

Society for Research Into Higher Education

Final Report

Open access data sharing policies: Implications for academic roles, practices and identities

**Dr Natasha Mauthner,
University of Aberdeen Business School
March 2013**

INTRODUCTION

Although the sharing of research data has been a commonplace practice, particularly amongst natural scientists, for well over half a century, only recently has it come to be seen as a matter for government regulation. From about the mid 1990s onwards, researchers carrying out publicly funded research have come under growing pressure to share their research data for the good of science and society. This data sharing imperative has been operationalized through a range of institutional mechanisms: data sharing policies adopted by funding agencies and scientific journals; research governance frameworks issued by Universities, professional associations, and funding bodies; new legislation (Freedom of Information); training and development in 'digital scholarship'; and investment in digital infrastructures necessary to support data sharing activities. These processes have been accompanied by discourses highlighting the economic, scientific, moral and political importance of digital data sharing as a new mode of science and research.

Academic communities have been responding in diverse and discipline-specific ways, reflecting variations in data types and data sharing cultures, traditions and practices. Some researchers are keen advocates of open data while others call for a more cautious approach to public release of raw data (e.g. Cauldfield et al. 2008). A common response has been support for the principle of open access data sharing but reluctance to put it into practice (Nelson 2009). The scientific, ethical, moral and political concerns raised by researchers suggest that data sharing imperatives may be reconfiguring their roles and responsibilities, research practices, and academic identities in critical ways that remain poorly understood. The purpose of this scoping study is to explore the emerging digital data sharing movement and its implications for academic roles, practices and identities.

SUMMARY OF ACTIVITIES AND EMERGING RESULTS

Objective 1: To undertake a review of recent data sharing policies and discourses relating to publicly-funded research undertaken by University-based researchers in the UK.

This objective has been achieved through a review of data sharing policies, guidelines and discourses issued by:

- government departments (e.g. Information Commissioner's Office)
- national and international science policy organisations (e.g. OECD 2007)
- universities (Manchester eResearch Centre)
- research funding agencies (ESRC, MRC, NERC, AHRC, RCUK, SHERPA, DDC)
- scientific journals (in the fields of Political Science, International Relations, Economics, Sociology, Psychology, Anthropology, Geography)
- data depositories (e.g. Inter-University Consortium for Political and Scientific Research; Economic and Social Data Service; ESDS Qualidata; Data Archiving and Networked Services; Council for European Social Science Data Archive)
- non departmental public bodies (e.g. Joint Infrastructures Systems Committee)
- advocacy groups (Digital Social Research; Alliance for Taxpayer Access)

Data sharing is being promoted as good practice and included within ethics and research governance frameworks, guidelines and good practice guides. The review has therefore also explored ethics and research governance frameworks issued by universities (e.g. Universities of Edinburgh and Oxford), funding agencies (ESRC Research Ethics Framework 2010), and professional associations (e.g. British Sociological Association; Social Research Association; British Psychological Society).

Research funders' data sharing policies

Research grant applicants are increasingly expected to specify data storage, access and management plans, and funding agencies are instituting measures to motivate researchers to share their data such as 'conditions being attached to funding schemes or ... services offered to recipients of funding' (Ruusalepp 2008: 3). Overall, funding agencies are encouraging (but increasingly requiring) their grant holders to share their data, with as few restrictions as possible, usually within a specified time frame after completion of the research (see SHERPA 2009; DDC 2010). In the UK, making research data available to other users is now a core part of the Research Councils' remit and RCUK's Common Principles on Data Policy. The UK's Economic and Social Research Council (ESRC), for example, requires its grant holders to make their data available for reuse (unless there are convincing reasons for not doing so), and attaches financial penalties to non-compliance (ESRC 2010a: 3-4). ESRC applications have a section on 'Data collection' where applicants have to justify the collection of new data and explain why existing datasets are inadequate for the proposed research. The application form also indicates that 'It is a requirement to offer data for archiving' and applicants are asked to include a statement on data sharing. Where

applicants believe that further sharing is not possible they are asked to present their argument to justify the case. (Data preparation and archiving costs can be included in an ESRC application). In 2011, the Engineering and Physical Sciences Research Council (EPSRC) announced its Policy Framework on Research Data concerning the management and provision of access to EPSRC-funded research data. The EPSRC now expects that the research institutions it awards research contracts to develop a clear roadmap to align their data sharing policies and processes with EPSRC's expectations by 1st May 2012; and be fully compliant with these expectations by 1st May 2015. The EPSRC will monitor progress and compliance on a case by case basis and investigate non-compliance. If it deems that proper sharing of research data is being obstructed, the EPSRC reserves the right to impose appropriate sanctions.

