

**Key Issues for Digital Research:  
A Social Science Perspective on Policy and Practice**

William H. Dutton, Marina Jirotko, Eric Meyer,  
Ralph Schroeder, and Cohen Simpson\*

University of Oxford

DRAFT 31 May 2012

A Forum Discussion Paper for the Oxford e-Social Science Project of the Oxford Internet Institute, Oxford e-Research Centre, and the Saïd Business School, University of Oxford.

\* The authors thank all the participants at the forum (Appendix B), whose contributions we have tried to reflect in this discussion paper. Cohen Simpson served as the rapporteur for the forum, writing initial drafts of this discussion paper in collaboration with his co-authors. Authors are listed alphabetically to denote their equal contributions.

## Preface

The Oxford e-Social Science (OeSS) project investigated the uses and impacts of digital research – what others have called e-Research or e-Science – from the perspective of the social sciences. The study examined the factors shaping new approaches to digital research across the sciences and humanities as well as its implications for the nature and quality of research, in addition to the ethical, legal, and institutional issues it raises in particular research areas. Appendix A provides an overview of the OeSS project.

As the OeSS project came to an end after six years of research across two phases, project members hosted two days of events focused on the challenges presented by digital research, and the issues raised for policy and practice. A key aim was to move beyond the specific findings of the project and reflect more broadly on the themes for future research in this area. These events focused on the challenges and opportunities presented by digital research for practitioners, social scientists and policymakers engaged with digital tools and methods, seeking to pool their insights in order to inform the field's future.

The two days began on 12<sup>th</sup> March 2012 with an open symposium at Oxford's Keble College where 50 attendees heard five keynote speakers<sup>1</sup> and a dozen paper presentations discussing data, methods, collaboration, interdisciplinary research, and new models of scholarly communication. On the second day, 13<sup>th</sup> of March 2012, OeSS investigator William Dutton moderated a forum discussion that provided much of the input that has shaped this paper. Invited participants engaged in lively and wide-ranging discussion around new problems for digital research. These discussions focused on topics such as the wealth of digital data, collaboration across disciplines and sectors, supporting computational approaches to social research, and training new generations of researchers in computational methods, among others. The 26 expert participants (see Appendix B) in the policy forum included not only academics, but also representatives from industry, government, and funding bodies including the Economic and Social Research Council (ESRC), Engineering and Physical Sciences Research Council (EPSRC), and the Joint Information and Systems Committee (JISC).

A number of recommendations developed out of the forum. We do not mean to imply that participants at the forum reached consensus. We offer them to provide a sense of the ideas generated by the discussion and to stimulate and inform future debate about policy and practice. Even though the OeSS project has come to an end, academics within the OII, OeRC, and Saïd Business School maintain a strong focus on understanding collaboration and computation in research across the disciplines. Indeed, the University of Oxford has a number of continuing and related projects focused on using digital research in the social sciences and analysing its various aspects. Thus OeSS researchers will continue conducting research in this area with current and future grant-funded projects arising out of the inspiration of this ESRC project, such as through the emerging focus on 'big data' and our continuing work on collaborative network organisations and the information and communication practices of researchers.

---

<sup>1</sup> <http://microsites.oii.ox.ac.uk/oess/webcasts>

## Executive Summary

Advances in information and communication technologies (ICTs) are reconfiguring the way scholarly research is conducted across all disciplines. The use of powerful and versatile networking and computational systems promises to change research activity as profoundly as the mobile phone, the Internet, and email have changed everyday life. Scientific collaboration is increasingly seen as dependent upon effective access to shared data, often in new forms such as links, video records, and sensor data, and the advanced information tools that enable data storage, search, retrieval, visualisation, and higher level analysis.

The growing role that the Internet and advanced information communication technologies play in the practice of research across the sciences and humanities promises to transform the way observations of the physical and social world are acquired, shared, analysed, and translated into useful knowledge. By looking at digital research from a social shaping perspective, social scientists can help practitioners and policy makers better understand how researchers mould the technologies they use, and how they in turn have their behaviour shaped by technological, social, ethical, legal and organisational forces.

This paper outlines key issues raised in the discussion of digital research. Key recommendations that emerged from these issues included:

### Implications for Research

- Researchers with a strong capacity for digital research should develop and offer instruction in digital methods in addition to mid-career training to up-date the skills of other academics.
- Academics without an awareness or confidence in the use, strengths and weaknesses of digital research should seek training in data analytics, networked collaboration, and computational approaches to keep up with both students and the field.

### Implications for Methodologists and Digital Researchers

- Ensure that the best methodological practices are translated to new forms of digital research, from understanding the bias of samples and populations to the reliability and validity of particular tools and indicators.
- Question taken-for-granted assumptions and ample hype about the value of digital research in each particular implementation in addition to questioning the validity of the ways in which digital methods shape research questions and designs.

