Swedish National Data Service's Strategy for Sharing and Mediating Data

Practices of Open Access to and Reuse of Research Data

Carina Carlhed and Iris Alfredsson[1]

Abstract

This paper begins with a description of the current key actors in Sweden, which are promoting research infrastructure and accessibility to research data, put into context. The Swedish National Data Service's (SND) organization, mission, and strategy to promote data sharing is also described. SND's strategy is a combination of top-down and bottom-up activities. An example of a top-down activity is to influence research funders to put higher demands on future open access data when studies are completed or to support researchers through the whole research process by providing guidelines on ethical and legal issues. Examples of bottom-up activities are to be present in different research contexts and to inform about the benefits of sharing data. One example of this is a joint project with SND and four university libraries. SND has conducted a national inventory survey, initiated in the fall of 2008, of existing databases and database research, as well as attitudes towards data sharing among researchers and university managements within social sciences and humanities departments at Swedish universities and university colleges.

In addition to the inventory process, two survey studies have been carried out in spring 2009, one targeting professors and the other doctoral students in the same domains of disciplines at Swedish universities and university colleges. The questionnaire contained 80 items covering the researchers' affiliations; domain of discipline; gender; age; familiarities with research policies and ventures; and use, reuse, and archiving practices of digital research data. Furthermore, there were questions about possible reasons for not using digital data, interventions and barriers to enhanced reuse and accessibility to data, possible agents in overcoming barriers, and willingness to share their digital research data. The surveys were carried out through email questionnaires sent to professors (N=549) and doctoral students (N=1147). The results from the surveys show that doctoral students in general expressed great uncertainty about questions of amounts of reusable digital data and effective interventions to enhance accessibility to digital research data. They identified research ethical aspects as important barriers to sharing digital research data, while professors emphasize lack of resources for researchers to document and make their data accessible for others as the most important obstacle. Concerning interventions to enhancing reuse of digital data, the majority of the doctoral students and the professors thought it should be effective to get more information about accessible research data in data archives or databases. Nearly 100% in both groups reported that more training in research methods, digital research databases, and information about accessible e-tools would be effective interventions. The most effective interventions for enhancing accessibility to digital data were that research grants should include funds for preparing the data for sharing and archiving and that archiving data for use by the scientific community is acknowledged to be of scientific merit. Surprisingly, when it comes to the degree of urgency in sharing their own data, the professors seem to be a bit more eager to share data than the doctoral students. The results are compared with the results from the parallel study of the professors and from a recent survey targeted at professors in various social sciences and humanities disciplines at Finnish universities (Kuula and Borg, 2008).

1. Introduction

1.1 Building a Swedish research infrastructure

The Swedish Research Council (VR) has, since its start in 2001, been focusing on the need to build a research infrastructure. As a part of this work the Committee for Research Infrastructure (KFI) was set up in 2004. The main purpose of KFI is to formulate long-term strategies and handle resource allocation for expensive scientific equipment, large research facilities, and extensive databases. The committee also deals with Swedish interests in, and funding of, various national and international research infrastructures. The overall aim is to provide better conditions for Swedish researchers by ensuring access to high-quality infrastructures.

The committee is also the producer of the Swedish roadmaps for research facilities to meet future scientific demands. The first *Swedish Research Council's Guide to Infrastructure* was published in 2006 and the second by the end of 2007 (The Swedish Research Council, 2007).

As part of the Swedish Research Council's major

infrastructure initiative, the Database Infrastructure Committee (DISC) was founded in 2006 (www.disc.vr.se). DISC's mission is to promote the development of an effective infrastructure for sharing research data resources in Sweden and it aims to ensure that researchers have rapid, easy, and free-of-charge access to research databases of high quality. High quality here means up-to-date, relevant, quality-assured, well-documented, and standardized databases meeting high international standards of quality, comparability, and security. The mission also includes creating new joint research data and facilitating access to existing data.

One of DISC's first key issues concerned transforming the Swedish Social Science Data Service (SSD) into the Swedish National Data Service (SND, www.snd.gu.se). The matter was studied during 2006, and in autumn 2006 there was a call for applications to host the new data service. The University of Gothenburg was, in competition with four other universities, appointed to host the SND. A five-year agreement to support the organization was signed by the research council and the university in November 2007.

The new organization should not only replace SSD, but also take responsibility for a broader area. According to Section 3 of the agreement, SND "shall meet the needs of the research community for data on empirical research in the areas of social science, humanities, and medicine. Actions include providing technical, legal, educational, and other administrative resources for collecting, storing, and distributing data for research."

SND is governed by a steering board and by a national reference group. The board of SND consists of five national representatives for the above sciences, appointed by the Swedish Research Council, the national reference group, and the University of Gothenburg. Formally the new organization started 1 January 2008. However, SSD performed the operational tasks during the first six months of 2008. On 1 July 2008 most of the SSD staff was transferred to the new organization. At the same time SSD was closed down, and the SSD data collection and equipment were taken over by SND.

1.2 The Swedish National Data Service (SND)
According to the guiding principles of SND, the main purposes of the data service are to mediate information on databases and other digital material collections for research, to facilitate access to research databases, and to serve as a knowledge node for documenting and managing research data in several knowledge fields. Thus, a very important task for SND is to strengthen the altruistic approach of the importance of data sharing and open access among researchers.

