Zoran Slavnic

Preservation and Sharing of Qualitative Data - Academic Debate and Policy Developments
Preservation and Sharing of Qualitative Data – Academic Debate and Policy Developments

Zoran Slavnic

In most advanced economies the digital archiving of quantitative data is as old as digital technology itself but the idea of digital archiving of qualitative data is of quite recent date¹. Two interrelated events are relevant in this context. The first is the establishment of the Qualidata Centre in the Department of Sociology at the University of Essex in the UK in 1994. The initial idea was to preserve data from pioneering examples of social research: in particular empirical data from the classical sociological studies of John Goldthorpe, Peter Townsend and Stan Cohen (Cheshire et al 2009). The second related event happened in 1996, when the UK’s largest provider of funds for social and economic research, the Economic and Social Research Council (ESRC) imposed new requirements on its grant-holders, namely to consider the issues of preservation and sharing of empirical data provided by their research projects (Mauthner and Parry, 2009). According to the agreement of 1996, between the ESRC, Qualidata Centre and the UK Data Archive (UKDA), Qualidata was appointed to provide a specialist archiving service for the UKDA, while at the same time the ESCR started to impose a requirement on all their award-holders to deposit copies of their qualitative data with Qualidata.

In other countries the issue of preserving and re-using qualitative data began to attract serious attention in the mid-2000s. This interest was initiated by the OECD’s Declaration on Access to Research Data from Public Funding, 2004 and the OECD’s Principles and Guidelines for Access to Research Data from Public Funding, 2006. According to these documents, the most important advantages of ready access to research data are: improving the transparency of research processes; recognizing the fact that products of publically funded research are public property; avoiding unnecessary duplication of field work and the burden on research participants; and making data available for other researchers (Fry et al. 2008, Bishop, 2009).

¹ The trend of archiving qualitative data has its historical precedents, among others, in the British Mass Observation Project started in 1937 (Cheshire et al, 2009; Mauthner and Parry, 2009), the Henry A. Murray Research Archive at Harvard University, collecting and archiving qualitative data since 1976 (Cheshire et al, 2009) and the Australian PARADISEC, containing songs, languages and dances and cultural rituals from the 1950s and 1960s (Cheshire et al 2009). This paper is however concerned in the first place with developments that have been taking place during the last 15 years or so.
Despite all these arguments the idea about archiving and open access to research data is viewed among some of the actors involved as problematic. While within quantitative research communities, data archiving and re-using are mainly perceived as trouble free, many qualitative researchers are skeptical. Most are generally reluctant to deposit their empirical data for sharing and re-use. This situation has led to an academic debate, primarily among British qualitative researchers so far, which is not surprising bearing in mind that the archiving policy was introduced in the UK approximately a decade earlier than elsewhere.

In the first part of this paper I will present and comment on the epistemological/methodological, ethical/legal, ideological/political, and practical/technical aspects of the ongoing British debate about data archiving and re-use. In the second part of the paper, the Swedish case will be briefly described and considered in light of the academic concerns that have been raised by the British debate.

**Epistemological/methodological issues**

*Context, researcher-participant relationship and misrepresentation of data*

According to skeptics of data archiving and sharing, the actual data preservation and sharing discourses and policies rest almost exclusively on so-called foundational epistemology, which is, as argued by Mauthner and Parry (2009), characterized by two essential properties. First, it is based on a positivist philosophy of knowledge that views the social world as part of an objective external material world that exists independently from the subjects who observe it. Objective knowledge about this world is, according to this view, possible if objective scientific methods are applied and if subjective interference by observers (researchers) is kept away.

Secondly, foundational epistemology tends to understand itself as the only right perspective on knowledge, denying at the same time any legitimacy to other epistemological perspectives. As such, foundational epistemology tends to ignore some of the most basic principles upon which that discipline is founded.

One of these is that qualitative research is both derived from and dependent on the relationship between a researcher and his/her informant(s). In other words, qualitative data is socially constructed through interpersonal relations between researcher and respondents and it is never just “out there” waiting to be collected, separate from the interpersonal, social, cultural and political context in which it is supposed to be collected (Hammersley, 1997; Mauthner et al, 1998). Besides, the constructed nature of qualitative data requires the researcher to demonstrate a high level of self-awareness, or reflexivity, while handling and analyzing the data. To be able to make any kind of sense of the data, the researcher needs permanently to evaluate and re-evaluate her/his own role in producing the data (Mauthner et al. 1998; Mauthner and Doucet, 2003). These epistemological principles are however incompatible not only with the foundational view of empirical data as discrete entities, independent of the specific conditions it
has been collected in, but also with the foundational claim that there is only one social reality – “epistemic monism”, denying at the same time all other realities - “epistemic pluralism” (Mauthner and Parry, 2009).

