M. (2010). *JISC Infrastructure for Integration in Structural Sciences Report*.  
<http://www.ukoln.ac.uk/projects/I2S2/documents/I2S2-WP1-D1.1-RR-Final-100707.pdf>
- Pennock, M. (2007). Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information. *Library & Archives Journal*, 1.  
[http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch\\_curation.pdf](http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf)

- Pepe, A., Mayernik, M., Borgman, C. L. & Van de Sompel, H. (2009). From artifacts to aggregations: Modelling scientific life cycles on the semantic Web. *Journal of the American Society for Information Science and Technology*. <http://doi.wiley.com/10.1002/asi.21263>
- Pickering, A. (1992). *Science as Practice and Culture*. Chicago: University of Chicago Press.
- Pienta, A. M., Alter, G. & Lyle, J. (2010). The enduring value of social science research: the use and reuse of primary research data. <http://deepblue.lib.umich.edu/handle/2027.42/78307>
- Podgórecki., A. (1997). *Higher faculties: a cross-national study of university culture*. Westport: Praeger.
- Polit, D. F. & Beck, C. T. (2010). Generalization in quantitative and qualitative research: myths and strategies. *International Journal of Nursing Studies*, 47(11), 1451–1458. <http://doi.org/10.1016/j.ijnurstu.2010.06.004>
- Poschen, P. (2007). *Developing the UK's e-infrastructure for Science and Innovation*. National e-Science Centre. <http://www.nesc.ac.uk/documents/OSI/report.pdf>
- Press, S. J. & Tanur, J. M. (2001). *The Subjectivity of Scientists and the Bayesian Approach*. Hoboken: John Wiley & Sons, Inc. <http://doi.org/10.1002/9781118150634>
- Prime Minister's Science Engineering and Innovation Council. (2006). *From Data to Wisdom: pathways to successful data management for Australian science*. <http://www.innovation.gov.au/science/PMSEIC/Documents/DataForScience.pdf>
- Pryor, G. (2009). Multi-scale data sharing in the life sciences: some lessons for policy makers. *The International Journal of Digital Curation*, 4(3), 71–82.
- Pryor, G. (2012). *Managing Research Data*. London: Facet.
- Pryor, G. & Donnelly, M. (2009). Skilling up to do data: whose role, whose responsibility, whose career? *International Journal of Digital Curation*, 4(2), 158–170. <http://www.ijdc.net/index.php/ijdc/article/view/126>
- Pryor, G., Jones, S. & Whyte, A. (Eds.). (2013). *Delivering Research Data Management Services*. London: Facet.
- RDMRose. (2013). Research Data Management for LIS. <http://rdmrose.group.shef.ac.uk/>
- Repko, A. F., Szostak, R. & Buchberger, M. P. (2013). *Introduction to Interdisciplinary Studies*. Thousand Oaks, CA: Sage.
- Research Assessment Exercise (RAE). (2008). *Geography and Environmental Studies: subject overview report*. <http://www.rae.ac.uk/panels/main/h/geog/>
- Research Councils UK (RCUK). (2005). *RCUK Position Statement on Access to Research Outputs*. <http://www.rcuk.ac.uk/documents/documents/2005statement.pdf>
- Research Councils UK (RCUK). (2011). *RCUK Common Principles on Data Policy*. <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>

- Research Councils UK (RCUK). (2016). *Concordat on Open Research Data*.  
<http://www.rcuk.ac.uk/media/news/160728/>
- Research Information Network (RIN). (2008). *To share or not to share: publication and quality assurance of research data outputs - Report commissioned by the Research Information Network*. <http://www.rin.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs>
- Research Information Network (RIN) & National Endowment for Science Technology and the Arts (NESTA). (2010). *Open to All? Case studies of openness in research*.  
[http://www.rin.ac.uk/system/files/attachments/NESTA-RIN\\_Open\\_Science\\_V01\\_0.pdf](http://www.rin.ac.uk/system/files/attachments/NESTA-RIN_Open_Science_V01_0.pdf)
- Richards, K. (2008). *Sub-panel 32: Geography and Environmental Studies Subject Overview Report*.  
<http://www.rae.ac.uk/pubs/2009/ov/>
- Richards, K., Batty, M., Edwards, K., Findlay, A., Foody, G., Frostick, L., Jones, K., Lee, R., Livingstone, D., Marsden, T., Petts, J., Philo, C., Simon, D., Smith, S. and Thomas, D. (2009), The nature of publishing and assessment in Geography and Environmental Studies: evidence from the Research Assessment Exercise 2008. *Area*, 41(3), 231–243. <http://doi.org/10.1111/j.1475-4762.2009.00908.x>
- Ritchie, J. & Lewis, J. (2014). *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. *Qualitative Research* (2nd ed.). London: Sage.  
<http://doi.org/10.4135/9781452230108>
- Rosdahl, C. B. & Kowalski, M. T. (2011). *Textbook of Basic Nursing* (10th ed.). Philadelphia: Lippincott Williams and Wilkins.
- Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), *Raw Data is an Oxymoron* (pp. 15–40). Cambridge, Mass.: MIT Press.
- Rowley, J. (2007). The Wisdom Hierarchy: representations of the DIKW hierarchy, 33(2), 163–180.  
<http://doi.org/10.1177/0165551506070706>
- Royal Geographical Society. (2013). What is Geography?  
<http://www.rgs.org/GeographyToday/What+is+geography.htm>
- Sands, A., Borgman, C. L., Wynholds, L. & Traweek, S. (2012). Follow the data: How astronomers use and reuse data. *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–3. <http://doi.org/10.1002/meet.14504901341>
- School of informatics University of Edinburgh. (2012). What is Informatics?  
<http://www.ed.ac.uk/schools-departments/informatics/about/vision/overview>
- Secor, A. J. (2010). Social Surveys, Interviews, and Focus Groups. In B. Gomez & J. P. Jones (Eds.), *Research Methods in Geography: A Critical Introduction*. Chichester: Wiley-Blackwell.
- Shearmur, R. (2015). Dazzled by data: Big Data, the census and urban geography. *Urban Geography*, 36(7), 965–968. <http://doi.org/10.1080/02723638.2015.1050922>
- Silverman, D. (2011). *Qualitative Research*. (3rd ed.). London: Sage.