Experiences from SND's predecessor SSD, show that this is not an easy task. Only a small proportion of data produced by Swedish social science researchers were deposited at SSD. SND's conditions are, however, better than SSD's: increased economic resources and an organization placed within a bigger network of infrastructural resources. However, an important factor for the result is the general attitudes towards data sharing among researchers. Is there simply no culture of data sharing and reuse of data among researchers in Sweden? Or does sharing and reuse exist, but not via a data service?

We have identified two major barriers for reaching our goals: legal barriers and possessive barriers. The legal barriers are obstacles in Swedish current laws and regulations. The possessive barriers are thresholds connected to unconscious attitudes of researchers.

1.3 Legal issues
Issues surrounding shared data infrastructures have important legal implications. For this reason, DISC has surveyed the legal regulations that apply. The survey includes an inventory of relevant regulations in the areas of integrity protection, copyright, and archiving (DISC, forthcoming). The report will provide a basis for determining the actions needed to facilitate the creation of a common data infrastructure. A working hypothesis at DISC is that the issues involved are so fundamental that they require a public investigation.

An example of legal obstacles pointed out by DISC is the regulation concerning the use of the Swedish person-identified population registers on health and social conditions. This very important source for Swedish research gives unique opportunities to create research databases for longitudinal research in medicine and social sciences. However, the current ban on creating a common research infrastructure with personalized data limits the use of these resources. The Personal Data Act, the Secrecy Act, and other regulations allow the use of research material only for specific projects. Universities may not collect and store data intended to serve a wide number of researchers in the same scientific area.

DISC also calls attention to the fact that the central Swedish administrative agencies, mainly Statistics Sweden, are not given the basic instruction to provide the research community with data from registers. Instead they sell research data to individual research projects as the need arises. This results in the fact that research funders over and over allocate funds to purchase the same research data.

While DISC is looking into the need for new regulations, SND will work on the task of informing researchers on legal issues. The impression is that there is a great deal of uncertainty among researchers when it comes to the legal aspects of data sharing. Starting in June 2009, one of the university jurists will support SND with legal advice. This project will include training the SND staff in legal matters

concerning the collection of research data, as well as the use and reuse of data. The jurist will also represent SND in the cooperation between DISC and SND on legal matters.

1.4. Possessive barriers
Data collected by a researcher or a research team are often considered to belong to the original investigator(s). This is not the case, as the ownership of the data most often is connected to the university where the researcher is employed. Nevertheless Swedish researchers often bring the data with them when changing workplaces.

Experience shows that data not used and taken care of rapidly get obsolete. Documentation gets lost and old data formats become unreadable.

When asked if they would consider depositing their data at SND, researchers often doubt that their data are of interest for other researchers. Other reasons for not sharing data with others are that data are not properly documented and organized or that reuse of data needs a lot of information from the principal investigator.

1.5. Activities to promote data sharing
The SND strategy to promote data sharing is a combination of top-down and bottom-up activities. An example of a top-down activity is to influence research funders to put higher demands on future open access data when studies are completed. To make the researcher aware of the complete life-cycle of data, the research plan always should include a plan for how to preserve and share the data. Another way of encouraging data sharing is to regard it as a merit to make your research data available for other researchers.

Another activity is to support researchers through the whole research process by providing guidelines on ethical and legal issues, on how to store and document data, etc. The SND web site will be the central place for this information, but it will also be published in different printed versions.

Examples of bottom-up activities are to be present in different research contexts and to inform about the benefits of sharing data. One example of this is a joint project with SND and the university libraries in Gothenburg, Lund, Linköping, and Malmö. Financed by the Royal Library's Open Access Program, the aim of the project is to look into open access within the humanities and arts. The one-year project, starting in September 2009, will try to answer the following questions: Where and how to store research data? Which parts can be published as open access? How to connect the open archives and the Swedish National Data Service? How to connect research data and publications?

1.6. Feedback from the research community
When working on a strategy to promote data sharing, you need to know the opinion of the target group. Inspired by our colleagues at the Finnish Social Science Data Archive (FSD), we decided to ask the professors within the humanities and social sciences about their opinion on open access and data sharing. To compare with another target group within the research community, we also asked the same questions of the Ph.D. students within the humanities and social sciences.

1.7. Outline
The Swedish National Data archive (SND) is currently an operative key actor in conducting a national inventory survey, initiated in the fall of 2008, of existing databases and database research as well as attitudes towards data sharing among researchers and university managements within social sciences and humanities departments at Swedish universities and university colleges. The aim of this ongoing inventory survey is twofold: first, to identify and coordinate existing data resources; and second, to identify barriers and enablers to using and depositing data to open repositories. Some preliminary findings from this inventory survey and follow-up interviews with researchers based at the participating departments have revealed a number of important issues that need further investigation: the general unwillingness towards sharing information about research data with coordinating institutions (such as SND); the reported time scarcity preventing researchers from collecting, coordinating and delivering information about research data; and the ethical concerns of how to handle commitments to research subjects and how to protect sensitive information. A number of issues also clearly related to the fact that there was a wide distribution among social sciences and humanities disciplines represented in the inventory survey. There were, for example, quite different views among the respondents on fundamental issues such as the nature of and purpose of research, research ethics, ownership of research data and research results, and how to best enhance research infrastructures. In addition to the inventory process, two survey studies have recently been carried out – one targeting professors and the other doctoral students at Swedish universities and university colleges with input from the above mentioned national inventory study by SND, and from a Finnish survey, which was carried out 2006 by the Finnish Social Science Data Archive (FSD) targeting professors in various social sciences and humanities disciplines at Finnish universities and the practices related to open access to research data (Kuula and Borg, 2008).