### Institutions of Higher Education

- Create one or more undergraduate courses to introduce digital research (e.g. such as a course on data analytics) as a necessary aspect of digital literacy,

and consider introducing digital research modules as a required aspect of curricula.

- Ensure that postgraduate opportunities for innovative interdisciplinary research are both supported and encouraged in addition to making sure postgraduates have a strong understanding of how data is collected, stored, archived, analyzed, and distributed in the information age. In other words, work to raise the 'data literacy' of new researchers.

#### Policy Makers and Research Councils

- Urge all areas of research to consider new digital research strategies in their approaches to empirical research when appropriate, realizing that older research strategies must remain central to our various disciplinary tool kits.
- Raise awareness, and observance of social, ethical and legal issues tied to digital research, from privacy to informed consent.

#### Implications for Research on the Public at Large

- Raise awareness of the situations in which researchers should anonymize data when reporting analyses of content from the public web. This is particularly crucial when dealing with sensitive topics as republication of public content removes it from the original content within which it was embedded, potentially resulting in issues around discoverability should the content be sensitive and its owner be easily (or inadvertently) identifiable.
- Be aware that an individual's actions online leave an array of digital traces which may be used by researchers and government for various purposes and by business to drive advertising. Thus members of the public must consider the advantages and disadvantages of sharing information online across different contexts, pushing for legal changes to protect privacy should it be necessary.
- Insist that products of digital research, such as maps and visualizations, be transparent regarding the sources, samples and limitations of data employed in addition to protecting the privacy of individuals.

These recommendations and an array of issues tied to the vitality of digital research are more fully developed in the body of this discussion paper. We hope this paper helps foster creative responses, in addition to consideration of some of the solutions posed by forum participants.

## **Key Issues for Digital Research**

There are a number of key issue areas that emerged in discussion of the social shaping and implications of digital research. They include the growing wealth of digital data, the potential for digital collaboration, new forms of scholarly communication, the ethical challenges of digital research, the reshaping of institutional boundaries, and the need for digital curricula. While there are many other issues, these surfaced as key in discussion about digital researchers and colleagues from business, industry and the policy communities with an interest in the vitality of this burgeoning area of research. The following sections of the report discuss each of these issues in turn before concluding with thoughts on moving ahead. Appendix A provides an overview of the Oxford e-Social Science Project, which provided a foundation for the forum discussion, and Appendix B a list of participants in the policy forum that informs this report.

## **The Growing Wealth of Digital Data**

The Internet and ICTs generate incredible amounts of behavioural data. The “data deluge” (Hey and Trefethen 2003) has figured prominently in debates about e-science since its inception. However, recent discussions have shifted away from the deluge as a ‘problem’ to the resources provided by the large-scale of information produced by individuals and organisations (Dutton and Jeffreys 2010; Nentwich and König 2012).<sup>2</sup> This ‘big data’ perspective provides a complement and contrast to the data generated by surveys—a tool social scientists have traditionally used to measure attitudes and behaviours in known populations. In recent years survey response rates have declined, while the costs to execute them offline have steadily increased.<sup>3</sup> Despite the costs for online execution being virtually non-existent, the proliferation of cheap and virtual questionnaires have resulted in a population of Internet users that are often over surveyed—in addition to being non-representative. Thus, what was once viewed as a core of social research is swiftly becoming harder to execute in a reliable and valid manner.

However, the exploitation of user-generated and other big data sources is not a fix to the problems facing traditional research methods, as big data sources carry some of their own risks to reliability and validity, as well as a host of legal and ethical issues. The evaluation of big data is difficult as it is often generated for very specific purposes, and frequently behind closed proprietary doors by private organisations and governmental bodies. This context collapse, i.e. data obtained for one purpose potentially being used for another, obfuscates the meaning behind users’ construction of these new units of analysis. In this new data deluge, researchers risk the loss of human elements and factors as they move increasingly further into large-scale abstraction. The implications of this loss of context must be better understood.

---

<sup>2</sup> View “Webometrics: The Evolution of a Digital Social Science Research Field”: [http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312\\_420](http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312_420)

<sup>3</sup> Tracking the experiences of over 30 nations involved in the World Internet Project, these trends have become apparent. Few countries can afford field interviewing and phone interview response rates are often too low to generate the confidence required in the resulting samples for analysis.

Additionally, with new data comes a focus on the utility of old social theory—explanations of the local and the global which have grown out of data collected with the survey, direct observation and the interview. To what extent does classical social theory speak to the large, macro-behaviour reflected in big data? A set of related issues is the shaping of social scientific inquiry by the allure and the limits of digital data. In what ways does this new data influence and impact research questions and research design and, more generally, what does this mean for the production of knowledge by means of digital methods? New techniques and data sources now drive the construction of some social science questions, rather than the research questions determining the methodological approach. One example here are research designs built around Twitter, which often focus on short time spans simply because they are the easiest to execute given the volume limitations of Twitter's Application Programming Interface (API).

