Digital technologies in the research process: lessons from the digital research community in the UK[1]

Dr Panayiota Tsatsou[2]

This paper reports on a qualitative study of the employment of digital tools, resources and services by social researchers in the UK and has a twofold aim. First, it presents the employment of digital means of research work from the stage of designing the research through to data collection and dissemination of the research results. Second, it assesses the complexity and multiplicity of digital tools, resources and services used in research as well as the complexity and range of such usage, also providing explanations as to why researchers in different disciplines use in different ways and for different purposes digital technologies of various ranges and degrees of complexity. The paper concludes that there are certain commonalities and differences in researchers' practices with digital technologies and that such practices are largely driven by researchers' expertise combined with associated disciplinary traditions and etiquette.

Keywords: complexity, digital research, humanities, multiplicity, research process, social science, UK

1. Introduction

Digital research is a rapidly growing area of development, deliberation and reflection. Digital research, Internet research, online research, e-research and e-science are concepts often used

[2] Department of Media and Communication, University of Leicester, Bankfield House, 132 New Walk, Leicester LE1 7JA, United Kingdom

interchangeably, and, although they are not identical, they all suggest the fast-developing and highly transformative role that information, communication and networking technologies play in the conduct of scientific study and research.

According to the Association of Internet Researchers, Internet technologies, tools and services comprise a 'social phenomenon, a tool, and also a (field) site for research. Depending on the role the internet plays in the research project or how it is conceptualized by the researcher, different epistemological, logistical and ethical considerations will come into play' (Markham and Buchanan, 2012, p. 3). From this one can infer that digital research, more broadly, involves the use of digital technologies, tools and services as objects of research (e.g., research into blogs, social networking sites, virtual worlds, virtual communities and instant messaging spaces), as tools for the creation of innovative methodological practices (e.g., hardware or software for devising, designing and executing methods of research) and as the actual sphere wherein research is positioned and from which researchers can draw research material and data (e.g., online datasets and repositories, search engines, data aggregators and automated means of data scraping). Digital technologies can serve as objects, tools and venues of research simultaneously (e.g., research into the affordances, content and users of online social networking sites), influencing research design, data collection and data analysis as well (Buchanan and Zimmer, 2012).

Also, digital research often suggests the collaboration of social, computer and web scientists, with knowledge elements from various disciplines being combined, influencing one another and boosting niche spaces for new knowledge networks and novel fields of study (e.g., artificial intelligence). This has led to new trends in research models (e.g., computational social science) and data (e.g., big data), as well as to new practices of collaboration that involve technology experts, funders, creative practitioners, industry actors and technology users. The multiplicity of actors, the dynamic role of digital technologies in

and for research, and the interdisciplinary and often cross-institutional nature of research collaborations portray a complicated and continuously shifting set of power relationships, dynamics, constraints, possibilities and synergies in digital research (Tsatsou, 2014, p. 166).

This paper draws from an EPSRC-funded qualitative study of the employment of digital tools, resources and services by social researchers in the UK, from the stage of designing the research through to data collection and the dissemination of results. The study aimed to map out the claimed, actual and potential role of digital technologies and to offer a critical assessment of the existing and potential innovation pathways signalled by the employment of digital technologies in social research in the UK. Specifically, it examined ongoing and prospective patterns of use of digital technologies in research contexts and shed light on associated skills and capacity challenges (Tsatsou, 2015). This paper presents and critically discusses the study's insights into the main patterns of use of digital technologies by social researchers in the UK and has a twofold aim:

- first, to present the stage(s) of the research process at which digital means of work are employed: the emphasis is on whether social researchers in the UK use digital technologies to design research, for data-gathering purposes, to analyse collected data or for dissemination, impact and knowledge transfer activities;

- second, to assess the complexity and multiplicity of digital tools, resources and services used in research as well as the complexity and range of usages accordingly, providing some explanation as to why researchers in different disciplines use in different ways and for different purposes digital technologies of various ranges and degrees of complexity.

In what follows, I offer a critical discussion of debates concerning the broader digital research domain. This is followed by an overview of the methodology of the study that this paper draws upon. This leads to a findings section, which addresses the twofold aim of the paper. The paper closes with some concluding remarks as well as reflections on the lessons to

be drawn from this study by digital research scholars and experts both within and outside the UK.

2. Digital research: novelty tied to messiness

Digital research developments are rapid. The unprecedented rate at which digital technologies are both diffusing through society and developing new research capacities seriously challenges the questions, phenomena and objects of research, as well as the ways in which researchers conduct research, thus creating a 'messy research landscape' (Karpf, 2012, p. 645). However, messiness goes hand in hand with novelty in digital research.

2.1. The 'digital as object of research

Researchers revisit existing research questions and generate new and often novel research questions in order to capture how existing or emerging phenomena take place in complex and rapidly evolving digital contexts. Initial considerations of the Internet as an object of research put forward the premise that the Internet can both expand existing research interests and yield new themes and areas of investigation (Costigan, 1999; Sterne, 2005). Costigan (1999, pp. xviii–xix) remarked that this can be done in two ways: first, through engaging in the search for, retrieval and analysis of vast information databases online; and second, by analysing unique communication and interaction phenomena online. I would add a third pathway, that of the study of the relationship of ordinary digital users with socio-technically founded affordances of digital technologies and the implications for existing and new communication-related phenomena. In the remainder of this section I discuss some examples, discourses and debates that shed light on digital technology as an object of research.

A first example is the study of hyperlinks, namely (hyper)link studies (Ackland and Gibson, 2013; Chang, Himelboim and Dong, 2009; De Maeyer, 2013; Park and Thelwall, 2003; Shumate and Lipp, 2008). De Maeyer (2013) suggests that (hyper)link studies are

present in various social science disciplines and can be split into two categories: the study of hyperlink networks and their properties so as to understand the web's structure; and the study of links as indicators of existing social phenomena, namely the social significance of hyperlinks. With respect to the latter, the Digital Methods Initiative examined hyperlinks as part of the study of 'how an actor may be characterized by the types of hyperlinks given and received', what types of associations an actor on the Internet can have and the 'everyday politics of association'.[3] Wilkinson, Thelwall and Xuemei (2003) studied hyperlinks in relation to informal scholarly communication via the web, while Ackland and Gibson (2013) examined how political actors use links as a new form of 'networked communication' to promote themselves, to reinforce their policy messages and to inflate the support they enjoy. Chang, Himelboim and Dong (2009) studied the political economy of hyperlinks and found that the flow of news and information through outgoing hyperlinks between countries remains mostly closed.

Websites and their content constitute another important object of research and have given rise to website analysis (Cai and Zhao, 2013; Das and Turkoglu, 2009; Kingston and Stam, 2013; McCluskey, 2013; Ortega, Aguillo and Prieto, 2006; Schweitzer, 2008). Content analysis is the foremost method of studying websites per se, but it can involve the examination of both website content and aesthetics (e.g., Das and Turkoglu, 2009; McCluskey, 2013; Ortega, Aguillo and Prieto, 2006). At the same time, website analysis can enable researchers to develop a better understanding of web or online phenomena, such as e-democracy, online advertising, online advocacy and others (e.g., Cai and Zhao, 2013; Kingston and Stam, 2013; Schweitzer, 2008). Website analysis is often conducted via webscraping and webmining tools,[4] but it also takes place in the form of 'web archive'

---

[3] For more information, see https://wiki.digitalmethods.net/Dmi/DmiAbout

[4] A compilation of social media/online scraping tools can be found at http://socialmediadata.wikidot.com/

research. A web archive is formed by the archiving process and 'embeds particular preferences for how it is used and for the type of research performed with it'.[5] Web archives allow the study of the history and content of the web and enable study to be organized by time as well as by website type. In addition, researchers are often interested in the analysis of one or more websites at a particular moment in time. Some studies look at web archiving per se and at associated processes and issues (e.g., Gresham and Higgins 2012; Wang, 2007). The Wayback Machine[6] is broadly used for the construction of a narrative around website history, whereas Karpf suggests that 'many publicly available types of content go unsaved and disappear forever. Researchers are limited to whatever the Wayback Machine happens to capture' (2012, p. 648).

