FORS
explore.understand.share.

Marieke Heers, Eliane Ferrez, and Emilie Morgan De Paula

Data sharing and re-use: Researcher practices, attitudes and needs

FORS survey of social science researchers in Switzerland

Lausanne, March 2017

Highlights

Between November 2016 and February 2017 FORS conducted a brief survey to better understand researchers' attitudes, practices, and needs specifically regarding data sharing and re-use. The insight gained from this survey will be used to improve existing services and prioritize new ones. Here are some key findings:

- Across disciplines, around 80 percent of researchers generally consider the sharing of research data to be very important.
- More than half of the respondents were involved in the production of both quantitative and qualitative data.
- More than two-thirds of the researchers stated that they had already shared data.
- Data sharing is most common amongst colleagues inside the same institution and within researchers' broader networks.
- Conditions most cited for sharing data include being informed about the projects that other researchers carry out based on one's data and the desire to be cited for the use of one's data by others.
- Re-using quantitative data is much more common than re-using qualitative data, at the same time most researchers indicated that they would consider re-using qualitative data for either their own research and/or for teaching purposes.
- There is a large interest in and strong need for additional services and materials in the area of data management, especially for data management planning, data preparation and cleaning, and security, storage and backup.

1 Introduction

FORS, the Swiss Center of Expertise in the Social Sciences, is a national research infrastructure at the service of the research community. The Data and Research Information Services (DARIS) of FORS solicits, preserves, and disseminates Swiss social science research data that can be used for secondary analyses. Between November 2016 and February 2017 we conducted a brief survey to better understand researchers' attitudes, practices, and needs specifically regarding data sharing and re-use. The insight gained from this survey will be used to improve existing services and prioritize new ones.[1]

A link to the online survey was sent to 10,692 potential respondents via email.[2] Respondents were promised confidentiality. Two reminders were sent, one of them in the FORS Bulletin. Respondents could fill in the survey in German, French, or English. The data from this survey and the documentation, including the questionnaire, can be accessed via our online data platform FORSbase. In the following sections, this short report describes the key survey results.[3] In total, 1,079 people participated in the survey, with a response rate of approximately 10%. 939 respondents provided a sufficient number of responses to questions, and are considered in the analyses presented below. [4]

Respondents represented a variety of disciplines, including sociology (20%), political science (13%), education science (12%), (social) psychology (12%), and economics (8%).[5] Most respondents were either professors (22%), researchers (22%), lecturers and post-docs (22%), or teaching or research assistants or doctoral students (17%). Moreover, more than half were affiliated with a university (57%), and about a quarter with a university of applied sciences or a university of teacher education (24%).

We asked respondents about the methodological approaches they have used in their research in recent years – 50% had used quantitative approaches, while 42% had used qualitative approaches. In addition, 39% said that they used mixed methods in their research. 5% reported that they do not carry out research themselves.[6]

We thank the researchers who took the time to participate in the survey.

---

[1] For this survey *data* were defined as "any observations, measurements, or facts that can be analyzed in order to draw conclusions and gain scientific knowledge. Data consist of information collected by researchers, and can be quantitative (e.g. survey results) or qualitative (e.g. texts, images, videos, and so on). Aggregated data are not included here".

[2] Potential respondents were social science researchers included in our FORSbase database.

[3] For a more in-depth analysis please contact the authors or download the data via FORSbase.

[4] Please note that due to filters and non-response for certain questions the number of respondents differs somewhat across questions.

[5] Additional disciplines included: social work (5%), anthropology (4%), communication sciences (3%), history (3%), demography (1%), linguistics (1%), and "other" (15%).

[6] Respondents could respond positively to more than one item within the question.

2 Results

The survey consisted of three parts. Respondents were asked to answer questions about (1) data sharing, (2) data re-use, and (3) using FORS services. At the end of the survey, users had the opportunity to provide us with some feedback. In the following sections we present the major results for each of these three sections.

2.1 Data sharing: opinions and practices

In the survey, *data sharing* was defined as “making data available to other researchers, so that they can (re-)use them for their own research projects and analyses. This could be either informally (e.g. with a colleague) or formally (e.g. via an archive)”.

When asked about the importance of data sharing, respondents generally judged this as very important or somewhat important. This holds – with slightly decreasing intensity – when asked concerning the social sciences in general (79% very important, 18% somewhat important), in their discipline (71% very important, 22% somewhat important), and in their own work (58% very important, 26% somewhat important).