University data sharing policies

This move on the part of the EPSRC (along with other developments such as the Finch Report on Open Access publishing) is pushing universities to review their research data management practices, develop Research Data Management policies, and investigate the resource and infrastructure implications. Some universities have already introduced data preservation and sharing policies requiring their researchers to address, at the outset of their projects, the question of data management and sharing (e.g. Universities of Edinburgh¹ and Oxford²); and some have developed institutional data repositories in which academics and PhD students are encouraged to deposit their data (e.g. 'Edinburgh DataShare'). One of the ways in which these new requirements are being institutionalised by Universities is by defining data sharing as 'good research practice' and incorporating data management and sharing into University ethics and research governance regulations and procedures.

Journal data sharing policies

A final trend is the implementation of data archiving policies by scientific journals requiring authors to publicly archive supporting datasets as a condition of publication. This is a well-established tradition in fields such as crystallography (Arzberger et al. 2004) and molecular biology (Benson et al. 2008), and has recently been introduced by several key journals in the fields of evolution and ecology (Bruna 2010; Whitlock et al. 2010). Some journals in the social sciences are also now requiring or encouraging data sharing (e.g. *British Journal of Political Science*; *Journal of Peace Research*; *International Studies Quarterly*; *Economic and Social Review*; *Canadian journal of Economics*).

Ethical guidelines and regulations

¹ <http://www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt/data-sharing>

² <http://www.admin.ox.ac.uk/rdm/dmp/checklist/>

The shift in policy is leading to increasingly prescriptive ethical guidelines and regulations requiring that researchers privilege open over restricted access to their data. This is illustrated in the Economic and Social Research Council's (2010) Framework for Research Ethics (FRE), which sets out ESRC ethics approval requirements, and its views on good practice for all social science research. Whilst it is mandatory for ESRC-funded research, compliance with the ESRC FRE has become part of the condition of funding of other research councils funding social research (Jennings 2012). The ESRC FRE specifies that researchers who collect primary data "should be aware that the ESRC expects that others will also use it, so consent should be obtained on this basis and the original researcher must take into account the long-term use and preservation of data" (p. 24). Seeking informed consent to data sharing is therefore becoming mandatory practice, over and above obtaining consent for participation in the study and for using data in publications. While the ESRC recognises that in some cases it may not be possible to sufficiently anonymise data in order for it to be available in a data archive, ESRC funded researchers nevertheless are reminded that they have *a responsibility and obligation to seek informed consent for data sharing* from their research participants *regardless* of whether and how the data are later archived or shared. In other words, researchers' primary responsibility is to follow the ESRC FRE guidelines and now-normative expectation and practice of seeking informed consent to data sharing, rather than consider whether such guidelines are ethically appropriate to the particular research context: "Normally, social scientists should ensure that research participants are aware of and consent to arrangements made with regard to the management and security of data, the preservation of anonymity, and any risk that might arise during or beyond the project itself, and how these might be minimised or avoided" (ESRC 2010: 21). This is despite the fact that, elsewhere in the Framework document, the ESRC recognizes that digital data sharing and reuse is a form of research that involves 'more than minimal risk' (p 8-9) because of it may breach confidentiality agreements; make data disclosure more likely; increase the possibility of identification through data linkage; and lead to data use not anticipated in the initial consent to deposit.

The moral case for data sharing is built on the assumption and expectation that following the recommended and prescribed ethical guidelines will ensure sharing data is both ethical and legal. The ESRC refers researchers to the UK Data Archive for advice on securing consent for secondary use. The UKDA houses the UK's largest collection of digital research data in the social sciences and humanities. It provides best practice guidelines (see Van den Eynden et al 2011), and funding agencies refer researchers to the UKDA for advice and protocols on dealing with ethical issues in digital data sharing. The UK Data Archive Principles of Research Ethics (2011) identifies a set of key principles of research ethics that have a bearing on sharing or archiving confidential research data, and these include: (1) a duty of confidentiality towards informants and participants; (2) a duty to protect participants from harm, by not disclosing sensitive information; (3) a duty to treat participants as intelligent beings, able to make their own decisions on how the information they provide can be used, shared and made public (through informed consent); (4) a duty to inform participants how information and data obtained will be used, processed, shared, disposed of, prior to obtaining consent; (5) a duty to wider society to make available resources produced by researchers with public funds (data sharing required by research funders).