Social media and their content constitute another new object of research. Kwon, Park and Kim (2014) studied the motivational factors for using social networking services and user acceptance of platforms such as Facebook and Twitter. Researchers also study the post demographics of social networking sites, namely user profiling, which allows a more insightful definition of self, tastes, interests, 'likes' and other profile characteristics (e.g., Hagger-Johnson, Egan and Stillwell, 2011; Krämer and Winter, 2008; Lorenzo-Romero, Alarcón-Del-Amo and Constantinides, 2012; Pfeil, Arjan and Zaphiris, 2009; Quercia et al., 2011). Whereas social media content frequently comprises a virtual version of pre-existing offline discourses and objects of research, the availability of multidirectional public dialogue, diverse conversations and even multiple threads of monologue in social media spaces poses new opportunities and challenges for researchers (Tsatsou, 2014, p. 170). Social media content can aggregate thousands, millions or even billions of discourses or big data from myriad locals and over an extended period of time, and it comprises a particularly popular

---

[5] See https://www.digitalmethods.net/Digitalmethods/TheWebsite

[6] The Wayback Machine can be accessed at http://www.archive.org

content option for analysis. However, researchers are often presented with challenges, such as the range of tools they can use to aggregate and analyse massive social media data[7] as well as associated literacy and skills requirements. Also, social media content complicates the research process due to the difficulties researchers encounter with labelling and categorising (e.g., many types of discourse under the same theme) as well as with archiving (e.g., how to archive massive social data) and inclusiveness (e.g., what content to include in the analysis and what to leave out) (Tsatsou, 2014, p. 170). Some of these issues and challenges are discussed below, in section 2.3.

2.2. The 'digital' and virtual methods of research

Digital technologies have driven researchers to revisit not only the themes and questions of their research but also their research methodologies (Tinati et al., 2014), often leading to the creation of novel methods of research (Fielding, Lee and Blank, 2008; Hine, 2005; Hughes, 2012; Johns, Chen and Hall, 2004; Jones, 1999; Markham and Baym, 2009; Roberts et al., 2013; Salmons, 2010, 2012; Sappleton 2013).

Early on, scholars acknowledged that tools, applications, content formatting and other Internet affordances not only repackage existing research methods, adapting them to the Internet, but also suggest completely new and often innovative approaches to and methods of research (Jones, 1999; Sosnoski, 1999; Sudweeks and Simoff, 1999). Schneider and Foot (2004) suggested that the analysis of websites and their content requires new analytical tools, since increasingly complex web applications alter traditional relationships between media form and content (Schneider and Foot, 2004, p. 116). More recently, Rogers (2013) separated

---

[7] For instance, to aggregate and analyse tweets, researchers use tools such as DiscoverText, a cloud-based, collaborative text-analytics solution for capturing, filtering, duplicating, clustering, searching, human coding and machine classifying large numbers of small, unstructured items of text. DiscoverText allows researchers to have access to Gnip-enabled data streams for import and, in this way, to data from the full Twitter firehose. For more information, see http://www.discovertext.com

digital methods from virtual methods. He defined virtual methods as imported and migrated from offline research and as adapted to the online environment wherein they are employed (e.g., online surveys, e-interviews). On the other hand, for Rogers, digital methods are those 'native' to the medium (i.e., digital technologies), such as hyperlink analysis, web engine diagnostics, web archival research, web content analysis and social media research. Rogers considers digital methods appropriate for the analysis of digital data (e.g., hyperlinks, web content) and distinguishes them from digitised or virtual methods, aiming to draw researchers' attention to the 'medium' so as to 'reorient Internet research to consider the Internet as a source of data, method, and technique' (2013, p. 27). Others disagree with such a separation, considering it restrictive and arguing that it misses the full range of prospects and possibilities in digital methods (Roberts et al., 2013, p. 6).

In any case, methodological approaches such as web crawling, social media data and text scraping and mining, sentiment analysis, hyperlink analysis, issue network mapping, visualisation, recommendation systems and folksonomy (Kennedy et al., 2015; Rogers, 2013) owe their existence to the tools and affordances of the digital technological domain. On the other hand, virtual versions of mainstream methods of research, such as online surveys, email or Skype interviews, virtual or online focus groups, virtual ethnography or netnography, web social network analysis, web text analysis, digital conversation analysis and Internet-based psychological experiments (Giles et al., 2015; Kozinets, 2010; Meho, 2006; Reips, 2002a, 2002b; Turney and Pocknee, 2005; for more, read Tsatsou, 2014) mostly showcase the added value of the involvement of technology in the design and/or execution of traditional methodological approaches:

> Web-based survey instruments are cheaper, faster and easier than postal service-based surveys. Blog and webpage hyperlinks provide ample traces of network ties, enhancing social network analysis. Several scholars have begun experimenting with Amazon's Mechanical Turk as a low-cost, rapid service for content analysis and other tasks…

Qualitative case analysis and process-tracing research also find fertile ground in the online environment, as time-stamped archives lower the barriers to constructing timelines of events and identifying how participants discussed a case as it unfolded. (Karpf, 2012, p. 648)

Nevertheless, such methodological breakthroughs come with advantages and disadvantages. On the one hand, they allow low-budget and quickly completed research while breaking down time and spatial barriers to participant recruitment (Tsatsou, 2014). In addition, digital methods in particular revolutionise existing methodologies, giving rise to novel approaches to research such as 'computational social science' (Lazer et al., 2009) and signalling the advent of a new era for interdisciplinary and cross-disciplinary research (Ruppert, Law and Savage, 2013). On the other hand, scholars have problematised digital methods, identifying possible problems and issues that researchers should be aware of and take action for (Karpf, 2012; Kennedy et al., 2015). Karpf (2012) argues that digital research methods are untested and often rely upon data sources that are incomplete and systematically flawed. In addition, lack of universal access to the Internet, varying levels of researchers' and research participants' literacy in appropriating technology for research purposes, sample representativeness (Barratt, Ferris and Lenton, 2015), and the wide range of ethical concerns that arise when technology is involved in research (British Psychological Society, 2013; Buchanan and Zimmer, 2012) all problematise both virtual and digital methods of research (Bryman, 2012, p. 658).

2.3. Digital, social or big data?

Research methods are tightly linked with the types, traits and analysis of collected data. In digital research such a link makes researchers who employ digital or virtual methods of research consider issues concerning digital data production, availability, retrieval and analysis.

Digital researchers use terms such as 'digital data', 'social data' and 'big data' interchangeably, without employing clear lines of distinction between them. Some speak about 'social data', referring to data produced through people's multi-formatted and wide-ranging interactions on social media and in other user-based spaces online (Kennedy et al., 2015, p. 172). Others speak about online-generated and web-based data, referring to them as 'big data' (boyd and Crawford, 2012; Coté, 2014; Halavais, 2013; Helles, 2013) and characterising big data as 'a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology that provokes extensive utopian and dystopian rhetoric' (boyd and Crawford, 2012, p. 662). Yet others regard digital data and associated tools and analyses as synonyms of digital methods (Kennedy et al., 2014; Kennedy et al., 2015), although I would argue that digital methods incorporate a vast realm of epistemological and methodological issues concerning digital technologies and research methods – among which are the making, shaping and appropriation of digital data – but they do not necessarily involve, nor are limited to, digital data.

A similar variety of approaches can be found in evaluations of the role and importance of digital data. On the one hand, it is suggested that digital technologies have increasingly enabled the rapid generation and availability of vast amounts and varieties of data (Floridi, 2012; Ruppert, Law and Savage, 2013). On the other hand, an increasing number of scholars are raising questions over the difficulties researchers encounter in uncovering the qualitative traits of such data, the 'small patterns' of big data, and spotting where new patterns with real added value lie in the immense databases and messy data spaces available online (Floridi, 2012). Others note the opposite as problems: the limited scope of web archived data and the lack of tools for producing metrics for web platforms and services of interest for study (Karpf, 2012). Many have also problematised the ephemerality of digital data, their limited accessibility as well as the economic or commercial interests that

determine the availability and quality of such data, hindering their democratic use and openness to public engagement (Baym, 2013; boyd and Crawford, 2012; Karpf, 2012; Kennedy et al., 2014; Williamson, 2014). This is why Karpf (2012, p. 642) refers to the existence of a 'data divide' and Kennedy et al. (2014, p. 9) suggest the 'democratisation of [digital data] analytics'.