- Spurling, N. (2012). Research strategies: “capital” and the goods of social science research. In P. Trowler, M. Saunders & V. Bamber (Eds.), *Tribes and Territories in the 21st century: Rethinking the significance of disciplines in higher education*. London: Routledge.
- Stephen, T. (2000). What are disciplines? And how is interdisciplinarity different? In P. Weingart & N. Stehr (Eds.), *Practising Interdisciplinarity* (pp. 46–65). Toronto: University of Toronto Press.
- Stevenson, J. (2003). Processing ARSF remote sensing data with open source GIS tools. <http://all-geo.org/volcan01010/2013/01/processing-arsf-remote-sensing-data-with-open-source-gis-tools/>
- Storer, N. (1966). *The Social System of Science*. New York: Holt, Rinehart and Winston.
- Storarr, R. D., Stokes, C. R. & Evans, D. J. a. (2013). A map of large Canadian eskers from Landsat satellite imagery. *Journal of Maps*, 9(3), 456–473. <http://doi.org/10.1080/17445647.2013.815591>
- Structural Reform Group. (2004). *Overview of the DDI Version 3.0 Conceptual Model*. Data Documentation Initiative. [http://opendatafoundation.org/ddi/srg/Papers/DDIModel\\_v\\_4.pdf](http://opendatafoundation.org/ddi/srg/Papers/DDIModel_v_4.pdf)
- Subramanyam, K. (1983). Bibliometric studies of research collaboration: A review. *Journal of Information Science*, 6(1), 33–38. <http://doi.org/10.1177/016555158300600105>
- Swan, A. (2008). Key concerns within the scholarly communication process. <http://www.jisc.ac.uk/whatwedo/topics/opentechnologies/openaccess/reports/keyconcerns>
- Tenopir, C., Sandusky, R. J., Allard, S. & Birch, B. (2013). Academic librarians and research data services: Preparation and attitudes. *IFLA Journal*, 39(1), 70–78. <http://doi.org/10.1177/0340035212473089>
- Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, et al. (2011) Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6(6): e21101. <http://doi.org/10.1371/journal.pone.0021101>
- The Royal Society. (2012). *Science as an open enterprise. The Royal Society Science Policy Centre report 02/12*. <http://royalsociety.org/policy/projects/science-public-enterprise/report/>
- Thelwall, M. (2004). *Link Analysis: an information science approach*. Amsterdam: Elsevier.
- Thelwall, M. (2010). Researching the public web. *eResearch Ethics*. <http://eresearch-ethics.org/position/researching-the-public-web/>
- Thrift, N. (2002). The future of geography. *Geoforum*, 33(3), 291–298. [http://doi.org/10.1016/S0016-7185\(02\)00019-2](http://doi.org/10.1016/S0016-7185(02)00019-2)
- Trowler, P., Saunders, M. & Bamber, V. (Eds.). (2012). *Tribes and Territories in the 21st century: rethinking the significance of disciplines in higher education*. London: Routledge.
- University for the Creative Arts. (2013). Research Governance. [http://www.ucreative.ac.uk/research\\_governance](http://www.ucreative.ac.uk/research_governance)

- University of Cambridge. (n.d.). Data Management Glossary.  
<http://www.lib.cam.ac.uk/dataman/glossary.html>
- University of East London. (2012). Research Data Management Policy for UEL.  
<http://www.uel.ac.uk/wwwmedia/services/library/lls/resources/rspresearchtools/Research-Data-Management-policy-for-UEL-FINAL.pdf>
- University of Hertfordshire. (2011). Data Management Policy.  
<http://sitem.herts.ac.uk/secreg/upr/pdf/IM12-Data>
- University of Illinois. (2012). Data curation.  
[http://www.lis.illinois.edu/academics/programs/specializations/data\\_curation](http://www.lis.illinois.edu/academics/programs/specializations/data_curation)
- University of Leicester. (2013). Data management support for researchers.  
<http://www2.le.ac.uk/services/research-data>
- University of Leicester. (2012). What would you do if you lost your research data tomorrow?  
[http://www2.le.ac.uk/services/research-data/documents/UoL\\_RDMPPoster2012\\_landscape.pdf](http://www2.le.ac.uk/services/research-data/documents/UoL_RDMPPoster2012_landscape.pdf)
- University of Oxford. (2012). Policy on the Management of Research Data and Records.
- University of Salford. (2016). Research Data Management Policy.  
[http://www.salford.ac.uk/\\_\\_data/assets/pdf\\_file/0009/674208/ResearchDataManagementPolicy.pdf](http://www.salford.ac.uk/__data/assets/pdf_file/0009/674208/ResearchDataManagementPolicy.pdf)
- University of Southampton. (2015). Research Data Management Policy.  
<http://www.calendar.soton.ac.uk/sectionIV/research-data-management.html>
- University of St Andrews. (2013). Research Data Management.  
<http://arts.standrews.ac.uk/rdm/index.shtml>
- University of St Andrews Library. (2016). Research Data Management Survey Results. <http://www.st-andrews.ac.uk/libraryblog/2016/08/research-data-management-survey-results/>
- Uribe, L. M. (2008). Research Data Management and Curation Services Framework. In *Scoping Digital Repository Services for Research Data Management project (Blog)*.  
<http://oxdrrc.blogspot.co.uk/2008/12/research-data-management-services.html>
- Van den Besselaar, P. & Heimeriks, G. (2006). Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. *Scientometrics*, 68(3), 377–393.  
<http://doi.org/10.1007/s11192-006-0118-9>
- Van den Eynden, V. & Bishop, L. (2014). Sowing the Seed: Incentives and motivations for sharing research data, a researchers' perspective. [http://repository.jisc.ac.uk/5662/1/KE\\_report-incentives-for-sharing-researchdata.pdf](http://repository.jisc.ac.uk/5662/1/KE_report-incentives-for-sharing-researchdata.pdf)
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L. & Horton, L. (2011). *Managing and Sharing Data: best practices for researchers*. Essex: UK Data Archive. <http://data-archive.ac.uk/media/2894/managingsharing.pdf>