The UKDA recognizes that ethical and legal uncertainty over what, and what is not, permissible to share is inhibiting data sharing by researchers. Its guidelines are aimed towards alleviating these concerns and designed to reassure researchers that they can share their data ethically and legally if the recommended guidelines are followed: “A combination of gaining consent for data sharing, anonymising data to protect people’s identities and controlling access to data will enable sharing people-related research data - even sensitive ones - ethically and legally”. In cases where anonymisation would result in too much loss of data content, they advise regulating access to the data. This can be done in a number of ways: secondary users may need specific authorisation from the data owner to access data; confidential data can be embargoed for a given period of time; access can be restricted to approved researchers only; and restrictions can be placed on data downloading. There is a reminder to researchers, however, that “Restricting access to data should never be seen as the only way to protect confidentiality. Obtaining appropriate informed consent and anonymising data enable most data to be shared”.

The ESRC and UKDA guidelines configure researchers’ ethical and moral responsibilities precisely along lines that are explicitly rejected by the ethical guidelines issued by professional associations such as the Social Research Association (SRA, 2003), British Sociological Association (BSA, 2002) and Association of Social Anthropologists of the UK and Commonwealth (ASA, 2011). Where the ESRC and UKDA guidelines are authoritarian and prescriptive, these professional bodies aim to be descriptive and informative. They emphasise that: guidelines cannot resolve ethical dilemmas in a vacuum; the reputation of the discipline and its researchers depends less the ethical norms issued by professional bodies and more on the conduct of individual researchers; social researchers’ individual ethical judgments and decisions need to be informed by shared values and experience, rather than to be imposed by the profession; and researchers need to take ethical responsibility for their own practices. The ASA, for example, endeavours to help anthropologists resolve their dilemmas rather than impose a rigid set of rules backed by institutional sanctions.

The professional associations highlight researchers’ multiple obligations towards different constituencies: research participants; sponsors, funders and employers; co-investigators, colleagues and the discipline; and wider society. The guidelines emphasise that researchers are faced with competing duties, obligations and conflicts of interest; and they make clear that they do not seek to allocate greater priority to one set of principles over another, or privilege one group or set of interests over another. The SRA (2003: 15) for example notes that “Concern for individual rights needs to be balanced against the benefits to society that may accrue from research activity. Such ethical conflicts are inevitable. Above all, however, researchers should not automatically assume that their priorities are shared by society in general”. And the ASA notes that anthropologists have a responsibility to resolve ethical problems, insofar as it is possible, without harming the research participants or the scholarly community. They are urged to do their utmost to ensure they leave the field in a state which permits future access and use by other researchers. Again, these guidelines contrast with those issued by the ESRC and UKDA, which implicitly prioritise the interests of ‘the public’ (whatever that means) and the agencies that fund social research, and which are further reinforced by freedom of information legislation, its presumption in favour of data release and its privileging of the public interest. Data sharing mandates mean that researchers are entering the field *a priori* privileging

those whose interests are served by data sharing over those whose interests may be harmed. This position has significant ethical, moral, ontological and political consequences.

Objective 2: To critically analyze the legal, ethical, moral and epistemic frameworks underpinning data sharing policies and discourses.

This objective has been achieved in two ways. First, I have undertaken a critical analysis of the ontological, epistemological, moral and political frameworks underpinning data sharing policies, guidelines and discourses. I draw on performative understandings of scientific practice (e.g. Barad 2007, Law 2004, Law and Ury 2004), and their critique of classical Cartesian and Newtonian metaphysical assumptions, to argue that the data sharing movement is premised on a Cartesian or representational understanding of data, their production, and their use in the making of knowledge. This metaphysical, moral and political framework understands data as given entities that are separate from the contexts and relationships through which they are produced; excludes the labour of data producers and the moral ownership rights that arise from it; and privileges the rights of those who make economic rather than material investments in the production of data.

Second, I have conducted a review of legislation relevant to the disclosure and ownership of research data, copyright and intellectual property (Freedom Information Act 2000; Data Protection Act 1998; Copyright, Designs and Patents Act 1988). In particular, the introduction of freedom of information legislation in the UK (*Freedom of Information Act (2000)*, *Freedom of Information (Scotland) Act (2002)*; *Environmental Information Regulations 2004*, *Environmental Information (Scotland) Regulations 2004*) has been a further driver for the development of data sharing policies, as well as a means of enforcing data release into the public domain. These Acts, which came into force in 2005, provide the public with a right to access information held by a UK public authority (including Universities) and are designed to ensure accountability and good governance in public authorities (Rusbridge and Charlesworth 2010). The information requested can include research data and must be provided unless an exemption or exception allows an institution not to disclose it. The exemptions, which may be absolute or qualified, generally relate to considerations such as national security, law enforcement, commercial interests or data protection. The exemptions in Scotland differ in certain important respects from those in the remainder of the UK. This legislation means that researchers can now be legally forced to release their data, an issue I return to below.