A discussion of digital data would be remiss without mention of its commoditisation. The ability of researchers to collect data is currently effectively trumped by those of industry. Companies such as Facebook and Google collect and own huge quantities of data about their users. Over time, there have been increases in costs to access digital data owned by companies in addition to an increase in the number of restrictions on and exclusivity of their APIs. This scenario favours other members of industry and business with the large budgets necessary for paid access. Is this the end of a short-lived era of open access to online relational and transactional data? And what does it mean for digital research when these data are kept behind paywalls? This is not to say that all data have become commercial: many sources of free data exist. Nor does it mean that all data should be public and open, since much is potentially sensitive and personal information. Still, the question of who controls the most useful or powerful data is increasingly on the agenda for researchers.

Another challenge lies with the curation infrastructure necessary to manage such a large amount of data, infrastructure that is almost nonexistent in digital research at present. Data repositories are rare, standards for data structures and metadata may not exist, and, when they exist, data handling practices are based on local and embedded knowledge. There tends to be a mismatch between those academics who understand the depth of this problem and the public who may have little sense of the importance of data collection and preservation—many academics and members of the public being oblivious to curatorial issues. Additionally, there are often differences in opinion between academic institutions and government over what kind of information should be collected and preserved (Borgman 2012).<sup>4</sup>

While scholars such as historians and archaeologists, have always been aware of large-scale loss of data and artefacts, e-researchers are only beginning to see how these issues apply to contemporary data. What are the cost and benefits associated with data curation, and to whom? What data are most worthy of curation and who decides? Who should fund this curation and by what means? Where should the data be maintained, by whom and for how long? The solutions, as Borgman argues, are hardly likely to be “one-size-fits-all”, different fields facing different challenges, infrastructure models, and a variety of researcher needs and time horizons.

---

<sup>4</sup> View “Reproducibility: Gold or Fool's Gold in Digital Social Research?”:  
[http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312\\_417](http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312_417)

## **The Potential for Digital Collaboration**

Digital research has contributed to, while benefitting substantially from, the globally distributed collaboration made possible by the Internet and ICTs. There is a sense that digital research inherently requires teams,<sup>5</sup> diverse and varied scholars coming together to focus on substantive problems. The ways in which technology allows researchers to compress space and time and cut across disciplinary and institutional boundaries begs the question of true transdisciplinarity – scholars engaged in e-Science that transcends disciplinary boundaries. The possibilities of digital research have fostered a desire for scholars to collaborate in ways previously believed to be impossible. The degree to which new networking capabilities allow new forms of collaboration such as distributed problem solving (i.e. crowdsourcing) has led e-researchers to increase efforts to partner with industry and government in new ways. Academics are now embedding themselves in policymaking, the private sector and the digitally connected economy, in addition to general efforts of interdisciplinarity within the academy itself.

There are major problems involved with ensuring that interdisciplinary collaborations work well, not the least of which is effective communication across disciplinary divides (see Cummings and Kiesler, 2008; 2003). Personalities and perspectives towards science and the research process come to a head when discussing collaboration. Considerations for collaboration in digital research include group-norms and culture around data curation, preservation and sharing which sit alongside traditional disciplinary differences around ethics and the nature of knowledge. Questions of appropriate venues for publishing research results are also vitally important. For instance, computer scientists, engineers, social scientists, and humanities scholars have very different norms regarding what constitutes a ‘top’ high-impact publication. Also, while multidisciplinary projects are increasingly favoured by funding agencies, discipline-specific publications are still the way research impact is judged in exercises like the UK REF (Research Excellence Framework) exercise, which has little provision for multidisciplinary outputs.

A major challenge for scholars engaged in collaborations outside of the academy is the different pace with which industry and government tend to operate. Digital researchers need an understanding of differences and similarities between logistical rhythms inside and outside of the academy. The facilitators of academic-industry and academic-government collaborations in addition to the various barriers to these types of collaboration need to be identified as early and as thoroughly as possible. The creation of incentives for interdisciplinary collaboration, such as the various boundary crossing funding programmes between the UK research councils, may counter and help overcome the traditional institutional incentives for disciplinary divides. Finally, e-researchers are increasingly working with datasets owned by organisations that are not in the academy. What does it mean to work with people who may have their own expectations about how to collect, manage, share and

---

<sup>5</sup> View “Digital Social Research: An Interdisciplinary Niche or the Future of the Social Sciences?": [http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312\\_419](http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20120312_419)

especially exploit data that does not necessarily align with those of the academy?  
Are researchers losing control over research instrumentation?