- Vaughan, L. (2001). *Statistical Methods for the Information Professional*. Medford, New Jersey: Information Today.
- Velden, T., Haque, A. & Lagoze, C. (2010). A new approach to analysing patterns of collaboration in co-authorship networks: mesoscopic analysis and interpretation. *Scientometrics*, 85(1), 219–242. <http://doi.org/10.1007/s11192-010-0224-6>
- Wallis, J.C., Rolando, E., Borgman, C.L. (2013). If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology. *PLoS ONE* 8(7): e67332. <http://doi.org/10.1371/journal.pone.0067332>
- Webb, E. J., Campbell, D. T., Schwartz, R. D. & Sechrest, L. (1966). *Unobtrusive Measures: Nonreactive measures in the social sciences*. Chicago: Rand McNally.
- Webb, T. (2013). Open research data = win-win presented at Open Access to Research Data event. University of Sheffield. [http://www.sheffield.ac.uk/polopoly\\_fs/1.275148!/file/OpenDataWebb.pdf](http://www.sheffield.ac.uk/polopoly_fs/1.275148!/file/OpenDataWebb.pdf)
- Weingart, P. & Stehr, N. (Eds.). (2000). *Practising Interdisciplinarity*. Toronto: University of Toronto Press.
- Wessels, B., Finn, R., Sveinsdottir, T. & Wadhwa, K. (2017). *Open Data and the Knowledge Society*. Amsterdam University Press. <http://doi.org/10.5117/9789462980181>
- Whitley, R. (2000). *The Intellectual and Social Organization of the Sciences*. Oxford: Oxford University Press.
- Whyte, A. & Pryor, G. (2011). Open Science in Practice: Researcher Perspectives and Participation. *The International Journal of Digital Curation*, 6(1), 199–213. <http://ijdc.net/index.php/ijdc/article/view/173/241>
- Whyte, A. & Tedds, J. (2011). Making the Case for Research Data Management. In *DCC Briefing Papers*. Edinburgh: Digital Curation Centre. <http://www.dcc.ac.uk/resources/briefing-papers#sthash.UwNli2RI.dpuf>
- Williams, R. & Pryor, G. (2009). Patterns of information use and exchange: case studies of researchers in the life sciences. <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/patterns-information-use-and-exchange-case-studie>
- Willinsky, J. (2005). The unacknowledged convergence of open source, open access, and open science. *First Monday*, 10(8). <http://ojphi.org/ojs/index.php/fm/article/view/1265>
- Wilson, J. A. J., Jeffreys, P., Patrick, M., Rumsey, S. & Jefferies, N. (2012). *Results of the 2012 University of Oxford Research Data Management Survey*. <http://blogs.it.ox.ac.uk/damaro/2013/01/03/university-of-oxford-research-data-management-survey-2012-the-results/>
- Winn, J. (2012). *Data Assets Framework Survey Summary*. <http://orbital.blogs.lincoln.ac.uk/2012/04/30/data-assets-framework-survey-summary/>

- Wiseman, V., Conteh, L. & Matovu, F. (2005). Using diaries to collect data in resource-poor settings: Questions on design and implementation. *Health Policy and Planning*, 20(6), 394–404. <http://doi.org/10.1093/heapol/czi042>
- Wynholds, L. A., Wallis, J. C., Borgman, C. L., Sands, A. & Traweek, S. (2012). Data, data use, and scientific inquiry. In *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries - JCDL '12* (p. 19). New York: ACM Press. <http://doi.org/10.1145/2232817.2232822>
- Young, M. (2008). *Bringing Knowledge Back In: from social constructivism to social realism in the sociology of education*. London: Routledge.
- Zelenkauskaite, A. & Bucy, E. P. (2016). A scholarly divide: Social media, Big Data, and unattainable scholarship. *First Monday*, 21(5), 2 May 2016. <http://doi.org/10.5210/fm.v21i5.6358>
- Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), 479–493. <http://doi.org/10.1002/asi.20508>