This review of freedom of information legislation provides insights into shifts currently taking place in relation to legal and moral ownership of research data. Three issues are worth highlighting here: (1) The introduction of freedom of information legislation in the UK means that researchers can now be legally forced (by the Information Commissioner's Office via Universities) to release their data, effectively eroding the de facto personal ownership rights they have enjoyed over their research data in favour of public ownership. (2) There is a tension between data protection legislation (and associated policies which encouraged as good practice the destruction of research data, particularly personal data, once they were analysed and written up) and data sharing legislation. (3) The status of research data under copyright law is

complex, and depends in part on how data are defined (e.g. if data are viewed as ‘facts’ they cannot be copyrighted; if they are viewed as ‘literary works’ they can). Furthermore, while an individual piece of data (a fact) has no protection, collections of data may enjoy database rights, copyright, or both sets of rights, depending on circumstances.

Objective 3: To examine the implications of these policies, discourses and frameworks for researchers’ roles and responsibilities, research practices, and academic identities.

The study has investigated the regulation of data sharing and its potential unintended consequences in the following areas:

Unethical practices vis-à-vis research participants

The practice of seeking informed consent (IC) is regarded as one of the main ethical strategies for protecting the rights of research participants. Traditionally, IC is considered to be specific to a particular research study. This restricts how data can be used as it requires researchers to contact participants and seek their consent for every new research project. This has prompted a move towards ‘broad consent’ (BC), or ‘open consent’ which is being proposed as best practice for biobanks and others types of data repositories (Cambon-Thomsen et al 2006). While some medical and health science researchers support this move (Kozlakidis et al 2012), opponents view BC as a dilution of ethics which could destroy public trust (Cauldfield 2007). Informed consent has been criticized as inadequate for addressing ethical concerns because in practice it is satisfied by ‘largely ignorant blanket permission’ (Greely 2007: 361). Achieving the level of understanding that is required for truly informed consent in a data sharing context is difficult. Researchers can provide only limited information and reassurances about potential future uses and users of shared data (Greely 2007). Placing data within a public digital archive can mean that both researchers and respondents lose control over usage of the data. Indeed, retrospective discovery by research participants of unknown or unconsented applications of their data has led to legal disputes (O’Brien 2009). Furthermore, broad consent requires proxy decision-making by secondary data users and/or research ethics committees and transfers ethical decision making to third parties who may lack relevant knowledge or expertise. A further issue concerns national variations in ethical norms, guidelines and practices which may render data reuse across national borders unethical and illegal. For example, the Western tradition of informed consent operates according to the principle that the most specific consent is the best consent. When data are exchanged across national boundaries, there may be little scope for regulating whether data reuse conforms to the specific purposes respondents have consented to (Cambon-Thomsen 2007).

In the context of social research, the ESRC mandates researchers to seek their participants’ informed consent for data preservation and sharing. Here, the same question arises as to how ethical is it to seek participants’ ‘informed’ consent for future unknown uses of their data by unknown third parties (Richardson and Godfrey

2003; Mauthner 2012), particularly within the context of international research where informed consent and data archiving and sharing may have culturally-specific meanings (Morrow 2009). The UKDA guidelines, however, claim that participants are protected from any potential harm caused by data reuse by virtue of the End Use Licence that data reusers are required to sign, which “has contractual force in law, in which they agree to certain conditions, such as not to disseminate any identifying or confidential information on individuals, households or organisations; and not to use the data to attempt to obtain information relating specifically to an identifiable individual”. But while this *might* protect participants against possible identification, it does not protect them against the potentially morally harmful effects of seeing their story misunderstood, misrepresented or used for different purposes, which are perhaps some of the greatest moral harms that can be done to participants (e.g. Oakes 2002). As Bond (2012: 101) explains, the excessive fixation on harm avoidance overshadows other, more likely and subtle forms of harm such as damage to “intangible social assets such as identity, relationships or reputation, which derive their significance from the social context and are contingent on the perceptions of the people affected”. This challenges the assumption that uninvolved third parties, who are far removed from research generation contexts, can or should make ethical and moral judgements about potential harms to participants (Mauthner 2012).