Epistemologically speaking, researchers have started to problematise the emphasis on numbers, quantity and large scale when collecting, retrieving and analysing digital, social or big data. Accordingly, some researchers have proposed a qualitative and small-scale approach that looks first at the people and then at the numbers– namely,more at the quality and less at the quantity (Baym, 2013; Helles, 2013; Karpf, 2012, p. 655; Kennedy et al., 2015). However, I would argue that small-scale qualitative analysis cannot fully address the problem of the low reliability and quality of much of the data we find online, since messiness and economic drivers will continue to determine the digital data landscape, while small-scale qualitative studies have their own limitations, such as small scope, and they are highly open to subjective interpretation. At the same time, some have questioned the value of big data as a whole, suggesting that researchers should never compromise the essence, direction or goals of their research simply to make it suitable to a big data approach, as a thorough analysis of smaller data sets can produce even more meaningful results:

> We argue that researchers need to consider whether the analysis of huge quantities of data is theoretically justified, given that it may be limited in validity and scope, and that small-scale analyses of communication content or user behavior can provide equally meaningful inferences when using proper sampling, measurement, and analytical procedures. (Mahrt and Scharkow, 2013, p. 20)

This section has reviewed some developments and arguments concerning digital technologies, services/content and tools as objects, methods and data of research and has tried

to demonstrate not only the novelty of digital research but also its diversity and even messiness at a number of levels. Although the concept of 'novelty tied to messiness' has led to a range of debates and disagreements among digital research scholars and practitioners, most of what has been argued to date offers an out-of-context and strictly technical account of how digital technologies influence various facets and areas of research work, hoping to report 'objective', 'observable' and 'meaningful' trends in digital research (e.g., Nicholas and Rowlands, 2011). On the other hand, there has been hardly any insight into researchers' own experiences and reflective accounts of the role and importance, if any, of digital technologies in their research practices. Even pleas for small-scale, case-study qualitative examination of digital methods and data of research have not looked at individual researchers' experiences and evaluations of digital technologies as objects, tools or facilitators of research work. Likewise, there is too little literature that adopts a comparative perspective and examines the role of digital technologies in more than one discipline or beyond the discipline-neutral level. Even studies that attempt to compare disciplines tend to examine just one particular platform or technology (e.g., Holmberg and Thelwall, 2014). In addition, existing scholarship considers digital means of research work in relation to specific tasks and stages of the research process, such as the recruitment as research participants of hidden populations whose geographical span goes beyond country, language or cultural borders (e.g., Barratt et al., 2015).

This paper aims to address these knowledge gaps by offering an insight into researchers' employment of digital technologies in various phases of the research process and in various scientific disciplines.

3. Methodology

Figure 1 shows the key concepts, scope and aims of the study that this paper draws upon.

Figure 1: Study framework
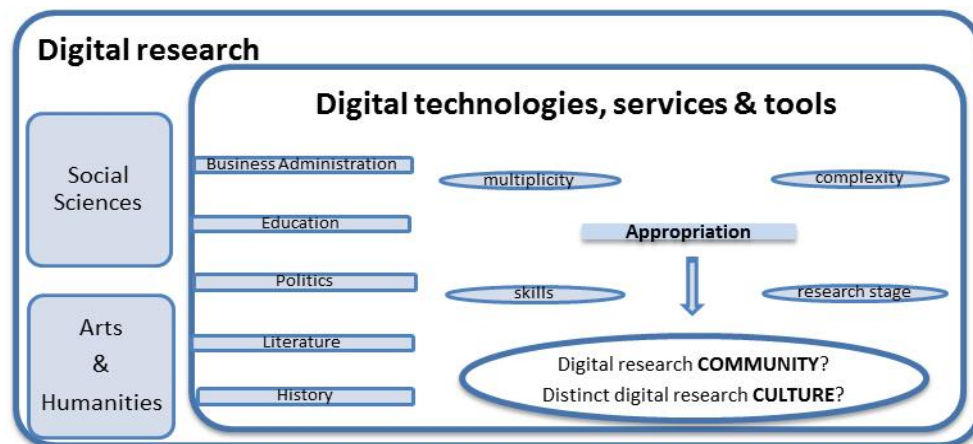
The study explored UK-based research in the disciplines of business administration, education, history, literature and politics so as to develop an understanding of how social researchers in five different disciplines appropriate digital tools, resources and services; the patterns of use of such tools, resources and services; the parameters of 'multiplicity' and 'complexity'; and associated skills and capacity challenges.

The selection of the five disciplines might suggest a too-broad scope of study, but it resonates with the ultimate goal of the study, which is to reach some broad conclusions about the rise of a digital research community and culture in the UK. In addition, the selection of the five disciplines can be justified on the grounds of their similarities and differences. On the one hand, all five disciplines are unfamiliar with the development and form of digital technologies, unlike disciplines such as media and communication, information science, information systems, technology studies and so on. This is useful because the study aimed to explore the employment of digital technologies in areas where digital technology is not an integral part or the very object of research. In addition, these five disciplines involve a large volume of what is broadly defined as 'social research', and a significant bulk of their research outputs derives from the study of society, economics, politics and culture. On the other hand, some of the selected disciplines have just recently started to incorporate digital means of work, and others have for some time provided fertile ground for the use of digital tools in

research. In addition, some of the selected disciplines – history and literature – lie in humanities, whereas others – education, business administration and politics – belong to the social sciences. Although the distinction between humanities and the social sciences is not always clear cut (e.g., many would position politics in humanities), it is a useful use, as it helped this study shed light both on the hypothetically dissimilar dialogues that humanities and the social sciences have developed with technology and on relevant developments (e.g., digital humanities) (Tsatsou, 2015).

The study examined two research cases for each discipline so as to cast light on analogous and dissimilar patterns of appropriation of digital technologies within each discipline as well as across disciplines.[8] The decision to study ten research projects from five different disciplines ensured a sufficiently wide scope and satisfactory disciplinary diversity within the timeframe of the study. The selected cases of research were all UK-based and externally funded (i.e., not merely self-funded) and they all satisfied the following two conditions: first, they employed digital means of work – at least to an extent – so that they were informative with regard to the employment of digital tools in the research process; and second, the research was ongoing (at the time), so that observation of actual research activities was possible and timely insights into the use of digital tools in research were extracted. Sampling started prior to the official start date of this study. It involved systematic mapping of the research field in the five pre-selected research disciplines and consisted of the following steps: (1) listing of ongoing (at the time) research projects (i.e., cases of research)[9] that fell in the five disciplines and made use of or promised to use digital technologies for research-related activities; (2) reviewing of the research specifics and generation of a brief review report for each research case eligible for selection; and (3) finalization of the sample

---

[8] A list of the studied research cases can be found in the Appendix.

[9] The study used the following databases to search and identify eligible cases of research: the Gateway to Research (GtR), the ESRC and AHRC websites, and the UK Data Archive database.

after extensive email, phone or Skype communication took place with the principal investigator (PI) of each research case so as to confirm their participation and also to agree on timeframes, availability and practical arrangements for participation.

Qualitative methods were employed for the study of all the research cases and disciplines. First, we collected observation data for all cases and for various activities in each case. Observation took place in a number of formats, depending on the specifics of each case as well as the preference and convenience of the observed researchers. For most research cases, we conducted non-participant observation. This decision was driven by the understanding that, in this study, the researchers conducting observation should not be involved in the observed activities and their presence should be as invisible as possible so that the observation records would be as unbiased as possible. This relates to the fact that having researchers observing other researchers through participation might introduce the factor of 'evaluation', which could result in certain biases and errors in the observation process as well as in the collected data. For instance, researchers participating in the activities of other researchers who are under observation might result in the former including evaluative comments on the performance of the latter into their observation notes, while the latter might act unnaturally due to the fear of being the subject of evaluation. However, when non-participant observation was not possible or suitable, purposeful demonstration or unstructured and reflective accounts of one or more elements of the research case took place. For instance, it was agreed with the business researchers at the Open University that one of them would provide an unstructured and reflective account of their research.

Second, we conducted semi-structured in-depth interviews with one or more researchers from each and every research case. The number of interviews for each case varied depending on the size of the research team and the availability of individual researchers. In total, we interviewed 19 researchers and our sample covered a range of roles and levels of

15

seniority in the studied cases (e.g., PI, co-investigator, research assistant). The interviews often deviated from the interview topic guide, according to the specifics of each research case and discipline as well as the stage the research had reached at the point of the interview. Inclusion of 19 interviewees can be considered to be a moderate sample for in-depth qualitative interviewing, which, however, came on top of a rather intense ethnographic data collection process for all ten cases of research and prior to the interviews. Therefore, the study collected considerably rich and lengthy qualitative data.