Trust is a fundamental aspect of ensuring the success of collaborations around digital research. Allowing others to shape research practices and products may pose a very threatening experience to a scholar. Here, collaboration is not so much about being able to talk to collaborators, it is working in, around, and through an issue. While researchers may attempt to avoid disciplinary clashes by sidestepping interdisciplinary research altogether, an approach that is less antithetical to innovation is to acknowledge and address difficulties while ensuring that priorities are negotiated and communication issues are recognised and solved. Thus, there is a need for intermediaries privy to the intricacies and nuances of disciplinary divides and who are able to translate the languages of respective fields.

## **New Forms of Scholarly Communication**

The research community has experienced profound changes in the ways in which it communicates with those within and outside of the academia. These changes are the result of a number of internal and external pressures. Scholars are expected to do research faster, share it widely and ultimately generate significant impact. Concurrently, there has been a rise in the number of tools for digital scholarly communication (DSC), including platforms such as arXiv.org, SSRN, and institutional and disciplinary pre-print repositories which provide new ways to share pre- and post-publication materials, while services such as CiteULike, Academia.edu, and Google Scholar which provide new, and often simpler, conduits for information discovery. In addition to allowing for greater establishment of researcher presence and what marketers increasingly refer to as clout, DSC is a tool which enables additional transparency in the research process and increased accountability of academics to society at large. Controversy over the validity of scientific results has become an ongoing concern in certain fields such as climate science and certain fast-moving areas of medical research. The growing public scepticism in some countries about the research process has led to calls for science to be as open as possible for public use and scrutiny, although motivations behind this perspective vary widely. Tools for DSC allow for the public to take a more active role in its consumption of scholarly material by enabling engagement and critical exchange in discussions of accountability and ethics.

Despite the new affordances of DSC, uptake of its various forms has been met with resistance, cynicism and disinterest by some researchers. Many raise questions and concerns regarding the utility of this form of communication in a world driven by a publish-or-perish mentality where many of the highest impact journals, which have developed their reputations over many years, are proprietary journals that are not open access. While there are good arguments for academic journals to be supported by a mix of different business models, since no model is ideal, issues of access to and control over academic publication have become prominent.

Beyond scholarly journal publication, while time constraints of academia are real, researchers across the disciplines are seen to have an increasing responsibility to be visible on the Web. In key research areas, however, academics can be lost in the



long-tail of online presence, when Web presence is dominated by entities such as think-tanks and non-governmental organisations that are using digital forms of outreach more strategically.

The Internet has become a key gateway to scholarship for members of the public that consume scholarly material, in addition to members of industry and government that shape policy. While the academic presence in these spaces might seem a non-issue at first thought, the Internet is not unbiased in the presentation of material by means of such things as web-site rankings and search engine algorithms. Thus, there is a possibility of public misinterpretation of research due to a bias in how materials are presented or become visible (and others less visible) online. One obvious solution is research, currently only in its infancy, into how DSC is changing what research is accessed and by whom, and how the visibility and rankings and citations of scholarly materials has transformed the landscape of online knowledge.

## **The Ethical Challenges of Digital Research**

It is not clear where potential harm lies in the digital model of research, while the human subjects model (based on medical research norms) makes it relatively clear where harm might be done to participants. Yet, some scholars have argued that the ideas advanced by the human-subjects model, particularly that of informed consent, is inadequate when applied to virtual space. The human subjects model, with its basic assumptions often embedded in sensitive medical research, has an unintended consequence of subtly encouraging scholars in less-sensitive research areas to simply “tick boxes” on human subjects forms that ask many questions unrelated to social research—digital or otherwise. Researchers are often diligent in attempting to preemptively identify all of the ways in which data will be used and manipulated, yet find this an impossible task in most digital contexts. e-Researchers are in need of a comprehensive framework from which to draw ethical principles for resolving concrete questions in an array of scenarios. How thorough and up-to-date are current guidelines such as those of the Association of Internet Researchers, last distributed in 2002, and how far should new guidelines go in addressing the ethical issues of digital research? While the drafting of ethical guidelines around digital methods is indeed a matter of increased discussion among e-researchers, this process requires the incorporation of public opinion around the collection and use of big data. The challenge lies with public engagement regarding the ethics of digital research and the adoption of human concerns in ways that do not stifle innovation.

As previously mentioned, means of DSC allow virtual spaces to be used as venues for the presentation of scholarship and for public engagement. Through DSC, academics may foster conversations around a model of Responsible Research and Innovation (RRI). This calls for a network of interlinking and overlapping responsibilities that need to be embedded in a general culture of responsible practice (Jirotko, Eden, & Stahl, 2012). Key questions for RRI are: who answers to whom in the research environment under inquiry? What do they answer for? And what are a piece of scholarship’s expected positive outcomes vs unexpected negative outcomes in the case of an innovation in digital tools, data sets or methods?<sup>6</sup> For example,

---

<sup>6</sup> View “Visioning Studies: A Socio-technical Approach to Designing the Future”:

growing popularity and use of social network analysis, among social scientists and industry professionals, require researchers to acknowledge their awareness of the implications of their mapping of an individuals network(s) as privacy stands to be infringed upon (see Zimmer, 2010).