While conventional wisdom is that anonymization of participant information, and the removal of obvious personal identifiers, provides effective privacy protection, recent evidence suggests this is not necessarily the case. Given that the protective power of anonymisation is weakened within a digital context where data linkage heightens the risk of identification (see Zimmer 2010; ESRC 2010a: 25), the ethics of following the recommended guidelines by anonymising narratives and telling respondents that this will protect their identity are questionable. Furthermore, in some cases respondents feel they lose ownership of their stories when these are anonymised and they want to be named and acknowledged in research outputs (Grinyer 2002). As Grinyer explains, participants will have different preferences and these issues can only be decided on a case-by-case basis.

The UKDA reminds researchers that, as well as a duty to seek consent for data sharing they also have a duty of confidentiality towards their participants. Under the new information sharing legislative context, however, providing assurances of confidentiality to our participants may be unethical given that these can be overridden. Under the Freedom of Information (FOI) legislation, if a person has provided consent for the disclosure of their information, a researcher (or the public institution that employs them) would be legally bound to disclose this information. Furthermore, such a breach of confidence on the part of the researcher or their institution would not be actionable and participants could not take legal action. This is because even though, under FOI, the exemption for information provided in confidence is an ‘absolute exemption’, a public interest test may still need to be applied. This is a complex legal matter which most researchers are probably not fully aware of, and are therefore unlikely to be explaining to their participants. It does mean, however, that promising confidentiality is a misleading and unethical practice in a data sharing context.

The recommended guidelines and procedures for securing ethical and legal data sharing can be seen as unethical because as researchers we can find ourselves making promises and assurances that are difficult if not impossible to keep when working in a digital data sharing context in which we have less control over the data and how they

are used. There is also the risk that even if researchers do not agree with the recommended practices, they will seek to comply with funders' requirements and expectations, through fear of jeopardising future funding. This in turn may lead researchers to subtly coerce their participants into agreements to have their data shared. This means that recommended guidelines may be more coercive than moral (see Haggerty 2004), and urges us to consider them within a wider context that takes into the account the politics of knowledge production.

The violation of trust between researchers and participants

The presumption in favour of data release privileges researchers' moral obligations and responsibilities towards the general public. In practice, however, researchers themselves tend to be more concerned about fulfilling their moral obligations and responsibilities towards respondents who are, after all, specific members of the public. Their practices are guided, not so much by universalist moral principles, but by contextual, situational, relational, contingent and practice-based ethical approaches (see Edwards and Mauthner 2002, 2012) such as the 'ethic of care' (Gilligan 1982). For example, researchers commonly uphold a moral imperative to honour relationships of trust they have developed with respondents who have entrusted them with personal information and confidences (whether DNA, medical histories, or personal narratives). They may use the data for the benefit of science and society, and to further their own careers, but many will do so only in a context where they feel they can safeguard respondents' moral interests. This is why many researchers want to retain some personal control over their data, and over respondent protection. The presumption in favour of data release, and its privileging of universal others, can be experienced as a violation of these personal and specific trust-based relationships and moral responsibilities. Indeed, such concerns were expressed by researchers in a recent FOI request in Scotland. In September 2011, the tobacco company Philip Morris International submitted a series of Freedom of Information requests to the University of Stirling to gain access to research data collected by a team of researchers over ten years exploring attitudes towards smoking amongst 6000 teenagers and young adults. The researchers who conducted the study feared that releasing confidentially shared data to the tobacco industry would breach the trust-based relationships that were established with the young people who took part in the studies (Christie 2011).³ Furthermore, there is an additional risk that breaching relationships with participants will in turn erode the perceived trustworthiness of researchers and deter the public from participation in research.

These risks tend not to be seen as legitimate concerns with data sharing because the normative ethical framework informing data sharing practices is focused on 'harm'. But as Jennings (2012) points out, perhaps the greatest risk of social science research lies not so much in its potential to inflict *harm* but rather in its capacity to *wrong* research participants: "The key risks involved are of breaching the trust, exploiting or

³ The University was reluctant to comply with the FOI request owing to concerns that the data would be used to market tobacco to young people, which could also have the effect of deterring sponsors. Scotland's Information Commissioner dismissed the University's case for failing to release the information and asked the university to respond to the demands made by Philip Morris (Christie 2011). My understanding is that the request was finally refused on the grounds of the cost of compliance (see <http://tinyurl.com/d95etvw>).

taking advantage of subjects. When you start researching people directly by accessing data about them, talking to them, and finding out about their lives, you enter into a relationship with them. That relationship engenders responsibilities just as other professional relationships do. To fail to meet those responsibilities is to wrong your research subjects even if it does not harm them in any physical sense and even if they never find out about it” (Jennings 2012: 90).