NVivo 10.0 was employed for the interview data analysis and the development of both textual and graphical analytic reflections on the key themes of the study. A coding framework was designed for the NVivo analysis, while new and revised codes were added through intercoder reliability tests and throughout the analysis. The NVivo analysis was complemented with insights from the observation data so as to enable the study to produce a rich set of qualitative insights. Specifically, the observation data were used so that the study could reflect, unpack and even challenge some of the interviewees' words and evaluations, while drawing out specific examples of activities and experiences in support of or questioning some of the interview insights.

On the whole, the interview analysis, combined with reflections and examples from the observation data, relied on a data analysis framework that incorporated the following three levels of analysis:

- *First level – case-level analysis*: main trends in each *research* case; similarities and differences across individual cases.

- *Second level – discipline-level* analysis: main trends in each discipline; similarities and differences between the two research cases in each discipline.

- *Third level – cross-disciplinary-level analysis:* main trends in the five disciplines; similarities and differences across all five disciplines.

4. Findings

4.1. Research case attributes and 'vocabulary'

In the initial phase of the analysis we mapped the attributes of the ten cases (see Table 1).

One interesting pattern is that most cases did not involve any other, second discipline than the

main discipline. Given that the selection of the studied cases was not determined by the

criterion of 'interdisciplinarity', this could say something about the strong disciplinary

attributes of social research in the UK that employs digital technologies. Another interesting

pattern is that the majority of the cases employed digital technology just as a means of

research and only the two education cases employed digital forms of technology as the actual

objects of their work. Specifically, the education case at the Institute of Education aimed to

design, implement and evaluate apps for smart phones that would assist community health

volunteers in Africa in their daily duties. The education case at the University of Edinburgh

aimed to develop a framework that would explore the relationship between children's speech,

gestures and interactions when they explain concepts by using materials manipulated through

digital interfaces. All in all, education research appeared to be concerned with digital

technology as an object of study more than any of the other four disciplines.

Table 1: Research case attributes

Research case	Discipline	Interdisciplinarity	Location	Project team	Seniority of PI	Technology's position
Project 1	Business	No other discipline	England (not London)	5+ persons	Professor	Means of research
Project 2	Business	No other discipline	England (not London)	Up to 2 persons	Professor	Means of research
Project 3	Education	Two or more other disciplines	London	3–5 persons	Reader	Object & means of research
Project 4	Education	Another discipline involved	Scotland	Up to 2 persons	Research staff	Object & means of research
Project 5	History	Another discipline involved	London	3–5 persons	Senior lecturer	Means of research
Project 6	History	No other discipline	England (not London)	3–5 persons	Lecturer	Means of research
Project 7	Literature	No other discipline	England (not London)	5+ persons	Professor	Means of research
Project 8	Literature	No other discipline	Scotland	5+ persons	Professor	Means of research
Project 9	Politics	No other discipline	Scotland	3–5 persons	Professor	Means of research
Project 10	Politics	No other discipline	England (not London)	Up to 2 persons	Professor	Means of research

In conducting an initial scan of themes and patterns of discussion among the participant researchers, the analysis captured the most frequent words the researchers used when discussing digital technologies. As shown in Figure 2, the researchers made reference to a range of words, with the following words (and words with the same etymological root) appearing in the interview texts more than 100 times in total (ordered here in descending frequency): 'think', 'project', 'work', 'see', 'use' 'make', 'know', 'research', 'training', 'data', 'technology', 'find', 'website', 'digital', 'public', 'communication', 'university', 'recording', 'material', 'information', 'documents', 'course', 'support', 'design', 'trying', 'online', 'blog', 'understanding', 'access', 'report', 'resources', 'email', 'social', 'tools', 'video', 'media', 'impact', 'dissemination', 'future', 'Twitter', 'Web', 'learning', 'database', 'outputs', 'archive', 'content', 'images', 'link', 'collection', 'exchange', 'activities', 'follow' and 'practice'.

Figure 2: Word frequency cloud

Whereas all these words belong to various thematic clusters, NVivo showed that most of them populate the following four themes: 'digital technology', 'action', 'reflection' and 'training'. This indicates that the participant researchers' focus was mostly on the active and

reflective appropriation of digital technologies in research, as well as on digital capacity enhancement through learning and training. The remainder of this paper unpacks the theme of active and reflective appropriation of digital technologies by examining trends of appropriation in and through the main phases of the research process and for the parameters of 'complexity' (i.e., level and quality of technology and/or its use) and 'multiplicity' (i.e., single or multiple technologies and/or usages). The theme of digital capacity enhancement, learning and training goes beyond the scope of this paper.

4.2. The 'digital' in the research process

For the analysis of the role of digital technologies in the research process, the study identified key elements and stages of the research process, as shown in Figure 3.

Figure 3: Research process model
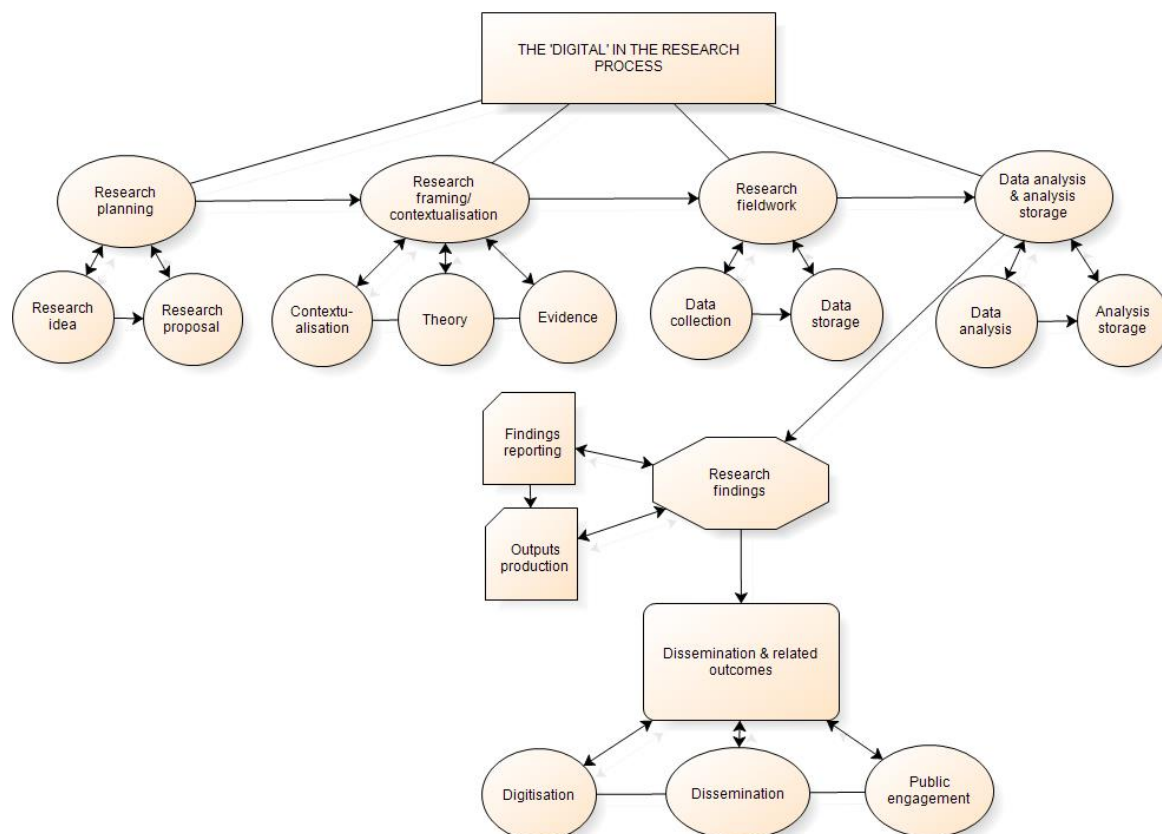
In general, we found that different stages of the research process involved different digital technologies and that, throughout the research process, the researchers mostly made

use of emails, Skype, online repositories and file-sharing systems. In what follows, I present the key findings for each stage of the research process.

4.2.1. The 'digital' in planning and designing research

The participant researchers did not reflect extensively on the use of digital technologies at the stage of research planning and when putting together a research proposal.