## **The Reshaping of Institutional Boundaries**

The needs of academics engaged in globally distributed research are often in conflict with those of the institutions within which scholars are embedded. Researchers may find that efforts to engage in interdisciplinary inquiry and inter-institutional collaboration are inhibited by existing organisational and administrative practices. Explicit interdisciplinarity may even in some instances serve as a drawback when moving to collaborate or achieve funding for interdisciplinary projects. The Centre for Creative Practice and Cultural Economy (University of Technology Sydney) found that overt interdisciplinarity hindered efforts to achieve funding at the national level as evidenced in reports on the Australian national research funding body and its attitudes toward interdisciplinary endeavours (Anderson, 2012).

Institutional constraints to collaboration can result in tensions amongst academics, and university and research administrators, as scholars are continually encouraged to launch large-scale collaborations to bring in large grants while not being supported on the levels of service and infrastructure required. Additionally, external pressures adversely affect collaboration. For example, software licensing works well for within-institution collaboration and breaks down for cross-institution collaboration due to limits imposed by legal contracts. Other external pressures include the rules governing access and data-sharing for cloud and server services. When support is offered, institutions tend to provide clunky tools, compared to commercial offerings, that are often more of a burden to researchers rather than being of value. How can institutions begin to compete against private industry in the economy of scale when creating the kinds of collaborative tools researchers need? This is particularly important in light of businesses like Facebook and Google often providing more effective, efficient, malleable and scaleable services than academics can find in their own institutions, which lack the economies of scale of major technological firms. Institutional IT professionals are in a position to offer the centralised support necessary for collaboration in digital research. However, concerns of IT professionals may pose challenges to collaboration and should be resolved. Infrastructure, level(s) of user access, and security are challenges to centralised IT support for collaborative efforts.

## **A Need for Digital Curricula**

The research skills that social science students currently learn are often very traditional, with little attention to understanding appropriate digital methods increasingly available for research. While foundational training in classical research methodology is crucial in an increasingly digital research environment, there is a growing mismatch between the digital skills students are interested in receiving and

the ability of their instructors to impart them. The most basic of these digital skills is programming—a foreign skill to most social science researchers. Students are taught by many faculty who are not actively engaged in digital research or interested in employing digital methods. As digital research continues to grow in popularity and becomes more attractive to employers, social scientists must be able to give students the digital skills needed to make them employable and ready to do cutting-edge digital research.

Undergraduate courses in issues and topics related to digital research may be implemented in a variety of ways, journalism and media and communication being prime examples of programs that can implement digital-centric curricula. Furthermore, some digital skills are largely transdisciplinary, with many undergraduate students planning to move into the workforce after their undergraduate degrees benefiting from instruction in such things as Web-Analytics and instruction in the usage of APIs with a user interface. This technology-aware curriculum draws attention to the need for institutions and academics to help address issues related to digital literacy. It should also be noted that while the vision for such a curriculum often exists, visionaries may meet barriers to execution in the form of institutional resistance growing out of a lack of understanding of the utility of digital methods or simply disagreement with such an approach.

For postgraduate students, joint supervision may address many of the issues around digital skills training. However, there is currently no comprehensive model of how to achieve this, and academics may be resilient to joint supervision due to logistical and compensation issues in addition to traditional concerns along the lines of epistemology and research culture. These issues have not been thoroughly investigated and resistance to this approach may largely be attributed to existing mindsets. While the monetary disincentives for joint supervision are primarily at the inter-departmental level, at the individual level the main issues relates to the demands of time and attention placed on joint supervisors. Academics may actually place postgraduate students with research interests that benefit from joint supervision and interdisciplinary in general at a disadvantage in regard to research innovation, skills development and employability. Thus, there is a need for systematic and empirical analysis comparing students with different supervision and recruitment scenarios. The programs in Box 1 lists some of a growing number of programmes that seek to provide interdisciplinary training in digital research.