The empirical part of this paper is based on two recently conducted survey studies, targeted at professors and doctoral students within humanities and social sciences at Swedish universities and university colleges, with the broader aim of investigating existing practices and attitudes when it comes to availability and reuse of research data. The results are tentative and descriptive and are discussed in a theoretical context in another conference paper (Axelsson & Carlhed, forthcoming).

2. Procedures of the surveys

The two surveys, one directed to Swedish professors in humanities and social sciences and the other directed to Swedish doctoral students in the same domains of disciplines, contained 80 items covering the researchers' affiliations; domain of discipline; gender; age; and familiarities with research policies and ventures, and use, reuse, archiving practices of digital research data. Furthermore, there were questions about possible reasons for not using digital data, interventions and barriers to enhanced reuse and accessibility of data, possible agents in overcoming barriers, and willingness to engage in promoteing changes in this area and to share their digital research data. The surveys were carried out through e-mail questionnaires and with lists of respondents based on retrievals from databases at the universities' offices for IT or personnel administration. In some of these lists it was easy to recognize respondents' disciplines; others were sorted by thematic or interdisciplinary departments and no information about discipline was accessible. Therefore, even departments that were within science and technology, educational sciences, and social medicine were included, but only departments that described themselves as interdisciplinary on their websites. Nevertheless, most of the departments were within humanities and social sciences. Because the population was broad and had somewhat non-distinct boundaries, we asked respondents to reply to us if they did not use perspectives of social science or humanities in their research. In those cases they were cancelled from the survey.

Initially, the survey was sent to 1589 professors from 35 universities/university colleges, and after the cancelling procedure of non-social science or non-humanities researchers (by the definition above) there were 1436 professors. The response rate was 38%, with 549 responses. The same procedure was carried out with the population of doctoral students. However, the lists from the universities that formed the respondent list had minor inaccuracy problems, due to some "natural" conditions, namely doctoral students becoming doctors. This affected the update status on information in the university personnel information systems, which had in some cases inaccurate information about the doctoral students. In addition, doctoral students at university colleges could also appear at a list from another university, hence with double mail addresses. A check up was made before the distribution of the e-mail questionnaire in order to avoid obvious doubles; however in some cases the e-mail addresses were abbreviated and impossible to relate to the names of the doctoral students. Similar to the professor survey, the population was broadly defined, which called for a similar procedure for cancelling, by respondents' reply stating their non-social sciences or non-humanities affiliation. Initially, the doctoral student population included 4697 potential subjects and after the cancelling procedure (mentioned above), 4065 remained. The response rate was 28%, with 1147 responses. When comparing how the professors' response rate patterns related to the distribution among a selection of the universities that received the largest proportion of questionnaires, we can conclude that the response rate from the larger respondent groups' universities alternated between 22 to 44 %. See table 1 below.

Table 1. Professors' response rate patterns

University	*Sent*	*Received*	*Response rate in %*	*Response rate in % compared with statistics*
Stockholm	286	123	43	53
Lund	279	63	22	40
Uppsala	217	80	37	45
Göteborg	213	75	35	42
Linköping	117	25	21	25
Umeå	84	37	44	47

Table 1 shows response rates based on the initial number of questionnaires sent before the cancelling procedure of non-social sciences or non-humanities affiliation. Because our method of selection was somewhat unstable, we found it necessary to investigate our precision further. The Swedish National Agency for Higher Education produces statistics about the universities and university colleges[2]. XXComparing statistics of professors and doctoral students and their affiliation to university and disciplinary domain from 2008 and our response patterns gives a view of how our survey succeeded in targeting the population. It seems that the population of professors (constructed from statistics, i.e., number of professors in different domains of disciplines and university), is well-covered by our group of professors who have participated in the survey. In concordance with this one can argue that our response frequencies are higher in reality (see table 1).

Table 2 shows the doctoral students' response rates

Table 2. Doctoral student response rate patterns

University	*Sent*	*Received*	*Response rate in %*	*Response rate in % compared with statistics*
Stockholm	1111	219	20	27
Lund	801	135	17	21
Uppsala	705	185	26	31
Göteborg	397	139	35	20
Linköping	220	58	26	36
Umeå	217	96	44	29

related to the distribution among a selection of the universities that received the largest proportion of questionnaires.

Like the professors' response rate patterns, the table above shows response rates based on the initial number of questionnaires sent before the cancelling procedure of non-social sciences or non-humanities affiliation. As argued above, the actual response rate is higher when comparing it with the statistical population, which we constructed for comparison reasons. For some universities, however, the response rate was lower in this comparison. It signals distortion in our precision about the doctoral students. In conclusion, our generalization opportunities are limited due to these aspects that have been discussed above. It seems that the ground for conclusion about the group of professors is more stable than the group of doctoral students. Nevertheless, a large number of professors (N=549) and doctoral students (N=1147) have participated in our studies, which implies considerable opportunities to assume valid conclusions.