In just three cases the researchers made reference to this initial phase of the research process, mostly mentioning the use of online search engines, online databases, web archives and file-sharing systems for research project planning and design purposes. For instance, the researchers in the politics case at the University of Edinburgh used online archives and other resources to retrieve background policy archives, documents and related evidence. In addition, the PI of the education case at the University of Edinburgh mostly used Dropbox to share project planning documents and to enjoy flexibility in working on the research proposal from different places and through various devices.

The participant researchers stated that they had to use online funding application systems, as the studied cases were supported by one or more forms of external funding. A couple of researchers referred to the Je-S application system as being quite straightforward and easy to use.

4.2.2. The 'digital' in framing and contextualising research

Regarding framing and contextualising research, the researchers in six cases referred to the use of digital resources to retrieve research evidence in their respective fields. In three cases the researchers had used such resources to contextualise research, whereas – rather surprisingly – not a single researcher had used digital resources to search theoretical texts.

More specifically, the researchers used online search tools, online databases, web-based archives and carefully selected websites to access vast amounts of evidence, documents

and information that were essential for mapping out relevant areas of research and, to a lesser extent, for contextualising their own research. The PI of the politics case at the University of Edinburgh relied heavily on the Internet to locate government and other policy documents, and she expressed the opinion that online archives are considerably rich and continuously expanding. Similarly, the researchers in the politics case at the University of Nottingham made use of specialised sites such as that of the International Trade Union Confederation site to find statistics that would inform their research. The PI of the Scottish literature case at the University of Glasgow appreciated the tremendous amount of time that can be saved when digital platforms and tools are used in retrieving previously hidden or undiscovered literary resources. On the other hand, he was cautious concerning the assumption that everything is online and digitised and he liked to remind himself that there may be materials that are not yet available online. Thus, he continued to make use of library archives and other offline sources of literary materials, as he thought that a lot of Burns' work is in private hands, and the Burns community demographically consists of senior people who barely make use of digital technologies.

The researchers stated that they were increasingly resorting to online resources to quickly and easily inform, map out and contextualise their research: 'I would say 75 percent I find online, whereas before it was 25 percent online and 75 percent looking at new periodical shelves… the amount of time devoted to finding out has gone down… to simply find things was a lot more laborious, now it's much faster' (PI, history case at UCL). At the same time, they argued that online resources can often be overwhelming in volume and range and not all of them are relevant to their research.

4.2.3. The 'digital' in fieldwork

In most cases, digital means were popular for the conduct of fieldwork. The researchers recognised the value of the wealth of data and repositories available online as well as how

21

data search and retrieval are sped up when technology is used.

The exact technologies used for data collection or retrieval varied by case and also depended on the type of data sought. The researchers in the business case at the University of Manchester used webscraping and webmining software to retrieve and analyse unstructured web data from company websites in China, the UK and the USA. They also made use of the Wayback Machine to access website archives and analyse past web data for the companies under study. In contrast, the education researchers made particular reference to the use of visual and multimedia tools (e.g., cameras, video recording tools) for the production of primary data. Specifically, the education project at the University of Edinburgh aimed to collect video-recorded data from sessions with school children, while the education project at the Institute of Education primarily used mobile-phone-recorded data and assessment results in combination with visual materials such as photographs. On the other hand, the literature, history and politics researchers were reliant mostly on online databases, directories and archives (e.g., LexisNexis, Hansard Society, Scottish music archive, museum archives, library archives) so as to access and retrieve data in textual, documentary or multimedia format. At the same time, the literature, history and politics researchers maintained a strong bond with offline sites (e.g., libraries, museums) to access data archives that were not digitised or accessible online.

Regarding data storage, researchers across cases and disciplines made use of technology to store data and make it broadly available. Some expressed concerns over losing data and used multiple technologies for back-up purposes. They also used file-sharing systems, such as Dropbox, and login-protected online repository systems or shared drives in order for the data to be available within the project team for cross-referencing and cross-analysis purposes. Finally, they used website-housed databases to make the data available to

the broader community. Interestingly, none of the researchers reported the use of cloud systems or Google Drive.

4.2.4. The 'digital' for data analysis and reporting of findings

In six cases, computer-based or online software was used to facilitate and systematise data analysis. For instance, the researchers in the business case at the University of Manchester used advanced data-manipulation and data-analysis software (e.g., Content Analytics, IBM and VantagePoint) to analyse large volumes of unstructured web data. The researchers in the politics case at the University of Edinburgh used NVivo to analyse qualitative interview data, while the video data in the education project at the University of Edinburgh were trimmed via QuickTime and then coded and analysed through specialist software.

As regards the reporting of research findings, in about half of the cases, the project-dedicated website was the main venue for findings to be reported, released or stored. In addition, regardless of copyright constraints, the literature and history researchers made extensive use of digitisation techniques and aspired to create web-based and highly interactive resources so as to present their data, findings and outputs. Interestingly, the history researchers at the University of Leeds aimed for the project's interactive website to be where research resources and outputs would become available. They viewed the interactive website as a research output in its own right and the project blog as a research archive that could potentially be valuable for other researchers in the field.

4.2.5. The 'digital' in dissemination and public engagement

Interestingly, the researchers did not distinguish between dissemination of research and public engagement activities, often referring to the two interchangeably. One exception was the literature case at the University of Leicester, as the project website contained a 'Get involved' link that invited the public to engage with project activities such as a book group, a Waugh Forum and various project events, as well as to help the researchers trace 'orphan

works', namely to uncover Evelyn Waugh's more elusive correspondence.

In general, though, project websites and social media platforms, primarily, and emailing and research summaries on various other websites, secondarily, were the main routes for dissemination and public engagement activities in the studied cases. Nonetheless, there were exceptions, as a couple of research cases did not have a project-dedicated website and the researchers did not think a website was needed for dissemination purposes: 'There are so many places we are supposed to put stuff up... we post our outputs on the ESRC [Economic and Social Research Council] site… we post our outputs on the University website and on individual websites… I'm not sure there's a need to maintain a project website' (PI, business case, University of Manchester).

Regarding social media, the researchers acknowledged the potential for viral and instant dissemination of research through social media tools such as blogs and Twitter, and they seemed to consider social media metrics (e.g., number of re-tweets, comments, posts and followers) an indicator of the breadth of dissemination of their work. A couple of exceptions were noted, though, as not all valued social media as necessary or efficient dissemination and public engagement tools: 'I think we have a very clear dissemination strategy but we never promised that we will spend all of our time on social media… we do some social media communication but I wouldn't say that it's central' (PI, business case, University of Manchester).

In a couple of cases, the researchers used videos in combination with social media platforms as a dissemination tool, thus employing less conventional means of technologically enabled dissemination of their work. For instance, the education researchers at the University of Edinburgh converted video data to a format that could be embedded into visual or multimedia presentations and also uploaded videos to YouTube so as to broadly disseminate their work. The history and literature researchers in particular made digitised resources (e.g.,

song recordings, digitised literary texts, interactive maps) available online in order to disseminate their work. In this regard, digitisation appeared to be not just one of the outputs of the literary and history research but also a key dissemination and research impact tool.

Rather unsurprisingly, in some projects, dissemination was pursued at the end rather than throughout the research. The politics case at the University of Nottingham is such an example, as the researchers were planning to move ahead with the creation of an interactive project website and the production of podcasts to disseminate their work and promote public engagement after the official end date of their project. At the same time, in some cases, such as the Scottish literature case at the University of Glasgow and the history case at the University of Leeds, the researchers noted the value of physical interaction with audiences and communities to which their research work aimed to speak. In fact, offline and online means of dissemination and public engagement appeared to complement and enhance each other. For instance, the project website, social media platforms, research repositories (e.g., SlideShare) and academic networking sites (e.g., Academia.edu) were used to share and link up offline research resources, outputs and events, while conferences and other offline dissemination and networking activities were used to promote the project website, online research outputs, and online or digitised research tools and resources.

4.3. The 'digital' and the parameters of 'complexity' and 'multiplicity'

The above findings demonstrate a degree of diversity as well as common trends in researchers' practices with digital technologies across the research process. The combined presence of diversity and common trends can be shown more clearly by examining the parameters of 'complexity' (i.e., level of complexity of technology and its use) and 'multiplicity' (i.e., single or multiple technologies and usages), at the levels of both the research case and the discipline.