<b>Box 1. Interdisciplinary Curriculum/Supervision for Research: Examples</b>
<ul style="list-style-type: none"><li>• B.A. in Information Technology and Informatics (Rutgers University)</li><li>• MSc in Social Science of the Internet (Oxford Internet Institute, University of Oxford)</li><li>• Goldsmiths Digital Studios (University of London)</li></ul>



Convincing mid-career researchers who lack digital research skills to learn and adopt digital methodologies is integral to addressing issues around the student-instructor expertise mismatch (Box 2). Here, pedagogy is a primary concern. Traditional models such as instruction in how to use software is not an effective means of

inspiring mid-career researchers to engage with digital tools and methodologies. These individuals are motivated to learn about the possibilities of digital research by witnessing innovative application of digital tools and methodologies by colleagues. To some degree, capturing the attention of mid-career researchers becomes more motivational than instructional. Socially informed and substantive examples in domain areas of breakthrough research must be shared and supplemented with clear and accessible explanations.

#### Box 2. Supporting Early and Mid-Career Researchers

- Summer Doctoral Program (Oxford Internet Institute)
- Digital Methods Summer School (University of Amsterdam)
- ESRC Methods Festival (National Centre for Research Methods, UK)

## Moving Forward

This overview suggests a number of key issues for digital research that have been informed by a social science perspective, and a lively forum with practitioners and policymakers. This paper is intended to provide a snapshot of the fast-moving domain of digital research, forgoing strong and decisive conclusions about a field that is ever changing. By way of next steps, we invite an ongoing dialogue between social researchers studying the implications of digital research, practitioners and policymakers in order to collaboratively and creatively address these and other issues tied to digital research.

As previously mentioned, the OeSS project has concluded, but a number of follow-on projects are continuing its work. One major potential for continuity is emerging around research on 'big data', which ushers in a new stage for digital research that has been explored by initiatives around e-Research and e-Science over the past decade, and that can be informed by extending many of the issues identified in this report, including the ethics of digital research. We have indicated various ways that this work can be taken forward, and hope that we have not only provoked discussion, but convinced you to join this effort to identify and address issues of digital research from perspectives of the social sciences.

## **Appendix A. OeSS and the Social Shaping of Technology**

The social shaping of technology has become a broad umbrella term to cover a variety of theoretical and methodological perspectives in the social sciences (Dutton 2012). These perspectives can be brought to bear on the programme of funded projects in the UK that have focused in part on a technical initiative originally called e-social science and later known by the more general term digital social research. The OeSS project was the main UK initiative tasked with understanding the social aspects of both digital social research and digital research in general. This is in contrast to the bulk of the funded projects that were devoted to building tools and creating data and thus advancing research tasks in various domains. Continuing advances in digital research make it ever more important to adopt a social scientific framework when viewing applications of digital research and their implications. This paper seeks to identify the present opportunities and challenges of digital research from the social shaping perspective, illuminating diverse concerns integral to its sustainability and continued growth.

Early visions of the internet as a space for collaborative research underpin contemporary efforts to increase the networking and computational capacity available to researchers, endeavours undergirded by technological innovations such as Grid and Cloud computing. These efforts have led to a wide range of initiatives aimed at using advanced technologies for applications in support of research through collaboratories, e-Science, e-Social Science, computational social science, e-Humanities, e-Research, and digital scholarship—some of the many labels assigned to the new wave of initiatives around digital research (Nentwich 2003; Borgman 2007; Dutton and Jeffreys 2010). While these labels continue to be fluid, they are similar in their description of the uses of digital tools and data by collaborative teams via networks and across different sites.

Since 2005, following initiatives focused on e-Science, a group of academics sought to study the social shaping of digital research. This approach to inquiry around digital research emerged from the establishment of the National Centre for e-Social Science (NCeSS) by the ESRC at the University of Manchester in 2004. This centre was the hub of the first phase of research on e-social science. Most projects that were funded within this program focused on building applications or tools for conducting digital social research, which originally centred around initiatives for creating applications and digital infrastructure to handle the integration, management, or analysis of very large data sets—recently popularised as “big data.”

An exception was a project led by several colleagues at the University of Oxford that proposed to study the social shaping of e-science, not solely e-social science, rather than designing and building tools for conducting digital social research. There had been several pilot projects within the social shaping framework, however, the Oxford e-Social Science (OeSS) was the only social shaping initiative among 11 others which were linked to NCeSS. In 2009, the ESRC revised its strategy for fostering e-social science and replaced the NCeSS hub with a National Strategic Directorate for e-Social Science, which developed a digital social research strategy to foster an “...inter-agency approach to maximise the uptake, use and impact of new digital

technologies across the social science community.” The OeSS project has remained the only social shaping initiative.

The OeSS project was an interdisciplinary collaboration between the Oxford Internet Institute (OII), the Oxford eResearch Centre (OeRC) and the Institute for Science, Innovation and Society (INSIS; Saïd Business School), drawing on connections with the humanities, social and computer sciences and engineering across the University of Oxford. When the OeSS project started in 2005, the UK e-Science and e-Research programmes were in their infancy, particularly in relation to our understanding of the social implications and challenges of digital research. OeSS was phased over two periods. In phase one, it focused on identifying problems, which included issues of confidentiality, anonymity, privacy, data protection, ownership, accountability, and distributed collaboration. In the second and final phase, it moved to identify possible solutions, and how competing approaches to these issues further redefine research practices, and the quality and sustainability of digital research.