Table 3. Amount of professors and doctoral students in different Domains of disciplines, survey participation, and data from statistics From the Swedish National Agency for Higher Education

	Professors statistics		*Professors SND survey*		*Doctoral students statistics*		*Doctoral students SND survey*	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*
HUM	*409*	*36*	*217*	*40*	*1268*	*31*	*376*	*33*
LAW	*83*	*7*	*29*	*5*	*181*	*4*	*29*	*2*
SOC	*655*	*57*	*290*	*53*	*2661*	*65*	*731*	*64*
*	-	-	*13*	*2*	-	-	*11*	*1*
Totals	*1147*	*100*	*549*	*100*	*4110*	*100*	*1147*	*100*

**For some participants no information about domain of discipline were available*

Based on the discussion above, our conclusion is that the results from our surveys could be treated as fairly valid, in spite of the relatively low response rate. The amount of responses from professors and doctoral students in different domains of disciplines, age, and gender corresponds to the official statistics that have been described and discussed.

2.1 Generalizability

In the professors' group, a majority of men answered the questionnaire, 73% compared to 27% women. This reflects the demographics of the larger population, whereas 23% of the professors in social sciences, humanities (and law) are women. In the group of doctoral students the conditions were opposite; 61% of the doctoral students in our survey were women. In comparing with the statistics from The Swedish National Agency for Higher Education, the larger population consisted of 56% women. In both cases we can conclude that women were slightly a bit more inclined to participate in our surveys than men.

Considering age, with our survey we seem to engage a larger part (25%) of the younger doctoral students (younger than 29 years old), than expected (16%). The same counts for the group of professors, but there were only a minor difference. Two percent more of professors participated that were younger than 50 years old (18%), compared to statistics from The Swedish National Agency for Higher Education (16%).

According to domains of disciplines, it seems that our groups of professors and doctoral students reflect the structure of the larger population (table 3).

3. Results

The Swedish Research Council has in a current venture made a long-term strategic plan - a roadmap *The Swedish Research Council´s Guide to Infrastructure* (The Swedish Research Council, 2007). In the questionnaire we asked the researchers about their knowledge about this venture and their opinions about it. Eleven percent of the professors were familiar with the venture and the guide and only 1 % of the doctoral students. Half of the professors' group did know about the venture but not its details and 40% did not have knowledge about it at all. This was also true for

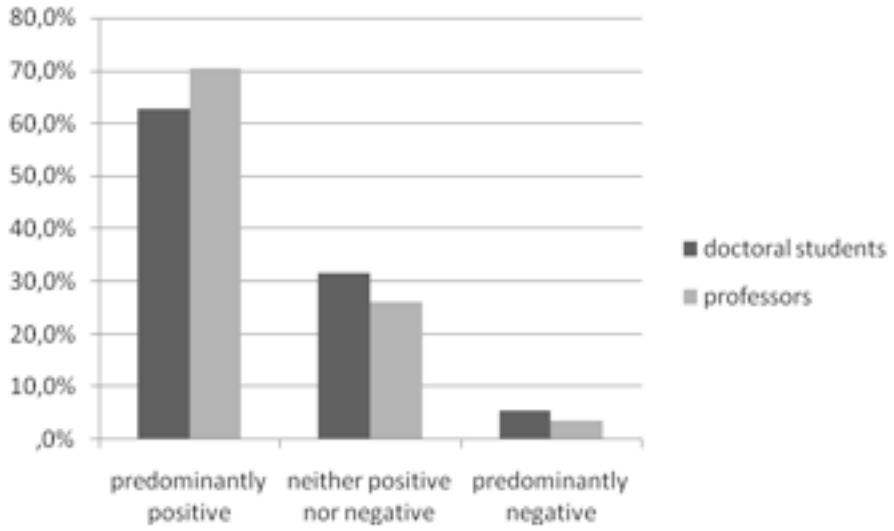


Figure 1. Opinions about the Swedish Research Council´s venture on research infrastructure (n=1109)

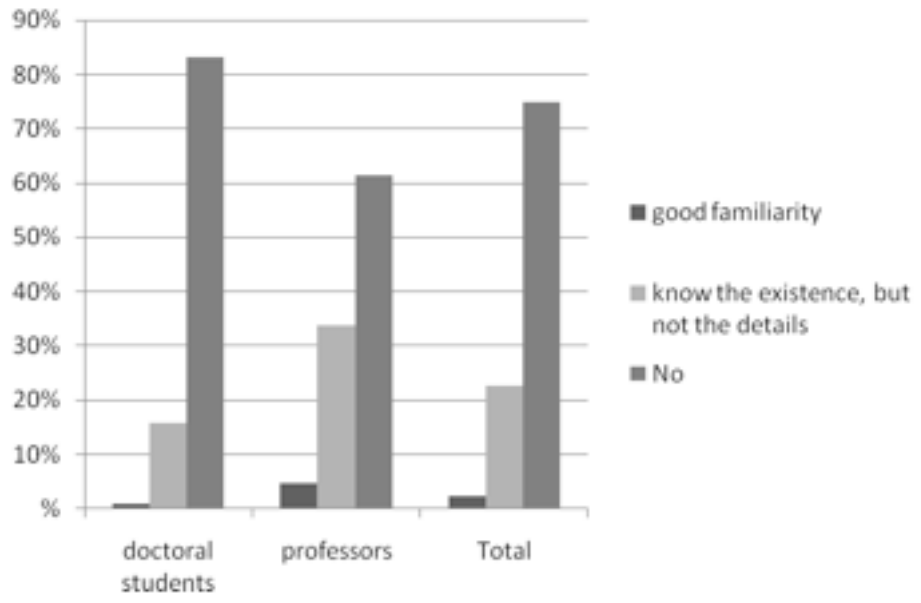


Figure 2. Knowledge about OECD Guidelines on Open Access to Research Data from Public Funding(n=1108)

the majority of the doctoral students (82%). Professors were more inclined to express positive opinions about the venture and the doctoral students followed the same pattern (Figure 1).