On the one hand, the study found commonalities in the range of digital technologies being used across cases and disciplines. This is so as the researchers made use of certain common technological platforms for the conduct of research, such as the project-dedicated website, web search tools, emails, Skype, social media (blogs in particular), Dropbox, online repositories and other file-sharing systems. In addition, commonalities appeared in the range and purpose of use of these digital means of work across projects and disciplines. For example:

o project-dedicated websites were used mostly as public-facing entities for research dissemination purposes;

o web search tools were used mostly for contextualisation and secondary data search purposes;

o emails were used mostly for internal and external project communication;

o Skype was used mostly for internal project communication;

o social media (e.g., blogs, Twitter, Facebook, LinkedIn, Academia.edu) served various purposes, such as dissemination, development of research dialogue, sharing of resources and so on.

On the other hand, we found differences in the complexity and level of use of digital technologies, mostly across disciplines and a lot less within the same discipline. As largely indicated by the cluster analysis below (Fig. 4), the research cases in the same discipline belonged to the same or to adjacent clusters for the parameter of 'complexity'.[10] More specifically:

---

[10] The politics case at the University of Edinburgh is an outlier, as the coding for the PI interview was assigned different codes from those assigned to the interview with the other three project researchers. However, this can be explained by the number of interviewees in the latter case, something that questions the rightness of the decision to interview three different researchers at the same time.

o Both literature cases made extensive use of advanced online repository systems and digitisation practices to preserve and share literary resources (e.g., online editions, audio material), thus engaging in advanced uses of complex pieces of software.

o Both history cases used interactive digital media to preserve historical resources and disseminate research, thus demonstrating skilled and advanced employment of some complex technologies and tools.

o Both education cases made use of web-based platforms, mostly for data collection and pedagogical/training purposes. Although they did not employ exactly the same technologies, both cases seemed to use tools and technologies of high complexity, even pursuing the production of new software (e.g., mobile assessment systems; software that 'speaks' to children's gestures and cognitive development).

o Both politics cases shed light on the importance of using digital tools that enhance dialogue and communication with stakeholders and appeared to aim at technologies with medium-level complexity and for medium-level usage (e.g., social media). At the same time, the two cases varied in terms of the specific tools and technologies they used and in terms of their timing and scheduling of their use of technology.

o The two business cases did not employ technologies of the same complexity. Unlike the case at the Open University, the researchers at the University of Manchester made advanced use of very complex digital means of work, such as webscraping, webmining, social media data analytics and other social computing techniques for searching, retrieving and analysing unstructured web data. On the other hand, both cases seemed to be quite selective, as neither employed some of the most popular and simple technological tools, such as websites and social media platforms.

Figure 4: Coding clusters by research case
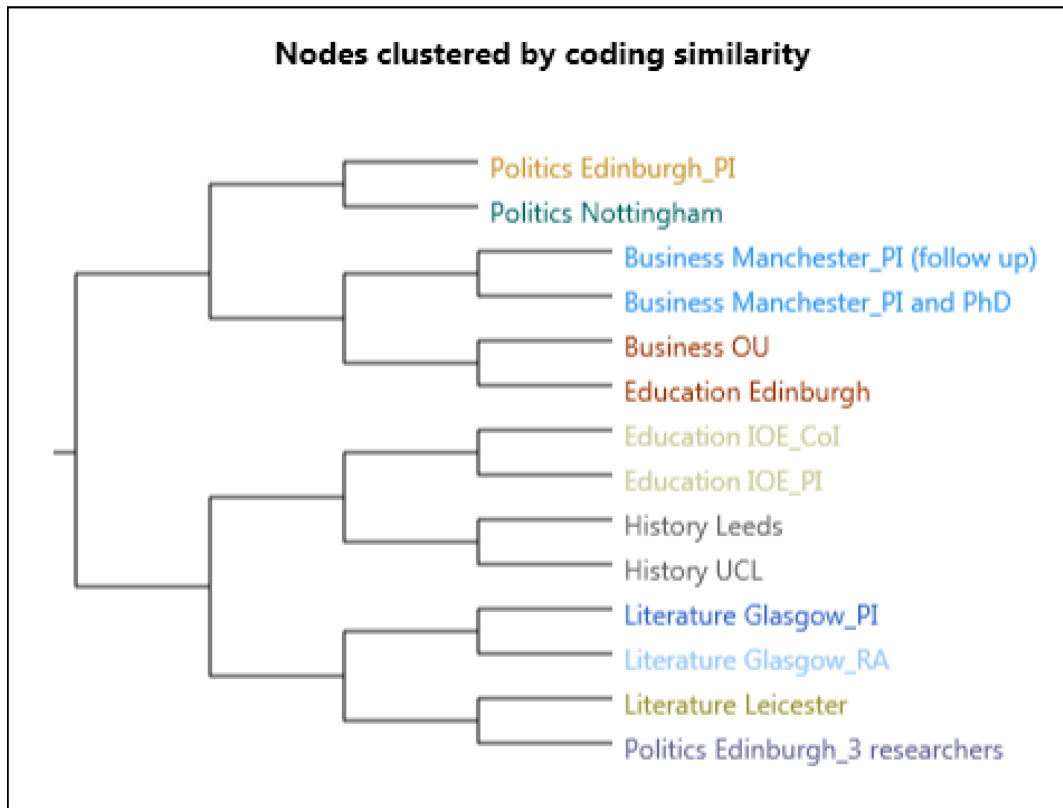
Note: One node is one interview text; a research case can consist of one or two nodes.

To further unpack the complexity and multiplicity trends, the study produced a set of coding matrices (see Table 2) to examine the associations of the parameters of complexity and multiplicity with the research case attributes presented in Table 1 (i.e., discipline, interdisciplinarity, location, size of research team, seniority of PI, whether technology was a means and/or object of research). Specifically, in looking at the association between the attribute of 'discipline' and the parameter of complexity, one could conclude that the researchers in the disciplines of business and education were a lot more concerned with making advanced use of complex forms of digital technologies, tools and services than the researchers in the other three disciplines. On the other hand, Table 2 indicates that the study did not identify any particular association between the attribute of 'discipline' and the parameter of multiplicity, something that confirms, across disciplines, the broadly common patterns that we found for multiplicity.

However, we need to be cautious in how we read Table 2 for the other research case attributes.[11] This is because the other attributes are not evenly distributed across the ten research cases. For instance, as mentioned, most of the studied cases were discipline-specific rather than interdisciplinary (i.e., interdisciplinarity), while eight cases used technology as a means rather than an object of research (i.e., technology's position in the research). Also, in six cases the PI was a professor (i.e., seniority of PI) while the majority of the cases were based at institutions outside London (i.e., location). While this caveat indicates the need for further and larger-scale studies into researchers' employment of digital technologies and associated experiences so as to achieve a relatively even distribution of these attributes across the sample of research cases under study, the findings here point out that the attribute of discipline does matter. The individual researchers and project teams appear to have developed their digital research practices largely on the grounds of what they regard as known and acceptable to their discipline, hence capitalising on what they consider familiar, namely on what resonates with their research expertise and disciplinary etiquette.

Table 2: Research case attributes by complexity and multiplicity

Complexity & Multiplicity Parameters		Research Case Attributes																			
		Discipline (business, education, history, literature, politics)					Interdisciplinarity (whether other disciplines involved)			Technology's position (in research)		Seniority of PI					Research team (number of persons)			Location		
		B	E	H	L	P	no other	one	two or more	means	object & means	research staff	Lect urer	SL	Rea der	Prof	up to 2	3-5	5+	Lond on	England (non-London)	Scotland
Complexity	technology	10	9	2	4	0	15	9	1	16	9	8	1	1	1	14	8	3	14	2	11	12
	usage	8	12	6	7	0	18	10	5	21	12	7	3	3	5	15	7	11	15	8	13	12
Multiplicity	technology	5	4	7	3	5	18	4	2	20	4	2	5	2	2	13	6	11	7	4	13	7
	usage	2	2	6	3	1	10	2	2	12	2	0	4	2	2	6	1	8	5	4	8	2

Note: The figures are the number of references (i.e., coding occurrence) in the interview texts.

---

[11] Table 2 indicates that research which involves just one discipline scores higher in complexity and multiplicity than interdisciplinary research, and also that researchers who use technology as a means or facilitator of research score higher in complexity and multiplicity than those who use technology as both the object and means of their work. In addition, Table 2 gives higher complexity and multiplicity scores for research cases where the PI is a senior researcher.

5. Concluding remarks

This paper reports on a study that aimed to offer insights into how UK-based researchers in five disciplines in humanities and the social sciences employ digital technologies in the research process and in relation to the parameters of 'complexity' and 'multiplicity'.