# Appendices



## Appendix 1 Disciplinary coverage in studies of research data management

	RIN- to share or not to Share (2008)	RIN Case studies of openness (2010)	DCC SCARP (2010)	JISC RDM Training materials	JISC RDM Planning	JISC RDM Infrastructure	Data Sharing by Scientists: Practices and Perceptions (Tenopir et al, 2011)	Data sharing, small science and institutional repositories (Cragin et al, 2010)	Sowing the seed: incentives and motivations for sharing research data: researcher's perspectives (Van den Eynden & Bishop, 2014)
Sciences and Engineering	Astronomy	Astronomy	Atmospheric sciences	Health studies (DATUM for Health)	Physics and Astronomy- Gravitational Waves (MRD-GW)	Biomedical research (MaDAM)	Environmental sciences & ecology	agronomy and soil science	Sociology
	Chemical crystallography	Image bioinformatics	Engineering		Bio-imaging (DMBI)	Life sciences (ADMIRAL)	Social sciences	biochemistry	Humanities
	Climate science	Chemistry and chemical biology	Life science		Engineering research (ERIM)	Palaeoclimate and environment data (PEG-BOARD)	Biology	biology	Biomedicine
	Genomics	Clinical neuroimaging	Clinical psychiatry				Physical science/engineering	civil engineering	Plant genetics, taxonomy, ecology
	Social & public health sciences	Epidemiology research	e-health			Structural sciences (I2S2)	Atmospheric science	earth and atmospheric sciences	
	Systems biology	Natural language processing					Medicine	electrical engineering	

Social sciences, Arts and Humanities	Classics		Architecture	Archaeology and social anthropology (DATATrain)	ESRC research data (DMP-ESRC)	Humanities (SUDAMIH)	Other	food sciences	
	Rural economy and land use ( Interdisciplinary )		Social studies of human interaction	Artistic research (CAiRO)	History Archaeology Linguistics Onomastics and genetics (HALOGEN)  -Cross disciplines			horticulture/plant sciences	
				Psychological studies (DMTPsych)				anthropology	
								geology	
								kinesiology	
								speech and hearing	

## Appendix 2 Seed URLs used in the link analysis feasibility study

---

<http://www.geog.cam.ac.uk/people/>  
<http://www.geog.ox.ac.uk/staff/>  
<http://www.dur.ac.uk/geography/staff/>  
<http://www.geog.qmul.ac.uk/staff/>  
<http://www.bris.ac.uk/geography/people/>  
<http://www.geog.leeds.ac.uk/people/>  
<http://www.sheffield.ac.uk/geography/staff>  
<http://www.kcl.ac.uk/sspp/departments/geography/people/>  
<http://www.geog.ucl.ac.uk/about-the-department/people/>  
<http://www.lse.ac.uk/geographyAndEnvironment/whosWho/>  
<http://www.aber.ac.uk/en/iges/staff/>  
<http://www.rhul.ac.uk/geography/staffdirectory/home.aspx>  
<http://www.nottingham.ac.uk/geography/people>  
<http://www.sussex.ac.uk/geography/people>  
<http://geography.exeter.ac.uk/staff/>  
<http://staffprofiles.humanities.manchester.ac.uk/StaffList.aspx?ou=I4044>

## Appendix 3 Definitions of codes

---

Epistemic 1. 1 Research topic	Interviewees describe their research interests, main area of research
1.1.1 Multiple facets of individual researchers	Interviewees' involvement in more than one research field
1.2 Keywords	The best terms or keywords to describe research areas of interviewees
1.3 Researcher's background	Researchers' educational background and career history
1.4 Pure and applied research	Interviewees label their own research as pure or applied
1.5 Methods	Methods used in research
1.6 Creating data	Collect or generate research data
1.7 Processing data	Processing and analysing data
1.8 Keeping data	The way researchers keep or store their data
1.9. Disciplinary structure and boundaries -within geography	Disciplinary structures within human and physical geography
2. Social 2.1 Consensus	Consensus in the discipline
2.2 Collaboration	Collaboration in research, including who interviewees collaborate with, the nature of collaboration etc.
2.3 Scale of the field	The scale of interviewees' research field in the uk and internationally
2.4 Sharing data with collaborators	Sharing data between collaborators
2.4.1 Sharing data with collaborators- problems	Problems of sharing data between collaborators
2.5 Data sharing practices	Sharing data beyond collaborators
2.5.1 Climategate	The climategate scandal in 2009
2.5.2 Issues in archiving data	Issues in depositing data in data archives

2.5.3 Journal data policy	Journals' data deposit requirements or expectations
2.6 Data requests	Data requests from other researchers or organisations
2.7 Views on data sharing	Researchers' views on data sharing
2.7.1 Data misuse	Misuse of research data
2.7.2 Data citation	Data citation
2.8 Data sharing in other countries or communities	Culture of data sharing in other countries or communities
2.9 Using or reusing data	Using or reusing data created from other research
3. Organisational / external? 3.1 Research groups	Research groups
3.2 Institutional support	Institutional support in data management
3.3. Research funders	Research funders' policies, funders' preferences, interviewees' funding sources etc.
4. Interdisciplinary, multidisciplinary	Interdisciplinary/ multi-disciplinary research

## Appendix 4 Revised code structure in NVivo

---

### 1. Epistemic

- 1.1 Research topic
  - 1.1.1 Multiple facets of individual researchers
- 1.2 Keywords
- 1.3 Researcher's background
- 1.4 Pure and applied research
- 1.5 Methods
- 1.6 Creating data
- 1.7 Processing data
- 1.8 Keeping data
- 1.9. Disciplinary structure and boundaries