At the same time the constructed nature of qualitative data and its dependence on the context of its production make each particular data set unique, which in its turn puts limits on the possibilities of (re)using anyone else’s data. Data produced by different researchers, with different cultural backgrounds and different theoretical perspectives, “cannot be treated as if they represent a common currency” (Hammersley, 1997:139). According to Mauthner and Doucet (2008) they cannot be used as equivalent currency even when the data have been generated within the same research team, working within the same research project, if the division of labour includes a separation of data collection and data analysis. Finally, re-using one’s own data can also be problematic, if data are to be used for a purpose different from the original one. This is because “…‘findings’ are not in the data but created through the interaction of a particular ... researcher with particular respondents in particular locations and at particular historical junctures...the meaning [of data] is made rather than found” (Mauthner et al, 1998:735)

In sum, according to skeptics, data are viewed within the foundational discourse, as separated from both the researchers and from the contexts that made the generation of data possible. Such data, taken out of the context of its production and away from the “original” researcher do not have the same epistemological and methodological legitimacy when re-used by others. Furthermore, researchers who approach the data without first-hand knowledge of the context of their production are not able to reach the real meaning of the data, and even run the risk of misrepresenting them i.e. using them in a way that may result in completely different conclusions, compared with the initial research intentions.

Let me now briefly summarize the most relevant arguments put forward by proponents of archiving and re-using of qualitative data. According to them, most of the skeptics give “almost obsessive attention to the context” in re-using qualitative data (Moore, 2007; Walters, 2009). In reality however, qualitative researchers practically never draw their conclusions exclusively on the narrow ground of the concrete research context. In one way or other they need to relate the meaning created in the particular research to the wider social context(s). As argued by Walters (2009), in the process of interpretation, the researcher is supposed to explain why his/her data are relevant in a wider social context and in which way the data help us understand these broader social structures and processes. If micro-context is in this way present in the macro-context, then the macro-context is also present in the micro-context. This also brings into question the unconditional privilege that the original researcher, as claimed by skeptics, has in relation to the micro-context. Thus, according to this argument, it is not only possible for the original data to be revisited by other researchers, but also “[t]he ability to revisit qualitative data in light of social change may allow the future researcher to at-
tribute those participants with a degree of presence about future social conditions that the original researcher was in no position to understand” (Walters, 2009:312).

The criticism that secondary data analysis ignores the importance of reflexivity in qualitative research is, according to Fielding (2004) also badly founded. Reflexivity cannot be an exclusive privilege of those who do primary data analysis. On the contrary, reflexivity has always been an essential part of all qualitative research and there is no logical incompatibility between reflexive exercises in primary or secondary data analysis. The lack of some essential contextual information about the data, of course, makes trouble for researchers doing secondary data analysis, but the same is often true even for the original researcher. Sometimes we simply cannot find the evidence for a given analytic point even in our own data. The problem is therefore not epistemological, but practical (Fielding, 2004:99).

Even the argument about the constructed nature of qualitative empirical data put forward by skeptics is, according to Van den Berg (2005) one-sided and overestimated. He basically agrees with skeptics that the naïve positivist concept of the objective and independent nature of empirical data, waiting “out there” to be observed and collected, needs to be rejected, but this does not mean that data are simply a reflection of researchers’ theoretical and methodological backgrounds. “The ‘empirical’ is not just ... a dress up of the ‘theoretical’” (Van den Berg, 2005:5). Empirical poses its own value or meaning that is often unpredictable, unexpected and surprising for the researcher.

Regarding the researcher-participant relationship, Bishop (2009) argues that although much of the qualitative research without a doubt depends on close/personal/unique relationships between researchers and their informants, it is certainly not true for all qualitative research. It is possible to do first rate qualitative research (for instance about right-wing extremism) without sympathizing or empathizing with research participants. Another issue addressed by Bishop (2009:264) is the role of the participant in the re-use of data. It is of course not possible for a participant to have a direct role in the secondary analysis of data, but this argument, according to Bishop (2009) cannot be used for criticising archiving and re-use, because it is taken for granted that the participant normally would have a role in data analysis. In fact however, this issue is far from being so clear. There is an ongoing academic debate on participants’ roles in the data analysis process, with arguments both pro et contra. Skeptics of data archiving and re-use prejudge the outcomes of this debate, taking for granted that participating in data analysis is necessary for all qualitative research.