Since its inception, computer and natural scientists have had difficulty understanding the social shaping of technology, often inquiring as to the utility of this approach to the advancement of digital research. Despite conceptions of the social shaping of technology or science and technology studies (STS) being prominent for decades, there is no simple explanation. Donald MacKensie and Judy Wajcman (1985) developed this concept in the introduction to their widely used reader “The Social Shaping of Technology”. Since then, many scholars have sought to clarify and apply this concept to the study of any number of information and communication technologies, most notably Bruno Latour’s (1987) Actor Network Theory. Even within the social sciences, however, there are many approaches to studying not just the social implications, but also influencing policy and practices in digital research (Fry & Schroeder, 2010).

The difficulty in conveying these approaches arises for several reasons. First, non-social scientists often expect simple answers to questions about the social aspects of their projects. The question most often asked by computer scientists and engineers, who might imagine that the STS is a single unified perspective, is how people, mainly users in the case of Information Communication Technologies (ICTs), influence the success or failure of technologies. From this perspective, if social scientists could provide better insights about users, systems could be designed and used more effectively. However, social science does not limit any definition of the STS to users alone as they are only one set of actors among many that shape technical change. Any characterisation of the STS that only takes account of the individual or the user can be rejected as overly simplistic. Conventional social science definitions of technology include people and their practices, and are not limited to technical artefacts. They also include organisations, policymakers, and changes at various scales and reaches.

Secondly, there is not a simple answer across the social sciences. Compared to fields with a high degree of task certainty and mutual dependence (Whitley, 2000), the social sciences have a wide variety of perspectives, methods, and interpretive frameworks. Thus, when asked for their own characterisations of something like the social impact of a technology initiative, social scientists have a difficult time

conveying a unified perspective. One group might be interested in ethical issues, another in organisational structures, a third in micro-sociological interactions, and a fourth in thickly described ethnographic meanings. The OeSS project itself focuses, not just on STS, but an array of social scientific approaches such as the economics of innovation, ethics, law, science policy, information sciences and the study of digital research more generally. Any such explanations upon which social scientists could agree are inevitably too general of a level to provide meaningful ways forward. This is because the STS is not a unified perspective but an interlinking set of notions covering a number of alternative perspectives on the ways in which people—users, developers, policy-makers, managers, and more—shape technology, which itself is a social as well as a technical system.

Despite —or because of—this diversity, this report adopted a broad approach in a discussion of the social dynamics around the future of digital research. The overall aim of the OeSS project has been a continued to approach digital research as an object of social scientific inquiry. Here, we have specifically discussed the present challenges to digital research while proposing solutions, in addition to key areas of concern as this field continues to grow.

## Appendix B. Conference Attendees

Dr. Lawrence Ampofo Founder, Semantica Research Limited

Dr. Theresa Anderson Centre for Creative Practice and Cultural Economy, University of Technology, Sydney

Sheila Anderson Director, Centre for e-Research, King's College London

Dr. Smiljana Antonijevic Assistant Professor, Roskilde University, Royal Netherlands Academy of Arts and Sciences

Dr. Judit Bar-Ilan Professor of Information Science, Bar-Ilan University

Jessica Bland Policy Advisor, The Royal Society

Dr. Jeff Boase Assistant Professor, School of Professional Communication, Ryerson University

Dr. Christine Borgman Professor and Presidential Chair in Information Studies, University of California, Los Angeles

Dr. Dave de Roure Professor of e-Research, Oxford e-Research Centre and National Strategic Director for Digital Social Research, ESRC

Matthew Dovey Programme Director, Digital Infrastructure, JISC

Daphne Duin Research Associate, Organisation Sciences, Vrije Universiteit (VU) Amsterdam

Dr. William Dutton Professor of Internet Studies, Oxford Internet Institute, University of Oxford

Dr. Grace Eden Senior Research Associate, Oxford e-Research Centre

Dr. Marina Jirotko Director of the Centre for Requirements Engineering and Associate Director of the Oxford e-Research Centre

John Kaye Lead Curator in e-Social Sciences, British Library

Dr. David King Centre for Research in Computing, Open University

Dr. Eric Meyer Research Fellow, Oxford Internet Institute, University of Oxford

Dr. Stasa Milojevic Assistant Professor of Library and Information Science, Indiana University Bloomington

Dr. Michael Popov Open Anthropology Cooperative

Dr. Cornelius Puschmann Berlin School of Library and Information Science, Humboldt University Berlin

Sonja Rewhorn University of Chester

Dr. Ralph Schroeder Director of Research and Senior Research Fellow, Oxford Internet Institute, University of Oxford