The knowledge about the *OECD Guidelines on Open Access to Research Data from Public Funding* (2007) was generally low; 75% of the researchers (both groups) did not know about it at all. Surprisingly, 61% of the professors were not aware of its existence (Figure 2). Breaking down results by domains of disciplines, it seems that professors within social sciences are the most informed about the OECD guidelines, and the group which was least informed was the doctoral students within law. Considering the situation of being a doctoral student, we are not surprised at the large amount of them not having knowledge about the guidelines and/or the research venture mentioned above.

3.1. Archiving practices and reuse of digital research data

The primary condition of archiving and reusing digital research data is that data are collected and compiled in some way. 73% of the professors stated that digital empirical data are used in research and 16% stated that the use of digital empirical data is unusual or are never used. The major part that did not use empirical digital data was the professors within humanities (56%). Among the doctoral students, they declared that digital empirical data was used (42%), but they expressed a great uncertainty about these questions generally.

According to the professors, the digital data are often kept by the researcher after analysis and reporting, without any actions to documentation (46%) but according to 15% of the researchers, they occasionally keep the digital data without any further documentation. Archiving practices where digital data always are kept and documented in a catalogue/database at the university is quite rare (11%). Almost half of the professors' group stated that these practices were unusual or never occurred. The same tendency showed concerning facilitating availability of digital research data at a data archive, that is, unusual practices. However, it seems that the research data are not regularly destroyed after analysis and reporting; at least 49% stated that destruction is rare

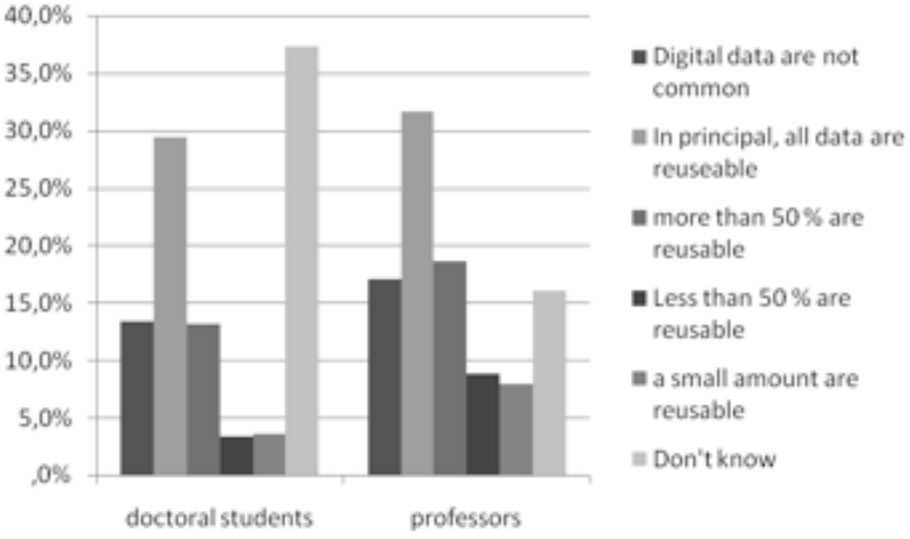
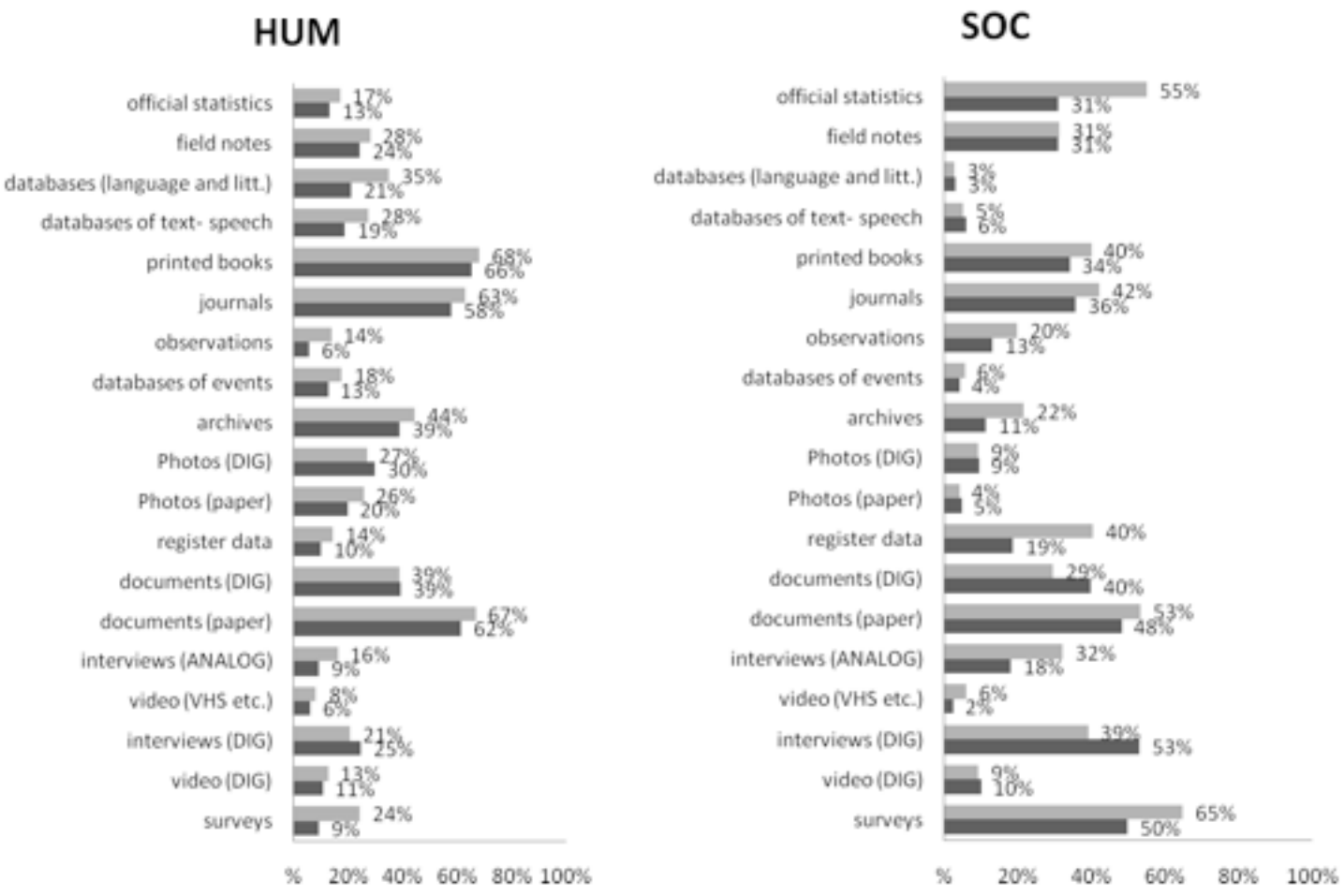