The study identified certain commonalities and differences in researchers' practices with digital technologies. On the one hand, it found that researchers across the board made extensive use of digital technologies for fieldwork, data analysis, reporting of findings and dissemination purposes and a lot less for planning, designing and contextualising research. In this sense, digital technologies appeared to be a lot more useful for empirical research work and outreach activities than at the start of the research process. On the other hand, even when making extensive use of digital technologies, history, literature and politics researchers would not abandon physical means and non-technological tools of research work, especially for the purpose of searching for and retrieving research resources. By the same token, not all the research cases in all the disciplines supported the usefulness of digital technologies for outreach activities. Specifically, the researchers in the two business cases did not think of social media or a project-dedicated website as important dissemination and outreach tools, while the politics researchers at the University of Nottingham planned to employ digital tools for outreach activities after the official end of the research project.

Similarities and differences were also found in relation to the parameters of 'multiplicity' and 'complexity'. On the one hand, all research cases across all five disciplines demonstrated the use of a range of digital means of work and most of them for very similar purposes or usages. On the other hand, researchers in the same discipline (but not across disciplines) appeared to use technologies and tools of analogous complexity as well as for similarly complex usage. The discipline of business administration was the only main exception, as the study found remarkable diversity with regard to the complexity of the

technologies business researchers were using and the levels of usage of those technologies, regardless of broader traditions and practices in the discipline.

The study found that 'complexity' was determined by researchers' expertise and associated disciplinary traditions and etiquette. The force of research expertise and disciplinary boundaries could explain why researchers in the same discipline showcased common 'complexity' trends and practices in the use of digital technologies. This confirms the conclusion of recent work (Holmberg and Thelwall, 2014) showing that different disciplines display different trends and patterns of usage of technology, showcasing the continuing rigour and influence of disciplinary values and traditions.

These findings could be used to inform the broader social research community and trigger conversations between researchers, research institutions, research funders and technology experts on existing patterns of use of digital technologies in social research and associated opportunities and challenges. Specifically, with just two of the studied research cases having an interdisciplinary scope, the study points out the continuing influence of disciplinary boundaries and traditions, with 'discipline' being a driver of the complexity of digital technologies and their level of usage. This is an important lesson for researchers: first, it refutes assumptions about the tight links and rather symbiotic relationship between interdisciplinarity and digital research; and, second, it invites researchers to move beyond rigid disciplinary boundaries so as to exchange knowledge and learn from each other's experiences with and usage of digital technologies in the research process.

Regarding the existing literature, the study does not dismiss the combined force of 'innovation and messiness' in digital research that the literature reports on, and it acknowledges the thorny problems as well as rapid advancements that digital researchers encounter today. At the same time, the study hopes to contribute to a better understanding of the digital research field. Specifically:

- The study hopes to support the importance of taking a bottom-up view, listening to researchers and their experiences, and making sense of their feedback on the value and effects of digital technologies in the actual research process. Only in this way, I would argue, can we develop a good understanding of the nuances of innovativeness and messiness in the field and also identify concrete problems and opportunities for researchers who aim to employ digital means of work.

- The study also points to the need to refrain from treating digital research as a novel 'research reality', with no links to or dependencies on past research traditions and practices. As the study showed, disciplinary principles and long-lived research traditions highly influence how and how much digital technologies are being used for research purposes. Hence, the study invites scholars to place digital research into a broader research history and discipline context so as to allow a better understanding of related novelties as well as continuities. This challenges approaches that only examine digital research and its novelty, disregarding the role of epistemic traditions, developments and legacies.

Nevertheless, these findings indicate the need for more study and should be subject to further elaboration and testing. More and larger-scale studies of social researchers' experiences with and evaluations of digital technologies are required in order to dig deeper into the factors that influence the use of digital technologies in the research process and associated multiplicity and complexity patterns. In addition, more and larger-scale studies are needed that will embrace a range of disciplines and that will carry out research at the levels of the individual researcher, the research case or project, and the discipline, comparing and contrasting trends in the employment of digital technologies, tools and services across disciplines and especially across humanities and the social sciences.

References

Ackland, R. & Gibson, R. (2013). Hyperlinks and networked communication: a comparative study of political parties online. *International Journal of Social Research Methodology*, *16*(3), 231–244.

Barratt, M.J., Potter, G.R., Wouters, M., Wilkins, C., Werse, B., Perälä, J., et al. (2015). Lessons from conducting trans-national Internet-mediated participatory research with hidden populations of cannabis cultivators. *International Journal of Drug Policy*, *26*(3), 238–249. doi: 10.1016/j.drugpo.2014.12.004

Barratt, M.J., Ferris, J.A. & Lenton, S. (2015). Hidden populations, online purposive sampling, and external validity: Taking off the blindfold. *Field Methods*. doi: 10.1177/1525822X14526838

Baym, N. (2013). Data not seen: The uses and shortcomings of social media metrics. *First Monday, 18*(10). doi: 10.5210/fm.v18i10.4873

boyd, d. & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679. doi: 10.1080/1369118X.2012.678878

British Psychological Society (2013). *Ethics Guidelines for Internet-mediated Research*. INF206/1.2013. Leicester: Author. Retrieved from

http://www.bps.org.uk/system/files/Public%20files/inf206-guidelines-for-internet-mediated-research.pdf

Bryman, A. (2012). *Social Research Methods*. 4th Edition. Oxford: Oxford University Press.

Buchanan, E.A. & Zimmer, M. (2012). Internet research ethics. *The Stanford Encyclopedia of Philosophy*, Fall 2012 Edition. Retrieved from

http://plato.stanford.edu/archives/fall2013/entries/ethics-internet-research

Cai, X. & Zhao, X. (2013). Online advertising on popular children's websites: structural features and privacy issues. *Computers in Human Behavior,* 29(4), 1510–1518.

Chang, T.-K., Himelboim, I. & Dong Dong. (2009). Open global networks, closed international flows: world system and political economy of hyperlinks in cyberspace. *International Communication Gazette, 71*(3), 137–159.Costigan, J.T. (1999). Forests, trees, and internet research. In Jones, S. (Ed.), *Doing Internet Research: Critical Issues and Methods for Examining the Net* (pp. xvii–xxiv). Thousand Oaks, CA: Sage.

Coté, M. (2014). Data motility: The materiality of big social data. *Cultural Studies Review, 20*(1). doi: 10.5130/csr.v20i1.3832

Das, R. & Turkoglu, I. (2009). Creating meaningful data from web logs for improving the impressiveness of a website by using path analysis method. *Expert Systems with Applications, 36*(3), 6635–6644.

De Maeyer, J. (2013). Towards a hyperlinked society: a critical review of link studies. *New Media &* Society, *15*(5), 737–751.

Fielding, N. Lee, R.M. & Blank, G. (2008). The *SAGE Handbook of Online Research Methods*. London: Sage.

Floridi, L. (2012). Big Data and Their Epistemological Challenge. *Philos. Technol.* 25, 435–437. doi: 10.1007/s13347-012-0093-4

Giles, D., Stommel, W., Paulus, T., Lester, J. & Reed, D. (2015). Microanalysis of online data: the methodological development of ''digital CA''. *Discourse, Context and Media, 7*(2015), 45–51.

Gresham, E. & Higgins, S. (2012). Improving browsability of archive catalogues using web 2.0. *Library Review, 61*(5), 309–326.

Hagger-Johnson, G., Egan, V. & Stillwell, D. (2011). Are social networking profiles reliable indicators of sensational interests? *Journal of Research in Personality*, *45*(1), 71–76.

Halavais, A. (2013). Home made big data? Challenges and opportunities for participatory social research. *First Monday, 18*(10). doi: 10.5210/fm.v18i10.4876

Helles, R. (2013). The big head and the long tail: An illustration of explanatory strategies for big data Internet studies. *First Monday, 18*(10). doi: 10.5210/fm.v18i10.4874

Hine, C. (Ed.) (2005). *Virtual Methods: Issues in Social Research on the Internet*. Oxford: Berg.

Holmberg, K. & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, 101, 1027–1042. doi: 10.1007/s11192-014-1229-3

Hughes, J. (Ed.) (2012). SAGE *Internet Research Methods*. London: Sage.

Johns, M.D., Chen, S-L. & Hall, G.J. (Eds.) (2004). *Online Social Research: Methods, Issues & Ethics*. New York: Peter Lang.

Jones, S. (Ed.) (1999). *Doing Internet Research: Critical Issues and Methods for Examining the Net*. Thousand Oaks, CA: Sage.