Dr. Diane Sonnenwald Professor of Library and Information Studies, University College Dublin



Dr. Mike Thelwall	Professor of Information Science and leader of the Statistical Cybermetrics Research Group, University of Wolverhampton
Dr. Lesley Thompson	Chief Executive, Engineering and Physical Sciences Research Council
Dr. Peter van den Besselaar	Professor, Organisation Sciences, Vrije Universiteit (VU) Amsterdam
Dr. Sarah Werts	Economic and Social Research Council

## **Appendix C. A Glossary of Training and Funding Resources**

Economic and Social Research Council (ESRC)	<a href="http://www.esrc.ac.uk/">http://www.esrc.ac.uk/</a>
Engineering and Physical Sciences Research Council (EPSRC)	<a href="http://www.epsrc.ac.uk/">http://www.epsrc.ac.uk/</a>
Joint Information and Systems Commission (JISC)	<a href="http://www.jisc.ac.uk">http://www.jisc.ac.uk</a>
National Science Foundation (NSF)	<a href="http://www.nsf.gov/">http://www.nsf.gov/</a>
Oxford e-Social Science Project (OeSS)	<a href="http://microsites.oii.ox.ac.uk/oe&lt;br/&gt;ss/">http://microsites.oii.ox.ac.uk/oe ss/</a>
Oxford e-Research Centre (OERC)	<a href="http://www.oerc.ox.ac.uk/">http://www.oerc.ox.ac.uk/</a>
ESRC e-Social Science Programme (Digital Social Research; DSR)	<a href="http://www.digitalsocialresearch&lt;br/&gt;.net/wordpress/">http://www.digitalsocialresearch .net/wordpress/</a>
ESRC National Centre for Research Methods	<a href="http://www.ncrm.ac.uk/">http://www.ncrm.ac.uk/</a>
European Research Council	<a href="http://erc.europa.eu/">http://erc.europa.eu/</a>
Research Information Network (RIN)	<a href="http://www.rin.ac.uk/">http://www.rin.ac.uk/</a>

## References

- Anderson, T. D. (2012). Dealing with the human factors in digital research. Paper presented March 13, 2012 at Digital Social Research: A Forum for Policy and Practice.
- Borgman, C. L. (2012 forthcoming). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*.
- Borgman, C. L. (2007) *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*, Cambridge, Massachusetts: MIT Press.
- Cummings, J. & Kiesler, S., (2008) Who collaborates successfully? Prior experience reduces collaboration barriers in distributed interdisciplinary research. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work CSCW '08*. NY: ACM Press
- Cummings, J. & Kiesler, S. (2003) Coordination and success in multidisciplinary scientific collaborations. *Proceedings of the International Conference on Information Systems. ICIS 2003*, Dec. 14 – 17, Seattle.
- Dutton (2012) *The Social Shaping of Digital Research*. Paper presented March 13, 2012 at Digital Social Research: A Forum for Policy and Practice.
- Dutton, W. H., and Jeffreys, P. W. (2010) (eds) *World Wide Research: Reshaping the Sciences and Humanities*, Cambridge, MA: The MIT Press.
- Fry, J. & Schroeder, R. (2010) *Disciplinary Differences in eResearch*. In W. H. Dutton & P. Jeffreys (Eds.), *World Wide Research: Reshaping the Sciences and Humanities* (pp. 257-275). Cambridge, MA: The MIT Press.
- Hey, T. and Trefethen, A. (2003) *The Data Deluge: An e-Science Perspective*, in *Grid Computing: Making the Global Infrastructure a Reality* (eds F. Berman, G. Fox and T. Hey), John Wiley & Sons, Ltd, Chichester, UK
- Jirotko, M., Eden, G. & Stahl, B. (2012). Position Paper on Responsible Research and Innovation. Paper presented March 13, 2012 at Digital Social Research: A Forum for Policy and Practice.
- Latour, B. (1987), *Science in Action: How to Follow Scientists and Engineers Through Society*. Milton Keynes: Open University Press.
- MacKenzie, D., and Wajcman, J. (1985), *The Social Shaping of Technology: How a Refrigerator Got its Hum* (Milton Keynes: Open University Press).
- Nentwich, M. (2003) *Cyberscience: Research in the Age of the Internet*, Vienna, Austria: Austrian Academy of Sciences.
- Nentwich, M., and König, R. (2012), *Cyberscience 2.0: Research in the Age of Digital Social Networks*. Frankfurt, New York: Campus Verlag.

Whitley, R. (2000) *The Intellectual and Social Organization of the Sciences*. Oxford University Press: Oxford

Zimmer, M. (2010). 'But the data is already public: On the ethics of research in Facebook', *Ethics and Information Technology*, 12(4), 313-325.