Figure 3. Amount of the digital data that are reusable (n=1346)



Figures 4 and 5. Researchers' use of different empirical research data within humanities and social sciences. Light bars represent the professors and dark bars represent the doctoral students. (n= 1677). Multiple response choices were available.

and only 3% reported that it was common. The reuse of digital data are relatively common; 59% of the professors stated that data are reused in Ph.D. works or other research projects and only 3% reported that it never happened. The use of reused digital data in teaching is also quite frequent according to 59% of the professors. Reusing all kinds of empirical data is most common in situations when researchers use the data themselves; approximately one-third stated that they pass data on to other researchers who are studying similar kinds of areas. Five percent of the professors reported that this never occurred. Regarding the amount of the digital data that are reusable, professors are more optimistic in general than the doctoral students, who seemed very uncertain and had difficulties expressing opinions of estimates (Figure 3). There were small differences between both groups and the domains of disciplines in these issues.

The wide range of empirical research data within social sciences and humanities is shown in figures 4 and 5.

In both domains of disciplines researchers use a broad empirical base, especially when it comes to use of non-digital empirical materials, where 74% of social sciences researchers and 86% of humanities always use several empirical sources. When considering the use of digital data it seems that the empirical data are less varied, according to 66% of the researchers in social science and 68% of the researchers within humanities.

Important reasons for not reusing digital data are mentioned by the professors as uncertainty about the quality of data (62%), ethical aspects (57%), technical issues (53%,) and juridical issues (49%). However, the professors' group is divided in opinions and the other part does not think that these factors are crucial (38%-50%). According to those who think ethical aspects are important, we found that these professors were mainly from social sciences. That is true

also for doctoral students in the same domain of discipline. The importance of juridical aspects is represented by the doctoral students in law, but not the law professors. Both professors and doctoral students in humanities deviated in general from the others in these issues, i.e., the technical issues were considered important. They also report other reasons for not reusing digital data, such as not using empirical or/and digital data at all, lack of knowledge and routines, decontextualized data having weak relevance for others, etc.

Concerning interventions to increase reuse of digital data, 95% of the doctoral students and 93% of the professors thought it should be effective to get more information about accessible research data in data archives or databases. Nearly 100% in both groups reported that also more of training in research methods, digital research databases, and information about accessible e-tools would be effective interventions (89%-95%). It seems that professors and doctoral students in humanities are most positive towards more education interventions and researchers in social sciences are the least positive, but all groups are generally positive to the interventions proposed.

3.2 Obstacles to sharing digital data

Our seven suggested obstacles to sharing digital data have been ordered in degree of perceived difficulty by the respondents. The professors regard deficiency of resources for researchers to document and arrange their data for reuse as the most difficult obstacle to sharing digital data. They also ranked lack of other resources like guidelines and directions for documentation as an important issue. Another obstacle highly ranked by the professors was doubt about the correct use of their data, i.e., risks of mistakes and misuses. An additional impediment was the fact that their respondents were not informed that their contributions should be used in the research society in general, only for a particular study. Juridical obstacles and loss of one's own advantage of competition in keeping data to oneself were not considered as crucial. The least important obstacle, according to the professors, was ethical aspects such as threats to confidentiality and delicate information. The doctoral students, however, thought that the ethical aspects mentioned above were the most difficult obstacles of all. After that, they considered the information given to the respondents and the use of their contributions to the research society in general was an important issue. Deficiency of resources for researchers to document and arrange their data to for reuse were also ranked as important, followed by juridical aspects. The least important obstacles according to the doctoral students were lack of other resources like guidelines and directions for documentation, loss of one's own advantage of competition in keeping data to oneself and doubt about the correct use of their data, i.e., risks of mistakes and misuses of data. The response pattern did not change depending on the researchers' use of digital data or not.