Karpf, D. (2012). Social science research methods in internet time. *Information, Communication & Society*, *15*(5), 639–661. doi: 10.1080/1369118X.2012.665468

Kennedy, H., Moss, G., Birchall, C. & Moshonas, S. (2015). Balancing the potential and problems of digital methods through action research: methodological reflections. *Information, Communication & Society*, *18*(2), 172–186, doi: 10.1080/1369118X.2014.946434

Kennedy, H., Moss, G., Birchall, C. & Moshonas, S. (2014). 'Digital Data Analysis, Public Engagement and the Social Life of Methods: Final Report'. *Working Papers of the Communities & Culture Network+* Vol.3, April 2014.

Kingston, L.N. & Stam, K.R. (2013). Online advocacy: analysis of human rights NGO websites. *Journal of Human Rights Practice*, *5*(1), 75–95.

Kozinets, R.V. (2010). *Netnography: Doing Ethnographic Research Online*. Thousand Oaks: Sage.

Krämer, N.C. & Winter, S. (2008). Impression management 2.0: the relationship of self-esteem, extraversion, self-efficacy, and self-presentation within social networking sites. *Journal of Media Psychology*, *20*(3), 106–116.

Kwon, S.J., Park, E. & Kim, K.J. (2014). What drives successful social networking services? A comparative analysis of user acceptance of Facebook and Twitter. *The Social Science Journal*, *51*(2014), 534–544. doi: 10.1016/j.soscij.2014.04.005

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., et al. (2009). Computational social science, *Science*, *323*(5915), 721–723.

Lorenzo-Romero, C., Alarcón-Del-Amo, M.-C. & Constantinides, E. (2012). Segmentation of users of social networking websites. *Social Behavior and Personality*, *40*(3), 401–414.

Mahrt, M. & Scharkow, M. (2013). The value of big data in digital media research. *Journal of Broadcasting & Electronic Media*, *57*(1), 20–33. doi: 10.1080/08838151.2012.761700

Markham, A.N. & Baym, N.K. (Eds.) (2009). *Internet Inquiry: Conversations about Method*: Thousand Oaks, CA: Sage.

Markham, A. & Buchanan, E. (2012). *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*. Retrieved from http://aoir.org/reports/ethics2.pdf

McCluskey, M. (2013). Website content persistence and change: longitudinal analysis of pro-white group identity. *Journal of Information Science*, *39*(2), 188–197.

Meho, L.I. (2006). E-mail interviewing in qualitative research: a methodological discussion. *J. Am. Soc. Inf. Sci. Technol*., *57*(10), 1284–1295.

Nicholas, D. & Rowlands, I. (2011). Social media use in the research workflow. I*nformation Services and Use*, *31*(1), 61–83.

Ortega, J.L., Aguillo, I. & Prieto, J.A. (2006). Longitudinal study of content and elements in the scientific web environment. *Journal of Information Science*, *32*(4), 344–335.

Quercia, D., Kosinski, M., Stillwell, D. & Crowcroft, J. (2011). Our twitter profiles, our selves: predicting personality with twitter. In *Proceedings of the Third IEEE International Conference on Social Computing*, Social Com' 11, 180–185.

Park, H.W. & Thelwall, M. (2003). Hyperlink analyses of the world wide web: a review**.** J*ournal of Computer-Mediated Communication*, *8*(4). Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2003.tb00223.x/full

Pfeil, U., Arjan, R. & Zaphiris, P. (2009). Age differences in online social networking – a study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior*, *25*(3), 643–654.

Reips, U.-D. (2002a). Standards for internet-based experimenting. *Exp. Psychol, 49*, 243–256.

Reips, U-D. (2002b). Internet-based psychological experimenting: five dos and five don'ts. *Soc. Sci. Comput. Rev.*, *20*, 241–249.

Roberts, S., Hine, C., Morey, Y., Snee, H. & Watson, H. (2013). *'Digital Methods as Mainstream Methodology': Building Capacity in the Research Community to Address the Challenges and Opportunities Presented by Digitally Inspired Methods*. NCRM Networks for Methodological Innovation Report, June 2013.

Rogers, R. (2013). *Digital Methods*. Cambridge, MA: MIT.

Ruppert, E., Law, J., & Savage, M. (2013). Reassembling social science methods: The challenge of digital devices. *Theory, Culture and Society*, *30*(4), 22–46.

Salmons, J. (2010). *Online Interviews in Real Time*. London: Sage.

Salmons, J. (2012). Designing and conducting research with online interviews. In Salmons, J. (Ed.), *Cases in Online Interview Research* (pp. 1–30). Thousand Oaks: Sage.

Sappleton, N. (Ed.) (2013). *Advancing Research Methods with New Technologies*. Hershey, PA: IGI Global.

Schneider, S.M & Foot, K.A. (2004). The web as an object of study. *New Media & Society*, *6*(1), 114–122.

Schweitzer, E.J. (2008). Innovation or normalization in e-campaigning? A longitudinal content and structural analysis of German party websites in the 2002 and 2005 national elections. *European Journal of Communication*, *23*(4), 449–470.

Shumate, M. & Lipp, J. (2008). Connective collective action online: an examination of the hyperlink network structure of an NGO issue network**.** *Journal of Computer-Mediated Communication*, *14*(1), 178–201. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2008.01436.x/abstract

Sosnoski, J.J. (1999). Configuring as a mode of rhetorical analysis. In Jones, S. (Ed.), *Doing Internet Research: Critical Issues and Methods for Examining the Net* (pp. 127–145). Thousand Oaks, CA: Sage.

Sterne, J. (2005). Digital media and disciplinarity. *The Information Society*, *21*(4), 249–256.

Sudweeks, F. & Simoff, S.J. (1999). Complementary explorative data analysis: the reconciliation of quantitative and qualitative principles. In Jones, S. (Ed.), *Doing Internet Research: Critical Issues and Methods for Examining the Net* (pp. 29–55). Thousand Oaks, CA: Sage.

Tinati, R., Halford, S., Carr, L. & Pope, C. (2014). Big data: methodological challenges and approaches for sociological analysis. *Sociology*, *48*(4), 663–681. doi: 10.1177/0038038513511561

Tsatsou, P. (2015). The Role of Digital Technologies in Social Research in the UK: An Emerging Digital Research Community. Final Report. *Working Papers of the Communities & Culture Network+* Vol. 5 (April 2015).

Tsatsou, P. (2014). *Internet Studies: Past, Present and Future Directions*. Surrey: Ashgate Publishing.

Turney, L. & Pocknee, C., (2005). Virtual focus groups: new frontiers in research. *Int. J. Qual. Methods*, *4*(2), 1–11.

Wang, L-C. (2007). Web archives: the concept and application of multi-level description model. *Journal of Educational Media and Library Science*, *44*(4), 455–471.

Wilkinson, D., Thelwall, M. & Xuemei, L. (2003). Exploiting hyperlinks to study academic web use**.** *Social Science Computer Review*, *21*(3), 340–351.

Williamson, B. (2014). The death of the theorist and the emergence of data and algorithms in digital social research. Impact of Social Sciences, LSE Blog. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2014/02/10/the-death-of-the-theorist-in-digital-social-research/

Appendix: research cases

Title	Institution	Principal investigator	Discipline
Home-based businesses and online home-based businesses	Open University	Prof Elizabeth Daniel Dr Muhammad Naveed Anwar	Business
Sustaining growth for innovative new enterprises	University of Manchester	Prof Philip Shapira	Business
mCHW: a mobile learning intervention for community health workers	Institute of Education (IoE)	Dr Niall Winters	Education
Interaction, embodiment, and technologies in early years learning	University of Edinburgh	Dr Andrew Manches	Education
Law and the end of empire 2: AD 800–1150	University College London (UCL)	Dr Richard Salway	History
Antique dealers: the British antique trade in the 20th century	University of Leeds	Dr Mark Westgarth	History
Editing Robert Burns for the 21st century	University of Glasgow	Prof Gerard Carruthers	Literature
The complete works of Evelyn Waugh	University of Leicester	Prof Martin Stannard	Literature
The politics of monitoring: information, indicators and targets in climate change, defence and immigration policy	University of Edinburgh	Prof Christina Boswell	Politics
Globalisation, national transformation and workers' rights: an analysis of Chinese labour within the global economy	University of Nottingham	Prof Andreas Bieler	Politics