Finally, as regards the criticism that original data in the context of re-using may be misrepresented, the skeptics’ argument is that the original research situation creates “intuitive”, “organic”, “intimate”, and “personal” relationships between the researcher and his/her data. Such understanding represents a kind of mystification of qualitative research. (Bishop, 2009). Scientific argument is and should be built on rational scientific procedure and not on the intuitive, organic and personal relation with the empirical data. In this context, conclusions derived from secondary analysis need not to be disadvantaged in relation to those derived from the original analysis (Bishop, 2009:266).
What is there to say about the above debate? In my view, the qualitative research community should pay full attention and show due respect to the arguments repeatedly presented by skeptics about data archiving. As we are going to show later on in this text, actual data preservation- and sharing discourses and policies are really dangerously characterized by foundational epistemology and based on a neo-positivistic paradigm, treating the qualitative and quantitative data in the same way. This fact however should not and cannot justify the lack of nuance in actual academic debate on these issues, which is still basically an all-or-nothing debate. Those in favour are for instance right in pointing out the epistemological problems of context and relationship between researcher and research participant. But things are not as simple as that. The “original” researcher of course has potentially a better insight into the context of producing the knowledge in a concrete interview situation, than the re-user of the data. But the final outcome always relies on the creative ability of the researchers (either the “original” or the re-user) to do something valuable with the data. “Being there” is not always a guarantee of good research outcomes. A bad researcher can ruin even the best “original” data, while at the same time a good researcher can produce high quality research outcomes by merely re-using the data. Furthermore, the quality of outcome depends not only on who is doing the analysis (“original” researcher or re-user) but also on how well archived data are documented and prepared for re-using (which is, for the most part, also a job that is supposed to be done by the “original” researcher).

In this context it is important to say something about the problems besetting the difference between “data” and “knowledge” that characterize Mauthner, Parry and their colleagues’ contributions to the debate. They tend to confuse the processes of “data production” and “knowledge production” (compare for instance Mauthner et al, 1998; and Mauthner and Parry, 2009). They certainly are right in emphasizing the importance of not separating data collection from its analysis (Mauthner and Doucet, 2008:973). As Walters (2009) pointed out, however, this is only micro-context, which is of course important, but not crucial to knowledge production. Macro-context is at least as important as micro-context. Data become knowledge only when they are placed in and related to the broader, existing knowledge in a given research field.

The other relevant issue is related to relational context, i.e. uniqueness of the relationship between the researcher and his/her informant. According to the contras, without this “original” insight into these relations, is not possible to reach the real meaning of data. These relations are surely very important in qualitative research, but if the relationship is too close the researcher may become blind to other interpretations of the data, and/or even fail to discover their real meaning. What she/he sees as the real meaning of the data may in fact be only a fata morgana. Re-users of data however may be less exposed to such risks. At the same time, Bishop (2009) is correct when she argues that high quality qualitative analysis is possible even without any close, sympathizing or empathizing relationships with the informants.

Another important circumstance that may help a re-user of data to get closer to the real meaning is when texts, already written and published using original data, do not
have an up-to-date focus on the debate. Academic excellence is in fact nothing more than the ability to reveal, understand and possibly explain, in the form of published texts the “real meaning of data”, based on the “original“ understanding of the situation in the field. By reading a relevant publication, a re-user of data is not only in a position to “reach” the real meaning of the data (certainly in a better way than otherwise), but also to re-examine critically the published text based on the data. This possibility should be welcomed by the academic community as a way of securing the scientific quality of research, but obviously many individual scholars are reluctant to open up to the possibility of this becoming an everyday academic routine. This reluctance is, according to Cheshire (2009) not a sign of some general disinclination among researchers to expose their research to critical scrutiny, but rather the result of the widespread academic culture within the qualitative community which sees qualitative research “... as a personal endeavour – largely due to the significant personal and emotional involvement of the researcher in the construction of data ... – and a corresponding fear that others may scrutinise something so personal and declare it inferior” (Cheshire, 2009:38).

Parry and Mauthner (2004:145) point out another aspect of this issue. Being forced to expose their field work to external scrutiny, researchers may become more careful and formal and less free and spontaneous, which may negatively affect the quality of interviews and research in general.