The ontological effects of data sharing mandates

Whether the question of data sharing arises before or after we have conducted our research and generated our data, either way data sharing mandates change the relationships of trust that we seek to establish with our respondents because they create new and specific moral conditions of possibility for participants’ storytelling. Seeking participants’ informed consent to share their data, for example, gives rise to different stories to those that might otherwise have been told because knowledge that their account might be lodged within a digital archive (consciously or unconsciously) may foster a ‘public’ rather than a ‘private’ narrative. Similarly, this might impact on a researcher’s interviewing style: on their willingness to share and expose personal experiences as part of building relationships of trust with their participants, knowing these may be shared with a broader audience. These changes *matter ontologically* because the relationships that we form with our participants are *constitutive of the narratives* that they tell us. Creating conditions that foster public narratives might, for example, render harder to reach (and tell) the muted, marginalised, and largely untold stories that challenge or trouble normative worldviews.

Another example of such ontological effects is the normative practice of anonymising archived data. In order to avoid the lengthy and resource-intensive process of anonymising data after they have been collected, the UKDA (2011) recommends that researchers keep some personal names and issues out of the interview conversation from the outset:

“Pre-planning and agreeing with participants during the consent process, on what may and may not be recorded or transcribed, can be a much more effective way of creating data that accurately represent the research process and the contribution of participants. For example, if an employer's name cannot be disclosed, it should be agreed in advance that it will not be mentioned during an interview. This is easier than spending time later removing it from a recording or transcript.”

Researchers, the UKDA suggests, might extend this to other personal identifiers: names of friends, relatives, places, institutions, etc. Like informed consent, asking informants to refrain from using personal identifiers in their narratives is an interview practice that gives rise to a specific relational context that in turn generates specific stories while excluding others. When we interview people, the degree of interest we express in the personal details and specificities of their lives and experiences, and the fact that we remember this information and use it throughout the interview to generate further questions and stories, has *ontological effects*. It plays an important part in building a particular kind of empathic, caring, understanding and responsive relationship, which in turn allows respondents to tell the *particular* stories they do, and not others.

A final example of such ontological effects is the advice being given to researchers to “avoid offering to keep confidential what is discussed between the researcher and the participant” so that they can share their data. This is the advice given by the Relu Data Support Service, which helps researchers of the ESRC Rural Economy and Land Use (Relu) Programme manage their data throughout the research lifecycle. With support from the UK Data Archive, it co-ordinates the archiving of Relu research data in established data centres to make them available for future research. (<http://relu.data-archive.ac.uk/data-sharing>) Accessed 14 September 2012).

The regulation of research, data sharing and research ethics is not neutral in its effects. It has ontological consequences because it changes our research methods and practices, our research relationships with research participants, and the stories that participants tell us about their lives. It makes a *material difference* not only to what we will discover about the social world; but perhaps more importantly to what will be systematically and constitutively excluded from and by our investigative practices (see also Stanley and Wise 2010; Haggerty 2004). Van den Hoonard (2001), for example, cites the case of a student who found that asking her participants for their written informed consent was intrusive and engendered an atmosphere of mistrust. Writing about one particular encounter she says: “I wondered just how different the interview might have gone had I not felt the need to be so formal with her”. This too points to the ontological effects of our practices.

The violation of researchers’ intellectual property rights

Prima facie, ownership of intellectual property is vested in the creator. However, when a creation is made by an employee in the course of employment, ownership may instead be vested in the employer. This applies to forms of intellectual property such as copyright (covering literary work, musical work, computer programs and other original works) and databases produced in the course of employment.⁴ In a research context, this means that Universities, as employers of researchers, have long had legal ownership of the research data that researchers produce. In practice, however, they have rarely exercised these rights and researchers have enjoyed moral ownership rights and control over their data. The introduction of freedom of information legislation is eroding these informal moral ownership regimes whilst strengthening legal and institutional ownership rights. This means that Universities are much more likely to, and can now more easily, exercise their ownership rights. This was illustrated in the high-profile case of Mike Baillie, a dendrochronologist from Queen’s University Belfast, who was forced to release tree-ring data under the Freedom of Information Act in April 2010. The Information Commissioner’s Office ruled that Queen’s University Belfast must release the data to the public because Baillie did all the work while employed at a public university. Baillie unsuccessfully contested this directive by claiming that the tree-ring data he had collected over a 40-year period were his own personal intellectual property (Baillie 2010). This is because freedom of information legislation stipulates that these intellectual property rights will not be infringed where an act is authorised by an Act of Parliament. Responding to a

⁴ The protection they afford is set out in the Copyright Design and Patents Act 1988 and the Copyright and Rights in Databases Regulations 1997.

freedom of information request is an act authorised by Parliament and so disclosures under freedom of information will not infringe intellectual property rights. In other words, FOI legislation can effectively over-ride intellectual property rights. The regulatory, policy and legislative changes of recent years therefore represent a significant moral and political shift in relation to the ownership and control of research data. The case of Mike Baillie illustrates how data sharing mandates have the potential to lead to coercive relationships between universities and their employees.