### 2. Social

- 2.1 Consensus
- 2.2 Collaboration
- 2.3 Scale of the field
- 2.4 Sharing data with collaborators
  - 2.4.1 Sharing data with collaborators- problems
- 2.5 Data sharing practices
  - 2.5.1 Climategate
  - 2.5.2 Issues in archiving data
  - 2.5.3 Journal data policy
- 2.6 Data requests
- 2.7 Views on data sharing
  - 2.7.1 Data misuse
  - 2.7.2 Data citation
  - 2.7.3 Culture of data sharing in other countries or communities
- 2.8 Using or reusing data
- 2.9 Interdisciplinary, multidisciplinary

### 3. Organisational / external

- 3.1 Research groups
- 3.2 Institutional support
- 3.3. Research funders
  - 3.3.1 REF



### Information Sheet

#### The nature of Geography and its research practices

##### Research aim

The aim of the research is to develop a better understanding of research practices in geography and how they vary across sub-disciplines - and the impact of this on data management and sharing practices.

##### Potential contributions

The study will enhance the conceptual theory of disciplinarity through detailed analysis of changes around research data in geography and its sub-disciplines, e.g. types of data, scale of data. In terms of practical contributions, the research will help institutions and funders to inform the development of data tools and technologies to support research and also to better support effective data management through policies and services.

This research will also provide a better understanding of the complexity of research data use and sharing, and aims to address the challenge of improving research data management. It will investigate how data practices are grounded in the underlying nature of disciplines/sub-disciplines. It will also develop insights needed to shape/ understand the current and emerging data environment.

##### Interview

Interviews are an integral part of the study. The primary purpose of the interviews is to understand researchers' research practices, especially their understandings and attitudes toward sharing data. The interview will also investigate the impact of current changes in research environment and how the discipline is changing as a result.

Topics to be addressed in the interview:

1. The nature of geography as a discipline (e.g. research topics, methods, collaboration)
2. Research practices of the interviewee (with a main focus on data use and sharing)
3. The impact of current changes in data environment: (1) Big data/ 3V's (2) Funders' policy (3) Open data (4) Demand for interdisciplinarity & collaboration
4. The Interviewee's views on how geography is changing (the future of the discipline)

The interview will last for about 45-60 minutes. The interview will be audio recorded. All information obtained from the interview will be treated confidentially. Data will be anonymised. No identifying information will be retained.

The results of this study will be included in my PhD thesis which will be publicly available. A summary of the results will also be available by contacting the primary investigator.

After you have read this information and asked any questions you may have, you will be asked to complete an Informed Consent Form. However if at any time, before, during or after the sessions you wish to withdraw from the study please just contact the main investigator. You can withdraw at any time, for any reason and you will not be asked to explain your reasons for withdrawing.

- **Contact information**

Winnie Tam, Centre for Information Management, Loughborough University, Loughborough LE11 3TU

Email: [w.w.t.tam@lboro.ac.uk](mailto:w.w.t.tam@lboro.ac.uk)

*Contact information of supervisors:*

Jenny Fry, Department of English and Drama, Loughborough University, Loughborough LE11 3TU

Email: [j.fry@lboro.ac.uk](mailto:j.fry@lboro.ac.uk)

Steve Proberts, Department of Computer Science, Loughborough University, Loughborough LE11 3TU

Email: [s.g.proberts@lboro.ac.uk](mailto:s.g.proberts@lboro.ac.uk)

Claire Creaser, Centre for Information Management, Loughborough University, Loughborough LE11 3TU

Email: [c.creaser@lboro.ac.uk](mailto:c.creaser@lboro.ac.uk)

If you are not happy with how the research was conducted, please contact Mrs Zoe Stockdale, the Secretary for the University's Ethics Approvals (Human Participants) Sub-Committee:

Mrs Z Stockdale, Research Office, Rutland Building, Loughborough University, Epinal Way, Loughborough, LE11 3TU. Tel: 01509 222423. Email: [Z.C.Stockdale@lboro.ac.uk](mailto:Z.C.Stockdale@lboro.ac.uk)

The University also has a policy relating to Research Misconduct and Whistle Blowing which is available online at [http://www.lboro.ac.uk/admin/committees/ethical/Whistleblowing\(2\).htm](http://www.lboro.ac.uk/admin/committees/ethical/Whistleblowing(2).htm)



**The nature of Geography and its research practices**

**INFORMED CONSENT FORM**

The purpose and details of this study have been explained to me. I understand that this study is designed to further scientific knowledge and that all procedures have been approved by the Loughborough University Ethics Approvals (Human Participants) Sub-Committee.

I have read and understood the information sheet and this consent form.

I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in the study.

I understand that I have the right to withdraw from this study at any stage for any reason, and that I will not be required to explain my reasons for withdrawing.

I understand that all the information I provide will be treated in strict confidence and will be kept anonymous and confidential to the researchers unless (under the statutory obligations of the agencies which the researchers are working with), it is judged that confidentiality will have to be breached for the safety of the participant or others.

I agree to participate in this study.

Your name \_\_\_\_\_

Your signature \_\_\_\_\_

Signature of investigator \_\_\_\_\_

Date \_\_\_\_\_

## Appendix 7 Sub-fields of Geography

	Biogeography	Cultural geography	Development geography	Economic geography	Environmental geography	Historical geography	Political geography	Population geography	Rural geography	Social geography	Urban geography
Cloke et al (2013)	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Matthews & Herbert, 2008		✓	✓	✓		✓	✓	✓	✓	✓	✓

Sub-fields of human geography identified in the literature.