In any case, the academic practice of data sharing and re-use is still not part of widespread academic routines. Full official statistics about the situation in the UK are not available. What is known is that Qalidata has archived 100 data sets, which have increased to 450 records by the year 2000 (Parry and Mauthner, 2005). What is not known however is how often the archive had been used and how many publications resulted from this secondary analysis (Parry and Mauthner, 2005:338). Obviously researchers are still not especially enthusiastic about re-using others’ data (Mauthner and Parry, 2009), and if they re-use qualitative data, then it is most often their own (Heaton, 2004).

What we also know is that there is a general reluctance among researchers to deposit their qualitative data for re-use, which is going to be the topic of the following sections of this paper.

**Ethical and legal issues**

*Copyright*

The copyright pertaining to qualitative data is currently rather unclear. In British legislation for instance, different kinds of copyright apply to any qualitative interview that is recorded (Parry and Mauthner 2004). Copyright regarding what was said during the interview belongs to the informant but copyright in what was recorded belongs to the institution that arranged the research project in question. The researcher, in fact,
does not have a legal right to his/her data, despite the fact that producing qualitative data is a joint venture between researcher and respondent (Parry and Mauthner 2004). In practice however, institutions, which formally own the copyright of the qualitative data, very seldom use these rights.

This has made possible a situation in which researchers traditionally perceive themselves, and have been perceived by others as exclusive owners and protectors of empirical data that they have gathered (ibid.). Not only in the UK is this the case. In the Swedish Science Council’s guidelines for Good Research Practice (Gustafsson et al. 2005:33) the situation is summarized as follows:

In many departments, researchers regard their source data as their personal property. That view, though, has to be questioned. The work of collecting the data has as a rule been done as part of a contract of employment, often with public funding; society may thus have invested major sums in the study in question. The value of material of this kind may be considerable, for later generations as well, and perhaps for reasons which cannot currently be foreseen.

At the same time, a national inventory survey (conducted in 2008) which studied attitudes towards data sharing among Swedish social scientists showed that only 11% of all data at the universities investigated is archived, and almost 50% of all professors interviewed stated that archiving is generally very unusual or never happens (Carlhed and Alfredsson, 2009; Axelsson and Carlhed, 2009). Moreover according the same study, many Swedish researchers take with them their data sets when they move to another university or research centre.

In sum, there are at least two kinds of inconsistency here. One is between researchers’ formal/legal position in relation to their research data (which is a problem in itself and needs a separate debate) and the actual situation where most researchers have de facto control over the data. The second inconsistency is between the general requirement for archiving the research data on the one hand and the equally general practice among researchers not to do so. This is an obstacle to further development of the infrastructure for qualitative data preservation and sharing. Another consequence of these contradictions, especially in situations when formal pressure on researchers (to obey rules) is amplified, is increased frustration and resistance within the research community. This will be discussed in more detail below.

Confidentiality

Protecting the identity of respondents in qualitative research and securing their anonymity may be rather problematic if data are to be archived and available to others. In the Swedish survey mentioned above, 65% of researchers stated that threats to confidentiality prevent them from depositing their data in archives (Carlhed and Alfredsson, 2009). A similar survey conducted in Finland in 2006 showed that 66% of researchers viewed the lack of informed consent as the biggest obstacle to depositing their data,
while 48% of them stated the confidentiality and research ethics as a reason for not doing so (Kuula and Borg, 2008).

Two possible solutions for these problems have been proposed by researchers who are in favor of data archiving and sharing. One is removing the identifying characteristics of informants. The other is falsifying unnecessary information about the informant in order to protect them.

Skeptics however argue that the first proposal, i.e. taking away the identifying characteristics of the informant, should in fact meet anonymity requirements, but on the other hand would seriously damage the integrity of data and under certain circumstances might even change their meaning (Parry and Mauthner, 2004). The other measure, namely falsifying the less important details about informants, is also problematic, because what is considered as less significant and important to the original analysis, may obtain a new significance and new meaning for other researchers with different research perspectives (Parry and Mauthner, 2004:144).

Bishop’’s (2009) criticism of this matter focuses in the extreme on the participants’ rights, as compared with the rights of other legitimate agents. There is also an obsessive focus among many researchers on their own right to protect their participants, while neglecting their duties towards other agents. Bishop argues that researchers have duties of transparency and respect for the research, academic and professional standards to their research colleagues, or duties of openness and effective use of public funds to the public.