The largest share of the respondents (51%) was involved in the production of both quantitative and qualitative data. 22% were involved in the production of quantitative data only, while 19% were involved in the production of qualitative data only. 9% were not involved in any production of data.

In the following part of the survey, we were interested in how researchers who produce quantitative or qualitative data judge the importance of data sharing in general: Researchers who have produced quantitative data only tended to judge data sharing as more important than researchers who have produced qualitative data only (respectively 91% and 69% said it is very important). 75% of those who have been involved in the production of both data types stated that sharing is very important.

Next, we turn from attitudes about sharing to actual sharing behaviors. Data producers were asked about their previous data sharing activities – 67% stated that they had already shared data. The results also show that 65% percent of those who had been involved in the production of quantitative data had already shared it, while 50% of those who had been involved in producing qualitative data had shared it. Sharing behaviors differ across disciplines (Figure 1). Sharing quantitative data is most common in (social) psychology and political science, while sharing qualitative data is most common in the educational sciences.

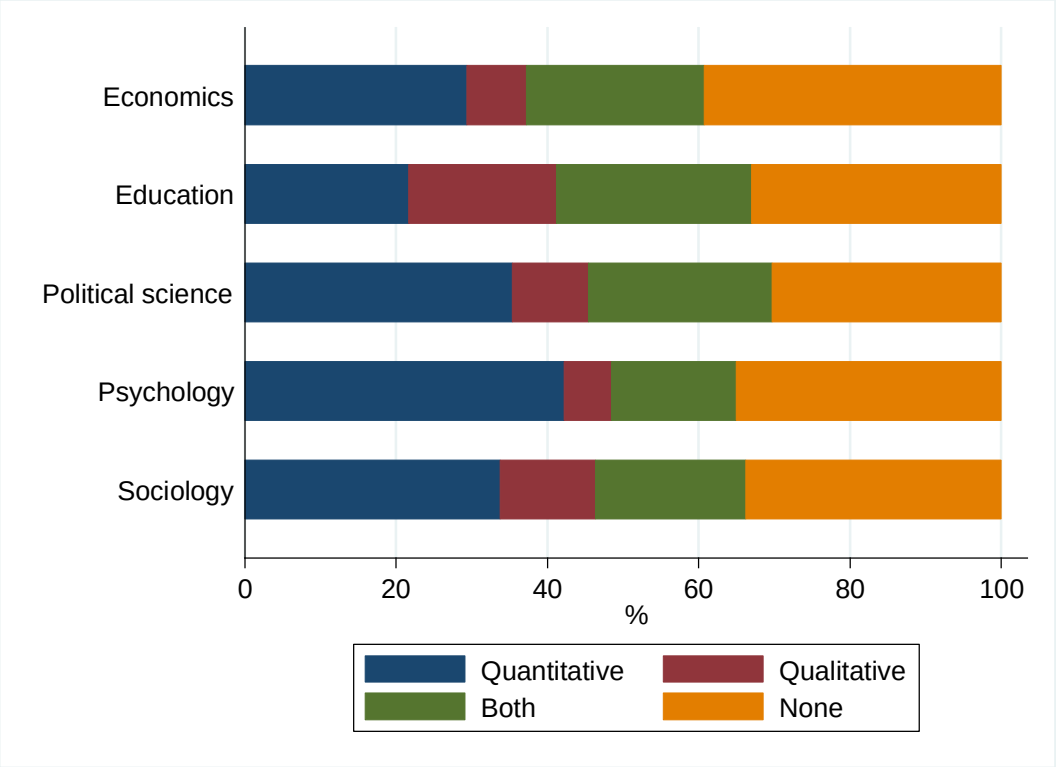

Figure 1: Data-sharing, by discipline (n= 545)

Among those who had been involved in the production of both data types, and who had shared their data, 54% had shared both types, 35% had shared only their quantitative data, and 11% had shared only their qualitative data.

Most had shared their data with colleagues inside their institution (83%) and with people from their broader network (69%). Sharing via an archive (24%) and through an open access or institutional online repository (19%) were less popular sharing options.

The reasons for sharing are diverse – a colleague's request (64%) is the most popular one, followed by the objective of creating new collaborations (53%). Almost half of the respondents share data 'by principle' (48%). Funders' (22%) and scientific journals' requirements (13%) are less common reasons for sharing data.