Power relations between and amongst researchers

The assumption that data are a shared resource also obscures the politics of knowledge production within research teams, and power relations amongst and between researchers (see Mauthner and Edwards 2007, 2010; Mauthner and Doucet 2008). In practice, most data are produced by junior researchers, PhD students and/or technicians who, because of their structural positions and/or career stage, may lack the time and resources that senior researchers have to make full and timely use of the data they generate. While their data collection efforts are usually, though not necessarily, recognized and rewarded within the team through, for example, joint publications, recognition of their labour is less likely within the context of open data sharing as currently constituted. Power and status differentials between team members risk leaving junior researchers open to exploitation as principal investigators have responsibility for making decisions about data sharing. Even where senior colleagues consult junior team members, by dint of their power and status they may privilege their own objectives. Within the context of international projects, there is a risk that data sharing becomes a form of scientific neo-colonialism. While turning data into open or common resource has the potential to provide postcolonial contexts with easy and cheap access to data generated elsewhere, they may lack the necessary scientific, technical, digital or cultural capital and resources to make full and speedy use of the data (see Luo and Olson 2008). In practice, it may be primarily well-resourced researchers and nations who stand to gain from the globalisation of research data, by reaping scientific and economic benefits and rewards from data generated by less well-resourced researchers and nations. From this perspective, the global data sharing project risks reproducing exploitative relations between nations, and between data users and data producers.

Objective 4: To disseminate the research findings through academic, policy, practice and media outlets.

Research publications

1. Mauthner, Natasha S. (In press). Should data sharing be regulated? In *The Ethics Rupture: Alternatives to formal Ethics Review in the Social Sciences*, ed. Will C. van den Hoonaard and Ann Hamilton. Toronto: University of Toronto Press.
2. Mauthner, Natasha S. and Parry, Odette (2013) Open access digital data sharing: Principles, policies and practices. *Social Epistemology*, 27:1, 47-67.

3. Mauthner, Natasha S. (2012) 'Accounting for our part of the entangled webs we weave': Ethical and moral issues in digital data sharing. *Second edition*. ed. Tina Miller, Maxime Birch, Melanie Mauthner, and Julie Jessop, 157-175. London: Sage Publications.
4. Mauthner, Natasha S. (2012) Are research data a 'common' resource? *Feminists @ Law*, 2(2): 1-22.
<http://journals.kent.ac.uk/index.php/feministsatlaw/article/view/60>.

Research presentations

1. 'Digital data sharing: Implications for academic roles, identities and practices'. Society for Research into Higher Education Annual Research Conference, 7-9 December 2011, Newport, Wales.
2. 'The power, politics and performativity of method: Feminist reflections on digital data sharing', Invited Plenary Presentation, Conference on Feminist Research Methods, University of East Anglia Business School, 9 December 2011.
3. 'Slow scholarship in a digital world', Invited Plenary Presentation, Doctoral Conference, School of Education and Social Work, Sussex University, 9 July 2012.
4. 'Should data sharing be regulated?', Invited Plenary Presentation, 'The Ethics Rupture: A Summit about Alternatives to Formal Research-Ethics Review', 26-28 October 2012, University of New Brunswick and Saint Thomas University, Fredericton, New Brunswick, Canada.
5. 'The performativity of Qualitative Longitudinal Research methods: Research governance, data sharing and the reconfiguration of research relationships', ESRC National Centre for Research Methods Network for Methodological Innovation on 'New Frontiers for Qualitative Longitudinal Methods' event on 'Research Relationships in Time', University of Cardiff, 7-8 February 2013

Four further dissemination activities are planned:

1. A paper to give at the World Social Science Forum conference on "Social transformation and the digital age", Montreal, Canada, 13-15 October 2013. Title of paper: 'Digital data sharing and its unintended consequences: A new materialist perspective'.
2. I have been asked to give a plenary address to a forthcoming conference on Secondary Data Analysis, organised by the Polish Academy of Science in Warsaw in early 2014.
3. A website – *Data Sharing in a Digital Age* - has been built and will be launched in the coming months.
4. An article will be published in the Aberdeen University Science Magazine.

Links, collaborations and research impact

In order to achieve and sustain research impact, I am engaged in the following outreach and networking activities.