I personally could not support Bishop’s sense of balance between protecting participant’s rights and duties towards other relevant agents. On my ethical apothecary’s scale, my informants’ rights would always outweigh the rights of all other agents. But Parry’s and Mauthner’s (and perhaps the majority of the other qualitative researchers’) somewhat rigid aim to protect the identity of participants by securing their anonymity in this context is also problematic. Most of the researchers take it for granted that their informants not only “deserve protection of their anonymity, but also desire it” (Grinyer, 2002). But it is more complicated than this. First, it is not true that most of the participants insist on anonymity. Grinyer (2002) has shown that even in the most ethically sensitive research projects, participants do not have anything against being referred to by their real names. In one such qualitative study of parents of young adults (18-25) who had been diagnosed with cancer, only 7 out of 30 participants wanted to have pseudonyms. Secondly, full anonymity may violate or prevent participants from using their other rights, for instance the copyright of their own interviews.

**Informed consent**

The principle of “informed consent” is one of the most important ethical principles in the context of gathering empirical data in social sciences. The researcher should normally inform his/her informants about the purpose of the research project, who is conducting and financing it and how the data provided are going to be used. The problem
with archiving and re-using the data is that the future purpose and use of data cannot be covered by the original “promise” given and therefore the secondary use of data cannot be covered by the original consent (Parry and Mauthner, 2004:146).

This criticism has been met by the argument that consent can never be fully informed, because of the nature of qualitative research, with an infinite variety of analytical paths that are impossible to predict in advance. That is why, according to some scholars (see for example Lawton, 2001), the traditional notion of consent should be replaced with something called “process consent”.

There is a special issue in this debate, namely what to do with those data-sets where researchers have neither discussed, nor got formal consent from their informants for archiving and re-using. There are two important aspects in this context that need to be discussed. First, there are many researchers who base their requests for consent on the promise that no one but the researcher in question is going to have access to the data. Such a promise then becomes one of the most important obstacles to archiving and sharing the data. The point is that such a promise is in conflict with good research practice. The Swedish guidelines for good research practice (Gustafsson et al. 2005:33) state:

It is important, therefore, to handle source data with great care and to retain and archive them in such a way that they can be made available to researchers other than the ones who gathered them. In the shorter term, this is important in order to allow published results to be verified following publication, for example to trace sources of error or to enable the researcher to counter accusations of research fraud. Investigators therefore cannot promise that no other researchers outside the group collecting the source data will ever, under any circumstances, be given access to those data...

The second relevant aspect is that many researchers are reluctant to contact their informants again in order to get consent to do this, because they take it for granted that the answers would be negative (Bishop, 2009). There are examples however that show quite the opposite (Corti et al. 2000; Kuula and Berg, 2008). In an example described in Kuula and Berg’s report, the Finnish Data Archive contacted a number of informants in order to re-gain consent. They did so on behalf of original researchers who themselves were skeptical about the whole thing. The result was that out of 169 informants who were contacted, 14 did not answer, 4 were negative, and as many as 151 or 89% were in favor of archiving and re-using the data.

Some of the problems related to confidentiality, informed consent and in general protection of the participants, as Cheshire (2009) proposes, may be resolved by imposing an embargo on the data in the short or medium term. In this case other methods of protecting the participants such as falsifying or removing the identifying details from the material would not be needed. Data in their original form would be accessible to future generations of researchers. As Walters (2009) reminds us, the most important scientific value of archived data becomes visible only when we put it in a broader historical perspective. Indeed, for us contemporary scholars, it would be scientifically and
intellectually quite exciting to have access to original qualitative data about some social processes or about some social groups for the time, say, between the two world wars. The sole but still important negative aspect of Cheshire’s proposal is that contemporary researchers would not be able to have access to data.

A final point that needs to be put forward in this section, however, is that ethical and legal issues discussed here cannot be resolved by presenting only ethical and legal arguments. I agree with Heaton (2004:86) who pleads for more research on these matters, not only exploring the researchers’ views, but also informants’ views on the conditions in which qualitative data can be re-used.