On the other hand, researchers in social sciences and women were more concerned with research ethical aspects and threats to confidentiality, etc., while researchers in humanities and men tend to stress lack of resources to document and arrange their data for reuse. According to age, older researchers tend to emphasize lack of resources and juridical issues. Younger researchers pointed out ethical aspects, threats to confidentiality, and doubts about incorrect use of their data. Thirty-eight percent of the professors and 34 % of the doctoral students stated that these obstacles did not prevent them from sharing data with The Swedish National Data Service (SND). However, 65% of the researchers indicated that these obstacles did prevent them from sharing data to SND. When we asked if, for example, SND could help them to overcome these obstacles, 43% (of the 65%) responded positively. The research funders were also regarded as important agents in overcoming the obstacles by 43% of those who expressed that the obstacles prevented them to share their data. Researchers who did not feel prevented to share data believed to a greater extent that SND could be of help (55 %) and research funders as well (59 %). There were however, minor differences, where the older researchers and the researchers in humanities were more optimistic about overcoming the suggested obstacles. There were very small differences between professors and doctoral students. On the other hand, if the researchers would consider engaging in promoting alterations in these areas, the doctoral students tend to embrace issues of research ethics and changing values in accessibility and practices in sharing data, while professors were inclined to issues of jurisdiction.

Surprisingly, when it comes to the degree of urgency in sharing their own data, the professors are a bit more eager to share data (30%) than the doctoral students (24%). In total, there were 53% of the researchers who thought it was urgent to *share* data (55% of the professors and 52% of the doctoral students), but only 26% in the total group reported that they *intended to share* their data. A large proportion of the total group expressed doubts about sharing data (40%). Researchers in law were the least keen on doing it and thought it was not so urgent. Researchers in humanities however, were those who distinguished themselves as potential "sharers." According to gender and age, men and older researchers expressed more willingness to share than others.

3.3 Promoting accessibility to digital research data

The results show which authorities the researchers point out as important key actors in promoting accessibility to publicly-funded digital research data and also to actively participate in shaping guidelines. The universities and university colleges were considered as the most important key actors according to 82% of the researchers and the second was the two largest research funders for social sciences and humanities, The Swedish Research Council

and The Swedish Council for Working Life and Social Research, with 80% of the researchers' responses. Fifty percent stated that Statistics Sweden would also be participating in shaping such guidelines. No significant differences between doctoral students and professors, or age, in this matter were observed. According to domains of disciplines there were very small differences, except for the opinion about participation of Statistics Sweden, where researchers within social sciences emphasized this as a key actor to a larger extent than the others. It seems also that male researchers are a bit more pessimistic about the importance and role the proposed key actors could play.

Furthermore, the most effective interventions for enhancing accessibility to digital data were that research grants should include funds for preparing the data for sharing and archiving (88% of the doctoral students and 83% of the professors) and that making data accessible for the use by the scientific community is acknowledged to be scientific merit (87% of the doctoral students and 83% of the professors). Generally, the doctoral students were more optimistic about the efficiency of interventions proposed, especially the issue of acknowledgement of promoting accessible data to be of scientific merit and except for those mentioned as top-ranked above, where the professors expressed beliefs in their efficiency in a higher degree. There were no or very small differences in response pattern among the domains of disciplines in these issues. According to gender it seems that women researchers believe in more education about life cycles of digital data (74%) and research ethics (77%) to a higher degree, than men. 59 % of the male group thought that more education about life cycles of digital data is needed and 63% of the male researchers believed that more education about research ethics is necessary.

4. Discussion

Our results are descriptive and have been presented tentatively in this article. Further statistical analyses are needed concerning impact of differences and in addition, scrutinized examination of all openended questions, where a lot of interesting comments are made by the researchers.

Overall we interpret the researchers' attitudes towards current ventures and strivings in research infrastructures as predominantly positive. A key actor is The Swedish Research Council that has, in a current venture made a long-term strategic plan - a roadmap *The Swedish Research Council´s Guide to Infrastructure (2007)*. The researchers' knowledge about this venture was minor. Professors were more inclined to express positive opinions about the venture and the doctoral students followed the same pattern.

The knowledge about *OECD Guidelines on Open Access to Research Data from Public Funding* was generally low and somewhat discouraging; professors within social sciences were the most informed, however. The least informed were the doctoral students within law. In comparison with the Finnish survey (Kuula & Borg 2008), where 81% of the professors did not know about the OECD recommendations, compared to 61% of the Swedish professors, it could be encouraging from the Swedish perspective. Nevertheless, time has passed between thses surveys and perhaps have the Finnish professors got more informed than they were in 2006.

A conclusion based on our results is that it seems important to raise issues of guidelines in social sciences and humanities concerning accessibility to digital research data and to engage researchers and relevant authorities in creating arenas for discussing and shaping research infrastructure for the future. According to the researchers, the universities, university colleges, The Swedish Research Council, and The Swedish Council for Working Life and Social Research are the most important key actors in promoting accessibility to digital research data from public funding through participation in shaping guiding principles.