- I used my research on open access data sharing in a response put forward by the University of Aberdeen's College of Arts and Social Sciences to the HEFCE consultation on open access and submissions to the REF post-2014. My work informed our response to the question of whether a requirement for open *data* (as well as open access *publications*) should be implemented in the next REF. The question was:

“We invite comment on whether respondents feel this is the appropriate approach or whether they feel that sufficient progress has in fact been made to implement a requirement for open data as well. We will consider any representations that such a requirement may reasonably now be developed but would also need advice on how this might be achieved.

The response put forward was:

“We are entirely opposed to incorporating blanket requirements for open data into the REF. There are circumstances in which it is desirable to share ‘data’ – particularly those where the collation of ‘data’ involves the construction of new quantitative datasets that might be usefully re-analysed. However, making all data (in the broadest sense of all of the information underpinning research) available is neither possible nor desirable across the range of disciplines encompassed within our College. Frequently such a requirement would jeopardise the production of valuable research.

In the case of human subject research, the ethical complexities and sensitivities involved in releasing data render open data mandates inappropriate because they may expose human subjects to potential ethical harm. One recent (2011) example of this comes from the University of Stirling, where researchers expressed concern that releasing confidential data about attitudes towards smoking to a tobacco company – under a Freedom of Information request - would breach the trust-based relationships that were established with the young people who took part in their research. A further example comes from the United States where Boston College is currently fighting a legal battle with the government which wants it to hand over oral history interviews conducted with former IRA paramilitaries in 2001 who were assured that the interviews would not be released until after their deaths. Such cases underline the ethical risks posed by data sharing mandates.

In humanities disciplines where qualitative archival research dominates, the concept of the ‘dataset’ has little meaning and ‘data’ in fact comprises notes, scans, photocopies from a plethora of sources – published, archival, statistical, and the like. A host of copyright restrictions may apply to such ‘data’. For example, archives frequently permit copying by individual researchers but prohibit reproduction and publication. Such copying is of particular benefit to those working in inaccessible or distant archives, vastly reducing the costs of archival trips. Similarly, for art historical research the images analysed are the property of galleries and individuals – and could not be made freely available (except perhaps at prohibitive cost). Finally, for virtually all humanities and social sciences, the process of research is cumulative across an entire career. A scholar approaching retirement will often still draw on information (data) collated as a graduate. The practices of publication in humanities disciplines

frequently require such evidence (or data) as is necessary to establish a scholar's argument to be presented within the body of their published work. In most humanities, Open Access to 'Data' is neither meaningful, practical, nor useful at present, nor is it likely to become so in the future.

Open Data should not be a blanket requirement in the REF. This is an area in which one size can never fit all."

- Links have been forged with the UK Data Archive (and senior data archivist) through participation in a workshop on *How to set up and run a data service: the challenges of social science data*, 13-14 October 2011, UK Data Archive, University of Essex.
- I am a core member of the ESRC National Centre for Research Methods Network for Methodological Innovation project on '*New Frontiers for Qualitative Longitudinal Methods*', which was launched in November 2012. The network brings together academics with expertise in Qualitative Longitudinal Research - the preservation and sharing of qualitative data over time - along with a new generation of researchers for whom temporal perspectives enable a range of interdisciplinary and emergent research practices. I am participating in the five events being hosted by institutions across mainland Britain (Southampton, Cardiff, Manchester, Birkbeck College, Sussex), and was invited to present my work on digital data sharing at the second event held at Cardiff University. The network brings together a multidisciplinary and international group of established, early career and doctoral researchers.
- In October 2012 I was invited to give a plenary presentation at the first Ethics Rupture Summit held in Fredericton, New Brunswick, Canada in October 2012. This was a major gathering of 40 experts in research ethics from Australia, Brazil, Canada, Italy, New Zealand, the United Kingdom and the United States, committed to enhancing ethical research practice and developing innovative alternatives to the regulation of research ethics that might achieve this. The summit resulted in the *New Brunswick Declaration: A Declaration on Research Ethics, Integrity and Governance*. It was issued on 4 February 2013 and has since been adopted by the UK's Social Research Association as part of its ethical guidelines for social researchers, and by the University of Sheffield's School of Education. A second summit is planned for 2014 in New Zealand.
- In March 2012 I became a founding member of the University of Aberdeen's Research Data Management Group, which together the University Librarian and Director, the Vice-Principals for Research and Knowledge Exchange, Head of Infrastructure Management, and a University Senior Policy Advisor to develop a researcher (rather than research funder or institutional) centred policy on open access to research data.