Ideological and political issues

Some scholars suggest that researchers are reluctant to deposit their data partly because of the feeling that their academic freedom is gradually but systematically being restricted by non-academic institutions according to non-academic criteria and values (Travers, 2009). It has already been mentioned in this paper that the qualitative data archiving movement in UK was initiated by the ESRC’s requirements for depositing data for archiving and sharing by other researchers (Mauthner and Parry, 2009). The ESRC has itself however a historical background that makes many social researchers in the UK treat it suspiciously. Due to the political pressure of the Thatcher government, the ESRC replaced the Social Science Research Council at the beginning of the 1980s. Griffin (1997:44-45) describes these changes as follows:

“Thatcherite logic dictated that most social scientists were seen … as Marxists, communists and dangerous left-wing agitators, so the ESRC gained a strong(er) emphasis on economics rather than sociology in order to purge itself of such unacceptable influences…The ESRC’s mission statement includes the aim of supporting research in the social sciences which will contribute towards ‘the economic competitiveness of the UK, the effectiveness of public services and policy, and quality of life’…”

This skepticism towards the ESRC may certainly be viewed as an internal issue affecting only UK social scientists. Nevertheless, some other, more general, trends are also relevant. Funding has been increasingly removed from universities to National Research Funds, applications for funding need to be adjusted to more or less non-academic requirements such as policy relevance, using methods and language that non-academic readers can understand.

Apart from these general troubles, scholars dealing with qualitative research have increasingly been affected by some specific problems. This dissatisfaction with the status of qualitative research is in my view best summarized by Denzin (2009:139-140):

“Like an elephant in the living room, the evidence-based model is an intruder, whose presence can no longer be ignored. Within the global audit culture propos-
These external threats, however, have still not led the qualitative research community to create its own standards of quality, its own criteria, but rather to growing frustration and individual strategies of avoidance and resistance to depositing their findings in databases. A common strategy deployed by British researchers is to encourage their informants not to give their consent for the interviews to be archived (Moore, 2006).

Indeed, this story about everyday troubles that many social scientists, not only in UK, are forced to deal with sounds so familiar. More high quality research, more vibrant academic debate and, not least, more organized and coordinated political action related to all issues mentioned in Denzin’s diagnosis above, are necessary. But what has all this to do with the methodological question of whether and under what circumstances it is possible to develop and use qualitative databases. These two things, our struggle for academic freedom and intellectual independence on the one hand, and our academic debate on possibilities and limitations of data archiving and re-using on the other, need to be kept separate. Instead of refusing to deposit our data in qualitative databases, just because we are frustrated with other things, we need a more nuanced debate about what is possible to do to overcome epistemological/methodological problems in order to allow existing and future qualitative data to be deposited and re-used without violating any of the ethical, epistemological or methodological principles that guide social sciences in general and qualitative research in particular.

**Administrative and practical/technical issues**

A final reason for a lack of enthusiasm among researchers towards depositing their data may be the fact that it involves additional work. Good management of a database means that the data included should be properly organized, carefully preserved and well documented.

A large part of this job, especially related to the data documentation, is supposed to be done by the researchers themselves. Many researchers however are not ready, or do not have time, or possibly even the skills to be able to do this. Besides, this job is viewed as an administrative task rather than academic work, and is most often not paid, which all together results in low motivation among scholars.

There is a growing consensus among all the actors involved that this job needs to be paid (Travers, 2009), possibly through including special funds in research grants that would cover the costs of preparing the data (Carlhed and Alfredsson, 2009).
interesting motivating measure, proposed by, among others, Carlhed and Alfredsson (2009), (see also, Axelsson and Carlhed, 2009; and Elman et al. 2010), is to make work related to data archiving acknowledged as having scientific merit, which would presumably make more researchers become interested and motivated towards do this job.

**Qualitative data archiving and “re-use” in Sweden**

The preceding part of this paper has primarily focused on the British debate on different aspects of qualitative data archiving and re-using. The debate was initiated by the ESRC’s requirements from 1996. A first impression suggests that there were only a few reactions to the new policy requirements from the qualitative research community during the 1990s (more influenced than others were Hammersley, 1997; Griffin, 1997 and Mauthner et al 1998). It took almost a decade, as Moore (2007) has noticed, before the debate moved into the mainstream of sociological debate, so that towards the end of the 2010s it has become increasingly vibrant.

Internationally, the OECD’s documents from 2004 and 2006 have initiated work directed towards promoting qualitative data archiving and re-use in most of the OECD countries. At the same time, however, these policy initiatives came nearly a decade later than those in UK. Apart from the odd exception, academic debate, similar to the British one, has still not started in these countries2.