The most effective interventions for enhancing accessibility to digital data that were identified were that research grants should include funds for preparing the data for sharing and archiving and that making data accessible for the use by the scientific community is acknowledged to be of scientific merit. More education about life cycles of digital data and research ethics were expressed as needs.

Considering archiving practices, use and reuse of digital research data, 16% of the Swedish professors stated that the use of digital empirical data is unusual or are never used. Eighteen percent of the Finnish professors reported a similar amount of digital data non-use. When comparing between the countries what happens to digital data after analysis and reporting, it seems that it was more common for Finnish professors to keep digital data without any further actions to documentation (56%) compared to Swedish professors (46%). Data are destroyed to a larger extent in Finland (20%) than in Sweden (3%). However, the saved data is reused by the researchers themselves to a greater extent in Finland (94%) than in Sweden (54 %). The opinions about amounts of reusable digital data differ also; 50% of the Swedish professors stated that more than half the amount of produced digital data is reusable, compared to 21% of the Finnish professors. In analyzing responses to important reasons for not reusing digital data, it appears that the Swedish researchers emphasize ethical, juridical, technical aspects, and quality of data as more problematic than the Finnish researchers. Again, it seems that there are a considerable amount of issues that need to be clarified and solved in order to develop a well-functioning research infrastructure with a high degree of re-using practices within social sciences and humanities. As information of importance, we think that the researchers' beliefs that promoting accessibility of their own data to

be acknowledged as a scientific merit and that research grants should include funds for preparing the data for sharing and archiving, points out legitimate measures with both force and enticement - like the stick and the carrot. The last mentioned intervention was also one ranked high by the Finnish professors (80%), but their top priority of effective interventions was establishment of guidelines and principles by the Finnish universities together (84 %).

The Swedish professors also point out other obstacles to sharing digital data, and regard deficiency of resources for researchers to document and arrange their data for reuse as the most difficult obstacle to sharing digital data together with lacking guidelines to documentation, while the Finnish professors reported that it was the situation when the respondents were not informed that their contributions should be used in the research society generally. They share this concern with the Swedish doctoral students. These results relate to the results mentioned in the former paragraphs, namely the need of different types of guidelines (ethical, technical, and juridical), earmarked resources to documentation and education in this area.

As we mentioned in the Results section we found that the professors seems to be more eager to share data than the doctoral students. A large proportion of the total group also expressed doubts in sharing data, probably because of uncertainty and lack of sufficient guidelines.

The Finnish questionnaire did not have a pushing question like we had, but the Finnish professors were asked their attitude to open access to digital research data collected in their own research and 76% of them expressed positive attitudes. One might conclude that professors in social sciences and humanities in Sweden and the Finnish professors differ a lot in opinions about digital research data. However, two years have passed with increasing focus on open access issues in research policies in these countries. It would be interesting to see if the Finnish professors have changed their minds since 2006.

About our own results, it is always interesting when the research is surprising. We were surprised that the professors were more positive and humble towards sharing and promoting accessibility to digital research data than were the doctoral students. But on the other hand, being a doctoral student means a lot in terms of keeping on one own's track, concentrating on the Ph.D. work, and having little time to orient oneself to ventures, research policies, and university practices. In conclusion and in spite of many prejudices about "conservative" professors, it seems that one has to acknowledge their positive orientation about e-science and put forward these survey results of barriers and opinions to be able to support and realize sharing of digital research data in the future. At last, the results of these surveys have to be acknowledged and seriously taken care of in understanding the obstacles and challenges we face, in order to achieve a sufficient and approved research infrastructure, adapted for the distinctive features of social sciences and humanities with their wide range of empirical materials.

5. Acknowledgments
The authors wish to thank the doctoral students and professors who participated in the surveys.

References

Axelsson, AS. & Carlhed, C. (forthcoming). Next Generation e-Researchers: Doctoral Students in Social Sciences and Humanities in Sweden and their Attitudes towards Open Access and Open Repositories. Paper presented at NCeSS National Centre for e-Social Science, the 5th International Conference on e-Social Science. Cologne, Germany, 24th - 26th June 2009.

Database Infrastructure Committee (DISC). (forthcoming). Juridical conditions for a research infrastructure.

Högskoleverket. (Swedish National Agency for Higher Education). http://www.hsv.se/statistik/statistikomhogsklan/personal.4.6df71dcd1157e43051580001770.html

Kuula, A. and Borg, S. (2008). *Open Access to and Reuse of Research Data – The State of the Art in Finland*. Finnish Social Science Data Archive 7, 2008. http://www.fsd.uta.fi/julkaisut/julkaisusarja/FSDjs07_OECD_en.pdf

OECD. (2007). *Principles and Guidelines for Access to Research Data from Public Funding* http://www.oecd.org/dataoecd/9/61/38500813.pdf

The Swedish Research Council. (2007). *The Swedish Research Council's guide to infrastructure*. http://www.vr.se/download/18.76ac7139118ccc2078b800011940/Rapport+5.2008.pdf

Notes

1 Carina Carlhed, Mälardalen University, Sweden, Email: carina.carlhed@mdh.se, and Iris Alfredsson, Swedish National Data Service, Sweden, Email: iris.alfredsson@snd.gu.se.

2 http://www.hsv.se/statistik/statistikomhogskolan/personal.4.6df71dcd1157e43051580001770.html. In this paper, comparisons have consistently been made between statistics from The Swedish National Agency for Higher Education for year 2008 and the background information about the participants in our surveys.