One of FORS' missions is to encourage data sharing within the Swiss social science research community. Therefore, we asked respondents who had not yet shared data if they would be willing to do so – 29% said that they would be willing to share their data, while 65% would be willing to do so under certain conditions. Some of the most commonly mentioned conditions are that respondents must be protected and that users of data have to cite the data producer. Only 6% of the respondents were not willing to share their data.

In order to improve its services, for FORS it is particularly important to understand the obstacles and barriers that prevent researchers from sharing their data[7]. Therefore, we asked respondents to what extent they agree with the following statements. In parentheses, we show the share of respondents that agreed "somewhat" or "strongly":

- I want to publish my results before anyone else uses my data (85%)
- I lack the time or resources to prepare my data for sharing (70%)
- My data are too sensitive to share with other researchers (42%)
- Other researchers would not fully understand my data or might use the data improperly (31%)
- My data are probably not of interest or use to others (20%)
- My data are my intellectual property and should not be used by others (19%)
- None outside the research team should ever have access to the data (13%)
- Data should be destroyed as soon as the purpose of their collection has been fulfilled (10%)

We also asked respondents about the support and encouragement they have experienced regarding data sharing. The results indicate that at the institutional level support for data sharing could be more strongly encouraged: 59% were encouraged very little or not at all by their institutions. Only 16% stated that they were very much encouraged and supported by their institution to share their data. The numbers are similar for encouragement and support by research funders and scientific journals. Colleagues seem to play a more powerful role in this respect – 36% of the respondents were somewhat encouraged and 16% were very encouraged by their colleagues to share their data.

2.2 Importance of data re-use

The next section of the survey dealt with data re-use, which was defined as "conducting analyses of the (non-aggregated) data produced by others in order to address new research questions".

40% of the respondents had re-used quantitative data, 8% had re-used qualitative data, and 21% had re-used both types. It is interesting to consider whether those researchers who have shared data are also those who have re-used data: 78% of the data sharers report to also have re-used data.

[7] Respondents could also mention additional reasons and obstacles. For an overview please contact the authors.

Interesting differences in re-using data emerge across disciplines (Figure 2). For instance, re-using quantitative data only seems much more common in sociology than in educational sciences. Compared to the other disciplines, in political sciences it is more common to re-use both quantitative as well as qualitative data.

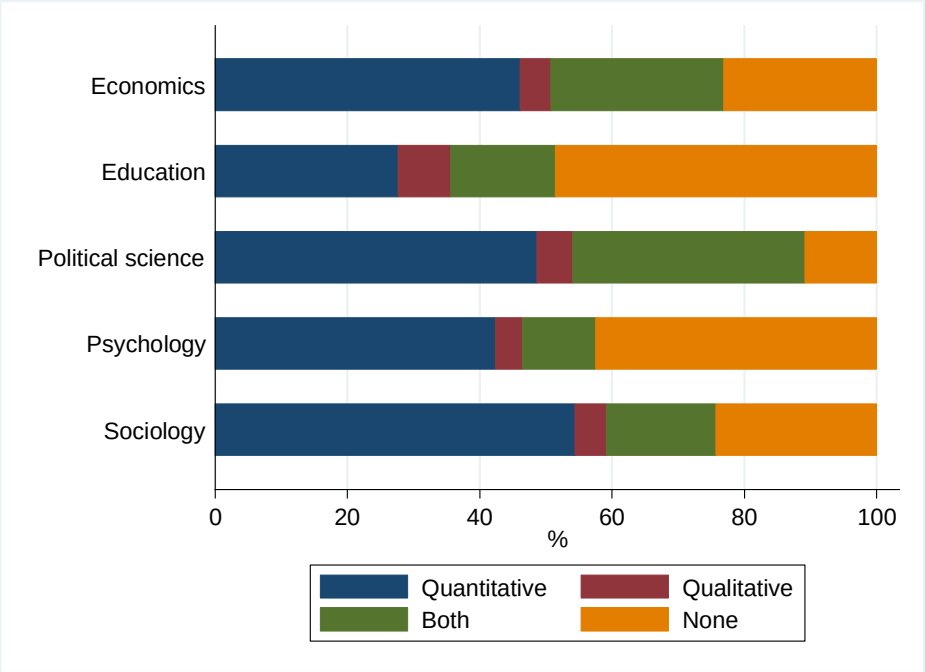

Figure 2: Re-use of data, by discipline (n=855)