	Geomorphology	Geocryology	Hydrology	Soil geography (sometime regards as part of biogeography)	Biogeography	Climatology
Matthews & Herbert, (2008)	✓	✓	✓	✓	✓	✓
Gregory, 2005	✓	✓ (Quaternary environmental change)	✓		✓	✓

Subfields of physical geography identified in the literature

## Appendix 8 Invitation Letter

---

Dear Prof. X,

### **The nature of Geography and its research practices**

I am a PhD student in information management at Loughborough University, undertaking research to develop a better understanding of research practices in geography and how they vary across sub-disciplines.

In my research, I am examining how geographers undertake their own research, e.g. in terms of their object of research, research methods and techniques, and the type of data they produce and use. My research aims to identify the cultural characteristics of sub-disciplines within geography and the ways in which research practices might be evolving in response to external factors, such as institutional and funder policies around sharing research data.

I have already undertaken some desk based research, including identifying institutional representations of geography on the Web and an analysis of collaboration patterns based on selected published outputs, however a further important element of my research is to carry out interviews with a number of senior researchers from across a range of sub-disciplines within Geography. Given your research interests in ..... I am very keen to interview you about the nature of your research.

Would it be possible to interview you in the near future? The interview will be at a time and place that is most convenient for you. (If you have a suitable time, please reply to this email and I will get back to you with a confirmation.) The interviews are primarily being conducted face-to-face, but we could arrange an interview via Skype if you prefer. I anticipate the interview lasting around 60 minutes. If it would be helpful, I can provide you with an information sheet that explains my research in more detail, and a list of the topic areas to be covered prior to the interview.

I look forward to hearing from you.

Yours sincerely,  
Winnie Tam

Centre for Information Management  
Loughborough University

## Appendix 9 Profile of Interview Participants

---

	Title	Physical / Human geography	Gender	Institutional Ranking	Interview Duration
1	Professor	Physical	M	Top	46 mins
2	Professor	Physical	M	Top	1 hr
3	Professor	Physical	M	Top	1 hr 3 mins
4	Senior lecturer	Physical	M	Top	45 mins
5	Senior lecturer	Human	F	Top	1 hr 31 mins
6	lecturer	Human	M	Top	1 hr 25 mins
7	research fellow	Human	M	Top	55 mins
8	Professor	Human	M	Top	1 hr 24 mins
9	Professor	Human	M	Top	51 mins
10	Professor	Physical	M	Middle	1 hr 09 mins
11	Professor	Physical	M	Middle	1 hr 07 mins
12	Professor	Physical	M	Middle	52 mins
13	Professor	Physical	M	Middle	1hr 7 mins
14	Lecturer	Human	M	Middle	43 mins
15	Professor	Human	M	Middle	1 hr 42 mins
16	Senior lecturer	Human	M	Middle	50 mins
17	Senior lecturer	Human	F	Middle	1 hr 42 mins
18	Professor	Human	M	Middle	43 mins
19	Senior Lecturer	Human	M	Middle	45 mins
20	Professor	Physical	M	Low	1 hr 06 mins
21	Senior lecturer	Physical	M	Low	1hr 11 mins
22	Lecturer	Physical	F	Low	1hr 17 mins
23	Professor	Human	F	Low	1 hr 08 mins

## Appendix 10 Interview guide

---

### Interview Guide

Date:

Thank you for agreeing to take part.

Have you read the information sheet?

Do you have any further questions?

So my research is about the nature of geography as a discipline and particularly issues around sharing data.

Is it ok if I record the interview?

Could you please sign the form?

Please let me know if you have any questions or you want to stop the interview at any point.

#### Research field

**1. Could you briefly outline your main area of research?**

-

**2. What are the best terms/ keywords to describe this area of research/specialisms?  
(Does this area belong to.. [social and cultural geographies?])**

**3. Is this research mostly theoretical or applied or both?**

- primarily theoretical / applied

**4. I am trying to find out the scale of the field.**

**Are there many people working in this area in the UK at the moment or are there just a few experts on this topic?**

**How about at an international level?**

**5. Would you describe it as an interdisciplinary field? (How interdisciplinary is this field)- Why?**

- Does the field borrow concepts/ methods from other disciplines?

~~6. Do people tend to agree on what the key methods and questions are in this area?~~

To what extent is there consensus in the field about the key research questions, approaches and methods?  
In other words, when you present at a conference or write a journal article how much effort do you need to make in order to justify your approach or methods?

7. Do you often do collaborative research with other researchers within geography or more widely from other disciplines, or both? [within the department/ UK/ international ?]

Why? (Is that a common practice?) (What sorts of disciplines are they from?)

8. I have seen your website you are in [research group].

How would you describe the importance of research group to you? (How does it affect your research?)

9. What are the key audiences of your research? (e.g. researchers in your field, other research field, the public, external organisation)

I'd now like to move on to discuss the nature of data and your data practices.

#### The nature of data and data practices

(10a) Could you briefly explain the research methods you usually use and the kinds of data you generate/ create for your research?

( primary data)

(11b) How do you process your data? (e.g. where do you get the data, what tools are being use to analyse/ process them...)

-

**11. (If the researcher do collaborative research)**

When you collaborate with other researchers on a research project, how do you share your data with collaborators? e.g. what tools do you use? , what kinds of data do you share?