This section of the paper is going to discuss the developments related to the qualitative data archiving and re-use in Sweden. It is difficult to estimate how the above-mentioned OECD documents affected development in Sweden, because the Swedish policy related to qualitative data archiving has been part of a much broader policy initiative directed towards extensive investments in research infrastructure, with the general aim of promoting Swedish research within all sciences. This policy was from the beginning part of and coordinated with the EU initiative for coordinated development of pan-European research infrastructures, which was carried on by the European Strategy Forum on Research Infrastructures (ESFRI). The first concrete measures in Sweden occurred in 2005, when the Board of the Swedish Research Council established the Committee for Research Infrastructures (KFI) with the aim of building a strategy for developing Swedish research infrastructure. The Committee’s work resulted in The Swedish Research Council’s Guide to Infrastructure (2006), which described necessary long-term preconditions and needs for Swedish research, in order to produce research of the highest quality in future. The document deals with six general scientific fields3, including Humanities and Social Sciences as one of them.

---

2 One such example is a special issue of the Australian Journal for Social Issues (Vol. 44, No 3), 2009, which was dedicated to various aspects of qualitative data preservation and sharing.

3 Astronomy and Subatomic Research, Energy Research, eScience, Earth and Environmental Sciences, Humanities and Social Sciences, Material Sciences, and Medicine and Life Sciences.
As part of the general infrastructure project, and within the field of the Humanities and Social Sciences, the Swedish Research Council also established in 2006 the Database Infrastructure Committee (DISC), with the aim of establishing databases and developing effective instruments for sharing research data among Swedish researchers. For more practical work in this context the Swedish National Data Service (SND) was established in 2007. SND is a service organisation for research in the humanities, social sciences and medicine, with the aim of securing technical, administrative, legal and educational resources for collecting and storing research data (both quantitative and qualitative) and making it accessible to other researchers.

Before I make any comment on the developments in Sweden, I have to explain the reason for putting the word “re-use” in quotation marks in the heading of this section. A most important formal obstacle to qualitative data sharing in Sweden is the fact that according to The Personal Data Act (1998:204) it is illegal to use research data other than in specific projects for which the data were collected. In a telephone interview with a representative of SND, I asked him how they dealt with this problem. His answer was that they formally informed the Government about the problem, but as to if and when the political process of adjusting legislation to the needs of further development of the research infrastructure will take place, it is impossible to get an answer. In the meantime, he said, the SND has already made it possible even for qualitative researchers to deposit their data in the SND (because it is not illegal to archive data) and does all the necessary administrative and technical preparation work to make these data accessible for re-use sometime in the future.

Thus we have the peculiar situation that the SND actively promotes data archiving without making it possible to re-use the data. However, even if we leave aside this particular detail there are other general aspects in Swedish policy related to the research infrastructure development that need to be mentioned.

It is clear that Swedish research infrastructure development is a huge, complicated, expensive and, perhaps more than anything, far-reaching project in terms of the consequences for research communities. At the same time, the research community (or at least large parts of it) seems to have been excluded from the process. As an illustration of this, I shall present some findings from the survey carried out by SND (already quoted above) and summarized in Axelsson and Carlhed (2009). The sample of 549 professors and 1147 PhD students in Swedish social sciences and humanities departments were asked about their knowledge, routines and practices in relation to data archiving and sharing. The fieldwork was done during 2008 and the study was in fact a replica of the Finnish study, which was conducted by the Finnish Social Science Data Archive (FSD) two years earlier, and whose results are presented in Kuula and Borg (2008).

---

4 SNDs home page (http://snd.gu.se/sv/forskningssupport) refered to 11-Aug.-11
Regarding the Swedish Research Council’s most important document related to long-standing plans for development of the national research infrastructure, the “The Swedish Research Council’s Guide to Infrastructure” (2007) the following results were obtained:

- Only 11% of the professors and 1% of the PhD students were familiar with the document.
- 50% of the professors had heard about it but not in detail:
- 40% of the professors and 82% of the PhD students had never heard about it;

Concerning knowledge about the OECD’s corresponding documents, 75% of all participants (whereof 61% were professors) had never heard about these documents. On the routine archiving of qualitative data, almost 50% of the professors said that such practices were rather unusual or completely absent at their institutions. 46% stated that research data were most often kept by researchers without any intention of archiving, and only 11% stated that research data are routinely documented, categorised and archived. A consoling result of the study was nevertheless the fact that 49% stated that at least it is very unusual to destroy data.

So what we can conclude here? We have a huge ongoing building project of the national research infrastructure on the one hand and on the other a widespread lack of knowledge/awareness on the part of those who are going to be most affected by this project.

This is, however, not all. SND is, as mentioned, the most important actor in the process of data-archiving and re-use. Looking at the composition of the SND’s five-member board, or fifteen-member scientific committee, a tentative conclusion, which may be drawn after a quick look at their academic backgrounds, is that almost none, or at least very few, of them belong to the qualitative research community.