An important aspect of re-using research data is how to obtain data that have been produced by others. This is particularly interesting for FORS, because with our online platform FORSbase we offer the largest data archive for the social sciences data in Switzerland. Most respondents in our sample had obtained data from the Swiss Federal Office of Statistics (64%) and from their network (62%). 43% had obtained data from FORS (via the archive, Nesstar, or the Swiss Household Panel). 41% had obtained data from institutional, regional, or national databases (also referred to as data repositories). Databases from other countries were less used (37%).

2.3 FORS services

FORS offers a number of support services to social scientists in Switzerland. To get a better idea of what researchers actually need, we asked to what extent they had used different existing services in the last 3 years.

60% had already explored the online catalogue of Swiss research projects at least once, with 12% four or more times. 47% had at least once searched for data made available by FORS, with 12% having searched four or more times. 34% had registered a research project at least once. 10% had deposited data in our data archive once or more often. 21% had asked for support or information from FORS staff at least once.

We are currently planning to expand our services in the area of *data management,* which was defined in the survey as “good practices for enhancing and ensuring usability of data in the long run”. To meet researchers’ needs we asked respondents what data management-related services they would find particularly useful. Overall, they were positive towards new services. The activities are presented in the order of their rated usefulness, along with the percentages of respondents who rated each activity either “very useful” or “somewhat useful”.

- Data management planning (80%)
- Data preparation and cleaning (78%)
- Security, storage and backup (76%)
- Documentation (73%)
- Anonymization (70%)
- Informed consent (69%)
- Folder and file organization and labelling (59%)

Another recent focus of FORS is the archiving and dissemination of qualitative research data. Therefore, we wanted to assess to what extent there is an interest in re-using this type of data. Our findings show that there is significant interest: 79% said they would re-use qualitative data for their own research, while 85% would re-use them for their teaching activities.

3 Conclusion

The results from this survey reveal that overall researchers see data sharing and re-use as highly important. This is the case across disciplines. When asked about more practical questions, researchers have more reservations. For example, the proportion of those who think a data sharing culture is important is higher than that of those who have actually shared data. It will be interesting to re-assess the situation in a few years.

Important conclusions can be drawn with regard to the existing services of FORS. First of all, regarding the barriers and obstacles for data sharing it would be worth reconsidering how the benefits of FORSbase could be better put forward. For example, many respondents state that they want to publish their research findings before they are willing to share their data. FORSbase allows doing this, and researchers can publish their data under an embargo for a certain period of time.

Another option in FORSbase that could be made more visible to researchers is that of *prior approval* – among the respondents, many expressed willingness to share their data under certain conditions stated that they want to be informed about the research projects of other researchers based on their data. In such cases, requiring prior approval before allowing access is an available option.

Similarly, many respondents mentioned that they want to be cited when their data are used in publications. Indeed, a requirement to cite data is part of the contract that researchers sign with FORS when they download data, and this could be made better known to data producers who might want to deposit their data at FORS.

The survey results also show that there is a large interest in and strong need for developing additional services and materials in the area of data management. Data management planning was rated as most useful. This fits well into FORS' current efforts to establish materials that aim at supporting researchers in their data management planning. An essential part of these efforts is to convince researchers that making data management a priority from the beginning of a project not only benefits the researchers during their own research process, but also facilitates data sharing at a later stage.

Respondents were also positive about the planned extension of offering more qualitative data at FORS, with many stating that they would be very interested in using them for research as well as for teaching. This is very encouraging for FORS, since we are currently making various efforts to reach out to qualitative researchers in order to raise awareness regarding data sharing and re-use.

It will be generally interesting to observe how the culture of data sharing and re-use further evolves in the coming years. To monitor these developments and others within

the Swiss social science research community, we will be carrying out similar surveys biannually.

Overall, the results presented above indicate that data sharing and re-use are not uncommon practices in the Swiss social science research community, although these might have been influenced by the definitions that we employed for “data sharing” and “re-use”, and by the fact that those who chose to participate in the survey may have been more positively oriented toward the topics. In order to better understand the subtleties in the current research context, a next step will be to carry out qualitative interviews with some of the respondents who expressed a willingness to discuss these issues with us.