Does it work well?

**12. Are you aware of any journals in your field require you (or provide you an option) to link your datasets with your publication?**

**13. After publication of journal articles/ reports or completion of the research project, what happens to the data?**

- Is it stored? Where? How long?

- Do you keep it for further research?

- Is it made it openly available?

(Do you re-use your own data in other research projects/ teaching? Why?)

The next topic I would like to discuss is data sharing (i.e. sharing your data with others).

**Data sharing**

**14. What do you think about the idea of sharing your own data beyond your research team?**

- *The willingness of sharing data*
- *>Competition in the field*
- *The value/ importance of data in their field*
- *Provenance*
- copyright
- authorship

**14(a) what types of data are you willing to share? (raw data)**

**15. Under what conditions are you willing to share your own data? ( e.g. you will share your data if.... your**

fundlers require you to share/ the person is from the same research field/ your publication are published )

Would you like to control who access it? (Is it something you have done?)  
(How about the idea of publishing data as a formal publication?)

- awareness of institutional RDM policy
- *Conditions to share your data e.g. attributions*
- *Issues/ barriers*

16. (Similar to Q. 13) Do you ever put your data in a repository? Where? Why/ Why not? Any difficulty?

Have you ever receive any request for you research data? ( e.g photos, transcripts, coding)

17. What would motivate you to share your data? ( e.g. funder mandates/ policies, data citation, disciplinary repositories, increasing research impact )

Another important topic I would like your views on is data re-use (using other researchers' data).

#### Data re-use

18. Do you use data generated / collected by other researchers?

What are that data? How do [or Why don't] you re-use those data?

(Is it a common practice in your field?)

- National (international) data sets e.g. meterological data, census data?
- are those data freely available?

- *what is the purposes of re-using such data?*

-*Experience of re-using data in the field*

19. Where do you get the data? Is it difficult to identify the right sort of dataset?

(Why/ What would improve in the process)

- *Data accessibility/ discovery*

20. What are the issues around re-using such data?

(Are there any issues of integrating your own data with other people's datasets? What are the issues? Data



quality?)

**21. (What are your criteria for re-using data? e.g. research area/ integrity)**

**Would you consider reusing data from other disciplines/ different types of data (e.g. [survey data])? Why?**

This is the last section. I want you to say something about the big trends affecting the discipline today.

According to the literature, these are some big trends affecting the discipline and the research data environment.

Could you rank them in terms of how relevant to you and your research?

How has it affected your research?

How is it changing research in human geography?

- **Technology: e.g. increasing scale of digital data, big data**
- **Research funders' policies (Were you aware of BA / ESRC's data policy?)**
- **Open data to public**
- **Demand for interdisciplinarity**

# Appendix 11 Ethics Checklist

**Ethics Approvals (Human Participants) Sub-Committee**



## Ethical Clearance Checklist

Has the Investigator read the 'Guidance for completion of Ethical Clearance Checklist' before starting this form?	Yes
---	-----

### Project Details

1. Project Title: The disciplinary shaping of research data management practices

### Applicant(s) Details

2. Name of Applicant 1: Winnie Tam	10. Name of Applicant 2: Click here to enter text JENNY FRY / STEPHEN PROBETS
3. Status: PGR student	11. Status: Choose an item SUPERVISOR / SUPERVISOR
4. School/Department: Centre for Information Management	12. School/Department: Click here to enter text. ENGLISH & DRAMA / COMPUTER SCIENCE
5. Programme (if applicable):	13. Programme (if applicable): Click here to enter text. N/A
6. Email address: w.w.t.tam@lboro.ac.uk	14. Email address: Click here to enter text. J.FRY@LBORO.AC.UK S.G.PROBETS@LBORO.AC.UK.
7a. Contact address: Apartment 102, West One Space, 8 Broomhall Street, Sheffield S3 7SY	15a. Contact address: Click here to enter text. ENGLISH & DRAMA, L.U.
7b. Telephone number: 0755 278 8487	15b. Telephone number: Click here to enter text. 223074
8. Supervisor: No	16. Supervisor: Choose an item N/A
9. Responsible Investigator: No	17. Responsible Investigator: Choose an item YES

### Participants

#### Positions of Authority

18. Are researchers in a position of direct authority with regard to participants (e.g. academic staff using student participants, sports coaches using his/her athletes in training)?	No
--	----

### Vulnerable groups

19. Will participants be knowingly recruited from one or more of the following vulnerable groups?	
Children under 18 years of age	No
Persons incapable of making an informed decision for themselves	No
Pregnant women	No
Prisoners/Detained persons	No
Other vulnerable group Please specify: Click here to enter text	No
<b>If you have selected No to all of Question 19, please go to Question 23.</b>	
20. Will participants be chaperoned by more than one investigator at all times?	Choose an item
21. Will at least one investigator of the same sex as the participant(s) be present throughout the investigation?	Choose an item
22. Will participants be visited at home?	Choose an item

### Researcher Safety

23. Will the researcher be alone with participants at any time?	Yes
<b>If Yes, please answer the following questions:</b>	
23a. Will the researcher inform anyone else of when they will be alone with participants?	Yes
23b. Has the researcher read the 'guidelines for lone working' and will abide by the recommendations within?	Yes