So the Swedish qualitative research community is not only uninformed about the policy measures related to building research infrastructure and inclusive qualitative data archiving and re-use, but they are also mostly excluded from these processes. To date the job has been done primarily by policymakers, IT experts and people who previously were responsible for developing and maintaining the national statistics (dominated, of course, by qualitative professionals and academics). The research community (especially qualitative researchers) for the most part, as mentioned, has been excluded from the process. They are now being invited in, when the process is practically over, or at least its designing and projecting phase. In the SND survey the scholars were even asked their opinion on what would be the most effective measures in enhancing the re-use of digital data. Over 93% of all participants stated that more information about the data sharing and data bases would be effective, almost 100% declared that more training in research methods and digital research databases could be effective, nearly 90% thought that research grants should include resources for data documentation before archiving, and finally the same percentage of all participants held the opinion that work on data documentation should be counted as academically meritorious.
The question is however whether these results may be taken seriously. If a majority of the participants have no, or very limited, knowledge about the actual policy measures related to data sharing, then they obviously know even less about the British academic debate. Consequently, they know very little about all those epistemological, ethical, political and technical problems that the British debate has pointed out. At the same time the SND’s questions are formulated as if these academic problems do not exist at all.

Yet another point is that a great part of the SND’s activities (how great is up to future research to find out) are based on a neo-positivistic paradigm, i.e. designing and building the database infrastructure in the same way for qualitative data as for quantitative data. As mentioned, references to or influences from the ongoing ethical and epistemological discussion about the differences between basic preconditions for archiving for qualitative respective quantitative data bases are completely absent. They imply a prevailing neo-positivistic understanding that all data should be treated in the same way.

Concluding remarks

This paper has considered policies for qualitative data archiving and re-use, and the reactions to these processes from the academic community. The first systematic policy measures were implemented in the UK in the middle 1990s. Reactions from the research community came shortly after that, but it took almost a decade before a full-scale academic debate on all relevant epistemological/methodological, ethical/legal, political/ideological and practical/technical issues of the phenomenon developed. At the same time the results of the archiving/sharing policy have been rather modest so far. Researchers are still reluctant to deposit their qualitative data in archives, yet there are still quite a few academic publications based on the analysis of others’ qualitative data.

Internationally, policy measures, which corresponded with the British ones, though with approximately a ten year delay, followed the OECD initiatives of 2004 and 2006. In the EU and Sweden these documents coincided with the development of the pan-European research infrastructures initiated by the European Strategy Forum for Research Infrastructures. Three points are sufficient to summarise progress in Sweden:

- There are many signs that the infrastructure for qualitative data archiving and re-use is being developed without the active participation of sufficient representatives of the qualitative research community;
- There are many signs that developing the infrastructure takes as a point of departure neo-positivistic or foundational epistemology, treating qualitative data in the same way as quantitative data, with all the methodological problems resulting from this;
There are very few, if any, signs that serious academic debate has started in relation to all the issues that have been raised by the corresponding debate in Britain.

If this paper does no more than help qualitative researchers in Sweden to become aware of some methodologically problematic consequences of actual policy trends, and even if it does not initiate an academic debate on these issues, it will still have more than fulfilled its purpose.

References


Denzin, NK (2009) The elephant in the living room: or extending the conversation about the politics of evidence. *Qualitative Research* 2009(9): 139-160.


Fry J, Lockyer S and Oppenheim C (2008) Identifying benefits arising from the curation and open sharing of research data produced by UK Higher Education and
research institutes, Department of Information Science, Loughborough University and Centre for Strategic Economic Studies, Victoria University, Melbourne.


The MES—occasional papers and reprints on migration and ethnic studies are published by REMESO, Institute for Research on Migration, Ethnicity and Society. REMESO is an Institute for research, post graduate and general education at the Department for Social and Welfare Studies, Campus Norrköping at Linköping University.

www.ep.liu.se

REMESO also publishes a REMESO REPORT series. This comprises reports resulting from stakeholder collaboration with Municipalities, NGO’s and civil society organizations, and business organizations. The REMESO REPORT series also disseminates the results of commissioned assignments from Government agencies.

www.ep.liu.se

TheMES No. 38-2011
ISSN 1651-8306

Series editor:
Erik Berggren
Tel. +46 11 36 32 66
E-mail: erik.berggren@liu.se