### Methodology and Procedures

24. Please indicate whether the proposed study:

Involves taking bodily samples (please refer to published guidelines)	No
Involves using samples previously collected with consent for further research	No
Involves procedures which are likely to cause physical, psychological, social or emotional distress to participants	No
Is designed to be challenging physically or psychologically in any way (includes any study involving physical exercise)	No
Exposes participants to risks or distress greater than those encountered in their normal lifestyle	No
Involves collection of body secretions by invasive methods	No
Prescribes intake of compounds additional to daily diet or other dietary manipulation/supplementation	No

Involves pharmaceutical drugs	No
Involves use of radiation	No
Involves use of hazardous materials	No
Assists/alters the process of conception in any way	No
Involves methods of contraception	No
Involves genetic engineering	No

Involves testing new equipment	No
--------------------------------	----

**Observation/Recording**

25a. Does the study involve observation and/or recording of participants?	Yes
If Yes:	
25b. Will those being observed and/or recorded be informed that the observation and/or recording will take place?	Yes

**Consent and Deception**

26. Will participants give informed consent freely?	Yes
---	-----

**Informed consent**

27. Will participants be fully informed of the objectives of the study and all details disclosed (preferably at the start of the study but, where this would interfere with the study, at the end)?	Yes
28. Will participants be fully informed of the use of the data collected (including, where applicable, any intellectual property arising from the research)?	Yes

29. For children under the age of 18 or participants who are incapable of making an informed decision for themselves:	
a. Will consent be obtained (either in writing or by some other means)?	N/A
b. Will consent be obtained from parents or other suitable person?	N/A
c. Will they be informed that they have the right to withdraw regardless of parental/guardian consent?	N/A
d. For studies conducted in schools, will approval be gained in advance from the Head-teacher and/or the Director of Education of the appropriate Local Education Authority?	N/A
e. For detained persons, members of the armed forces, employees, students and other persons judged to be under duress, will care be taken over gaining freely informed consent?	N/A

### Deception

30. Does the study involve deception of participants (i.e. withholding of information or the misleading of participants) which could potentially harm or exploit participants?	No
If Yes:	
31. Is deception an unavoidable part of the study?	Choose an item
32. Will participants be de-briefed and the true object of the research revealed at the earliest stage upon completion of the study?	Choose an item
33. Has consideration been given on the way that participants will react to the withholding of information or deliberate deception?	Choose an item

### Withdrawal

34. Will participants be informed of their right to withdraw from the investigation at any time and to require their own data to be destroyed?	Yes
--	-----

### Storage of Data and Confidentiality

35. Will all information on participants be treated as confidential and not identifiable unless agreed otherwise in advance, and subject to the requirements of law?	Yes
36. Will storage of data comply with the Data Protection Act 1998?	Yes
37. Will any video/audio recording of participants be kept in a secure place and not released for any use by third parties?	Yes
38. Will video/audio recordings be destroyed within ten years of the completion of the investigation?	Yes
39. Will full details regarding the storage and disposal of any human tissue samples be communicated to the participants?	N/A
40. Will research involve the sharing of data or confidential information beyond the initial consent given?	No
41. Will the research involve administrative or secure data that requires permission from the appropriate authorities before use?	No

### Incentives

42. Will incentives be offered to the investigator to conduct the study?	No
43. Will incentives be offered to potential participants as an inducement to participate in the study?	No

### Work Outside of the United Kingdom

44. Is your research being conducted outside of the United Kingdom?	Yes
<b>If Yes:</b>	
45. Has a risk assessment been carried out to ensure the safety of the researcher whilst working outside of the United Kingdom?	Yes
46. Have you considered the appropriateness of your research in the country you are travelling to?	Yes
47. Is there an increased risk to yourself or the participants in your research study?	No
48. Have you obtained any necessary ethical permission needed in the country you are travelling to?	Yes

### Information and Declarations

#### Checklist Application Only:

If you have completed the checklist to the best of your knowledge, and not selected any answers marked with an \* or †, your investigation is deemed to conform with the ethical checkpoints. Please sign the declaration and lodge the completed checklist with your Head of Department/School or his/her nominee.

#### Checklist with Additional Information to the Secretary:

If you have completed the checklist and have only selected answers which require additional information to be submitted with the checklist (indicated by a †), please ensure that all the information is provided in detail below and send this signed checklist to the Secretary of the Sub-Committee.

#### Checklist with Generic Protocols Included:

If you have completed the checklist and you have selected one or more answers in which you wish to use a Generic Protocol (indicated by #), please include the Generic Protocol reference number in the space below, along with a brief summary of how it will be used. Please ensure you are on the list of approved investigators for the Generic Protocol before including it on the checklist. The completed checklist should be lodged with your Head of Department/School or his/her nominee.

#### Full Application needed:

If on completion of the checklist you have selected one or more answers which require the submission of a full proposal (indicated by a \*), please download the relevant form from the Sub-Committee's web page. **A signed copy of this Checklist should accompany the full submission to the Sub-Committee.**

**Space for Information on Generic Proposals and/or Additional Information as requested:**

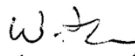
**For completion by Supervisor**

Please tick the appropriate boxes. The study should not begin until all boxes are ticked.

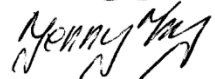
- The student has read the University's Code of Practice on investigations involving human participants
- The topic merits further research
- The student has the skills to carry out the research or are being trained in the requires skills by the Supervisor
- The participant information sheet or leaflet is appropriate
- The procedures for recruitment and obtaining informed consent are appropriate

Comments from supervisor:

Signature of Applicant: Winnie Tam



Signature of Supervisor (if applicable):



Signature of Head of School/Department or his/her nominee:



Date:

26/10/13